

Arabic Sentiment Analysis on Chewing Khat Leaves using Machine Learning and Ensemble Methods

Wael M.S. Yafooz

Computer Science Department
Taibah University
Madinah Munawarah, Saudi Arabia
wyafooz@taibahu.edu.sa

Essa Abdullah Hezzam

Information Systems Department
Taibah University
Madinah Munawarah, Saudi Arabia
essa_alkadasi@yahoo.com

Waseem Alromema

Information Systems Department
Taibah University
Madinah Munawarah, Saudi Arabia
wromema@taibahu.edu.sa

Abstract-Sentiment analysis plays an important role in obtaining speakers' opinions or feelings towards events, products, topics, or services, helping businesses to improve their products. Moreover, governments and organizations investigate and solve current social issues by analyzing perspectives and feelings. This study evaluated the habit of chewing Khat (qat) leaves among the Yemeni society. Chewing Khat plant leaves, is a common habit in Yemen and East Africa. This paper proposes a model to detect information about the Khat chewing habit, how people explore it, and the preference for Khat leaves among Arabic people. A dataset consisting of user comments on 18 youtube videos was prepared through several natural language processing techniques. Several experiments were conducted using six machine learning classifiers and four ensemble methods. Support Vector Machine and Linear Regression had almost 80% accuracy, whereas xgboost was the most accurate ensemble method reaching 77%.

Keywords-sentiment analysis; machine learning; classification; ensemble methods

I. INTRODUCTION

Nowadays, the study of user opinions has attracted substantial attention in social perspectives, focusing on services, products, and habits through many data mining applications, recommender systems, and business intelligence applications. The analysis and interpretation of user opinions are altogether known as sentiment analysis, which is an area of natural language processing, also known as the voice of the customer in business intelligence [1, 2]. Business owners need to be aware of feedback to improve future performance. Such a time-consuming and difficult task is used in the analysis of huge unstructured data gathered through social media or internet comments. Several studies have been conducted, classifying sentiments as positive, negative, or neutral [3, 4, 6]. More complex sentiment analysis [7-10], often referred to as fine-grained, classify datasets into five classes, namely very positive, positive, neutral, negative, and very negative. Moreover, aspect-based sentiment analysis [11-14] classifies datasets by extracting entities from text.

Users' comments are often an outcome of their opinions, and they can be considered as the main factor in evaluating services or products. Some studies focused on education [15, 16], while other researches focused on detecting health

misinformation on social media users [18, 19]. Khat is a type of plant that pleases and stimulates, and chewing Khat leaves is a commonly seen habit in Yemen and East Africa [21-23]. Although it is customary in these countries, several experiments showed its direct impact on human organs. This paper presents a model to study consumers' opinions regarding the habit of chewing Khat leaves in Yemeni and East African society. At first, the dataset was collected by extracting user comments from 18 youtube videos. The annotation process classified the data into positive and negative fractions. Several NLP processes were executed to prepare the data for Machine Learning Classifiers (MLCs) and Ensemble Methods (EMs).

II. RELATED WORKS

Several studies utilized sentiment analysis in different ways. Multilingual student comments, obtained through student feedback, were used to evaluate online courses' effectiveness and teachers' performance in [3, 15-17]. In [3], the dataset was collected using approximately 4000 student comments through surveys conducted on 25 university courses to evaluate the performance of a professor who had been teaching for 10 years, while the sentiment analysis was directed including positive, negative, and eight more emotions. Similarly, authors in [15] proposed a system to evaluate a lecturer's performance by collecting data through student surveys via a rating system in a form of numerical data. The MLC Naïve Bays was employed to predicate the positive and negative students' sentiments toward the lectures. A recurrent neural network of long and short term memory in deep learning was utilized in [16]. The dataset was collected from 3000 positive, negative, and neutral student comments on 30 courses. The performance improved when using the softmax activation function, reaching 89%, 99%, and 90% during training, testing, and validation, respectively. Deep learning was applied on a course evaluation dataset with 3000 student comments using three predefined classes in [17], while the results showed that relu and softmax performed better.

Sentiment analysis is used to identify the main factors affecting the success of businesses, particularly start-ups. In [1], user comments were extracted from Twitter using topic modeling and applying supervised vector machine learning to divide comments into three main classes. The textual analysis was applied based on the entities trained in the previous phase

Corresponding author: Wael M.S. Yafooz

using Nvivo software. In [4], an analysis of a massive amount of user comments (approximately 1.6 million) from the Yelp Challenge Dataset was conducted. The dataset was divided to 20% for testing and 80% for training, using four machine learning classifiers. The best accuracy rate reached 92.6% and 92.3% under Stochastic Gradient Descent and Linear Support Vector Classification respectively. Similarly, the same dataset was utilized in [2] to analyze restaurant reviews through a hybrid classifier ensemble method using Naïve Bayes, Support Vector Machines, and Genetic Algorithms.

Some health sector studies have also been conducted [18, 19]. Authors in [18] focused on tweets on breast cancer, collecting user comments from approximately 845 cancer patient accounts with 48,000 posts. The logistic regression classifier and a Convolutional Neural Network was utilized in the process, and the model's performance accuracy was 97.6%. Besides, it was found that positive experiences had more shares, providing more awareness to the general public. Descriptive statistics of text mining and topic modeling were utilized in [19]. Unstructured data from 3 million news articles on Reuters assisted in identifying the 10 major health issues published in news articles from 2007 to 2017. On the contrary, the analysis of user reviews on mobile health applications was prioritized in [8], collecting data from 104 mobile health applications with approximately 88,125 user reviews. The data were categorized based on each comment's functionality (such as usability, content, customer support, and ethics), the polarity concept was divided into three classes, and five machine classifiers were applied. The best accuracy was recorded at 89.42% through Stochastic Gradient Descent.

III. METHODS

This section describes the main model phases, as shown in Figure 1. There are four phases: data acquisition, pre-processing, machine learning classifiers, and model evaluation.

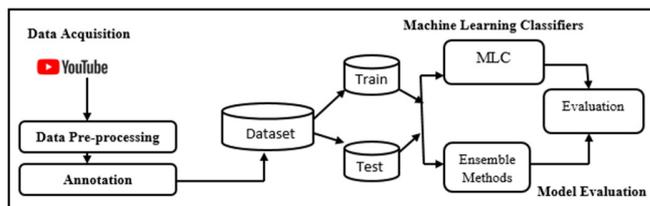


Fig. 1. Model architecture.

A. Phase 1: Data Acquisition

The dataset was collected using Python 3.8 programming language and YouTube API (googleapiclient package), for information extracting from 18 videos related to chewing Khat. The criteria for selecting videos were: published date between 2015-2020, more than 50K views, more than 10K likes, and focus on Arabic speakers. Moreover, some keywords were used to locate the videos, such as Khat, Khat is dangerous, and disadvantages of Khat. The main attributes for the extracted video information were: commenter_id, commenter_name, comment, video_id, number of views, number of likes, and date. Table I shows the dataset description and the minimum and maximum length of user comments.

TABLE I. DATASET DESCRIPTION

Items	Description	Max length	Min length	Average
Negative	1436	427	1	17
Positive	1296			
Total	2732			

The initial step of data preprocessing was carried out, removing English or duplicate comments. The next step, data annotation, was a manual process conducted with the assistance of three annotators that were Ph.D. holders, Arabic native speakers, and computer science specialists. Data annotation classified the comments into negative and positive. Some unrelated, unclear, or ambiguous comments were removed. If two annotators classified comments as either positive or negative then comments were considered respectively, otherwise, the comments were removed.

B. Phase 2: Pre-processing

The natural language pre-processing steps were: data cleaning, tokenization, normalization of Arabic words, lemmatization, deletion of special characters, and removal of repeating characters. Then, the annotation was performed by three annotators into positive and negative. These pre-processing steps increased accuracy by removing "TSHKEEL", "TATWEEL", and "HAMZAH" using Python 3.6 and a package called "tashaphyne".

C. Phase 3: Machine Learning Classifiers

Two types of MLCs were used: classic MLCs and Ensemble Methods (EMs). The MLCs were Linear Regression (LR), Naïve Bayes (NB), Support Vector Machine (SVM), K-nearest Neighbor (KNN), Stochastic Gradient Descent (SGD), and Decision Tree (DT). The EMs were Random Forest (RF), Adaboost (ADA), Gradient Boosting (BG), and xgboost (XG).

D. Phase 4: Model Evaluation

The model's performance was verified using Precision (1), Recall (2), F-Score (3), Accuracy (4), and 5-fold cross-validation on the dataset.

$$\text{Precision} = \frac{\text{Retrieved and Relevant Documents}}{\text{All Retrieved Documents}} \quad (1)$$

$$\text{Recall} = \frac{\text{Retrieved and Relevant Documents}}{\text{All Relevant Documents}} \quad (2)$$

$$F - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

$$\text{Accuracy} = \frac{\text{Number of correct predications}}{\text{Total number of predications}} \quad (4)$$

IV. RESULTS AND DISCUSSION

This section presents the experiments and the results. Two experiments were conducted on the dataset: classic MLCs and EMs. In both experiments, the dataset was divided into 70% for training and 30% for testing, while 5-fold cross-validation was applied.

A. Classic Machine Learning Classifiers

As mentioned above, six classic MLCs were used. The n-gram with Unigram, Bigram, and Trigram was used with all six classifiers to examine their performance. Table II shows the results using Unigram in the six classifiers. It can be observed

that the SVM classifier had the highest performance accuracy (80.12%), while the lowest accuracy was noted on KNN (65%).

TABLE II. CLASSIFIERS' PERFORMANCE USING UNIGRAM

MLCs	Class	Precision	Recall	F-score	Accuracy
LR	Negative	84%	78%	81%	79.39%
	Positive	75%	81%	78%	
NB	Negative	63%	78%	70%	71.95%
	Positive	81%	68%	74%	
SVM	Negative	84%	79%	81%	80.12%
	Positive	76%	82%	79%	
KNN	Negative	94%	60%	73%	65.00%
	Positive	34%	84%	49%	
SGD	Negative	81%	77%	79%	77.80%
	Positive	74%	79%	76%	
DT	Negative	74%	71%	73%	71.34%
	Positive	68%	72%	70%	

Table III shows the MLCs performance using bigram. SVM had the highest accuracy (79.76%), whereas the lowest performance was noted on KNN (65.24%).

TABLE III. CLASSIFIERS' PERFORMANCE USING BIGRAM

MLCs	Class	Precision	Recall	F-score	Accuracy
LR	Negative	83%	78%	80%	78.90%
	Positive	75%	80%	77%	
NB	Negative	63%	78%	70%	72.07%
	Positive	82%	68%	74%	
SVM	Negative	83%	79%	81%	79.76%
	Positive	76%	81%	78%	
KNN	Negative	91%	61%	73%	65.24%
	Positive	37%	81%	51%	
SGD	Negative	79%	78%	79%	77.80%
	Positive	76%	78%	77%	
DT	Negative	77%	69%	73%	70.49%
	Positive	64%	72%	68%	

Table IV shows the MLCs performance using trigram. The highest accuracy was 79.51% using SVM, whereas the lowest was noted again for KNN (65.12%). Figure 2 depicts the overall MLCs results for Unigram, Bigram, and Trigram. Although SVM had the highest accuracy, it was followed closely by both LR and SGD at almost 80%. NB's and DT's accuracies were near 70%, whereas KNN was less accurate.

TABLE IV. CLASSIFIERS PERFORMANCE USING TRIGRAM

MLCs	Class	Precision	Recall	F-score	Accuracy
LR	Negative	83%	78%	80%	78.90%
	Positive	75%	80%	77%	
NB	Negative	63%	78%	69%	71.59%
	Positive	81%	67%	73%	
SVM	Negative	83%	78%	81%	79.51%
	Positive	76%	81%	78%	
KNN	Negative	91%	61%	73%	65.12%
	Positive	37%	80%	51%	
SGD	Negative	76%	78%	77%	76.71%
	Positive	77%	75%	76%	
DT	Negative	78%	67%	72%	69.15%
	Positive	59%	72%	65%	

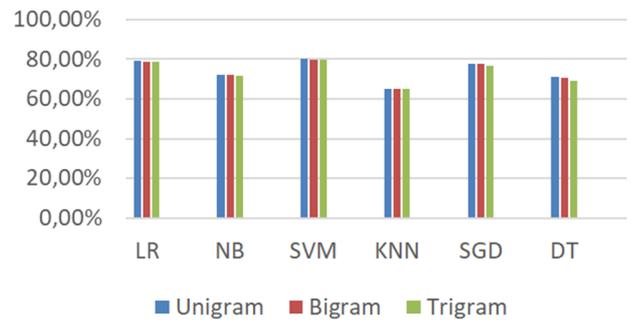


Fig. 2. Accuracy comparison for 6 MLCs on three n-grams.

B. Ensemble Methods

The four mentioned above common methods were used. Table IV shows the accuracy of these EMs. The highest accuracy was recorded for XG, while the lowest was noted for GB. Figure 3 demonstrates the accuracy of the DT classifier using Unigram, Bigram, and Trigram compared to RF. As it can be noted, RF outperformed DT.

TABLE V. ENSEMBLE CLASSIFIERS' PERFORMANCE

EMs	Class	Precision	Recall	F-score	Accuracy
RF	Negative	83%	73%	78%	75%
	Positive	68%	79%	73%	
ADA	Negative	80%	75%	77%	75%
	Positive	71%	77%	74%	
GB	Negative	81%	73%	77%	74%
	Positive	68%	77%	72%	
XG	Negative	82%	76%	79%	77%
	Positive	72%	79%	75%	

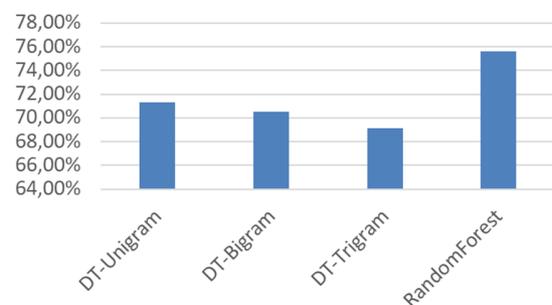


Fig. 3. Accuracy comparison between Decision Tree and Random Forest.

V. CONCLUSION

This paper presented a study on users' opinions on chewing Khat in Yemen and East Africa, using a dataset collected from YouTube comments. Several natural language processing steps were carried on the dataset to get the best performance using classifiers. Classic MLCs and Ems were applied. The best performance in terms of accuracy was recorded when using SVM, followed by Linear Regression. The best accuracy using EMs was recorded for XG.

REFERENCES

[1] J. R. Saura, P. Palos-Sanchez, and A. Grilo, "Detecting Indicators for Startup Business Success: Sentiment Analysis Using Text Data Mining."

- Sustainability*, vol. 11, no. 3, Jan. 2019, Art. no. 917, <https://doi.org/10.3390/su11030917>.
- [2] M. Govindarajan, "Sentiment analysis of restaurant reviews using hybrid classification method," in *Proceedings of 2nd IRF International Conference*, Chennai, India, Feb. 2014, pp. 127–133.
- [3] S. Rani and P. Kumar, "A Sentiment Analysis System to Improve Teaching and Learning," *Computer*, vol. 50, no. 5, pp. 36–43, May 2017, <https://doi.org/10.1109/MC.2017.133>.
- [4] A. Salinca, "Business Reviews Classification Using Sentiment Analysis," in *2015 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, Timisoara, Romania, Sep. 2015, pp. 247–250, <https://doi.org/10.1109/SYNASC.2015.46>.
- [5] U. P. Gurav and S. Kotrappa, "Sentiment Aware Stock Price Forecasting using an SA-RNN-LBL Learning Model," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6356–6361, Oct. 2020, <https://doi.org/10.48084/etasr.3805>.
- [6] J. Carrillo-de-Albornoz, J. R. Vidal, and L. Plaza, "Feature engineering for sentiment analysis in e-health forums," *PLOS ONE*, vol. 13, no. 11, 2018, Art. no. e0207996, <https://doi.org/10.1371/journal.pone.0207996>.
- [7] M. Madhukar and S. Verma, "Hybrid Semantic Analysis of Tweets: A Case Study of Tweets on Girl-Child in India," *Engineering, Technology & Applied Science Research*, vol. 7, no. 5, pp. 2014–2016, Oct. 2017, <https://doi.org/10.48084/etasr.1246>.
- [8] O. Oyeboade, F. Alqahtani, and R. Orji, "Using Machine Learning and Thematic Analysis Methods to Evaluate Mental Health Apps Based on User Reviews," *IEEE Access*, vol. 8, pp. 111141–111158, 2020, <https://doi.org/10.1109/ACCESS.2020.3002176>.
- [9] S. Angelidis and M. Lapata, "Multiple Instance Learning Networks for Fine-Grained Sentiment Analysis," *Transactions of the Association for Computational Linguistics*, vol. 6, pp. 17–31, Aug. 2018, https://doi.org/10.1162/tacl_a_00002.
- [10] Z. Wang, C. S. Chong, L. Lan, Y. Yang, S. B. Ho, and J. C. Tong, "Fine-grained sentiment analysis of social media with emotion sensing," in *2016 Future Technologies Conference (FTC)*, Dec. 2016, pp. 1361–1364, <https://doi.org/10.1109/FTC.2016.7821783>.
- [11] J. Luo, S. Huang, and R. Wang, "A fine-grained sentiment analysis of online guest reviews of economy hotels in China," *Journal of Hospitality Marketing & Management*, vol. 30, no. 1, pp. 71–95, Jan. 2021, <https://doi.org/10.1080/19368623.2020.1772163>.
- [12] C. Yang, H. Zhang, B. Jiang, and K. Li, "Aspect-based sentiment analysis with alternating coattention networks," *Information Processing & Management*, vol. 56, no. 3, pp. 463–478, May 2019, <https://doi.org/10.1016/j.ipm.2018.12.004>.
- [13] M. Song, H. Park, and K. Shin, "Attention-based long short-term memory network using sentiment lexicon embedding for aspect-level sentiment analysis in Korean," *Information Processing & Management*, vol. 56, no. 3, pp. 637–653, May 2019, <https://doi.org/10.1016/j.ipm.2018.12.005>.
- [14] W. Xue and T. Li, "Aspect Based Sentiment Analysis with Gated Convolutional Networks," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia, Jul. 2018, vol. 1, pp. 2514–2523, <https://doi.org/10.18653/v1/P18-1234>.
- [15] F. F. Balahadia, M. C. G. Fernando, and I. C. Juanatas, "Teacher's performance evaluation tool using opinion mining with sentiment analysis," in *2016 IEEE Region 10 Symposium (TENSymp)*, May 2016, pp. 95–98, <https://doi.org/10.1109/TENCONSpring.2016.7519384>.
- [16] I. A. Kandhro *et al.*, "Sentiment Analysis of Students' Comment by using Long-Short Term Model," *Indian Journal of Science and Technology*, vol. 12, no. 8, pp. 1–16, Feb. 2019, <https://doi.org/10.17485/ijst/2019/v12i8/141741>.
- [17] I. A. Kandhro, S. Z. Jumani, F. Ali, Z. U. Shaikh, M. A. Arain, and A. A. Shaikh, "Performance Analysis of Hyperparameters on a Sentiment Analysis Model," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6016–6020, Aug. 2020, <https://doi.org/10.48084/etasr.3549>.
- [18] E. M. Clark *et al.*, "A Sentiment Analysis of Breast Cancer Treatment Experiences and Healthcare Perceptions Across Twitter," *arXiv e-prints*, vol. 1805, p. arXiv:1805.09959, May 2018.
- [19] M. Zolnoori *et al.*, "Mining news media for understanding public health concerns," *Journal of Clinical and Translational Science*, pp. 1–10, Oct. 2019, <https://doi.org/10.1017/cts.2019.434>.
- [20] F. Saeed, W. M.S. Yafouz, M. Al-Sarem, and E. A. Hezzam, "Detecting Health-Related Rumors on Twitter using Machine Learning Methods," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, 2020, <https://doi.org/10.14569/IJACSA.2020.0110842>.
- [21] A. Al-Alimi, E. Halboub, A. K. Al-Sharabi, T. Taiyeb-Ali, N. Jaafar, and N. N. Al-Hebshi, "Independent determinants of periodontitis in Yemeni adults: A case-control study," *International Journal of Dental Hygiene*, vol. 16, no. 4, pp. 503–511, 2018, <https://doi.org/10.1111/idh.12352>.
- [22] M. Hijazi, H. Jentsch, J. Al-Sanabani, M. Tawfik, and T. W. Remmerbach, "Clinical and cytological study of the oral mucosa of smoking and non-smoking qat chewers in Yemen," *Clinical Oral Investigations*, vol. 20, no. 4, pp. 771–779, May 2016, <https://doi.org/10.1007/s00784-015-1569-2>.
- [23] M. A. Al-Duais and Y. S. Al-Awthan, "Association between qat chewing and dyslipidaemia among young males," *Journal of Taibah University Medical Sciences*, vol. 14, no. 6, pp. 538–546, Dec. 2019, <https://doi.org/10.1016/j.jtumed.2019.09.008>.
- [24] B. Kalakonda, S. A. Al-Maweri, H.-M. Al-Shamiri, A. Ijaz, S. Gamal, and E. Dhaifullah, "Is Khat (Catha edulis) chewing a risk factor for periodontal diseases? A systematic review," *Journal of Clinical and Experimental Dentistry*, vol. 9, no. 10, pp. e1264–e1270, Oct. 2017, <https://doi.org/10.4317/jced.54163>.

Stability of the Data-Model Fit over Increasing Levels of Factorial Invariance for Different Features of Design in Factor Analysis

Deyab Almaleki

Department of Evaluation, Measurement, and Research
Umm Al-Qura University
Makkah, Saudi Arabia
damaleki@uqu.edu.sa

Abstract—The aim of this study is to provide an empirical evaluation of the influence of different aspects of design in the context of factor analysis in terms of model stability. The overall model stability of factor solutions was evaluated by the examination of the order for testing three levels of Measurement Invariance (MIV) starting with configural invariance (model 0). Model testing was evaluated by the Chi-square difference test ($\Delta\chi^2$) between two groups, and Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI). Factorial invariance results revealed that the stability of the models was varying over increasing levels of measurement as a function of Variable-To-Factor (VTF) ratio, Subject-To-Variable (STV) ratio, and their interactions. There were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved. For VTF ratios 4:1, 7:1, and 10:1, the models started to show stability over the levels of measurement when the STV ratio was 4:1. Yet, the frequency of stability models over 1000 replications increased (from 77% to 91%) as the STV ratio increased. The models showed more stability at or above 32:1 STV.

Keywords—model stability; factorial invariance; level of measurement invariance; model design

I. INTRODUCTION

Confirmatory Factor Analysis (CFA) is a form of Factor Analysis (FA) that tests hypotheses regarding how well the measured indicator variables represent the number of constructs [1]. CFA is a confirmatory method a researcher can use to examine, evaluate, and/or test a number of hypothesized factors underlying the variance/covariances in a set of measured indicator variables. CFA allows the researcher to test hypothetical and plausible alternative latent variable structures for the observed indicator variance/covariances [2]. More recently, CFA has also been used in exploratory analysis too [3-9]. Three major concerns have emerged repeatedly in the literature related to the use and interpretation of FA in social science research: (a) determining an adequate number of indicator variables to describe the latent trait, (b) factoring a sufficient sample size to have reasonable confidence in the stability of the model estimate, and (c) establishing minimum communality levels to determine which indicator variables can

represent a latent trait, especially in simulation studies [8, 10-13]. FA assumes that the indicator variables used should be linearly related to one another. Otherwise, the number of extracted factors will be the same as the number of original variables [2, 15]. Survey instrument length and the number of variables differ based on discipline, purpose, sample frame, and method of data collection. Recently, the online survey has become an important method of data collection for a variety of reasons (e.g. online surveys are easy to design, conduct, and often they are the only option for data collection). According to SurveyMonkey the median length of its paid surveys was 10 questions [9]. Industry-specific surveys and market-research surveys tend to have more questions while event surveys and just-for-fun surveys tend to be shorter (see Figure 1) [9]. If the length of the survey is about 10 questions or fewer, it can lead to a higher completion rate and increase the likelihood that people will choose to take more of the researcher's surveys in the future. More recent studies of factor analysis do not include the VTF ratio 10:1 in their investigations [5, 12, 14-17], nor the way this number is relative to the sample size or the communality magnitude when factor analysis is conducted.

II. LITERATURE REVIEW

A. Observed Variables

In FA, the observed indicator variables can be viewed as representing a sample of potential variables, all of which measure the same construct or factor [1, 14, 16, 17]. Authors in [18] examined the magnitude of the correlation between the observed variables and the factor components by manipulating sample size, number of variables, number of components, and component saturation. They concluded that the VTF ratio was important for factor stability, with more variables per factor yielding a more stable result. Authors in [19], partially confirmed this conclusion. They found that the necessary minimum indicator variables to attain factor solutions that are adequately stable relative to population factors are dependent on several aspects of any given study, including the level of communality and sample size. Similarly, authors in [20] found that when the VTF ratio increased the factor analysis solution improved.

Corresponding author: Deyab Almaleki

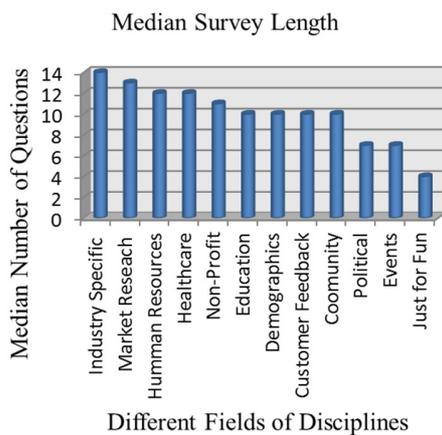


Fig. 1. Median survey length in different fields of disciplines.

The issue of variable sampling has been used extensively in conceptual development but has received almost no empirical evaluation of those that has sampled indicator variables at random from the universe of variables. The assumption of random sampling is useful to minimize sampling issues and for developing generalizability rather than a prescription for applied research procedures. Authors in [21] examined the quality of factor analytic research published between 1999 and 2009 in five leading developmental disabilities journals. They found 35% of the studies used some form of FA. However, the guidelines for using FA were largely ignored and failed to account for levels of overdetermination and communalities among measured variables. Furthermore, the authors in [19] found that there was a lack of validity in some common practice rules used in FA. Thus, anything that influences or changes variance may affect the conclusions related to FA. Researchers should determine an adequate number of indicator variables that is required to produce a stable and precise model in order to describe the latent trait. Authors in [21] investigated the effects of indicator variables on pattern recovery to determine the sufficient number of indicator variables that is likely to produce patterns that closely approximate the population pattern. They reported that the number of indicator variables can strongly affect the degree to which a sample pattern reproduces the population pattern, and that a minimum of three variables per factor is critical. The information about the adequate number of indicator variables that is required to produce a stable and precise model can be used in the design of a study and, retrospectively, in the evaluation of an existing study.

B. Adequate Sample Size

Determining sample size requirements for FA is complicated because it is dependent on other aspects of design, such as VTF and h^2 . Previous studies in FA revealed several approaches that have been used to propose guidelines for the sample size. However, most of these approaches were concerned with identifying either the Subject-To-Variable (STV) ratio or the absolute sample size, regardless of the effect of these rules on WSV. The examples below describe some reported results about sample size in the context of FA. A larger sample size is better than a smaller sample size because

it is minimizing misfit and the probability of errors. In many cases, increasing the sample size may not be possible. In medical research, it is very difficult to collect a large sample of patients suffering from a certain disease [19-21]. Investigating the minimum STV ratio or small absolute sample size to obtain the stability of the model is necessary. Only a very limited number of studies on the role of sample size in FA have investigated real or simulated small sample size. Authors in [17] investigated the minimum sample size necessary to obtain reliable factor solutions under various conditions. They concluded that under the conditions of high communality, high number of observed variables, and small number of factors, FA yields a stable estimate model for sample sizes below 50. Selecting the adequate sample size is an important decision in study design. A researcher must determine how large the sample should be and what is the most appropriate sampling frame. One problem is that the proposed recommendations vary dramatically. Clearly, the wide range in these recommendations causes them to be of rather limited value to empirical researchers. Yet, there is a need to conduct studies examining systematically the model estimate stability latent variable variance with different facets of study Experimental Design (ED) and Sampling Design (SD).

Previous research has investigated the stability of factor solutions by the examination of chi-square value (χ^2) and Overall Model Fit (OMF) indices such as Goodness-of-Fit Index (GFI), Adjusted GFI (AGFI), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Root Mean Square Residual (RMSR) [8, 10, 22-25]. OMF indices examined global measures of data-model fit. Examinations of Measurement Invariance (MIV) (configural, weak, and strong) were used to evaluate model stability. The effects of communality magnitude in FA have been mostly vague. Studies have revealed a varied range of communality magnitude and common practice rules [1, 3, 4, 16, 26, 27]. The communality measures the percent of variance in a given variable explained by the factors. If communalities are high, model stability in the sample data is normally very good [18, 26, 28, 29]. Authors in [20] investigated the quality of factor solutions. They found that when the communalities were high, sample size tended to have less influence on the quality of factor solutions than when communalities are low. Authors in [30] confirmed that communality magnitudes play an important role in determining the adequate sample size. Moreover, authors in [21] found that the communality magnitudes became most relevant in determining the sufficient sample size and the number of variables per component.

III. METHODS

A. Simulation Data

Simulation data are used in social science to answer a particular research question, solve a statistical problem, or improve analysis procedure techniques. Statistical program developers and research designers usually perform simulation data techniques for several reasons: gathering real data may be difficult, time-consuming, expensive, or real data sometimes violate distributional assumptions. Simulation data often lead to greater understanding of an analysis and the results one can

expect from various oddities of real-life data [31]. Simulation may approximate real-world results yet requires less time and effort and gives the researcher a chance to experiment with data under various conditions. Data can be simulated by several methods. The Monte Carlo technique is one popular method that has been used in social science since the 1940s [28]. A Monte Carlo simulation is a numerical technique that can be used to conduct experiments and repeated random sampling to simulate data for a given mathematical model. The key point of the simulation model was the development of the matrices for a 5-factor domain as an example of the factor. The statistical software package SAS was used, and the syntax code was written by the researcher.

B. Estimation Methods

Maximum Likelihood (ML) is the most common method of factor extraction that estimates population values for FA by calculating loading that maximizes the probability of sampling the observed correlation matrix from a population [12] and is often used in CFA. The current study used ML as a method of factor extraction. Authors in [21] concluded that if the data are normally distributed, ML is the best estimation option because it allows the computation of a mixed range of indexes of the goodness of fit of the model. The ML estimation method assumes that the data are independently sampled from a multivariate normal distribution with mean μ and variance-covariance matrix that takes this form: $\Sigma = LL^t + \Psi$, where L is the matrix of factor loadings and ψ is the diagonal matrix of specific variances. Authors in [32] indicated that the ML estimation method is the most precise when the data are continuous and normally distributed, but it does not provide accurate results with ordinal data or when the data violate the assumption of multivariate normality. Table I illustrates the procedure for testing model stability starting with a CFA model relative to a known factor structure for each condition involved in the study separately.

TABLE I. PROCEDURE FOR TESTING MODEL STABILITY

Test name	Symbols	Statistics guidelines
Chi-square value	χ^2	
Tucker-Lewis index	TLI	≥ 0.96 good fit
Comparative fit index	CFI	> 0.95 good fit
Root mean square error of approximation	RMSEA	0.00-0.05 very good fit 0.05-0.08 fair fit 0.08-0.10 mediocre fit

C. Procedure for Testing Stability Across Models

The procedure for testing measurement invariance was performed in order to evaluate the variation over increasingly levels of measurement invariance among models. There are different approaches that could be used to evaluate the measurement invariance among groups. The present study used a Multiple-Group Confirmatory Factor Analysis (MG-CFA) model to test invariance among the levels of STV ratios. Table II illustrates the order for testing measurement invariance starting with configural invariance (model 0). Model testing was evaluated by the chi-square difference test ($\Delta\chi^2$) between two groups [20, 30, 33], and RMSEA, CFI, and TLI were used to evaluate all the model fits. As referenced above, the criteria values suggested in [20, 34] were used in

this study: RMSEA: 0.00 - 0.05 very good fit, CFI > 0.95 good fit, and TLI ≥ 0.96 good fit. Three levels of MIV were tested.

TABLE II. PROCEDURE FOR STABILITY TESTING AMONG MODELS

M		Test name	H_0	Symbol	$\Delta\chi^2$ test	Test statistics guide
M0	Measurement Invariance (MIV)	Configural invariance	$H_0: \lambda_{group1}^1 = \lambda_{group2}^2 = \dots = \lambda_{groupg}^g$	λ : The number of factor patterns across g groups		If $\Delta\chi^2$ NS, the model shows configural factorial invariance in place
M1		Weak measurement invariance	$H_0: \lambda_j^{group1} = \lambda_j^{group2} = \dots = \lambda_j^{groupg}$	λ_j^{group1} : The factor loading of the j th indicator variable in the group	$\Delta\chi^2_{M1-M0}$	If $\Delta\chi^2$ NS, the model shows weak factorial invariance in place
M2		Strong measurement invariance	$H_0: \tau_j^{group1} = \tau_j^{group2} = \dots = \tau_j^{groupg}$	τ : The indicator variables intercept (means) of the j th indicator variable in the group	$\Delta\chi^2_{M2-M1}$	If $\Delta\chi^2$ NS, the model shows strong factorial invariance in place

Configural invariance (M0) indicates that across groups, the pattern of group1 and group2 is equivalent $\lambda_{group1}^1 = \lambda_{group2}^2 = \dots = \lambda_{groupg}^g$ where, λ represents the number of factor patterns across the gth group. Configural invariance at best indicates that the group factors are similar but gives no indication that they hold measurement equivalence. $\Delta\chi^2$ was used to judge configural invariance. For instance, if the χ^2 was not significant, the indicator variables loaded to the same factors across the groups. In other words, there were no differences in factor construct between the groups. Weak Measurement Invariance (M1) indicates that across groups, the corresponding factor loadings are equivalent: $\lambda_j^{group1} = \lambda_j^{group2} = \dots = \lambda_j^{groupg}$, where λ_j^{group1} represents the factor loading of the jth indicator variable in the group. Factor loadings represent the direct effect of the latent construct on each indicator variable, and factor loadings are represented by lambda (λ) [11]. At this level (weak measurement invariance), variables were loaded to the same factors across the group and the factor loadings across groups were equal. $\Delta\chi^2$ was used to judge weak invariance. For instance, if the χ^2 was not significant, the factor loadings across groups were equal. In other words, there were no differences in the factor loading of the indicator variables and their construct between the groups. Strong measurement invariance (M2) means that across groups, corresponding indicators' means are equivalent, $\tau_j^{group1} = \tau_j^{group2} = \dots = \tau_j^{groupg}$, where τ represents the intercept (means) of the jth indicator variable in the group. At this level (strong measurement invariance), indicator variables were loaded to the same factors across the groups. Factor loading, and indicators intercepts across groups were equal. $\Delta\chi^2_{M2-M1}$ was used to judge strong invariance. For instance, if the χ^2 was not significant, the variables intercepts across groups were equal.

D. Design

The main research question was if the stability of the simulated models varies over increasing levels of measurement invariance as a function of the following conditions and their interactions:

- Variable-To-Factor (VTF) ratio
- Subject-To-Variables (STV) ratio
- Communalities magnitude (h²)

MGCFA was used to test the measurement invariance among the levels of STV ratios in order to evaluate model stability. Figure 2 illustrates the study design. Measurement invariance of the STV levels, as shown in Table III, were examined in each cell of the h² *VTF study.

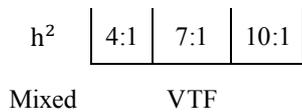


Fig. 2. Design of interaction conditions.

IV. RESULTS

Table III presents the complete findings of measurement invariance for mixed communality among levels of STV over 1000 replications where each significant p-value is marked with "*".

TABLE III. TEST OF FACTORIAL INVARIANCE FOR MIXED COMMUNALITY ACROSS VTF AND STV RATIOS AVERAGED OVER 1000 REPLICATIONS

VTF	Between groups	χ ²	M	Δχ ²	p-value	RMSEA CFI TLI
4:1	STV=2:1 & STV=32:1	369.18	M0	---	0.0301*	0.0201 0.9893 0.9874
		---	M1-M0	---	---	---
		---	M2-M1	---	---	---
	STV=2:1 & STV=16:1	371.33	M0	---	0.0252*	0.0285 0.9791 0.9754
		---	M1-M0	---	---	---
		---	M2-M1	---	---	---
	STV=2:1 & STV=8:1	375.77	M0	---	0.0172*	0.0399 0.9602 0.9530
		---	M1-M0	---	---	---
		---	M2-M1	---	---	---

4:1	STV=2:1 & STV=4:1	386.26	M0	---	0.0065*	0.0573 0.9245 0.9104
		---	M1-M0	---	---	---
		---	M2-M1	---	---	---
	STV=4:1 & STV=32:1	340.29	M0	---	0.2084	0.0117 0.9952 0.9950
		356.07	M1-M0	15.78	0.3968	0.0117 0.9950 0.9951
		370.63	M2-M1	14.56	0.4835	0.0114 0.9950 0.9954
	STV=4:1 & STV=16:1	342.43	M0	---	0.1859	0.0164 0.9907 0.9902
		358.02	M1-M0	15.59	0.4098	0.0163 0.9904 0.9904
		372.60	M2-M1	14.58	0.4820	0.0159 0.9905 0.9910
STV=4:1 & STV=8:1	346.89	M0	---	0.1444	0.0237 0.9825 0.9808	
	362.81	M1-M0	15.92	0.3873	0.0236 0.9819 0.9810	
	377.46	M2-M1	14.65	0.4769	0.0229 0.9820 0.9820	
STV=8:1 & STV=32:1	329.85	M0	---	0.3402	0.0082 0.9971 0.9978	
	345.07	M1-M0	15.22	0.4356	0.0081 0.9970 0.9979	
	360.13	M2-M1	15.06	0.4471	0.0080 0.9969 0.9979	
STV=8:1 & STV=16:1	331.88	M0	---	0.3120	0.0115 0.9946 0.9957	
	346.96	M1-M0	15.08	0.4456	0.0113 0.9945 0.9958	
	362.06	M2-M1	15.1	0.4442	0.0111 0.9945 0.9960	
STV=16:1 & STV=32:1	325.34	M0	---	0.4066	0.0063 0.9980 0.9990	
	340.40	M1-M0	15.06	0.4471	0.0062 0.9980 0.9990	
	355.28	M2-M1	14.88	0.4600	0.0061 0.9980 0.9991	
7:1	STV=2:1 & STV=32:1	1257.15	M0	---	0.0006*	0.0152 0.9942 0.9938
		---	M1-M0	---	---	---

		---	M2-M1	---	---	---	---
STV=2:1 & STV=16:1	1262.68	M0	---	0.0004*	0.0213 0.9888 0.9879	---	---
	---	M1-M0	---	---	---	---	---
STV=2:1 & STV=8:1	1262.65	M0	---	0.0004*	0.0213 0.9888 0.9879	---	---
	---	M1-M0	---	---	---	---	---
STV=2:1 & STV=4:1	1310.82	M0	---	<0.0001*	0.0425 0.9579 0.9544	---	---
	---	M1-M0	---	---	---	---	---
STV=4:1 & STV=32:1	1170.54	M0	---	0.0686	0.0093 0.9975 0.9973	---	---
	1200.54	M1-M0	30	0.4656	0.0092 0.9975 0.9974	---	---
	1230.60	M2-M1	30.06	0.4625	0.0090 0.9975 0.9975	---	---
STV=4:1 & STV=16:1	1176.06	M0	---	0.0549	0.0130 0.9952 0.9949	---	---
	1206.12	M1-M0	30.06	0.4625	0.0128 0.9951 0.9950	---	---
STV=4:1 & STV=8:1	1236.47	M2-M1	30.35	0.4478	0.0127 0.9951 0.9951	---	---
	1191.49	M0	---	0.0279*	0.018 0.9905 0.9898	---	---
STV=8:1 & STV=32:1	---	M1-M0	---	---	---	---	---
	---	M2-M1	---	---	---	---	---
STV=8:1 & STV=16:1	1137.82	M0	---	0.2086	0.0062 0.9986 0.9987	---	---
	1167.98	M1-M0	30.16	0.4574	0.0061 0.9986 0.9987	---	---
	1197.93	M2-M1	29.95	0.4682	0.0061 0.9986 0.9987	---	---
	1143.33	M0	---	0.1773	0.0086 0.9975 0.9975	---	---

		1173.42	M1-M0	30.09	0.4610	0.0085 0.9975 0.9976	---
		1203.61	M2-M1	30.19	0.4559	0.0084 0.9974 0.9976	---
STV=16:1 & STV=32:1	1122.39	M0	---	---	0.3133	0.0044 0.9991 0.9993	---
	1152.28	M1-M0	29.89	0.4712	---	0.0044 0.9991 0.9993	---
		1182.31	M2-M1	30.03	0.4641	0.0044 0.9991 0.9994	---
STV=2:1 & STV=32:1	2640.05	M0	---	<0.0001*	---	0.01241 0.9948 0.9946	---
	---	M1-M0	---	---	---	---	---
		---	M2-M1	---	---	---	---
STV=2:1 & STV=16:1	2654.33	M0	---	<0.0001*	---	0.0174 0.9899 0.9894	---
	---	M1-M0	---	---	---	---	---
		---	M2-M1	---	---	---	---
STV=2:1 & STV=8:1	2681.11	M0	---	<0.0001*	---	0.0244 0.9806 0.9796	---
	---	M1-M0	---	---	---	---	---
		---	M2-M1	---	---	---	---
STV=2:1 & STV=4:1	2748.81	M0	---	<0.0001*	---	0.0345 0.9621 0.9601	---
	---	M1-M0	---	---	---	---	---
		---	M2-M1	---	---	---	---
STV=4:1 & STV=32:1	2464.25	M0	---	<0.0001*	---	0.0076 0.9978 0.9978	---
	---	M1-M0	---	---	---	---	---
		---	M2-M1	---	---	---	---
STV=4:1 & STV=16:1	2478.53	M0	---	<0.0001*	---	0.0108 0.9958 0.9956	---
	---	M1-M0	---	---	---	---	---
		---	M2-M1	---	---	---	---

STV=4:1 & STV=8:1	2505.31	M0	---	0.0059*	0.0154 0.9918 0.9914
	---	M1-M0	---	---	---
	---	M2-M1	---	---	---
STV=8:1 & STV=32:1	2396.55	M0	---	0.1648	0.0047 0.9989 0.9990
	2441.69	M1-M0	45.14	0.4661	0.0047 0.9989 0.9990
	2486.29	M2-M1	44.6	0.4470	0.0046 0.9989 0.9990
STV=8:1 & STV=16:1	2410.83	M0	---	0.1190	0.0069 0.9980 0.9980
	2455.99	M1-M0	45.16	0.4652	0.0068 0.9980 0.9980
	2500.68	M2-M1	44.69	0.4432	0.0067 0.9980 0.9980
STV=16:1 & STV=32:1	2369.77	M0	---	0.2782	0.0034 0.9993 0.9995
	2413.98	M1-M0	44.21	0.5053	0.0033 0.9993 0.9995
	2458.99	M2-M1	45.01	0.4715	0.0033 0.9993 0.9995

VTF (4:1). The examination of the measurement invariance, begins with configural (M0) to weak (M1) and strong (M2). The findings revealed that for VTF ratios (4:1), $\chi^2_{M_0}$ showed statistically significant results when testing configural invariance: (2:1 with 32:1); (2:1 with 16:1); (2:1 with 8:1); and (2:1 with 4:1). Thus, non-invariance was established precluding further invariance testing, e.g., weak, strong, and structural. However, at higher STV ratios, e.g. groups (4:1 with 32:1); (4:1 with 16:1); (4:1 with 8:1); (8:1 with 32:1), (8:1 with 16:1); (8:1 with 32:1); and (16:1 with 32:1), $\chi^2_{M_0}$ was not statistically significant indicating configural invariance was established. Given the presence of configural invariance, testing for weak invariance was conducted. Again, chi-square difference between $\Delta\chi^2_{M1-M0}$ was not statistically significant supporting the hypothesis of weak factorial invariance between the two groups. After weak invariance was supported, the examination of the indicator intercepts was conducted. The results again supported the finding of strong invariance, e.g. the $\Delta\chi^2_{M2-M1}$ was not statistically significant. In conclusion, there were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved as described above.

VTF (7:1). The model begins with configural (M0) to weak (M1) and strong (M2). The findings revealed that for VTF ratio (7:1), $\chi^2_{M_0}$ showed statistically significant results when testing

configural invariance: (2:1 with 32:1); (2:1 with 16:1); (2:1 with 8:1); (2:1 with 4:1); and (4:1 with 8:1). Thus, non-invariance was established precluding further invariance testing, e.g. weak, strong, and structural. However, at higher STV ratios, e.g. groups (4:1 with 32:1); (4:1 with 16:1); (8:1 with 32:1), (8:1 with 16:1); (8:1 with 32:1); and (16:1 with 32:1), $\chi^2_{M_0}$ was not statistically significant indicating configural invariance was established. Given the presence of configural invariance, testing for weak invariance was conducted. Again, the chi-square difference between $\Delta\chi^2_{M1-M0}$ was not statistically significant supporting the hypothesis of weak factorial invariance between the two groups. After weak invariance was supported, the examination of the indicator intercepts was conducted. The results supported again the finding of strong invariance, e.g. the $\Delta\chi^2_{M2-M1}$ was not statistically significant. In conclusion, there were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved as described above.

VTF (10:1). The findings revealed that for VTF ratios (10:1), $\chi^2_{M_0}$ showed statistically significant results when testing configural invariance: (2:1 with 32:1); (2:1 with 16:1); (2:1 with 8:1); (2:1 with 4:1); (4:1 with 8:1); (4:1 with 16:1); and (4:1 with 32:1). Thus, non-invariance was established precluding further invariance testing. However, at higher STV ratios, e.g. groups (8:1 with 16:1); (8:1 with 32:1); and (16:1 with 32:1), $\chi^2_{M_0}$ was not statistically significant indicating configural invariance was established. Given the presence of configural invariance, testing for weak invariance was conducted. Again, chi-square difference between $\Delta\chi^2_{M1-M0}$ was not statistically significant supporting the hypothesis of weak factorial invariance between the two groups. After weak invariance was supported, examination of the indicator intercepts was conducted. The results supported again the finding of strong invariance, e.g. the $\Delta\chi^2_{M2-M1}$ was not statistically significant. In conclusion, there were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved as described above.

Table IV presents a frequency analysis of a chi-square p -values > 0.05 was tested against the null proportion $p = 0.05$ to determine if there was a statistically significant number of invariance failures. The factorial invariance result revealed that the stability of the models was varying over increasing levels of measurement as a function of VTF, STV, and their interactions.

V. DISCUSSION

The study findings refuted some of the guidelines found in the literature, e.g. authors in [18] reported that the sample size was not an important factor in determining model stability, and authors in [35] reported that the STV ratio should be no lower than 5.

The results of the current study revealed that sample size did have a strong effect on stability of the simulated models. For instance, when VTF ratio was 4:1, the mean values related to data-model fit indices were adequate at STV ratio $\geq 4:1$. However, looking at the frequency of rejections based on

conventional thresholds over the 1000 replications depicted a different conclusion. The percentage of stable (invariant) models ranged from 77% at 4:1 STV to 91% at 32:1 STV clearly indicating that larger STV ratios are related to higher stability levels with a model. These findings validated the results in [16] who reported that the percentage of invariance tests varied based on the sample size per group.

TABLE IV. χ^2 FREQUENCY AGAINST THE NULL PROPORTION OF STRUCTURAL MEAN INVARIANCE IN MIXED COMMUNALITY

VTF	Nested models	Invariance status	Frequency & percent	p-value
4:1	STV=4:1 & STV=32:1	Successful	799 79.90	<0.0001*
		Failures	201 20.10	
	STV=4:1 & STV=16:1	Successful	765 76.50	<0.0001*
		Failures	235 23.50	
	STV=4:1 & STV=8:1	Successful	712 71.20	<0.0001*
		Failures	288 28.80	
	STV=8:1 & STV=32:1	Successful	899 89.90	<0.0001*
		Failures	101 10.01	
	STV=8:1 & STV=16:1	Successful	878 87.80	<0.0001*
		Failures	122 12.20	
	STV=16:1 & STV=32:1	Successful	933 93.90	<0.0001*
		Failures	67 6.70	
7:1	STV=4:1 & STV=32:1	Successful	569 56.90	<0.0001*
		Failures	431 43.10	
	STV=4:1 & STV=16:1	Successful	513 51.30	<0.0001*
		Failures	487 48.70	
	STV=8:1 & STV=32:1	Successful	790 79.00	<0.0001*
		Failures	210 21.00	
	STV=8:1 & STV=16:1	Successful	752 75.20	<0.0001*
		Failures	248 24.80	
	STV=16:1 & STV=32:1	Successful	868 86.80	<0.0001*
		Failures	132 13.20	
10:1	STV=8:1 & STV=32:1	Successful	734 73.40	<0.0001*
		Failures	266 26.60	
	STV=8:1 & STV=16:1	Successful	670 67.00	<0.0001*
		Failures	330 33.00	
	STV=16:1 & STV=32:1	Successful	840 84.00	<0.0001*
		Failures	160 16.00	

The findings of the current study do agree with [36] where in some models an STV of 30:1 was needed to produce stable model and minimize the amount of misfit. The study findings also contradict some previous research that investigated the effect of VTF ratio on stability of factor solutions. The authors in [18] concluded that the VTF ratio was important for factor stability with more indicator variables per factor yielding more stable result.

VI. CONCLUSION

In general, this study provided an empirical evaluation of the stability of the data-model fit over increasing levels of factorial invariance for different features of design in FA. The study concluded that the stability of the models was varying over increasing levels of measurement as a function of VTF, STV, and their interactions. There were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved. For VTF ratios of 4:1, 7:1, and 10:1 the models started to show stability over the levels of measurement when the STV ratio was 4:1. Yet, the frequency of stability models over 1000 replications increased (from 77% to 91%) as STV ratio increased. The models showed more stability at or above 32:1 STV.

REFERENCES

- [1] J. S. Tanaka, "How Big Is Big Enough?: Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables," *Child Development*, vol. 58, no. 1, pp. 134–146, 1987, <https://doi.org/10.2307/1130296>.
- [2] J. J. Hox, C. J. Maas, and M. J. Brinkhuis, "The effect of estimation method and sample size in multilevel structural equation modeling," *Statistica Neerlandica*, vol. 64, no. 2, pp. 157–170, 2010, <https://doi.org/10.1111/j.1467-9574.2009.00445.x>.
- [3] T. A. Holster, J. W. Lake, and W. R. Pellowe, "Measuring and predicting graded reader difficulty," *Reading in a Foreign Language*, vol. 29, no. 2, pp. 218–244, Oct. 2017.
- [4] S. Doi, M. Ito, Y. Takebayashi, K. Muramatsu, and M. Horikoshi, "Factorial validity and invariance of the Patient Health Questionnaire (PHQ)-9 among clinical and non-clinical populations," *PLOS ONE*, vol. 13, no. 7, 2018, Art. no. e0199235, <https://doi.org/10.1371/journal.pone.0199235>.
- [5] T. A. Brown, *Confirmatory Factor Analysis for Applied Research, Second Edition*. New York; London: Guilford Publications, 2015.
- [6] P. J. Ferrando and U. Lorenzo-Seva, "Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis - Pere J. Ferrando, Urbanoa, 2018," *Educational and Psychological Measurement*, vol. 78, no. 5, 2018, <https://doi.org/10.1177/0013164417719308>.
- [7] U. Lorenzo-Seva *et al.*, "Psychometric properties and factorial analysis of invariance of the Satisfaction with Life Scale (SWLS) in cancer patients," *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, vol. 28, no. 5, pp. 1255–1264, May 2019, <https://doi.org/10.1007/s11136-019-02106-y>.
- [8] D. Almaleki, "The Precision of the Overall Data-Model Fit for Different Design Features in Confirmatory Factor Analysis," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6766–6774, Feb. 2021, <https://doi.org/10.48084/etasr.4025>.
- [9] D. Almaleki, "Empirical Evaluation of Different Features of Design in Confirmatory Factor Analysis," Ph.D. dissertation, Western Michigan University, Kalamazoo, MI, USA, 2016.
- [10] O. P. John and S. Srivastava, "The Big Five Trait taxonomy: History, measurement, and theoretical perspectives," in *Handbook of personality: Theory and research, 2nd ed*, New York, NY, US: Guilford Press, 1999, pp. 102–138.

- [11] B. Thompson, "Ten commandments of structural equation modeling," in *Reading and understanding MORE multivariate statistics*, Washington, DC, US: American Psychological Association, 2000, pp. 261–283.
- [12] K. Coughlin, "An Analysis of Factor Extraction Strategies: A Comparison of the Relative Strengths of Principal Axis, Ordinary Least Squares, and Maximum Likelihood in Research Contexts that Include both Categorical and Continuous Variables," Ph.D. dissertation, University of South Florida, Tampa, FL, USA, 2013.
- [13] A. H. Blasi and M. Alsuwaiket, "Analysis of Students' Misconducts in Higher Education using Decision Tree and ANN Algorithms," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6510–6514, Dec. 2020, <https://doi.org/10.48084/etasr.3927>.
- [14] A. J. S. Morin, N. D. Myers, and S. Lee, "Modern Factor Analytic Techniques," in *Handbook of Sport Psychology*, 4th ed., M. Tenenbaum and R. C. Eklund, Eds. Hoboken New Jersey, USA: Wiley, 2020, <https://doi.org/10.1002/9781119568124.ch51>.
- [15] R. K. Henson and J. K. Roberts, "Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice," *Educational and Psychological Measurement*, vol. 66, no. 3, pp. 393–416, Jun. 2006, <https://doi.org/10.1177/0013164405282485>.
- [16] A. W. Meade and D. J. Bauer, "Power and Precision in Confirmatory Factor Analytic Tests of Measurement Invariance," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 14, no. 4, pp. 611–635, Oct. 2007, <https://doi.org/10.1080/10705510701575461>.
- [17] J. C. F. de Winter, D. Dodou, and P. A. Wieringa, "Exploratory Factor Analysis With Small Sample Sizes," *Multivariate Behavioral Research*, vol. 44, no. 2, pp. 147–181, Apr. 2009, <https://doi.org/10.1080/00273170902794206>.
- [18] E. Guadagnoli and W. F. Velicer, "Relation of sample size to the stability of component patterns," *Psychological Bulletin*, vol. 103, no. 2, pp. 265–275, 1988, <https://psycnet.apa.org/doi/10.1037/0033-2909.103.2.265>.
- [19] R. C. MacCallum and J. T. Austin, "Applications of Structural Equation Modeling in Psychological Research," *Annual Review of Psychology*, vol. 51, no. 1, pp. 201–226, 2000, <https://doi.org/10.1146/annurev.psych.51.1.201>.
- [20] K. Y. Hogarty, C. V. Hines, J. D. Kromrey, J. M. Ferron, and K. R. Mumford, "The quality of factor solutions in exploratory factor analysis: The influence of sample size, communalities, and overdetermination," *Educational and Psychological Measurement*, vol. 65, no. 2, pp. 202–226, 2005, <https://doi.org/10.1177/0013164404267287>.
- [21] L. R. Fabrigar, D. T. Wegener, R. C. MacCallum, and E. J. Strahan, "Evaluating the use of exploratory factor analysis in psychological research," *Psychological Methods*, vol. 4, no. 3, pp. 272–299, 1999, <https://psycnet.apa.org/doi/10.1037/1082-989X.4.3.272>.
- [22] J. C. Westland, "Lower bounds on sample size in structural equation modeling," *Electronic Commerce Research and Applications*, vol. 9, no. 6, pp. 476–487, 2010, <https://doi.org/10.1016/j.elerap.2010.07.003>.
- [23] S. Zitzmann and M. Hecht, "Going Beyond Convergence in Bayesian Estimation: Why Precision Matters Too and How to Assess It," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 26, no. 4, pp. 646–661, Jul. 2019, <https://doi.org/10.1080/10705511.2018.1545232>.
- [24] R. Jacobucci, A. M. Brandmaier, and R. A. Kievit, "A practical guide to variable selection in structural equation modeling by using regularized multiple-indicators, multiple-causes models," *Advances in Methods and Practices in Psychological Science*, vol. 2, no. 1, pp. 55–76, 2019, <https://doi.org/10.1177/2515245919826527>.
- [25] A. B. Costello and J. Osborne, "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis," *Practical Assessment, Research, and Evaluation*, vol. 10, 2005, Art. no. 7, <https://doi.org/10.7275/jyj1-4868>.
- [26] G. P. Brooks and G. A. Johanson, "TAP: Test Analysis Program," *Applied Psychological Measurement*, vol. 27, no. 4, pp. 303–304, Jul. 2003, <https://doi.org/10.1177/0146621603027004007>.
- [27] X. An and Y.-F. Yung, "Item Response Theory: What It Is and How You Can Use the IRT Procedure to Apply It." SAS Institute Inc, 2014.
- [28] B. S. Everitt, "Multivariate analysis: The need for data, and other problems," *The British Journal of Psychiatry*, vol. 126, no. 3, pp. 237–240, 1975.
- [29] P. M. Bentler and D. G. Bonett, "Significance tests and goodness of fit in the analysis of covariance structures," *Psychological Bulletin*, vol. 88, no. 3, pp. 588–606, 1980, <https://doi.org/10.1037/0033-2909.88.3.588>.
- [30] G. D. Garson, "StatNotes: Topics in Multivariate Analysis," *North Carolina State University*. <https://faculty.chass.ncsu.edu/garson/PA765/statnote.htm> (accessed Feb. 02, 2021).
- [31] D. L. Bandalos and P. Gagné, "Simulation methods in structural equation modeling," in *Handbook of structural equation modeling*, New York, NY, US: The Guilford Press, 2012, pp. 92–108.
- [32] H. W. Marsh, K.-T. Hau, and D. Grayson, "Goodness of Fit in Structural Equation Models," in *Contemporary psychometrics: A festschrift for Roderick P. McDonald*, Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2005, pp. 275–340.
- [33] R. O. Mueller and G. R. Hancock, "Structural equation modeling," in *The reviewer's guide to quantitative methods in the social sciences*, New York, NY, USA: Routledge, 2018, pp. 445–456.
- [34] N. O'Rourke and L. Hatcher, *A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling*, 2nd ed. Cary, NC, USA: SAS Institute, 2014.
- [35] F. B. Bryant and P. R. Yarnold, "Principal-components analysis and exploratory and confirmatory factor analysis," in *Reading and understanding multivariate statistics*, Washington, DC, USA: American Psychological Association, 1995, pp. 99–136.
- [36] I. P. Cobham, "The effects of subject-to-variable ratio, measurement scale, and number of factors on the stability of the factor model," Ph.D. dissertation, ProQuest Information & Learning, USA, 1999.

CP-SDN: A New Approach for the Control Operation of 5G Mobile Networks to Improve QoS

Youcef Djeldjeli

Department of Technology
Nour Bachir University Center
El-Bayadh, Algeria and

Laboratory of Electromagnetism, Photonics and Optonics
Djillali Liabes University of Sidi Bel-Abbès
Sidi Bel Abbès, Algeria
youcef.djeldjeli@univ-sba.dz

Mahdjoub Zoubir

Laboratory of Electromagnetism, Photonics and Optonics
Djillali Liabes University of Sidi Bel-Abbès
Sidi Bel Abbès, Algeria
zoubir.mahdjoub@univ-sba.dz

Abstract—Today, the Software Defined Network (SDN) technology gives more efficiency and flexibility to the 5G mobile networks that are expected to support an enormous amount of data relating to various constrained services. The 5G network should implement newer approaches and technologies that allow supporting the scalability and mobility of the network. The SDN approach consists of decoupling between the control operation and the networking operation, where the control operation is held by the SDN controller that is responsible for defining the management and the control rules. Data forwarding is performed by switches that apply rules defined by their controllers. In the current study, we have proposed and defined a new approach named CP-SDN: Cooperative Protocol-SDN, as an extension to the existing Software Defined Networks, especially when the network experiences saturation due to the huge amount of exchanged data. This congestion may affect the constrained flow and leads to an undesired delay that affects the network Quality of Service (QoS). CP-SDN consists of a cooperation technique between neighboring controllers that aims to relieve the congested centers and redirect the extra flow through neighbors. CP-SDN processing keeps controller databases updated and assures the optimized path for the extra flow when network congestion occurs. The performed simulations on calculating the e-Mbb and M-iOT delay performances for various probability densities show that CP-SDN brings more reliability and efficiency in reducing the transmission delay and overcome the existing SDN scheme. This makes it a prime candidate for the evolved high scalable 5G networks.

Keywords—5G; SDN; CP-SDN; QoS; IoT; m-IoT; eMBB; MFT

I. INTRODUCTION

A huge increase of exchanged data is conducted globally by the deployment of 5G networks. These networks are expected to support several media services that need very high speeds, even Internet of Things (IoT) networks that contain a dense amount of connected devices with different requirements and constraints. 5G networks are deployed to satisfy the need of multiple vertical industries over a shared infrastructure with the use of concepts inclusive of community slicing [1]. Regarding the transport stream held by 5G, video-primarily based totally clients represent a big portion of the data stream carried, with

approximately 79% of cell information [2]. The significant increase of the connected smart devices with greedy smart applications like video streaming, requires a higher spectral efficiency and represents a big challenge of 5G [3]. Authors in [4] illustrate that 5G is expected to reach 4.7 times more data amount than the regular 4G, with more than 12.3 billion users, which exceeds the world population, in the next decade. The new 5G network systems need to deliver capacity a thousand times more than the current cellular 4G systems, synchronizing with the expansion of the application development with personal communications, knowing that mobile devices will reach an immense number timing with the 5G commercial initiation. Owing to the mobile data explosion, current mobile networks suffer from scalability and performance degradation problems[4].

The main goal of the SDN technology resides in separating of the software network management and the hardware based management network (packet forwarding). SDN in its new version leads to an opened standard with optimization in resources with the virtualization development concept. Even LTE disruption level is less critical than SDN deployment. 5G could reap SDN technology features and benefit from combining other compaction technologies such as mm-wave, DAS (Distributed Antenna System) to achieve a 1000-fold capacity, 100-fold rate and 100-fold more active connections than LTE existing networks [5]. SDN, as the up-and-coming technology, dissociates the forwarding logic from the management logic in the existing network. SDN abstracts all management functions into the centralized controller, whereas sanctioning programmability for the network [6, 7]. SDN/NFV is an emerging technology that intends to deploy flexible, more efficient, and cost-effective networks with a high security level [8]. Moreover, SDN philosophy splits the network in two main planes, a management plane and user plane [9]. The management plane is considered as the brain of the network with a centralized controller that runs a software package and that is responsible for controlling the user plane. The latter consists of low-price network devices, such as common switches. The controller runs the Open Flow (OF) protocol in order to configure routing tables and monitor packet statistics

Corresponding author: Youcef Djeldjeli

of the user plane. In the other side, NFV offers operators the virtualization of the network operations and functionalities by implementing standard servers instead of high-expensive and non-standard appliances [10]. Therefore, the SDN design pattern should keep the network infrastructure transparent for the applications level [11].

Various resource allocation and management solutions have been proposed to contribute with 5G and 6G networks. Virtual Network Embedding (VNE) is such a strategy that meets with the knowledge of the management plane for 5G/6G. The Radio Atmosphere Map (REM) is mainly considered as one of the pioneers of psychological feature systems. However, most of these schemes are close to GPP and GTP, and their quality management planes meet with some particular conditions. The future need of 5G networks needs a development of new patterns that satisfy, not only the decoupling between the management plane and the user plane, but should support hot-spots with high-capacity and coverage functionalities [12]. In addition, the virtualization technology aims to emulate the hardware platforms running in software packages. These packages are installed in industrial servers named as COTS (Commercial Off-The-Shelf) servers, which are accessible via Virtual Machines (VMs), instead of hardware dedicated appliances. These VMs can be mounted over different platforms, which makes virtualization more flexible and the technology versatile [13]. More research articles and case studies have been proposed to depict the SDN features and what they can bring to the 5G mobile industry. Authors in [10] proposed a solution for the adjustability support in SDN-based mobile networks. Authors in [14] developed a protocol that evaluates the handover execution time. Moreover, they focused on the open flow protocol to achieve S5/S8 constraint while keeping the main standard rules of SDN. SDN is presently known as a trending technology in the networking domain. In such networks, the controller runs as a network monitor and defines the forwarding rules of the data flows. Switches perform the forwarding and dropping of packets according to the rules defined by the controller. Thus, the controller should be able to implement a centralized control strategy for every exchanging data flow by using the header packet reporting technique [15].

The main advantage of the proposed approach in the current paper resides in the slicing feature that splits the processing functionality into more than a dedicated core for each service. This increases the delivering flexibility and ensures the QoS requirements, instead of having ad-hoc U/C plane functionalities for each business case. Moreover, this feature will assure high scalability and will lead to augmented efficiency in terms of signaling and latency, by avoiding the unnecessary processing.

II. MODELS AND METHODS

A. SDN for 5G Networks

5G mobile networks are expected to fulfill the users' and applications' increasing requirements in terms of bandwidth and QoS [16]. The Network Function Virtualization (NFV) is taking a part in the new IT world and evolves the networking strategy. Both SDN and NFV can operate together and lead to

the deployment of sophisticated future networks [17]. Despite the growth of several mobile networks in 2020, various challenges should be taken into account: high scalability with IoT implementation, energy constraints for the devices and mobile stations, delay and bandwidth optimization, optimization of resource allocation, reduction of deployment costs, etc. [18]. Although several types of mobile networks will dominate the 2020's, there are also many challenges, such as the extenuating power consumption in devices and base stations, better resource allocation, higher data rates, ensuring lower round trip times, decreasing all costs, developing and boosting the mobility management policies, elasticity, agility, and scalability [18].

SDN is an innovative technology within the networking domain that consists on defining the network functionalities in software packages. It enhances the efficiency of the network configuration and its resources allocation. By the separation between the control plane and the data plane, SDN guarantees the transparency between the network infrastructure and the running applications. Thus, each controller manages various network devices as switches and expects more flexible and intelligent control of the exchanged data [19]. Unlike the trivial pattern networking which experiences a difficulty in the resource allocation and its arrangement, SDN is a favorable candidate for such unmanaged networks by decoupling the management and the user pane [20]. To do so, SDN implements OpenFlow as a communication protocol that is standardized by ONF (Open Networking Foundation) [21]. In SDN, the topology discovery is considered as the masterpiece of the SDN controllers and it enables controlling various applications like data routing, network virtualization, live migration, resource optimization, etc. The network topology is not only intended to determine routing tables for forwarding data, but also to manage the network resources [8, 22]. Within the SDN network, the data plane is responsible for forwarding data between the connected devices and the controllers. Switches running the OF protocol form the information plane, while the controllers are distributed and connected together in order to form the control plane [23]. This new organization may have a positive impact in terms of flexibility and programmability [24]. Thus, SDN is a way to introduce the virtualization within the classical IP networks [25], offering a new vision and gathering new features in terms of knowledge and deployment. Furthermore, mobile networks can in turn benefit from SDN advantages, leading to enhanced network control and management [26]. At the front haul part of the network, the present SDN technology does not define a clear vision due to the particular characteristics of this part in power consumption and cost of the elements that spread on a defined geographical area [27].

Figure 1 illustrates the common SDN Architecture. It can be seen that the network is divided into two main planes. The connected devices that constitute the data plane are responsible for forwarding the network traffic. To do so, the open flow enabled-switches should follow switching rules and strategies received from the SDN controller that forms the control plane.

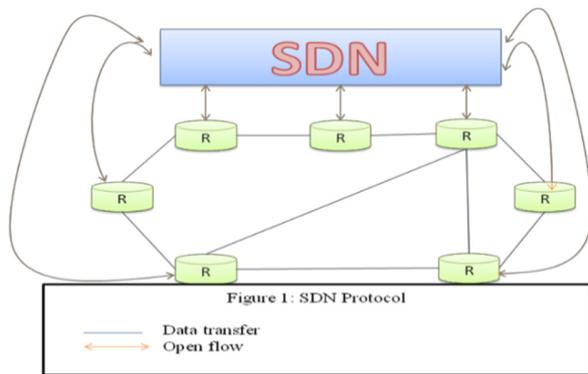


Fig. 1. SDN architecture.

B. Switching in SDN

SDN enables more flexibility in the routing operation than a traditional mobile network. Unlike the existing cellular networks where the traffic passes through the middle boxes using the same routing path, the SDN strategy offers a flexible routing, where only the necessary traffic is conducted via the middle boxes reducing their complexity. Moreover, this strategy allows a dynamic network topology with a flexible routing path that fulfills the change of the network behavior and the user needs [28]. To do so, data packets are forwarded by switches according to an existing matching rule. Otherwise, the concerned switch sends a packet-in message to the controller which determines a forwarding rule for the present packet. Then, the controller sends a flow-mod message to the switches concerned by the new route in order to update their routing tables. Various studies focused on the pipeline MFT process. In [29], authors proposed a framework that divides a large flow table into a number of smaller ones. An SDN switch implementing such method can be seen in [28]. It forwards all the data coming from mobile terminals and m-IOT devices implementing the respective rules, enabling connectivity between the edge devices, and empowering the transport traffic between the core network and other edge clouds [30]. Other studies [12, 30] focused on implementing learning techniques to improve the control performance while selecting the optimized links into the satellite communication system.

C. CP-SDN

This section describes the proposed framework, as a new scheme for enhancing the control flexibility of the existing SDN networks.

CP-SDN is considered as an extension to the existing SDN technology. CP-SDN aims to optimize the delay and enhance the flexibility in the control operation when the network experiences saturation. This may happen when the network is very dense and supports a huge amount of traffic data, as it is expected in 5G technology. The network may contain millions of fixed and mobiles devices, or even IoT systems. The idea consists on using two SDN controllers instead of one, as shown in Figure 2. In the normal case, both controllers are responsible for defining and broadcasting the switching rules to the switches that belong to them. When the network experiences saturation, the saturated controller forwards the superfluous data packets to its CP-SDN neighboring controller in order to

assure load balancing in terms of control. The second controller will be responsible of defining rules to assure the forwarding of the remaining data belonging to the first controller. This aims to relieve the saturated controller, to avoid packet loss and to optimize the data forwarding process. CP-SDN is a proactive protocol that allows controllers to periodically exchange their switching rules for updating the control operation.

D. CP-SDN Control

In SDN, the controller is considered as the core of the system, by gathering the traffic requirements and the status of the network devices. Based on this information, it defines the control rules to the data plane in order to provide an optimized and efficient network for forwarding data services [30]. In case of saturation due to the amount of exchanged data, the controller may take more time to perform the control operation and thus, the transmission delay may increase. This represents one of the main problems experienced by an SDN network. Thereby, real-time services will be affected and the network QoS will be degraded. CP-SDN as a proposed scheme, takes part when the network experiences a saturation issue. As depicted in Figure 2, the CP-SDN consists on liaising between two or more controllers in order to unload the saturated point. The CP-SDN protocol operates between the two SDN centers by allowing the controllers to exchange control information and update their rules. Thus, both centers will have the whole information of the connected data planes.

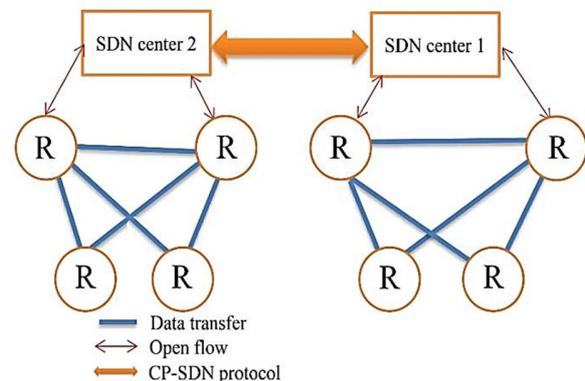


Fig. 2. CP-SDN architecture.

E. CP-SDN Control Switching

The SDN controller is updated to track each TCP connection state established between the switches belonging to its data plane [31]. The switching in CP-SDN is performed in two separate levels. The first level is between the SDN switches, while the second level is between the SDN controllers. CP-SDN implements an algorithm based on two tables (i,j). The first table contains the MFT pipeline processing data for the first level (between switches), and the second table contains the respective data for the second level. As illustrated in Figure 3, the SDN controllers are updated by the same control databases allowing each controller to take decisions about the traffic rules of data packets exchanged in their belonging data planes. Furthermore, in case of an overload of a

part of the network, the controllers should relieve each other and decide about the optimized path to route data packets from source to destination. The diagram illustrated in Figure 4, shows the CP-SDN functioning. Thus, in case of a constrained traffic such as a real-time service, the data will not be queued due to saturation, but processed by the relieved controller. By this way, network optimization is guaranteed and a certain network QoS level for the constrained traffic is assured.

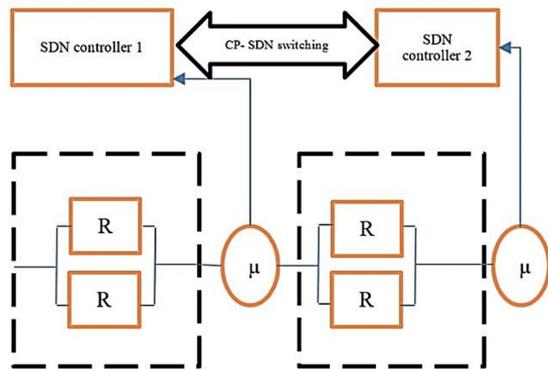


Fig. 3. CP-SDN switching.

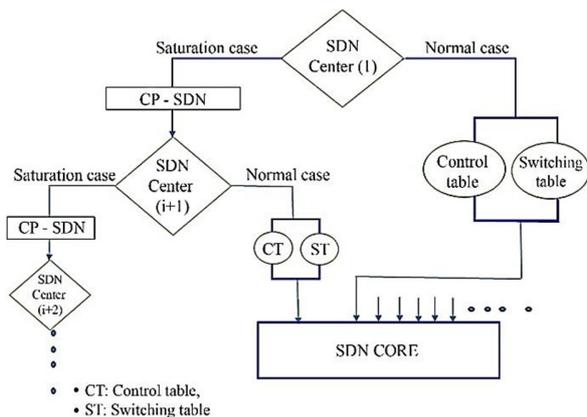


Fig. 4. CP-SDN operational diagram.

III. RESULTS AND DISCUSSION

Let us consider a sub-urban cell covering area $A=2\text{km}^2$ and having on average a $\beta_1=1000$ eMBB (enhanced Mobile Broadband) users, with only a percentage $p_A=0.5, 0.6$, and 0.7 active at a time, and an average of a $\beta_2= 1000$ active m-IoT (massive IoT) devices, in each respective service slice. The slicing should be deployed across all network domains, i.e. a network slice instance consists of network slice subnet instances from different domains [32]. The data rates are kept at 50Mbps for eMBBs and 100kbps for m-IoT, following the standards [31]. We take $\mu=1024\text{pkt/ms}$ with a length $L=2\text{kbytes}$. These values are taken from one node from our network. Figure 5 shows the eMBB packet control time in a single cell by varying number of eMBB slices (n) for the three probabilities $p_A=0.5, 0.6$ and 0.7 .

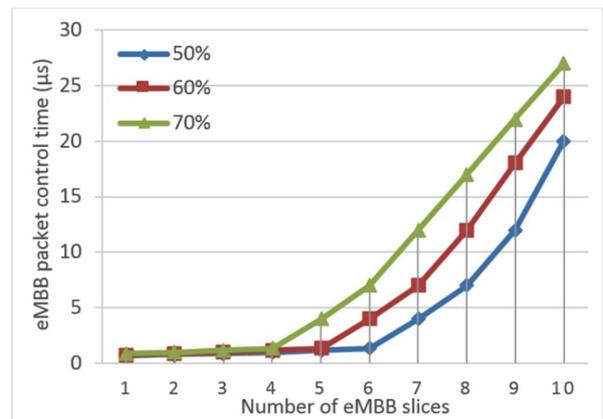


Fig. 5. eMBB delay performance for different numbers of slices.

The eMBB delay performance for different numbers of slices shows the control time of the eMBB packets based on the number of slices (n). It can be seen that for n from 1 to 4, almost the same time is needed, which is $<1\mu\text{s}$. Starting from $n=4$, the case $p=0.7$ shows a different behavior that indicates a start of saturation, where it shows a continuous increase, to reach $27\mu\text{s}$ for $n=10$. This needed control time may lead to network latency. The same behavior is shown in the case of $p=0.6$ and 0.5 , where an increase is noticed from $n=6$ and $n=5$, to reach 24 and $20\mu\text{s}$ respectively for $n=10$.

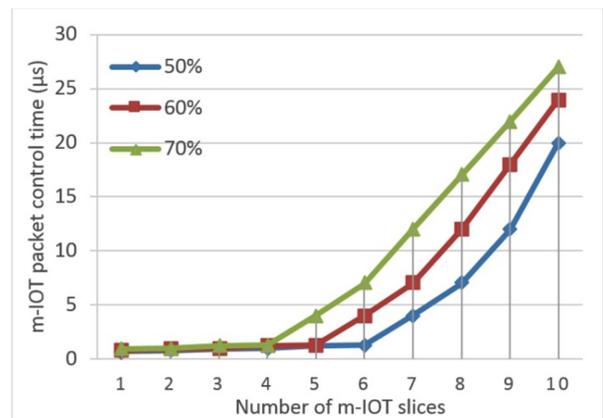


Fig. 6. m-IoT delay performance for different numbers of slices.

Figure 6 illustrates the m-IoT packet control time in a single cell by varying number of m-IoT slices (n) for three probabilities $p_A=0.5, 0.6$ and 0.7 . The m-IoT delay performance shows the control time of the m-IoT packets based on n . It can be seen that for n from 1 to 4, almost the same time is required, which is $<1\mu\text{s}$. Beyond $n=4$, the case $p_A=0.7$ shows a different behavior exhibiting the start of saturation, where it shows an increase to $27\mu\text{s}$ for $n=10$. This produced control time may affect the network latency. In the same diagram, in the case of $p_A=0.6$ and 0.5 , a growth is noticed from $n=6$ and $n=5$, to respectively reach 24 and $20\mu\text{s}$ for $n=10$.

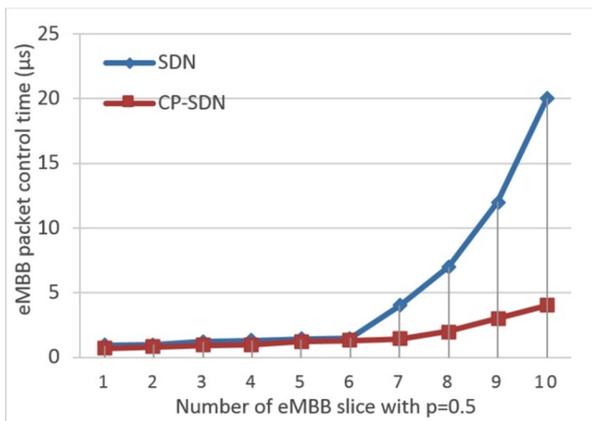


Fig. 7. eMBB delay performance for different values of n using the SDN and the CP-SDN at pA=0.5.

Figure 7 depicts a comparison of the eMBB delay performance for different values of n using SDN and CP-SDN with the probability of 0.5. The simulation results show that both SDN and CP-SDN have the same performances for $n \leq 6$, where the time spent in the control is equal to $1 \mu s$. The CP-SDN shows a slight increase compared to the SDN, for values of n from 6 to 10, where the induced delay in the case of SDN is 5 times greater than in CP-SDN. This enhancement is the result of the use of a second controller, to share the control operation. The difference in the behavior, noticed from $n=6$ (the beginning of the saturation), is due to the collaboration with a second station. The use of a second station is not always enough, where the CP-SDN shows a saturation from $n=8$, which requires a third collaborative station.

Figure 8 illustrates the comparison between the eMBB delay performance variation for different values of n when using SDN and CP-SDN.

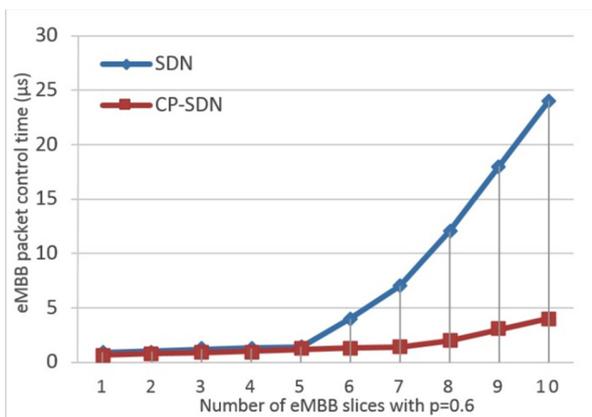


Fig. 8. eMBB delay performance for different values of n using the SDN and the CP-SDN at p=0.6.

The simulation results indicate that both SDN and CP-SDN, express the same behavior for $n \leq 5$, where the time exhausted in the control operation is equal to $1 \mu s$. The CP-SDN shows an obvious improvement compared to the SDN, for values from 5 to 10, where the produced delay in the case of

SDN is 6 times more than the CP-SDN. This efficiency in the case of CP-SDN is the result of using a second station, to divide control. The difference in the behavior noticed from $n=5$ that indicates the beginning of a saturation state, is due to the alliance with the second station. The use of a second station is not usually sufficient, where the CP-SDN shows a saturation from $n=8$, that requires a third corporate station.

In Figure 9, the same results are gained as in Figure 7 and 8 with the change of saturation initiation, which occurs when $n=4$. The second saturation state remains $n=8$, the same as for $pA=0.5$ and $pA=0.6$. As depicted by the various simulations performed during our work, the CP-SDN acts better than SDN, especially when saturation occurs. This happens when the network is dense and supports a huge amount of data exchanged between various segments and supporting multiple services. CP-SDN implementation leads to an optimization in terms of the delay for processing packets and therefore, an enhancement in the QoS of the 5G network. Table I summarizes the CP-SDN performance compared to the basic SDN.

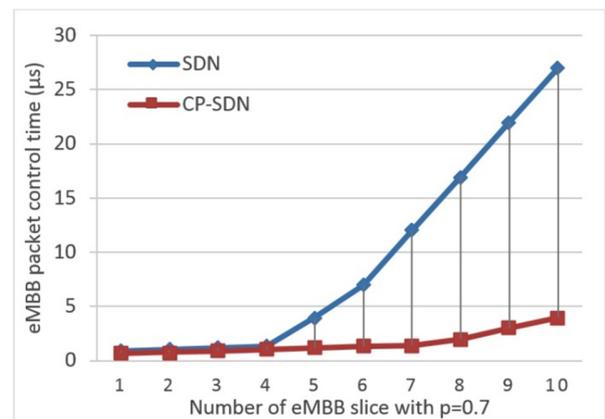


Fig. 9. eMBB delay performance for various values of n using the SDN and the CP-SDN at pA=0.7.

TABLE I. CP-SDN AND SDN PERFORMANCE COMPARISON

Packet control time (µs)	Network probability density					
	0.5		0.6		0.7	
	eMBB	m-IOT	eMBB	m-IOT	eMBB	m-IOT
SDN, current work results	20	19	24	22.5	27	26
SDN [31]	15	10.5	18	13.8	23.5	16.4
Proposed CP-SDN	3.5	3.4	3.6	3.5	4	3.9

In the case of CP-SDN, the control time doesn't exceed $4 \mu s$ for both e-MBB and m-IOT slices, for various probabilities of network density. However, for the SDN case the control time is more significant and it increases when the network becomes dense and experiences congestion. This illustrates the outcome of the CP-SDN for dense networks and its efficiency for real-time traffic, compared to the basic SDN. This quality improvement is directly related to the new mechanism implementation of CP-SDN which relieves controllers by their neighbors. The main drawback of the CP-SDN that may be observed is the increase in the energy needed for the transmission. In the case of cooperation between two or more

controllers, we can perceive more energy consumption in the cooperative station. More energy consumed means more cost charging of power consumption.

IV. CONCLUSION

SDN technology involved in 5G networks has a leading role by giving more flexibility and efficiency in terms of network management and data traffic, knowing that 5G networks are very scalable and connect billions of stationed and mobile devices with different requirements and characteristics. The traffic flow passing through the 5G network has different throughput and priority needs that should be in charge by the networks elements. This is achieved by splitting the network into two planes, the control and the data plane. The proposed extension scheme named CP-SDN, gives more efficiency and optimization to the existing SDN network, by relieving centers that experience saturation due to the huge supported data traffic. The simulation results show that CP-SDN offers a greater outcome for the 5G networks in terms of processing delay optimization and thus can assure a high QoS level, especially for the constrained services. The obtained results make CP-SDN a strong candidate for future dense and mobile 5G networks.

ACKNOWLEDGEMENT

The authors express our sincere thanks to the General Directorate of Scientific Research and Technological Development (DGRSDT) for their support in the development of this work.

REFERENCES

- [1] J. Baranda *et al.*, "Automated deployment and scaling of automotive safety services in 5G-Transformer," in *IEEE Conference on Network Function Virtualization and Software Defined Networks*, Dallas, USA, Nov. 2019, pp. 1–2, <https://doi.org/10.1109/NFV-SDN47374.2019.9039990>.
- [2] J. Aires, P. Duarte, B. Parreira, and S. Figueiredo, "Phased-vCDN Orchestration for flexible and efficient usage of 5G edge infrastructures," in *IEEE Conference on Network Function Virtualization and Software Defined Networks*, Dallas, USA, Nov. 2019, pp. 1–6, <https://doi.org/10.1109/NFV-SDN47374.2019.9040097>.
- [3] A. A. Barakabitze, A. Ahmad, R. Mijumbi, and A. Hines, "5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges," *Computer Networks*, vol. 167, Feb. 2020, Art. no. 106984, <https://doi.org/10.1016/j.comnet.2019.106984>.
- [4] H. Ko, I. Jang, J. Lee, S. Pack, and G. Lee, "SDN-based distributed mobility management for 5G," in *IEEE International Conference on Consumer Electronics*, Las Vegas, U.S.A, Jan. 2017, pp. 116–117, <https://doi.org/10.1109/ICCE.2017.7889250>.
- [5] Z. Zaidi, V. Friderikos, Z. Yousof, S. Fletcher, M. Dohler, and H. Aghvami, "Will SDN Be Part of 5G?," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3220–3258, Oct. 2018, <https://doi.org/10.1109/COMST.2018.2836315>.
- [6] T. Hu, P. Yi, Y. Hu, J. Lan, Z. Zhang, and Z. Li, "SAIDE: Efficient application interference detection and elimination in SDN," *Computer Networks*, vol. 183, Dec. 2020, Art. no. 107619, <https://doi.org/10.1016/j.comnet.2020.107619>.
- [7] M. F. Hyder and M. A. Ismail, "INMTD: Intent-based Moving Target Defense Framework using Software Defined Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5142–5147, Feb. 2020, <https://doi.org/10.48084/etasr.3266>.
- [8] I. H. Abdulqadder, S. Zhou, D. Zou, I. T. Aziz, and S. M. A. Akber, "Multi-layered intrusion detection and prevention in the SDN/NFV enabled cloud of 5G networks using AI-based defense mechanisms," *Computer Networks*, vol. 179, Oct. 2020, Art. no. 107364, <https://doi.org/10.1016/j.comnet.2020.107364>.
- [9] M. H. H. Khairi, S. H. S. Ariffin, N. M. A. Latiff, A. S. Abdullah, and M. K. Hassan, "A Review of Anomaly Detection Techniques and Distributed Denial of Service (DDoS) on Software Defined Network (SDN)," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2724–2730, Apr. 2018, <https://doi.org/10.48084/etasr.1840>.
- [10] J. J. Prados-Garzon, O. Adamuz-Hinojosa, P. Ameigeiras, J. J. Ramos-Munoz, P. Andres-Maldonado, and J. M. Lopez-Soler, "Handover implementation in a 5G SDN-based mobile network architecture," in *IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, Valencia, Spain, Sep. 2016, pp. 1–6, <https://doi.org/10.1109/PIMRC.2016.7794936>.
- [11] X. Huang, P. Shi, Y. Liu, and F. Xu, "Towards trusted and efficient SDN topology discovery: A lightweight topology verification scheme," *Computer Networks*, vol. 170, Apr. 2020, Art. no. 107119, <https://doi.org/10.1016/j.comnet.2020.107119>.
- [12] Q. Long, Y. Chen, H. Zhang, and X. Lei, "Software Defined 5G and 6G Networks: a Survey," *Mobile Networks and Applications*, Nov. 2019, <https://doi.org/10.1007/s11036-019-01397-2>.
- [13] M. Condoluci and T. Mahmoodi, "Softwarization and virtualization in 5G mobile networks: Benefits, trends and challenges," *Computer Networks*, vol. 146, pp. 65–84, Dec. 2018, <https://doi.org/10.1016/j.comnet.2018.09.005>.
- [14] L. M. Contreras, L. Cominardi, H. Qian, and C. J. Bernardos, "Software-Defined Mobility Management: Architecture Proposal and Future Directions," *Mobile Networks and Applications*, vol. 21, no. 2, pp. 226–236, Apr. 2016, <https://doi.org/10.1007/s11036-015-0663-7>.
- [15] H. Wang, H. Xu, C. Qian, J. Ge, J. Liu, and H. Huang, "PrePass: Load balancing with data plane resource constraints using commodity SDN switches," *Computer Networks*, vol. 178, Sep. 2020, Art. no. 107339, <https://doi.org/10.1016/j.comnet.2020.107339>.
- [16] S. Hu, X. Wang, and M. Z. Shakir, "A MIH and SDN-based Framework for network selection in 5G HetNet: Backhaul requirement perspectives," in *IEEE International Conference on Communication Workshop*, London, UK, Jun. 2015, pp. 37–43, <https://doi.org/10.1109/ICCW.2015.7247072>.
- [17] B. Gero *et al.*, "The orchestration in 5G exchange — A multi-provider NFV framework for 5G services," in *IEEE Conference on Network Function Virtualization and Software Defined Networks*, Berlin, Germany, Nov. 2017, pp. 1–2, <https://doi.org/10.1109/NFV-SDN.2017.8169865>.
- [18] C. Bouras, A. Kollia, and A. Papazois, "SDN NFV in 5G: Advancements and challenges," in *20th Conference on Innovations in Clouds, Internet and Networks*, Paris, France, Mar. 2017, pp. 107–111, <https://doi.org/10.1109/ICIN.2017.7899398>.
- [19] X. Cui, X. Gao, and Y. Ma, "An Optimized Controller Placement Algorithm in 5G Based on SDN," in *International Wireless Communications and Mobile Computing*, Limassol, Cyprus, Jun. 2020, pp. 816–819, <https://doi.org/10.1109/IWCMC48107.2020.9148091>.
- [20] S. K. Tayyaba and M. A. Shah, "5G cellular network integration with SDN: Challenges, issues and beyond," in *International Conference on Communication, Computing and Digital Systems*, Islamabad, Pakistan, Mar. 2017, pp. 48–53, <https://doi.org/10.1109/C-CODE.2017.7918900>.
- [21] M. K. Forland, K. Kralevska, M. Garau, and D. Gligoroski, "Preventing DDoS with SDN in 5G," in *IEEE Globecom Workshops*, Waikoloa, USA, Dec. 2019, pp. 1–7, <https://doi.org/10.1109/GCWkshps45667.2019.9024497>.
- [22] A. Hussein, I. H. Elhadj, A. Chehab, and A. Kayssi, "SDN VANETs in 5G: An architecture for resilient security services," in *Fourth International Conference on Software Defined Systems*, Valencia, Spain, May 2017, pp. 67–74, <https://doi.org/10.1109/SDS.2017.7939143>.
- [23] Y. Qin, L. Zhang, F. Xu, and D. Luo, "Interference and Topology-Aware VM Live Migrations in Software-Defined Networks," in *IEEE 21st International Conference on High Performance Computing and Communications*, Zhangjiajie, China, Aug. 2019, pp. 1068–1075, <https://doi.org/10.1109/HPCC/SmartCity/DSS.2019.00152>.

- [24] Z. Wu, Q. Wei, K. Ren, and Q. Wang, "A Dynamic Defense Using Client Puzzle for Identity-Forgery Attack on the South-Bound of Software Defined Networks," *KSI Transactions on Internet and Information Systems (TIIS)*, vol. 11, no. 2, pp. 846–864, 2017.
- [25] D. K. Luong, Y. Hu, J. Li, and M. Ali, "Metaheuristic Approaches to the Joint Controller and Gateway Placement in 5G-Satellite SDN Networks," in *IEEE International Conference on Communications*, Dublin, Ireland, Jun. 2020, pp. 1–6, <https://doi.org/10.1109/ICC40277.2020.9149373>.
- [26] O. Awobuluyi, J. Nightingale, Q. Wang, and J. M. Alcaraz-Calero, "Video Quality in 5G Networks: Context-Aware QoE Management in the SDN Control Plane," in *IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, Liverpool, UK, Oct. 2015, pp. 1657–1662, <https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.250>.
- [27] C. Bouras, A. Kollia, and E. Maligianni, "The techno-economic models for CR and SDN in 5G," in *12th IFIP Wireless and Mobile Networking Conference*, Paris, France, Sep. 2019, pp. 39–46, <https://doi.org/10.23919/WMNC.2019.8881823>.
- [28] P. Iovanna and F. Ubaldi, "SDN solutions for 5G transport networks," in *International Conference on Photonics in Switching*, Florence, Italy, Sep. 2015, pp. 297–299, <https://doi.org/10.1109/PS.2015.7329032>.
- [29] J. Zhang, W. Xie, and F. Yang, "An architecture for 5G mobile network based on SDN and NFV," in *6th International Conference on Wireless, Mobile and Multi-Media*, Beijing, China, Nov. 2015, pp. 87–92, <https://doi.org/10.1049/cp.2015.0918>.
- [30] C. Wang, K. T. Kim, and H. Y. Youn, "PopFlow: a novel flow management scheme for SDN switch of multiple flow tables based on flow popularity," *Frontiers of Computer Science*, vol. 14, no. 6, Jul. 2020, Art. no. 146505, <https://doi.org/10.1007/s11704-019-8417-5>.
- [31] A. Chilwan and Y. Jiang, "Modeling and Delay Analysis for SDN-based 5G Edge Clouds," in *IEEE Wireless Communications and Networking Conference*, Seoul, Korea, May 2020, pp. 1–7, <https://doi.org/10.1109/WCNC45663.2020.9120849>.
- [32] A. Esmaily, K. Kravetska, and D. Gligoroski, "A Cloud-based SDN/NFV Testbed for End-to-End Network Slicing in 4G/5G," in *6th IEEE Conference on Network Softwarization*, Ghent, Belgium, Jul. 2020, pp. 29–35, <https://doi.org/10.1109/NetSoft48620.2020.9165419>.

Complexity and Limitations of GNSS Signal Reception in Highly Obstructed Environments

Arif Hussain

Department of Electrical Engineering
Sukkur IBA University
Sukkur, Pakistan
arif.hussain@iba-suk.edu.pk

Faheem Akhtar

Department of Computer Science
Sukkur IBA University
Sukkur, Pakistan
fahim.akhtar@iba-suk.edu.pk

Zahid Hussain Khand

Department of Computer Science
Sukkur IBA University
Sukkur, Pakistan
zahid@iba-suk.edu.pk

Asif Rajput

Department of Computer Science
Sukkur IBA University
Sukkur, Pakistan
asif.ali@iba-suk.edu.pk

Zeeshan Shaukat

Faculty of Information Technology
Beijing University of Technology
Beijing, China
wakeupzee@yahoo.com

Abstract—Multipath (MP) and/or Non Line-Of-Sight (NLOS) reception remains a potential vulnerability to satellite-based positioning and navigation systems in high multipath environments, such as an urban canyon. In such an environment, satellite signals are reflected, scattered or faded, and sometimes completely blocked by roofs and walls of high-rise buildings, fly-over bridges, complex road structures, etc. making positioning and navigation information inaccurate, unreliable, and largely unavailable. The magnitude of the positioning error depends on the satellite visibility, geometric distribution of satellites in the sky, and received signal quality and characteristics. The quality of the received signal (i.e. its statistical characteristics) can significantly vary in different environments and these variations can reflect in signal strength or power, range measurements (i.e. path delay and phase difference), and frequency, all of which distort the correlation curve between the received signal and receiver-generated replicas, resulting in range errors of tens of meters. Therefore, in order to meet stringent requirements defined for the Standard Positioning Service (SPS), the characterization of distortions that could significantly affect a Global Navigation Satellite System (GNSS) signal is essentially important. The scope of this paper is to detect possible imperfections/deviations in the GNSS signal characteristics that can occur due to MP or NLOS reception and analyze its effects. For this purpose, analysis of fading patterns in received signal strength (i.e. Carrier-to-Noise Ratio and strength fluctuations) is carried out in both clear LOS and high MP environment and then its impact on satellite lock state (i.e. tracking) is assessed. Furthermore, phase fluctuations and range residuals are computed to analyze the effects of path delays. The results show that significant variations can occur in GNSS signal

characteristics in the MP environment that may result in loss of lock event and inaccurate/faulty range measurements.

Keywords—blockage; availability; continuity; NLOS; accuracy

I. INTRODUCTION

Satellite-based navigation systems (i.e. GNSS/GPS) are providing accurate and reliable Positioning, Navigation, and Timing (PNT) services to hundreds of civilian and military applications across the globe. Some of these include road, rail, and aerial transport, precision agriculture, electric power grids, stock exchange systems, autonomous navigation, Intelligent Transportation Systems (ITS), etc. [1, 2]. Although GNSS is now widely used and providing PNT services with an acceptable level of accuracy in clear open-sky views, the precise positioning and navigation is quite a challenging task in urban canyons. In such an environment, the satellite signals are reflected, scattered, fluctuated (i.e. amplitude and phase) and sometimes completely blocked by roofs and walls of high-rise buildings, fly-over bridges and complex road scenarios, making positioning information inaccurate, unreliable and largely unavailable [3]. Specifically, obstructions in urban canyons can affect the availability and accuracy of GNSS based positioning and navigation services in two ways. At first, they can completely block signals from Line-of-Sight (LOS) satellites resulting in a reduced number of visible satellites. Secondly, they can give rise to a phenomenon commonly known as Multipath (MP) that results in biased range measurements. MP refers to the combination of LOS and a number of NLOS signal

Corresponding author: Faheem Akhtar

components reflected off nearby obstacles one or more times before reaching the receiver [2]. The MP effects when severe, can significantly affect the quality of the received signal (i.e. statistical characteristics) and thus result in poor positioning performance of the GNSS receiver. In order to ensure that a receiver in a field has achieved the Required Navigation Performance (RNP) threshold, precise analysis of signal characteristics and noise residuals is essentially important [4] and MP/NLOS detection and error analysis under realistic conditions is often a necessary part of the receiver's performance assessment.

The main contribution of this paper is the analysis of the performance of commercial grade GNSS receivers in urban canyons by detecting the possible anomalies in the GNSS signal reception and analyzing its impact. For this purpose, carefully planned field experiments were conducted on pre-surveyed candidates to acquire the actual GNSS data. For a comprehensive analysis of multipath/NLOS effects, five fundamental GNSS signal parameters were investigated and compared. These parameters include Carrier-to-Noise Ratio (CNR) and its variations, satellite lock time, phase fluctuations, and range residuals. Furthermore, the fading patterns in CNR are analyzed in MP/NLOS conditions and their impact on satellite lock state (i.e. tracking) is assessed.

II. GNSS FUNDAMENTALS AND OBSERVATION MODEL

The positioning and navigation problem considered in GNSS is based on estimating the position of a receiver from signals sent by several satellites containing positioning and timing information. More precisely, by measuring the signal propagation time between a specific satellite and the ground receiver, the receiver is able to compute the distance or geometric range. GNSS signals have very low power, and hence they are prone to several errors and biases. The geometric range measured by the receiver is contaminated by these errors or biases, which is why it is called the pseudorange [5, 6]. The general pseudorange observation [7] equation is given as:

$$\rho = \|P^k - P_u\|_2 + c(dT_u - dT^k) + T^k + I^k + \xi_{MP} + \eta \quad (1)$$

where, ρ represents the pseudorange between the k^{th} satellite and the receiver (u), $P^m = (x^k, y^k, z^k)^T$ is the satellite position available in navigation messages, $P_u = (x_u, y_u, z_u)^T$ represents the receiver position to be estimated, $\|P^k - P_u\|_2 = \sqrt{(x^k - x_u)^2 + (y^k - y_u)^2 + (z^k - z_u)^2}$ is the true distance or actual geometric range between the satellite and the receiver, T^k and I^k are the tropospheric delay and ionospheric delay respectively, $c(dT_u - dT^k)$ is the error due to clock biases whereas T_u is the receiver's clock bias, T^k is the satellite's clock bias, c represents the speed of light, ξ_{MP} is the error factor due to MP/NLOS reception, and η represents the receiver or measurement noise.

III. METHODOLOGY

In an urban area, GNSS signals may arrive at the receiver not only through the direct (LOS) path, but also through several indirect paths due to the presence of different obstructions (high rise buildings, trees, bridges, etc.) in the way of the

signal. The superposition of direct LOS and one or more indirect signals yields a compound signal (i.e. MP signal) [7]. The MP/NLOS can affect the quality of the navigation signal in a number of ways, e.g. it can introduce random fluctuations in amplitude and phase of the received signal or it can result in biased range measurements. In this research, the detailed analysis of MP/NLOS effects is carried out by investigating and comparing the five fundamental parameters of the GNSS signal. These parameters include CNR, variations in CNR, satellite lock time, phase fluctuations, and range residuals as shown in Figure 1. Generally, the MP signal [8, 9] can be modeled as:

$$S_{MP} = A_d \times \cos(\phi) + A_r \times \cos(\phi + \Delta\phi) \quad (2)$$

where A_d and A_r indicate the amplitudes of direct (LOS) and reflected (NLOS) signals respectively, ϕ denotes the phase and $\Delta\phi$ denotes the phase shift in the NLOS signal [10, 11]. The amplitude composite signal (S_{MP}) and phase shift can be estimated as:

$$A_m = \sqrt{A_d^2 + A_r^2 + 2A_d A_r \cos\Delta\phi} \quad (3)$$

$$\Delta\phi_m = \arctan\left(\frac{A_r \sin\Delta\phi}{A_d + A_r \cos\Delta\phi}\right) \quad (4)$$

where $\Delta\phi_m$ and A_m denote the phase shift and amplitude influence by the multipath (i.e. the composite signal).

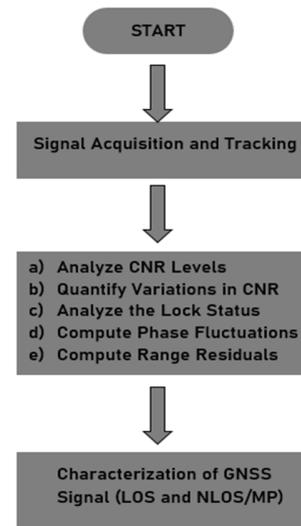


Fig. 1. Navigation signal analysis methodology.

A. Carrier-to-Noise Ratio

CNR is the fundamental quality indication factor of any RF (i.e. GNSS) signal and it represents the strength/power of the received signal. In GNSS, the quality of the signal is directly linked with positioning and navigation accuracy. Strong CNR values with minimal fluctuations result in better positioning performance [12, 13]. However in urban canyons, the CNR tends to be lower and fluctuates around the nominal asynchronously due to superposition of direct LOS and one or more NLOS signals. The CNR can be estimated as:

$$CNR = A_d^2 + A_r^2 + A_\xi^2 + 2A_d^2 A_r^2 A_\xi^2 \cos\Delta\theta \quad (5)$$

It is obvious that CNR can be significantly influenced by MP effects. The multipath can either be destructive or constructive depending upon the phase shifts.

B. Variations Patterns in CNR

As discussed above, in MP environment, the CNR may fluctuate asynchronously around the nominal. These fluctuations and their impact depend upon the severity of the MP effect. The variations in CNR can be estimated as:

$$\Delta CNR = \frac{10 \times (CNR_{max}) - 10 \times (CNR_{min})}{10 \times (CNR_{max}) + 10 \times (CNR_{min})} \quad (6)$$

where CNR_{max} and CNR_{min} are the maximum and minimum CNR values of the received signal respectively.

C. Satellite Lock State

In a GNSS receiver, the signals are initially acquired, tracked and then a navigation message is decoded to estimate the positioning solution. During the tracking stage, the receiver keeps track of the acquired satellite and maintains it in lock state. The MP/NLOS reception can interrupt the signal tracking due to the degraded signal quality and may result in the loss of lock events. The loss of lock interrupts the connection between the satellite and the receiver leading to non-visibility of the satellite even if the satellite is present at the particular time and place. The duration and occurrence of loss of lock events depends upon the strength of the signals. Lower CNR values and/or significant CNR variations may result in frequent loss of locks.

D. Phase Deviations

According to (4), MP signal can introduce phase shifts in the received signal which ultimately contaminates the range

measurements. These shifts/variations in phase can be quantified by computing the standard deviation of phase over a certain interval (60s). In MP free environment, these variations are observed to be negligible.

E. Range Residuals

Pseudorange represents the distance between the satellite and the ground receiver, measured by the estimation of the propagation time of the satellite signal. The propagation time can be affected by several factors like the ionospheric delay, the tropospheric delay, clock biases, and longer travel time due to MP/NLOS resulting in extended pseudorange and thus inaccurate and unreliable positioning solution. MP/NLOS remains an unmodeled and major source of error due to its nondeterministic nature, while other effects can be compensated or modeled. The MP error in the pseudorange can be quantified by range residuals. The range residuals are simply the change in pseudorange from one epoch to the next. In case of a static receiver, the range residuals remain almost zero for direct LOS signals, however this may not be the case in MP environment [14].

IV. FIELD EXPERIMENTS AND DATA COLLECTION

In order to precisely investigate the complexity and limitation of GNSS signal reception in real time, a carefully planned field experiment is conducted on a pre-surveyed candidate site to acquire and record (log) the GNSS signals. The field experiment was performed in the Sukkur IBA University, Pakistan. The site, as shown in Figure 2, has a high degree of realism, being surrounded by buildings and dense trees. In such an environment, there are significant chances of MP and NLOS receptions that can degrade the positioning and navigation accuracy.

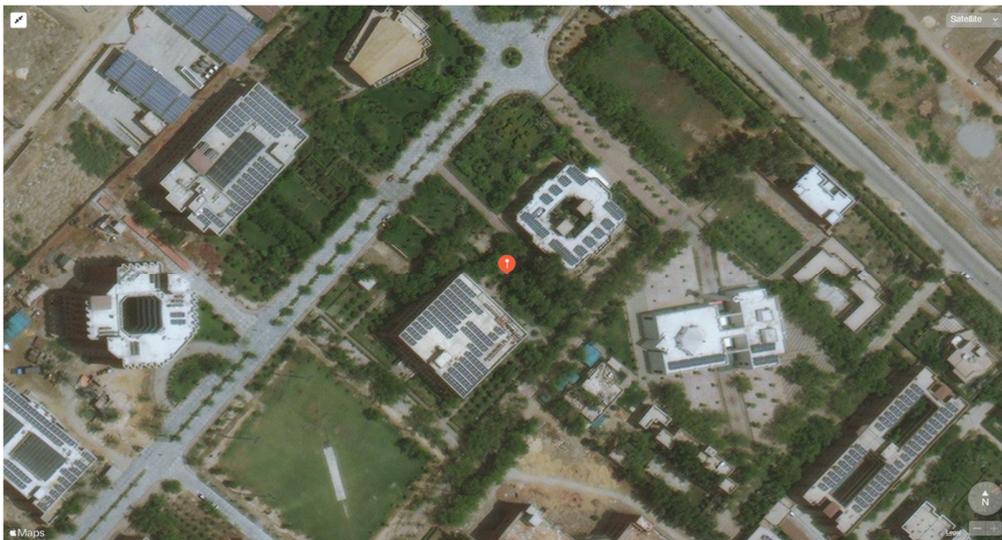


Fig. 2. Experimental site for GNSS raw data/signal acquisition (Google maps, © CNES / Airbus, Maxar Technologies)..

During the field experiment, real time GNSS signals/data were logged by a Septentrio Polarx5S receiver connected to a PolaNt Choke Ring B3/E6 Antenna and the post processing

was done in Matlab. The setup is shown in Figure 3. The Septentrio Polarx5S GNSS receiver is multi-frequency, multi-constellation, capable of receiving 544 hardware channels to

instantaneous track all satellites signals and can log data at 50Hz. The Septentrio receiver locked the GPS (L1C/A, L1P, L2P, L2C, L5), GLONASS (L1, L2, L3) GALILEO (E1, E5ab, AltBoc, E6), BEIDOU (B1, B2, B3), SBAS (L1, L5), QZSS (L1, L2, L5) signals [15]. The letters show the frequency bands related to each particular GNSS system. It has the capability of logging real time GNSS data for all frequencies. The PolaNt Choke Ring antenna is a high precision wideband antenna that is compatible with the PolaRx group and has the capability of receiving all GNSS signals [16].



Fig. 3. Experimental setup for capturing GNSS data.

V. RESULTS AND DISCUSSION

The precise positioning and navigation is highly dependent on the quality/characteristics of the GNSS signal. The comprehensive analysis of GNSS signal characteristics in NLOS or Multipath environment is presented in Figures 4 and 5. This study is performed on actual GNSS data acquired in MP environment in static mode and it is focused on identifying/showing the signal parameters affected by MP/NLOS reception. Figure 4 shows the fundamental signal characteristics (CNR, CNR variations in relationship with elevation angle and lock time) of one satellite PRN 12 for an observation period of 5 hours (from 05:30 to 10:30 UTC). Based on the signal characteristics, the observation period is categorized into two segments: MP/NLOS and LOS. Figure 4(a) shows the CNR of the L1 and L2 signals. It is obvious that CNR is significantly influenced by the MP effect. Under MP conditions, the acquired CNR is much lower (it drops to 30dB) and significantly fluctuates between 30 and 40dB even at good elevation angles ($\theta \geq 20$). Furthermore, the receiver is not able to acquire/track when the CNR drops below 30dB. For better analysis of the severity of the NLOS/MP effects on CNR of the received signal, the variations in the CNR are analyzed and quantified with respect to the elevation angle. Figure 4(b) shows the CNR variations (ΔCNR) in relationship with the elevation angle. Due to the MP/NLOS, significant variations in CNR even at higher elevation angles ($20 \leq \theta \leq 40$) are observed. The results show that when the ΔCNR is greater than 0.3, there are significant chances of the loss of lock event. This can be seen in Figure 4(c). In GNSS, when a connection between the satellite and the receiver is established (i.e. the satellite is in

tracking) then the satellite is in lock state. Figure 4(c) shows that MP/NLOS can interrupt the lock state of the satellite. It can be clearly seen that the lock state of the satellite is frequently interrupted when CNR variations are higher. Frequent loss of locks may affect the continuity of GNSS service and overburden the processing load because the receiver has to reinitiate the acquisition process.

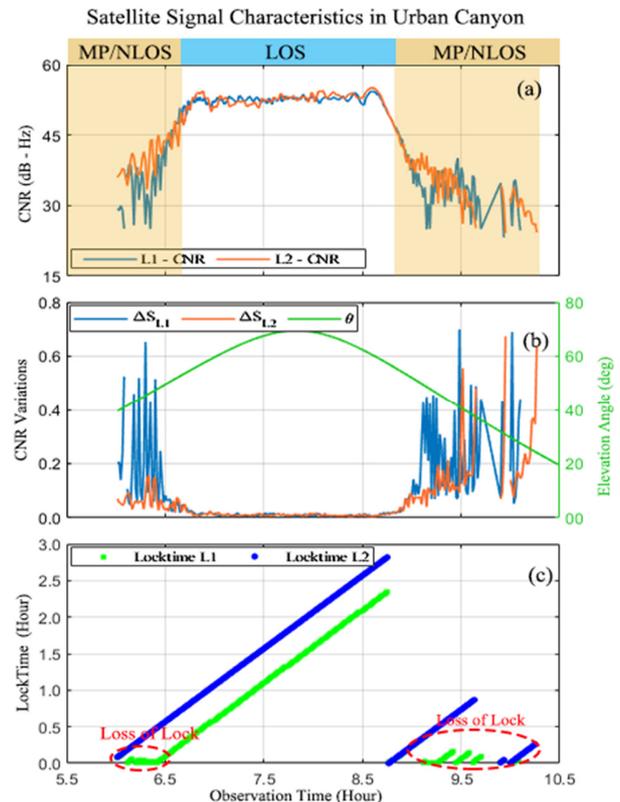


Fig. 4. Satellite signal characteristics in urban canyon.

Multipath (MP) signals are reflected by obstacles one or more times before reaching the GNSS antenna. Since we are estimating distances to the satellites, and that is the foundation of the positions in GNSS, a signal that bounces from an obstruction before it reaches the GNSS antenna causes a problem. If there is a bend in the signal, it will affect the range measurement. These effects can be analyzed by observing phase fluctuations/deviations and computing range residuals. Figure 5 shows the range characteristics (phase deviations and range residuals) of the same satellite, and for the same observation period. It can be observed that MP/NLOS obviously affects both parameters which are highly correlated with the signal characteristics presented in Figure 4. When the MP effect is significant, it may cause rapid fluctuations in phase and/or results in large range residuals. The results show that during the window from 09:00 to 10:00, phase deviations up to 0.3 were observed along with range residual jumps to 20m. This may result in large positioning error.

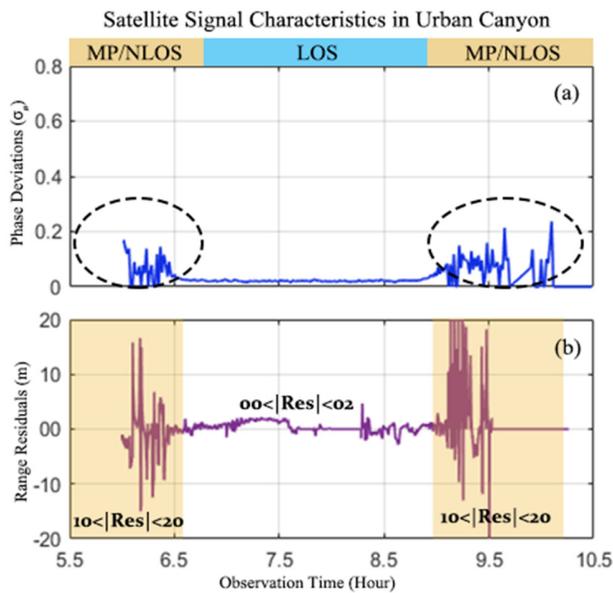


Fig. 5. Satellite range characteristics in urban canyon.

VI. CONCLUSION

In this paper, the characteristics of GNSS signal in both LOS and NLOS conditions were investigated and compared. This study is focused in detecting possible imperfections/deviations in the GNSS signal characteristics that can occur due to MP or NLOS reception only and analyze their effects. Based on the comprehensive analysis of the fundamental characteristics of the signal, it is established that MP/NLOS reception can significantly affect the quality of the received signal. The severe MP effects can result in frequent loss of locks and range errors of more than 10m. Furthermore, the statistical characteristics of direct LOS signal significantly differ from the indirect NLOS signal's.

REFERENCES

- [1] "GNSS Market Report," *EU GNSS*, Jul. 03, 2012. <https://www.gsa.europa.eu/market/market-report> (accessed Feb. 03, 2021).
- [2] J. Guo *et al.*, "Multi-GNSS precise point positioning for precision agriculture," *Precision Agriculture*, vol. 19, no. 5, pp. 895–911, Oct. 2018, <https://doi.org/10.1007/s11119-018-9563-8>.
- [3] C. Stallo *et al.*, "GNSS-based Location Determination System Architecture for Railway Performance Assessment in Presence of Local Effects," in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, Apr. 2018, pp. 374–381.
- [4] D. Lu, S. Jiang, B. Cai, W. Shangguan, X. Liu, and J. Luan, "Quantitative analysis of GNSS performance under railway obstruction environment," in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, Apr. 2018, pp. 1074–1080, <https://doi.org/10.1109/PLANS.2018.8373489>.
- [5] J. B. Tsui, "Satellite constellation," in *Fundamentals of GPS receivers: A software approach*, Hoboken New Jersey, USA: John Wiley & Sons, 2002, pp. 30–50.
- [6] S. Jin, *Global Navigation Satellite Systems: Signal, Theory and Applications*. London, UK: IntechOpen, 2012.
- [7] T. L. Dammalage, "The Effect of Multipath on Single Frequency C/A Code Based GPS Positioning," *Engineering, Technology & Applied*

Science Research, vol. 8, no. 4, pp. 3270–3275, Aug. 2018, <https://doi.org/10.48084/etasr.2206>.

- [8] B. Parkinson, J. S. Jr, P. Axelrad, and P. Enge, *Global Positioning System: Theory and Applications, Volume II*. Washington, DC, USA: AIAA, 1996.
- [9] "Error Sources," *Novatel*, 2015. <https://novatel.com/an-introduction-to-gnss/chapter-4-gnss-error-sources/error-sources> (accessed Feb. 03, 2021).
- [10] Z. Zhang, B. Li, Y. Gao, and Y. Shen, "Real-time carrier phase multipath detection based on dual-frequency C/N0 data," *GPS Solutions*, vol. 23, no. 1, Nov. 2018, Art. no. 7, <https://doi.org/10.1007/s10291-018-0799-6>.
- [11] J. B. Tsui, "Basic GPS concept," in *Fundamentals of Global Positioning System Receivers: A Software Approach*, Hoboken, NJ, USA: John Wiley & Sons, 2005, pp. 7–27.
- [12] P. Xie and M. G. Petovello, "Measuring GNSS Multipath Distributions in Urban Canyon Environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 2, pp. 366–377, Feb. 2015, <https://doi.org/10.1109/TIM.2014.2342452>.
- [13] L. Jordanova, L. Laskov, and D. Dobrev, "Influence of BCH and LDPC Code Parameters on the BER Characteristic of Satellite DVB Channels," *Engineering, Technology & Applied Science Research*, vol. 4, no. 1, pp. 591–595, Feb. 2014, <https://doi.org/10.48084/etasr.394>.
- [14] T. Suzuki, "Mobile robot localization with GNSS multipath detection using pseudorange residuals," *Advanced Robotics*, vol. 33, no. 12, pp. 602–613, Jun. 2019, <https://doi.org/10.1080/01691864.2019.1619622>.
- [15] "Polarx5S," *Septentrio*. <https://www.septentrio.com/en/products/gnss-receivers/reference-receivers/polarx5s> (accessed Feb. 03, 2021).
- [16] "PolaNt Choke Ring B3/E6," *Septentrio*. <https://www.septentrio.com/en/products/antennas/chokering-b3-e6> (accessed Feb. 03, 2021).

Storage Optimization using Adaptive Thresholding Motion Detection

Muhammad Atif Memon

Computer Science Department
Sukkur IBA-IET
Khairpur, Pakistan
muhammadatif.cs@iba-suk.edu.pk

Sajid Khan

Computer Science Department
Sukkur IBA University
Sukkur, Pakistan
sajidkhan@iba-suk.edu.pk

Zahid Hussain Khund

Computer Science Department
Sukkur IBA University
Sukkur, Pakistan
zahid@iba-suk.edu.pk

Faheem Akhtar

Computer Science Department
Sukkur IBA University
Sukkur, Pakistan
fahim.akhtar@iba-suk.edu.pk

Asif Rajput

Computer Science Department
Sukkur IBA University
Sukkur, Pakistan
asifali@iba-suk.edu.pk

Abstract-Data storage is always an issue, especially for video data from CCTV cameras that require huge amounts of storage. Moreover, monitoring past events is a laborious task. This paper proposes a motion detection method that requires fewer calculations and reduces the required data storage up to 70%, as it stores only the informative frames, enabling the security personnel to retrieve the required information more quickly. The proposed method utilized a histogram-based adaptive threshold for motion detection, and therefore it can work in variable luminance conditions. The proposed method can be applied to streamed frames of any CCTV camera to efficiently store and retrieve informative frames.

Keywords-storage optimization; adaptive threshold; motion detection; video mining

I. INTRODUCTION

Image and video processing involves the transmission and storage of large data volumes in the form of frames and images. Several researchers [1-3] worked on reducing bandwidth requirements, whereas others [4] worked on optimizing storage utilization. In today's modern world, a huge amount of data is being generated [5], therefore, extracting the required information from large data volumes is a difficult task, while the storage of large data volumes is an issue [6]. Multimedia mining is a subfield of data-mining that addresses the issues of extracting information such as audio, text, links, images, or digital writing from multimedia files. Many researchers [7-12] worked on video mining, motion detection in CCTV cameras, and crime rate reduction. The importance of CCTV cameras and their placement was discussed in [7], concluding that although placing cameras in both random and expert-suggested sites reduced crime rates, the latter had more impact on crime rate reduction.

Many methods have been proposed for motion detection for security purposes. In [8], MATLAB was used for detecting

motion after getting video images, converting to grayscale, using erosion to remove noise, and finally comparing two consecutive images by using absolute difference. A threshold value between 0.5 and 2 was suggested initially, but further experiments indicated that the ideal threshold value was between 1.3 to 1.5 in most cases. As the use of an image's morphology slows down the processing of every frame, this method cannot be used in real-time video processing. In [9], an external motion sensor and a wireless camera were controlled by an android device. The observer was notified by an email with photos when the motion sensor detected any activity. The suggested system incorporated both an alarm based alert system and a password-based camera power control system. This method's drawbacks were the need for an external sensor for motion detection and the numerous emails in case of continuous motion. A genetic algorithm was utilized for motion detection in [12] to reduce surveillance data storage. The use of genetic algorithms for motion detection resulted in increased complexity and slow frame processing. In [11], motion detection was achieved by calculating the difference between a reference and the current frame. If their difference exceeded a threshold, the frame was marked and transmitted to the observer. Two methods were proposed in [4] for motion-based storage optimization. The first method used a sensor to detect and store the frames containing significant motion. This method had two problems: the use of an external sensor and its range. If the sensor is less sensitive, several sensors are required to cover the whole region of interest. However, if the sensor is more sensitive, it could consider motions out of the region of interest. The second method optimized storage by using neighboring frame differences and a fixed threshold. Fixed threshold-based motion detection will not work well on variable luminance conditions [14-15].

This paper proposes a novel method for optimizing motion detection and video storage without using external sensors. The proposed method uses a single color channel for detecting

motion instead of the RGB. Moreover, an adaptive threshold, calculated every 10,000 frames using a frame's histogram, was utilized to detect motion based on significant changes in neighboring frames. The proposed method was faster than the previously proposed methods as it used one color channel and the frame's histogram. Furthermore, the proposed method is dynamic in nature due to its adaptive threshold.

II. METHODOLOGY

Figure 1 shows a summary of the proposed method, while Figure 2 shows the proposed algorithm's flowchart. The adaptive threshold was recalculated every 10,000 frames, and it was applied to the red layer difference of each frame.

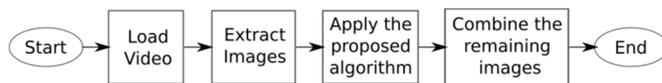


Fig. 1. Summary of the proposed method.

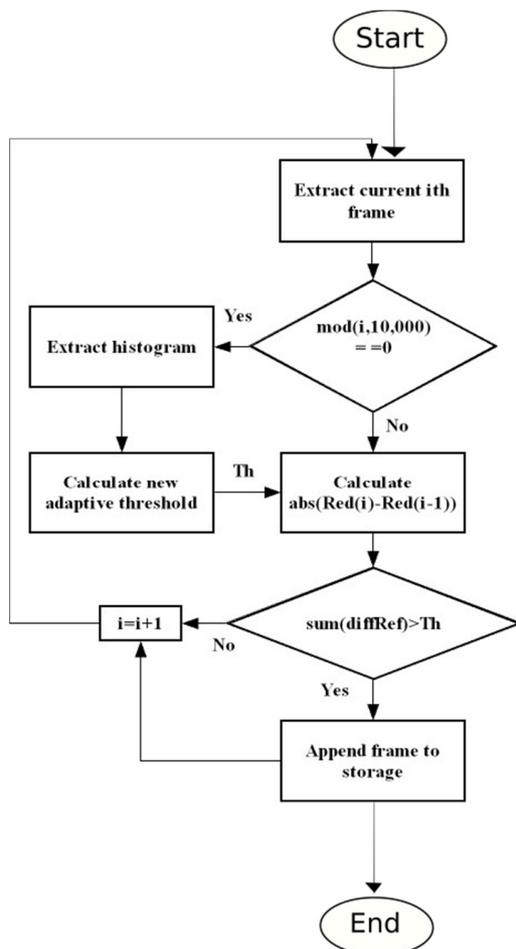


Fig. 2. Flowchart of the proposed algorithm.

A. Using Red Channel for Processing

An RGB frame can be converted to grayscale when the color of the scene is unnecessary. The simpler RGB to grayscale conversion requires two additions and one

assignment per pixel, whereas advanced conversion requires several arithmetic operations [7]. As the proposed method was based on the simple differences of neighboring frames, converting each RGB frame to grayscale would not simplify the algorithm significantly. For example, the conversion of an input frame sized 100×100×3 using exhaustive loops would require 20,000 additions, 30,000 multiplications, and 10,000 assignments per frame. Moreover, the difference calculation requires 10,000 subtractions and 10,000 assignments per frame. So, conversion and difference calculation require in total 30,000 additions/subtractions and 20,000 assignments. On the other hand, using an RGB frame without converting it to grayscale requires 30,000 subtractions and 30,000 assignments. Hence, converting to grayscale would save 10,000 assignments in total for exhaustive differences' calculation.

The proposed method used only the red channel instead of the three RGB or grayscale conversion, as color information was unnecessary. Using the red channel required only 10,000 subtractions and respective assignments for a 100×100×3 sized frame, even if using exhaustive loop-based subtraction instead of vectorization. Moreover, the green channel may be used instead of the red. On the other hand, using the blue channel may not provide good results, as human skin and natural luminance contain fewer blue channel contributions.

B. Adaptive Threshold Calculation

This study aimed to propose a real-time luminance-invariant optimization for motion detection. The luminance of a scene captured by a CCTV camera changes throughout the day due to different atmospheric conditions and daylight changes. If a CCTV camera captures video at 30 frames per second, it takes slightly less than six minutes to capture a ten thousand frame video, while this duration is reduced to three minutes when using 60 frames per second. The first frame after every ten thousand is used for calculating the next threshold. The histogram of this frame is used to calculate the next threshold by using:

$$T_{new} = \frac{\alpha_T}{100} \times B_{80} \times M \times N \quad (1)$$

where M and N are the frame's amount of rows and columns, respectively. B_{80} is the value of k that can be calculated as the index value that minimizes the difference, given by:

$$B_{80} = \min_{k \in \{20,25,30,\dots,125\}} |N_{80} - \sum_{h=-k}^k F_{P+h}| \quad (2)$$

This corresponds to the frequency of the intensity level in the image's histogram, and it is a count of the 80% of the image's pixels obtained by using:

$$N_{80} = M \times N \times 0.8 \quad (3)$$

P in (2) refers to the index of the peak of the histogram. α_T is a parameter that controls the sensitivity of motion detection. A smaller value of α_T will result in more motion detection and vice versa. The possible range of α_T is: $0 < \alpha_T \leq 0.17$. Based on the experience of extensive simulations, the recommended value of α_T was 0.11. It should be noted that all calculations were made for a non-normalized 8-bit image having intensities between 0-255.

C. Application of Threshold

A frame is marked to have motion if:

$$\Sigma(diffRef) > T_{new} \quad (4)$$

where $diffRef$ is given by:

$$diffRef = I_R(i) - I_R(i - 1) \vee \quad (5)$$

where $I_R(i)$ and $I_R(i - 1)$ are the red channel layers for the current and the previous frame respectively. Each frame having motion is stored in the database, where the others are neglected. Figure 3 shows a test case during the daytime displaying an original frame, a grayscaled frame, and the red channel's histogram. Figure 4 shows a test case during the nighttime, displaying an original frame, a grayscaled frame, and the red channel's histogram.

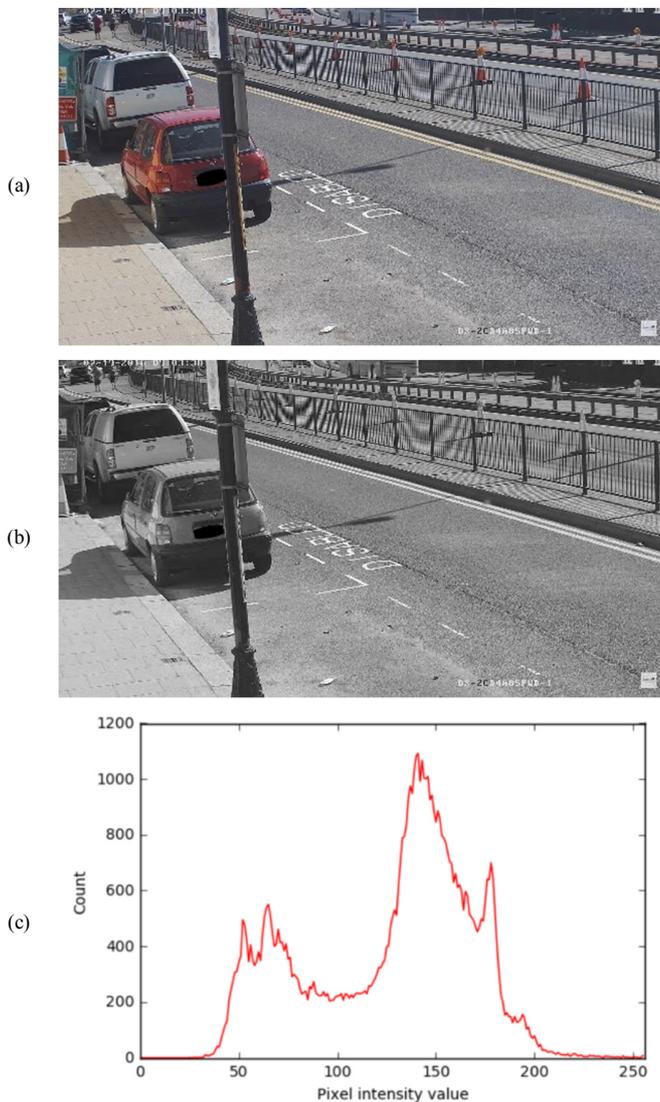


Fig. 3. Daytime case test: (a) Original colored image, (b) grayscaled image, and (c) pixel intensity count of the red channel.

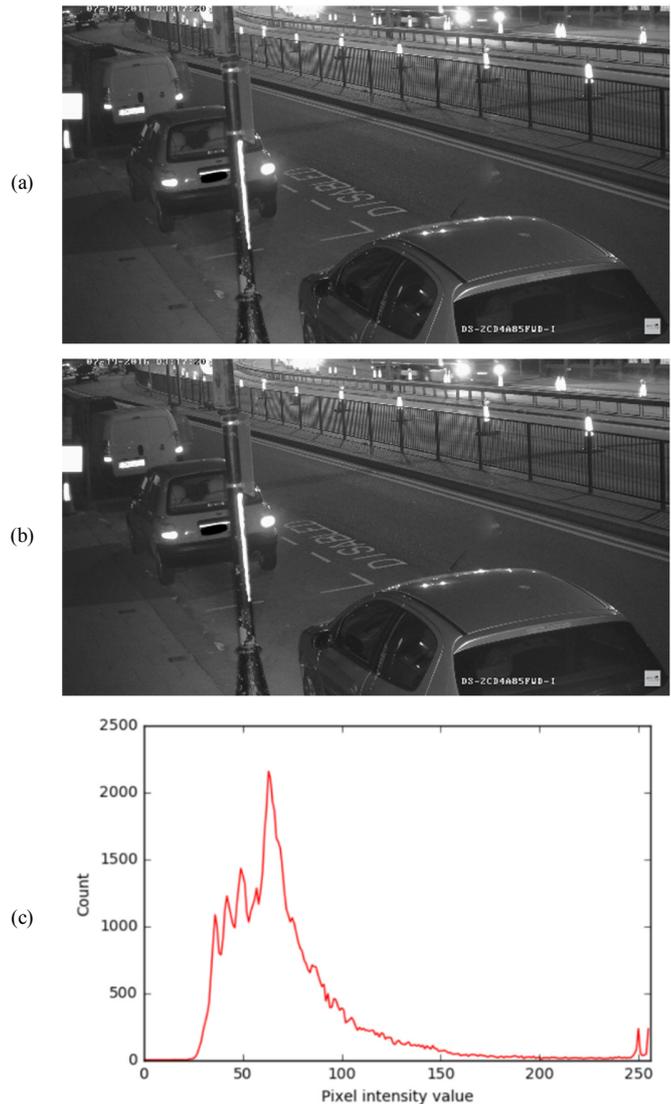


Fig. 4. Nighttime case test: (a) Original colored image, (b) grayscaled image, and (c) pixel intensity count of the red channel.

III. RESULTS AND DISCUSSION

The proposed method was applied on datasets obtained from Sukkur IBA IET (Institute of Emerging Technologies), Khairpur CCTV Cameras, YouTube, and the Venom movie. Results were obtained for both the threshold calculation and storage optimization processes. Figure 3 shows a test case during daytime, used for calculating the threshold. As it can be noted, the histogram was well distributed. Any object moving during the daytime would result in a significant change of intensities and it could be detected using a higher threshold. On the other hand, Figure 4 shows a test case during nighttime. As it can be noted, the histogram is not well-distributed. Any object moving during the nighttime would produce less change of intensities and hence it would require a smaller threshold for motion detection. Moreover, the size of the frame plays a significant role, as a high-resolution frame would require a

higher threshold when compared to a lower-resolution frame of the same scene.

As both 3(a) and 4(a) Figures had 1920×1080 resolution, the total frame's pixels were 2,073,600. The thresholds calculated for the day and nighttime frames were 182,476 and 68,428, respectively. The values of were 62 and 30 for day time and night time frames respectively. When these frames were downscaled to 960×540, the values of thresholds were 45,619 and 17,107 for day and nighttime, respectively.

Table I shows the results of applying the proposed method to the Venom movie and a CCTV camera footage. Venom movie's total duration was 92 minutes, and after applying the proposed method it was reduced to 52 minutes. The total frames of the Venom movie were 11,040, while 6,112 frames were preserved. The frame reduction of this movie was 44.63%.

TABLE I. PERFORMANCE COMPARISON OF THE PROPOSED SCHEME WITH THE ORIGINAL VIDEO

Size	Video sample of venom movie	Video sample of CCTV recording
Actual size	807 MB	14.4 GB
Size after applying the proposed method	220 MB	251 MB

Regarding the application of the proposed method to a CCTV camera stored footage, significant optimization was observed. A video having 268.741 minutes duration was reduced to 32.25 minutes since the CCTV footage had rare motion. The total frames of the input video were 32,249, while after ignoring the redundant frames only 3,714 were preserved. In terms of percentage, the total frame storage was reduced up to 88.4%. The proposed method worked faster than conventional methods by saving many computations due to the use of a single color channel for motion detection. Moreover, the proposed method expedited the calculation process by utilizing frame histograms. Furthermore, the proposed method is dynamic in nature, due to its adaptive threshold and it achieved a significant reduction in storage needs.

IV. CONCLUSION

This paper presented a video optimization method that reduced storage needs by ignoring frames without significant information. A significant informative frame is decided based on the proposed motion detection method. The conventional motion detection methods use fixed thresholds that make them sensitive in variable illumination. For efficient processing, the new threshold was calculated using a histogram every 10,000 frames, considering the illumination of the current scene. The absolute difference was calculated between consecutive frames, and its sum was compared with the threshold to decide whether the frame should be stored in memory or not. Furthermore, the red channel of each frame was used, instead of grayscale or RGB in conventional methods. The proposed method provided efficient video optimization, and it can be applied to any streaming CCTV camera irrespective of the radiance of any scene's light source or the illumination of the CCTV camera.

REFERENCES

- [1] S. Khan and D. Lee, "Efficient deinterlacing method using simple edge slope tracing," *Optical Engineering*, vol. 54, no. 10, Oct. 2015, Art. no. 103108, <https://doi.org/10.1117/1.OE.54.10.103108>.
- [2] S. Khan, D. Lee, M. A. Khan, A. R. Gilal, and G. Mujtaba, "Efficient Edge-Based Image Interpolation Method Using Neighboring Slope Information," *IEEE Access*, vol. 7, pp. 133539–133548, 2019, <https://doi.org/10.1109/ACCESS.2019.2942004>.
- [3] S. Khan *et al.*, "Image Interpolation via Gradient Correlation-Based Edge Direction Estimation," *Scientific Programming*, vol. 2020, Apr. 2020, Art. no. 5763837, <https://doi.org/10.1155/2020/5763837>.
- [4] S. Arora, K. Bhatia, and V. Amit, "Storage optimization of video surveillance from CCTV camera," in *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, Oct. 2016, pp. 710–713, <https://doi.org/10.1109/NGCT.2016.7877503>.
- [5] M. Hazas, J. Morley, O. Bates, and A. Friday, "Are there limits to growth in data traffic? on time use, data generation and speed," in *Proceedings of the Second Workshop on Computing within Limits*, Irvine, CA, USA, Jun. 2016, pp. 1–5, <https://doi.org/10.1145/2926676.2926690>.
- [6] M. H. Padgavankar, "Big Data Storage and Challenges," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 2, pp. 2218–2223, Apr. 2014.
- [7] J. M. Caplan, L. W. Kennedy, and G. Petrossian, "Police-monitored CCTV cameras in Newark, NJ: A quasi-experimental test of crime deterrence," *Journal of Experimental Criminology*, vol. 7, no. 3, pp. 255–274, Sep. 2011, <https://doi.org/10.1007/s11292-011-9125-9>.
- [8] F. N. Tawfeeq, "Real Time Motion Detection in Surveillance Camera Using MATLAB," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 9, pp. 622–626, Sep. 2013.
- [9] M. Rushambwa, T. Chamunorwa, and K. Nyachionjeka, "Real Time Wireless Surveillance System with Motion Detection and Device Control over Internet," *International Journal of Scientific and Research Publications*, vol. 6, no. 3, Mar. 2016, Art. no. 470.
- [10] S. Huang, "An Advanced Motion Detection Algorithm With Video Quality Analysis for Video Surveillance Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 1, pp. 1–14, Jan. 2011, <https://doi.org/10.1109/TCSVT.2010.2087812>.
- [11] A. Upasana, B. Manisha, G. Mohini, and K. Pradnya, "Real Time Security System using Human Motion Detection," *International Journal of Computer Science and Mobile Computing*, vol. 4, no. 11, pp. 245–250, 2015.
- [12] J. C. M. Butil, M. L. F. Magsisi, J. H. Pua, P. K. Se, and R. Sagum, "The Application of Genetic Algorithm in Motion Detection for Data Storage Optimization," *International Journal of Computer and Communication Engineering*, vol. 3, no. 3, pp. 199–202, May 2014.
- [13] C. Saravanan, "Color Image to Grayscale Image Conversion," in *2010 Second International Conference on Computer Engineering and Applications*, Mar. 2010, vol. 2, pp. 196–199, <https://doi.org/10.1109/ICCEA.2010.192>.
- [14] M. V. Sarode and P. R. Deshmukh, "Image Sequence Denoising with Motion Estimation in Color Image Sequences," *Engineering, Technology & Applied Science Research*, vol. 1, no. 6, pp. 139–143, Dec. 2011, <https://doi.org/10.48084/etasr.54>.
- [15] A. B. Altamimi and H. Ullah, "Panic Detection in Crowded Scenes," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5412–5418, Apr. 2020, <https://doi.org/10.48084/etasr.3347>.

Maximum Credible Earthquake Ground Motions with Focus on Site Amplification due to Deep Subsurface

Takashi Nagao

Research Center for Urban Safety and Security
Kobe University
Kobe City, Japan
nagao@people.kobe-u.ac.jp

Abstract—Since an Earthquake Ground Motion (EGM) is amplified from the propagation through the ground, different models are required for each ground type in the seismic design of structures. While the shallow subsurface indicators are used for the classification of ground types, a deep subsurface has a significant impact on the amplification of the EGMs. This study discusses the maximum credible EGMs for seismic design reflecting seismic amplification due to deep subsurface. The design spectra, reflecting the site amplification factor of the target location, are presented by the calculation of the EGMs with the same source and path characteristics and different site amplification factors as recent major Japanese earthquake records have shown, from the perspective of establishing the maximum credible EGMs that may occur in the future at a target site. The present design spectra, which are based on the natural period of a shallow subsurface, are compared with those based on the site amplification factors, considering the effect of deep subsurfaces. Although there are almost no differences in the design spectra with the present design methods according to the surface ground type, the proposed method provides significantly different design spectra for each site amplification factor.

Keywords—maximum credible earthquake ground motion; site amplification factor; deep subsurface; earthquake response analysis

I. INTRODUCTION

Earthquake Ground Motions (EGMs) for seismic design are classified into two categories: moderate EGMs and Maximum Credible EGMs (MCEGMs). The former is used for the assessment of structural serviceability and the latter is used for the assessment of structural safety. MCEGMs are evaluated successfully by adequately taking into account source, path, and site amplification characteristics. A state-of-the-art method is highly applicable to the evaluation of MCEGMs because the said characteristics can be properly considered [1]. However, the application of the state-of-the-art method is often challenging and expensive in practical design. MCEGMs of seismic design codes [2–7] are determined by considering the seismicity and seismic amplification characteristic at the target site, however, conventional methods employed in seismic design codes have been pointed out to overestimate or underestimate MCEGMs at many sites due to the insufficient consideration of source, path, and site amplification characteristics [1]. Of the three characteristics, site

amplification characteristic is known to greatly differ from place to place, thereby precise evaluation of the site amplification characteristic is of utmost importance for MCEGMs. The problem of the conventional methods for the evaluation of site amplification characteristic is that the methods only consider seismic amplification factor by shallow subsurface. The ground comprises of both shallow and deep subsurfaces. The boundary between the two is the engineering bedrock. It has high rigidity and a negligible nonlinear response during major earthquake events. Seismic amplification factors by a deep subsurface are significantly larger than those by a shallow subsurface. Therefore, considering amplification factors due to shallow subsurface alone significantly underestimates the site amplification factors [1, 8–10].

To evaluate the site amplification characteristics of a deep subsurface, various methods utilizing the strong-motion record have been proposed. The spectral ratio of the horizontal component of the target site to the reference site (Standard Spectral Ratio-SSR) [11] and the spectral ratio of the horizontal component to the vertical component at the target site (Receiver Function Method-RF) [12] are taken as examples of the simplified method. Another method is the spectral inversion which separates the source, path, and site characteristics from the earthquake record observed at multiple sites [13–15]. These differences in the evaluation methods of site amplification factors result in variations in the calculated amplification factors. Comparative studies of the differences between these techniques [16–20] have revealed that the spectral inversion method and the SSR method produce almost the same amplitude, while the RF method results in lower amplitude compared with other approaches. It should be noted that the SSR method requires a hard-rock site as a reference location that is in close proximity to the target site. However, identifying the ideal reference site is not always possible, especially at sedimentary sites. Furthermore, surface-rock sites have their specific amplification because of the effects of near-surface weathering or cracking [21]. Consequently, the SSR method may result in an underestimation of the site amplification factor [8, 19].

Considering the above points, it could be said that the spectral inversion method is highly applicable for the precise evaluation of the site amplification factors for arbitrary sites.

Corresponding author: Takashi Nagao

This article discusses MCEGMs considering site-specific amplification factors obtained by spectral inversion, while conforming to the method for establishing the MCEGMs of the Japanese Specifications for Highway Bridges (JSHB) [7].

II. MCEGMs OF JSHB

A. MCEGMs of JSHB and their Problems

The JSHB stipulates six kinds of standard MCEGMs. These encompass two earthquake types (types 1 and 2) and three kinds of ground conditions (types 1, 2, and 3). A type 1 earthquake is the MCEGM from a large-scale inter-plate earthquake in the subduction area, while type 2 is the MCEGM from a shallow crustal earthquake. The type 1 MCEGMs were established considering EGM records, such as the 1923 Kanto Earthquake, the 2003 Tokachi-oki Earthquake, and the 2011 off the Pacific coast of Tohoku Earthquake. The type 2 MCEGMs were established mainly with reference to the EGM records from the 1995 Kobe Earthquake. The standard MCEGMs are multiplied by a regional coefficient in the range of 0.7–1.2 to reflect the seismicity of the site of interest. The ground condition is classified by the natural period of the shallow subsurface using the quarter-wavelength approximation:

$$T_G = 4 \sum_{i=1}^n \frac{H_i}{V_{si}} \quad (1)$$

where T_G is the natural period of the shallow subsurface, and H_i and V_{si} are the thickness and the S-wave velocity of the i -th layer for the shallow subsurface respectively. The ground conditions are classified in accordance with the natural period of the site of interest as follows: type 1 when T_G is 0.2s or less, type 3 when T_G is 0.6s or more, and type 2 otherwise.

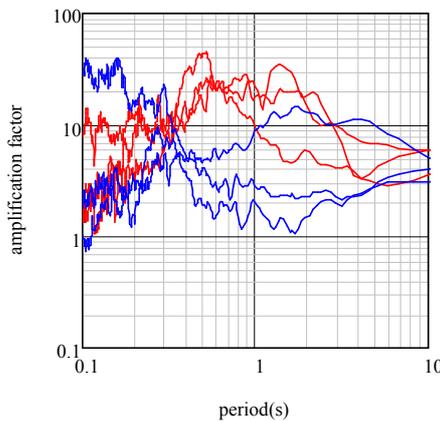


Fig. 1. Site amplification factor comparisons.

The importance of the amplification characteristics by deep subsurfaces is briefly demonstrated. Figure 1 shows the site amplification factors of three type 1 ground sites (ISKH01, MYZ017, and SZOH01) and three type 3 ground sites (KGWH01, TTR005, and YMT006) from strong-motion observation locations in the K-NET [22] and KiK-net [23] networks. The site amplification factor is the amplification

magnitude from the seismic bedrock to the ground surface determined by the spectral inversion [24]. The red line expresses the type 1 ground, while the blue line denotes the type 3 ground. In the 0.6–1.0s period, the amplification factor of type 1 ground with a shallow subsurface natural period of 0.2s or less greatly exceeds the amplification factor of type 3 ground with a shallow subsurface natural period of 0.6s or more. In addition, the amplification factor of the type 1 ground exceeds that of the type 3 ground in the period up to 2.5s, except for MYZ017. This is caused by the differences in the amplification factor due to the deep subsurface. It is evident that the ground type classification of the JSHB is not appropriate.

B. Comparison of JSHB's MCEGMs and Recent Major EGMs

The intensity of the present MCEGMs of the JSHB was compared with recently recorded major EGMs. Three period bands of 0.2–0.6, 0.6–1.0, and 1.0–2.0s were defined by referring to the range of the natural periods of highway bridges with span lengths of 200m or less [7]. Then, records showing remarkably large spectral response acceleration in the above period ranges were extracted from recent major earthquake records. The focus was on the sites where the site amplification factor was calculated by spectral inversion [24].

TABLE I. TARGET SEISMIC WAVES

Period (s)	Earthquake	Station
0.2-0.6	2003 Tokachi-oki Earthquake	HKD086 (2)
	2004 Mid Niigata Prefecture Earthquake	NIG021 (1)
	2008 Iwate-Miyagi Nairiku Earthquake	AKTH04 (1)
	2011 off the Pacific coast of Tohoku Earthquake	IBR003 (1), IBRH11 (2)
	2016 Kumamoto Earthquake	KMMH16 (2)
0.6-1.0	1995 Southern Hyogo Prefecture Earthquake	Port Island (3)
	2000 Western Tottori Prefecture Earthquake	TTRH02 (1), TTR008 (3)
	2003 Tokachi-oki Earthquake	HKD086 (2)
	2004 Mid Niigata Prefecture Earthquake	NIG019 (1)
	2007Noto Hanto Earthquake	ISK005 (3)
	2008 Iwate-Miyagi Nairiku Earthquake	AKTH04 (1)
	2011 off the Pacific coast of Tohoku Earthquake	MYG013 (2)
	2018 Hokkaido Eastern Iburi Earthquake	HKD128 (2)
1.0-2.0	2000 Western Tottori Prefecture Earthquake	TTRH02 (1)
	2003 Tokachi-oki Earthquake	HKD086 (2), HKD098 (1)
	2004 Mid Niigata Prefecture Earthquake	NIG019 (1)
	2007 Niigataken Chuetsu-oki Earthquake	NIG018 (2)
	2011 off the Pacific coast of Tohoku Earthquake	MYG006 (2)

The selected seismic records are presented in Table I. The numbers in parentheses are the ground types according to the JSHB. By evaluating the ground types, it is evident that the K-NET site has the largest P-S logging data down to 20m below ground surface. Consequently, the ground conditions leading to the engineering bedrock are unknown and the ground types

cannot be determined at several locations. For those sites, the ground type was assessed with the following method. At first, it was assumed that the engineering bedrock with an S-wave velocity of 400m/s appears below the deepest point for which P-S logging data have been obtained and that no soil layer with an S-wave velocity exceeding 400m/s appears down to 30m below the ground surface. Next, in accordance with [25], average shear-wave velocity of the top 30m (V_{s30}) was estimated from the average S-wave velocity of less than 20m in the surface layer, and the ground structures from the deepest position of P-S logging to a depth of 30m were evaluated. The EGMs of the observation records in Table I were compared with the MCEGMs of the JSHB in Figure 2. The thin gray line represents the individual EGM, the solid blue line expresses the average, the broken blue line denotes the average and standard deviation EGM. The solid red, black, and green lines are the MCEGMs of ground types 1, 2, and 3 of the JSHB respectively. The damping constant was established at 5% of the standard value.

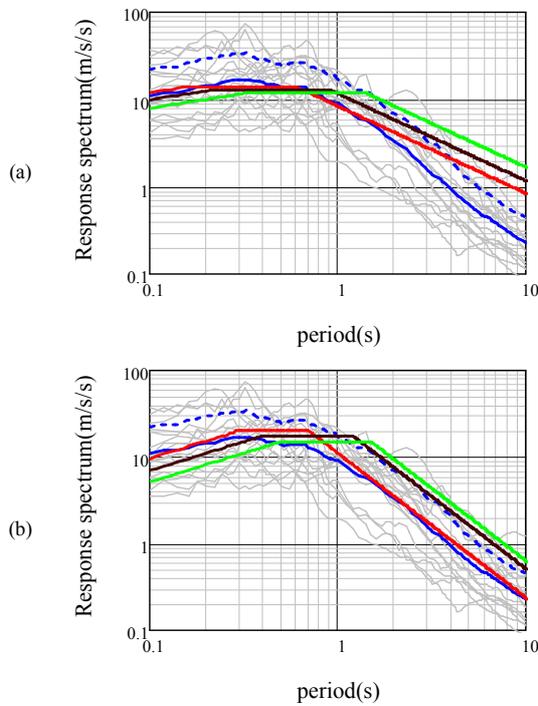


Fig. 2. Comparison of major earthquakes and design spectra: (a) type 1 EGM, (b) type 2 EGM.

From the comparison of the observed EGMs and the MCEGMs of the JSHB, the average of the observed spectra of recent major earthquakes in the period range of less than 1s and the average and standard deviation of the observed spectra in the period range of more than 1s were found to generally correspond to the MCEGMs of the JSHB. However, MCEGMs are not necessarily the largest observed EGMs, and the reference motions are selected from a statistical perspective. Furthermore, if the difference in levels due to the difference in the period bands is considered, it can be said that a highway bridge with a short natural period is a small-scale bridge with a relatively low degree of importance. Thus, an average level is

regarded as reasonable, while a highway bridge with a long natural period is a long-span bridge with a high degree of importance, and an average and standard deviation level is assumed by engineering judgement. As different MCEGMs are established for different earthquake types, the variations in EGMs are compared according to the type of earthquakes in Figure 3. This research focuses on site-specific MCEGMs in terms of amplification. The specific type of earthquake is not considered, and the differences in EGMs between the interplate earthquakes and the crustal earthquakes are not remarkable. Therefore, the variations in earthquake types are not considered.

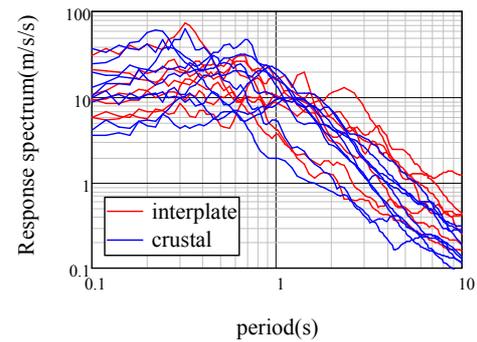


Fig. 3. Differences by earthquake type.

III. METHODS

A. Evaluation Methods of Site-Specific MCEGMs

As described above, the MCEGMs of the JSHB are established with reference to the records of past major earthquakes. This is because earthquakes of the same intensity as historic earthquake events may reoccur in any part of Japan in areas of high seismicity, and many earthquake records for reference in the assessment of MCEGMs have been accumulated. This paper discusses the MCEGMs considering site-specific amplification factors while conforming to the method for establishing the MCEGMs of JSHB. The EGM is determined by the characteristics of the source, path, and site amplification. The observed spectrum is represented by:

$$O(f) = S(f) P(f) G(f) \quad (2)$$

where $O(f)$ is the observed spectrum, $S(f)$ is the source spectrum, $P(f)$ is the path spectrum, $G(f)$ is the site amplification spectrum, and f is the frequency.

This study considers MCEGMs with the same source and path characteristics of past major earthquakes, and with the site amplification factor at the site of interest. The reference points from which seismograms are obtained are represented by subscript R , and the target sites are represented by subscript T . Thus, the seismic ground motion at the site to be examined is obtained by:

$$O_T(f) = O_R(f) \cdot \frac{G_T(f)}{G_R(f)} \quad (3)$$

During major earthquakes, shallow subsurfaces demonstrate nonlinearity. To consider the effect of this

nonlinearity, this study conducted an equivalent-linear earthquake response analysis considering the frequency dependency of the ground's nonlinear characteristics by [26] and evaluated an incident waveform at the engineering bedrock of the reference site. The obtained spectrum was divided by the site amplification factor $G_R(f)$ from the seismic bedrock to the engineering bedrock at the reference site and was multiplied by the site amplification factor $G_T(f)$ from the seismic bedrock to the engineering bedrock at the target site. Finally, an equivalent-linear earthquake response analysis was conducted to determine the $O_T(f)$ at the ground surface. The nonlinear characteristics of the soil used in the earthquake response analysis were taken from [27], while the site amplification factor from the seismic to the engineering bedrock was calculated by [24]. Since this amplification factor is the one from the ground surface to the seismic bedrock, the amplification factor from the ground surface to the engineering bedrock was eliminated by the application of the SH-wave multiple reflection theory.

B. Relationship between the Site Amplification Factor and the Response Spectrum

Figure 4 contains the site amplification factor, the acceleration Fourier amplitude spectrum, and the acceleration response spectrum of the seismic waveforms presented in Table I, extracting HKD098, NIG018, and TTR008 from ground type 1, 2, and 3 sites respectively. The acceleration Fourier amplitude spectrum was smoothed with the use of a Parzen window with a bandwidth of 0.2Hz. The peak frequencies of the Fourier spectrum and the site amplification factor do not necessarily correspond due to the influence of the nonlinearity of the shallow subsurface during strong earthquakes. However, both the Fourier amplitude and the response spectrum reveal a large value in the periodic band with a large site amplification factor in general. Therefore, when discussing the magnitude of the spectrum, the amplitude of the site amplification factor in the relevant periodic band must be considered.

C. Selection of Target Sites

A typical site amplification factor was selected from the following perspectives based on the nationwide site amplification factors determined by [24]. First, three period bands of 0.2–0.6, 0.6–1.0, and 1.0–2.0s were established. The nationwide site amplification factors were classified into large, medium, and small groups for these periodic zones. A total of 100 points were sampled from logarithmically spaced site amplification factors in a period of 0.1 to 10.0s, and Average Site Amplification Factors (ASAF) in the corresponding period bands were calculated. The frequency distributions of the ASAF for the respective bands are presented in Figure 5. The distribution of ASAF can be approximated to a lognormal distribution in every periodic band. The solid blue line in Figure 5 is the probability density function of the lognormal distribution. The cumulative distribution of ASAF was calculated by the lognormal distribution approximation, as shown in Figure 6. Since the distributions of 0.6–1.0 and 1.0–2.0s are very close to the mean and standard deviations, the cumulative distributions for the two bands almost overlap.

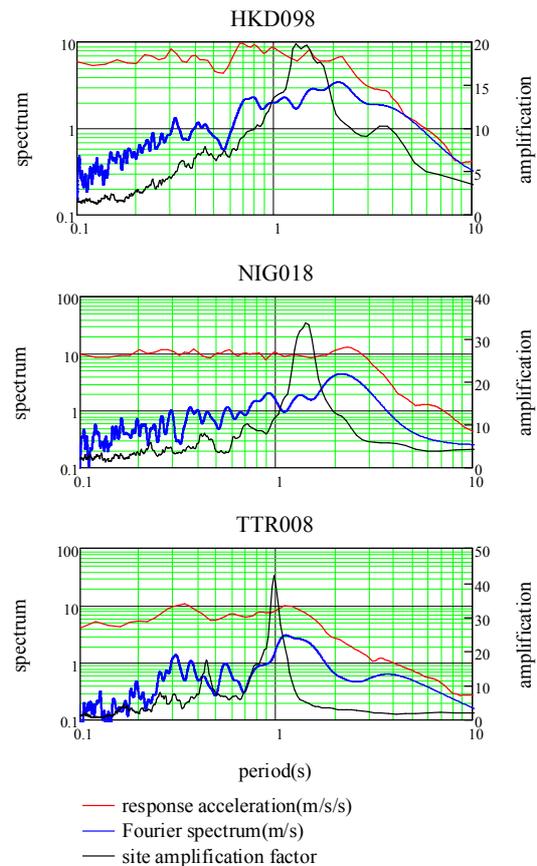


Fig. 4. Relationship between the site amplification factor, the Fourier spectrum, and the response spectrum.

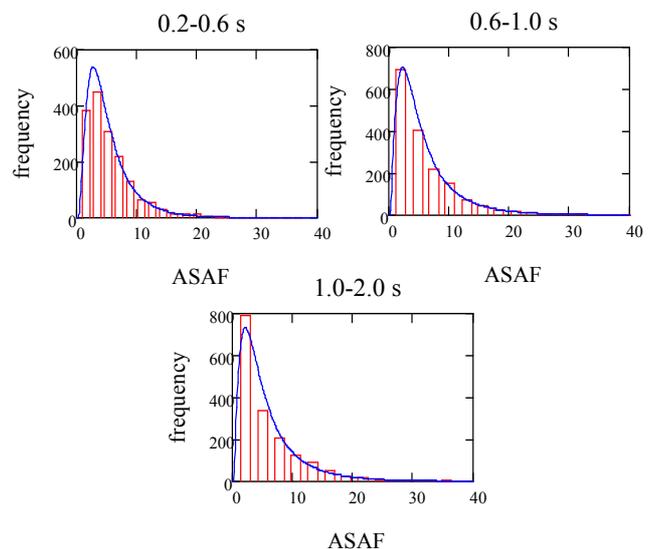


Fig. 5. ASAF frequency distributions.

Based on the results shown in Table II, the range of the cumulative distribution 0–1/3 is classified as the small site amplification factor. The range of 1/3–2/3 is classified as the medium site amplification factor, and the range 2/3–1 is

categorized as the large site amplification factor. Among the presented 1695 nationwide site amplification factor sites [24], 1067 are type 1 ground sites, 439 are type 2 sites, and 81 are type 3 sites. For each ground type, 3 sites were selected from the large, medium, and small ASAF in each periodic zone to eliminate the bias in the region. From this, a total of 81 sites, with 9 sites from each ground type and each periodic zone, were used as objects for the present study.

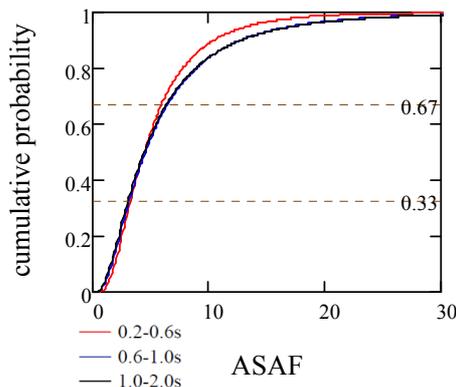


Fig. 6. ASAF cumulative distributions.

TABLE II. ASAF CLASSIFICATIONS

Period (s)	0.2-0.6	0.6-1.0	1.0-2.0
Small	ASAF < 3.30	ASAF < 3.19	ASAF < 3.08
Medium	3.30 ≤ ASAF < 5.88	3.19 ≤ ASAF < 6.39	3.08 ≤ ASAF < 6.30
Large	5.88 ≤ ASAF	6.39 ≤ ASAF	6.30 ≤ ASAF

TABLE III. TARGET SITES

	Type 1	Type 2	Type 3
0.2-0.6s			
small	GIFH24, KYT005, TKS011	CHBH14, MYG007, YMN011	CHB008, KSRH04, TTR005
medium	GNM007, HYG009, SMN006	KYT004, MYG013, NIG018	AKT015, HKD161, KYT002
large	IBUH06, KMM007, MYZ017	FKO015, ISK002, OITH07	ISK005, SBSH03, SZOH42
0.6-1.0s			
small	NGNH20, TCGH07, TTR003	HYGH11, IBRH11, OKY004	KGWH01, TTR005, YMT006
medium	AICH16, HKD134, SMN006	KOC002, TCGH16, WKY006	FKI003, KNG009, SRCH03
large	ISKH01, MYZ017, SZOH26	AOM024, KGS007, OIT016	IBUH03, ISK005, KYTH05
1.0-2.0s			
small	FKIH01, TCGH07, TTR003	AIC010, EHMH09, IWT004	AICH12, HYGH10, TTR005
medium	HKD057, KNG003, OIT007	KMM010, SZO025, TTRH04	FKIH05, ISK005, TTR008
large	ISK003, MYG005, SZOH26	KGS012, MYZ009, SZOH28	FKS020, IBUH03, KGS012

Table III shows a list of the selected points. When the shallow subsurface P-S logging at the selected site did not reach the engineering bedrock, the subsurface structure was determined in the same manner described in the previous section. The site amplification factors of the selected sites are shown in Figure 7. The red, blue, and brown lines correspond to the large, medium, and small ASAF respectively.

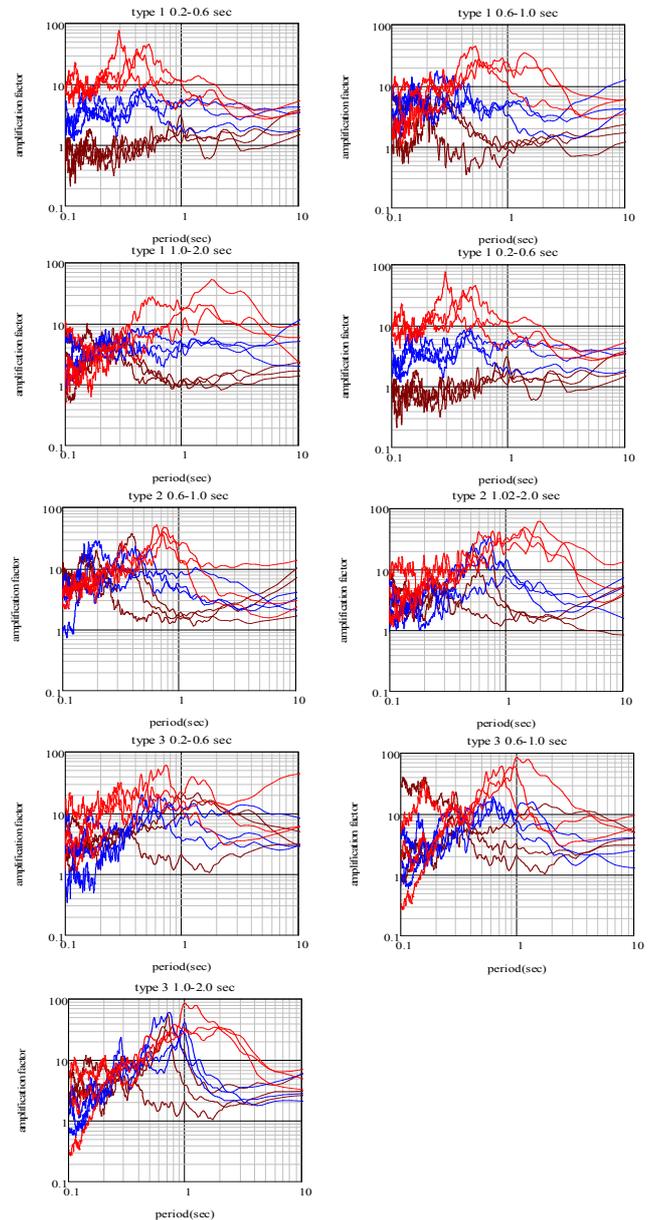


Fig. 7. Site amplification factor.

IV. RESULTS AND DISCUSSION

The 15 seismic waves shown in Table I are the reference EGMs. The sites' amplification factors were replaced, and the seismic ground motions were obtained at the 81 points shown in Table III. From this, a total of 1215 hypothetical EGMs were

acquired. Several sites are selected multiple times along with locations identical to the reference seismic ground motion observed sites in Table III. In this study, however, the statistical analyses focus on the period zone, ground type, and site amplification factor. Therefore, duplication is permitted with emphasis on the fact that the number of data is the same for each item. Figure 8 presents the results of a series of processes using the reference recording of the Noto Hanto Earthquake in 2007 at ISK005 and the target site KNG009 as examples. When a conventional equivalent-linear earthquake response analysis is conducted for a major earthquake, the amplification factor from the engineering bedrock to the ground surface tends to become significantly smaller than 1.0, especially in the high-frequency range, resulting in an underestimation of amplification. In such cases, the deconvolved EGM at the engineering bedrock becomes excessively large in amplitude, which generates unreliable results. This study conducted equivalent-linear earthquake response analyses, considering the frequency dependency of the nonlinear characteristics of the ground by [26], as described above, and no such phenomenon was observed.

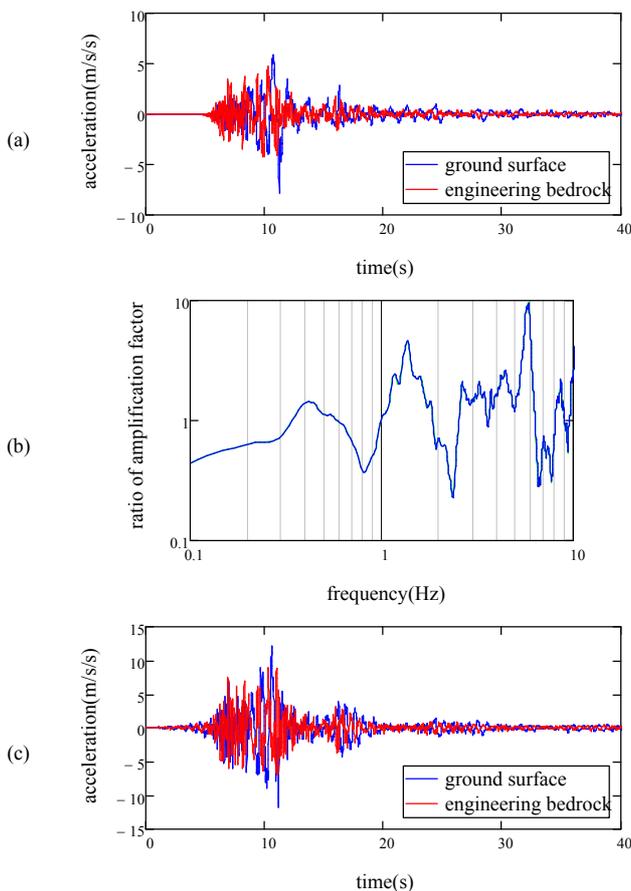


Fig. 8. Example of seismic ground motion calculation. (a) Deconvolution at ISK005, (b) ratio of amplification factor, (c) convolution at KNG009

The obtained acceleration response spectra for each ground type are shown in Figure 9. The thin gray line denotes the

individual response spectrum, the bold red line expresses the logarithmic mean, and the bold blue line signifies the logarithmic mean and logarithmic standard deviation. Figure 10 displays the response spectra of the logarithmic mean and the logarithmic mean + logarithmic standard deviation for each ground type. The black, blue, and red lines correspond to ground types 1, 2, and 3, respectively, the solid lines represent the logarithmic mean + logarithmic standard deviation, and the broken lines are the logarithmic mean spectra.

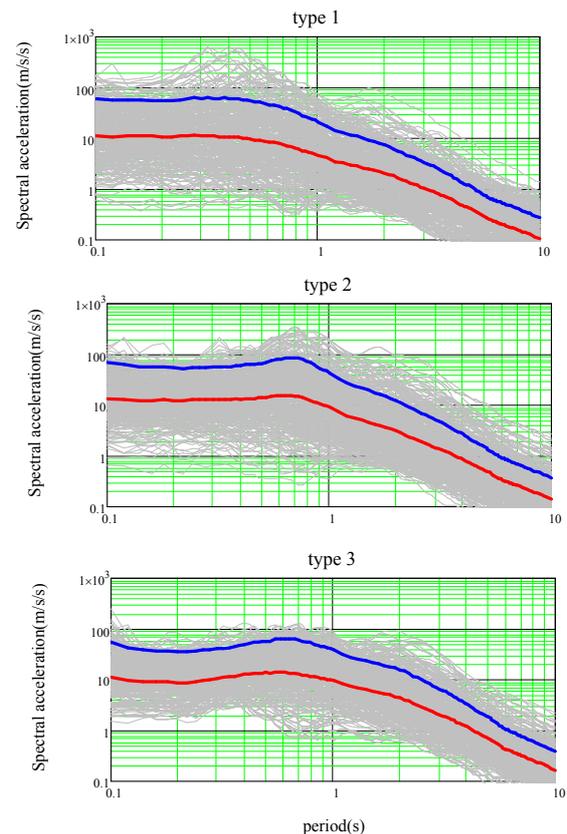


Fig. 9. Response spectrum (by ground type).

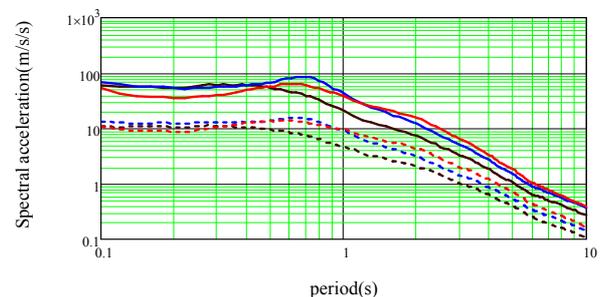


Fig. 10. Response spectrum (representative value).

Comparing the logarithmic mean and the logarithmic mean and logarithmic standard deviation response spectra, the type 1 ground spectra tend to be slightly smaller than the values of type 2 and 3 ground spectra in the frequency band of 0.5s or more. However, the difference is unremarkable. Comparing the

logarithmic mean and logarithmic standard deviation response spectra, the value of the type 3 ground tends to be slightly smaller in the periodic zone of 0.4s or less compared with the values of the types 1 and 2, but the difference is still unremarkable.

The obtained response spectra are shown in Figure 11 for large, medium, and small ASAF. The thin gray, bold red, and bold blue lines are the same as in Figure 8. Figure 12 presents the response spectra of the logarithmic mean and the logarithmic mean and logarithmic standard deviation spectra for each ASAF. The black, blue, and red lines are the small, medium, and large ASAF respectively. The solid line represents the logarithmic mean and logarithmic standard deviation, and the dashed line denotes the logarithmic mean spectra. According to the ASAF classification, the larger the ASAF, the larger the response spectra. For example, the values of the logarithmic mean and logarithmic standard deviation of the spectral response acceleration in a period of 1 s are 80, 30, and 10m/s² in the order of large, medium, and small ASAF. A remarkable difference is observed. The same applies to the logarithmic mean spectra, where the magnitude of the ASAF and the magnitude of the response spectrum correspond. Thus, depending on the site, rational design spectra can be established by classifying the target site with the magnitude of the ASAF.

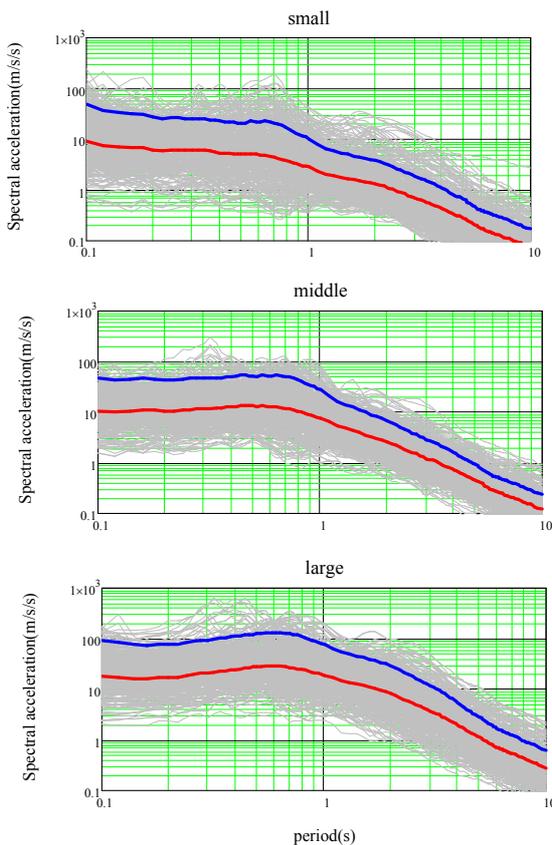


Fig. 11. Response spectrum (by ASAF)

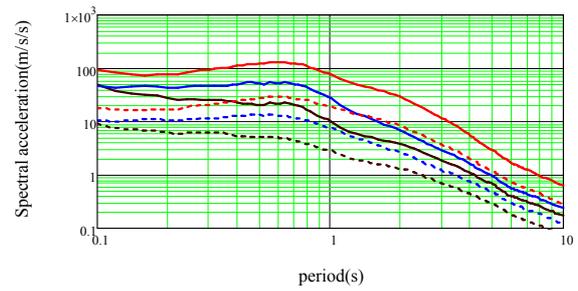


Fig. 12. Response spectrum (representative value)

Based on the results obtained, this study proposes design spectra according to the classification of ASAF. From the analysis of the present state of the design spectra in the JSHB, as described above, the proposed design spectra are the average spectra for the period range smaller than 1.0s and the average and standard deviation spectra for the period range larger than 1.0s. When the average spectrum and the average and standard deviation spectrum are combined in a period of 1s, spectral discontinuities occur. Therefore, the average spectrum of the short period side is extended in excess of 1s, and both are combined in a position corresponding with the average and standard deviation spectrum value. Finally, the function of the design spectrum is created to be the maximum envelope of the obtained result. Figure 13 contains the results along with the current design spectra.

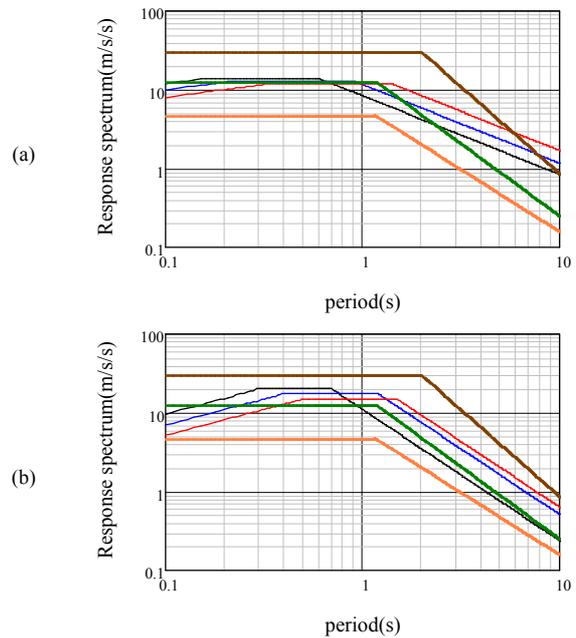


Fig. 13. Proposed spectrum: (a) Type 1 EGM, (b) Type 2 EGM

TABLE IV. PROPOSED SPECTRUM

ASAF	Spectral acceleration(m/s/s); S ₀	
Small	$T < 1.2; S_0 = 6.0$	$1.2 \leq T; S_0 = 8 / T^{19/12}$
Medium	$T < 1.3; S_0 = 12.0$	$1.3 \leq T; S_0 = 18 / T^{5/3}$
Large	$T < 2.0; S_0 = 29.0$	$2.0 \leq T; S_0 = 133 / T^{11/5}$

T is the natural period (s)

In the diagram, the thick solid brown, green, and orange lines are the proposed spectra at the points where the ASAF is respectively large, medium, and small. The solid blue line is the spectrum of the mean and standard deviation, and the solid black, blue, and red lines are the design spectra of the current design methods of ground types 1, 2, and 3, respectively. When the proposed design spectrum and the present design spectrum are compared, the proposed spectrum coincides with the present design spectrum in the case of type 1 seismic motion in medium ASAF. It increases in the case of large ASAF and decreases in the case of small ASAF. Therefore, it can be said that a reasonable design can be conducted with the same construction cost as conventional national average costs. Regarding the type 2 EGM, in the present design spectra of all ground types in the period bands of less than 1s and of more than 1s, the present design spectra of ground types 2 and 3 are between those proposed for the middle and large ASAF. In terms of the national average cost, a more rational design can be produced while suppressing more the construction cost. The function of the proposed design spectrum is shown in Table IV.

V. CONCLUSION

From the perspective of establishing rational site-specific MCEGMs, this study proposed design spectra using ASAF as indices. The main obtained conclusions of this research are:

- From the comparison of the design spectra for the current level 2 EGMs with the spectra of recent strong-motion records of major earthquakes, it was found that the average spectra for a period of less than 1s and the average and standard deviation spectra for a period of more than 1s of recent major earthquakes generally correspond to the design spectra of the JSHB, which is the MCEGMs assumed at the construction site. It is presumed that this is based on engineering judgment that considers the importance of highway bridges.
- The response spectra of recent major earthquakes with different site amplification factors are evaluated by categorizing the ASAF period ranges as large, medium, and small. The design spectra corresponding to the mean and the mean and standard deviation are proposed for a period of less than 1s and a period of more than 1s respectively. From the comparison of the proposed spectrum with the current design spectrum, it becomes possible to conduct a rational design almost as economical as or slightly more economical than the present design method.

ACKNOWLEDGEMENTS

The authors would like to thank Yuta Kinoshita for helping in conducting the seismic response analysis. The ground structure data and the strong-motion observation records at K-NET and KiK-net sites were provided by the National Research Institute for Earth Science and Disaster Resilience.

REFERENCES

- [1] T. Nagao and Y. Fukushima, "Source- and Site-Specific Earthquake Ground Motions: Application of a State-of-the-Art Evaluation Method," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5882–5888, Aug. 2020, <https://doi.org/10.48084/etasr.3612>.
- [2] *NEHRP recommended provisions for seismic regulations for new buildings and other structures*. Washington DC, USA: Building Seismic Safety Council for the Federal Emergency Management Agency, 2001.
- [3] *AASHTO Guide Specifications for LRF D Seismic Bridge Design, 2nd Edition, with 2012 and 2014 Interim Revisions*. Washington, DC, USA: AASHTO, 2011.
- [4] *Eurocode 8, Design of structures for earthquake resistance, Part1: general rules, seismic actions and rules for buildings*. Brussels, Belgium: European Committee for Standardization, 2003.
- [5] *Recommended Provisions for Seismic Regulations for New Buildings and Other Structures (FEMA 450), Part 1: Provisions*. Washington DC, USA: Building Seismic Safety Council for the Federal Emergency Management Agency, 2003.
- [6] *Minimum Design Loads for Buildings and Other Structures*, ASCE/SEI 7-10. ASCE, 2013, <https://doi.org/10.1061/9780784412916>.
- [7] *Specifications for highway bridges, Part 5, Seismic design, ver. 2012*. Japan Road Association, 2016.
- [8] D. M. Boore and W. B. Joyner, "Site amplifications for generic rock sites," *Bulletin of the Seismological Society of America*, vol. 87, no. 2, pp. 327–341, Apr. 1997.
- [9] Y. Fukushima, T. Nagao, J. Oshige, and I. Suetomi, "Ground motion evaluation for intra-plate earthquake by different site amplification factors and source models," in *7th International Conference on Earthquake Geotechnical Engineering*, Roma, Italy, Jun. 2019, pp. 2484–2492.
- [10] T. Nagao, "Seismic Amplification by Deep Subsurface and Proposal of a New Proxy," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5157–5163, Feb. 2020, <https://doi.org/10.48084/etasr.3276>.
- [11] R. D. Borcherdt, "Effects of local geology on ground motion near San Francisco Bay," *Bulletin of the Seismological Society of America*, vol. 60, no. 1, pp. 29–61, Feb. 1970.
- [12] J. Lermo and F. J. Chávez-García, "Site effect evaluation using spectral ratios with only one station," *Bulletin of the Seismological Society of America*, vol. 83, no. 5, pp. 1574–1594, Oct. 1993.
- [13] D. J. Andrews, "Objective Determination of Source Parameters and Similarity of Earthquakes of Different Size," in *Earthquake Source Mechanics*, Washington, USA: American Geophysical Union, 1986, pp. 259–267.
- [14] T. Iwata and K. Irikura, "Source Parameters of the 1983 Japan Sea Earthquake Sequence," *Journal of Physics of the Earth*, vol. 36, no. 4, pp. 155–184, 1988, <https://doi.org/10.4294/jpe1952.36.155>.
- [15] J. Boatwright, J. B. Fletcher, and T. E. Fumal, "A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes," *Bulletin of the Seismological Society of America*, vol. 81, no. 5, pp. 1754–1782, Oct. 1991.
- [16] L. F. Bonilla, J. H. Steidl, G. T. Lindley, A. G. Tumarkin, and R. J. Archuleta, "Site amplification in the San Fernando Valley, California: Variability of site-effect estimation using the S-wave, coda, and H/V methods," *Bulletin of the Seismological Society of America*, vol. 87, no. 3, pp. 710–730, Jun. 1997.
- [17] E. H. Field and K. H. Jacob, "A comparison and test of various site-response estimation techniques, including three that are not reference-site dependent," *Bulletin of the Seismological Society of America*, vol. 85, no. 4, pp. 1127–1143, Aug. 1995.
- [18] S. Parolai *et al.*, "Comparison of Different Site Response Estimation Techniques Using Aftershocks of the 1999 Izmit Earthquake," *Bulletin of the Seismological Society of America*, vol. 94, no. 3, pp. 1096–1108, Jun. 2004, <https://doi.org/10.1785/0120030086>.
- [19] L. Malagnini, K. Mayeda, A. Akinci, and P. L. Bragato, "Estimating Absolute Site Effects," *Bulletin of the Seismological Society of America*, vol. 94, no. 4, pp. 1343–1352, Aug. 2004, <https://doi.org/10.1785/012003161>.
- [20] S. Drouet, P. Triantafyllidis, A. Savvaïdis, and N. Theodulidis, "Comparison of Site-Effects Estimation Methods Using the Lefkas, Greece, 2003 Earthquake Aftershocks," *Bulletin of the Seismological*

- Society of America*, vol. 98, no. 5, pp. 2349–2363, Oct. 2008, <https://doi.org/10.1785/0120080004>.
- [21] J. H. Steidl, A. G. Tumarkin, and R. J. Archuleta, "What is a reference site?," *Bulletin of the Seismological Society of America*, vol. 86, no. 6, pp. 1733–1748, Dec. 1996.
- [22] S. Kinoshita, "Kyoshin Net (K-NET)," *Seismological Research Letters*, vol. 69, no. 4, pp. 309–332, Jul. 1998, <https://doi.org/10.1785/gssrl.69.4.309>.
- [23] S. Aoi, K. Obara, S. Hori, K. Kasahara, and Y. Okada, "New strong-motion observation network: KiK-net," *Eos, Transactions of the American Geophysical Union*, vol. 81, 2000.
- [24] A. Nozu and T. Sugano, "Site Amplification Factor for Strong-Motion Sites in Northern Hokkaido, Japan, Based on Spectral Inversion Technique," Technical Note of the Port and Airport Research Institute, No.1102, 2005.
- [25] D. M. Boore, E. M. Thompson, and H. Cadet, "Regional Correlations of VS30 and Velocities Averaged Over Depths Less Than and Greater Than 30 Meters," *Bulletin of the Seismological Society of America*, vol. 101, no. 6, pp. 3046–3059, Dec. 2011, <https://doi.org/10.1785/0120110071>.
- [26] N. Yoshida, S. Kobayashi, I. Suetomi, and K. Miura, "Equivalent linear method considering frequency dependent characteristics of stiffness and damping," *Soil Dynamics and Earthquake Engineering*, vol. 22, no. 3, pp. 205–222, Apr. 2002, [https://doi.org/10.1016/S0267-7261\(02\)00011-8](https://doi.org/10.1016/S0267-7261(02)00011-8).
- [27] S. Yasuda and I. Yamaguchi, "Dynamic shear modules obtained in the laboratory and in-situ," in *Proceedings of the Symposium on Evaluation of Deformation and Strength of Sandy Grounds*, 1985, pp. 115–118.

Seamless Transition between Islanded and Grid Connected Three-Phase VSI-based Microgrids

Mubashir Hayat Khan

Department of Electrical Power Engineering
Faculty of Electrical and Electronic Engineering
University Tun Hussein Onn Malaysia
Johor, Malaysia
mubashir.uthm@gmail.com

Erum Pathan

Electronic Engineering Department
Quaid-e-Awam University of Engineering, Science and
Technology
Nawabshah, Pakistan
erumasad79@gmail.com

Muhammad Asad

Commissioning Services Division-Central
National Grid SA
Saudi Arabia
csd_sec@yahoo.com

Muhammad Ahsan Sadiq

Department of Electrical Engineering
University of Poonch
Rawalakot, Pakistan
enr.ahsan@upr.edu.pk

Amjad Ammar Qureshi

NER Engineering & Capital Projects
Trafigura Nyrstar, Australia
enr.amjadammar@gmail.com

Muhammad Shahid

Protection & Automation
Siemens Ltd Saudi Arabia
enr.shahid26@gmail.com

Amanullah Khan Pathan

Power and Distribution Department
Larsen and Toubro Saudi Arabia LLC
Dammam, Saudi Arabia
amanullahpathan@yahoo.com

Nadim Imtiyaz Shaikh

Construction Manager
Larsen and Toubro Saudi Arabia LLC
Dammam, Saudi Arabia
nahids12@yahoo.com

Abstract-Microgrids (MGs) are the emergent solution to overcome the current electricity demand. The MGs provide the facility to operate in both isolated and grid-connected modes. For both operating modes, Distributed Generation (DG) inverters are operating under grid forming or grid following control modes. During mode switching, the MG experiences enormous fluctuations, which occur due to the unidirectional islanding event. This paper presents a control strategy by using the modified power control scheme, current controller, and DC linked voltage controller scheme to ensure the operational mode transfer smoothly from the grid-connected to the islanded mode and vice versa. The proposed control scheme is applied to a three-phase distributed energy resource-based MG system with fixed loads. The simulation results validate the effectiveness of the control technique while tested at the point of common coupling and also at the time of mode transfer.

Keywords-grid forming; grid following; seamless transition; droop control

I. INTRODUCTION

Microgrids (MGs) provide a promising solution to overcome the electricity shortage in a reliable way. Distributed

Generation (DG) utilized in MGs is pretty common nowadays. DGs like solar, wind, and Combined Heat Power (CHP), that are attached to Grid Connected Mode (GCM) or Islanded Mode (IM) protocols with connecting loads have many advantages over the traditional grid [1-2]. Although the MGs provide an alternative with bidirectional communication, there are some control issues, particularly regarding power, voltage, and frequency deviations that may occur when the MG changes its operational mode from GCM to IM [3-4]. Generally, the control schemes in terms of transition are divided into two categories. The first one is the single control scheme for regulating voltage [5-10], which is non-linear theory-based such as the Lyapunov-based control. The second one is the control with two schemes with pre-allotted intents [11-16], based on the model predictive control scheme. However, in both control techniques, high computational values and the complexity of the schemes do not allow efficient implementation of the control schemes. On the contrary, linear control schemes are less complicated and can be easily implemented without any complex computational burden [5]. Feedback and feed-forward procedures can be a good alternative to make the control structure simple and convincing

Corresponding author: Mubashir Hayat Khan

with the provision of seamless transition either from GCM to IM or IM to GCM [17-18]. During the mode transition from grid following to grid-forming mode, frequency deviations may occur which may lead the MG system to instability. Moreover, voltage and current errors may also occur during the mode transition, while the voltage during the switching process must follow the IEEE standard criteria [19]. A proper synchronization algorithm is needed in order to have smooth switching from one mode to another. Switching between modes can be done with the help of an operator or automatically. In the first case, the control of transition intensity can be controlled with the MG functional point's readjustment. In the second case, unidirectional islanding is used to readjust the operational points of the MG, but faces various issues related to frequency and voltage errors and system stability concerns. In [8-9, 14, 16, 20-23], different solutions have been presented to compensate the errors that occur during transition. Most of the control solutions are based on switching, but the droop-based control scheme [8, 19] does not need any kind of switching. This control scheme can be used for the transition from IM to GCM or GCM to IM.

In [5, 9, 16, 21], nonlinear control schemes with adaptive feedback stepping technique have been introduced to solve the voltage frequency deviations and were applied to both operation modes. The Model Predictive Control (MPC) technique [5] is applied to the single phase multi bus inverter-based MG system with automatic tuning functionality for seamless transition. The interaction between different inverter controllers used in the same system still needs to be investigated because during the IM, frequency and power sharing errors occur. Hence, the damping ratio needs to be improved in this case. In order to achieve smooth operation the frequency deviations are mitigated by utilizing the virtual inertia [23]. Virtual Synchronous Generators (VGSs) with droop control are used to improve the dynamic performance of the system. However, active power sharing accuracy is not achieved [24]. If the droop control parameters are selected accurately as per system requirements, the active power sharing errors can be controlled and the control strategy could be more efficient as compared to the control using VGS [25]. The Proportional Resonant (PR) controller is implemented to mitigate the frequency errors during mode transition in [26]. Moreover, phase angle correction with autonomous operation was also ensured in the control scheme. Modified PQ control technique with tracking V/f controller [27-28] needs constant communication between the IM and the main grid. A two separate synchronization compensator-based control system was implemented in [29] to mitigate the voltage and phase angle deviations but the system stability was not ensured. To get an effective MG operation during mode switching, a fast seamless transient control technique should be implemented which will not compromise the permissible standards in terms of power sharing accuracy, voltage, and frequency.

This paper presents a control strategy that uses a modified power control scheme, consisting of the current controller and a DC linked voltage controller scheme to ensure the operational mode transfer smoothly from GCM to IM transitions and vice versa. The proposed control scheme is applied to a three phase Distributed Energy Resources-based MG system with fix loads.

The simulation results validate the effectiveness of the control technique in terms of voltage, frequency, and power sharing behavior at the Point of Common Coupling (PCC) and at the time of mode transfer.

II. MICROGRID CONFIGURATION AND OPERATION

A three phase MG system with three DGs is shown in Figure 1. A constant DC source is connected separately to all the DGs. An LCL filter is connected with each DG and the loads are connected to the PCC. All the Voltage Source Inverters (VSIs), explored for voltage, power sharing, and frequency, are connected to the main grid with the help of a circuit breaker installed just after the PCC.

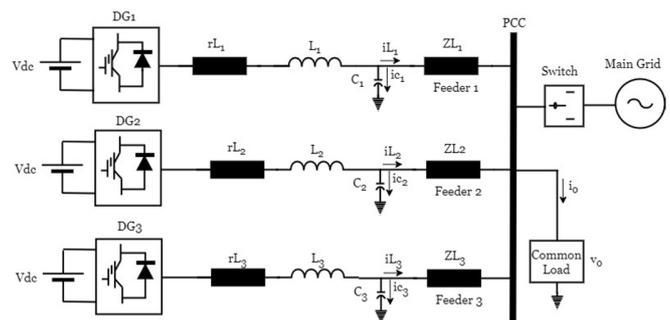


Fig. 1. A microgrid with three DG inverters connected in parallel.

A common load is connected to the PCC and the parallel connected inverters can be detached from the main grid for IM operation. GCM and IM operational modes are further explained below.

A. Grid Connected Mode

Generally the GCM control is implemented in a grid-feeding control scheme. A control structure based on a PI controller with a dq reference frame work is established by using the convention current control structure [30-31]. The current control structure block arrangement is shown in Figure 2.

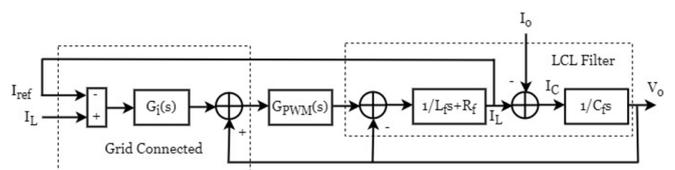


Fig. 2. Current control.

B. Islanded Mode

During the IM control mode, when a change occurs between frequency and voltage because they are designed separately for load and generation, ultimately the P and Q sharing accuracy is decreased [30-31]. It is therefore essential to overcome this mismatch according to the load requirements. In a parallel connected MG system, when multiple DGs are connected in the same network, the load sharing also needs to be equally distributed to all the DGs as per their capacities so the droop characteristics need to be incorporated in the system to cope up with the load sharing issue as shown in Figure 3.

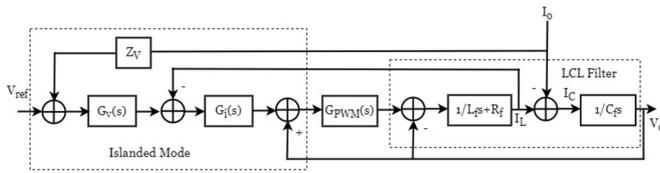


Fig. 3. Voltage control.

Active and reactive powers are calculated by using LPF with very small cutoff frequency. The block diagram for grid forming is shown in Figure 3. The output voltage can be expressed as:

$$V_o = G(s)V_{ref} - Z_o I_o \quad (1)$$

III. THE PROPOSED CONTROL SCHEME

The control loops consists of a droop controller, a current controller, and a linked voltage controller. For these controllers, the control design procedure is given below which is further categorized into GCM and IM control schemes.

A. Modified Power Control

Droop control technique provides P- ω and Q-f droop control with decentralized and communication free facilities. The active and reactive powers of each inverter are depending on voltage and frequency at the PCC which is managed by conventional droop control in the IM mode control scheme. Moreover, another PI controller is also applied for ensuring proper tracking of V_{ref} . The power calculation mechanism is given in Figure 4.

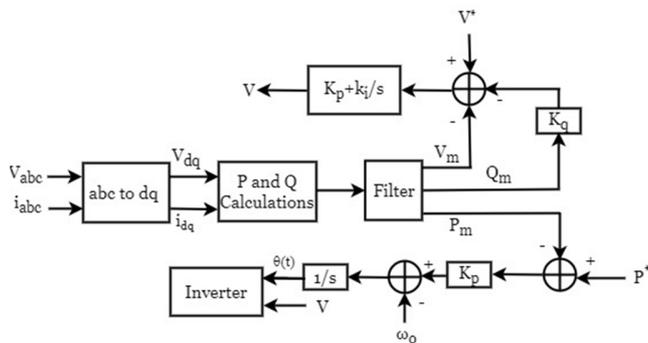


Fig. 4. Active and reactive power calculations.

B. Current Controller

Higher bandwidth is required in the design of the current controller. It is preferred to get faster response, so low switching frequency is needed. In the closed loop system, the bandwidth is taken 10 times smaller than the switching frequency. The block diagram in Figure 5 shows the current controller configuration. The transfer function for the current control strategy is:

$$\frac{G_c(s)G_m(s) \cdot \frac{1}{R_f + sL_f}}{\frac{1}{R_f + sL_f}} = \frac{1}{\tau s + 1} \quad (2)$$

where τ can be considered as a constant of the PI controller transfer function G_c .

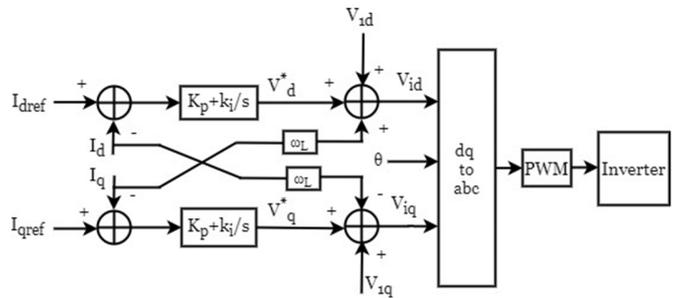


Fig. 5. Current controller.

C. DC Linked Voltage Controller

Both side inverter currents are equalized for the voltage DC linked controller before designing the controller:

$$V_{dc} I_{dc} = 3I_{ph} V_{ph} \quad (3)$$

where $I_d = \sqrt{2}I_{rms}$ and can be is expressed as:

$$I_{dc} = \sqrt{3/2} \frac{V_L - L}{V_{dc}} I_d = K_{DC} I_d \quad (4)$$

and I_{dc} relates to the capacitor current which is given as:

$$I_{dc} = C \frac{dV_{dc}}{dt} \quad (5)$$

So, the basic functionality of the voltage controller is to maintain the voltage. The configuration of the DC linked voltage controller is given in Figure 6.

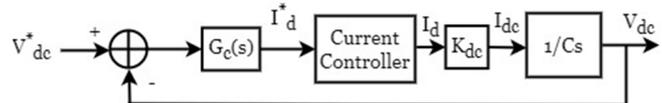


Fig. 6. DC lined voltage controller.

IV. SIMULATION RESULTS

The validation of the proposed control mechanism for the seamless transition of MGs between the two operational modes was carried out in MATLAB/Simulink. The system specifications are given in Table I, and the overall MATLAB/Simulink model is given in Figure 11.

A. Case 1: Grid Connected Mode

In the GCM control scheme it is necessary for a MG system to operate in constant PQ mode. To attain constant PQ, the inverter is operated in dq-reference framework to get current control. In the transformation dq-abc, the observation time scale is very high and so it is preferred to consider the average model to get the voltage at the abc domain. In the GCM mode,

the imbalance occurs due to the generated power and load at the point of connection and the excess power is supplied by the DC bus capacitor. The occurring reduction in the bus voltage due to the supply needs to be maintained again. However, an inner loop is established to balance the DC bus voltage and current controller. When the breaker of the proposed model is closed towards the GCM, the grid connection mode operation is established. In GCM the voltage and frequency at the PCC must always be in a fixed allowable range. As shown in Figure 7(a), voltage and frequency are in the permissible range. The frequency is almost constant throughout the operation from $t=0$ to $t=10$ s. At $t=0.8$ s, voltage maintains the fixed magnitude. The inverter current is successfully tracking the reference current as shown in Figure 7(b). The active and reactive powers are shared accurately as shown in Figure 7(c) just after some delay at $t=0.6$ s.

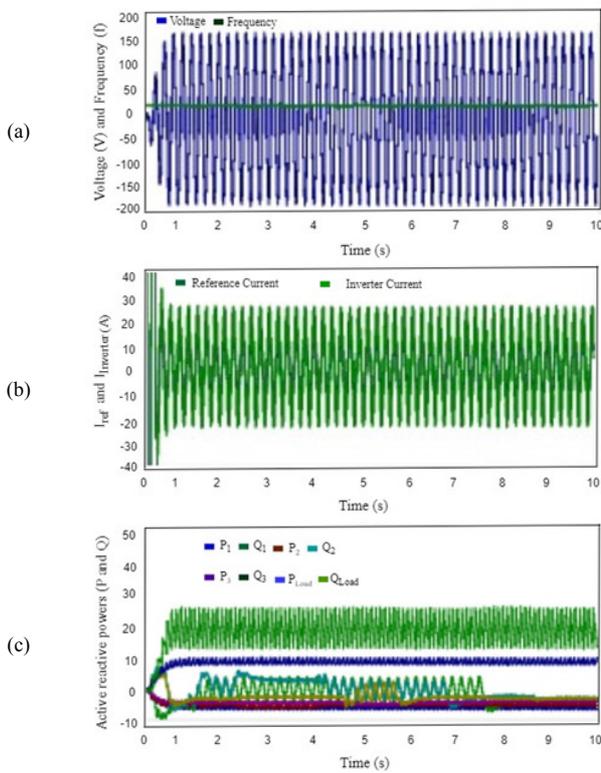


Fig. 7. GCM results: (a) Voltage and frequency, (b) reference and inverter currents, (c) active and reactive powers.

B. Case 2: Transition between GCM and IM

This section shows the results when transition occurs from GCM to IM. The PCC voltage and frequency waveform are shown in Figure 8(a). At $t=5$ s there is a little deviation which is compensated in a very short time interval. The same happens in the case of deviating frequency. Figure 8(b) shows the behavior of the currents. $I_{inverter}$ decreases at $t=5$ s which is successfully tracked by the reference current. Figure 8(c) shows the power sharing accuracy of the DG inverters during the change in the operational mode. From $t=5$ s the active and reactive powers are shared equally.

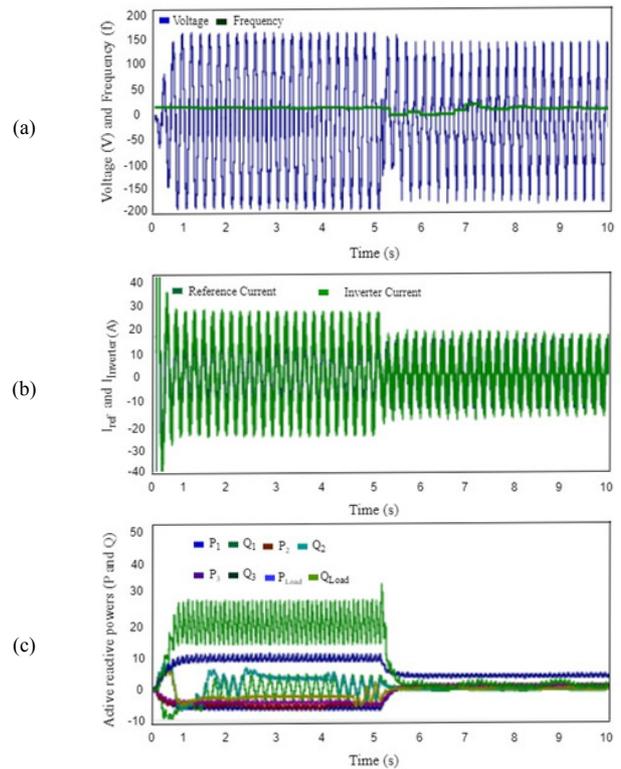


Fig. 8. Transition from GCM to IM: (a) Voltage and frequency, (b) reference and inverter currents, (c) active and reactive powers.

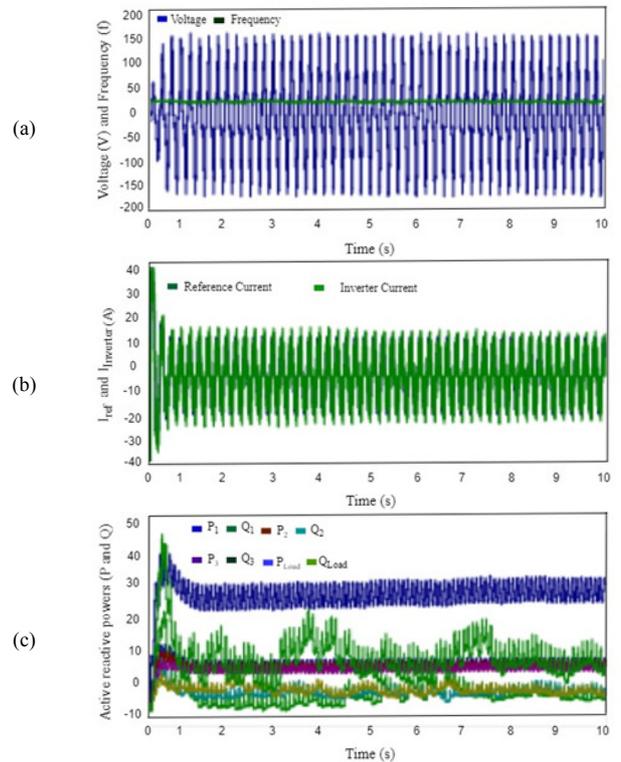


Fig. 9. IM results: (a) Voltage and frequency, (b) reference and inverter currents, (c) active and reactive powers.

C. Case 3: Islanded Mode

In Case 3, the IM of the MG is discussed with respect to the behavior of voltage, current, and powers. As shown in Figure 9(a), voltage and frequency are in the permissible range. The frequency is constant throughout the simulation and the voltage magnitude is constant just after 0.8s. The inverter current is successfully tracking the reference current as shown in Figure 9(b). Figure 9(c) shows the PQ sharing in IM.

D. Case 4: Transition from IM to GCM

It is shown in Figure 10(a) that when the reconnection of the MG from IM to GCM takes place, there is a small deviation in the voltage waveform and the frequency shows some deviations before returning to a smooth level just after $t=5s$. Figure 10(b) shows the behavior of the inverter and reference currents which are tracking each other successfully at $t=5s$. In Figure 10(c) the power sharing accuracy of the proposed control scheme is shown when the MG is shifted to GCM from IM operation. The results show that the active and reactive powers are shared accurately from $t=5s$.

TABLE I. SYSTEM PARAMETERS

Parameter	Value	Parameter	Value
V_{DC}	400 V	Load	130 Ω , 0.22 H
f_{Sw}	10kHz	Z_{grid}	0.4+j0.6 Ω
f	50 Hz	C	15 μ F
Z_{L1}, Z_{L2} and Z_{L3}	0.3+j0.5 Ω , 1.3+j2.65 Ω & 2.4+j3.54 Ω		

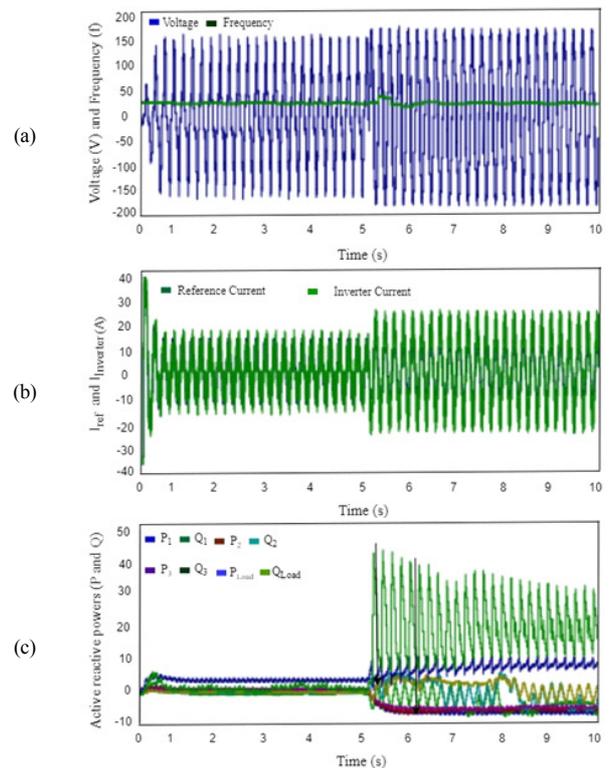


Fig. 10. Transition from IM to GCM: (a) Voltage and frequency, (b) reference and inverter currents, (c) active and reactive powers..

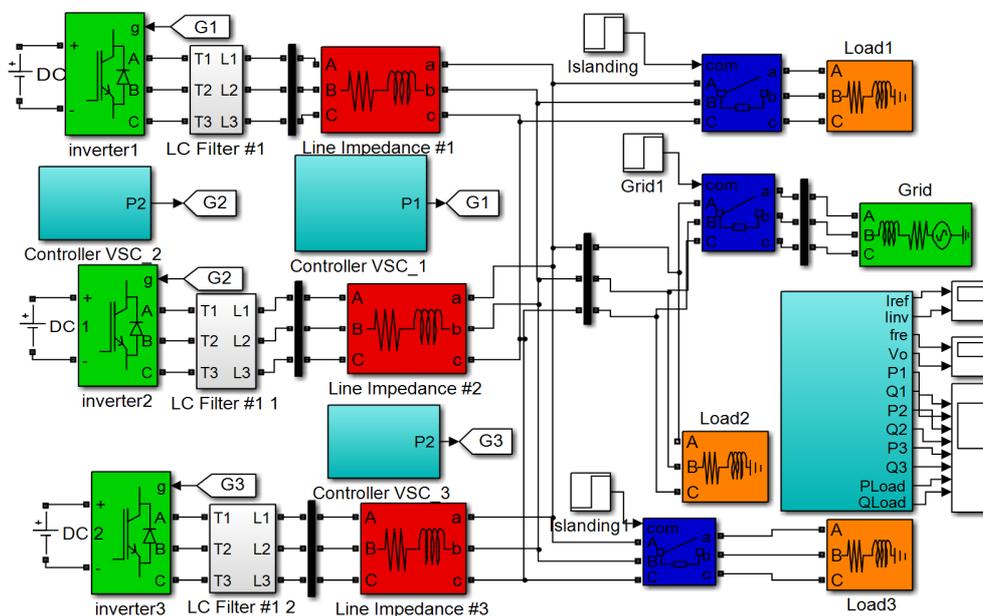


Fig. 11. MATLAB/Simulink model.

V. CONCLUSION

In this paper, the transition from GCM to IM and vice versa was studied in detail. A simple approach by modifying the power, current, and DC linked voltage controllers was

presented. A smooth transfer from one operational mode to another was conducted without affecting the active and reactive power sharing among the DG inverters connected in parallel and the voltage, frequency, reference current, and inverter

current the at the PCC have been explored. The proposed control scheme is applied to a three Distributed Energy Resource system with fixed load. The simulation results validate the effectiveness of the control scheme in terms of voltage, frequency, and power sharing behavior at the PCC and at the time of mode transfer.

REFERENCES

- [1] P. Singh, D. P. Kothari, and M. Singh, "Integration of Distributed Energy Resources," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 7, no. 1, pp. 91–96, 2014.
- [2] M. Castilla, L. G. de Vicuna, and J. Miret, "Control of Power Converters in AC Microgrids," in *Microgrids Design and Implementation*, A. C. Zamboni de Souza and M. Castilla, Eds. Cambridge, UK: Springer, 2019, pp. 139–170.
- [3] H. Etemadi and R. Iravani, "Supplementary mechanisms for smooth transition between control modes in a microgrid," *Electric Power Systems Research*, vol. 142, pp. 249–257, Jan. 2017, <https://doi.org/10.1016/j.epsr.2016.09.033>.
- [4] M. B. Delghavi and A. Yazdani, "A control strategy for islanded operation of a Distributed Resource (DR) unit," in *IEEE Power Energy Society General Meeting*, Calgary, Canada, Jul. 2009, pp. 1–8, <https://doi.org/10.1109/PES.2009.5275592>.
- [5] X. Li, H. Zhang, M. B. Shadmand, and R. S. Balog, "Model Predictive Control of a Voltage-Source Inverter With Seamless Transition Between Islanded and Grid-Connected Operations," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 7906–7918, Oct. 2017, <https://doi.org/10.1109/TIE.2017.2696459>.
- [6] Y. A. I. Mohamed and A. A. Radwan, "Hierarchical Control System for Robust Microgrid Operation and Seamless Mode Transfer in Active Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 352–362, Jun. 2011, <https://doi.org/10.1109/TSG.2011.2136362>.
- [7] J. Wang, N. C. P. Chang, X. Feng, and A. Monti, "Design of a Generalized Control Algorithm for Parallel Inverters for Smooth Microgrid Transition Operation," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 8, pp. 4900–4914, Aug. 2015, <https://doi.org/10.1109/TIE.2015.2404317>.
- [8] J. Kim, J. M. Guerrero, P. Rodriguez, R. Teodorescu, and K. Nam, "Mode Adaptive Droop Control With Virtual Output Impedances for an Inverter-Based Flexible AC Microgrid," *IEEE Transactions on Power Electronics*, vol. 26, no. 3, pp. 689–701, Mar. 2011, <https://doi.org/10.1109/TPEL.2010.2091685>.
- [9] S. M. Ashabani and Y. A. I. Mohamed, "A Flexible Control Strategy for Grid-Connected and Islanded Microgrids With Enhanced Stability Using Nonlinear Microgrid Stabilizer," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1291–1301, Sep. 2012, <https://doi.org/10.1109/TSG.2012.2202131>.
- [10] S. Sajadian and R. Ahmadi, "Model Predictive Control of Dual-Mode Operations Z-Source Inverter: Islanded and Grid-Connected," *IEEE Transactions on Power Electronics*, vol. 33, no. 5, pp. 4488–4497, May 2018, <https://doi.org/10.1109/TPEL.2017.2723358>.
- [11] G. G. Talapur, H. M. Suryawanshi, L. Xu, and A. B. Shitole, "A Reliable Microgrid With Seamless Transition Between Grid Connected and Islanded Mode for Residential Community With Enhanced Power Quality," *IEEE Transactions on Industry Applications*, vol. 54, no. 5, pp. 5246–5255, Sep. 2018, <https://doi.org/10.1109/TIA.2018.2808482>.
- [12] Z. Guo, D. Sha, and X. Liao, "Voltage magnitude and frequency control of three-phase voltage source inverter for seamless transfer," *IET Power Electronics*, vol. 7, no. 1, pp. 200–208, Jan. 2014, <https://doi.org/10.1049/iet-pel.2012.0723>.
- [13] K. Lim, I. Song, and J. Choi, "Indirect Current Control for Seamless Transfer of Utility Interactive Inverter," in *International Power Electronics Conference*, Niigata, Japan, May 2018, pp. 803–808, <https://doi.org/10.23919/IPEC.2018.8507999>.
- [14] C. Chen, Y. Wang, J. Lai, Y. Lee, and D. Martin, "Design of Parallel Inverters for Smooth Mode Transfer Microgrid Applications," *IEEE Transactions on Power Electronics*, vol. 25, no. 1, pp. 6–15, Jan. 2010, <https://doi.org/10.1109/TPEL.2009.2025864>.
- [15] L. G. Meegahapola, D. Robinson, A. P. Agalgaonkar, S. Perera, and P. Ciufo, "Microgrids of Commercial Buildings: Strategies to Manage Mode Transfer From Grid Connected to Islanded Mode," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 4, pp. 1337–1347, Oct. 2014, <https://doi.org/10.1109/TSTE.2014.2305657>.
- [16] H. Mahmood and Jin Jiang, "A control strategy of a distributed generation unit for seamless transfer between grid connected and islanded modes," in *IEEE 23rd International Symposium on Industrial Electronics*, Istanbul, Turkey, Jun. 2014, pp. 2518–2523, <https://doi.org/10.1109/ISIE.2014.6865016>.
- [17] Y. Chen, J. M. Guerrero, Z. Shuai, Z. Chen, L. Zhou, and A. Luo, "Fast Reactive Power Sharing, Circulating Current and Resonance Suppression for Parallel Inverters Using Resistive-Capacitive Output Impedance," *IEEE Transactions on Power Electronics*, vol. 31, no. 8, pp. 5524–5537, Aug. 2016, <https://doi.org/10.1109/TPEL.2015.2493103>.
- [18] J. C. Vasquez, J. M. Guerrero, A. Luna, P. Rodriguez, and R. Teodorescu, "Adaptive Droop Control Applied to Voltage-Source Inverters Operating in Grid-Connected and Islanded Modes," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 4088–4096, Oct. 2009, <https://doi.org/10.1109/TIE.2009.2027921>.
- [19] "IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems," *IEEE Std 1547-2003*, pp. 1–28, Jul. 2003, <https://doi.org/10.1109/IEEESTD.2003.94285>.
- [20] M. Syahril, M. A. Roslan, and B. Ismail, "Microgrid synchronization using power offset through a central controller," *Journal of Physics: Conference Series*, vol. 1432, Jan. 2020, Art. no. 012019, <https://doi.org/10.1088/1742-6596/1432/1/012019>.
- [21] G. Lou, W. Gu, J. Wang, J. Wang, and B. Gu, "A Unified Control Scheme Based on a Disturbance Observer for Seamless Transition Operation of Inverter-Interfaced Distributed Generation," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5444–5454, Sep. 2018, <https://doi.org/10.1109/TSG.2017.2769675>.
- [22] E. Pathan, A. A. Bakar, S. A. Zulkifli, M. H. Khan, H. Arshad, and M. Asad, "A Robust Frequency Controller based on Linear Matrix Inequality for a Parallel Islanded Microgrid," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6264–6269, Oct. 2020, <https://doi.org/10.48084/etasr.3769>.
- [23] J. Liu, Y. Miura, and T. Ise, "Comparison of Dynamic Characteristics Between Virtual Synchronous Generator and Droop Control in Inverter-Based Distributed Generators," *IEEE Transactions on Power Electronics*, vol. 31, no. 5, pp. 3600–3611, May 2016, <https://doi.org/10.1109/TPEL.2015.2465852>.
- [24] U. Tamrakar, D. Shrestha, M. Maharjan, B. P. Bhattarai, T. M. Hansen, and R. Tonkoski, "Virtual Inertia: Current Trends and Future Directions," *Applied Sciences*, vol. 7, no. 7, Jul. 2017, Art. no. 654, <https://doi.org/10.3390/app7070654>.
- [25] X. Meng, Z. Liu, J. Liu, T. Wu, S. Wang, and B. Liu, "Comparison between virtual synchronous generator and droop controlled inverter," in *IEEE 2nd Annual Southern Power Electronics Conference*, Auckland, New Zealand, Dec. 2016, <https://doi.org/10.1109/SPEC.2016.7846007>.
- [26] B. Shoeiby, R. Davoodnezhad, D. G. Holmes, and B. P. McGrath, "A resonant current regulator based microgrid control strategy with smooth transition between islanded and grid-connected modes," in *IEEE 5th International Symposium on Power Electronics for Distributed Generation Systems*, Galway, Ireland, Jun. 2014, pp. 1–8, <https://doi.org/10.1109/PEDG.2014.6878631>.
- [27] E. Pathan *et al.*, "Virtual Impedance-based Decentralized Power Sharing Control of an Islanded AC Microgrid," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6620–6625, Feb. 2021, <https://doi.org/10.48084/etasr.3946>.
- [28] L. Xia and L. Hai, "Comparison of dynamic power sharing characteristics between virtual synchronous generator and droop control in inverter-based microgrid," in *IEEE 3rd International Future Energy Electronics Conference and ECCE Asia*, Kaohsiung, Taiwan, Jun. 2017, pp. 1548–1552, <https://doi.org/10.1109/IFEEC.2017.7992276>.
- [29] Z. Chen, W. Zhang, J. Cai, T. Cai, Z. Xu, and N. Yan, "A synchronization control method for micro-grid with droop control," in

- IEEE Energy Conversion Congress and Exposition*, Montreal, Canada, Sep. 2015, pp. 519–524, <https://doi.org/10.1109/ECCE.2015.7309733>.
- [30] G. Shen, X. Zhu, J. Zhang, and D. Xu, "A New Feedback Method for PR Current Control of LCL-Filter-Based Grid-Connected Inverter," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 2033–2041, Jun. 2010, <https://doi.org/10.1109/TIE.2010.2040552>.
- [31] G. Shen, D. Xu, L. Cao, and X. Zhu, "An Improved Control Strategy for Grid-Connected Voltage Source Inverters With an LCL Filter," *IEEE Transactions on Power Electronics*, vol. 23, no. 4, pp. 1899–1906, Jul. 2008, <https://doi.org/10.1109/TPEL.2008.924602>.

Examinee Characteristics and their Impact on the Psychometric Properties of a Multiple Choice Test According to the Item Response Theory (IRT)

Deyab Almaleki

Department of Evaluation, Measurement, and Research
Umm Al-Qura University
Makkah, Saudi Arabia
damaleki@uqu.edu.sa

Abstract—The aim of the current study is to provide improvement evaluation practices in the educational process. A multiple choice test was developed, which was based on content analysis and the test specification table covered some of the vocabulary of the applied statistics course. The test in its final form consisted of 18 items that were reviewed by specialists in the field of statistics to determine their validity. The results determine the relationship between individual responses and the student ability. Most thresholds span the negative section of the ability. Item information curves show that the items provide a good amount of information about a student with lower or moderate ability compared to a student with high ability. In terms of precision, most items were more convenient with lower ability students. The test characteristic curve was plotted according to the change in the characteristics of the examinees. The information obtained by female students appeared to be more than the information obtained by male students and the test provided more information about students who were not studying statistics in an earlier stage compared with students who did. This test clearly indicated that, based on the level of the statistics course, there should be a periodic review of the tests in line with the nature and level of the course materials in order to have a logical judgment about the level of the students' progress at the level of their ability.

Keywords—item response theory; item characteristics; multiple-choice; psychometric properties

I. INTRODUCTION

A test is an educational tool that is frequently used to evaluate students' academic achievement and progress. Tests also provide an opportunity to verify students' skills in many educational situations when it is not possible to use other assessment methods. Despite the known problems of indirect measurement, a lot of traits such as mathematical abilities, verbal skills, resistance to stress, intelligence, dissatisfaction, different opinions about a particular topic, etc. cannot be directly observed and measured [1-3]. These are known as the latent traits, and they can be measured only indirectly, often using specially prepared questionnaires where the responses are closely related to the specific traits being studied. Tests are frequently used to assess students' cognitive progress and to

build question banks. As a result, the so-called latent trait models have been developed and are used to estimate the parameter values associated with the human personality [4-5]. These models provide a different type of information that in turn helps to develop and improve tests accordingly. Many researchers rely on the information from data that are analyzed as a result of the subjects' responses. But it is important to ask whether the formulation of stimuli (questions) may provide another type of data or information that serves the research process [1, 6-8].

Many researchers and specialists in the field of measurement and evaluation, are interested in the basic concepts and organizational theoretical frameworks of measurement and evaluation and ways to apply them [9-11], due to the great role this science plays in various fields of scientific research in general and educational and psychological research in particular. Research activity according to the needs of educational institutions will positively affect development and improvement in accordance with Saudi Arabia's Vision 2030. Furthermore, the means of developing tests and measurement methods are extremely important because the data issued from the measurement processes have to be valid and accurate, as some crucial decisions such as admission or promotion may be based upon them [12-16]. In addition, it is the responsibility of specialists in the field of measurement and evaluation to enrich the literature, develop tests used in the educational field, and reduce the possibility of potential measurement errors during the evaluation process [17, 18]. Postgraduate tests in Saudi universities have not yet been subjected to much scrutiny by local and international evaluation institutions, because most of the quality assurance agencies, such as the National Commission for Evaluation and Accreditation (NCAAA), has only recently included postgraduate programs in its plans, providing the opportunity for universities to apply for accreditation for these programs. Midterm and final exams and the way they are administered are some of the indicators used by the NCAAA or other agencies to accurately judge the progress of an academic program. Therefore, improving these tests has become a necessary requirement.

Corresponding author: Deyab Almaleki

II. SIGNIFICANCE OF THE CURRENT STUDY

The significance of the current study stems from the importance of the evaluation processes in the educational process. The process of improving tests and identifying their psychometric characteristics is the task of those working in the field of measurement and evaluation in order to provide a comprehensive understanding and a deep descriptive analysis of the advantages and disadvantages observed in those tests [19-21]. Furthermore, the scarcity of this type of scientific studies has widened the gap between the tools currently used to measure the level of achievement of students and what is hoped that these tools should be like. The quality of these tools has not been determined or reviewed, and therefore they have not been assessed or evaluated [22-25]. The practical significance of this type of research lies in the use of Item Response Theory (IRT) models in analyzing students' responses to the achievement test in a more objective way, showing whether there is an effect of the multiplicity of characteristics of the participants on the test items (multiple choice) in terms of the accuracy of the estimates of the items' parameters and the individuals' ability parameters. It can also guide the composers of the test questions to take into account some points that may affect the psychometric properties of the items and the test and the accuracy of its results [26-29].

III. ITEM RESPONSE THEORY

There are many ways (i.e. models) to determine the relationship between individual responses and student ability. Within the framework of modern measurement theory, many models and applications have been formulated and applied to real test data. This includes the measurement assumptions about the characteristics of the test item, the performance of the subject, and how this performance is related to knowledge [27-37-40]. Tests and evaluation processes in general form the basis of the education system, and their importance lies in improving educational planning, developing a mechanism for enhancing curricular content, measuring learners' competence, and comparing student performance or achievement data. Evaluations also have a role for schools and teachers [30-33]. Tests are a tool of assessment, and their quality depends on a large extent on the nature and quality of the information collected during the preparation of the assessment. Over the decades, the test building system has undergone a lot of development through the emergence of many test building theories focused in many types of tests, such as oral tests, standardized tests, and realistic evaluation. Until today, theories have continued to develop in order to keep up with the changes in policies and new educational practices [2]. The modern theoretical methods were largely developed in the sixties to the late eighties. The IRT is a general statistical theory considering the characteristics of the test item, the subject's performance on the item, and how the performance is related to the abilities that are measured by the test items [12, 34-36]. The IRT provides a rich statistical tool for analyzing educational tests and psychometric measures. The IRT assumes the following:

- The test performance of the subjects can be predicted (or explained) by a set of factors called traits or latent traits and abilities.

- The relationship between the subject's performance and the properties of the test item can be described through a monotonic increasing function called the item information function.
- The response on the test item can be either isolated or continuous and it can be binary or bimodular. Item score categories can be ordered or unordered, and there can be one or many abilities behind the test performance.

IV. CHARACTERISTICS OF THE IRT MODELS

- The IRT model should be defined as the relationship between the observed response and the unobserved infrastructure (latent trait).
- The model should provide a method for estimating the degrees of the latent trait.
- The subjects' scores will be the basis for the assessment of the basic construction of the model.
- The IRT model assumes that the subject's performance can be predicted or explained by one latent trait or more.

In IRT is often assumed that the examinee has some unobservable latent trait (also called latent ability), which cannot be studied directly. The purpose of the IRT is to propose models that allow linking these underlying traits to some of the characteristics that can be observed on the subject [41]. There are many models in the IRT and they have been classified into two types: models that use the cumulative natural curve and logistic models. Logistic models are currently more widespread, they are suitable for two-stage items, and differ according to the number of the estimated item parameters [6, 31, 42-45]. However, there are three commonly used models for binary data, which use (1) for the correct response and (0) for the wrong response, and these models are the one parameter and the two-parameter logistic models, which will be examined below.

A. One-Parameter Logistic Model

The concept of information availability plays an important role in IRT as it can be used to evaluate how the item included in the test accurately measures the level of the latent trait with (parameter value θ_i). This latent trait could include, for example, the level of the student's knowledge, intelligence, ability, satisfaction, stress, etc. For example, in educational tests, the item parameter represents the difficulty of the item while the subject parameter represents the ability level of the people being evaluated. The greater the subject's ability in relation to the difficulty of the item (the parameter α_j describes the degree of difficulty of the item and the level of influence of the item on the subject), the greater the probability of a correct response to that item. Whereas, when the subject's position on the latent trait is equal to the difficulty of the item, according to Rasch's model, there is a 0.5 probability that a subject's response is correct. Accurate information about the value of θ_i depends on a number of factors, the most important of which is the properties of the questions (items) used to evaluate the parameter (the latent trait) [2, 30, 33, 46].

B. Two-Parameter Logistic Model

In the Two-Parameter Logistic (2PL) model, the situation is different from the one in the one-parameter model. The one-parameter model assumes that questions differ only with respect to item difficulty, whereas, in the two-parameter logistic model, two parameters are assumed to be connected to the test item: the parameter α_j which describes the difficulty of the item (question) and the additional parameter β_j , which describes the discrimination of the item. The parameter β (the slope of the curve) describes the degree to which the question helps to distinguish between the subjects with the highest level of a trait compared to those with a lower level of the same trait. This parameter also shows the extent of the relevance of the item to the overall score of the test. The higher the value of that parameter, the greater the discrimination of the items (and the easier it is to select subjects with a high level and those with a low concentration of the same trait). It should also be noted that the most difficult test item is not necessarily the test item with the highest potential to discriminate between the subjects [2, 19, 36, 47, 48].

C. Three-Parameter Logistic Model

The Three-Parameter Logistic (3PL) model is used in IRT, and it determines the probability of a correct response for a dichotomously scored multiple-choice item as a logistic distribution. The 3PL model is an extension of the 2PL logistic model as it introduces the guessing parameter. Items now differ in terms of discrimination, difficulty, and probability of guessing the correct response [47]. After adding the guessing parameter, denoted C_i , in the 3PL model, this parameter is the lower asymptote of the item characteristic curve and represents the probability of subjects who have a low ability to answer the item correctly. The parameter is included in the model to account for item response data from low-ability subjects, where guessing is a factor in test performance [48-50]. The basic equation for the 3PL model is the probability that a randomly selected examinee with a certain proficiency level on scale k will respond correctly to item j , which is characterized by discrimination (α_j), difficulty (β_j), and guessing probability (C_i) [27, 35, 37, 38, 51].

V. MULTIPLE CHOICE TEST ANALYSIS

Understanding how to interpret and use the information based on students' test scores is just as important as knowing how to create a well-designed test. An essential part of building tests is using the feedback from a good test analysis. Among the most important statistical information provided by a good analysis of a multiple-choice test are the following:

A. Item Difficulty

The test item difficulty factor β_j represents the percentage of the respondents who answered the item correctly. The difficulty factor ranges from 0.0 to 1.00. The higher the value of the difficulty factor, the easier the test item is. For example, when the value of the difficulty factor β_j is higher than 0.90, the test item is described as very easy and should not be used again in subsequent tests since almost all students are able to properly respond to it. Whereas when the value of the β_j is less than 0.20, the test item is described as extremely difficult and should be reviewed in subsequent tests. The optimal test item

difficulty factor is 0.50, and it insures maximum discrimination between high and low ability [52-54]. To maximize item discrimination, the desired difficulty levels are slightly higher than halfway between the probability of answering correctly by chance (1.00 divided by the number of alternatives for the item) and the ideal score for the item (1.00) [55-58]. For example, if the test item contains four alternatives to the answer, the probability of answering it correctly by chance would be 0.25 (1.00/4=0.25), and the ideal degree of difficulty for the item can be calculated by substituting in the following rule:

((Ideal score for item - probability of a correct answer by chance) / 2) + probability of a correct answer by chance

TABLE I. THE IDEAL DEGREE OF DIFFICULTY GUIDELINE

Design of the test item	Ideal degree of difficulty for the test item
Multiple choice (5 alternatives)	0.60
Multiple choice (4 alternatives)	0.62
Multiple choice (3 alternatives)	0.66
Multiple choice (2 alternatives)	0.75

B. Item Discrimination

The test item discrimination factor is referred to using the symbol α_j , as it represents the point relationship between the respondent's performance on the item and the respondents' total scores. The discrimination factor value ranges from -1.00 to 1.00. When the value of the test item discrimination factor is high, it indicates that the test item is able to distinguish between respondents. It distinguishes between those who scored high in the tests and were able to answer the test item correctly and those who obtained low test scores and were not able to respond to the item correctly [54, 59]. Test items that have point values close to or less than zero should be removed. Moreover, further consideration should be given to the item which was responded to better by those who generally performed poorly on the test than those who performed better on the test as a whole. The test item may be confusing in some way to top-performing respondents [52, 53, 58, 59].

TABLE II. THE IDEAL DEGREE OF DISCRIMINATION GUIDELINE

Discrimination factor value	Description of the test item
≥ 0.4	A very good test item
0.3 - 0.39	A good test item. Possible improvements may be considered
0.2 - 0.29	A fairly good test item. It is recommended to improve it
≤ 0.2	A weak test item, with the recommendation of deleting it
≤ 0	It is recommended to directly delete the item

VI. PSYCHOMETRIC PROPERTIES OF THE TEST

Psychometric properties are a statistical mechanism to verify the fairness, objectivity, and relevance of the test for the phenomenon to be measured. The most important psychometric properties are the following:

- The individual data series for all test components.
- The characteristics of the data collected for all test components. The fairness of the test lies in its freedom from

any bias and its suitability for the target group, regardless of gender, race, and religion. The psychometric properties of the test are tested to verify that it is objectively constructed and free of any bias.

Studying the rules for formulating multiple-choice test items is important because it has an impact on the level of performance on the items or the test as a whole. This means that the good construction of the test and the verification of all its psychometric properties ensures that the test avoids any violations in the structure of the items, which in turn affects the individual's performance on the test items [36, 60-62].

VII. METHODS

The descriptive survey method was used to obtain data from a real-life scenario of giving postgraduate-level midterm and final exams to assess master's students' achievement level in the subject of applied statistics. This analytical study aimed to determine the quality of the test, its efficiency, and the reliability of its results despite the varying circumstances in which it is given.

A. Measurements

A Criterion Referenced Test (CRT) was used to evaluate students' achievement in the course of applied statistics in a master's degree level in order to verify the quality of the test as a tool to evaluate the level of students' achievement. The test was developed and based on content analysis. The test specification table covered some of the course vocabulary for the applied statistics course. The test in its first form consisted of 25 test items that were reviewed by specialists in the field of statistics to determine their face validity, and 7 items were omitted as a result. So, in its final form, the test had 18 items, and they were applied to the study sample to verify their quality (Table III). The results of the thorough analysis of the test items were handed over to the central question bank in order to compare the performance of the test items and verify their lifespan when re-performing statistical operations on them later. To verify the test reliability, the Kuder Richardson (KR-20) method was used because the binary data are coded using 0 and 1 after correcting the items, and because the test items differ in their difficulty parameter. The results indicated that the test has a high reliability coefficient of $KR-20 = 0.842$.

B. Sample

The current study population consisted of all students of the applied statistics course at the master's level at Umm Al-Qura University on the main campus in Makkah and all branches of the University. The size of the population was rather large, estimated at about 400 male and female students registered during the second semester of the academic year 2020. It was difficult to reach all the sample members because of the financial cost, time, and effort required. Moreover, the educational and population environment conditions for all students are very similar and the previous studies using samples from students in the Saudi universities' population did not show any clear bias. Therefore, the current study used a random sample consisting of 338 students, equivalent to the 84.5% of the study population, studying different disciplines in the College of Education.

TABLE III. TEST ITEMS

Test Items	Alternatives (answers)
1. Which of the following is true for the interval scale level?	- Classification of individuals - Ranking of individuals and identifying differences - Both
2. If the value of the correlation coefficient is equal to (-0.8) this is an indication that the relationship is	- Weak - Nonexistent - Strong
3. Estimates that are calculated by studying the sample members are called	- Variables - Parameters - Statistics
4. When the population is homogeneous and its number is very large, we use the following type of sampling	- Simple Random Sampling - Stratified Random Sampling - Cluster Random Sampling
5. When studying the relationship between job performance and job satisfaction, job satisfaction is called	- Dependent variable - Independent variable - Intruder valuable
6. Which of the following statements is true for the relationship between the sample and the population?	- Population parameters are a good estimate of sample statistics - Population parameters are a good estimate of raw scores - Sample statistics are a good estimate of population parameters
7. If the relationship between chronological age and academic achievement is ($r = 0.91$) then this is evidence of	- The greater the age, the greater the achievement - The younger the age, the lesser the achievement - The one or the other
8. If the sum of the squares of the deviations from the mean = 80 and the number of students = 21 then the standard deviation equals	- 2 - 4 - 6
9. The range of the relationship that exists between two quantitative variables is called	- Slope - Connection - Range
10. The mode for the values (1, 4, 9, 12) is	- Zero - 12 - No mode
11. The median of the values (9, 15, 7, 10, 12) is	- 7 - 9 - 10
12. In one of the regions in KSA, a study was conducted on the pros and cons of the e-learning system at the undergraduate level. In this study the academic level is	- Variable - Constant - Other
13. The measure of central tendency not affected by outliers is	- The median - The mode - The one or the other
14. The number of training courses during a whole semester is a	- Continuous quantitative variable - Discrete quantitative variable - Descriptive variable
15. Achievement tests depend on the following level of measurement scale	- Ordinal - Interval - Ratio
17. When the sample size is increased	- It increases the probability of a normal distribution - It makes the sample to not comply with the normal distribution - There is no relation between them
18. Setting the confidence level at 95% is a prerequisite for educational sciences.	- Yes - It may take different values according to the nature of the study - Other

The sample distribution when responding to the test is represented in Tables IV and V. The members of the study

sample were contacted by e-mail, and a test link was created and made available on the student's electronic page. The link was made available for one hour, representing the testing period, and prior coordination with the study sample was made to select a time appropriate for everyone. Cases which had any type of technological problems were not recorded. Electronic reminders via the university's electronic system were used before the test to alert the study sample about the test time.

TABLE IV. DISTRIBUTION OF THE STUDY SAMPLE ACCORDING TO GENDER AND CURRENT MAJOR

Gender	Current major in the master's study					
	Psychology	Islamic Education	Curriculum and Instruction	Educational Administration	Special Education	Total
Male	45	38	37	30	30	180
Female	29	31	29	32	37	158
Total	74	69	66	62	67	338

TABLE V. DISTRIBUTION OF THE STUDY SAMPLE ACCORDING TO STUDYING STATISTICS IN EARLIER STAGES AND CURRENT MAJOR

Studied statistics in earlier stage	Current major in the master's study					
	Psychology	Islamic Education	Curriculum and Instruction	Educational Administration	Special Education	Total
Yes	56	51	52	42	34	235
No	18	18	14	20	33	103
Total	74	69	66	62	67	338

VIII. RESULTS

Figure 1 show the eigenvalue scree plot. It is clear that the first eigenvalue is much greater than the others, suggesting that a unidimensional model is reasonable for this data.

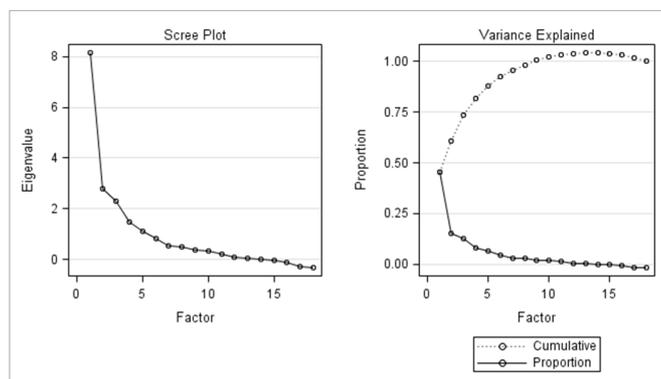


Fig. 1. Eigenvalue scree plot.

This test has items with one correct alternative answer that worths a single point, the item difficulty was simply the percentage of students who answer an item correctly, which is

equal to the item mean. In this case, the item difficulty index in Tables VI to X shows the ranges, based on the examinees' characteristics. The item difficulty ranged from 0.647 to 0.928. For students who studied statistics in earlier stages, the items' difficulties ranged from 0.878 to 0.970, however for students who did not, the items' difficulties ranged from 0.310 to 0.893, while the items' difficulties based on gender ranged from 0.650 to 0.966 for male and 0.645 to 0.924 for the female students. Regarding the students' current major, the item difficulty varied over subjects: for special education, it ranged from 0.522 to 0.985, for educational administration it ranged from 0.50 to 0.870, for curriculum and instruction it ranged from 0.651 to 0.999, for Islamic education it ranged from 0.521 to 0.927, and for psychology it ranged from 0.837 to 0.999. For higher GPA the items' difficulties ranged from 0.857 to 0.994, whereas for moderate GPA they ranged from 0.400 to 0.936, and for low GPA they ranged from 0.021 to 0.869.

TABLE VI. OVERALL ITEM DIFFICULTY

Items	Overall
Q1	0.837
Q2	0.668
Q3	0.843
Q4	0.751
Q5	0.647
Q6	0.928
Q7	0.798
Q8	0.763
Q9	0.917
Q10	0.857
Q11	0.757
Q12	0.792
Q13	0.784
Q14	0.739
Q15	0.695
Q16	0.825
Q17	0.786
Q18	0.828

TABLE VII. ITEM DIFFICULTY ACCORDING TO STUDYING STATISTICS IN AN EARLIER STAGE

Items	Studying statistics in earlier stages	
	Yes	No
Q1	0.927	0.631
Q2	0.787	0.398
Q3	0.970	0.553
Q4	0.842	0.543
Q5	0.795	0.310
Q6	0.944	0.893
Q7	0.897	0.572
Q8	0.765	0.757
Q9	0.974	0.786
Q10	0.944	0.660
Q11	0.834	0.580
Q12	0.868	0.621
Q13	0.880	0.563
Q14	0.872	0.436
Q15	0.825	0.398
Q16	0.910	0.631
Q17	0.842	0.660
Q18	0.893	0.679

Figures 2 - 19 present the combined curve for the 18 items based on the overall data. Each item has one threshold. Most thresholds span the negative section of the ability. Item information curves show that the items provide a good amount of information for students who had a lower or moderate ability compared to high ability students. In terms of precision, most of the items were more convenient to lower ability students (e.g. items 9, 10, 17), while items 2, 5, 14 gathered more information for students who had moderate ability.

TABLE VIII. ITEMS DIFFICULTY ACCORDING TO GENDER

Items	Gender	
	Male	Female
Q1	0.844	0.829
Q2	0.661	0.677
Q3	0.872	0.810
Q4	0.761	0.740
Q5	0.650	0.645
Q6	0.933	0.924
Q7	0.827	0.765
Q8	0.688	0.848
Q9	0.966	0.860
Q10	0.861	0.854
Q11	0.772	0.740
Q12	0.816	0.765
Q13	0.850	0.708
Q14	0.777	0.696
Q15	0.755	0.626
Q16	0.888	0.753
Q17	0.850	0.715
Q18	0.838	0.816

TABLE IX. ITEMS DIFFICULTY ACCORDING TO THE CURRENT MAJOR

Items	Current master's study major				
	Special Education	Educational Administration	Curriculum and Instruction	Islamic Education	Psychology
Q1	0.761	0.822	0.939	0.768	0.891
Q2	0.522	0.645	0.712	0.521	0.918
Q3	0.567	0.838	0.999	0.840	0.959
Q4	0.731	0.693	0.848	0.623	0.851
Q5	0.507	0.500	0.651	0.623	0.918
Q6	0.985	0.725	0.954	0.956	0.999
Q7	0.776	0.741	0.772	0.753	0.932
Q8	0.820	0.822	0.893	0.608	0.689
Q9	0.865	0.838	0.969	0.927	0.972
Q10	0.656	0.870	0.984	0.782	0.986
Q11	0.656	0.661	0.939	0.521	0.986
Q12	0.761	0.709	0.954	0.637	0.891
Q13	0.776	0.741	0.787	0.652	0.945
Q14	0.567	0.596	0.863	0.710	0.932
Q15	0.582	0.516	0.742	0.739	0.864
Q16	0.791	0.629	0.878	0.869	0.932
Q17	0.776	0.677	0.772	0.855	0.837
Q18	0.791	0.725	0.893	0.840	0.873

Figure 20 represents the test characteristic curve that was the functional relation between the true score and the ability scale. As we can see the probability of the correct response was near to 1 at the lowest levels of ability and it increased until it

came to the highest levels of ability. The probability of correct response in this test is about 18 for high ability students. Figure 21 presents the total amount of the information that has been obtained from the test. It appears clearly that the test gives good indicators to assess lower levels of agreement.

TABLE X. ITEM DIFFICULTY ACCORDING TO THE GPA

Items	GPA		
	≥ 3.75	3.30-3.75	2.5 – 3.30
Q1	0.950	0.809	0.456
Q2	0.950	0.427	0.130
Q3	0.989	0.872	0.195
Q4	0.928	0.581	0.456
Q5	0.956	0.400	0.021
Q6	0.994	0.845	0.869
Q7	0.939	0.700	0.478
Q8	0.857	0.645	0.673
Q9	0.983	0.936	0.608
Q10	0.939	0.854	0.304
Q11	0.972	0.554	0.282
Q12	0.956	0.627	0.543
Q13	0.956	0.627	0.478
Q14	0.972	0.627	0.086
Q15	0.939	0.527	0.130
Q16	0.950	0.745	0.521
Q17	0.923	0.700	0.456
Q18	0.934	0.781	0.521

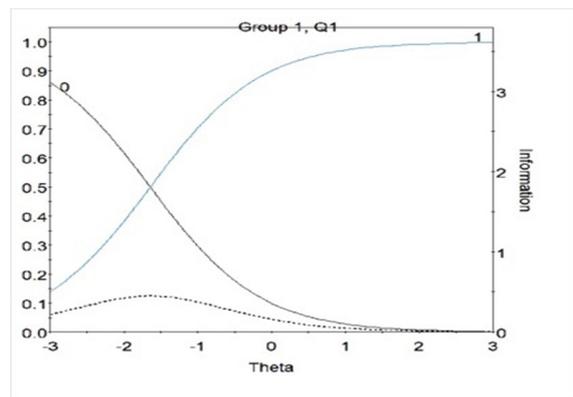


Fig. 2. Characteristic curve of item 1.

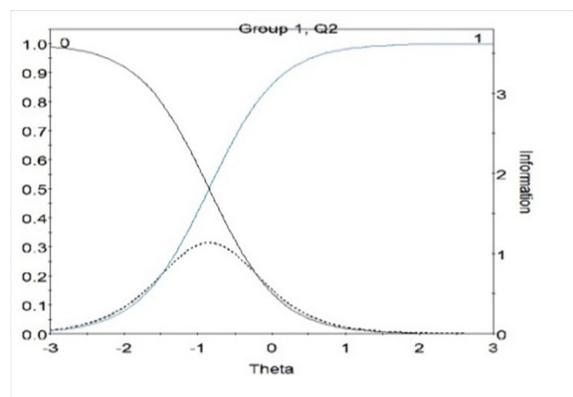


Fig. 3. Characteristic curve of item 2.

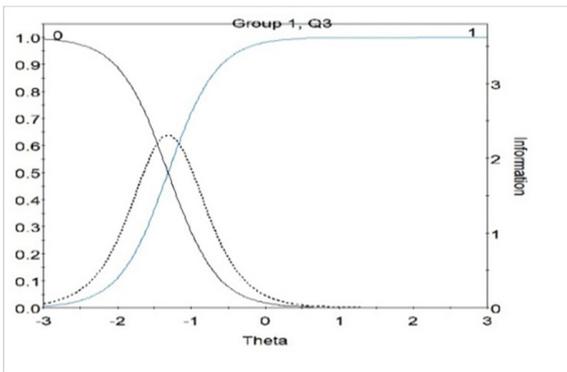


Fig. 4. Characteristic curve of item 3.

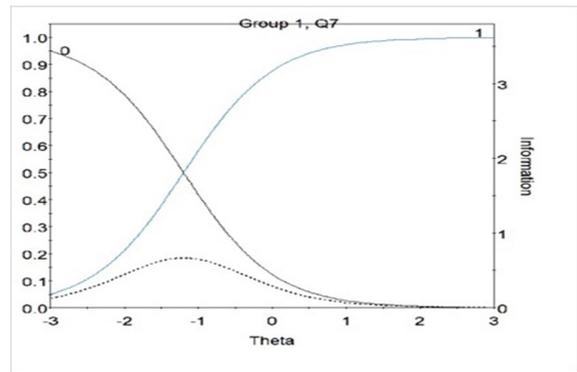


Fig. 8. Characteristic curve of item 7.

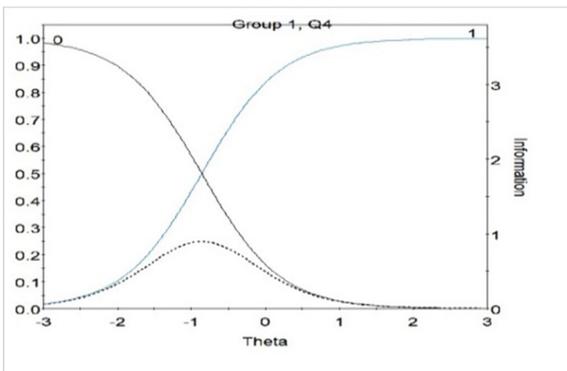


Fig. 5. Characteristic curve of item 4.

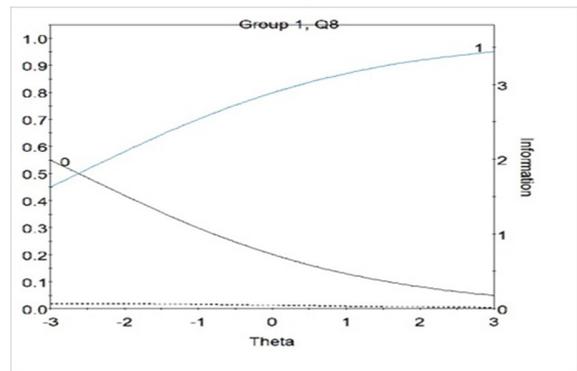


Fig. 9. Characteristic curve of item 8.

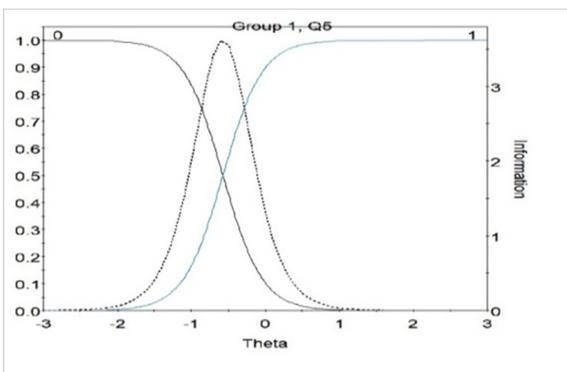


Fig. 6. Characteristic curve of item 5.

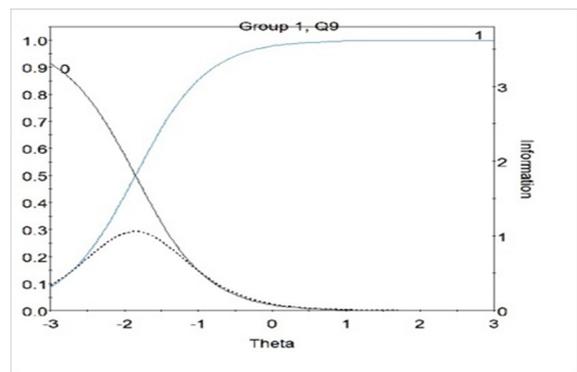


Fig. 10. Characteristic curve of item 9.

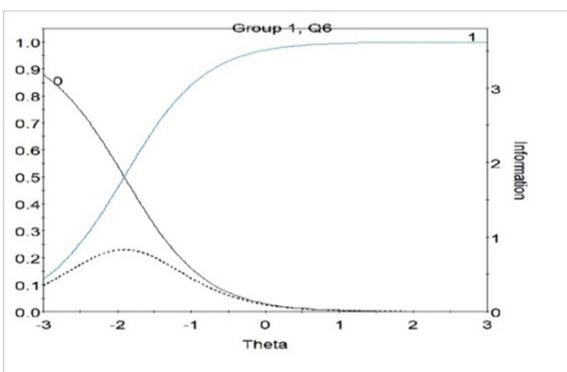


Fig. 7. Characteristic curve of item 6.

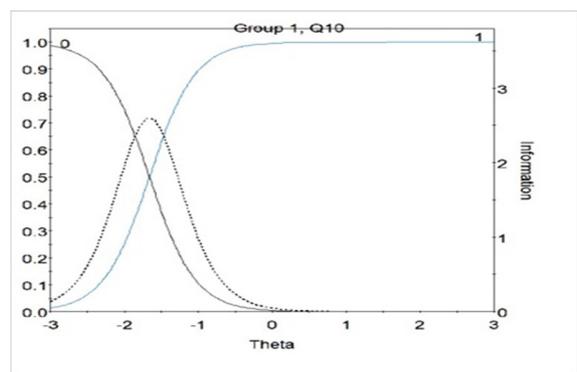


Fig. 11. Characteristic curve of item 10.

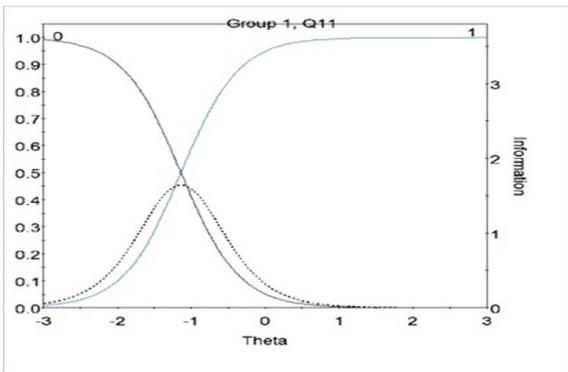


Fig. 12. Characteristic curve of item 11.

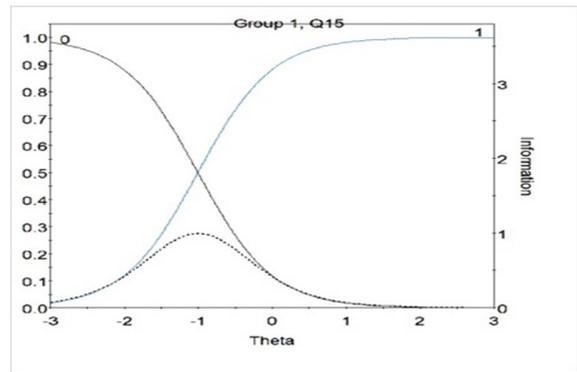


Fig. 16. Characteristic curve of item 15.

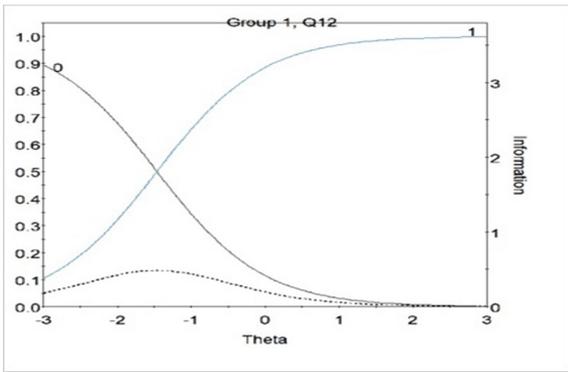


Fig. 13. Characteristic curve of item 12.

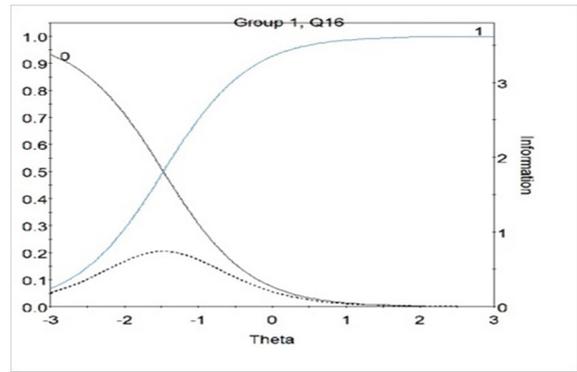


Fig. 17. Characteristic curve of item 16.

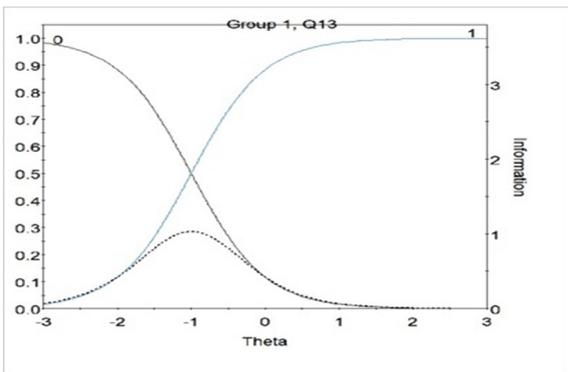


Fig. 14. Characteristic curve of item 13.

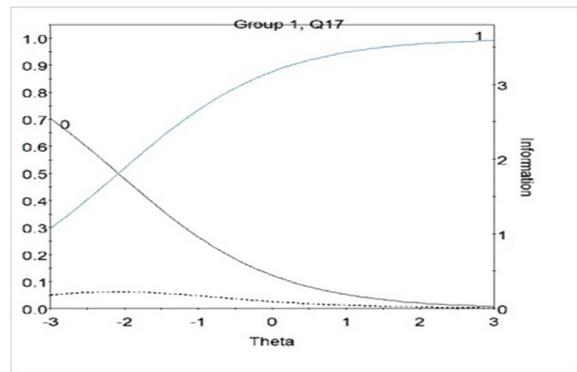


Fig. 18. Characteristic curve of item 17.

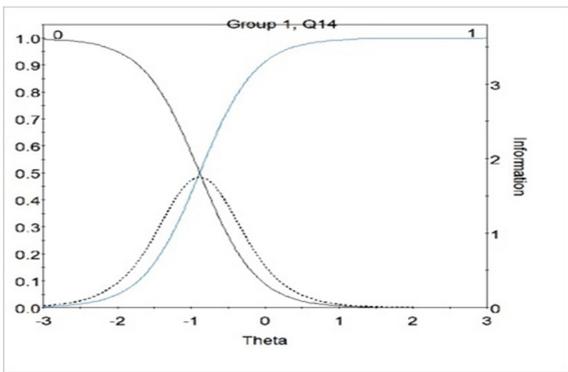


Fig. 15. Characteristic curve of item 14.

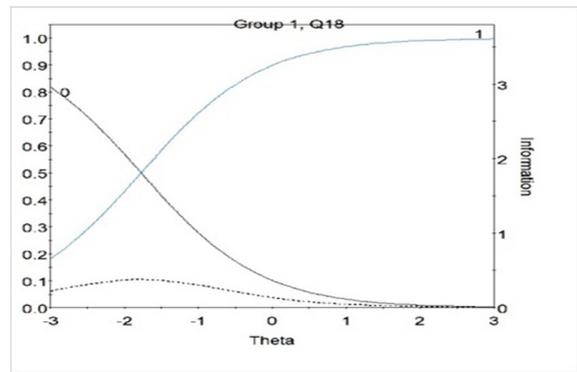


Fig. 19. Characteristic curve of item 18.

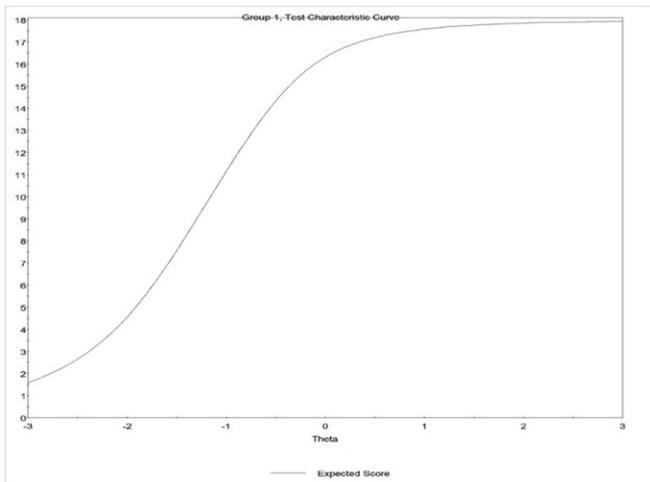


Fig. 20. Test characteristic curve between the true score and the ability scale.

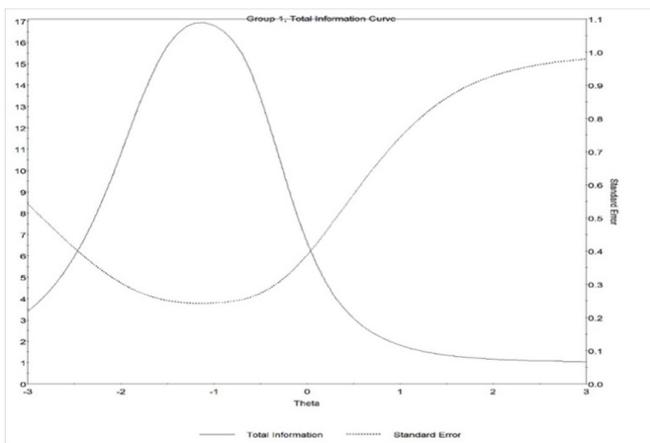


Fig. 21. Test characteristic curve.

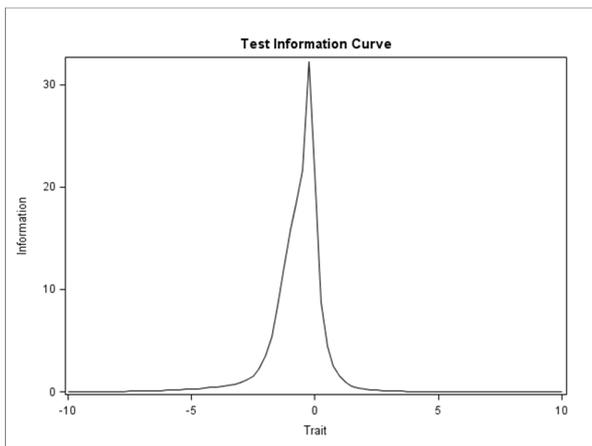


Fig. 22. Test information curve according to male examinees.

Figures 22 to 33 represent the test characteristic curve according to the change in the examinees' characteristics. The amount of information that has been obtained by female students appeared to be more than the information extracted by male and the test provided more information about students

who were not studying statistics in earlier levels compared to students who were. Figures 27 and 28 also confirm that the test provides a good amount of information for students who had a lower or moderate ability compared to students who had high ability.

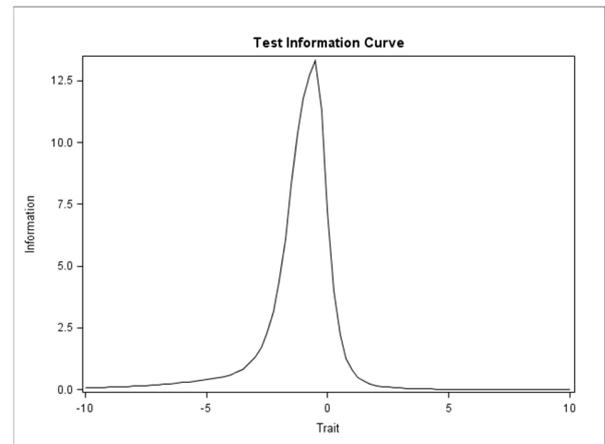


Fig. 23. Test information curve according to female examinees.

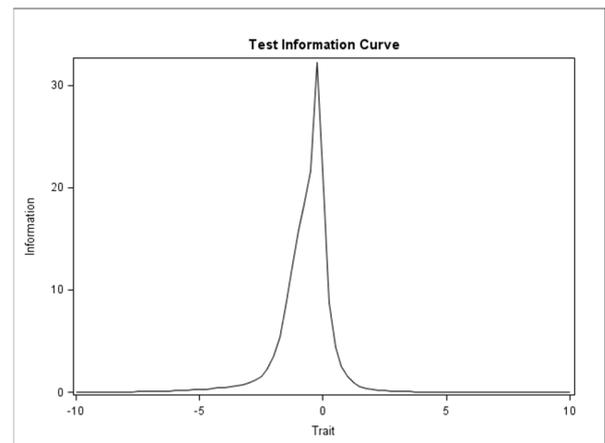


Fig. 24. Test information curve according to examinees who were studying statistics in earlier stage.

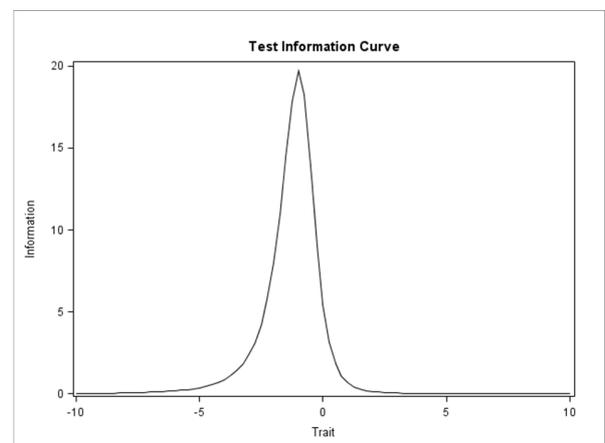


Fig. 25. Test information curve according to examinees who were not studying statistics in an earlier stage.

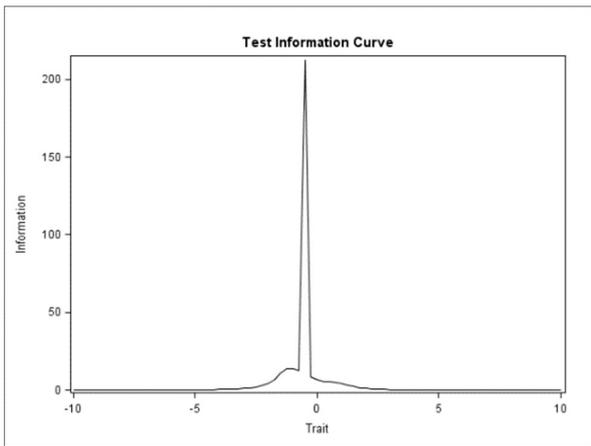


Fig. 26. Test information curve according to examinees who had a GPA of 3.75 and above.

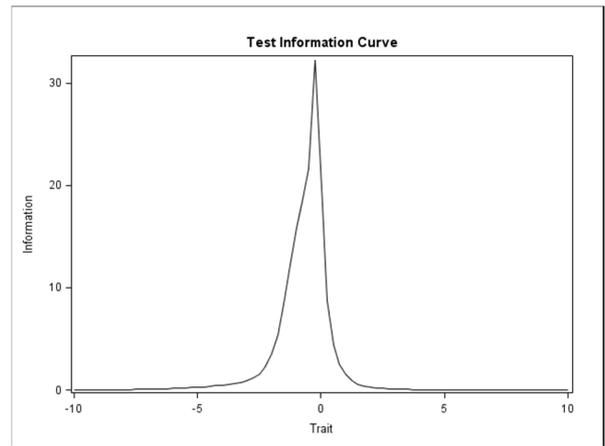


Fig. 29. Test information curve according to current major (Special Education).

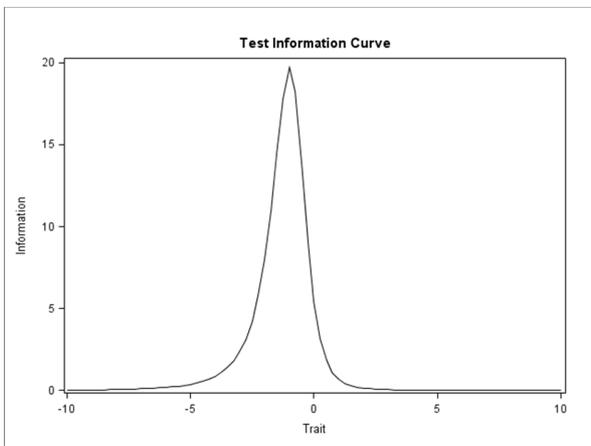


Fig. 27. Test information curve according to examinees who had a GPA in the range from 3.30 to less than 3.75.

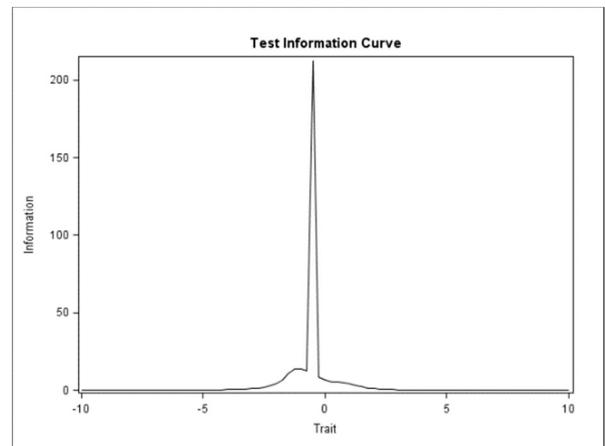


Fig. 30. Test information curve according to current major (Educational Administration).

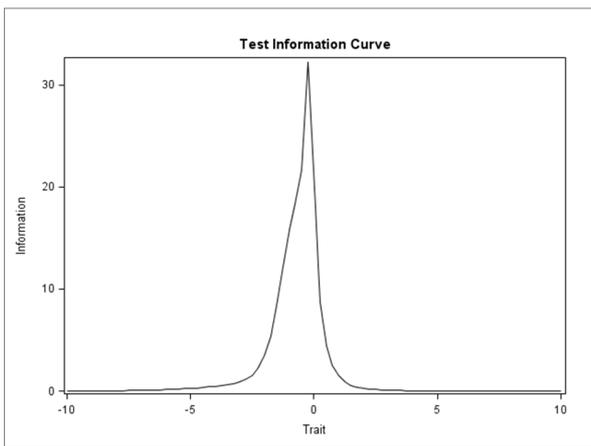


Fig. 28. Test information curve according to examinees who had a GPA in the range from 2.50 to less than 3.30.

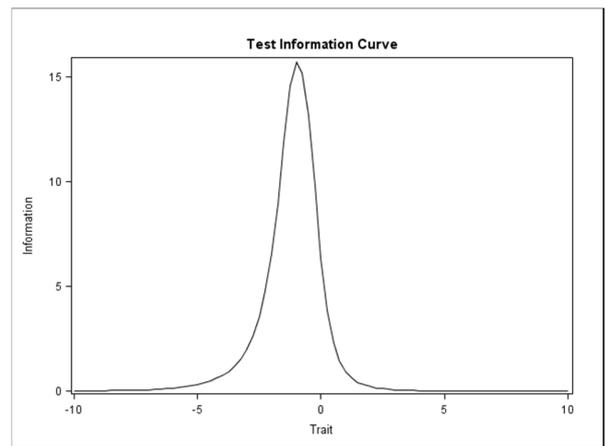


Fig. 31. Test information curve according to current major (Curriculum and Instruction).

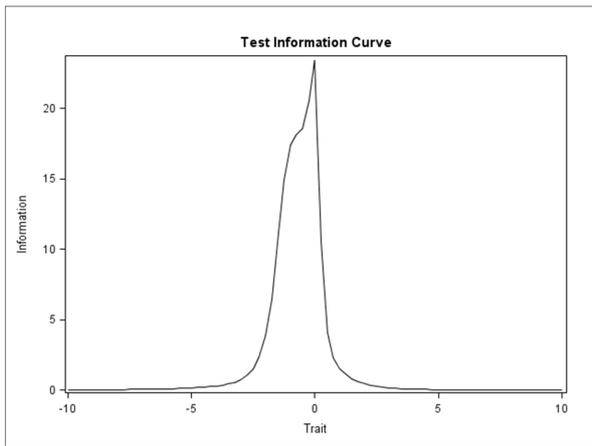


Fig. 32. Test information curve according to current major (Islamic Education).

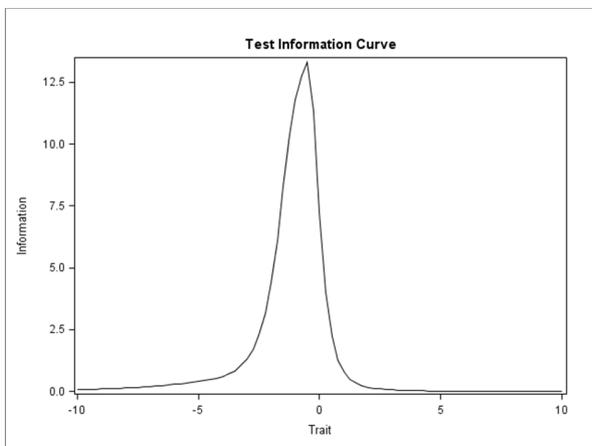


Fig. 33. Test information curve according to current major (Psychology).

IX. DISCUSSION

The study findings support previous researches [5-7, 63] for the use of IRT models in analyzing students' responses to an achievement test in a more objective way, showing whether there is an effect of the multiplicity of characteristics of the participants on the test items (multiple choice) in terms of the accuracy of the estimates of the items' parameters and the individuals' ability parameters. The results also help determine the relationship between individual responses and their basic ability. Within the framework of modern measurement theory, many models and applications have been formulated and applied to real test data. This includes the measurement assumptions about the characteristics of the test items, the performance of the subjects, and how performance is related to knowledge. This test clearly indicated that based on the level of the statistics course, there should be a periodic review of the tests in line with the nature and level of the course materials in order to have a logical judgment about the level of students' progress and the level of their ability. This conclusion is consistent with the findings of [50, 64].

X. CONCLUSION

In general, this study seeks to further improve the evaluation practices in the educational process. The tests that describe the students' progress in the educational process should be subject to review by evaluation and measurement specialists in order to ensure that we have valid and reliable evaluation tools. The administrators of the educational system need to find a mechanism to review question banks and align them with the requirements of the scientific material of the courses that are subject to continuous development.

REFERENCES

- [1] B. Zhuang, S. Wang, S. Zhao, and M. Lu, "Computed tomography angiography-derived fractional flow reserve (CT-FFR) for the detection of myocardial ischemia with invasive fractional flow reserve as reference: systematic review and meta-analysis," *European Radiology*, vol. 30, no. 2, pp. 712–725, Feb. 2020, <https://doi.org/10.1007/s00330-019-06470-8>.
- [2] Y. A. Wang and M. Rhemtulla, "Power Analysis for Parameter Estimation in Structural Equation Modeling: A Discussion and Tutorial," in *Advances in Methods and Practices in Psychological Science*, California, USA: University of California, 2020.
- [3] H. Zhu, W. Gao, and X. Zhang, "Bayesian Analysis of a Quantile Multilevel Item Response Theory Model," *Frontiers in Psychology*, vol. 11, Jan. 2021, Art. no. 607731, <https://doi.org/10.3389/fpsyg.2020.607731>.
- [4] M. R. Szeles, "Examining the foreign policy attitudes in Moldova," *PLOS ONE*, vol. 16, no. 1, 2021, Art. no. e0245322, <https://doi.org/10.1371/journal.pone.0245322>.
- [5] D. Almaleki, "The Precision of the Overall Data-Model Fit for Different Design Features in Confirmatory Factor Analysis," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6766–6774, Feb. 2021, <https://doi.org/10.48084/etasr.4025>.
- [6] D. Almaleki, "Empirical Evaluation of Different Features of Design in Confirmatory Factor Analysis," Ph.D. dissertation, Western Michigan University, MC, USA, 2016.
- [7] C. S. Wardley, E. B. Applegate, A. D. Almaleki, and J. A. Van Rhee, "A Comparison of Students' Perceptions of Stress in Parallel Problem-Based and Lecture-Based Curricula," *The Journal of Physician Assistant Education*, vol. 27, no. 1, pp. 7–16, Mar. 2016, <https://doi.org/10.1097/JPA.0000000000000060>.
- [8] C. Wardley, E. Applegate, A. Almaleki, and J. V. Rhee, "Is Student Stress Related to Personality or Learning Environment in a Physician Assistant Program?," *The Journal of Physician Assistant Education*, vol. 30, no. 1, pp. 9–19, Mar. 2019, <https://doi.org/10.1097/JPA.0000000000000241>.
- [9] A. C. Villa Montoya *et al.*, "Optimization of key factors affecting hydrogen production from coffee waste using factorial design and metagenomic analysis of the microbial community," *International Journal of Hydrogen Energy*, vol. 45, no. 7, pp. 4205–4222, Feb. 2020, <https://doi.org/10.1016/j.ijhydene.2019.12.062>.
- [10] N. M. Moo-Tun, G. Iniguez-Covarrubias, and A. Valadez-Gonzalez, "Assessing the effect of PLA, cellulose microfibrils and CaCO₃ on the properties of starch-based foams using a factorial design," *Polymer Testing*, vol. 86, Jun. 2020, Art. no. 106482, <https://doi.org/10.1016/j.polymertesting.2020.106482>.
- [11] K. M. Marcoulides, N. Foldnes, and S. Grønneberg, "Assessing Model Fit in Structural Equation Modeling Using Appropriate Test Statistics," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 27, no. 3, pp. 369–379, May 2020, <https://doi.org/10.1080/10705511.2019.1647785>.
- [12] M. D. H. Naveiras, "Using Auxiliary Item Information in the Item Parameter Estimation of a Graded Response Model for a Small to Medium Sample Size: Empirical versus Hierarchical Bayes Estimation," Ph.D. dissertation, Vanderbilt University, Nashville, TN, USA, 2020.

- [13] M. N. Morshed, M. N. Pervez, N. Behary, N. Bouazizi, J. Guan, and V. A. Nierstrasz, "Statistical modeling and optimization of heterogeneous Fenton-like removal of organic pollutant using fibrous catalysts: a full factorial design," *Scientific Reports*, vol. 10, no. 1, Sep. 2020, Art. no. 16133, <https://doi.org/10.1038/s41598-020-72401-z>.
- [14] W. van Lankveld, R. J. Pat-El, N. van Melick, R. van Cingel, and J. B. Staal, "Is Fear of Harm (FoH) in Sports-Related Activities a Latent Trait? The Item Response Model Applied to the Photographic Series of Sports Activities for Anterior Cruciate Ligament Rupture (PHOSA-ACLR)," *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, Sep. 2020, Art. no. 6764, <https://doi.org/10.3390/ijerph17186764>.
- [15] C. Shin, S.-H. Lee, K.-M. Han, H.-K. Yoon, and C. Han, "Comparison of the Usefulness of the PHQ-8 and PHQ-9 for Screening for Major Depressive Disorder: Analysis of Psychiatric Outpatient Data," *Psychiatry Investigation*, vol. 16, no. 4, pp. 300–305, Apr. 2019, <https://doi.org/10.30773/pi.2019.02.01>.
- [16] C. W. Ong, B. G. Pierce, D. W. Woods, M. P. Twohig, and M. E. Levin, "The Acceptance and Action Questionnaire – II: an Item Response Theory Analysis," *Journal of Psychopathology and Behavioral Assessment*, vol. 41, no. 1, pp. 123–134, Mar. 2019, <https://doi.org/10.1007/s10862-018-9694-2>.
- [17] A. Acevedo-Mesa, J. N. Tendeiro, A. Roest, J. G. M. Rosmalen, and R. Monden, "Improving the Measurement of Functional Somatic Symptoms With Item Response Theory," *Assessment*, Aug. 2020, Art. no. 1073191120947153, <https://doi.org/10.1177/1073191120947153>.
- [18] J. Xia, Z. Tang, P. Wu, J. Wang, and J. Yu, "Use of item response theory to develop a shortened version of the EORTC QLQ-BR23 scales," *Scientific Reports*, vol. 9, no. 1, Feb. 2019, Art. no. 1764, <https://doi.org/10.1038/s41598-018-37965-x>.
- [19] Y. Liu and J. S. Yang, "Interval Estimation of Latent Variable Scores in Item Response Theory," *Journal of Educational and Behavioral Statistics*, vol. 43, no. 3, pp. 259–285, Jun. 2018, <https://doi.org/10.3102/1076998617732764>.
- [20] U. Gromping, "Coding invariance in factorial linear models and a new tool for assessing combinatorial equivalence of factorial designs," *Journal of Statistical Planning and Inference*, vol. 193, pp. 1–14, Feb. 2018, <https://doi.org/10.1016/j.jspi.2017.07.004>.
- [21] P. J. Ferrando and U. Lorenzo-Seva, "Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis," *Educational and Psychological Measurement*, vol. 78, no. 5, pp. 762–780, Oct. 2018, <https://doi.org/10.1177/0013164417719308>.
- [22] X. An and Y.-F. Yung, "Item Response Theory: What It Is and How You Can Use the IRT Procedure to Apply It," SAS Institute Inc., Paper SAS364-2014.
- [23] K. Coughlin, "An Analysis of Factor Extraction Strategies: A Comparison of the Relative Strengths of Principal Axis, Ordinary Least Squares, and Maximum Likelihood in Research Contexts that Include both Categorical and Continuous Variables," Ph.D. dissertation, University of South Florida, Tampa, FL, USA, 2013.
- [24] D. L. Bandalos and P. Gagne, "Simulation methods in structural equation modeling," in *Handbook of structural equation modeling*, New York, NY, USA: The Guilford Press, 2012, pp. 92–108.
- [25] J. C. F. de Winter, D. Dodou, and P. A. Wieringa, "Exploratory Factor Analysis With Small Sample Sizes," *Multivariate Behavioral Research*, vol. 44, no. 2, pp. 147–181, Apr. 2009, <https://doi.org/10.1080/00273170902794206>.
- [26] J. D. Kechagias, K.-E. Aslani, N. A. Fountas, N. M. Vaxevanidis, and D. E. Manolakos, "A comparative investigation of Taguchi and full factorial design for machinability prediction in turning of a titanium alloy," *Measurement*, vol. 151, Feb. 2020, Art. no. 107213, <https://doi.org/10.1016/j.measurement.2019.107213>.
- [27] G. Kuan, A. Sabo, S. Sawang, and Y. C. Kueh, "Factorial validity, measurement and structure invariance of the Malay language decisional balance scale in exercise across gender," *PLOS ONE*, vol. 15, no. 3, 2020, Art. no. e0230644, <https://doi.org/10.1371/journal.pone.0230644>.
- [28] M. J. Allen and W. M. Yen, *Introduction to measurement theory*. Monterey, CA, USA: Cole Publishing, 1979.
- [29] O. P. John and S. Srivastava, "The Big Five trait taxonomy: History, measurement, and theoretical perspectives," in *Handbook of personality: Theory and research*, New York, NY, USA: Guilford Press, 1999, pp. 102–138.
- [30] S.-H. Joo, L. Khorrarnadel, K. Yamamoto, H. J. Shin, and F. Robin, "Evaluating Item Fit Statistic Thresholds in PISA: Analysis of Cross-Country Comparability of Cognitive Items," *Educational Measurement: Issues and Practice*, Nov. 2020, <https://doi.org/10.1111/emip.12404>.
- [31] H. Bourdeaud'hui, "Investigating the effects of presenting listening test items in a singular versus dual mode on students' critical listening performance," in *Upper-primary school students' listening skills: Assessment and the relationship with student and class-level characteristics*, Ghent, Belgium: Ghent University, 2019.
- [32] D. M. Dimitrov and Y. Luo, "A Note on the D-Scoring Method Adapted for Polytomous Test Items," *Educational and Psychological Measurement*, vol. 79, no. 3, pp. 545–557, Jun. 2019, <https://doi.org/10.1177/0013164418786014>.
- [33] J. Suarez-Alvarez, I. Pedrosa, L. Lozano, E. Garcia-Cueto, M. Cuesta, and J. Muniz, "Using reversed items in Likert scales: A questionable practice," *Psicothema*, vol. 30, no. 2, pp. 149–158, 2018, <https://doi.org/10.7334/psicothema2018.33>.
- [34] J. P. Lalor, H. Wu, and H. Yu, "Learning Latent Parameters without Human Response Patterns: Item Response Theory with Artificial Crowds," in *Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, Hong Kong, China, Nov. 2019, pp. 4240–4250, <https://doi.org/10.18653/v1/D19-1434>.
- [35] B. Couvy-Duchesne, T. A. Davenport, N. G. Martin, M. J. Wright, and I. B. Hickie, "Validation and psychometric properties of the Somatic and Psychological Health Report (SPHERE) in a young Australian-based population sample using non-parametric item response theory," *BMC Psychiatry*, vol. 17, no. 1, Aug. 2017, Art. no. 279, <https://doi.org/10.1186/s12888-017-1420-1>.
- [36] P. M. Bentler and D. G. Bonett, "Significance tests and goodness of fit in the analysis of covariance structures," *Psychological Bulletin*, vol. 88, no. 3, pp. 588–606, 1980, <https://doi.org/10.1037/0033-2909.88.3.588>.
- [37] A. Schimmenti, L. Sideli, L. L. Marca, A. Gori, and G. Terrone, "Reliability, Validity, and Factor Structure of the Maladaptive Daydreaming Scale (MDS-16) in an Italian Sample," *Journal of Personality Assessment*, vol. 102, no. 5, pp. 689–701, Sep. 2020, <https://doi.org/10.1080/00223891.2019.1594240>.
- [38] C.-Y. Lin, V. Imani, M. D. Griffiths, and A. H. Pakpour, "Validity of the Yale Food Addiction Scale for Children (YFAS-C): Classical test theory and item response theory of the Persian YFAS-C," *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, Jul. 2020, <https://doi.org/10.1007/s40519-020-00956-x>.
- [39] L. Jiang *et al.*, "The Reliability and Validity of the Center for Epidemiologic Studies Depression Scale (CES-D) for Chinese University Students," *Frontiers in Psychiatry*, vol. 10, 2019, Art. no. 315, <https://doi.org/10.3389/fpsy.2019.00315>.
- [40] S. Doi, M. Ito, Y. Takebayashi, K. Muramatsu, and M. Horikoshi, "Factorial validity and invariance of the Patient Health Questionnaire (PHQ)-9 among clinical and non-clinical populations," *PLOS ONE*, vol. 13, no. 7, 2018, Art. no. e0199235.
- [41] T. Tsubakita, K. Shimazaki, H. Ito, and N. Kawazoe, "Item response theory analysis of the Utrecht Work Engagement Scale for Students (UWES-S) using a sample of Japanese university and college students majoring medical science, nursing, and natural science," *BMC Research Notes*, vol. 10, no. 1, Oct. 2017, Art. no. 528, <https://doi.org/10.1186/s13104-017-2839-7>.
- [42] S. C. Smid, D. McNeish, M. Miocevic, and R. van de Schoot, "Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 27, no. 1, pp. 131–161, Jan. 2020, <https://doi.org/10.1080/10705511.2019.1577140>.
- [43] M. K. Cain and Z. Zhang, "Fit for a Bayesian: An Evaluation of PPP and DIC for Structural Equation Modeling," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 26, no. 1, pp. 39–50, Jan. 2019, <https://doi.org/10.1080/10705511.2018.1490648>.

- [44] D. Garson, "StatNotes: Topics in Multivariate Analysis," *North Carolina State University*. <https://faculty.chass.ncsu.edu/garson/PA765/statnote.htm> (accessed Feb. 10, 2021).
- [45] H. W. Marsh, K.-T. Hau, and D. Grayson, "Goodness of Fit in Structural Equation Models," in *Contemporary psychometrics: A festschrift for Roderick P. McDonald*, Mahwah, NJ, USA: Lawrence Erlbaum Associates Publishers, 2005, pp. 275–340.
- [46] I. Williams, "A speededness item response model for associating ability and speededness parameters," Ph.D. dissertation, Rutgers University, New Brunswick, NJ, USA, 2017.
- [47] B. Shamshad and J. S. Siddiqui, "Testing Procedure for Item Response Probabilities of 2Class Latent Model," *Mehran University Research Journal of Engineering and Technology*, vol. 39, no. 3, pp. 657–667, Jul. 2020, <https://doi.org/10.22581/muet1982.2003.20>.
- [48] K. M. Williams and B. D. Zumbo, "Item Characteristic Curve Estimation of Signal Detection Theory-Based Personality Data: A Two-Stage Approach to Item Response Modeling," *International Journal of Testing*, vol. 3, no. 2, pp. 189–213, Jun. 2003, https://doi.org/10.1207/S15327574IJT0302_7.
- [49] D. Tafiadis *et al.*, "Using Receiver Operating Characteristic Curve to Define the Cutoff Points of Voice Handicap Index Applied to Young Adult Male Smokers," *Journal of Voice*, vol. 32, no. 4, pp. 443–448, Jul. 2018, <https://doi.org/10.1016/j.jvoice.2017.06.007>.
- [50] L. Lina, D. Mardapi, and H. Haryanto, "Item Characteristics on PROTEFL Listening Section," presented at the First International Conference on Advances in Education, Humanities, and Language, ICEL 2019, Malang, Indonesia, 23-24 March 2019, Jul. 2019, <https://dx.doi.org/10.4108/eai.11-7-2019.159630>.
- [51] D. L. Moody, "The method evaluation model: a theoretical model for validating information systems design methods," in *European Conference on Information Systems*, Naples, Italy, Jun. 2003, pp. 1–17.
- [52] H. Davis, T. M. Rosner, M. C. D'Angelo, E. MacLellan, and B. Milliken, "Selective attention effects on recognition: the roles of list context and perceptual difficulty," *Psychological Research*, vol. 84, no. 5, pp. 1249–1268, Jul. 2020, <https://doi.org/10.1007/s00426-019-01153-x>.
- [53] L. Sun, Y. Liu, and F. Luo, "Automatic Generation of Number Series Reasoning Items of High Difficulty," *Frontiers in Psychology*, vol. 10, 2019, Art. no. 884, <https://doi.org/10.3389/fpsyg.2019.00884>.
- [54] T. O. Abe and E. O. Omole, "Difficulty and Discriminating Indices of Junior Secondary School Mathematics Examination; A Case Study of Oriade Local Government, Osun State," *American Journal of Education and Information Technology*, vol. 3, no. 2, pp. 37–46, Oct. 2019, <https://doi.org/10.11648/j.ajeit.20190302.12>.
- [55] G. Nelson and S. R. Powell, "Computation Error Analysis: Students With Mathematics Difficulty Compared To Typically Achieving Students," *Assessment for Effective Intervention*, vol. 43, no. 3, pp. 144–156, Jun. 2018, <https://doi.org/10.1177/1534508417745627>.
- [56] H. Retnawati, B. Kartowagiran, J. Arlinwibowo, and E. Sulistyarningsih, "Why Are the Mathematics National Examination Items Difficult and What Is Teachers' Strategy to Overcome It?," *International Journal of Instruction*, vol. 10, no. 3, pp. 257–276, Jul. 2017.
- [57] T. A. Holster, J. W. Lake, and W. R. Pellowe, "Measuring and predicting graded reader difficulty," vol. 29, no. 2, pp. 218–244, Oct. 2017.
- [58] S. Gaitas and M. A. Martins, "Teacher perceived difficulty in implementing differentiated instructional strategies in primary school," *International Journal of Inclusive Education*, vol. 21, no. 5, pp. 544–556, May 2017, <https://doi.org/10.1080/13603116.2016.1223180>.
- [59] J. L. D'Sa and M. L. Visbal-Dionaldo, "Analysis of Multiple Choice Questions: Item Difficulty, Discrimination Index and Distractor Efficiency," *International Journal of Nursing Education*, vol. 9, no. 3, pp. 109–114, 2017.
- [60] A. H. Blasi and M. Alsuwaiket, "Analysis of Students' Misconducts in Higher Education using Decision Tree and ANN Algorithms," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6510–6514, Dec. 2020, <https://doi.org/10.48084/etasr.3927>.
- [61] N. Sharifi, M. Falsafi, N. Farokhi, and E. Jamali, "Assessing the optimal method of detecting Differential Item Functioning in Computerized Adaptive Testing," *Quarterly of Educational Measurement*, vol. 9, no. 33, pp. 23–51, Oct. 2018, <https://doi.org/10.22054/jem.2019.11109.1323>.
- [62] J. J. Hox, C. J. M. Maas, and M. J. S. Brinkhuis, "The effect of estimation method and sample size in multilevel structural equation modeling," *Statistica Neerlandica*, vol. 64, no. 2, pp. 157–170, 2010, <https://doi.org/10.1111/j.1467-9574.2009.00445.x>.
- [63] G. Makransky, L. Lilleholt, and A. Aaby, "Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach," *Computers in Human Behavior*, vol. 72, pp. 276–285, Jul. 2017, <https://doi.org/10.1016/j.chb.2017.02.066>.
- [64] J. A. Costa, J. Maroco, and J. Pinto-Gouveia, "Validation of the psychometric properties of cognitive fusion questionnaire. A study of the factorial validity and factorial invariance of the measure among osteoarticular disease, diabetes mellitus, obesity, depressive disorder, and general populations," *Clinical Psychology & Psychotherapy*, vol. 24, no. 5, pp. 1121–1129, 2017, <https://doi.org/10.1002/cpp.2077>.

Predicting Flux Rates against Pressure via Solution-Diffusion in Reverse Osmosis Membranes

Hisham A. Maddah

Department of Chemical Engineering
King Abdulaziz University
Rabigh, Saudi Arabia
hmaddah@kau.edu.sa

Abstract-This paper suggests a new method of predicting flux values at Reverse Osmosis (RO) desalination plants. The solution-diffusion model is utilized to determine the osmotic pressure drops for seawater sources. The same technique was applied to the groundwater source at the Abqaiq plant (500 RO plant) to calculate the osmotic pressure. The calculated osmotic pressures were utilized to determine the appropriate flux rates and membrane resistances of different BWRO Toray membranes and a performance comparison between various membranes has been established. The model results confirm an inverse relationship between membrane thickness and water flux rate. Also, a proportional linear relation between the overall water flux and the applied pressure is identified. Higher flux rates and lower salinity indicate lower membrane resistance yielding higher production. The modeled data predict that BWRO Toray TM720D-440 with an 8" membrane is the optimal choice for treating waters from the three water sources at the Abqaiq plant.

*Keywords-*reverse osmosis; treatment; desalination; modeling

I. INTRODUCTION

The solution-diffusion model is a popular expression used to explain the transport in dialysis, reverse osmosis, gas permeation, and pervaporation. Previous experimental data and modeling results verified that the flux rate is proportional to the gradient in the chemical potential [1]. There are two different models that describe and control the permeation in membranes for better separation. The first model is the solution-diffusion model where permeants dissolve (sorption) in the membrane material at the upstream interface in the presence of a concentration gradient that allows permeants to diffuse through the membrane and desorb on the downstream interface side. The separation between different permeants occurs because each material has a different diffusion rate in the membrane. The solution-diffusion model has been used since 1940 to explain the transport of gases across polymeric membranes. A second model, called the pore-flow model, depends on the presence of a pressure gradient for a convection flow of permeants through the membrane's tiny pores, and is more limited compared to the first model. Exclusion or filtration of larger permeant's pores is the separation technique explained via the pore-flow model [1, 2]. There is a major difference between the solution-diffusion model and the pore-flow model in expressing the chemical potential. In the solution-diffusion

model, the pressure within a membrane is uniform and the chemical potential gradient is expressed only as a concentration gradient. Solution-diffusion membranes transmit pressure in the same way as liquids, which is the reason for expressing the pressure difference across the membrane as a concentration gradient only. On the other hand, the chemical potential gradient in the pore-flow model is expressed only as a pressure gradient since the concentrations of both solvent and solute within a membrane are uniform. Comparisons between the two models for a one-component solution in a pressure-driven permeation system were conducted in [1, 2].

The objective of this work is to estimate the osmotic pressure drop value of the high rejection brackish water RO membrane (Toray TM720D-400 with 8") by using the solution-diffusion model that is applied to the Abqaiq plant (500 RO plant) for Shedgum/Abqaiq groundwater at Saudi Aramco, Dhahran, Saudi Arabia. Osmotic pressure drops have been calculated for the groundwater, the Arabian Gulf water, and the Red Sea water at the same plant configuration and operating conditions of the Abqaiq plant in Aramco. The calculated osmotic pressures are utilized to determine the applied pressure drop across the membrane and the applicability of using different BWRO Toray membrane types for the treatment of seawater. The maximum achievable water flux values are determined for the various suggested BWRO membranes for the three water sources. Also, the membrane resistance values have been investigated for comparison purposes. The ideal membrane for the treatment of various water sources at a RO plant with the same configuration of the Abqaiq plant has been selected. The feasibility of using BWRO membranes in the desalination of Red Sea water in Jeddah, Saudi Arabia is studied at the same flux rate of the Arabian Gulf water source and the same plant conditions. Osmotic pressure drop, applied pressure drop, flux rates, and membrane resistance values for the Red Sea water source were compared with those of Shedgum/Abqaiq groundwater and Arabian Gulf water.

II. REVERSE OSMOSIS

In reverse osmosis, water flows from the salt solution to the pure waterside by applying pressure (Δp) that is greater than the osmotic pressure ($\Delta \pi$) [1]. Generally, in reverse osmosis, the condition $\Delta p > \Delta \pi$ must be satisfied all the time to allow water to pass through the membrane and reach the permeate

Corresponding author: Hisham A. Maddah

side [1, 2]. Reverse osmosis membranes are preferred over ultrafiltration and nanofiltration since they are capable of removing 90 to 99% of TDS in water [3]. The osmotic pressure $\Delta\pi$ is defined as the pressure difference $p_o - p_f$ across the membrane. If a pressure higher than the osmotic pressure is applied to the feed side of the membrane [1]. This process is called the reverse osmosis. The driving forces in a reverse osmosis membrane according to the solution-diffusion and the pore-flow models are visualized in [1]. μ_i and γ_i are the chemical potential and activity coefficient, respectively, of component i [1].

III. METHOD AND EQUATIONS

Collected Abqaiq 500 RO plant data (Table I) have been used to determine osmotic pressure drop values for the RO membrane (Toray TM720D-400 with 8") from (1) and (2). However, in order to calculate the osmotic pressure for seawater sources, the same information of Shedgum/Abqaiq groundwater at Abqaiq 500 RO plant was applied, except for the flux and salinity values, for the treatment of either the Arabian Gulf or the Red Sea waters as listed in Table I [1, 4]. Water permeability is approximately determined to be $9.5 \times 10^{-7} \text{ cm}^2/\text{s}$ [8]. For water-salt solution, reverse osmosis permeation expression can be simplified as [1, 5]:

$$J_i = A(\Delta p - \Delta\pi) \quad (1)$$

$$A = \frac{P_i c_{io} v_i}{RT\ell} \quad (2)$$

where J_i is the membrane flux of component i , water (gfd), Δp is the applied pressure drop across the membrane (psi), $\Delta\pi$ is the osmotic pressure drop across the membrane (psi), A is the water permeability constant ($\text{cm}^2/\text{atm}\times\text{s}$), P_i is the permeability of component i , water (cm^2/s), c_{io} is the initial mole concentration of water (ppm), v_i is the water molar volume (cm^3/mol), T is the water temperature (K), R is the gas constant, and ℓ is the membrane thickness which is assumed to be similar to spacer thickness (mil).

Membrane resistance [8] constants for each BWRO Toray membrane have been calculated by using (3).

$$J_i = \frac{\Delta p}{\kappa \mathcal{R}_m} \quad (3)$$

where J_i is the membrane flux of component i , water (gfd), Δp is the applied pressure across the membrane (psi), κ is the dynamic viscosity of water ($\text{lb s}/\text{ft}^2$), and \mathcal{R}_m is the membrane resistance (1/ft).

$$\pi = \mathcal{M}RT \quad (4)$$

In (4), \mathcal{M} is the molar concentration of dissolved species (mol/L), R is the ideal gas constant ($0.08206 \text{ L}\cdot\text{atm}/\text{mol}\cdot\text{K}$), and T is the water temperature (K).

Equation (5) defines the ability of a membrane to separate salt from the feed solution which is known as membrane removal percentage (χ) and it increases with the applied pressure. The feed TDS concentration is taken from the three studied sources, as shown in Table I, while the outlet TDS concentration is determined by using (5) at a similar removal percentage of Toray TM720D-400 with 8" membrane that is 99.8% (Table IV). The water molecular weight (18g/mol) should be used to convert the ppm values to molar concentrations of TDS.

$$\chi = \left(\frac{c_{jo} - c_{je}}{c_{jo}} \right) \times 100 \quad (5)$$

where χ is the membrane removal percentage, c_{jo} is the initial concentration of component j , salt (ppm), and c_{je} is the final concentration of component j (ppm).

Table II shows the applied pressure drop must be at 20psi or below per element (RO module) and 60psi or below per vessel [4, 6]. The assumption of having equal pressure on membranes per vessel would simplify our calculations. Altaee's study showed that permeate flow, pressure and recovery rate are distributed almost equally to membranes per RO vessel [10]. A field study confirmed an improved performance by rearranging the elements in pressure vessels in order to reduce the pressure drop and permeate conductivity across the vessel [11]. Typical flux rates and maximum recovery values for the groundwater and the two studied water source scenarios (the Arabian Gulf and the Red Sea waters) at Abqaiq 500 RO plant are given in Table III.

TABLE I. DATA OF RO MEMBRANE PROCESS AT ABQAIQ 500 RO PLANT AND THE TWO SEAWATER STUDIED SCENARIOS [1, 4, 6, 7]

Parameter	Shedgum/Abqaiq groundwater	Arabian Gulf water	Red Sea water
Membrane type	Toray TM720D-400 with 8"		
RO module	72 parallel membranes \times 8 units		
Membrane thickness (ℓ) [3]	Assumed to be similar to spacer thickness of 34 mil		
Membrane area (Area) [3]	400ft ²		
Max pressure drop per vessel (Δp)	~60psi		
Max pressure drop per membrane (Δp)	~20psi		
Water salinity (c_{io})*	~2800 [4]	~41070 [6]	~42070 [7]
Membrane water flux (J_i)*	~18gfd	~12 gfd	~12 gfd
Water temperature (T)	25C		
Water permeability constant (P_i)**	$9.5 \times 10^{-7} \text{ cm}^2/\text{s}$		
Water molar volume (v_i)	18cm ³ /mol		
Gas constant (R)	$8.2057 \times 10^{-5} \text{ m}^3 \cdot \text{atm}/\text{mol}\cdot\text{K}$		

*Averaged values

** Taken from [2], regardless of the temperature effect on permeability. Can be calculated at different temperatures [16]

TABLE II. OPERATING DESIGN LIMITS OF THE OVERALL RO MODULE AT ABQAIQ 500 RO PLANT [4, 12, 13]

Operating limits	
Maximum operating pressure	600psi (4.1MPa)
Maximum feed water temperature	113°F (45°C)
Maximum feed water SDI15	5
Feed water chlorine concentration	Not detectable
Feed water pH range, continuous operation	2-11
Feed water pH range, chemical cleaning	1-12
Maximum pressure drop per element	20psi (0.14MPa)
Maximum pressure drop per vessel	60psi (0.4MPa)

TABLE III. CHARACTERISTICS OF GROUNDWATER SOURCE AND STUDIED WATER SOURCES AT THE ABQAIQ 500 RO PLANT [4]

Water source	Shedgum/Abqaiq groundwater	Arabian Gulf	Red Sea
Feed silt density index	SDI < 3	SDI < 4	SDI < 4
Typical target flux (gfd)	18	12	12
Max. element recovery (%)	19	14	14

The determined osmotic pressure values for the RO membrane (Toray TM720D-400 with 8") of the groundwater and the two studied water sources are used again in (1) to calculate the applied pressure drop and suggested flux values. The same osmotic pressure drop for each case is utilized to determine the results of different Toray BWRO membrane types at high, low, and standard operating pressure as shown in Table IV. TM-720-370 and TM720-440 are standard BWRO membranes and TM720C-440, TM720L-400, and TM720L-440 are low-pressure BWRO membranes whereas TM720DA400, TM720D-400, and TM720D-440 are high-pressure BWRO membranes. It is worth mentioning that our applied pressure drop must be higher than the calculated osmotic pressure in order to have a positive flux. TS-diagrams [7] are used to determine the exact value of water densities at different feed sources from the average water temperature and water salinity (Table V). The exact water densities allow us to convert gas constant values from $m^3 \cdot atm/mol \cdot K$ to $kg \cdot atm/mol \cdot K$ to progress calculations.

TABLE IV. TORAY BRACKISH WATER RO 8" DIAMETER MEMBRANES [13, 14]

Category	Type	Rejection (%)	Thickness (mil)*
Standard BWRO	TM720-370	99.7	31
	TM720-440	99.7	28
High-pressure BWRO	TM720DA400	99.8	31
	TM720D-400	99.8	34
	TM720D-440	99.8	28
Low-pressure BWRO	TM720C-440	99.2	28
	TM720L-400	99.5	31
	TM720L-440	99.5	28

* The membrane thickness is assumed to be the same as spacer thickness

TABLE V. WATER DENSITIES FROM TS-DIAGRAMS [4, 6, 7, 15]

Water Source	Temperature (°C)	Salinity (ppm)	Density (kg/m^3)
Shedgum/Abqaiq groundwater	25	2800 [2]	999.19
Arabian Gulf	25	41070 [4]	1027.97
Red Sea	25	42070 [6]	1028.67

IV. RESULTS AND DISCUSSION

Equations (1) and (2) allow us to calculate the osmotic pressure drop for each water source (Table VI). The osmotic pressure of the groundwater source is less than the Arabian Gulf and the Red Sea water sources which is related to the flux rates and water salinity. Flux rates for the Arabian Gulf and the Red Sea waters are approximately the half of the groundwater source. However, the water salinity of the groundwater source is much lower than the other sources. Therefore, the required applied pressure drop must be larger in the case of seawater sources due to their higher determined osmotic pressure values. Since the plant configuration has 8 elements per vessel, we should have a maximum osmotic pressure of 60psi or less per vessel which is equivalent to a max pressure of 7.5psi per membrane, assuming that the pressure is distributed equally on membranes per vessel. The selected applied pressure range for our study is 6.5 to 7.5psi. Maximum pressure values are assigned to the different membranes based on their category as illustrated in Table VII.

TABLE VI. CALCULATED OSMOTIC PRESSURE DROP ($\Delta\pi$) FOR EACH WATER SOURCE

Water source	A (cm/atm.s)	J_i (cm/s)	J_i/A (atm)	$\Delta\pi$ (atm)	$\Delta\pi$ (psi)	$\Delta\pi$ per vessel < 60 (psi)
Shedgum/Abqaiq groundwater	0.00808	0.00083	0.1028	0.44	6.48	51.84
Arabian Gulf	0.00755	0.00056	0.0742	0.47	6.90	55.21
Red Sea	0.00754	0.00056	0.0743	0.47	6.90	55.20

TABLE VII. ASSIGNED PRESSURE VALUES FOR TORAY BWRO MEMBRANES

Category	Type	Δp range (psi)*
Standard BWRO	TM720-370	6.50 - 7.25
	TM720-440	6.50 - 7.25
High-pressure BWRO	TM720DA400	6.50 - 7.50
	TM720D-400	6.50 - 7.50
	TM720D-440	6.50 - 7.50
Low-pressure BWRO	TM720C-440	6.50 - 7.00
	TM720L-400	6.50 - 7.00
	TM720L-440	6.50 - 7.00

* High and low-pressure values are taken relative to the standard pressure

The relationship between the applied pressure drops and the overall water flux rates for the groundwater source are obtained in Figure 1(a)-(c) for standard, high-pressure, and low-pressure Toray BWRO membranes. Figure 1 shows that the maximum possible flux for the groundwater in the standard membranes is around 11gfd for TM720-440 membrane. In Figures 3(b) and 3(c) the highest observed groundwater flux in the high-pressure and low-pressure membranes are 14.7gfd for TM720D-440 and 7.5gfd for TM720C-440 and TM720L-440 respectively (blue and green lines overlap). This observation is associated with the membrane thickness in which the least membrane thickness (28mils) has been capable to achieve the highest flux. This confirms an inverse relationship between the membrane thickness and the water flux rate. Further, there is a linear relationship between the applied pressure drop and the overall water flux.

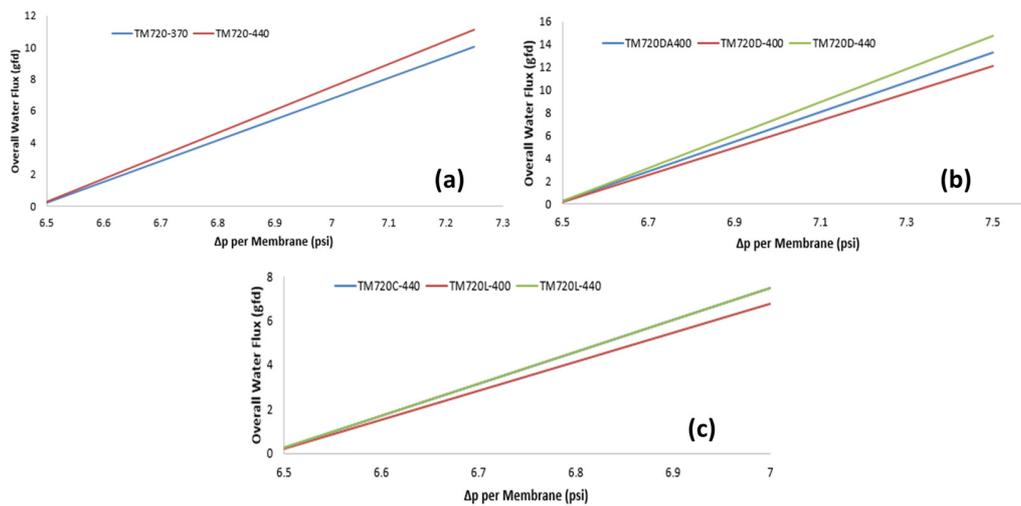


Fig. 1. Effect of different applied pressures on the groundwater flux for (a) Toray standard BWRO membranes, (b) Toray high-pressure BWRO membranes, and (c) Toray low-pressure BWRO membranes.

Figure 2 identifies a proportional relationship between the water flux and the applied pressure across the membrane. The highest recorded flux is accounted for TM720D-440 for Shedgum/Abqaiq groundwater because water TDS is low for groundwater and TM720D-440 has the lowest thickness and the highest pressure range. The Arabian Gulf and the Red Sea water sources almost have similar flux rates at the same applied pressures due to the similarities in their water salinity levels. TM720C-440, TM720L-400, and TM720L-440 membranes reserved the lowest flux values since they are categorized as low-pressure BWRO membranes.

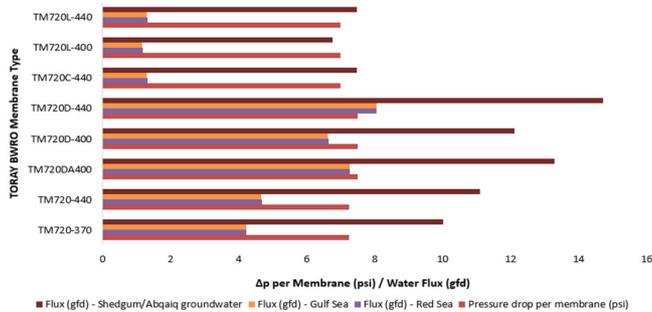


Fig. 2. Observed water flux for various water sources at different applied pressures using Toray BWRO membranes.

Figure 3 demonstrates the membrane resistance for the three studied water sources. Seawater sources have higher membrane resistances than the groundwater source because of their lower flux and higher TDS. TM720L-400 has the highest membrane resistance since it is in the low-pressure category and has the highest membrane thickness of 31mils. Equation (4) calculations are shown in Table VIII. The study predictions estimated that the overall osmotic pressure drops required for seawater and groundwater treatment plants are approximately 55psi and 830psi respectively. The higher the salinity difference between the fed and the produced water, the more the osmotic pressure drop we need to overcome in order to produce treated water (positive flux).

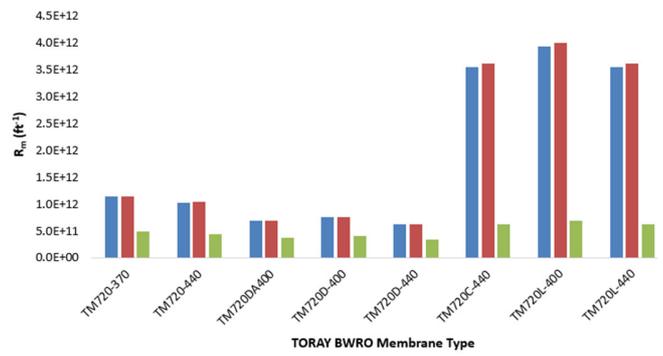


Fig. 3. Observed membrane resistance of various water sources in Toray BWRO membranes.

TABLE VIII. VAN'T HOFF CALCULATIONS FOR THE REQUIRED OSMOTIC PRESSURES

Water source	Concentration (mol/L)		Membrane removal (%)	Osmotic pressure (atm)			Δπ (psi)
	TDS _{in}	TDS _{out}		π _{in}	π _{out}	Δπ	
Shedgum/Abqaiq groundwater	0.156	0.00031	99.8	3.81	0.01	3.80	55.80
Arabian Gulf	2.282	0.005	99.8	55.82	0.11	55.71	818.41
Red Sea	2.337	0.005	99.8	57.18	0.11	57.07	838.34

V. CONCLUSION

The application of the solution-diffusion model to the Abqaiq plant (500 RO plant) is initiated by using various parameters to calculate the osmotic pressure of Toray TM720D-400 with 8" membrane for Shedgum/Abqaiq groundwater treatment. For the same membrane, the osmotic pressure values are determined for the Arabian Gulf and the Red Sea waters to predict flux rates in other membranes for seawater situations. Low, standard, and high pressure BWRO Toray membranes performances have been compared to identify the optimal membrane for treating saline water from the three studied water sources at the Abqaiq 500 RO plant.

The assumption of having a membrane thickness that is similar to its spacer thickness may not seem very accurate. However, it is true that we should have a proportional relation between both thicknesses which suggests that our results are still valid. A linear relationship has been observed between the water flux and the applied pressure drops. It is proved that membrane flux decreases with the increase in membrane thickness at constant pressure drop. Modeling results endorse that BWRO Toray TM720D-440 with 8"membrane is the optimum membrane choice for the water treatment from the three water sources at Abqaiq 500 RO plant since it has the lowest membrane resistance and the highest overall water flux.

REFERENCES

- [1] J. G. Wijmans and R. W. Baker, "The solution-diffusion model: a review," *Journal of Membrane Science*, vol. 107, no. 1, pp. 1–21, Nov. 1995, [https://doi.org/10.1016/0376-7388\(95\)00102-1](https://doi.org/10.1016/0376-7388(95)00102-1).
- [2] D. R. Paul, "Reformulation of the solution-diffusion theory of reverse osmosis," *Journal of Membrane Science*, vol. 241, no. 2, pp. 371–386, Oct. 2004, <https://doi.org/10.1016/j.memsci.2004.05.026>.
- [3] H. A. Maddah and A. M. Chogle, "Applicability of low pressure membranes for wastewater treatment with cost study analyses," *Membrane Water Treatment*, vol. 6, no. 6, pp. 477–488, Nov. 2015, <http://dx.doi.org/10.12989/mwt.2015.6.6.477>.
- [4] *Abqaiq plants: 500 RO plant, Operator Handbook*. Dhahran, Saudi Arabia: Saudi Aramco Company: Training & Career Development Center, 2008.
- [5] A. Sagle and B. Freeman, "Fundamentals of membranes for water treatment," *The future of desalination in Texas*, vol. 2, no. 363, 2004, Art. no. 137.
- [6] R. Smith, A. Purnama, and H. H. Al-Barwani, "Sensitivity of hypersaline Arabian Gulf to seawater desalination plants," *Applied Mathematical Modelling*, vol. 31, no. 10, pp. 2347–2354, Oct. 2007, <https://doi.org/10.1016/j.apm.2006.09.010>.
- [7] E. A. Abdel-Aal, M. E. Farid, F. S. M. Hassan, and A. E. Mohamed, "Desalination of Red Sea water using both electro dialysis and reverse osmosis as complementary methods," *Egyptian Journal of Petroleum*, vol. 24, no. 1, pp. 71–75, Mar. 2015, <https://doi.org/10.1016/j.ejpe.2015.02.007>.
- [8] J. C. Crittenden, R. R. Trussell, D. W. Hand, K. J. Howe, and G. Tchobanoglous, *MWH's Water Treatment: Principles and Design*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2012.
- [9] J. Prausnitz, R. Lichtenthaler, and E. G. de Azevedo, *Molecular Thermodynamics of Fluid-Phase Equilibria*, 3rd edition. Upper Saddle River, NJ, USA: Pearson, 1998.
- [10] A. Altaee, "A Computational Model to Estimate the Performance of 8 inches RO Membranes in Pressure Vessel," *Journal of Membrane and Separation Technology*, vol. 1, no. 1, pp. 60–71, Oct. 2012.
- [11] S. S. Madaeni, M. Afshar, N. Jaafarzadeh, F. Tarkian, and K. Ghasemipناه, "Rearrangement of membrane elements in the pressure vessels for optimum utilization of reverse osmosis process," *Chemical Engineering Research and Design*, vol. 89, no. 1, pp. 48–54, Jan. 2011, <https://doi.org/10.1016/j.cherd.2010.04.021>.
- [12] "TorayTM720D-400 with 8", in *ROMEMBRA® Toray RO Membrane*, Tokyo, Japan: Toray Corporation, 2011.
- [13] "Brackish Water & Sea Water Membranes, TM (inner connector type) Datasheets," in *ROMEMBRA® Toray RO Membrane*, Tokyo, Japan: Toray Corporation, 2016.
- [14] *Maintenance and handling manual for membrane elements (rev. 105, 106)*. Poway, CA, USA: Toray Membrane USA, 2016.
- [15] T. Beer, *Environmental Oceanography*, 2nd ed. Boca Raton, FL, USA: CRC Press, 1996.
- [16] H. A. Maddah, "Application of Finite Fourier Transform and Similarity Approach in a Binary System of the Diffusion of Water in a Polymer,"

Journal of Materials Science and Chemical Engineering, vol. 4, no. 4, pp. 20–30, Apr. 2016, <https://doi.org/10.4236/msce.2016.44003>.

AUTHORS PROFILE

Hisham Maddah is an independent researcher and a faculty member working with the Chemical Engineering Department at King Abdulaziz University in Rabigh (KAU-Rabigh). He has been awarded a full scholarship and received his PhD in "Naturally-Sensitized Photoanodes for Molecular Photovoltaics" from the University of Illinois at Chicago (UIC) by Summer 2020. He has completed his MS and BS in Chemical Engineering in the University of Southern California (USC) in 2017 and KAU in 2012 respectively. He is an expert in organic-inorganic dye-sensitized solar cells for solar energy harvesting. His research interests include mini-passive solar stills, machine learning for solar-desalination systems, membrane separation, capacitive deionization, and activated carbons for water treatment, separation technologies, and higher education administration. He has published more than 50 articles and 3 book chapters in reputed international peer-reviewed journals and conferences and co-arranged 5 outreach activities on third-generation solar cells.

Thai Water Buffalo Disease Analysis with the Application of Feature Selection Technique and Multi-Layer Perceptron Neural Network

Sumitra Nuanmeesri
Faculty of Science and Technology
Suan Sunandha Rajabhat University
Bangkok, Thailand
sumitra.nu@ssru.ac.th

Wongkot Sriurai
Faculty of Science
Ubon Ratchathani University
Ubon Ratchathani, Thailand
wongkot.s@ubu.ac.th

Abstract—This research aims to develop an analysis model for diseases of the water buffalo with the application of the feature selection technique along with the Multi-Layer Perceptron Neural Network (MLP-NN). The data used for analysis were collected from books and documents related to water buffalo diseases and the official website of the Department of Livestock Development. The data consist of the characteristics of 6 water buffalo diseases, including anthrax, hemorrhagic septicemia, brucellosis, foot and mouth diseases, parasitic diseases, and mastitis. Since the amount of the collected data was limited, the synthetic minority over-sampling technique was also employed to adjust the imbalance dataset. The adjusted dataset was used to select the disease characteristics towards the application of two feature selection techniques, correlation-based feature selection and information gain. Subsequently, the selected features were then used for developing the analysis model for water buffalo diseases towards the use of the MLP-NN. The evaluation results given by 10-fold cross-validation, showed that the analysis model for water buffalo diseases developed by correlation-based feature selection and MLP-NN provided the highest level of effectiveness with an accuracy of 99.71%, precision of 99.70%, and recall of 99.72%, implying that the analysis model is effectively applicable.

Keywords—water buffalo diseases; feature selection; multi-layer perceptron; neural network; synthetic minority over-sampling

I. INTRODUCTION

In Thailand, water buffaloes play an important role in the livestock economy. Thai water buffaloes are the most common farm animals in Asia and farmers typically use them for agricultural labor and as a source of food. To domesticate water buffaloes productively, it is vital to pay attention to their nutrition, habitats, sanitation, signs, behavior, and disease symptoms [1]. At present, water buffalo farmers are confronting many kinds of water buffalo diseases due to the seasonal changes, disease carriers, the lack of expertise and knowledge among farmers themselves, and the lack of experts who can analyze and diagnose these diseases. Moreover, the internet only offers basic information about the diseases gained from statistical surveys, resulting in the retrieval of incorrect or inadequate data for disease analysis and, ultimately, misunderstanding or incorrect analysis. Water buffalo farmers

who lack attention may overlook the signs and symptoms of serious infectious diseases, which could be spread to other animals and cause sickness or even casualties [2-3].

The Multi-Layer Perceptron Neural Network (MLP-NN) is one of the most popular techniques used to classify complex data. This research aims to develop an analysis model for water buffalo diseases towards the application of the feature selection technique and MLP-NN. Two feature selection techniques were employed, Correlation-based Feature Selection (CFS) and Information Gain (IG). After selecting the features, the data were then used for developing the analysis model by using the MLP-NN. The developed model can be applied to the development of a water buffalo disease analysis system, which is expected to help farmers timely analyze the diseases.

II. BACKGROUND

A. Data Imbalance Resolution

This research applies the Synthetic Minority Over-sampling Technique (SMOTE) to resolve the data imbalance problem. SMOTE helps to resynthesize data by increasing the dataset's size with a small amount of class data [4] to be compatible with the biggest dataset. This is done by randomizing a value and calculating the distance between the selected value with other values to find the nearest value [5].

B. Feature Selection Techniques

The goal of feature selection is to select the most significant features of each dataset in order to synthesize the model rapidly and increase the effectiveness of data classification. In this study, two feature selection methods are employed:

1) Correlation-based Feature Selection (CFS)

CFS is a feature selection method based on the relationship between the collections of features gained from the evaluation of feature prediction capacity used for data classification and irrelevant data management. CFS can rank the data subsets based on the data dimensions and select the data subsets based on the data dimensions with regard to high and low relationships between classes. Any irrelevant data or any data with a low level of relationship will be excluded. The same will

Corresponding author: Sumitra Nuanmeesri

occur with complex data dimensions which shall be excluded from the data dimensions with a high level of relationship. The formula for evaluating the subsets of CFS data dimensions is shown in (1) [6]:

$$M_s = \frac{k\bar{r}_{cf}}{\sqrt{k+k(k-1)\bar{r}_{ff}}} \quad (1)$$

where k refers to the data dimension or features, M_s refers to the value of S data dimension subset which composes of k data dimensions, \bar{r}_{cf} refers to the average value of the relationship between the variables and classes ($f \in S$), and \bar{r}_{ff} refers to the average value of the relationship of data dimension.

2) Information Gain (IG)

IG is a feature selection method in which the gain value of each node is evaluated. If a node has the highest gain value, it will be chosen as the root node, and the rest of the data will be reassessed in order to find the gain value of the next node. The formula for finding the IG value is [7]:

$$Gain(Y; X) = H(Y) - H(Y | X) \quad (2)$$

where Y refers to the feature value, which is a data class belonging to the $\{Y_1, Y_2, \dots, Y_n\}$ set where n is the number of features, X refers to the value of other features that are not classes ranging between $\{X_1, X_2, \dots, X_n\}$, $Gain(Y; X)$ refers to the score value gained from sample randomization ranging between 0 and 1, $H(Y)$ refers to the probability value gained from the randomization of Y samples, and $H(Y | X)$ refers to the probability value gained from the randomization of Y samples when compared to X . $H(Y)$ and $H(Y | X)$ are calculated in (3) and (4), respectively:

$$H(Y) = -\sum_{i=1}^{i=k} P(Y = y_i) \log_2 P(Y = y_i) \quad (3)$$

$$H(Y | X) = -\sum_{i=1}^{i=k} P(X = x_i) H(Y | X = x_i) \quad (4)$$

where $P(Y = y_i)$ refers to the probability value from y_1 to y_k , $P(X = x_i)$ refers to the probability value from x_1 to x_k , and k refers to the number of features.

C. Multi-Layer Perceptron Neural Network

The MLP-NN, illustrated in Figure 1, consists of an input layer, hidden layers, and an output layer [8]. In each layer, there are a collection of nodes. Possibly, there are more than one hidden layers [9]. The MLP-NN operates by inserting data into the input layer to estimate and deliver the results to the output layer. The estimation requires the sum total of input data multiplied by weight values, as shown in (5). After that, the output is used for the calculation with the sigmoid function, as shown in (6).

$$n = \sum_{i=1}^k P_i W_i \quad (5)$$

where n refers to the sum total of input P_i multiplied by the weight W_i and i refers to the number of inputs or weight value. In (6), x refers to the input value.

$$f_x = \frac{1}{1 + e^{-x}} \quad (6)$$

The output of the hidden layer is delivered to the output layer, where there is a comparison between the estimated and the target outputs. If there are different values that cannot be accepted, the outputs will get into the backpropagation process and go back to the hidden and input layer. Simultaneously, there is the weight adjustment process, which will find the most acceptable value after testing with the data. Subsequently, the output is estimated with the sigmoid function once again [8].

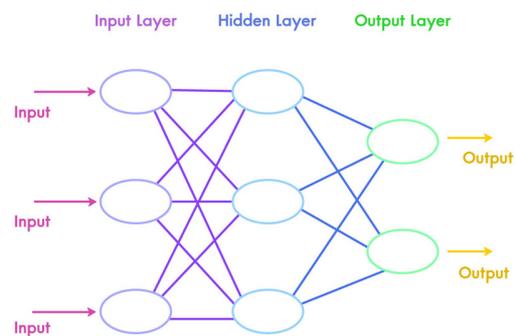


Fig. 1. MLP-NN.

D. Similar Studies

Authors in [13] compared the effectiveness of data imbalance resolution techniques by using diabetes patients' data. The research team compared four different methods, which include oversampling, undersampling, hybrid method, and SMOTE. Two data classification techniques were applied, Multinomial Logistic Regression Analysis and Decision Tree, to classify diabetes patients. The research findings showed that the combination of the data adjusted by SMOTE and the data classification with the decision tree technique provided the best results for the classification of the diabetes patients. Authors in [14] studied the classification of heart diseases using MLP-NN and IG as a feature selection method. The findings showed that the number of features could be reduced from 13 to 8, while the accuracy of the training dataset increased by 1.1% and the accuracy of the trial dataset increased by 0.82%. Authors in [15] studied the classification of ovarian cancer towards the application of SMOTE and MLP-NN. The findings indicated that the SMOTE technique could adjust the data balance, and after using the adjusted data to construct the model with MLP-NN, the model's effectiveness increased. The experiment results showed that the application of SMOTE+MLP provided a data classification accuracy of 96%, which was higher than the one gained from the application of SMOTE+RBF. Authors in [16] applied feature selection along with MLP to predict chronic diseases. The research findings showed that applying these two methods provided higher effectiveness in terms of chronic disease prediction than the application of Support Vector Machine (SVM) and Decision Tree.

The developed model of the current research can be applied to the development of a water buffalo disease analysis system. Two feature selection techniques were employed, CFS and IG.

E. Effectiveness Evaluation

The effectiveness evaluated by the confusion matrix is an evaluation method of discriminants' accuracy, which means the discriminants can be classified in accordance with their genuine value. The accuracy can be calculated by (7) [10], while precision and recall can be calculated in (8) and (9) [11-12], with the values represented in Figure 2.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \times 100 \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (9)$$

where TP refers to when the target class is "Yes," and the model predicts it as "Yes" (True Positive), FP refers to when the target class is "Yes," but the model predicts it as "No" (False Positive), TN refers to when the target class is "No," and the model predicts it as "No" (True Negative), and FN refers to when the target class is "No," but the model predicts it as "Yes" (False Negative).

		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

Fig. 2. Confusion matrix.

III. RESEARCH METHODOLOGY

The research methodology for the analysis of Thai water buffalo diseases consists of 1) data collection and preparation, 2) data imbalance adjustment using SMOTE, 3) feature selection by CFS and IG, 4) model development using MLP-NN, and 5) model effectiveness evaluation (Figure 3).

A. Data Collection and Preparation

This research collected data from books and documents related to water buffalo diseases [3] and the Department of Livestock Development's official website. There are totally 480 records of data. The data involves information about six water buffalo disease classes, namely anthrax, hemorrhagic septicemia, brucellosis, foot and mouth diseases, parasitic diseases, and mastitis. These data were used for developing the disease analysis model. There are 33 attributes and six classes, as illustrated in Table I.

TABLE I. DATA CHARACTERISTICS USED FOR DEVELOPING THE ANALYSIS MODEL OF THAI WATER BUFFALO DISEASES

Symptoms	Data symbol	Data value
Fever	Symptom 1	High: AA1 Moderate: AA2 No Fever: AA3
Drooling	Symptom 2	Normal: Ab1 Excessive: AB2
Breathing characteristics	Symptom 3	Fast and frequent breathing: AC1 Difficulty in breathing: AC2 Normal breathing: AC3
Bad breath	Symptom 4	Bad breath: AD1 Normal: AD2
Swollen breasts	Symptom 5	Little swollen: AE1 Red, hard, and swollen: AE2 Normal: AE3
Endometritis	Symptom 6	Inflammatory: AF1 Not inflammatory: AF2
Drowsiness	Symptom 7	Drowsy: AG1 Not drowsy: AG2
Eating habits	Symptom 8	Appetite loss: AH1 Normal: AH2
Swelling characteristics	Symptom 9	Body or neck swelling: AI1 Joint swelling: AI2 Not swollen: AI3
Urine characteristics	Symptom 10	Blood in urine: AJ1 Normal urine: AJ2
Testicle characteristics	Symptom 11	Swollen: AK 1 Normal: AK2
Stool characteristics	Symptom 12	Stools contained blood: AL1 Melena: AL2 Smelly stools: AL3 Normal: AL4
Walking	Symptom 13	Stiff: AM1 Normal: AM2
Milk characteristics	Symptom 14	Sticky or purulent: AN1 Normal: AN2
Retained placenta	Symptom 15	Retained: AO1 Not: AO2
Emaciation	Symptom 16	Emaciated: AP1 Not: AP2
Blisters inside the mouth and on the feet	Symptom 17	Has: AQ1 Has not: AQ2
Vesicular state (within mouth, tongue, beneath tongue, cheeks, hoofs, ankles, nipples, and breast)	Symptom 18	Has: AR1 Has not: AR2
Wounds around the hoofs and/or ankles	Symptom 19	Has: AS1 Has not: AS2
Weight	Symptom 20	Weight loss: AT1 Normal weight: AT2
Hard and rough hair	Symptom 21	Hard and rough: AU1 Normal: AU2
Sleep habits	Symptom 22	Excessive sleep: AV1 Normal: AV2
Muscle characteristics	Symptom 23	Shaky or twitching: AW1 Normal: AW1
Bleeding from nose, mouth, ears, eyes	Symptom 24	Has: AX1 Has not: AX2
Lip characteristics	Symptom 25	Pale: AY1 Normal: AY2
Oral mucositis and red eye sockets	Symptom 26	Has: AZ1 Has not: AZ2
Forward head posture	Symptom 27	Has: BA1 Has not: BA2
Tongue-tie	Symptom 28	Has: BB1 Has not: BB2
Cough	Symptom 29	Has: BC 1 Has not: BC2
Vaginal discharge	Symptom 30	Has: BD 1 Has not: BD2
Amount of milk	Symptom 31	Stopped milk production: BE1 Lower amount of milk: BE2 Normal: BE3
Excretion characteristics	Symptom 32	Frequent: BF1 Normal: BF2
Recurrent seizures	Symptom 33	Has: BG 1 Has not: BG2
Water buffalo diseases	Class	Anthrax Hemorrhagic septicemia Brucellosis Foot and mouth disease Parasitic disease Mastitis

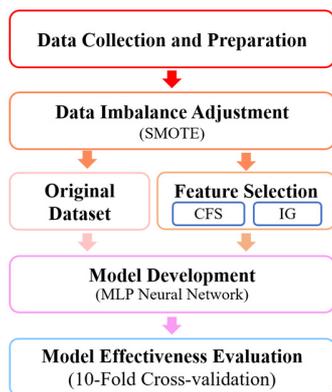


Fig. 3. The model framework.

After collecting the data illustrated in Table I, the research team rechecked their reliability and accuracy to ensure that no attribute was incorrect or missing. For example, there are no data out of range for each attribute. Then, the data were converted into .CSV file format in order to be operated with the Weka version 3.9, as shown in Figure 4.

#	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	
1	Symptom	class																		
2	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	antiradiase	
3	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	antiradiase	
4	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	antiradiase	
5	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	antiradiase	
6	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	antiradiase	
7	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	hemorrhagiceptemia	
8	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	hemorrhagiceptemia	
9	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	hemorrhagiceptemia	
10	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	hemorrhagiceptemia	
11	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	hemorrhagiceptemia	
12	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	brucellosis	
13	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	brucellosis	
14	AR2	AR2	AS2	AT2	AV2	AW2	AX2	AY2	AZ2	BA2	BB2	BC2	BD2	BE2	BF2	BG2	BH2	BI2	brucellosis	
15	AR1	AR1	AS1	AT1	AV1	AW1	AX1	AY1	AZ1	BA1	BB1	BC1	BD1	BE1	BF1	BG1	BH1	BI1	footandmouthdisease	
16	AR1	AR1	AS1	AT1	AV1	AW1	AX1	AY1	AZ1	BA1	BB1	BC1	BD1	BE1	BF1	BG1	BH1	BI1	footandmouthdisease	
17	AR1	AR1	AS1	AT1	AV1	AW1	AX1	AY1	AZ1	BA1	BB1	BC1	BD1	BE1	BF1	BG1	BH1	BI1	footandmouthdisease	
18	AR1	AR1	AS1	AT1	AV1	AW1	AX1	AY1	AZ1	BA1	BB1	BC1	BD1	BE1	BF1	BG1	BH1	BI1	footandmouthdisease	
19	AR1	AR1	AS1	AT1	AV1	AW1	AX1	AY1	AZ1	BA1	BB1	BC1	BD1	BE1	BF1	BG1	BH1	BI1	footandmouthdisease	

Fig. 4. Data used for model construction.

B. Data Imbalance Adjustment using SMOTE

Since the prepared data were found to be imbalanced in class/label, the research team decided to adjust the data imbalance of the datasets by increasing the number of datasets with a small size of classes. The best result was provided after increasing the *k*-nearest neighbor value from 1 to 5. It was experimentally found that *k* = 5 and *randomSeed* = 1 give the best result. The data size was then increased from 100% until the highest level of effectiveness could be gained (as evaluated by 10-fold cross-validation). The experiment results showed that the data could be balancing and upsizing to 300%. Thus, the new dataset increased to 528, 688, and 768 records for data balancing using SMOTE at 100%, 200%, and 300% of its original size respectively.

C. Feature Selection by CFS and IG

The data with 33 attributes and 6 classes were brought into the feature selection process using the CFS and the IG in Weka. In this work, there are 8 groups of data applied in this process, namely 1) the original dataset through the CFS, 2) the original dataset through IG, 3) 100% of SMOTE through CFS, 4) 200% of SMOTE through CFS, 5) 300% of SMOTE through CFS, 6) 100% of SMOTE through IG, 7) 200% of SMOTE through IG,

and 8) 300% of SMOTE through IG. These resulting datasets will be used at the next step.

D. Model Development Using MLP neural network

After the data imbalance had been adjusted, the data were transferred to the learning process to construct the model by which the research team applied two feature selection techniques, CFS and IG, along with the MLP-NN. In this research, the input layer consisted of 33 neurons. The output layer consisted of 6 neurons. Therefore, the optimum parameters for the MLP-NN model set in Weka were: Hidden Layer = 4, Training Time = 500, Learning Rate = 0.3, Momentum = 0.2, and 21 epochs= 500. These values provided the highest level of effectiveness after being evaluated by 10-fold cross-validation. Besides, the model was also generated from the original dataset that SMOTE, or CFS, or IG had not processed in any way, for effectiveness comparison with the other models that have undergone balancing and or feature selection. Thus, 9 models were built from the original dataset and 8 feature selection datasets.

IV. RESULTS

According to the experiment results, the most appropriate data size adjusted by SMOTE was 200%. Then, the features of the complete dataset were selected by CFS and IG. Afterwards, the MLP-NN was developed and its effectiveness was evaluated by 10-fold cross-validation, as illustrated in Table II and Figure 5.

TABLE II. MODEL'S EFFECTIVENESS EVALUATION

Method	Data (rows)	Precision (%)	Recall (%)	Accuracy (%)
MLP (Original data)	480	98.77	98.78	98.79
CFS+MLP	480	98.80	98.82	98.81
IG+MLP	480	91.59	91.54	91.52
SMOTE (100%) +CFS+MLP	528	98.85	98.87	98.86
SMOTE (200%) +CFS+MLP	688	99.70	99.72	99.71
SMOTE (300%) +CFS+MLP	768	99.60	99.61	99.62
SMOTE (100%) +IG+MLP	528	92.48	92.52	92.53
SMOTE (200%) +IG+MLP	688	93.62	93.59	92.61
SMOTE (300%) +IG+MLP	768	93.05	93.08	93.06

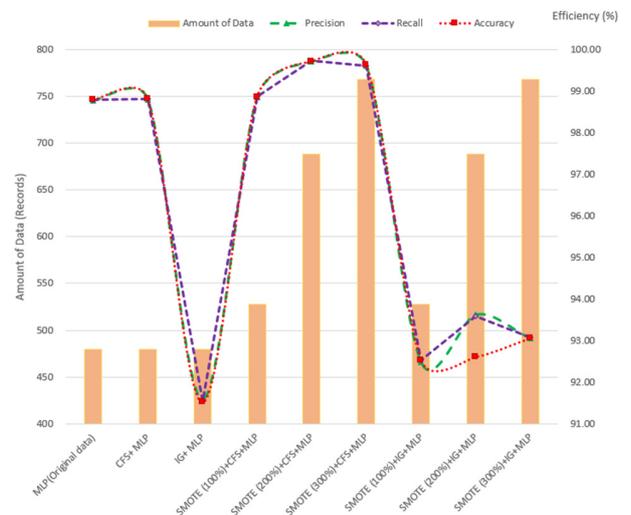


Fig. 5. The comparison of the model's effectiveness evaluation.

According to Table II and Figure 5, the 10-fold validation data classification method of the SMOTE (200%)+CFS+MLP dataset provided an accuracy of 99.71%, a precision of 99.70%, and a recall of 99.72%. These were the highest values. The results of the effectiveness comparison between the outputs of CFS and IG given by the MLP-NN showed that after being adjusted by SMOTE, the CFS method provided better feature selection than the IG method.

V. CONCLUSION

This research aimed to analyze the water buffalo diseases towards the application of feature selection techniques along with the MLP-NN. The data imbalance was adjusted by the SMOTE method. Two feature selection methods were employed: CFS and IG. After that, the data were classified by MLP-NN, and the model's effectiveness was evaluated by 10-fold cross-validation. The research findings showed that the most suitable data size after the data imbalance adjustment was 200%. After using the obtained data to construct the model, it was found that the model whose data size was adjusted by SMOTE and developed by CFS and MLP-NN provided the highest level of effectiveness in data classification with an accuracy of 99.71%. So, the developed model can be applied to the development of an analysis system for water buffalo diseases. The results of this study conform to the research conducted in [15], in which the SMOTE+MLP method was applied for data classification and a high level of effectiveness was reached with an accuracy of 90%, and also are in accordance with [16], in applied feature selection techniques were which alongside the MLP for data classification and gained a higher level of effectiveness.

ACKNOWLEDGMENT

The authors are grateful to the Institute for Research and Development, Suan Sunandha Rajabhat University and the Faculty of Science at Ubon Ratchathani University, for supporting this research.

REFERENCES

- [1] *Thailand's livestock statistics*. Bangkok, Thailand: Information and Communication Technology Center, Department of Livestock Development, 2015.
- [2] N. Hongboonmee and P. Sorrungrung, "Applying decision tree classification techniques for diagnose the disease in cow on mobile phone," *Journal of Science and Technology*, vol. 20, pp. 44–58, 2018.
- [3] P. Booranamanas, *Water buffaloes and treatments*, Bangkok, Thailand: Thaiwattanapanich, 1988.
- [4] M. G. Tsiouras, "Uterine EMG Signals Spectral Analysis for Pre-Term Birth Prediction," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3310–3315, Oct. 2018, <https://doi.org/10.48084/etasr.2146>.
- [5] P. Paranya, "Improving decision tree technique in imbalanced data sets using SMOTE for internet addiction disorder data," *Information Technology Journal*, vol. 12, no. 1, pp. 54–62, Jun. 2016.
- [6] A. H. Mark, "Correlation-based feature selection for machine learning," Ph.D. dissertation, The University of Waikato, Hamilton, New Zealand, 1999.
- [7] T. Puripat, "Ensemble algorithm for feature selection", M.S. thesis, Thammasat University, Bangkok, Thailand, 2016.
- [8] S. Boubaker, S. Kamel, and M. Kchaou, "Prediction of Daily Global Solar Radiation using Resilient-propagation Artificial Neural Network and Historical Data: A Case Study of Hail, Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5228–5232, Feb. 2020, <https://doi.org/10.48084/etasr.3278>.
- [9] A. Montaphan, "Comparison of feature selection methods to improve breast cancer prediction," *Royal Thai Air Force Medical Gazette*, vol. 65, no. 2, pp. 49–56, 2019.
- [10] M. B. Ayed, "Balanced Communication-Avoiding Support Vector Machine when Detecting Epilepsy based on EEG Signals," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6462–6468, Dec. 2020, <https://doi.org/10.48084/etasr.3878>.
- [11] S. Nuanmeesri, "Mobile application for the purpose of marketing, product distribution and location-based logistics for elderly farmers," *Applied Computing and Informatics*, 2019, <https://doi.org/10.1016/j.aci.2019.11.001>.
- [12] A. N. Saeed, "A Machine Learning based Approach for Segmenting Retinal Nerve Images using Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5986–5991, Aug. 2020, <https://doi.org/10.48084/etasr.3666>.
- [13] W. Kesornsit, V. Lorchirachoonkul, and J. Jitthavech, "Imbalanced data problem solving in classification of diabetes patients," *Khon Kaen University Research Journal*, vol. 18, no. 3, pp. 11–21, Jul. 2018.
- [14] A. Khempila and V. Boonjing, "Heart Disease Classification Using Neural Network and Feature Selection," presented at the Proceedings - ICSEng 2011: International Conference on Systems Engineering, Las Vegas, NV, USA, Sep. 2011, vol. 64, pp. 406–409, <https://doi.org/10.1109/ICSEng.2011.80>.
- [15] M. A. Hambali and M. D. Gbolagade, "Ovarian cancer classification using hybrid synthetic minority over-sampling technique and neural network," *Journal of Advances in Computer Research*, vol. 7, no. 4, pp. 109–124, Nov. 2016.
- [16] S. K. Hegde and R. Hedge, "Symmetry Based Feature Selection with Multi layer Perceptron for the prediction of Chronic Disease," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 3316–3322, Jul. 2019, <https://doi.org/10.35940/ijrte.B2658.078219>.

AUTHOR PROFILES



Sumitra Nuanmeesri received her Ph.D. in Information Technology at the King Mongkut's University of Technology North Bangkok, Thailand. She is an Assistant Professor in the Information Technology Department, Faculty of Science and Technology at Suan Sunandha Rajabhat University, Thailand. Her research interests include speech recognition, data mining, deep learning, image processing, mobile application, supply chain management system, internet of things, robotics, augmented reality, and virtual reality.



Wongkot Sriurai received her Ph.D. in Information Technology at the King Mongkut's University of Technology North Bangkok, Thailand. She is an Assistant Professor in the Mathematics, Statistics and Computer Department, Faculty of Science, Ubon Ratchathani University, Ubon Ratchathani Province, Thailand. Her research interests include data mining, text mining, web mining, recommender system, information filtering, information retrieval, decision support systems, expert systems, multimedia technology, and computer education.

Early Detection of Parkinson's and Alzheimer's Diseases Using the VOT_Mean Feature

Ahlem Kehili

Research Unit of Processing and Analysis
of Electrical and Energetic Systems
Faculty of Sciences
University of Tunis El Manar
Tunis, Tunisia
ahlemkehili@gmail.com

Karim Dabbabi

Research Unit of Processing and Analysis
of Electrical and Energetic Systems
Faculty of sciences
University of Tunis El Manar
Tunis, Tunisia
dabbabikarim@hotmail.com

Adnen Cherif

Research Unit of Processing and Analysis
of Electrical and Energetic Systems
Faculty of sciences
University of Tunis El Manar
Tunis, Tunisia
adenen.2fr@gmail.com

Abstract-Alzheimer's (AD) and Parkinson's diseases (PD) are two of the most common neurological diseases in the world. Several studies have been conducted on the identification of these diseases using speech and laryngeal disorders. Those symptoms can appear even at the early stages of AD and PD, but not in very specific and prominent ways. Voice Onset Time (VOT) is an acoustic specification of the stopping consonant that is commonly discussed in studies of phonetic perception. In this study, the VOT_Mean feature was explored to identify AD and PD early using /pa/, /ka/, and /ta/ syllables for the diadochokinetic task (DDK). VOT_Mean was calculated as the average of the first and the second VOT values (VOT_1 and VOT_2), corresponding to the second and the penultimate VOT measurement cycles. Experimental tests were performed on Tunisian Arabic and Spanish databases for the early detection of AD and PD respectively. The results showed a very high significance of VOT_Mean on the early detection of AD and PD. Moreover, the best results were achieved using the XGBoost (XGBT) algorithm as a classifier on the VOT_Mean feature.

Keywords-Alzheimer's disease (AD); Parkinson's disease (PD); early detection; neurological disorders; VOT_Mean; DDK; Tunisian Arabic database; Spanish database

I. INTRODUCTION

Associated speech disturbances caused by disturbances in the speech mechanism, particularly in muscle control, are grouped under the single definition called dysarthria. These disturbances result from disorders of the basic motor processes involved in the production of speech [1]. The difficulties in verbal production caused by dysarthria are due to weaknesses of muscle speech, lack of coordination, and paralysis. Moreover, dysarthria can cause peripheral and central nervous system damages. These lesions can appear from birth, as in the cases of cerebral palsy and muscular dystrophy, or later in life engendered by different conditions that can disrupt the nervous system, involving injuries, Parkinson's disease (PD), cerebral moods, multiple sclerosis, Huntington's disease, etc. The first studies related to the perceptual characteristics of dysarthria were carried in [2-3]. A total of 212 patients having symptoms of joint dysarthria with different neurological disorders, such as Parkinsonism, amyotrophic lateral sclerosis, chorea and

cerebellar ataxia, bulbar palsy, dystonia, and pseudobulbar palsy were examined, identifying 38 distinct characteristics of speech and categorizing them into seven modalities: articulation, respiration, pitch, prosody, loudness, resonance, and vocal quality. These studies specified the characteristics of each neurogenic group in addition to those shared by more than one. Once the perceptual characteristics of each neurogenic group were identified and grouped, the characteristics of dysarthric speech were classified into the following dysarthria types: hyperkinetic, ataxic, unilateral upper motor neuron, mixed and flaccid, and hypokinetic. Hypokinetic dysarthria is considered to be one type of dysarthria primarily associated with alterations in the functions of the control circuits of the basal ganglia. These disorders are the iterations that occur between motor and cognitive functions in the hippocampus. Hypokinetic dysarthria includes several aetiologies. Idiopathic PD is one such prototypical and common disease linked to this type of dysarthria. As the rate value of the estimated prevalence of PD in people aged over 65 years is 1.5%, it is considered to be the second most common neurodegenerative disorder [4], following Alzheimer's disease (AD) [5-6]. Moreover, cognitive impairment can be observed in 80% of people with PD [7]. The epidemiological study demonstrated in [8] showed that 80% of the patients with PD developed cognitive impairment, while 50% of them had PD dementia. The risk of developing dementia was higher in people with PD than in healthy ones [9]. Additionally, PD dementia contributes to the increased burden of care in addition to morbidity or mortality [10]. Thus, it is important to detect early Parkinson's Disease Dementia (PDD).

Further studies are needed to identify, evaluate, and distinguish PDD indices from other types of dementia, such as AD. Although the discriminating characteristics of PDD were manifested in damage on the executive function from the cognitive function test and problems related to the free recall in the memory test, PD without dementia may have diminished executive functions, such as worsening the verbal fluency [11]. However, problems with memory and executive functions can also be evidenced by AD and other types of cognitive impairments at the early-stage, complicating their distinction based on executive functions. Additionally, there may still be a

Corresponding author: Ahlem Kehili

significant overlap between the cognitive impairments of AD and PD patients, although the former primarily suffer from deterioration of the cortical profile (i.e. memory and ability language) and the latter have defects in the subcortical profile (i.e. executive functions and visuospatial ability) [12]. Moreover, similar cognitive disorders have been reported for AD and the 26% of the PDD cases [13]. Furthermore, 50% of PD patients without dementia exhibited amyloid accumulation [14], which is a widely known characteristic of AD. It should be noted that greater patient's age and later onset of PD increase the risk of dementia [15].

PD can lead to motor limitations affecting the speech production process causing several disorders, broadly categorized into three main dimensions: Linguistic, prearticulatory (resonance and phonatory characteristics), and prelinguistic (prosodic fluctuations and articulatory changes) [16]. In this context, sound production from the physical point of view is referred to prearticulatory, keeping in mind only the characteristics linked to the phonation process at the level of modulations integrated by the vocal tract (such as a resonant structure invariant in time) and phonation. Prelinguistic dimension refers to variations in the prosody of speech and is rendered as variations in f_0 and intensity, as well as articulatory variations appearing due to changes in the coordination frequency, the position of articulators, and a result of their interconnection with excitement. Thus, it mentions the mechanisms of speech production as forms varying in time. The linguistic dimension is linked to changes in the expected acoustic content of speech, mainly due to substitutions and/or repetitions. To sum up, the linguistic dimension is related to the expressive capacities of the language, while the prelinguistic and prearticulatory dimensions are related to the speech [16].

AD involves the whole process of the first pathological variations in the brain before its symptoms manifest as dementia [17]. AD includes patients with dementia, patients with Mild Cognitive Impairment (MCI), and asymptomatic persons having positive AD biomarkers [16]. Brain variations occur when specific neurons are deteriorated or destroyed. The progression of symptoms varies more quickly when the diagnosis of AD is made at an early stage [17]. Language issues are considered as one of the most specific symptoms of AD, which manifest as an unavoidable and direct outcome of cognitive impairment [18]. Primary Progressive Aphasia (PPA), defined as a clinical syndrome specified by a progressive language disorder with a neurodegenerative etiology, could be a manifestation mode of several neurodegenerative deficiencies such as AD [19]. Communication and language problems are evident in interactions between neurologists and patients, and interactional feedback can be explored to distinguish cognitive difficulties caused by neurodegenerative or functional memory impairments [20]. The use of different linguistic tests has shown poor performance for patients with AD [21], as they have more difficulty in naming tests where they often repeat the same ideas, use simpler forms of speech, and use empty or longer pauses. Moreover, people having AD show monotonous prosody and less cohesive, informative, and consistent speech compared to control groups [22].

Methods based on the automatic processing of the voice signal from its forms have been recently explored for the detection of neurodegenerative impairments. The traditional and common features of AD detection using voice signals are linear and considered to be the easiest to interpret clinically [23-28]. Other innovative and current techniques have explored non-linear features [29]. Both types of features are primary indicators of expressive forms of language based on the tasks performed [24]. Besides the feature extraction process, where mathematical models and/or statistical analysis are used to quantify the characteristics of speech, there is the classification process based on the exploration of machine learning techniques. Different types of classifiers have been constructed to perform the classification procedure, such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and K-Nearest Neighbour (KNN). Deep learning appears as a more sophisticated learning technique making possible the advance on automatic AD detection from the voice signal.

The DDK task is defined as a clinical test explored in the evaluation of the articulatory system's functional capacities [30]. Its concept is that motor deficits in vocal abilities of people suffering from PD appear more strongly in circumstances requiring motor execution and planning over long sequences of motor production (i.e. repeating the same sequence several times or in cases where a given fragment appears in a sentence context). Correcting PD results leads to simplifying the articulation at the cost of ease of processing [31], indicating important signs on speech production from a motor point of view. Fine articulatory precision is required for the DDK test, including the alternating production of syllable sequences, as well as the ability to quickly change articulators between consecutive segments. This is usually performed by asking the patient to produce combinations comprising of a voiceless consonant and vowels with velar, bilabial, and alveolar places of articulation. More precisely, the subjects are usually asked to constantly repeat the sequence of syllables /pa-/ta-/ka/ for about 10 seconds as quickly and clearly as possible. This test needs fast movements of the articulators using the soft palate (back of the mouth), the lips (the front), and the tip of the tongue (middle) continuously and sequentially. Regardless of its simplicity, this task reveals some signs of the speaker's ability to assess subphonemic durations or syllable-to-syllable stability and produce the speech at an appropriate rate [31]. The syllable rates of the DDK test are used in this sense to quickly alternate speech movements and assess the patient's ability [30]. The test was also used to evaluate imprecise consonantal coordination determining the VOT, typically measured as the duration between the initial burst and the vowel onset [32]. Moreover, in [33], it was proposed that the imprecisions of articulations on stop consonants were mainly due to low-frequency frictional noise or spirantization substituting stop gaps as a consequence of reduced closure. In this sense, the change in VOT for voiceless (/p/, /t/, /k/) and voiced (/b/, /d/, /g/) stops was identified as an indicator for the presence of PD [33-37]. Furthermore, VOT measurements were explored and compared for analyzing the speech production [p, t, k] of bilingual adults. VOT values were higher in Brazilian Portuguese (BP) than in English [38]. In [39], the differences between the VOT duration measures

were studied in the case of plosive consonants [p, t, k] for 14 bilingual subjects (Spanish-English) aged 18-24 years. Lower VOT values were found for subjects who learned both languages when they were less than 6 years old [40]. In [41], VOT measurements were studied in the case of non-voiced plosive sounds [p, t, k] for five bilingual children (English and BP) aged 8-9 years, reporting the following VOT values: for English, [p]=40ms, [k]=65ms, and [t]=56ms, while for BP [p]=48ms, [k]=70ms, and [t]=60ms. A few studies were conducted using VOT measurements in Turkish [42-44], while the VOTs in Lebanese Arabic were shorter compared to other languages [45].

This study adopted the VOT_Mean measure as an indicator to identify PD on a Tunisian Arabic dataset and AD on a Spanish dataset, using different machine learning algorithms and voiceless consonants (/pa/, /ta/, and /ka/) for the DDK task. The measurement of VOT_Mean was carried out by taking the average value of the second (VOT_1) and the penultimate (VOT_2) VOT measurement cycles.

II. METHODS AND DATASETS

A. Methods

The proposed method was composed of five phases: feature extraction, feature selection, feature scaling, data distribution, and classification.

1) Feature Extraction

The VOT measurements were carried out manually for /pa/, /ta/, and /ka/ syllables by utilizing the Praat software. As shown in Figure 1, the VOT_Mean was calculated by averaging the values of VOT_1 and VOT_2 which corresponded to the second and the penultimate VOT measurement cycles. The second cycle of the VOT measurement was selected due to the insufficient speech energy of the speaker, during the first VOT cycle. The penultimate VOT measurement cycle was selected due to its sufficient energy. The time cycle measurement was given in ms. The VOT_Mean values were consistent with the production of the stop consonant.

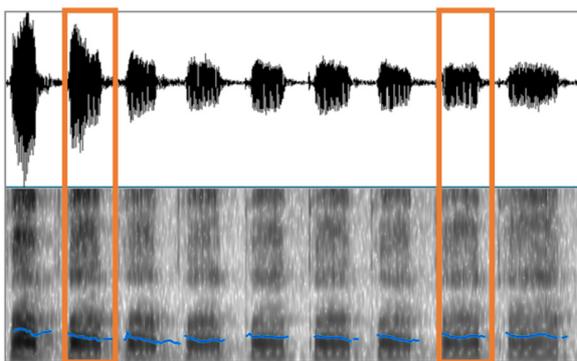


Fig. 1. The extraction of VOT_1 and VOT_2 using Praat software.

2) Feature Selection

In this phase, Principal Component Analysis (PCA) was used due to its high performance on feature extraction and feature dimension reduction. This approach was based on the

assumption that features with high variance included most of the information relating to certain classes. Furthermore, it permitted the conversion of observation groups with correlated variables into smaller groups with linear correlated variables by exploiting an orthogonal transformation.

3) Feature Scaling

This step was applied for the samples remaining after the feature selection. MinMax was used as a scaling function to scale the samples between -1 and 1. This phase's objective was to restrict the values of the concerned features in this range, avoiding large variations. Thus, the computational cost could be decreased ameliorating the model's performance.

4) Data Distribution

The explored database was split into training and test sets using the Python train_test_split function. As shown in Table I, 20% of the database was used for testing, while 80% was used for training. This distribution reduced computation time for the first cycles of modeling.

TABLE I. INFORMATION ABOUT THE EXPLORED DATABASE

Data Repartition	Test	Training	Number of samples
	20%	80%	1024

5) Classification

After the previous steps, the desired classifiers were selected and applied. Many Machine Learning (ML) algorithms were explored, such as Decision Tree (DT), KNN, XGBoost (XGBT), Random Forest (RF), and Multi-Layer Perceptron (MLP).

B. Datasets

1) Tunisian Arabic Dataset for Alzheimer's Disease

The dataset included numerous vocal recordings of male and female individuals of the AD group, where 20 suffered from AD, and 20 were healthy (HC). Each group was equally divided between the two sexes. Each speaker was requested to pronounce the sequence of syllables /pa/, /ka/, /ta/, or /pataka/ in the Tunisian Arabic dialect with a comfortable height and intensity. The goal behind the repetition of each syllable was the preservation of high reliability and quality in speech production for each participant. The number of repetitions was linked to the ability of each participant to repeat the requested sequence, so the duration of the recordings differ. All samples were recorded at 16-bit resolution, 22KHz sample rate, and saved in WAV format. The extracted features were saved in CSV format. Healthy participants' samples were labeled with "0", whereas participants suffering from AD were marked with "1". The mean age of the participants was 76 years, while its standard deviation was 8.17 years. The youngest and the oldest participants were 61 and 88 years old, respectively. The mean age of the healthy participants was 74.125, while its standard deviation was 5.79.

2) Spanish Dataset for Parkinson's Disease

This dataset was composed of voice recordings from 50 people suffering from PD and 50 HC. Each group consisted of 25 men and 25 women. All participants were native Colombian

Spanish speakers. Women aged between 44-75 years (mean 60.1±7.8), whereas men aged between 33-77 years (mean 62.2±11.2). In HC, women aged between 43-76 years (mean 60.7±7.7), while men aged between 31-86 years (mean 61.2±11.3). It should be noted that the participants suffering from PD did not suffer from any other neurological disease. Additionally, all participants were diagnosed and labeled by expert neurologists using the UPDRS and H&Y scales [46]. The samples were recorded at 16-bit resolution and 44.1KHz sample rate using a dynamic omnidirectional microphone.

3) Data Analysis

The SPSS software was utilized for performing Multivariate Analysis of Variance (MANOVA) on the repeated tests of VOT_Mean data. As shown in Table II, the VOT_Mean demonstrated a very high significant effect on the AD and PD groups ([F(1, 63)=44.773; p<0.000] and [F(1.296)=94.346; p<0.000], respectively), proving that the AD and HC groups' VOT_Means differ significantly. The sum and the average of squares of VOT_Means for inter and intra-groups are also presented in Table II.

TABLE II. ONE WAY ANOVA FOR VOT_MEANS

VOT Mean for AD					
	Sum of squares	ddl	Average of squares	F	Significance
Inter-groups	2877.19	1	2877.1	44.77	0.000
Intra-groups	3984.24	62	64.2		
Total	6861.43	63			
VOT Mean for PD					
Inter-groups	0.006	1	0.006	94.34	0.000
Intra-groups	0.019	295	0.000		
Total	0.25	296			

4) Evaluation Metrics

Different evaluation metrics were explored for verifying the predictability of the different models. The accuracy and the F-measure offered the opportunity to measure the capacity of differentiation of one class among others, even if they were unbalanced. Nevertheless, accuracy in such a case can lead to misleading results. The precision, sensitivity, and measurement metrics F were expressed as:

$$Precision = \frac{tp}{tp+fp} \quad (1)$$

$$Sensitivity = \frac{tp}{tp+fn} \quad (2)$$

$$F - Measure = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

where *tp*, *fp*, and *fn* represent true positive, false positive, and false negative, respectively. Precision and specificity were determined as:

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (4)$$

$$Specificity = \frac{tn}{tn+fp} \quad (5)$$

where *tn* represents true negative. Moreover, the Matthews Correlation Coefficient (MCC) was used to prove the quality of the binary classification. This metric was expressed as:

$$MCC = \frac{tp \cdot tn - fp \cdot fn}{\sqrt{(fn+tp)(fp+tn)(fp+tp)(fn+tn)}} \quad (6)$$

III. RESULTS AND DISCUSSION

A. Results

Table III shows the results of using different classifiers on the Tunisian Arabic dataset to detect AD. The best performance in all terms was achieved by XGBT, followed by RF. The XGBT algorithm had 92% precision, 92% accuracy, 92% sensitivity, 92% F1-measure, and 0.84 MCC. The RF algorithm had 90% precision, 90% sensitivity, 90% accuracy, 92% F1-measure, and 0.80 MCC.

TABLE III. RESULTS OF DIFFERENT CLASSIFIERS ON AD

		PRES (%)	SENS (%)	F1 (%)	ACC (%)	MCC
MLP	HC	83	86	88	87	0.74
	AD	90	88	86		
	Avg	87	87	87		
KNN	HC	86	85	88	88.3	0.77
	AD	92	92	89		
	Avg	89	88	88		
DT	HC	89	80	84	88	0.75
	AD	88	95	91		
	Avg	88	88	88		
RF	HC	89	89	89	90	0.80
	AD	90	90	90		
	Avg	90	90	90		
XGBT	HC	91	91	91	92	0.84
	AD	93	93	93		
	Avg	92	92	92		

Table IV shows the results of using different classifiers on the Spanish dataset for detecting PD. The best results were achieved by using the XGBT algorithm on the /ka/ syllable. The XGBT algorithm achieved the best average for precision (92%), followed by RF (88%). Regarding accuracy, the best average value was reached by XGBT (92%), followed by the RF (90%). The best sensitivity was achieved by XGBT (92%), followed by RF (90%). Concerning the F1, the best average values were achieved by XGBT (92%), followed by RF (87%), while the best MCC was achieved by XGBT (0.88), followed by RF (0.73). Overall, it can be concluded that XGBT was the best classifier on both AD and PD, in terms of results.

B. Discussion

This study aimed to examine whether subtle early symptoms of AD and PD exist in the acoustic signal of Arabic and Spanish speech respectively, by comparing the VOT_Mean feature between HC, AD, and PD groups. In [47], it was noted that the overall variation in VOT production could be revealed in individuals with moderate AD. Also, relying on results from various measurements on older people, there was an expectation that people at this stage of AD can show an overall change around the smaller VOT values. According to this study's results, the differences in VOT_Mean values were statistically very significant in people with AD or PD compared to a healthy control group, on both Arabic and Spanish databases. To the best of our knowledge, no published studies have tested VOT_Mean measurements in people with AD in Arabic languages. Nevertheless, some studies have been performed on Arabic databases using the VOT measure for

other tasks, such as testing the impact of contrast on F0 (CF0) caused by consonants in Lebanese Arabic [48]. Testing VOT measures on English [48] showed that there was no statistical significance in the differences of people with AD for the early detection of this disease.

TABLE IV. RESULTS OF DIFFERENT CLASSIFIERS ON PD

VOT Mean /ka/						
		PRE (%)	ACC (%)	SENS (%)	F1 (%)	MCC
MLP	HC	86	80	93	85	0.58
	PD	78		60	71	
	Avg	82		77	78	
KNN	HC	89	80	91	81	0.62
	PD	74		70	78	
	Avg	82		80.5	79.5	
DT	HC	85	84	90	89	0.59
	PD	80		61	87	
	Avg	82		76	88	
RF	HC	88	88	93	91	0.73
	PD	88		79	82	
	Avg	88		86	87	
XGBT	HC	94	92	94	94	0.88
	PD	90		89	89	
	Avg	92		92	92	
VOT Mean /pa/						
MLP	HC	86	80	90	84	0.56
	PD	76		63	71	
	Avg	81		67	77	
KNN	HC	89	88	90	91	0.68
	PD	83		75	77	
	Avg	86		83	84	
DT	HC	82	82	89	86	0.6
	PD	80		68	73	
	Avg	81		87	97	
RF	HC	88	88	94	91	0.71
	PD	86		75	80	
	Avg	87		84	85	
XGBT	HC	92	91	92	93	0.84
	PD	90		87	89	
	Avg	91		90	91	
VOT Mean /ta/						
MLP	HC	77	75	78	80	0.4
	PD	67		71	35	
	Avg	74		74	74	
KNN	HC	77	79	83	80	0.59
	PD	82		75	78	
	Avg	79		79	79	
DT	HC	80	81	89	84	0.61
	PD	83		71	77	
	Avg	82		80	81	
RF	HC	82	83	90	87	0.63
	PD	86		70	75	
	Avg	84		80	81	
XGBT	HC	90	88	90	90	0.73
	PD	83		83	83	
	Avg	87		87	87	

For early detecting PD in the Spanish language, it was found that utilizing the VOT Mean measure produced better results using the /ka/ syllable than /ta/ or /pa/. This is in agreement with [49], who proved that the proposed VOT was accurate for estimating its limits with the /ka/ syllable for healthy and PD-affected individuals. Moreover, the highest discrimination ability (94.4% with the leave-one-out method and 92.2% with the 10-fold cross-validation) was achieved

using the PD detection approach proposed in that study with /ka/ syllable. Furthermore, good results were obtained on the early detection of AD and PD using the VOT_Mean measurement, proving that this approach is effective regardless of the languages explored.

IV. CONCLUSION

This paper studied the early detection of Alzheimer’s and Parkinson’s diseases using the VOT_Mean feature of the DDK task. The early detection of AD test was performed on a Tunisian Arabic dataset using /pa/, /ta/, and /ka/ syllables, and the best results were obtained using the XGBT algorithm (precision=92%, accuracy=92%, sensitivity=92%, F1=92%, MCC=0.84). The early detection of PD test was carried out on a Spanish dataset using separately the /pa/, /ta/ and /ka/ syllables, and the best average results (precision=92%, accuracy=92%, sensitivity=92%, F1=92, MCC=0.88) were reached with XGBT on the /ka/ syllable. Moreover, the very high significance of the VOT_Mean feature was noted concerning early detection of AD and PD. Future work should test the VOT_Mean feature on the early detection of PD on Arabic databases, as there is no such database to this date. Furthermore, the proposed model for PD detection should be tested on other syllables, such as /pata/, /pakata/, etc. Regarding the early detection of AD, it is important to test the suggested approach on different language datasets.

ACKNOWLEDGMENT

The authors would like to express their cordial thanks to Pr. Adnen Cherif for his valuable advice. Also, they want to address their warm salutations to Dr. Afef Hammami, director of the AFA center, and all its members for their cooperation in creating the AD dataset from Arabic voices.

REFERENCES

- [1] A. E. Aronson and J. R. Brown, *Motor Speech Disorders*, 1st edition. Philadelphia: Saunders, 1975.
- [2] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential Diagnostic Patterns of Dysarthria," *Journal of Speech and Hearing Research*, vol. 12, no. 2, pp. 246–269, Jun. 1969, <https://doi.org/10.1044/jshr.1202.246>.
- [3] F. L. Darley, A. E. Aronson, and J. R. Brown, "Clusters of Deviant Speech Dimensions in the Dysarthrias," *Journal of Speech and Hearing Research*, vol. 12, no. 3, pp. 462–496, Sep. 1969, <https://doi.org/10.1044/jshr.1203.462>.
- [4] M.C. de Rijk *et al.*, "Prevalence of Parkinson’s disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group," *Neurology*, vol. 54, no. 11 Suppl 5, pp. S21-3, Jan. 2000.
- [5] L. M. de Lau and M. M. Breteler, "Epidemiology of Parkinson’s disease," *The Lancet Neurology*, vol. 5, no. 6, pp. 525–535, Jun. 2006, [https://doi.org/10.1016/S1474-4422\(06\)70471-9](https://doi.org/10.1016/S1474-4422(06)70471-9).
- [6] I. McKeith, "Dementia with Lewy bodies," in *Handbook of Clinical Neurology*, vol. 84, W. C. Koller and E. Melamed, Eds. Elsevier, 2007, pp. 531–548.
- [7] D. Aarsland, K. Andersen, J. P. Larsen, and A. Lolk, "Prevalence and Characteristics of Dementia in Parkinson Disease: An 8-Year Prospective Study," *Archives of Neurology*, vol. 60, no. 3, pp. 387–392, Mar. 2003, <https://doi.org/10.1001/archneur.60.3.387>.
- [8] M. A. Hely, W. G. J. Reid, M. A. Adena, G. M. Halliday, and J. G. L. Morris, "The Sydney multicenter study of Parkinson’s disease: The inevitability of dementia at 20 years," *Movement Disorders*, vol. 23, no. 6, pp. 837–844, 2008, <https://doi.org/10.1002/mds.21956>.

- [9] M. Emre, "Dementia associated with Parkinson's disease," *The Lancet Neurology*, vol. 2, no. 4, pp. 229–237, Apr. 2003, [https://doi.org/10.1016/S1474-4422\(03\)00351-X](https://doi.org/10.1016/S1474-4422(03)00351-X).
- [10] M. Nussbaum, T. A. Treves, R. Inzelberg, J. M. Rabey, and A. D. Korczyn, "Survival in Parkinson's disease: the effect of dementia," *Parkinsonism & Related Disorders*, vol. 4, no. 4, pp. 179–181, Dec. 1998, [https://doi.org/10.1016/S1353-8020\(98\)00039-X](https://doi.org/10.1016/S1353-8020(98)00039-X).
- [11] P. Péran *et al.*, "Deficit of verb generation in nondemented patients with Parkinson's disease," *Movement Disorders*, vol. 18, no. 2, pp. 150–156, 2003, <https://doi.org/10.1002/mds.10306>.
- [12] L. Bherer, S. Belleville, and C. Hudon, "Executive function deficits in normal aging, Alzheimer's disease, and frontotemporal dementia," *Psychologie & Neuropsychiatrie du Vieillessement*, vol. 2, no. 3, pp. 181–189, Sep. 2004.
- [13] I. U. Song, J. S. Kim, J. Y. Yoo, H. J. Song, and K. S. Lee, "Cognitive Dysfunctions in Mild Parkinson's Disease Dementia: Comparison with Patients Having Mild Alzheimer's Disease and Normal Controls," *European Neurology*, vol. 59, no. 1–2, pp. 49–54, 2008, <https://doi.org/10.1159/000109261>.
- [14] D. M. A. Mann and D. Jones, "Deposition of amyloid (A4) protein within the brains of persons with dementing disorders other than Alzheimer's disease and Down's syndrome," *Neuroscience Letters*, vol. 109, no. 1, pp. 68–75, Feb. 1990, [https://doi.org/10.1016/0304-3940\(90\)90539-L](https://doi.org/10.1016/0304-3940(90)90539-L).
- [15] R. Inzelberg, D. Paleacu, J. Chapman, and A. D. Korczyn, "Apolipoprotein E and Parkinson's disease," *Annals of Neurology*, vol. 44, no. 2, pp. 294–294, 1998, <https://doi.org/10.1002/ana.410440230>.
- [16] J. I. Godino-Llorente, S. Shattuck-Hufnagel, J. Y. Choi, L. Morovelázquez, and J. A. Gómez-García, "Towards the identification of Idiopathic Parkinson's Disease from the speech. New articulatory kinetic biomarkers," *PLOS ONE*, vol. 12, no. 12, p. e0189583, 2017, <https://doi.org/10.1371/journal.pone.0189583>.
- [17] "2017 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 13, no. 4, pp. 325–373, 2017, <https://doi.org/10.1016/j.jalz.2017.02.001>.
- [18] L. M. Gonnerman, J. M. Aronoff, A. Almor, D. Kempler, and E. S. Andersen, "From Beetle to Bug: Progression of Error Types in Naming in Alzheimer's Disease," *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 26, no. 26, 2004.
- [19] J. A. Matías-Guiu and R. García-Ramos, "Primary progressive aphasia: From syndrome to disease," *Neurología (English Edition)*, vol. 28, no. 6, pp. 366–374, Jul. 2013, <https://doi.org/10.1016/j.nrleng.2012.04.018>.
- [20] B. Mirheidari *et al.*, "Toward the Automation of Diagnostic Conversation Analysis in Patients with Memory Complaints," *Journal of Alzheimer's Disease*, vol. 58, no. 2, pp. 373–387, Jan. 2017, <https://doi.org/10.3233/JAD-160507>.
- [21] S. Weintraub, A. H. Wicklund, and D. P. Salmon, "The Neuropsychological Profile of Alzheimer Disease," *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 4, , Apr. 2012, Art. no. a006171, <https://doi.org/10.1101/cshperspect.a006171>.
- [22] V. Nieves López, "Evaluación del discurso de las personas con Enfermedad de Alzheimer: una revisión," B.S. Thesis, Universidad de Castilla-La Mancha, 2016.
- [23] J. B. Alonso, J. Cabrera, M. Medina, and C. M. Travieso, "New approach in quantification of emotional intensity from the speech signal: emotional temperature," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9554–9564, Dec. 2015, <https://doi.org/10.1016/j.eswa.2015.07.062>.
- [24] L. Brabenec, J. Mekyska, Z. Galaz, and I. Rektorova, "Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation," *Journal of Neural Transmission*, vol. 124, no. 3, pp. 303–334, Mar. 2017, <https://doi.org/10.1007/s00702-017-1676-0>.
- [25] P. Gómez-Vilda *et al.*, "Biomechanical characterization of phonation in Alzheimer's Disease," in *3rd IEEE International Work-Conference on Bioinspired Intelligence*, Jul. 2014, pp. 14–20, <https://doi.org/10.1109/IWOBI.2014.6913931>.
- [26] K. López-de-Ipiña *et al.*, "New Approaches for Alzheimer's Disease Diagnosis Based on Automatic Spontaneous Speech Analysis and Emotional Temperature," in *Ambient Assisted Living and Home Care*, Berlin, Heidelberg, 2012, pp. 407–414, https://doi.org/10.1007/978-3-642-35395-6_55.
- [27] K. López-de-Ipiña *et al.*, "Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach," *Computer Speech & Language*, vol. 30, no. 1, pp. 43–60, Mar. 2015, <https://doi.org/10.1016/j.csl.2014.08.002>.
- [28] H. Al-Dossari, F. A. Nughaymish, Z. Al-Qahtani, M. Alkahlifah, and A. Alqahtani, "A Machine Learning Approach to Career Path Choice for Information Technology Graduates," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6589–6596, Dec. 2020, <https://doi.org/10.48084/etasr.3821>.
- [29] J. Mekyska *et al.*, "Robust and complex approach of pathological speech signal analysis," *Neurocomputing*, vol. 167, pp. 94–111, Nov. 2015, <https://doi.org/10.1016/j.neucom.2015.02.085>.
- [30] S. G. Fletcher, "Time-by-Count Measurement of Diadochokinetic Syllable Rate," *Journal of Speech and Hearing Research*, vol. 15, no. 4, pp. 763–770, Dec. 1972, <https://doi.org/10.1044/jshr.1504.763>.
- [31] J. Kegl, H. Cohen, and H. Poizner, "Articulatory Consequences of Parkinson's Disease: Perspectives from Two Modalities," *Brain and Cognition*, vol. 40, no. 2, pp. 355–386, Jul. 1999, <https://doi.org/10.1006/breg.1998.1086>.
- [32] R. D. Kent, G. Weismer, J. F. Kent, H. K. Vorperian, and J. R. Duffy, "Acoustic studies of dysarthric speech: Methods, progress, and potential," *Journal of Communication Disorders*, vol. 32, no. 3, pp. 141–186, Jan. 1999, [https://doi.org/10.1016/S0021-9924\(99\)00004-0](https://doi.org/10.1016/S0021-9924(99)00004-0).
- [33] K. Chenausky, J. MacAuslan, and R. Goldhor, "Acoustic Analysis of PD Speech," *Parkinson's Disease*, Oct. 2011, Art. no. 435232, <https://doi.org/10.4061/2011/435232>.
- [34] M. Asgari and I. Shafran, "Predicting severity of Parkinson's disease from speech," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Aug. 2010, pp. 5201–5204, <https://doi.org/10.1109/IEMBS.2010.5626104>.
- [35] G. L. Dorze, L. Ouellet, and J. Ryalls, "Intonation and speech rate in dysarthric speech," *Journal of Communication Disorders*, vol. 27, no. 1, pp. 1–18, May 1994, [https://doi.org/10.1016/0021-9924\(94\)90007-8](https://doi.org/10.1016/0021-9924(94)90007-8).
- [36] G. Weismer and J. Wildermuth, "Formant trajectory characteristics in persons with Parkinson, cerebellar, and upper motor neuron disease," *The Journal of the Acoustical Society of America*, vol. 103, no. 5, pp. 2892–2892, May 1998, <https://doi.org/10.1121/1.421814>.
- [37] A. A. Alasadi, T. H. Aldhayni, R. R. Deshmukh, A. H. Alahmadi, and A. S. Alshebami, "Efficient Feature Extraction Algorithms to Develop an Arabic Speech Recognition System," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5547–5553, Apr. 2020, <https://doi.org/10.48084/etasr.3465>.
- [38] P. D. A. Rocca, "O desempenho de falantes bilingües: evidências advindas da investigação do VOT de oclusivas surdas do inglês e do português," *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, vol. 19, no. 2, pp. 303–328, 2003, <https://doi.org/10.1590/S0102-44502003000200004>.
- [39] P. Piccinini and A. Arvaniti, "Voice onset time in Spanish–English spontaneous code-switching," *Journal of Phonetics*, vol. 52, pp. 121–137, Sep. 2015, <https://doi.org/10.1016/j.jwocn.2015.07.004>.
- [40] M. T. R. Lofredo-Bonatto, M. A. A. e Silva, M. T. R. Lofredo-Bonatto, and M. A. A. e Silva, "Comparison of plosive sounds in monolingual and bilingual children, using the voice onset time acoustic parameter: cases report," *Revista CEFAC*, vol. 20, no. 5, pp. 680–687, Oct. 2018, <https://doi.org/10.1590/1982-021620182052118>.
- [41] M. C. Zimmer and M. H. T. Bandeira, "A dinâmica do multilingüismo na transferência de padrões de aspiração de obstruintes iniciais entre o pomerano (L1), o português (L2) e o inglês (L3)," presented at the X Congresso Nacional de Fonética e Fonoologia, Niterói, Brazil, Nov. 2008.
- [42] İ. T. Batuk, "Türkçe Konuşan Okul Çağı Çocuklarında Ses Çıkış Zamanlarının Değerlendirilmesi," M.S. thesis, University of Hacettepe, Ankara, Turkey, 2015.
- [43] A. Dalgic, T. Kandogan, and G. Aksoy, "Voice Onset Time for Turkish Stop Consonants in Adult Cochlear Implanted Patients," *Indian Journal*

- of *Otolaryngology and Head & Neck Surgery*, vol. 67, no. 3, pp. 308–313, Sep. 2015, <https://doi.org/10.1007/s12070-014-0784-5>.
- [44] H. Kopkalli-Yavuz, I. Mavis, and D. Akyildiz, "Analysis of VOT in Turkish speakers with aphasia," *Clinical Linguistics & Phonetics*, vol. 25, no. 4, pp. 287–301, Apr. 2011, <https://doi.org/10.3109/02699206.2010.529541>.
- [45] T. Cho, D. H. Whalen, and G. Docherty, "Voice onset time and beyond: Exploring laryngeal contrast in 19 languages," *Journal of Phonetics*, vol. 72, pp. 52–65, Jan. 2019, <https://doi.org/10.1016/j.wocn.2018.11.002>.
- [46] J. R. Orozco-Arroyave, J. D. Arias-Londoño, J. F. Vargas-Bonilla, M. C. González-Rátiva, and E. Nöth, "New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease," in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, May 2014, pp. 342–347.
- [47] J. Ryalls, M. Simon, and J. Thomason, "Voice Onset Time production in older Caucasian- and African-Americans," *Journal of Multilingual Communication Disorders*, vol. 2, no. 1, pp. 61–67, Mar. 2004, <https://doi.org/10.1080/1476967031000090980>.
- [48] J. Al-Tamimi and G. Khattab, "Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops," *Journal of Phonetics*, vol. 71, pp. 306–325, Nov. 2018, <https://doi.org/10.1016/j.wocn.2018.09.010>.
- [49] D. Montaña, Y. Campos-Roca, and C. J. Pérez, "A Diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of Parkinson's disease," *Computer Methods and Programs in Biomedicine*, vol. 154, pp. 89–97, Feb. 2018, <https://doi.org/10.1016/j.cmpb.2017.11.010>.

Markov Chain Monte Carlo Analysis of the Variable-Volume Exothermic Model for a Continuously Stirred Tank Reactor

Jean Pierre Muhirwa

Department of Applied Mathematics and Computational Science (AMCS), Nelson Mandela African Institution of Science and Technology Tanzania (NM-AIST)
Arusha, Tanzania and

Department of Mathematics, University of Rwanda, College of Science and Technology (UR-CST), Kigali, Rwanda
muhirwaj@nm-aist.ac.tz

Isambi Sailon Mbalawata

Research Department
African Institute for Mathematical Sciences (AIMS)
Kigali, Rwanda
imbawata@nexteinstein.org

Verdiana Grace Masanja

Department of Applied Mathematics and Computational Science (AMCS)
Nelson Mandela African Institution of Science and Technology Tanzania (NM-AIST)
Arusha, Tanzania
verdiana.masanja@nm-aist.ac.tz

Abstract-In this paper, a variable-volume Continuously Stirred Tank Reactor (CSTR) deterministic exothermic model has been formulated based on the Reynold Transport Theorem. The numerical analysis of the formulated model and the identifiability of its physical parameters are done by using the least squares and the Delayed-Rejection Adaptive Metropolis (DRAM) method. The least square estimates provide the prior information for the DRAM method. The overall numerical results show that the model gives an insight in describing the dynamics of CSTR processes, and 14 parameters of the CSTR are well identified through DRAM convergence diagnostic tests, such as trace, scatter, autocorrelation, histograms, and marginal density plots. Global sensitivity analysis was further performed, by using the partial rank correlation coefficients obtained from the Latin hypercube sampling method, in order to study and quantify the impact of estimated parameters, uncertainties on the model outputs. The results showed that 7 among the 14 estimated model parameters are very sensitive to the model outcomes and so those parameters need to be handled and treated carefully.

Keywords-parameter identifiability; variable volume; exothermic; CSTR; RTT; MCMC; DRAM

I. INTRODUCTION

During the past decades, Continuously Stirred Tank Reactors (CSTRs) have gained research momentum as important industrial and chemical production tools. For controlling the reactors, various methods have been proposed to tackle the complexity and the non-linearity operational behaviors that are present in the tank reactor during the production processes [1-6]. Discussions about CSTRs seem to

be broad and range from general to specific purposes. For example, mathematical modeling and numerical simulations of two-phases which are gas-liquid flow in the CSTR can be found in [7] and the Fokker-Plank Equation was applied for a two-state stochastic CSTR system in [8]. In [9], the robust feedback linearization of an isothermal CSTR was conducted by using the mixed sensitivity synthesis and iteration approaches in the presence of uncertainties. A one state variable, temperature, of a non-isothermal CSTR was analyzed by using Proportional Integral Derivative (PID) and fuzzy logic controllers, and the results from simulations and temperature control show that fuzzy logic can be adopted as a good controller of the process compared with the PID controller [10, 17]. The effects of hydrodynamic shear on biogas production in the CSTR using the Metzner-Otto method were analyzed and discussed in [11]. The Bayesian approach was used in [12] as the sorption parameter identifiability tool. The research outputs showed that the Bayesian inference is a more preferable method for the analysis of CSTR experiments as per numerical identifiability as well as the sorption parameter identifiability. The efficient Azo Dye color identification in the CSTR with the built-in bio-electrochemical system was developed for Azo dye alizarin Yellow R (AYR) which in turn can help in wastewater treatment [13]. Authors in [14] investigated the performance of CSTR as bioreactor for producing biohydrogen from water melon waste in the anaerobic digester. The Lyapunov-based stochastic non-linear model predictive control was used in [15] to shape the state probability density functions

in the CSTR with the exothermic reaction $A \xrightarrow{k} B$, where k is the reaction rate, A is the reactant and B is the product. The

Corresponding author: Jean Pierre Muhirwa

Luenberger fuzzy observer with and without sliding modes, the Walcott-Zak fuzzy observer, and the Utkin fuzzy observer were adopted and used as fault detection sensors of the CSTR in [16]. A two state CSTR stochastic model was studied and analyzed in [18] by using the Approximate Expectation Maximization (AEM) and Bayesian algorithms. It was revealed that the Bayesian algorithm is an effective method to be applied on the CSTR's stochastic model since it provided more accurate parameter estimates compared with AEM, and it is even more applicable for an unknown system with a small number of data sets. Authors in [19] used Monod and Haldane kinetics methods to perform the stability analysis of a system that models the formation of biofilms inside the CSTR during the waste-water treatment process. Even though both methods performed well, still the Monod kinetics provided biofilms formation in a shorter time compared to the opponent method.

The effect of operating conditions on the CSTR's performance with a saponification experiment was conducted in [20] and the results showed that the increase in conversion scale may depend on the increase in CSTR's volume. The neural network approach was used in [21] in order to identify the dynamics of two states, namely the temperature and the concentration of the CSTR's model, and reasonable and precise results were obtained. Three-dimensional Computational Fluids Dynamics (CFD) simulations were carried out [22] in order to identify the flow behaviors in the CSTR. Authors in [23] used the cascade control strategy to control the temperature of the exothermic CSTR with a cooling jacket. The stability analysis of the system was investigated with the Routh-Hurwitz and Argand diagram techniques. As a result, the cascade control was pointed out to be an efficient control method for the CSTR processes. A modified CSTR model for the neutralization process was studied and analyzed in [24]. This model has been used to assess the effects of strong acid (HCl) and strong base (NaOH) on the flow rates of ionic concentrations [24]. First-order and higher-order sliding mode observer methods were used in [25] in order to design and estimate states and unknown inputs of the CSTR, and it was shown that higher-order sliding modes may be adopted to reduce the noise into the system compared with the first-order sliding mode. Parameter estimation of non-linear chemical and biological processes with non-measured variables from a number of data sets was performed in [26] using the Bayesian approach. Furthermore, a mathematical model and simulation of reactors with the production experiment of hexane from benzene were conducted in [27] and the numerical investigation of phenol oxidation from waste-water inside a reactor was carried out in [28]. Non-parametric and non-linear stochastic dynamical model along with the behavior analysis of a class of the single state isothermal CSTR was also studied and analyzed in [29]. The adaptive method with recursive identification and the polynomial synthesis with placement of poles were applied on the CSTR system to determine its dynamics, however this method provided inappropriate control responses and overshoots [30]. The problem of characterizing the global dynamics of a single state non-linear stochastic CSTR system was addressed in [29] using the Fokker-Plank as the state probability distribution function, but the study of a several-state non-linear stochastic system is paramount as recommended in

this article. In the same way, different types of reactors and types of reactions in the chemical engineering processes are widely defined, explained and described in [31]. The modeling and control of the CSTR were also conducted based on a mixed logical dynamical model which resulted in a satisfactory performance of the tank [32]. In [33], the general model of the CSTR was developed and the transient behavior for irreversible non-linear polymerization process in the CSTR was studied and analyzed. Chemical process hazards, causes and proposed measures of safety of batch and semi-batch processes were discussed in [34]. The dynamical behavior of the CSTR through a single first order reaction was researched and analyzed in [35]. The limitations of CSTRs' performance due to cooling jacket dynamics with both open and closed loops are well explained and discussed in [36]. Authors in [37] performed multivariate character and stability analysis of irreversible exothermic CSTRs. The signal flow diagram and the equilibrium states were determined by taking into consideration the first and the second-order reactions.

Often, the volume is treated as a constant [4, 25, 38, 39], however, in reality the tanks may expand and lead to variable volume. Moreover, the parameters are randomly assumed and the estimation of physical parameters that are very influential in the system variation using MCMC methods is lagging behind. It is with these reasons this paper aims to treat the CSTR volume as a variable and identify 14 CSTR physical parameters.

II. MODEL FORMULATION

The assumptions on the parameters and physical properties inside and outside the tank have been considered during model formulation and are adopted from [25, 38-40], except for the volume which is considered to be a variable. To depict the problem, the CSTR dynamics are schematically illustrated using the CSTR and the Reynold Transport Theorem (RTT) diagrams.

A. Assumptions

The CSTRs under investigation assume perfect mixing to avoid spatial gradients of velocity, temperature, concentration and of other properties of the mixture.

- The CSTRs assume non-viscous liquid and gas phases which are frequently supplied in it and a static mixer which makes the shaft work produced by the stirring process to be negligible.
- No pressure drop is taking place in the CSTRs, i.e. CSTRs work at a constant pressure.
- Kinetic energy, potential energy, and other forms of external energy are infinitesimally small compared to the heat exchange and the heat from the chemical reactions.
- The wall is isolated and its temperature is negligible. Only the heat exchange is channelled through the designed area.
- The CSTRs assume the volume variation.
- The CSTRs assume density ρ and specific heat capacity C_p to be constants.

- Since there is negligible external stress acting on the system, it is assumed that there is also a negligible momentum on the system.

B. Formulation

Based on schematic diagrams presented in Figures 1 and 2, considering the above listed assumptions, and according to the transport phenomena, mass and energy conservation, the variable-volume exothermic CSTR model formulated is given by the system of equations (1):

$$\begin{cases} \frac{dV}{dt} = F_{in} - F_{out}, \\ \frac{dC}{dt} = \frac{F}{V}(C_{in} - C) - k_0 e^{-\frac{E}{R}(\frac{1}{T} - \frac{1}{T_{mean}})} C, \\ \frac{dT}{dt} = \frac{F}{V}(T_{in} - T) + \frac{-H^* k_0 e^{-\frac{E}{R}(\frac{1}{T} - \frac{1}{T_{mean}})} C}{\rho C_p} - \frac{UA(T - T_c)}{V \rho C_p}, \\ \frac{dT_c}{dt} = \frac{F_c}{V_c}(T_{c_{in}} - T_c) + \frac{UA(T - T_c)}{V_c \rho_c C_{p_c}}, \end{cases} \quad (1)$$

where V is the tank volume, $F_{in} = F$ is the feeding velocity, F_{out} is the outlet velocity, C_{in} is the feeding concentration, C is the mixture concentration, E is the activation energy, T is the temperature of the tank, T_{in} is the feeding temperature, T_{mean} is the reference temperature, H^* is the reaction enthalpy, ρ is the density, C_p is the heat capacity, U is the heat transfer coefficient, A is the cross-sectional area, T_c is the cooling temperature, F_c is the cooling velocity, V_c is the jacket volume, $T_{c_{in}}$ is the feeding coolant temperature, ρ_c is the coolant density, C_{p_c} is the jacket's heat capacity, and $k = k_0 e^{-\frac{E}{R}(\frac{1}{T} - \frac{1}{T_{mean}})}$ is the Arrhenius equation [41], with R being the gas law constant and k_0 the pre-Arrhenius frequency factor.

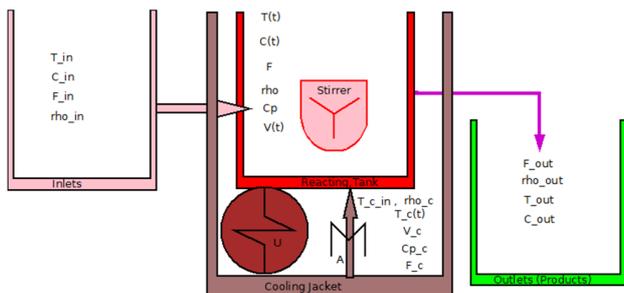


Fig. 1. Schematic diagram that describes the dynamics of CSTR.

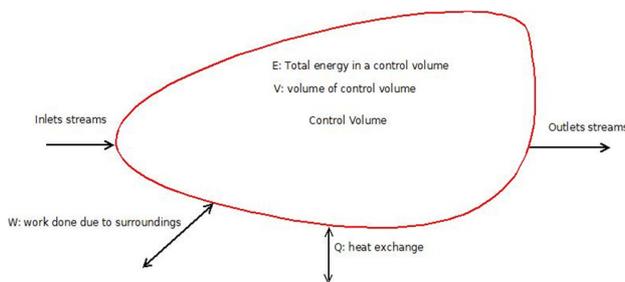


Fig. 2. Schematic illustration of the RTT.

III. MATERIALS AND METHODS

A. Markov Chain Monte Carlo Method

The Markov Chain Monte Carlo (MCMC) is a numerical analysis method used as a statistical and Bayesian technique to identify the complex ordinary differential equations' parameters that fit the dynamics of chemical and biological models [42-44]. In this paper, the DRAM method is used to estimate 14 parameters of the model (1). The Bayesian inference, qualified to be a very powerful statistical technique, has been widely used to identify the model's parameters θ which are obtained after evaluating the parameters' posterior density $p(\theta / X_{p1}, X_{p2}, \dots, X_{pN})$, where $X_{p1}, X_{p2}, \dots, X_{pN}$ are the measurement points of the chemical process. For the model (1), $\theta = [F_{out}, F, k_0, E, T_{mean}, H^*, \rho, C_p, U, A, F_c, V_c, \rho_c, C_{p_c}]$ and $X_p = (V, C, T, T_c)$.

Further, the MCMC method is a sampling technique that combines Monte Carlo integration and Markov chains [45]. The overall implementation process starts from proposing a suitable distribution, called proposal distribution, and drawing samples from it [46]. The proposal distribution sometimes depends on the present value to form the chain, which in turn is considered as a Markov chain. The acceptance or rejection mechanisms are employed in the simulation to rectify the trial proposal distribution which ends up with the target distribution. In the end, a simulated chain of parameters (drawn samples, $\theta_1, \theta_2, \dots, \theta_N$) can be used to approximate the intractable integral (distribution), as:

$$E[f(\theta) / X_{p1}, X_{p2}, \dots, X_{pN}] \approx \frac{1}{N} \sum_{i=1}^N f(\theta_i)$$

where $E[f(\theta) / X_{p1}, X_{p2}, \dots, X_{pN}]$ is the expectation and $f(\theta_i)$ is the density function.

The Metropolis algorithm, Metropolis Hastings, Hamiltonian Monte Carlo, Gibbs Sampler, Reversible Jump Monte Carlo, Metropolis Adjusted Langevin, Slice Sampling, Multiple Try Metropolis, and Delayed Rejection Metropolis are among the most used MCMC algorithms [47, 48]. However, the method considered in this paper is the DRAM, presented in Algorithm 1 since it has the desirable feature of tuning the reliable proposal distribution without defining it manually [45, 46]. This method overcomes the tedious task of trial and error of tuning the suitable proposal distribution that may appear in the Metropolis-Hastings technique [49]. In this case, the Gaussian distribution is mostly used as a proposal distribution. For a startup, MCMC needs the initial parameter values and a proposal distribution. The DRAM method is used to propose a distribution and its initial parameter values are computed using the classical least square method. The details of the numerical results obtained from the least squares and the MCMC methods are well explained and discussed in Section IV.

B. Least Squares Method

The model in (1) is a system of time dependent differential equations of the form:

$$\frac{dX_p}{dt}(X_p, \theta_j, t_i) = 0$$

where p is the number of state variable and X_p is the vector of

state variables. For the model (1), $p = 4$, t_i is the discrete time sampling, and $X_p = (V, C, T, T_c)$, θ_j is the set of parameters of the model to be identified, where $j = 1, 2, \dots, 14$. So, in this paper, the 14 parameters to be identified are:

$$\theta_j = [F_{out}, F, k_0, E, T_{mean}, H^*, \rho, C_p, U, A, F_c, V_c, \rho_c, C_{p_c}]$$

The least squares method one of the classical optimization methods with the purpose of minimizing the sum of squared residuals. To obtain the residuals, a predictive model is described as:

$$Y_{pi} = g(X_{pi}, \theta_j, t_i) + r_{pi}$$

where Y_{pi} represents the observations, $g(X_{pi}, \theta_j, t_i)$ are the solutions of the model (1) at time t_i starting from a fixed set of parameters θ_0 , and r_{pi} is the residual term.

The sum of squared residuals function is obtained by taking the sum of squares of r_{pi} :

$$S.S.R = \sum_{i=1}^m (r_{pi})^2 = \sum_{i=1}^m (Y_{pi} - g(X_{pi}, \theta_j, t_i))^2$$

Hence, the least square method searches the best fitting parameters that minimize the $S.S.R$ taken as the likelihood function. Thus, the fitting set of parameters θ_j can be obtained after solving (2):

$$\frac{\partial (S.S.R)}{\partial \theta_j} = 0 \equiv \frac{\partial (\sum_{i=1}^m (Y_{pi} - g(X_{pi}, \theta_j, t_i))^2)}{\partial \theta_j} = 0 \quad (2)$$

simultaneously. Almost all chemical processes are intractable and complex due to their non-linearity behaviors. In fact, stepping back to the model (1) with 14 parameters of interest, if we could manually solve (2), we would end up with solving 14 nonlinear equations simultaneously, which is a complicated task. As a simplification, the numerical simulations become a usual way of solving the problem.

C. Delayed Adaptive Markov Chain Monte Carlo

In this paper, the delayed-rejection adaptive metropolis (DRAM) algorithm is used for the parameter identifiability of the deterministic variable-volume exothermic CSTR model described in (1). The DRAM algorithm is chosen since it combines and utilizes both the Delayed-Rejection (DR) and the Adaptive Metropolis (AM) methods to efficiently improve and speed-up the computation process for a slow-start adaptation. The inputs of the algorithm are the initial values of the parameters and all of them are the least square optimal parameter values shown in Table II. The proposed distribution which is a Gaussian with mean 0 and standard deviation Σ_0 taken as the initial covariance matrix, $\Sigma_0 = \frac{10^{-4} I_j}{\sqrt{j}}$, has been used, where j is the length of model's parameters to be estimated, and $I_j = I_{j \times j}$ which is an $j \times j$ identity matrix.

1) Algorithm 1: DRAM

Draw the initial point θ_0 , from initial distribution $p_0(\theta)$. Set an initial non-adaptive period N_0 and initial covariance matrix Σ_0 .

For $i = 1, 2, \dots$ perform the following:

(i) Sample a current point $\hat{\theta}$ from the proposal distribution

$q(\hat{\theta}/\theta_{i-1})$

(ii) Compute the acceptance probability using:

$$\alpha(\theta_{i-1}, \hat{\theta}) = \min(1, \frac{p(\hat{\theta}/X_{p1}, X_{p2}, \dots, X_{pN})q(\theta_{i-1}/\hat{\theta})}{p(\theta_{i-1}/X_{p1}, X_{p2}, \dots, X_{pN})q(\hat{\theta}/\theta_{i-1})})$$

(iii) Acceptance/rejection rule by setting:

$$\theta_i = \begin{cases} \hat{\theta}, & \text{if } u < \alpha(\theta_{i-1}, \hat{\theta}) \text{ where } u \sim \mathbb{U}(0, 1) \\ \theta_{i-1}, & \text{otherwise.} \end{cases}$$

(iv) If $n_0 \ll i$ (or after every n_0^{th} iterations), then update the covariance matrix by using:

$$\Sigma_i = Cov(\theta_0, \theta_1, \dots, \theta_i) + \epsilon I_d$$

where I_d is an $d \times d$ identity matrix and ϵ is a small positive number that makes the matrix Σ_i to be non-singular [46, 52, 53].

(v) $i \leftarrow i+1$

TABLE I. VARIABLES, CONSTANTS, AND PARAMETERS VALUES

Parameters and variables				
Symbol	Unit	Physical meaning	Values	Source
C_{in}	$\frac{kmol}{min\ m^3}$	Feeding concentration	316.8	[50]
C_0	$\frac{kmol}{min\ m^3}$	Initial concentration	316.8	[50]
C	$\frac{kmol}{min\ m^3}$	Mixture concentration	State variable	Simulated
T_{in}	K	Feeding temperature	298.35	[50]
T_0	K	Initial temperature	298.35	[50]
T	K	Mixture temperature	State variable	Simulated
H^*	$\frac{kcal}{kmol}$	Enthalpy	-1004.3×10^3	[50]
T_{co}	K	Initial cool temperature	288.15	[50]
T_{cin}	K	Feeding cool temperature	293	[38]
T_c	K	Jacket temperature	State variable	Simulated
R	$\frac{kJ}{K\ kmol}$	Gas constant	8.314	[50]
V_0	m^3	Initial tank volume	100	[51]
V	m^3	Tank volume	State variable	Simulated
E	$\frac{kJ}{kmol}$	Activation energy	0.5	[50]
T_{mean}	K	Reference temp	298.15	[50]
C_p	$\frac{kcal}{K\ kg}$	Heat capacity	4186	[50]
U	$\frac{kJ}{K\ min\ m^2}$	Heat transfer coefficient	100000	Assumed
F_c	$\frac{m^3}{min}$	Cooling velocity	46.5×10^{-6}	[50]
V_c	m^3	Jacket volume	50×10^{-6}	[50]
ρ_c	$\frac{kg}{m^3}$	Density of the coolant	1000	[51]
C_{p_c}	$\frac{kcal}{K\ kg}$	Heat capacity of jacket	4186	[50]
F_{out}	$\frac{m^3}{min}$	Outlet velocity	130×10^{-6}	[50]
F	$\frac{m^3}{min}$	Feeding velocity	130×10^{-2}	Assumed
k_0	min^{-1}	Pre-Arrhenius factor	0.9	[50]
ρ	$\frac{kg}{m^3}$	Density	1000	[51]
A	m^2	Area	0.015	[50]

IV. NUMERICAL ANALYSIS OF THE MODEL AND DISCUSSION

The formulated model (1) is solved numerically by the 4th order Runge-Kutta method which is the ode45 solver available in Matlab R2016b software package. The parameter identifiability is done using the least squares and MCMC.

A. Numerical Simulations

Due to lack of actual information (real data) about the system, the model (1) is simulated by using literature and assumed values (Table I), the discrete sampling time points are 100, and 100×4 numerical solutions of the model Y_{pi} presented in (1) are obtained. The numerical results from the subplot 1 of Figure 3 reveal that the volume of tank reactor increases from 100m^3 to approximately 126m^3 and this is an indication of having non-constant flow rates of reactants due change at both the inlets and outlets.

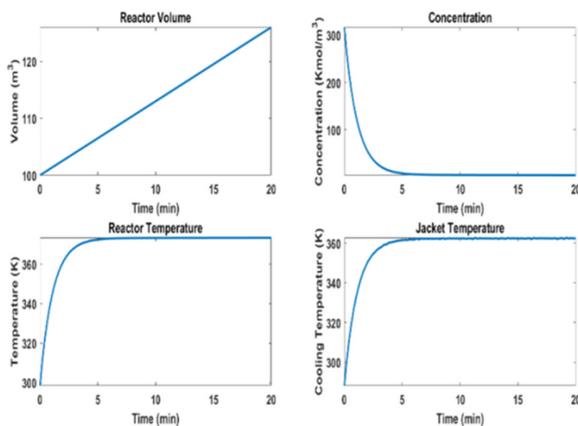


Fig. 3. Numerical solutions of the model.

Figure 3 also shows that the reactants are consumed inside the reacting tank as its concentration approaches zero, which means the complete mixing and at the same time symbolizes a non-partial conversion of reactants into products that may lead to time residence distribution analysis as one of the inconveniences of CSTRs. Along this process, there is a covering cooling jacket that communicates with the reacting tank through a designed cross-sectional area of 0.015m^2 to cool down the rising temperature inside the reacting tank. The covering cooling jacket contains the substance whose temperature is initially lower than the starting temperature of the reacting tank to disable the explosion of the reaction. During this process, the temperature of the reacting tank rises from 298.35K to its operating working temperature that is 373.48K . Meanwhile, the cooling temperature of the covering cooling jacket also rises from its initial temperature 288.15K until it reaches 363.14K . As a result, if this scenario is selected to be a piloting tank, then all the simulated information and the working conditions that are described above have to be taken into consideration quantitatively.

B. Least Square Results

The deterministic variable-volume exothermic CSTR model presented in (1) has been numerically analyzed by using the least square method and 14 parameters of the model were

optimized from the noisy measurements of the system. The noisy measurements are obtained after corrupting the obtained empirical data with Gaussian noise of 0.05 standard deviation. The obtained results are presented in Figure 4, and the 14 estimated parameters and their corresponding initial values are also presented in Table II. According to the results, all parameters relatively converge to their corresponding initial values and the measurements fit the exothermic CSTR model as can be viewed from fitted state variables shown in Figure 4.

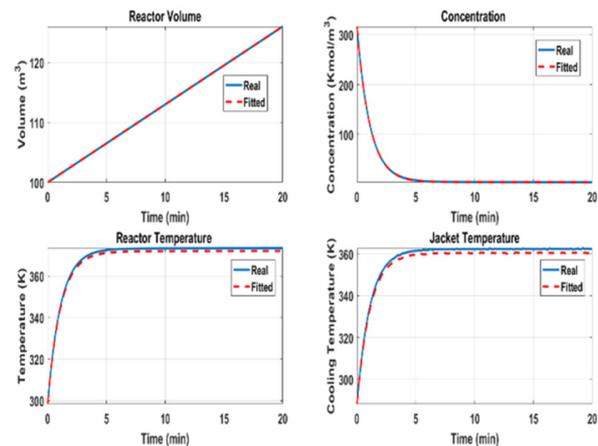


Fig. 4. Least square fitting model. We can see that the measurements fit the model well as parameters vary.

C. Markov Chain Monte Carlo Results

The initial values for the MCMC are estimated using the least squares method. Since we have fourteen parameters to be estimated, a sample consisting of $100,000 \times 14$ parameters has been generated during simulations and the method has been adapted 100 times. Also the initialized covariance matrix of MCMC is chosen to be $\Sigma_0 = \frac{10^{-4} I_{14 \times 14}}{\sqrt{14}}$. Details of the MCMC results are discussed by the following MCMC diagnostic tests.

1) MCMC Diagnostics

The identifiability of the model parameters is mainly based on the convergence of the MCMC. There are several statistical and graphical convergence tests for the MCMC methods [54-57]. In this paper, we use the common statistical inference and graphical analyses of the trace (time series plots), scatter plots, autocorrelation plots, histograms, and the marginal density distributions for each drawn sample of the parameters to identify the model and diagnose the convergence of the generated MCMC samples.

2) Trace Plots

One way to analyze the convergence of the MCMC is to check the mixing of the generated sample of posteriors through trace plots. If the generated chain of posteriors becomes stationary for several initial values, and there are no obvious spikes, it is an indication of having a good mixing which is a good sign of convergence. In Figure 5 we can see that the mixing of samples is very good so the chain converges, and the 14 sampled values of posteriors are the means (centers) of the samples.

TABLE II. THE STATISTICAL INFERENCE ON THE ESTIMATED PARAMETERS BY USING THE LEAST SQUARE AND MCMC METHODS

Param	LSQ and MCMC estimates										
	Initial values	LSQ estim	MCMC posterior mean	MCMC posterior median	Cred interv	Std	MCerr	tau	geweke	Kurtosis	Skewness
F_{out}	130×10^{-6}	130.016×10^{-6}	130.015×10^{-6}	130.016×10^{-6}	$[130.015 \times 10^{-6}, 130.016 \times 10^{-6}]$	7.32×10^{-8}	1.0633×10^{-9}	41.003	0.99994	2.9626	-0.0630
F	130×10^{-2}	130.09×10^{-2}	$129.971139 \times 10^{-2}$	$129.9704569 \times 10^{-2}$	$[129.9705396 \times 10^{-2}, 129.9716882 \times 10^{-2}]$	0.00093	3.9735×10^{-5}	62.713	0.99997	3.0915	-0.0225
k_0	0.9	0.900553	0.900592376	0.900596750	$[0.900586482, 0.900598270]$	0.000951	2.2856×10^{-5}	50.329	0.99994	3.0930	-0.0016
E	0.5	0.493183	0.493486288	0.493470050	$[0.493480384, 0.493492191]$	0.0009525	1.7754×10^{-5}	46.313	0.9985	3.1854	0.0934
T_{mean}	298.15	298.147	298.151888471	298.151879315	$[298.151882554, 298.151894388]$	0.00095466	1.9926×10^{-5}	58.812	1	3.0988	-0.0220
H^*	-1004.3×10^3	1026.03×10^3	-1028032.329702401	-1028032.329677610	$[-1028032.32970, -1028032.3296966452]$	0.00095973	3.2217×10^{-5}	57.28	1	3.1340	0.0125
ρ	1000	1017.16	1026.121280804	1026.121304498	$[1026.121274963, 1026.121286644]$	0.00094236	2.9081×10^{-5}	59.644	1	3.0811	-0.0299
C_p	4186	4231.14	4217.368624072	4217.368601824	$[4217.368617960, 4217.368630185]$	0.00098619	2.1133×10^{-5}	48	1	3.3398	0.0578
U	100,000	100681	100098.502552738	100098.502551528	$[100098.502547056, 100098.502558419]$	0.00091663	4.2388×10^{-5}	69.113	1	3.2362	0.0790
A	0.015	0.0151979	0.014959065	0.014962113	$[0.014953136, 0.014964995]$	0.00095668	3.0706×10^{-5}	53.978	0.99458	3.0965	-0.0279
F_c	46.5×10^{-6}	46.7705×10^{-6}	74.7936×10^{-5}	62.4641×10^{-5}	$[74.4298 \times 10^{-5}, 75.1574 \times 10^{-5}]$	0.00058698	3.7635×10^{-5}	85.568	0.21093	3.9661	1.0355
V_c	50×10^{-6}	49.6872×10^{-6}	76.2673×10^{-5}	63.7067×10^{-5}	$[75.8967 \times 10^{-5}, 76.6379 \times 10^{-5}]$	0.00059791	3.2879×10^{-5}	69.882	0.41264	3.8227	1.0014
ρ_c	1000	987.751	994.859555082	994.859566589	$[994.859549347, 994.859560817]$	0.00092534	2.464×10^{-5}	63.819	1	3.2461	0.0145
C_{Pc}	4186	4134.6	4163.918586616	4163.918608754	$[4163.918580689, 4163.918592542]$	0.00095621	1.7679×10^{-5}	45.604	1	3.0466	-0.0150

Note: Param=Parameters, LSQ=least squares, estim=estimates, Cred interv=Credible interval, Std=Standard deviation, MCerr=MCMC error

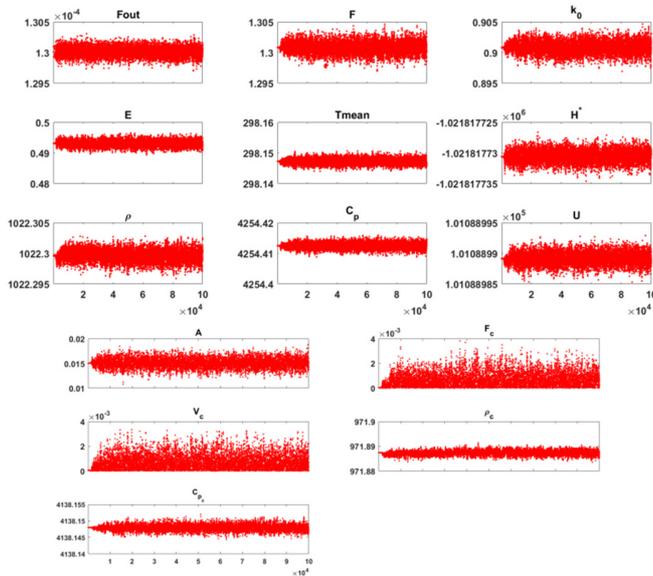


Fig. 5. The trace plots of the samples.

3) Scatter Plots (pairs)

A poor convergence of MCMC method inventively leads to high correlation between estimated parameters. Since there are

14 parameters to be identified, then there are 91 scatter plots which we need to investigate if there are strong correlations between them. Figure 6 shows the scatter plots for the first 10 sampled parameters equivalent to 45 scatter plots and how these posterior samples correlate to each other. We observe that there is no high correlation between the pairs of the estimated parameters and so the parameters of the model (1) are adequately identified.

4) Autocorrelation Plots

Figure 7 determines and examines the correlation between consecutive samples during posteriors chain sampling. We can see that the coefficients of autocorrelation functions of all generated samples (x-axis) tend to zero as the number of lags (y-axis) increases and get stationary around zero after 100 lags. That is an indication of having a good mix.

5) Marginal Distribution (Density) of the Sampled Parameters

Another diagnostic test is to plot the posteriors density distributions. Normally for better mixing we expect the distributions of all density estimations to follow a Gaussian distribution. Figure 8 depicts how density distributions of all estimated parameters follow Gaussian distribution and their values are taken as means of distributions.

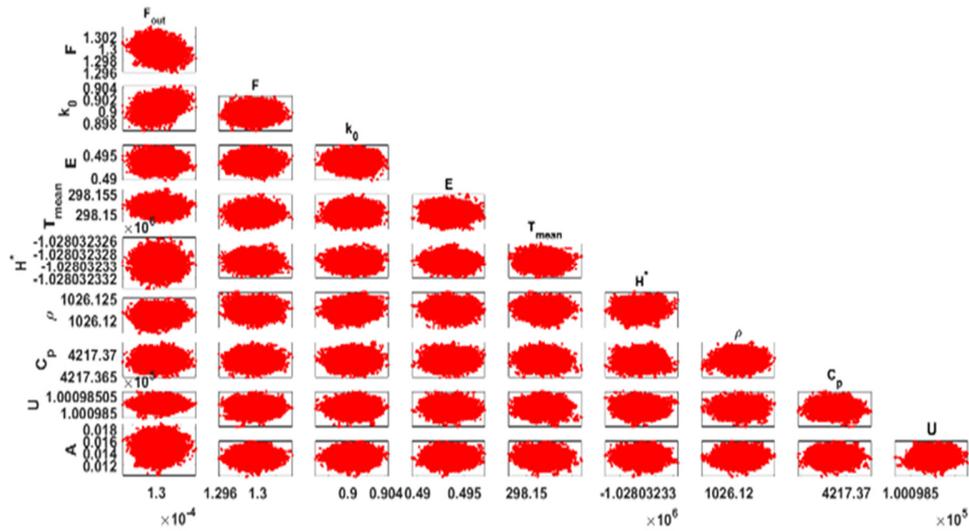


Fig. 6. The scatter plots of pairs of sampled parameters. We can see that none of the parameters $[F_{out}, F, k_0, E, T_{mean}, H^*, \rho, C_p, U, A]$ is correlated to the others because there are no correlation patterns observed.

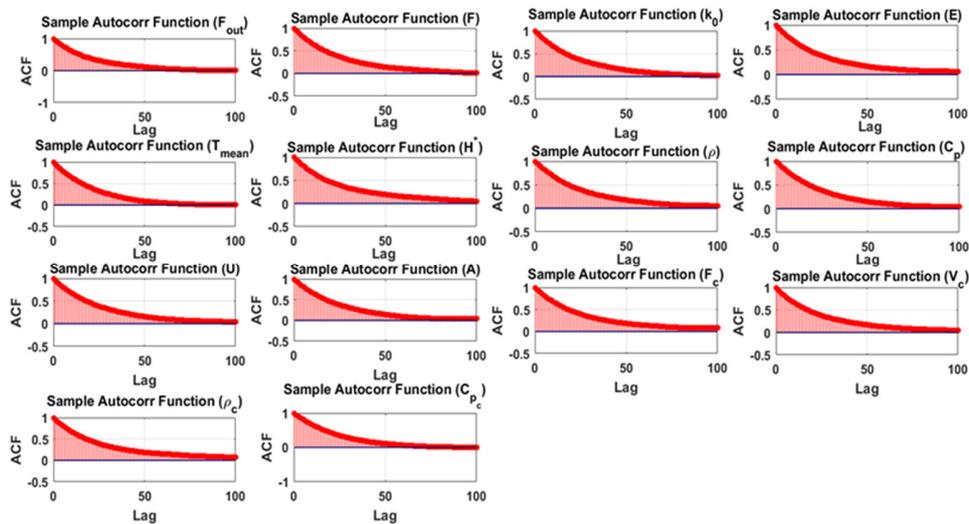


Fig. 7. Autocorrelation plots of the sampled parameters.

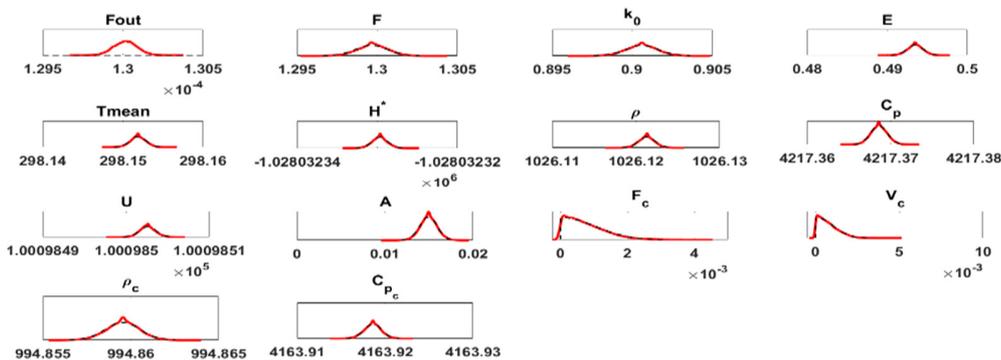


Fig. 8. The marginal density distribution of the sampled parameters.

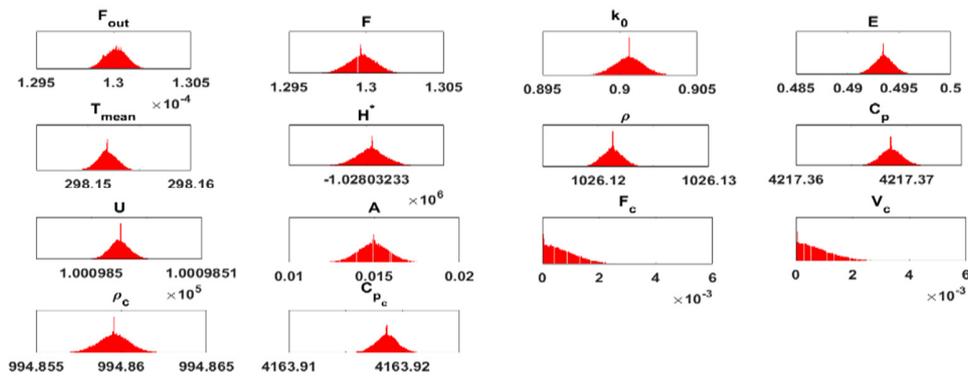


Fig. 9. Histograms of the 14 sampled parameters.

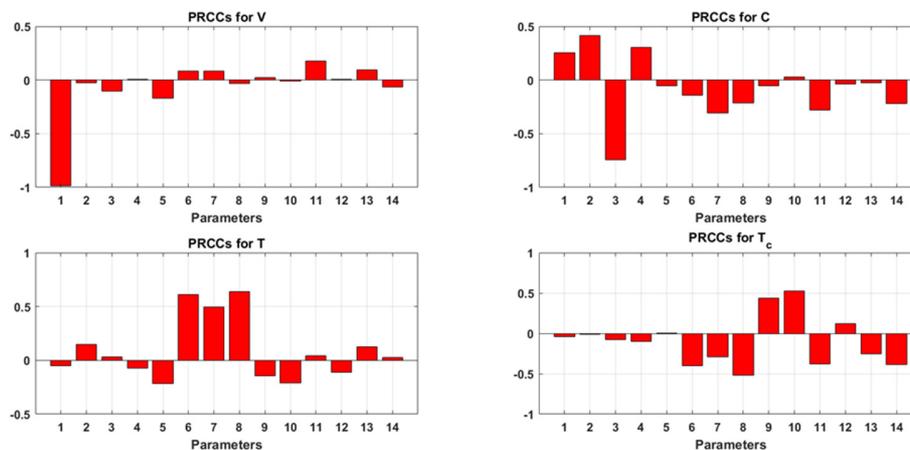


Fig. 10. Parameter PRCCs' plot for the model outputs.

6) Histograms of the Sampled Parameters

From Figure 9, we can see that almost all parameter values follow the normal distribution. The MCMC's convergence tends to be bad for F_c and V_c . However, the other 12 parameters are fine. This could be caused by different and uncertain reasons. It may stem from either the complexity of the model, the little information about their mean values, or the method itself has failed to sample well these two parameters from the proposed distribution.

The results in Table II show that almost all parameters converge to their initial values and posterior means are within their credible intervals. We can see that the posterior means and posterior medians are nearly equal. We can in addition observe that the Markov chain errors and the standard deviation of the model parameters are significantly small which shows that the method performed well in identifying the best estimates. The computed autocorrelation time (τ) values are small with the significance of having independence in sampling the successive posteriors as well as the convergence of the method. All Geweke values are almost 1, except 0.21093 for F_c and 0.41264 for V_c with the indication of having non-stationary means values by default for the start (10%) and the end (50%)

of sampling of these two parameters. In reality, the skewness and kurtosis values for the Gaussian distribution are approximately expected to be 0 and 3 respectively. So, the skewness and kurtosis values for the obtained samples show that only F_c and V_c values are somehow skewed in the right hand side and do not incorporate the exact features of the Gaussian distribution.

D. Global Sensitivity Analysis

To quantify the uncertainty effects of estimated parameters on the model outputs, we performed global sensitivity analysis by using Latin hypercube sampling method. This method computes the Partial Rank Correlation Coefficients (PRCCs) and determines which parameters are globally very sensitive to the model. PRCCs values vary between -1 and +1 with high sensitivity index for values approaching ± 1 and low sensitivity index for values which are far from ± 1 . The PRCCs values are graphically displayed in Figure 10 and their results can be seen in Table III.

From Figure 10, we observe that the 1st parameter significantly and negatively affects the model volume outputs. The increase in that parameter decreases the volume. From subplot 2, the 3rd parameter, which is k_0 , has high negative effect on the concentration outputs of the model. The increase

of the temperature profiles positively depends on the increase in the 6th, 7th and 8th parameter values whereas the increase of the 8th and 10th parameter values decreases and increases the cooling temperature profiles respectively. So, in total there are 7 identified parameters that may have a great impact on the model response and their values are represented by (*) in Table III.

TABLE III. PRCC RESULTS

Parameter	Partial rank correlation coefficient values for each variable			
	V	C	T	T _c
F _{out}	-0.9847(*)	0.2545	-0.0479	-0.0429
F	-0.0269	0.4127	0.1482	-0.0119
k ₀	-0.1060	-0.7418(*)	0.0302	-0.0761
E	0.0070	0.3054	-0.0743	-0.0973
T _{mean}	-0.1713	-0.0515	-0.2141	0.0032
H*	0.0813	-0.1444	0.6114(*)	-0.3979
ρ	0.0819	-0.3102	0.5000(*)	-0.2924
C _p	-0.0345	-0.2137	0.6391(*)	-0.5186(*)
U	0.0232	-0.0547	-0.1470	0.4348
A	-0.0091	0.0298	-0.2086	0.5232(*)
F _c	0.1761	-0.2802	0.0415	-0.3810
V _c	0.0038	-0.0401	-0.1139	0.1187
ρ _c	0.0928	-0.0246	0.1285	-0.2502
C _{p,c}	-0.0655	-0.2189	0.0294	-0.3894

V. CONCLUSION

This paper focused on the mathematical formulation of the deterministic variable-volume exothermic CSTR model, and its numerical simulations were tested to supplement the theoretical results as applications using literature values. The identifiability of the parameters required real system measurements, but due to the absence of real data, the system measurements were simulated, i.e. the numerical solutions were empirical data corrupted with Gaussian noise with a standard deviation of 0.05.

The identifiability of the physical parameters of the formulated model was numerically carried out by using the least squares and DRAM. The least square estimates converged to the literature values and were treated as prior information for the DRAM method. The generated DRAM samples were graphically and statistically analyzed to test the convergence. The results show that the model was solved and its physical parameters were well identified. To study and quantify the associated parameter uncertainties to the model outcomes, global sensitivity analysis was performed by using the Latin hypercube sampling method. Seven parameters among the 14 estimated model parameters were found to be very sensitive to the model outputs and therefore, these parameters need to be controlled effectively.

ACKNOWLEDGEMENT

The authors acknowledge the financial support received from the German Academic Exchange Service (DAAD) and the research department at African Institute for Mathematical Sciences (AIMS)-Rwanda Center.

REFERENCES

- [1] A. S. Al-Araji, "Modeling of Continuous Stirred Tank Reactor based on Artificial Neural Network," *Al-Nahrain Journal for Engineering Sciences*, vol. 18, no. 2, pp. 202–207, 2015.
- [2] A. O. Ahmed, G. A. Gasmelseed, A. B. Karama, and A. E. Musa, "Cascade Control of a Continuous Stirred Tank Reactor (CSTR)," *Journal of Applied and Industrial Sciences*, vol. 1, no. 4, pp. 16–23, 2013.
- [3] H. A. Maddah, "Numerical Analysis for the Oxidation of Phenol with TiO₂ in Wastewater Photocatalytic Reactors," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3463–3469, Oct. 2018, <https://doi.org/10.48084/etasr.2304>.
- [4] A. Simorgh, A. Razminia, and V. I. Shiryaev, "System identification and control design of a nonlinear continuously stirred tank reactor," *Mathematics and Computers in Simulation*, vol. 173, pp. 16–31, Jul. 2020, <https://doi.org/10.1016/j.matcom.2020.01.010>.
- [5] A. Uppal, W. H. Ray, and A. B. Poore, "On the dynamic behavior of continuous stirred tank reactors," *Chemical Engineering Science*, vol. 29, no. 4, pp. 967–985, Apr. 1974, [https://doi.org/10.1016/0009-2509\(74\)80089-8](https://doi.org/10.1016/0009-2509(74)80089-8).
- [6] A. S. Ibrehem, "Modified Mathematical Model For Neutralization System In Stirred Tank Reactor," *Bulletin of Chemical Reaction Engineering & Catalysis*, vol. 6, no. 1, pp. 47–52, May 2011, <https://doi.org/10.9767/brec.6.1.825.47-52>.
- [7] A. Z. Al-Khazaal, F. Ahmad, and N. Ahmad, "Study on the Removal of Thiosulfate from Wastewater by Catalytic Oxidation," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 4053–4056, Apr. 2019, <https://doi.org/10.48084/etasr.2553>.
- [8] B. G. Osorio, H. B. Castro, and J. D. S. Torres, "State and unknown input estimation in a CSTR using higher-order sliding mode observer," in *IEEE IX Latin American Robotics Symposium and IEEE Colombian Conference on Automatic Control*, Bogota, Colombia, Oct. 2011, pp. 1–5, <https://doi.org/10.1109/LARC.2011.6086829>.
- [9] C. G. Hill, *An Introduction to Chemical Engineering Kinetics and Reactor Design*. New York, USA: Wiley, 1977.
- [10] J. Du, C. Song, and P. Li, "Modeling and Control of a Continuous Stirred Tank Reactor Based on a Mixed Logical Dynamical Model," *Chinese Journal of Chemical Engineering*, vol. 15, no. 4, pp. 533–538, Aug. 2007, [https://doi.org/10.1016/S1004-9541\(07\)60120-7](https://doi.org/10.1016/S1004-9541(07)60120-7).
- [11] D. P. Karadimou, P. A. Papadopoulos, and N. C. Markatos, "Mathematical modelling and numerical simulation of two-phase gas-liquid flows in stirred-tank reactors," *Journal of King Saud University - Science*, vol. 31, no. 1, pp. 33–41, Jan. 2019, <https://doi.org/10.1016/j.jksus.2017.05.015>.
- [12] D. Ndanguza, J. P. Muhirwa, and A. Uwimana, "Modeling and parameters estimation of a Spatial Predator-Prey distribution," *Rwanda Journal of Engineering, Science, Technology and Environment*, vol. 2, no. 1, pp. 1–17, Jul. 2019, <https://doi.org/10.4314/rjeste.v2i1.5>.
- [13] D. Tamboli and R. Chile, "Multi-model approach for 2-DOF control of nonlinear CSTR process," *International Journal of Modelling, Identification and Control*, vol. 30, no. 2, pp. 143–161, Jan. 2018, <https://doi.org/10.1504/IJMIC.2018.094208>.
- [14] E. Shakeri, G. Latif-Shabgahi, and A. E. Abharian, "Design of an intelligent stochastic model predictive controller for a continuous stirred tank reactor through a Fokker-Planck observer," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 10, pp. 3010–3022, Jun. 2018, <https://doi.org/10.1177/0142331217712583>.
- [15] E. Vlahakis and G. Halikias, "Temperature and concentration control of exothermic chemical processes in continuous stirred tank reactors," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 15, pp. 4274–4284, Nov. 2019, <https://doi.org/10.1177/0142331219855591>.
- [16] E. A. Buehler, J. A. Paulson, and A. Mesbah, "Lyapunov-based stochastic nonlinear model predictive control: Shaping the state probability distribution functions," in *American Control Conference*, Boston, MA, USA, Jul. 2016, pp. 5389–5394, <https://doi.org/10.1109/ACC.2016.7526514>.

- [17] E. H. Karimi and K. B. McAuley, "A Bayesian Method for Estimating Parameters in Stochastic Differential," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 147–152, Jan. 2015, <https://doi.org/10.1016/j.ifacol.2015.08.172>.
- [18] F. Remo, L. S. Luboobi, I. S. Mabalawata, and B. K. Nannyonga, "A mathematical model for the dynamics and MCMC analysis of tomato bacterial wilt disease," *International Journal of Biomathematics*, vol. 11, no. 1, Aug. 2017, Art. no. 1850001, <https://doi.org/10.1142/S1793524518500018>.
- [19] G. I. Valderrama-Bahamondez and H. Frohlich, "MCMC Techniques for Parameter Estimation of ODE Based Models in Systems Biology," *Frontiers in Applied Mathematics and Statistics*, vol. 5, 2019, Art. no. 55, <https://doi.org/10.3389/fams.2019.00055>.
- [20] G. L. Foutch and A. H. Johannes, "Reactors in Process Engineering," in *Encyclopedia of Physical Science and Technology*, 3rd ed., R. A. Meyers, Ed. New York, NY, USA: Academic Press, 2003, pp. 23–43.
- [21] H. Ballesteros-Moncada, E. J. Herrera-Lopez, and J. Anzurez-Marin, "Fuzzy model-based observers for fault detection in CSTR," *ISA Transactions*, vol. 59, pp. 325–333, Nov. 2015, <https://doi.org/10.1016/j.isatra.2015.10.006>.
- [22] H. Haario, E. Saksman, and J. Tamminen, "Adaptive proposal distribution for random walk Metropolis algorithm," *Computational Statistics*, vol. 14, no. 3, pp. 375–395, Sep. 1999, <https://doi.org/10.1007/s001800050022>.
- [23] H. Haario, E. Saksman, and J. Tamminen, "An adaptive Metropolis algorithm," *Bernoulli*, vol. 7, no. 2, pp. 223–242, Apr. 2001, <https://doi.org/bj/1080222083>.
- [24] I. A. Stepanov, "Exact Calculation of the Internal Energy of the Ideal Gas in Statistical Mechanics," *Physical Science International Journal*, vol. 14, no. 5, pp. 1–5, Apr. 2017, <https://doi.org/10.9734/PSIJ/2017/32822>.
- [25] I. S. Mbalawata, S. Sarkka, and H. Haario, "Parameter estimation in stochastic differential equations with Markov chain Monte Carlo and non-linear Kalman filtering," *Computational Statistics*, vol. 28, no. 3, pp. 1195–1223, Jun. 2013, <https://doi.org/10.1007/s00180-012-0352-y>.
- [26] I. S. Mbalawata, S. Sarkka, M. Vihola, and H. Haario, "Adaptive Metropolis algorithm using variational Bayesian adaptive Kalman filter," *Computational Statistics & Data Analysis*, vol. 83, pp. 101–115, Mar. 2015, <https://doi.org/10.1016/j.csda.2014.10.006>.
- [27] J. C. Etchells, "Process Intensification: Safety Pros and Cons," *Process Safety and Environmental Protection*, vol. 83, no. 2, pp. 85–89, Mar. 2005, <https://doi.org/10.1205/psep.04241>.
- [28] J. Jiang, J. Wu, S. Poncin, and H. Z. Li, "Effect of hydrodynamic shear on biogas production and granule characteristics in a continuous stirred tank reactor," *Process Biochemistry*, vol. 51, no. 3, pp. 345–351, Mar. 2016, <https://doi.org/10.1016/j.procbio.2015.12.014>.
- [29] J. Vojtesek and P. Dostal, "Simulation analysis of continuous stirred tank reactor," in *22nd European Conference on Modelling and Simulation*, Nicosia, Cyprus, Jun. 2008, pp. 1–6.
- [30] J. Vojtesek and P. Dostal, "Simulation of adaptive control of continuous stirred tank reactor," *International Journal of Simulation Modelling*, vol. 8, no. 3, pp. 133–144, 2009.
- [31] J. P. Muhirwa and D. Ndanguza, "Effect of random noise, quasi random noise and systematic random noise on unknown continuous stirred tank reactor (cstr)," *Applied Mathematical Sciences*, vol. 11, no. 62, pp. 3051–3071, 2017.
- [32] K. Lopez Buritica, S. Casanova Trujillo, C. D. Acosta, and H. A. Granada Diaz, "Dynamical Analysis of a Continuous Stirred-Tank Reactor with the Formation of Biofilms for Wastewater Treatment," *Mathematical Problems in Engineering*, vol. 2015, Jun. 2015, Art. no. e512404, <https://doi.org/10.1155/2015/512404>.
- [33] K. Cahyari, Sarto, S. Syamsiah, and A. Prasetya, "Performance of continuous stirred tank reactor (CSTR) on fermentative biohydrogen production from melon waste," *IOP Conference Series: Materials Science and Engineering*, vol. 162, no. 1, Nov. 2016, Art. no. 012013, <https://doi.org/10.1088/1757-899X/162/1/012013>.
- [34] K. J. Keesman et al., "Aquaponics Systems Modelling," in *Aquaponics Food Production Systems: Combined Aquaculture and Hydroponic Production Technologies for the Future*, S. Goddek, A. Joyce, B. Kotzen, and G. M. Burnell, Eds. Cambridge, England: Springer International Publishing, 2019, pp. 267–299.
- [35] L. P. Russo and B. W. Bequette, "Cstr performance limitations due to cooling jacket dynamics," in *Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes*, J. G. Balchen, Ed. Oxford, UK: Pergamon, 1993, pp. 149–154.
- [36] M. Danish, M. K. Al Mesfer, and M. M. Rashid, "Effect of operating conditions on cstr performance: an experimental study," *International journal of engineering research and applications*, vol. 5, no. 2, pp. 74–78, 2015.
- [37] M. Laine, *Adaptive MCMC methods with applications in environmental and geophysical models*. Helsinki, Finland: Finnish Meteorological Institute, 2008.
- [38] D. F. A. M. N. Esmaeel, "Fuzzy logic Control of Continuous Stirred Tank Reactor," *Tikrit Journal of Engineering Sciences*, vol. 20, no. 2, pp. 70–80, 2013.
- [39] M. A. S. Aboelela and R. H. M. Hennas, "Development of a fractional order pid controller using adaptive weighted pso and genetic algorithms with applications," in *Fractional Order Systems*, A. T. Azar, A. G. Radwan, and S. Vaidyanathan, Eds. Cambridge, Massachusetts, MA, USA: Academic Press, 2018, pp. 511–551.
- [40] M.-H. Cui, D. Cui, L. Gao, H.-Y. Cheng, and A.-J. Wang, "Efficient azo dye decolorization in a continuous stirred tank reactor (CSTR) with built-in bioelectrochemical system," *Bioresource Technology*, vol. 218, pp. 1307–1311, Oct. 2016, <https://doi.org/10.1016/j.biortech.2016.07.135>.
- [41] N. M. Ramli and M. S. Mohamad, "Modelling for Temperature Non-Isothermal Continuous Stirred Tank Reactor Using Fuzzy Logic," *International Journal of Environmental and Ecological Engineering*, vol. 11, no. 2, pp. 169–175, Jan. 2017.
- [42] R. B. Bird, W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena*, 2nd edition. New York, NY, USA: Wiley, 2009.
- [43] R. Kandiyoti, *Fundamentals of Reaction Engineering*. London, UK: Bookboon, 2009.
- [44] R. C. S. Dias and M. R. P. F. N. Costa, "Transient Behavior and Relation of Free Radical Polymerizations in Continuous Stirred Tank Reactors," *Macromolecular Theory and Simulations*, vol. 14, no. 4, pp. 243–255, 2005, <https://doi.org/10.1002/mats.200400086>.
- [45] R. M. Sudhanan and D. P. Poongodi, "Comparative Analysis of Various Control Strategies for a Nonlinear CSTR System," *International Journal of Nonlinear Sciences and Numerical Simulation*, vol. 18, no. 3–4, pp. 269–276, Jun. 2017, <https://doi.org/10.1515/ijnsns-2015-0125>.
- [46] S. Brooks, "Markov chain Monte Carlo method and its application," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 47, no. 1, pp. 69–100, 1998, <https://doi.org/10.1111/1467-9884.00117>.
- [47] S. S. Jang and R. B. Gopaluni, "Parameter estimation in nonlinear chemical and biological processes with unmeasured variables from small data sets," *Chemical Engineering Science*, vol. 66, no. 12, pp. 2774–2787, Jun. 2011, <https://doi.org/10.1016/j.ces.2011.03.029>.
- [48] S. Masoumi, T. A. Duever, and P. M. Reilly, "Sequential Markov Chain Monte Carlo (MCMC) model discrimination," *The Canadian Journal of Chemical Engineering*, vol. 91, no. 5, pp. 862–869, 2013, <https://doi.org/10.1002/cjce.21711>.
- [49] S. Nanda, "Reactors and Fundamentals of Reactors Design for Chemical Reaction," Ph.D. dissertation, Maharshi Dayanand University, Haryana, India, 2008.
- [50] S. Sharma, "Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy," *Annual Review of Astronomy and Astrophysics*, vol. 55, pp. 213–259, Aug. 2017, <https://doi.org/10.1146/annurev-astro-082214-122339>.
- [51] S. Sinharay, "Assessing Convergence of the Markov Chain Monte Carlo Algorithms: A Review," *ETS Research Report Series*, vol. 2003, no. 1, pp. i–52, 2003, <https://doi.org/10.1002/j.2333-8504.2003.tb01899.x>.
- [52] S. Tronci, M. Grosso, J. Alvarez, and R. Baratti, "Stochastic dynamical nonlinear behavior analysis of a class of single-state CSTRs," *IFAC Proceedings Volumes*, vol. 42, no. 11, pp. 697–702, Jan. 2009, <https://doi.org/10.3182/20090712-4-TR-2008.00113>.

- [53] S. Zhang, D. Muller, H. Arellano-Garcia, and G. Wozny, "CFD simulation of the fluid hydrodynamics in a continuous stirred-tank reactor," *Chemical Engineering Transactions*, vol. 32, pp. 1441–1446, Jun. 2013.
- [54] S. N. Naikwad and S. V. Dudul, "Identification of a Typical CSTR Using Optimal Focused Time Lagged Recurrent Neural Network Model with Gamma Memory Filter," *Applied Computational Intelligence and Soft Computing*, vol. 2009, Jan. 2010, Art. no. 385757, <https://doi.org/10.1155/2009/385757>.
- [55] S. R. Tofighi, F. Bayat, and F. Merrikh-Bayat, "Robust feedback linearization of an isothermal continuous stirred tank reactor: H_∞ mixed-sensitivity synthesis and DK-iteration approaches," *Transactions of the Institute of Measurement and Control*, vol. 39, no. 3, pp. 344–351, Mar. 2017, <https://doi.org/10.1177/0142331215603446>.
- [56] T. Niederberger, "Markov Chain Monte Carlo Methods for Parameter Identification in Systems Biology Models," Ph.D. dissertation, Ludwig Maximilian University of Munich, Bad Reichenhall, Germany, 2012.
- [57] T. Rajagopalan and V. Seshadri, "Analysis of continuous stirred tank reactor as a multivariable process and algorithms for computer determination of the equilibrium states," *International Journal of Control*, vol. 15, no. 3, pp. 497–507, Mar. 1972, <https://doi.org/10.1080/00207177208932165>.
- [58] V. Nicoulaud-Gouin, L. Garcia-Sanchez, M. Giacalone, J. C. Attard, A. Martin-Garin, and F. Y. Bois, "Identifiability of sorption parameters in stirred flow-through reactor experiments and their identification with a Bayesian approach," *Journal of Environmental Radioactivity*, vol. 162–163, pp. 328–339, Oct. 2016, <https://doi.org/10.1016/j.jenvrad.2016.06.008>.
- [59] V. Roy, "Convergence Diagnostics for Markov Chain Monte Carlo," *Annual Review of Statistics and Its Application*, vol. 7, no. 1, pp. 387–412, 2020, <https://doi.org/10.1146/annurev-statistics-031219-041300>.
- [60] Y. Lu, Z. Fang, and C. Gao, "Stabilization of (state, input)-disturbed CSTRs through the port-Hamiltonian systems approach," *arXiv:1707.01560 [math]*, Jun. 2017, Accessed: Feb. 20, 2021. [Online]. Available: <http://arxiv.org/abs/1707.01560>.
- [61] Z. Prokopova and R. Prokop, "Modelling and Simulation of Chemical Industrial Reactors," in *23rd European Conference on Modelling and Simulation*, Madrid, Spain, Jun. 2009, pp. 378–383.

Evaluation Framework of the Deficit and Reliability Quality Measures of the Transmission System

Awad Saud Alshammari
Department of Electrical Engineering
College of Engineering
University of Ha'il
sa.awwad.0923@outlook.sa

Badr M. Alshammari
Department of Electrical Engineering
College of Engineering
University of Ha'il
bms.alshammari@uoh.edu.sa

Tawfik Guesmi
Department of Electrical Engineering
College of Engineering
University of Ha'il
t.guesmi@uoh.edu.sa

Rabeh Abbassi
Department of Electrical Engineering
College of Engineering
University of Ha'il
r.abbassi@uoh.edu.sa

Abstract-Power system planning faces various issues related to reliability and quality evaluation. The power system network planning is by nature a complex, huge-scale, and mixed-objective optimization problem, especially when concerning its non-linear behavior and the requirements of future unknown loads. In this regard, the electric power utilities attempt to maintain a balance between the generation energy, the transmission capacity, and the needed demand. The main purpose of the current paper is to utilize modern modeling techniques and computational procedures, including the advanced deficit transmission system evaluation method and sparse-matrix network analysis algorithms, in order to evaluate, with sufficient accuracy, the deficit and reliability levels in practical real-life large-scale power systems. The new evaluation methodology is based on three quantities representing the relationship between the generation push in the grid, the maximum limitation of the transmission capacity, and the needed load. The main contribution of the paper is assessing the deficit transmission system index with novel formulas.

Keywords-reliability evaluation; deficit index; transmission system; quality measures

I. INTRODUCTION

The main objective of electric power utilities is to maintain a continuous and sufficient power supply to customers at a reasonable cost [1]. Cost-effectiveness, security, adequacy, and reliability analyses of a power system have changed over the years to become a vital branch in today's highly competitive business environment of power utility planning and operations. Reliability and quality are considered two vital measures of power system planning and operational procedures. The engineers and utility technicians seek to design and synthesis a large-scale power system, they consider reliability as one of the key design factors [2-4]. Power system development and planning, which include an essential part of reliability and quality assessment, form a complex, large-scale, and multi-objective optimization problem [5, 6], which is related to

nonlinear characteristics, future demand uncertainty requirements, and system component availability. The issues of reliability and quality evaluation have seen a growing interest in power system reliability and quality assessment by power utilities and engineers. Several authors were involved in the research aiming at conducting reliability and quality assessment in an efficient and accurate manner and with as much realization of the practical circumstances of the power utility as possible. In this regard, the reliability of an electric supply system is defined as the probability of providing the customers with continuous and satisfactory service.

Generation system reliability indices have been used as a standard for measuring system reliability in [7]. A model was proposed based on a convolutional algorithm that predicts how these indices vary as system annual peak increases. The proposed model was used to calculate and test the same indices of the Benghazi North Power Plant (BNPP) before and after installing additional capacity. A power system by nature is complex, therefore there are multiple random events at different stages of power systems such as uncertainty in customer demand, intermittent outages of power generators and its related units, intermittent electricity production and its impact on the adequacy assessment [8]. Therefore, probabilistic methods have given detailed and more realistic information about random events and their negative impact on power system supply and demand. In this direction, an analytical model of the inverse power system reliability evaluation problem was proposed and formulated to find the Unknown Component Reliability Parameter (UCRP) [9]. In addition, an application to power system planning was tested to examine the way the component reliability parameters can be modified quantitatively in order to achieve the desired system reliability improvement.

The main purpose of this paper is to utilize modern modeling techniques and computational procedures, and focus on the deficit transmission system evaluation method and

Corresponding author: Badr M Alshammari

sparse-matrix network analysis algorithms, in order to evaluate, with sufficient accuracy, the deficit and reliability levels in practical real-life large-scale power systems. The new evaluation methodology is based on three quantitative parameters of the power network, which are representing the relationship between generation push in the grid, the maximum limitation of the transmission capacity, and the needed load. The main contribution of the current paper is the assessment of the deficit transmission system index with novel formulas.

II. PROBLEM FORMULATION

A. Power System Network Model

The novel framework applied in this paper is based on the original work of [10], in which 3 dimensions were introduced to represent the relationship between certain system generation capacity and the demand. These dimensions relate to the following demand fulfillment issues:

- The need of capacity for demand fulfillment.
- The existence of capacity (availability for demand fulfillment).
- The ability of capacity to reach the demand.

The first dimension defines whether or not the capacity is needed, the second one defines whether or not the capacity exists, and the last one defines whether or not the capacity can reach (delivered to) the demand. The 8 possible combinations associated with the 0/1 (Yes/No) values of the 3 dimensions, are illustrated in Table I. The generation and transmission quality indices are defined in terms of the previously defined 1/0 states indicating the (Needed, Exists, Can-reach) true/false values associated with each quality index. We shall use the symbol Q_{gijk} to indicate the generation quality index state. Also, in the following expressions, we shall use $\text{Min}\{x, y, \dots, z\}$ to indicate the minimum of x, y, \dots, z . The notation $\langle x \rangle$ will be used to denote $\text{Max}\{0, x\}$, i.e. is the maximum of x and zero ($= x$ if $x > 0$, or 0 otherwise). Table I summarizes the considered quality indices, namely the Utilized Generation Capacity ($Q111$), Bottled Generation Capacity ($Q110$), Shortfall Generation Capacity ($Q101$), Deficit Generation Capacity ($Q100$), Surplus Generation Capacity ($Q011$), Redundant Generation Capacity ($Q010$), Spared Generation Capacity ($Q001$) and Saved Generation Capacity ($Q000$).

TABLE I. ILLUSTRATION OF QUALITY ASSESEMENT INDICES

	Needed (L > 0)		Not needed (L = 0)	
	Can reach	Cannot reach	Can reach	Cannot reach
Exist (C > 0)	Utilized Q_g111	Bottled Q_g110	Surplus Q_g011	Redundant Q_g010
Not exist (C = 0)	Short-fall Q_g101	Deficient Q_g100	Spared Q_g001	Saved Q_g000

The evaluation of the above quality indices requires the knowledge of the following data types for the demand and various system facilities:

- The value of the demand required to be supplied.
- The value of the generation capacity and the maximum site

capacity (the limit of potential increase in existing generation capacity).

- The value of transmission capacity.

B. Linear Program Formulation

In the computational scheme of [10], the integrated system quality assessment is performed via solving a master linear programming problem in which a feasible power flow is established which minimizes the total system Load Not Served (LNS) subject to capacity limits and flow equations. The master linear program, which utilizes the network bus incidence matrix A , is formulated as:

$$\left. \begin{aligned} &\text{Objective function} = (\text{Minimize } f = \sum_{l=1}^{n_L} (-P_l)) \\ &\text{with respect to } P_L, P_G \text{ and } P_T \\ &\text{Subject to} \end{aligned} \right\} \quad (1)$$

$$A \cdot P_T = \begin{bmatrix} -P_L \\ P_G \end{bmatrix}$$

$$\begin{aligned} P_L &\leq \bar{P}_L, -P_L \leq 0 \\ P_G &\leq \bar{P}_G, -P_G \leq 0 \\ P_T &\leq \bar{P}_T, -P_T \leq \bar{P}_T \end{aligned}$$

The optimization software package CPLEX has been used to solve the Master Linear Program. The overall process of the evaluation of power systems reliability and quality measures is summarized in a flowchart shown in Figure 1.

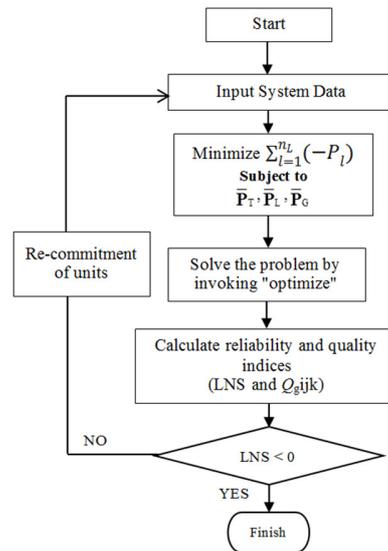


Fig. 1. Flow chart of the proposed methodology.

C. Application of Performance Quality Assessment

Consider the 3-bus sample power system in Figure 2, where a load of 130pu is supplied by two generators that have an available capacity of 170pu. The power system has three transmission lines having an available capacity of 60, 50, and 10pu respectively. For this simple system, the reliability and quality indices can be evaluated as shown in Table II. This paper concerns the indexes related with the deficit of the transmission system.

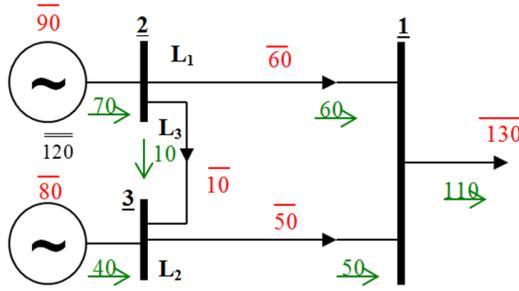


Fig. 2. The 3-bus sample power system.

TABLE II. RELIABILITY AND QUALITY INDICES FOR THE SAMPLE SYSTEM

Index	Q111	Q110	Q101	Q100	Q011	Q010	Q001	Q000
Value	110	0	0	20	0	0	0	35

D. Large-Scale Implementation

On [10], the formulation of 5 reliability and quality performance indices is considered, namely the Load Not-Served (LNS), Utilized Generation Capacity ($Qg111$), Bottled Generation Capacity ($Qg110$), Surplus Generation Capacity ($Qg011$), and Redundant Generation Capacity ($Qg010$) are presented while the current research work is concerned with the Deficit Transmission Capacity ($Qt100$) and new formulas are established as follows:

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g = 0]$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$

$$Qt100 = [\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] - [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t] \quad (2)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g \neq 0]$ and $\sum_{g=1}^{nG} \bar{P}_g \leq \sum_{t=1}^{nT} \bar{P}_t$ and $\sum_{l=1}^{nL} \bar{P}_l \leq \sum_{g=1}^{nG} \bar{P}_g$

$$Qt100 = \sum_{l=1}^{nL} \bar{P}_l - \sum_{t=1}^{nT} \bar{P}_t \quad (3)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g \neq 0]$ and $\sum_{g=1}^{nG} \bar{P}_g \leq \sum_{t=1}^{nT} \bar{P}_t$ and $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$

$$Qt100 = \sum_{g=1}^{nG} \bar{P}_g - \sum_{t=1}^{nT} \bar{P}_t \quad (4)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g \neq 0]$ and $\sum_{g=1}^{nG} \bar{P}_g > \sum_{t=1}^{nT} \bar{P}_t$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$

$$Qt100 = \sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g \quad (5)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g \neq 0]$ and $\sum_{g=1}^{nG} \bar{P}_g > \sum_{t=1}^{nT} \bar{P}_t$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$

$$Qt100 = \sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t \quad (6)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l > [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$ $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g = 0]$

$$Qt100 = [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l] - [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t] \quad (7)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l > [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$ $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g = 0]$ and $\sum_{t=1}^{nT} \bar{P}_t \leq \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] \leq [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$

$$Qt100 = \sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g \quad (8)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l > [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$ $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g = 0]$ and $\sum_{t=1}^{nT} \bar{P}_t \leq \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$

$$Qt100 = \sum_{l=1}^{nL} \bar{P}_l - \sum_{g=1}^{nG} \bar{P}_g \quad (9)$$

If: $\sum_{l=1}^{nL} \bar{P}_l > \sum_{g=1}^{nG} \bar{P}_g$ and $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g] > [\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l]$ and $\sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l > [\sum_{t=1}^{nT} \bar{P}_t - \sum_{t=1}^{nT} P_t]$ $[\sum_{g=1}^{nG} \bar{P}_g - \sum_{g=1}^{nG} P_g = 0]$ and $\sum_{t=1}^{nT} \bar{P}_t > \sum_{g=1}^{nG} \bar{P}_g$ and $\sum_{t=1}^{nT} \bar{P}_t \leq \sum_{l=1}^{nL} \bar{P}_l$

$$Qt100 = \sum_{l=1}^{nL} \bar{P}_l - \sum_{l=1}^{nL} P_l \quad (10)$$

where \bar{P}_l is the required value of the load, P_l is the actual value of the load, \bar{P}_g is the sit capacity of generation, \bar{P}_g is the available capacity of generation, P_g the actual capacity of generation, \bar{P}_t the capacity of the transmission lines, and P_t the actual flow in the transmission lines.

III. APPLICATION ON A REAL-LIFE NETWORK

The main purpose of the presented applications is to illustrate the implementation of the developed methodology on real-life systems and to demonstrate the applicability of the theoretical and computational developments of this work to practical power systems. The Wadi Aldawasir Network represents an isolated zone of the SEC (Saudi Electrical Company) system. Figure 3 shows the single-line diagram of the Wadi Aldawasir network. The system under investigation contains 1 generator, 9 branches (transmission lines and power transformers), and 7 loads. The results (Figure 4) show that the Deficit Transmission Capacity ($Qt100$) stays at zero value for all required load levels up to 198MW, where it starts to increase continuously to reach 80MW at a required load level of 270MW. Since the available generation at Wadi Aldawasir is 263MW, this situation indicates that the unsupplied load (between 180-263MW of load levels) is essentially due to transmission limitations rather than generation availability. $Qg111$ has a similar pattern with the $Qt100$, where it increases to reach 73MW at a required load level of 270MW. On the other hand, the $Qg111$ starts at 90MW and increases continuously to reach 180MW, when the required load level reaches 180MW, after which the $Qg111$ increases at a slower rate and saturates at about 190MW when the required load

level is 270MW. At this point, no more available generation capacity can be utilized. Figure 5 shows a 3-dimensional graph depicting the variation of the Q_{t100} with the required load and the available generation capacity levels of the Wadi Aldawasir network. It is noted from Figure 4 that the Q_{t100} stays at zero value for all available generation capacity levels between 50% and 150% of nominal as long as the required load level is

below 100% of nominal. This situation, however, changes after a region between 50% and 150% of nominal for the available generation capacity levels and between 100% and 150% of nominal for the required load levels, where more required load levels would increase the amount of the Q_{t100} . This situation indicates that the unsupplied load is now caused by the generation unavailability and transmission limitations.

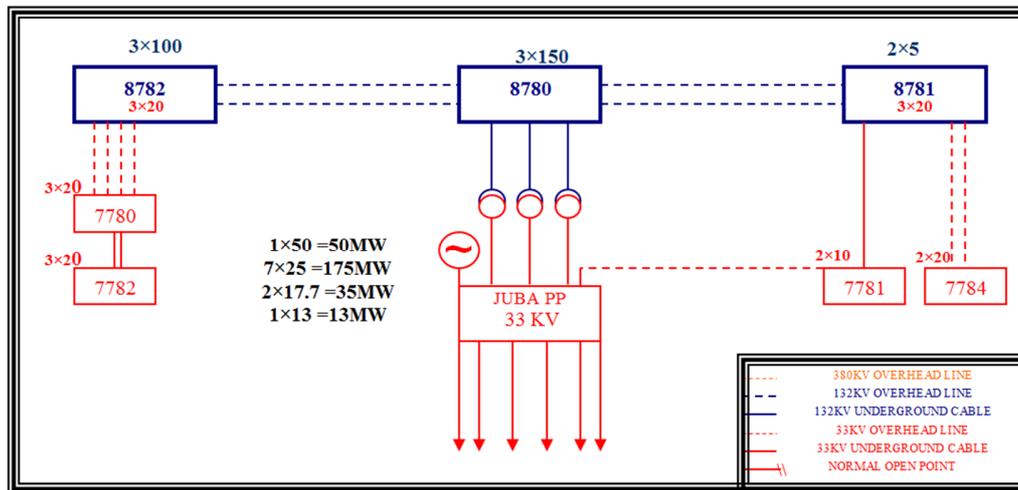


Fig. 3. Single-line diagram of the Wadi Aldawasir network.

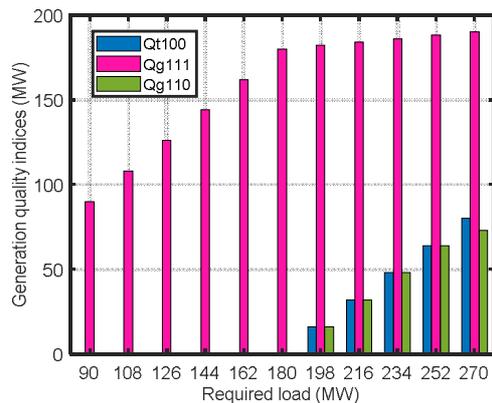


Fig. 4. Variation of generation capacity indices Q_{t100} , Q_{g111} , and Q_{g110} with the variation of required load levels of the Wadi Aldawasir network.

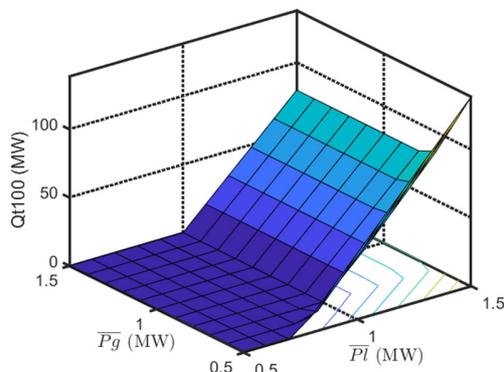


Fig. 5. 3-D graph of the variation of Q_{t100} with the required load levels and the available generation capacity levels of the Wadi Aldawasir network.

IV. CONCLUSION

In this paper, the evaluation of a new reliability and quality measures methodology is based on 3 quantities of the power network, which are representing the relationship between the generation push in the grid, the maximum limitation of the transmission capacity, and the needed load. At the same time, the maximum future expanded generation capacity that could be available at the same generation site, was discussed. The main purpose of the paper the utilization of modern modeling techniques and computational procedures, including the advanced deficit transmission system evaluation method and sparse-matrix network analysis algorithms, in order to evaluate with sufficient accuracy the deficit and reliability levels in practical real-life large-scale power systems. The main contribution of the paper is the assessment of the deficit transmission system index with novel formulas. The research work also includes a practical application to several operating scenarios in the Saudi electricity system. Most of the deficits on the power system are caused by generation unavailability and transmission limitations (contingency state).

ACKNOWLEDGMENT

The authors would like to thank the Deanship of the Scientific Research of the University of Ha'il, Saudi Arabia for funding and supporting this research project (# RG-191245).

REFERENCES

[1] S. de la Torre, A. J. Conejo, and J. Contreras, "Transmission Expansion Planning in Electricity Markets," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 238-248, Feb. 2008, <https://doi.org/10.1109/TPWRS.2007.913717>.

- [2] M. A. El-Kady, M. S. El-Sobki, and N. K. Sinha, "Reliability Evaluation for Optimally Operated, Large, Electric Power Systems," *IEEE Transactions on Reliability*, vol. 35, no. 1, pp. 41–47, Apr. 1986, <https://doi.org/10.1109/TR.1986.4335340>.
- [3] M. A. El-Kady, M. S. El-Sobki, and N. K. Sinha, "Loss of load probability evaluation based on real-time emergency dispatch," *Canadian Electrical Engineering Journal*, vol. 10, no. 2, pp. 57–60, Apr. 1985, <https://doi.org/10.1109/CEEJ.1985.6592019>.
- [4] Q. Chen, Y. Lin, and J. D. McCalley, "The Risk of High-Order Transmission Contingencies," in *2007 IEEE Power Engineering Society General Meeting*, Tampa, FL, USA, Jun. 2007, <https://doi.org/10.1109/PES.2007.385902>.
- [5] B. M. Alshammari, "Assessment of Reliability and Quality Performance using the Impact of Shortfall Generation Capacity Index on Power Systems," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4937–4941, Dec. 2019, <https://doi.org/10.48084/etasr.3190>.
- [6] B. M. Alshammari, T. Guesmi, M. Eleiwa, R. Abbassi, and M. Alghassab, "Optimal Technique of reliability and quality Assessment of two-area interconnected power systems," *Przegląd Elektrotechniczny*, vol. 96, no. 9, pp. 167–170, 2020, <https://doi.org/10.15199/48.2020.09.34>.
- [7] B. M. Alshammari, "Probabilistic Evaluation of a Power System's Reliability and Quality Measures," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5570–5575, Apr. 2020, <https://doi.org/10.48084/etasr.3441>.
- [8] M. K. Sharma, P. Phonrattanasak, and N. Leeprechanon, "Improved bees algorithm for dynamic economic dispatch considering prohibited operating zones," in *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, Bangkok, Thailand, Nov. 2015, <https://doi.org/10.1109/ISGT-Asia.2015.7386972>.
- [9] B. Hu, K. Xie, and H. Tai, "Inverse Problem of Power System Reliability Evaluation: Analytical Model and Solution Method," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6569–6578, Nov. 2018, <https://doi.org/10.1109/TPWRS.2018.2839841>.
- [10] B. M. Alshammari, M. A. El-Kady, and Y. A. Al-Turki, "Power system performance quality indices," *European Transactions on Electrical Power*, vol. 21, no. 5, pp. 1704–1710, 2011, <https://doi.org/10.1002/etep.515>.

Design and Study of an Adaptive Fuzzy Logic-Based Controller for Wheeled Mobile Robots Implemented in the Leader-Follower Formation Approach

Halima Medjoubi
Department of Electronics
University of Batna
Batna, Algeria
halimaautomatique@hotmail.com

Abdessemad Yassine
Department of Electronics
University of Batna
Batna, Algeria
yabdes@yahoo.fr

Hassam Abdelouahab
Department of Electronics
University of Setif
Setif, Algeria
abdelhassam@univ-setif.dz

Abstract-This paper presents a new design of an adaptive fuzzy logic control by implementing the leader-follower approach. The principle is to modify the feedback control of non-holonomic Wheeled Mobile Robot (WMR) to be adaptive according to a fuzzy controller in the control loop, in order to adjust the feedback control gains according to the distance error between the reference path and the real position. The trajectory tracking control for a single WMR is extended to the formation control for two WMRs in which the first one is the leader and the second is the follower. Simulation results are presented to demonstrate the effectiveness of the proposed controller.

Keywords-adaptive fuzzy control; feedback control; Wheeled Mobile Robot (WMR); leader-follower formation; trajectory tracking control; non-holonomic

I. INTRODUCTION

Wheeled mobile robots are very useful devices that perform a wide range of tasks [1-3]. For many years, single mobile robots have been used to execute simple tasks and to achieve desired goals, such as tracking a desired path or avoiding obstacles [4-6]. Recently, robotics research has been focused its interest to multi-robot systems. Certainly, this kind of systems could accomplish several complex tasks more economically and efficiently than a single-robot system [7, 8]. A team of mobile robots that performs a complex mission in short time, is necessary to work coordinately and cooperatively with respect to a specific formation. Various formation control approaches have been mentioned in the literature, such as the behavior-based approach [9], virtual structure approach [10], and leader-follower approach [11-13]. In the behavior-based approach, different desired behaviors (e.g. formation keeping, target seeking, etc.) are specified for each robot, where its action is depended on its role in the team. Since this approach is

complicated, exact formation control is hard to ensure [14-16]. On the other hand, in the virtual structure approach [7, 9], the entire formation is considered as a single rigid body, which means that all robots work as a single complex robot. This method can control the motion of the whole robot formation, but cannot change the formation when the environment is changed, which limits the scope of its application. The leader – follower method is also called as master-slave mode [15]. In the group shaped by multi-mobile robots, one robot is designed as the leader while the others are the followers. In this approach, the leader moves along a predefined trajectory while the followers are controlled to maintain a desired distance and bearing angle [16]. In robotics field, this is the most used technic, which allows the formation control to be transformed into two simple problems, a trajectory tracking by the leader and a control to keep the formation by the followers. Here, we must ensure permanent communication between the leader and all the followers in order to maintain the formation [4, 5, 7].

This paper deals with the leader-follower formations of non-holonomic mobile robots. A new, adaptive fuzzy controller is proposed to ensure tracking and formation control for two mobile robots. The modeling and control of two mobile robots where one is modeled as a leader and the second one as a follower are presented. The obtained simulation results for tracking and formation control are discussed and a comparison with conventional methods is made.

II. MULTI MOBILE ROBOT MODELING AND CONTROL

Usually, a mobile robot is presented in a Cartesian 2D frame with its coordinates, such as position coordinates (x, y) and its orientation θ [17]. In this study, a kinematic model is considered, taking into account linear and angular velocities for

both leader and follower mobile robots, and a controller for each one is designed.

A. Leader-Follower Formation Modeling

In order to model the kinematics of the leader robot, we use the configuration $P_l = [x_l, y_l, \theta_l]^T$ as shown in Figure 1. The control input for the leader robot is the linear velocity v_l and the angular velocity ω_l .

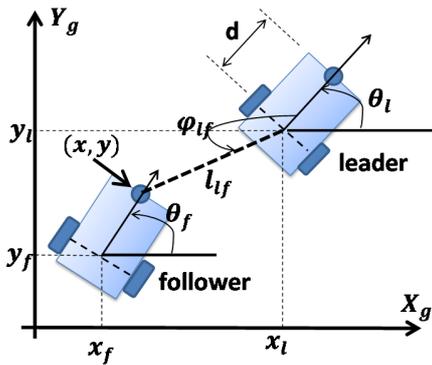


Fig. 1. Mobile robot in 2D frame.

The equations of the leader robot can be written as:

$$\begin{bmatrix} \dot{x}_l \\ \dot{y}_l \\ \dot{\theta}_l \end{bmatrix} = \begin{bmatrix} \cos \theta_l \\ \sin \theta_l \\ 0 \end{bmatrix} \cdot v_l + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \omega_l \quad (1)$$

The mobile robot wheels exhibit purely rolling motion and that assumed without slip. The nonholonomic constraint is expressed as:

$$-\dot{x}_l \cdot \sin \theta_l + \dot{y}_l \cdot \cos \theta_l = 0 \quad (2)$$

The leader robot model under Matlab/Simulink is shown in Figure 2.

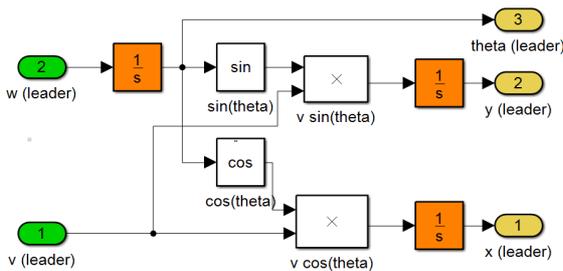


Fig. 2. Leader robot model.

The same constraint is observed for the follower robot. The follower robot is modeled relatively to the leader robot. It follows the leader while maintaining a desired relative distance and desired separation bearing angle with respect to the leader robot in a strategy known as separation bearing control (Figure 1), where:

φ_{lf} is the separation bearing angle between the leader and the follower robot.

l_{lf} denotes the separation distance between the center of the rear wheels of the leader, and the front castor of the follower robot.

(x, y) is the position coordinates of the front castor of the follower robot.

d is the distance between the front castor wheel and the mid point of the rear wheels for each robot.

(x_l, y_l) is the midpoint on the axis between the rear wheels.

$P_l = [x_l, y_l, \theta_l]^T$ is the leader robot position.

$u_l = [v_l, \omega_l]^T$: control inputs of leader robot.

$P_f = [x_f, y_f, \theta_f]^T$: follower robot position.

$u_f = [v_f, \omega_f]^T$: control inputs of follower robot.

The follower position is calculated as:

$$\begin{aligned} x_f &= x + d \cdot \cos \theta_f \\ y_f &= y + d \cdot \sin \theta_f \end{aligned} \quad (3)$$

The follower robot model is represented with the state vector $P_{lf} = [l_{lf}, \varphi_{lf}, \theta_f]^T$ expressed as:

$$\begin{aligned} l_{lf} &= \sqrt{(x_l - x_f - d \cdot \cos \theta_f)^2 + (y_l - y_f - d \cdot \sin \theta_f)^2} \\ \varphi_{lf} &= \pi - \arctan2(y_f + d \cdot \sin \theta_f - y_l, x_l - x_f - d \cdot \cos \theta_f) - \theta_l \end{aligned} \quad (4)$$

Finally, after differentiating the P_{lf} state vector we obtain the kinematic model for the follower mobile robot:

$$\begin{aligned} \dot{l}_{lf} &= v_f \cdot \cos \gamma - v_l \cdot \cos \varphi_{lf} + d \cdot \omega_f \cdot \sin \gamma \\ \dot{\varphi}_{lf} &= \frac{v_l \cdot \sin \varphi_{lf} - v_f \cdot \sin \gamma - \omega_l \cdot l_{lf} + d \cdot \omega_f \cdot \cos \gamma}{l_{lf}} \\ \dot{\theta}_f &= \omega_f \end{aligned} \quad (5)$$

with:

$$\gamma = \theta_l - \theta_f + \varphi_{lf} \quad (6)$$

Figure 3 illustrates the follower detailed model under Matlab/Simulink. The condition (7) is considered in order to avoid collision between the leader and the follower and must always be verified.

$$l_{lf} > 2 \cdot d \quad (7)$$

B. Control Strategy

The leader is controlled to follow a predefined path such as a circle or an eight shape. The control loop is presented in Figure 4. Considering a desired path, we can calculate the desired linear and angular velocities, and the kinematic controller is derived. In addition, the proposed adaptive controller is compare with a well-known kinematic controller presented in [18]. The kinematic controller explained with (8) is presented in detailed form in Matlab/Simulink in Figure 5.

$$\begin{bmatrix} u_{ref}^c \\ \omega_{ref}^c \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\frac{1}{d} \sin \theta & \frac{1}{d} \cos \theta \end{bmatrix} \begin{bmatrix} \dot{x}_d + l_x \cdot \tanh(\frac{k_x}{l_x} \tilde{x}) \\ \dot{y}_d + l_y \cdot \tanh(\frac{k_y}{l_y} \tilde{y}) \end{bmatrix} \quad (8)$$

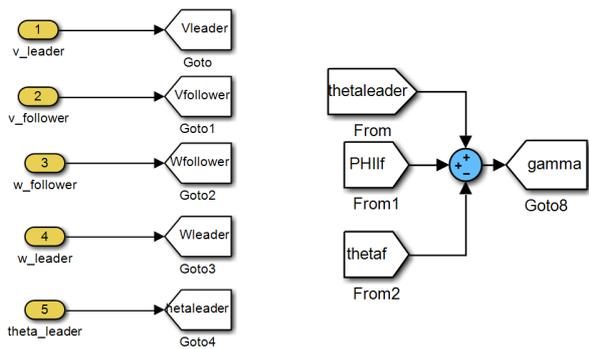


Fig. 3. Detailed model of the follower robot.

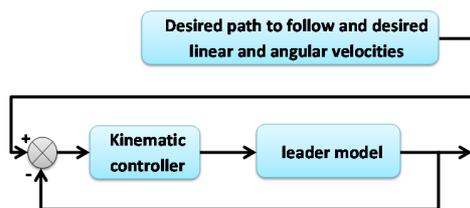


Fig. 4. Leader control loop.

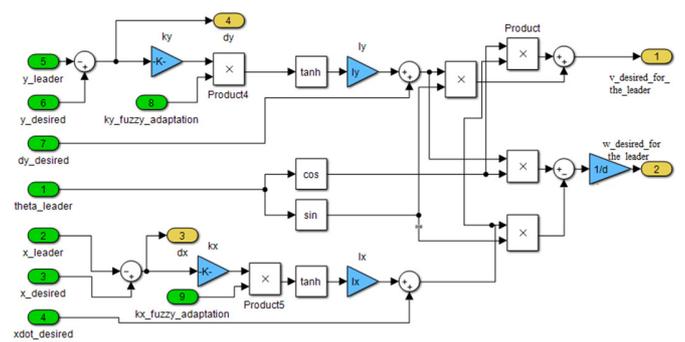


Fig. 5. Detailed kinematic controller of the leader.

In (8), $\tilde{x} = x_d - x_l$ and $\tilde{y} = y_d - y_l$ are the current position errors in the axes X and Y , respectively, $k_x > 0$ and $k_y > 0$ are the gains of the controller, $L_x \in \mathfrak{R}$ and $L_y \in \mathfrak{R}$ are saturation constants, and (x_l, y_l) and (x_d, y_d) are the current and the desired coordinates of the point of interest, respectively. The detailed model for the follower kinematic controller from [19] is given by (9) and is presented in Figure 6.

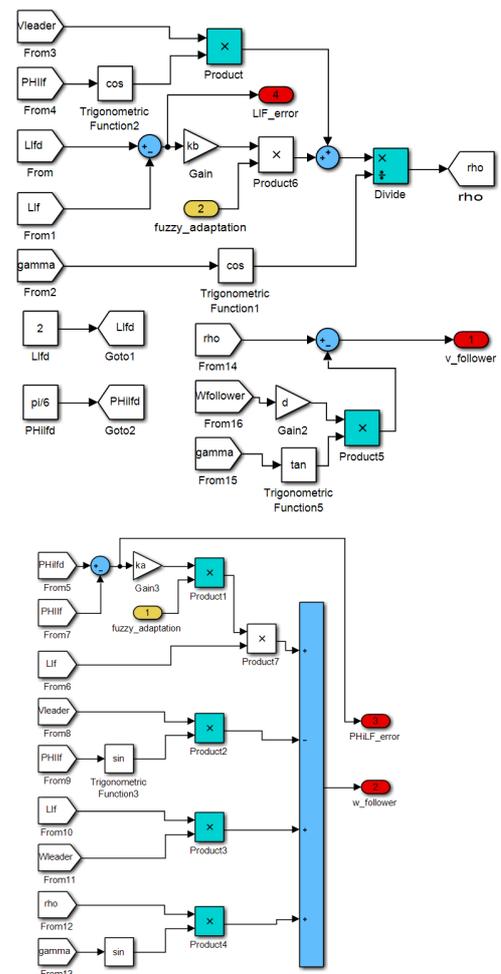


Fig. 6. Follower kinematic controller in Matlab/Simlink.

$$v_f = \rho - d \cdot \omega_f \cdot \tan \gamma$$

$$\omega_f = \frac{\cos \gamma}{d} \cdot \{k_a l_{lf} (\varphi_{lf}^d - \varphi_{lf}) - v_l \cdot \sin \varphi_{lf} + \beta\} \quad (9)$$

where:

$$\beta = l_{lf} \omega_l + \rho \cdot \sin \gamma$$

$$\rho = \frac{k_b (l_{lf}^d - l_{lf}) + v_l \cos \varphi_{lf}}{\cos \gamma} \quad (10)$$

$$\gamma = \varphi_{lf} + \theta_l - \theta_f$$

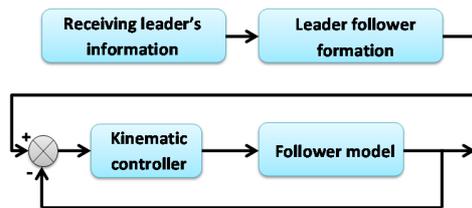


Fig. 7. Follower control loop.

It is noticed that the kinematic gain controllers $k_x, k_y, k_a,$ and k_b are constant during the mobile robot motion. So if one of the operation conditions changes, the actual gain value may no be adequate and may need to be adjusted. The main idea is to adjust the gain controller values in real time during the mobile robot operation.

III. ADAPTIVE FUZZY CONTROL FOR A MULTI-ROBOT SYSTEM

A. Fuzzy Systems and Fuzzy Control

Fuzzy logic provides human reasoning capabilities to capture uncertainties, which cannot be described by precise mathematical models. Traditionally, fuzzy logic controllers have been applied to the development of a complete navigation problem of a mobile robot [3]. Fuzzy logic is widely used to control various types of systems and also to develop adaptive controllers [20, 21]. The overall structure of a fuzzy controller is shown in Figure 8.

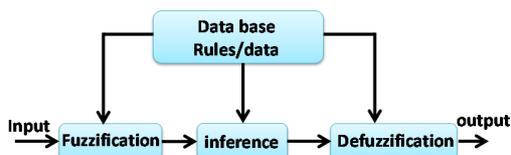


Fig. 8. Internal structure of a fuzzy controller.

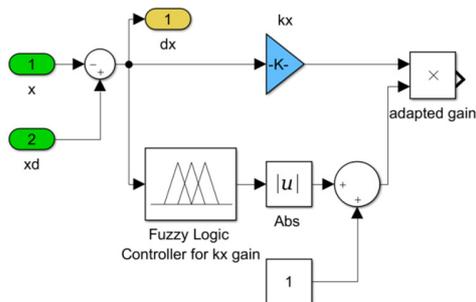


Fig. 9. Proposed adaptive controller.

In this paper, the proposed fuzzy controller will be added to the main controller where it can adjust controller gains according to the distance error as presented in Figure 9. The same structure is used for all controller gains (leader and follower).

B. Adaptive Controller Parameter Calculation

Both leader and follower robots are controlled by a kinematic controller. The controller gains are unique in value in order to give best results in a specific system status. In practice, the system statuses are not the same, and we come to a point for what adaptive gain controllers are necessary to satisfy several statuses. The idea here is to make changes in gain controllers in real time when the distance error between the desired path and actual robot position is different from zero. The larger distance error, the higher adaptation value will be applied to the gain controller value. In Figure 9, we notice that if the distance error is equal to zero then we don't need the fuzzy logic adaptation, which means that if it's output is zero we must keep the original gain controller, therefore we add a constant to the proposed controller loop to avoid zero multiplication when the output of fuzzy controller is equal to zero. The fuzzy logic controller was chosen to be Mamdani inference and the membership functions for input and output are presented in Figure 10. The input is the distance error and the output is the adaptation gain.

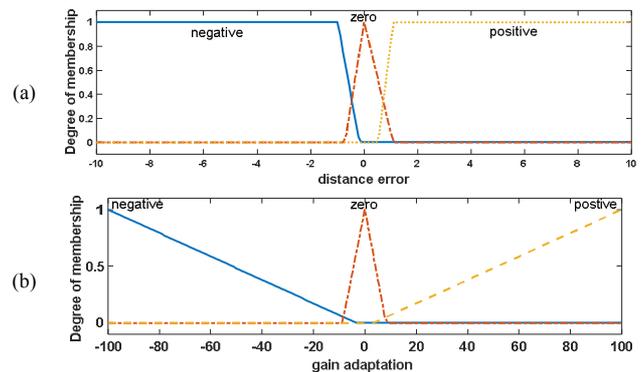


Fig. 10. Fuzzy controller: (a) input, (b) output.

IV. RESULTS AND DISCUSSION

To show the performance of the proposed controller, let's consider a group of two mobile robots in which one is the leader and one is the follower. The adaptation loop is inserted into the main control loop [19] for both. Circular and eight-shaped paths are used to test the proposed controller. The simulations were carried out in Matlab/Simulink. All robots and controller equations were modeled in Matlab/Simulink blocs. The proposed fuzzy adaptive controller was implemented with the Fuzzy Logic Toolbox of Simulink. The simulation parameters were:

- Mobile robot parameter: $d=0.2m$
- Radius of desired circular path: $r_{circle}^{desired}=5m$
- Radius of desired circular path : $r_{eight}^{desired}=10m$

Leader-follower formation parameters:

$$l_{lf}^{desired} = 2 \text{ m}, \varphi_{lf}^{desired} = \frac{\pi}{6} \text{ rad.}$$

The desired path generator modeled in Matlab/Simulink is shown in Figure 11.

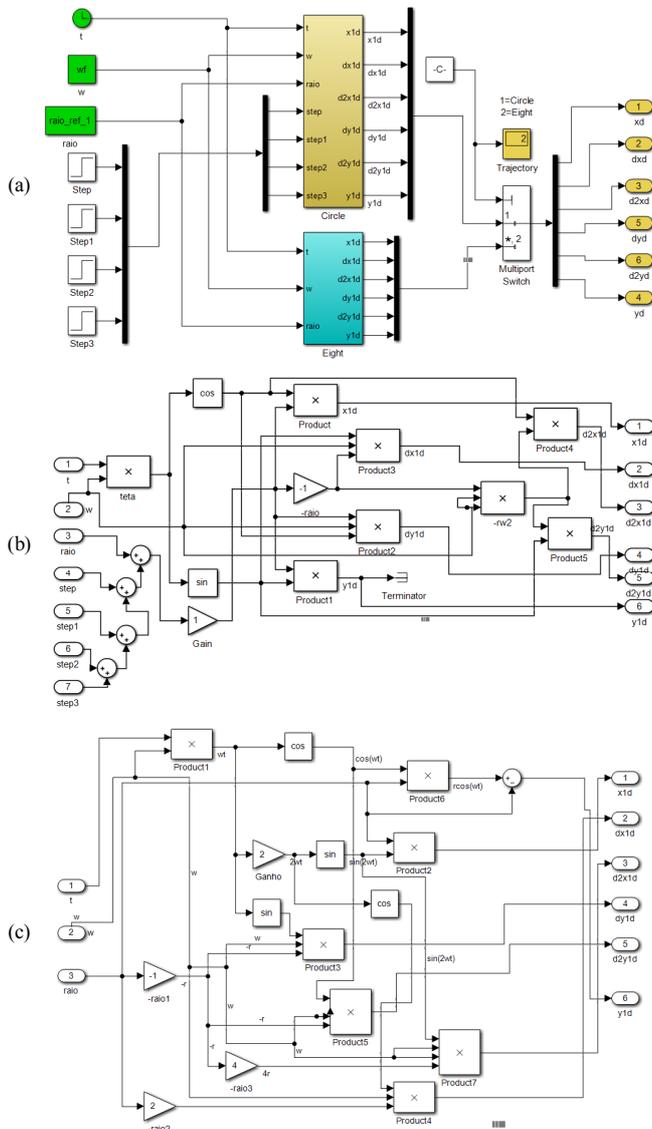


Fig. 11. Detailed desired path generation model. (a) general structure, (b) the eight generator model, and (c) circle generator model.

A comparison study was carried out between the kinematic controller ([18] for the leader and [19] for the follower) and the proposed adaptive kinematic fuzzy controller, considering paths and distance error. The results of the leader with and without adaptive controller are displayed in Figure 12 which shows the results of following a trajectory with circular path (Figure 12(a)), the error distance (Figure 12(b)), and the gain adaptation for x and y gain kinematic controller (Figure 12(c) and 12(d) respectively).

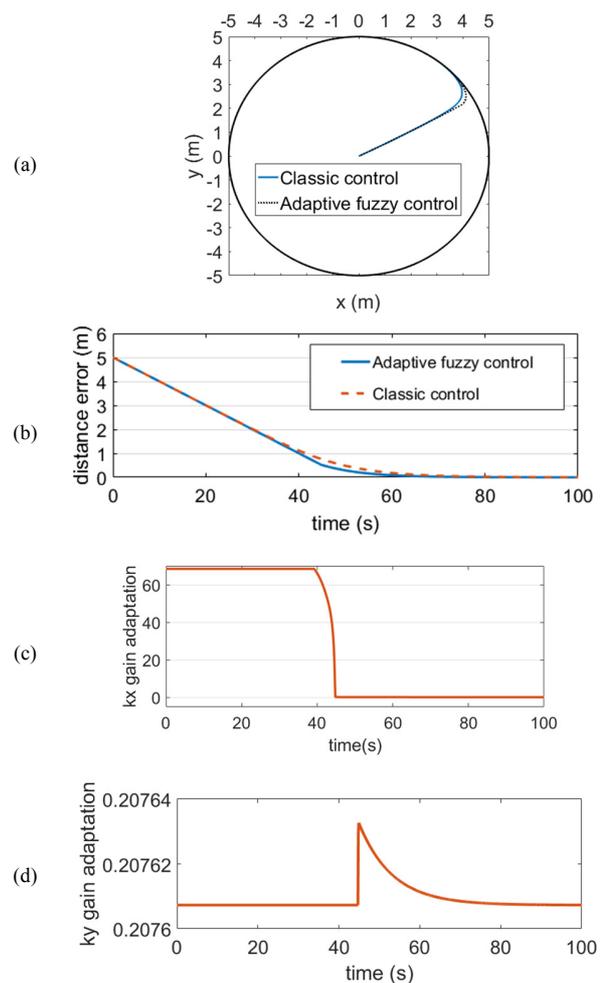


Fig. 12. Results of leader robot for circular path

It is noticed that the adaptation makes results better with a fast time response. The gain controllers are adapted to the distance error. The same remarks are observed for the eight-shaped path in Figure 13(a)-(c). The adaptation ensures that the system will not go out of the desired path once it reaches it because the gain controller will be decreased to a lower value while approaching the desired path. From Figure 13(b) and 13(c), we can see that at the beginning, the gain controllers are at high values because the robot is far from the desired path and decrease as the robot comes close to it. This technique will ensure that the system will not lose stability once it reaches its desired path.

As the next step, we give the results of the formation control for leader and follower with kinematic controller and our adaptive fuzzy kinematic controller. Similarly to the leader result, the two paths are tested. Figure 14(a)-(c) represent respectively the circular path, the distance between the leader and the follower, and the bearing angle between the leader and the follower. Gain controller adaptation (for l_{lf} and φ_{lf}) is displayed in Figure 15.

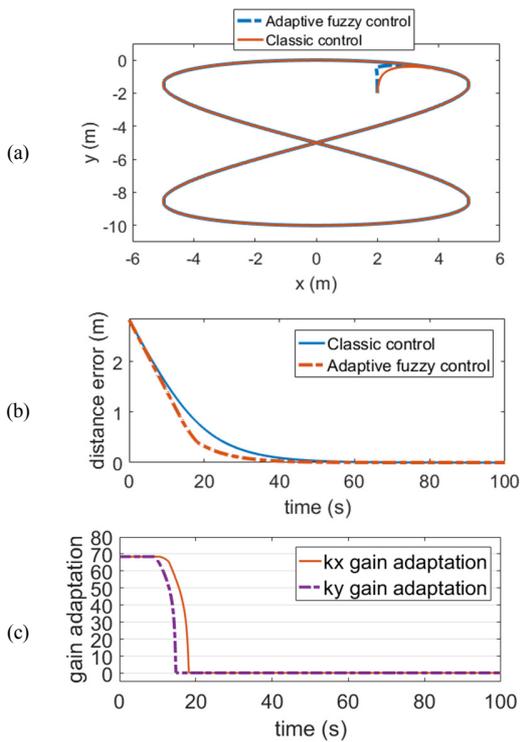


Fig. 13. Results of leader robot for the eight-shape path.

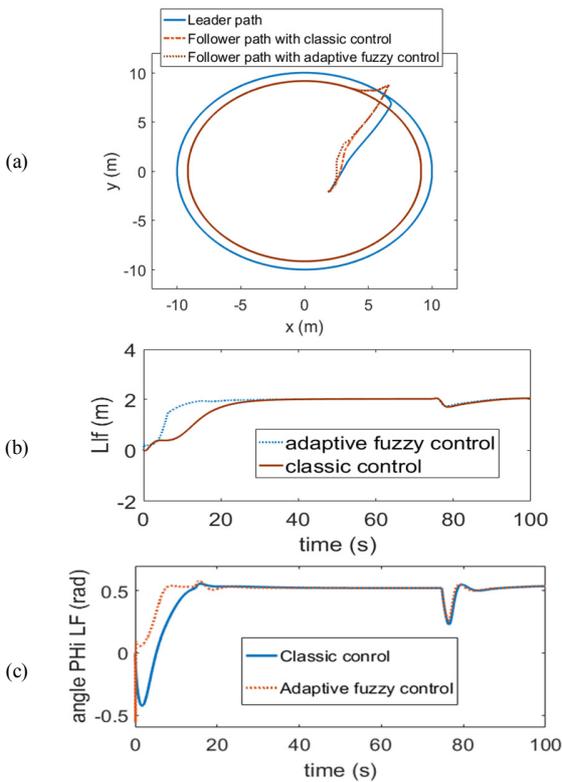


Fig. 14. Results of leader-follower for the circular path.

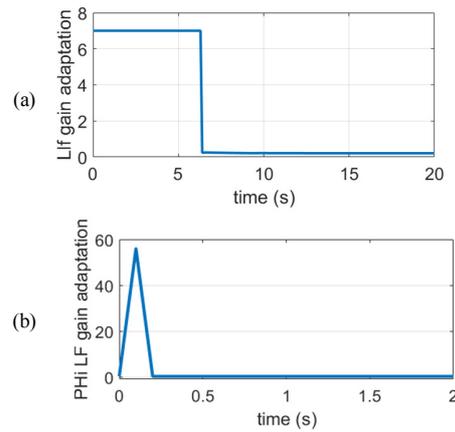


Fig. 15. Gain adaptation of leader-follower for the circular path.

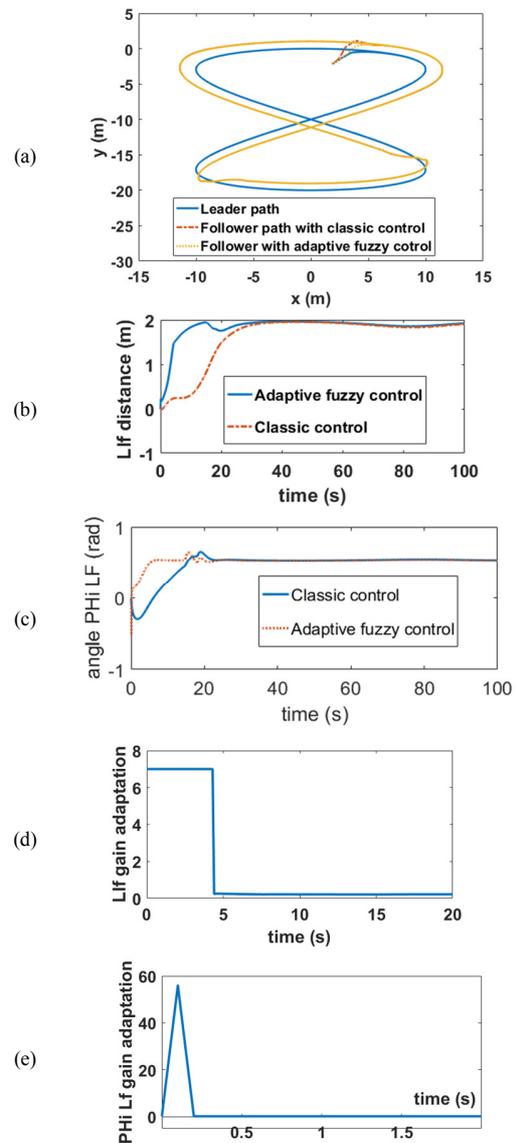


Fig. 16. Results of leader-follower for the eight-shaped path.

It is clear from the results shown in Figure 15 that the adaptive controller performs better than the kinematic one. The follower reached quickly the desired distance and bearing angle with the leader. The same observations are seen in Figure 16. The obtained results show an improvement with the proposed adaptive kinematic fuzzy-based controller compared to the ones in [18, 19]. Comparing the results, we can see that the system remains stable in our tests due to gains adaptation which will avoid excess value in gain controllers especially when the robot reaches its desired path. The time response will be faster and no oscillation can be seen after reaching the desired path.

V. CONCLUSION

An adaptive fuzzy kinematic controller has been proposed in order to solve the tracking control and leader-follower formation problems. An auxiliary fuzzy controller was inserted into the main kinematic control loop to make the adaptive controller gains according to distance and bearing angle errors. Simulations were carried out using two paths (circular and eight shaped) to test the effectiveness of the proposed controller. The obtained results showed that the proposed controller was better when compared with the kinematic controller in terms of time response for both tracking and formation control. For future work and to enlarge the present work, dynamics and obstacles avoidance will be considered.

APPENDIX

The follower position is calculated geometrically according to the leader position using the model presented in Figure 17.

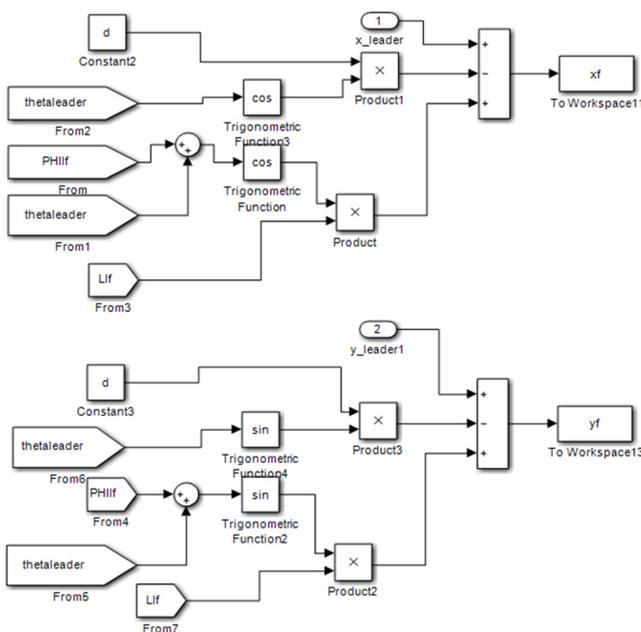


Fig. 17. Follower position calculation model.

REFERENCES

- [1] S. Sang, H. Wu, J. Zhao, and Q. An, "Fuzzy Logic Control for Wheeled Mobile Robots," in *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, Tianjin, China, Aug. 2009, vol. 6, pp. 237–241, <https://doi.org/10.1109/FSKD.2009.241>.

- [2] R. B. Meshram, "Motion Control of Wheeled Mobile Robots Using Fuzzy Logic," *International Journal of Recent Technology and Engineering*, vol. 2, no. 3, pp. 89–93, 2013.
- [3] T. Das and I. N. Kar, "Design and implementation of an adaptive fuzzy logic-based controller for wheeled mobile robots," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 3, pp. 501–510, May 2006, <https://doi.org/10.1109/TCST.2006.872536>.
- [4] T. T. Mac, C. Copot, R. D. Keyser, T. D. Tran, and T. Vu, "MIMO Fuzzy Control for Autonomous Mobile Robot," *Journal of Automation and Control Engineering*, vol. 3, no. 6, pp. 65–70, 2015, <https://doi.org/10.12720/joace.4.1.65-70>.
- [5] H. Omrane, M. S. Masmoudi, and M. Masmoudi, "Fuzzy Logic Based Control for Autonomous Mobile Robot Navigation," *Computational Intelligence and Neuroscience*, vol. 2016, Sep. 2016, Art. no. 9548482, <https://doi.org/10.1155/2016/9548482>.
- [6] A. Chatraei and H. Javidian, "Formation control of mobile robots with obstacle avoidance using fuzzy artificial potential field," in *2015 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*, Liberec, Czech Republic, Jun. 2015, pp. 1–6, <https://doi.org/10.1109/ECMSM.2015.7208710>.
- [7] M. A. Molina-Villa, D. R. Avendaño-Florez, L. E. Solaque-Guzman, and N. F. Velasco-Toledo, "Fuzzy logic controller for cooperative mobile robotics implemented in leader-follower formation approach," *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 76, 2015, <https://doi.org/10.17533/udea.redin.n76a03>.
- [8] Kim C. Ng and M. M. Trivedi, "Multirobot convoying using neuro-fuzzy control," in *Proceedings of 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996, vol. 4, pp. 417–421, <https://doi.org/10.1109/ICPR.1996.547600>.
- [9] T. Balch and R. C. Arkin, "Behavior-based formation control for multirobot teams," *IEEE Transactions on Robotics and Automation*, vol. 4, no. 6, pp. 926–939, Dec. 1998, <https://doi.org/10.1109/70.736776>.
- [10] Kar-Han Tan and M. A. Lewis, "Virtual structures for high-precision cooperative mobile robotic control," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS '96*, Osaka, Japan, Nov. 1996, vol. 1, pp. 132–139, <https://doi.org/10.1109/IROS.1996.570643>.
- [11] J.-L. Wang and H.-N. Wu, "Leader-following formation control of multi-agent systems under fixed and switching topologies," *International Journal of Control*, vol. 85, no. 6, pp. 695–705, Jun. 2012, <https://doi.org/10.1080/00207179.2012.662720>.
- [12] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "A Geometric Characterization of Leader-Follower Formation Control," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, Rome, Italy, Apr. 2007, pp. 2397–2402, <https://doi.org/10.1109/ROBOT.2007.363678>.
- [13] T. Gustavi and X. Hu, "Observer-Based Leader-Following Formation Control Using Onboard Sensor Information," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1457–1462, Dec. 2008, <https://doi.org/10.1109/TRO.2008.2006244>.
- [14] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Leader-follower formation control of nonholonomic mobile robots with input constraints," *Automatica*, vol. 44, no. 5, pp. 1343–1349, May 2008, <https://doi.org/10.1016/j.automatica.2007.09.019>.
- [15] Z. Ying and L. Xu, "Leader-follower formation control and obstacle avoidance of multi-robot based on artificial potential field," in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, Qingdao, China, May 2015, pp. 4355–4360, <https://doi.org/10.1109/CCDC.2015.7162695>.
- [16] Y. Dai and S.-G. Lee, "The leader-follower formation control of nonholonomic mobile robots," *International Journal of Control, Automation and Systems*, vol. 10, no. 2, pp. 350–361, Apr. 2012, <https://doi.org/10.1007/s12555-012-0215-x>.
- [17] M. Boubezoula, A. Hassam, and O. Boutalbi, "Robust-flatness Controller Design for a Differentially Driven Wheeled Mobile Robot,"

- International Journal of Control, Automation and Systems*, vol. 16, no. 4, pp. 1895–1904, Aug. 2018, <https://doi.org/10.1007/s12555-017-0408-4>.
- [18] F. N. Martins, W. C. Celeste, R. Carelli, M. Sarcinelli-Filho, and T. F. Bastos-Filho, "An adaptive dynamic controller for autonomous mobile robot trajectory tracking," *Control Engineering Practice*, vol. 16, no. 11, pp. 1354–1363, Nov. 2008, <https://doi.org/10.1016/j.conengprac.2008.03.004>.
- [19] S. Ahmed and M. N. K. and R. N. K. Loh, "Control Analysis and Feedback Techniques for Multi Agent Robots," *Multiagent Systems*, Jan. 2009, <https://doi.org/10.5772/6597>.
- [20] V. H. Nguyen, H. Nguyen, M. T. Cao, and K. H. Le, "Performance Comparison between PSO and GA in Improving Dynamic Voltage Stability in ANFIS Controllers for STATCOM," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4863–4869, Dec. 2019, <https://doi.org/10.48084/etasr.3032>.
- [21] M. Fouzia, N. Khenfer, and N. E. Boukezzoula, "Robust Adaptive Tracking Control of Manipulator Arms with Fuzzy Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6131–6141, Aug. 2020, <https://doi.org/10.48084/etasr.3648>.

Investigation of Swirl Stabilized CH₄ Air Flame with Varied Hydrogen Content by using Computational Fluid Dynamics (CFD) to Study the Temperature Field and Flame Shape

A. Bouziane

Laboratory of Materials and Reactive Systems (LMSR)
Faculty of Technology
University of Djillali Liabes
Sidi Bel Abbes, Algeria
abouziane@yahoo.fr

A. Alami

Renewable Energies Department
Saad Dahlab Blida University 1
Blida, Algeria
alami.ahmed@univ-blida.dz

M. Zaitri

Laboratory of Gas Combustion and Environment
University of Sciences and Technology of Oran
Oran, Algeria
zaitri_mohammed@yahoo.fr

B. Bouchame

Laboratory of Mechanics of Structures and Solids (LMSS)
Faculty of Technology
University of Djillali Liabes
Sidi Bel Abbes, Algeria
boucham1@yahoo.com

M. Bouchetara

Laboratory of Gas Combustion and Environment
University of Sciences and Technology of Oran
Oran, Algeria
mbouchetara@hotmail.com

Abstract—In the current paper, numerical simulations of the combustion of turbulent CH₄-H₂ are presented employing the standard k-epsilon and the RNG k-epsilon for turbulence closure. The Fr-ED concept is carried out to account for chemistry/turbulence interaction. The hydrogen content is varied in the fuel stream from 0% to 100%. The numerical solutions are validated by comparison with corresponding experimental data from the Combustion Laboratory of the University of Milan. The flow is directed radially outward. This method of fuel injection has been already been explored experimentally. The results show that the structure of the flame is described reasonably and both standard k-ε and RNG k-ε models can predict the flame shape. The general aspect of the temperature profiles is well predicted. The temperature profiles are indicating a different trend between CH₄ and CH₄/H₂ fuel mixtures.

Keywords—RNG k-epsilon; swirl; hydrogen; CFD; CH₄

I. INTRODUCTION

Lean combustion emerges as a potential and an advantageous concept to attain efficiently the combustion process and control NO_x emissions in power-generation gas turbines [1–3]. The best approach to obtain a stable lean flame

is to add a highly reactive fuel to a traditional hydrocarbon fuel. Due to its high chemical reactivity, hydrogen could be used as a useful fuel addition to enhance the combustion of natural gas. The use of hydrogen in blended fuel can be a good option regarding combustion characteristics, such as extending the lean flammability limits [4, 5] and the environmental advantages by decreasing the emissions [6]. The research on the combustion of NG–H₂ fuel mixtures is open and growing. Studies [5, 7–9] on NG–H₂ hybrid flame stability behavior are at the center of attention. Authors in [10–11] conducted an experimental investigation on laminar LPG–H₂ jet diffusion flame and NG–H₂ turbulent flame respectively and affirmed that a very small fluctuation in the visible flame length, in up to 20% of hydrogen addition, enhances the flame length when higher hydrogen concentrations are injected. In the other hand, different investigations [5, 7, 8, 12, 13] were undertaken on the height of the blue cone and pointed out the influence of hydrogen injection in the increment of the combustion rate, the reduction of CO, CO₂ and HC emissions, and the extension of the stable lean limit of natural gas burner [1, 9–12, 14–16]. Authors in [17] developed an experimental investigation in optically accessible propane/ hydrogen burner and found that

Corresponding author: A. Bouziane

the flame pattern of the hydrogen is of a weak turbulent flame [18]. A significant reduction in both CO emissions and soot concentration with hydrogen addition was reported in [4, 11, 12] when turbulent jet diffusion flame was characterized. Authors in [19-21] report an increase in flame adiabatic temperature and a reduction in flame thickness. Authors in [15] investigated the counterflow of NG/H₂-air partially premixed flame numerically and experimentally. They stated that the influence of prompt NO_x was important in the total NO_x formation. One of the purposes of this paper is to support certain theories and assumptions drawn from experiments in swirl flames. The goal was to obtain qualitatively correct predictions. This work focuses on the numerical simulation of turbulent flames utilizing well known turbulence and combustion models [22, 23].

Leaner flames due to the hydrogen addition with different various hydrogen-hydrocarbon mixtures lead to thermo-diffusive instabilities which have an impact on flame propagation and stability [24, 25]. More details are needed about the interaction of these types of flames with turbulence in practical devices (swirl-stabilized burner). A better understanding of the fuel improvement on the fundamental flame properties is critical to the development of combustion systems and the role of hydrogen on these systems appears to be an important parameter. The goal of the current study is the investigation of flame behavior by the use of numerical simulations with the RANS model to treat turbulence coupling (the standard /RNG k-epsilon) and the Fr-ED model to represent the turbulence/chemistry interactions. For this purpose, the effect of hydrogen addition in methane flames has been examined by modeling a swirl-stabilized combustor in radial injection. The aim is to give more information on the influence of hydrogen enrichment on the CH₄ flame properties and a comparison is presented between the experimental results and the numerical solutions by considering a flame shape, temperature field, and velocity vectors.

II. NUMERICAL METHODOLOGY

The conservation equations governing reactive flows solved by code FLUENT [26] are mass, momentum, species, and energy using several models described with finite volume methods [27]. The schematic of the partially premixed swirl-stabilized burner configuration can be seen in [5]. The grid is shown in Figure 1. The geometry is open to atmospheric pressure and the meshing of all computational domain examined in this work was set up with GAMBIT [24]. The meshing is non-uniform, with a fine meshing in zones of great interest. The numerical simulations carried out in this study were generated using a segregated, implicit solver. The coupled equations were solved with second-order accuracy in time and momentum, continuity, and turbulence parameters. The algorithms PRESTO and SIMPLE were respectively used for the pressure interpolation and for the coupling of pressure and velocity. A uniform velocity profile was assumed for the inlet conditions, and the inlet turbulence intensity was set to 11%. At the inlet, the boundary conditions were validated by the experiments. Theoretically, zero gradient condition is taken at the outlet boundary. These boundary conditions can be listed as follows: inlet, outlet, and wall boundary conditions. The DO

model was used to calculate the radiating heat transfer of the unconfined flames [28]. The experimental apparatus is a laboratory scale swirl burner [5]. It consists of two concentric pipes with the annulus (outer radius, $R_b = 18\text{mm}$) supplying the central pipe ($r_{int} = 7.5\text{mm}$) injecting the fuel mixtures (swirled air). The fuel mixture is radially injected into the swirled air stream through 8 circular holes. The holes are distributed on the circumferential section located 3mm upstream from the exit of the burner.

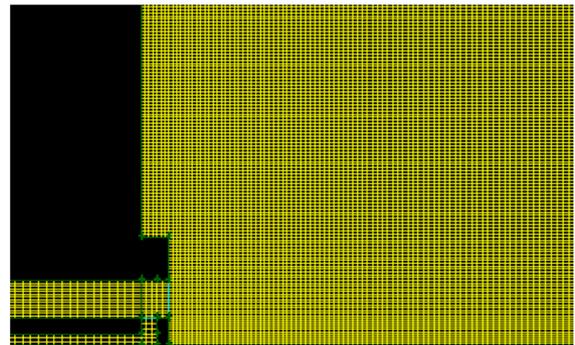


Fig. 1. Numerical grid of the burner/furnace system.

III. TURBULENCE MODEL

To solve the turbulence problem, the two equation-based RANS standard k- ϵ turbulence model was used [29]. It is a semi-empirical turbulence model based on isotropic eddy-viscosity hypotheses. The RNG k- ϵ model is considered as a new version of the k- ϵ family of turbulence models. It is derived by a statistical technique called renormalization group method [30].

IV. RESULTS AND DISCUSSION

A. Flame Shape

The consequence of hydrogen addition is studied using CFD and compared against measurement data [31, 32]. Streamlines were obtained by two-dimensional (2-D) PIV measurements, in a region located immediately downstream of the exit throat of the burner. Flame imaging of the spontaneous flame emission is used to gain a general understanding (see Figure 2(a)). We can see the simulation plot from the right side. Although these diagnostics provide a somewhat different view of the flame, they provide complementary information. In general, the simulated image of the investigated flame (100% CH₄) shows a very similar shape compared to the experimental ones, typical of high swirl processes in transverse injection. The comparison shows that standard k- ϵ is slightly over-predicted at most locations whereas the RNG k- ϵ model results are quite good [33]. Figure 2(a) shows the sudden expansion nozzle at the lower part of the combustion chamber and two locations of high product formation rates are located: one is on centerline in an upside-down V-shape, which can also be clearly seen in the photo of the flame in the experiment and the second in red color indicates a high product formation rate, which means an area with high chemical reaction production rates. The use of finite-rate-Eddy dissipation model generates a reaction at the mixing zone.

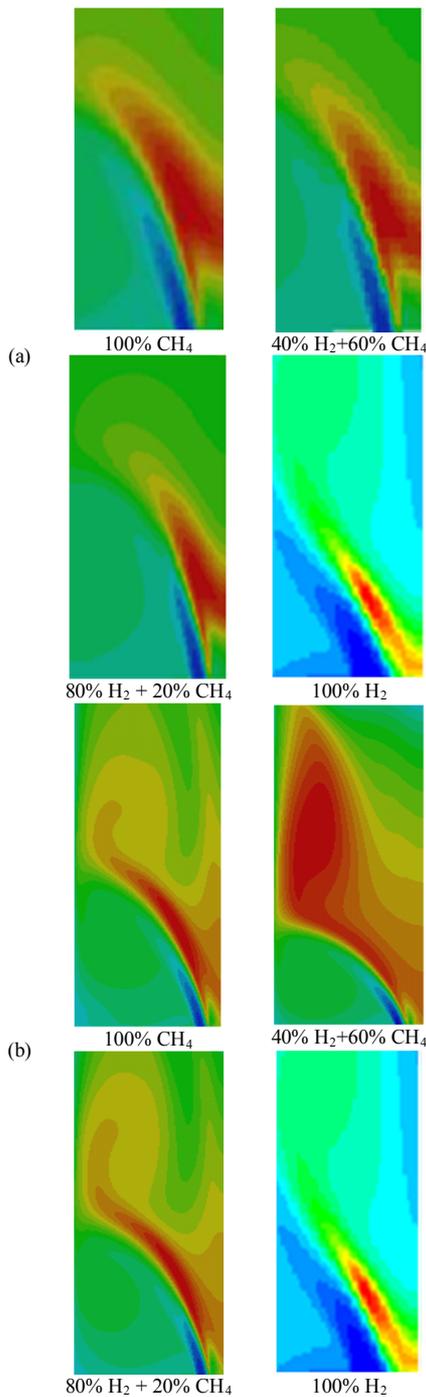


Fig. 2. Flame shape and position comparison of the simulation (right) and the flame photo from the experiment (left): (a) standard k-ε model, (b) RNG k-ε model.

The experimental reaction takes place at the center of the zone where a strong inner recirculation zone appears, whereas reaction occurs with a certain delay, when simulation is considered. The impact of the different percentage of H₂ blending to the CH₄-air flame on the global flame characteristics is shown in Figure 2 (different flame shapes and positions).

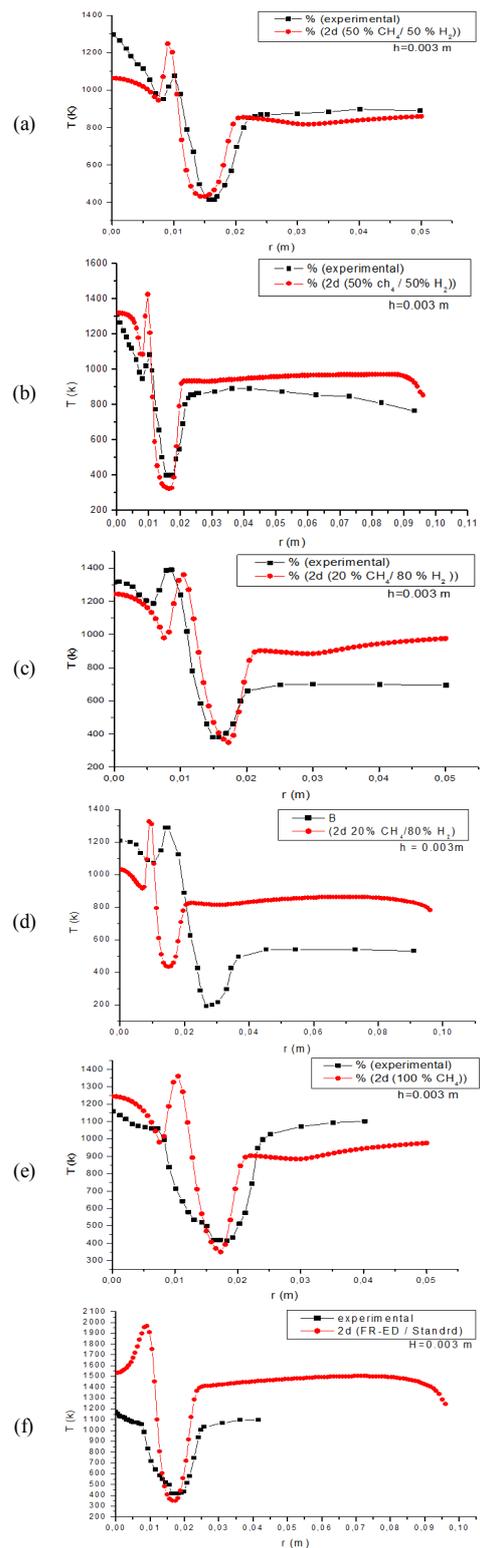


Fig. 3. Radial distribution of mean temperature: (a) RNG k-ε 50% CH₄ - 50% H₂, (b) standard k-ε 50% CH₄ - 50% H₂, (c) RNG k-ε 20% CH₄ - 80% H₂, (d) standard k-ε 20% CH₄ - 80% H₂, (e) RNG k-ε 100% CH₄, and (f) standard k-ε 100% CH₄.

It can be observed both from the experiments and the numerical simulation that the structure of the flame, i.e. central recirculation zone shape, changes with H_2 addition. Enhances in H_2 concentration in the mixture increase the axial velocity and decreases the size and strength of the recirculation region. Similar behavior was reported in [34].

B. Temperature Field

To provide more information in this section, the computational solutions were compared with the measurements. To examine the role of turbulence closure models, the same configuration was modeled using the standard $k-\epsilon$ and RNG $k-\epsilon$ turbulence models. The mean temperature profiles are shown in Figure 3, reporting a different trend between NG and H_2 /NG fuel mixtures.

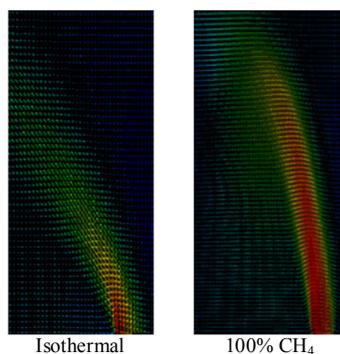


Fig. 4. Experiment (left) and simulation (right) velocity vectors.

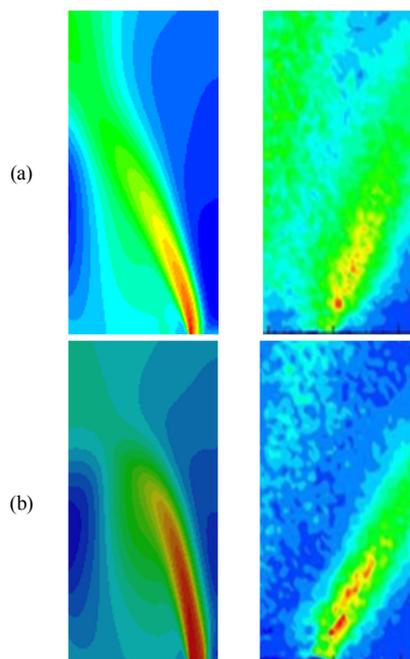


Fig. 5. Velocity fluctuations. Comparison of the experimental data and simulation results: (a) 100 % CH_4 , (b) 20% CH_4 + 80% H_2 flame.

The calculated temperature profiles are in agreement with the experimental results. When moving on radial direction, the

numerical simulation agrees well with the trend of the experimental data. The calculated temperature indicating the existence of the hot area of the flame and then finally to the limit which assigns to the recirculation zone. The temperature profiles can be split in into three separate areas as reported in Figures 5 and 6.

The same results were obtained in [35]. The highest temperature shows that the flame takes place in the inner recirculation zone. When burning 100% NG, the peak temperature of the flame is quite close to 1150K, while its experimental value is estimated as 1250K. The peak temperature of the 80% H_2 flame increases to about 1800K, as we seen in the experiment [31, 32]. An unexpected temperature value is determined by the CFD when the $k-\epsilon$ model is considered. At the periphery, a region characterized by quite uniform temperature values is established. This is related with the formation of a corner recirculating flow induced by the radial expansion of the air stream and by the wall confinement. This model succeeds to predict the hot zone of the flame while over predict the calculated temperature. When $r = 0.05m$, the model agrees with the experimental results. On the other hand, a good agreement is obtained when the RNG $k-\epsilon$ model is used. The RNG $k-\epsilon$ model results are in good accordance with the experimental measurements, and are comparable to the results obtained with the $k-\epsilon$ model, despite the lack of accounting for non-isotropic turbulent shear stresses expected in the swirling flow. It can be observed that the predicted radial profiles of temperature of the two models are reasonable at most of the axial locations.

C. Mean Flow Fields

To provide more information on the nature of the flow, we focus our solutions on the behavior of this type of flame by showing the existence of the Central Recirculation Zone (CRZ). The numerical solution shows an open swirl flow core. The Figure points out a more closed form with a small but strong recirculation region. The reaction zone estimated by simulation is not located on the axis as indicated by the measurements. The dynamic flow field is affected greatly by the combustion that redefines a velocity distribution. This discrepancy is mainly the result of shortcoming of the turbulence model which might overestimate the decay of turbulence intensity for these swirling flames. Under-combustion regime (Figure 4), in a region located immediately in the exit of the burner clearly indicates the existence of a CRZ. The width of the CRZ is smaller than under when burning 100% NG. Similar behavior was reported in [29]. As the H_2 percentage increases, the above regions reduce in size and move closer to the reactant outlet. When burning 100% NG, the axial flow fields by introducing velocity vectors obtained by the simulation are more large than the 80% H_2 flame.

Another important parameter describing the flow field, is the velocity fluctuations (Figure 5). Both results show an increase of the intensity a few mm downstream of the respective flame zone. A CRZ can be found, where the axial velocity is smaller than zero. This recirculation increases the flow divergence in the combustion chamber. In general, the numerical solutions agree well with the experimental results

(100% methane flame), but we failed to predict the inner recirculation zone with precision.

V. CONCLUSION

Numerical simulations of a turbulent (CH₄-H₂)/air flame were conducted in this study. The H₂ percentage varied in the mixture from 0% to 100%. The numerical investigation involved simulations employing the standard k-ε and the RNG k-ε for turbulence closure. The Fr-ED model is used to represent the turbulence/chemistry interactions. Numerical solutions were compared to experimental measurements which were used for model validation. The results show that the structure of the flame is described reasonably in both standard k-ε and RNG k-ε models which can predict satisfactorily the flame shape with the RNG k-ε model providing better performance than the k-ε model. From the flame simulations, the hydrogen addition alters considerably the global flame characteristics by different flame shapes and flame positions. It can be observed from the experiments and the simulations that the structure of the flame varies with hydrogen blending.

In general, the temperature data were well represented. The RNG k-ε model consistently demonstrated superiority over the standard k-ε turbulence model. The dynamic flow field is affected greatly by the combustion. A region located immediately downstream of the exit throat of the burner clearly indicates the existence of a CRZ. Under combustion conditions the width of the CRZ is smaller than under when burning 100% NG. It was concluded that the use of a 2D axial symmetric grid is a promising toolbox for modeling turbulent combustion and can be used for predicting the operating conditions of complex industrial furnaces.

NOMENCLATURE

Y_R	Mass fraction of a particular reactant R
$u'_{i,r}$	Stoichiometric coefficient for reactant i in reaction r
$u'_{i,r'}$	Stoichiometric coefficient for product i in reaction r'
N	Number of chemical species in the system
$M_{w,i}$	Molecular weight of species i
$M_{w,j}$	Molecular weight of species j
$M_{w,R}$	Molecular weight of a particular reactant R
k/ϵ	Large-eddy mixing time scale
A	Empirical constant equal to 4.0
B	Empirical constant equal to 0.5.
α	Thermal diffusivity [$\text{m}^2 \cdot \text{s}^{-1}$]
β	Thermal expansion coefficient [K^{-1}]
ϕ	Solid volume fraction
Θ	Dimensionless temperature
μ	Dynamic viscosity [$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-1}$]

REFERENCES

- [1] R. W. Schefer, "Hydrogen enrichment for improved lean flame stability," *International Journal of Hydrogen Energy*, vol. 28, no. 10, pp. 1131–1141, Oct. 2003, [https://doi.org/10.1016/S0360-3199\(02\)00199-4](https://doi.org/10.1016/S0360-3199(02)00199-4).
- [2] S. O. Bade Shrestha and G. A. Karim, "Hydrogen as an additive to methane for spark ignition engine applications," *Hydrogen as an additive to methane for spark ignition engine applications*, vol. 24, no. 6, pp. 577–586, 1999.
- [3] A. A. Konnov, G. P. Alvarez, I. V. Rybitskaya, and J. D. Ruyck, "The Effects of Enrichment by Carbon Monoxide on Adiabatic Burning Velocity and Nitric Oxide Formation in Methane Flames," *Combustion Science and Technology*, vol. 181, no. 1, pp. 117–135, Dec. 2008, <https://doi.org/10.1080/00102200802380173>.
- [4] A. R. Choudhuri and S. R. Gollahalli, "Combustion characteristics of hydrogen-hydrocarbon hybrid fuels," *International Journal of Hydrogen Energy*, vol. 25, no. 5, pp. 451–462, May 2000, [https://doi.org/10.1016/S0360-3199\(99\)00027-0](https://doi.org/10.1016/S0360-3199(99)00027-0).
- [5] F. Cozzi and A. Coghe, "Behavior of hydrogen-enriched non-premixed swirled natural gas flames," *International Journal of Hydrogen Energy*, vol. 31, no. 6, pp. 669–677, May 2006, <https://doi.org/10.1016/j.ijhydene.2005.05.013>.
- [6] A. J. Akande, "Production of hydrogen by reforming of crude ethanol," Ph.D. dissertation, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, 2005.
- [7] S. S. Shy, Y. C. Chen, C. H. Yang, C. C. Liu, and C. M. Huang, "Effects of H₂ or CO₂ addition, equivalence ratio, and turbulent straining on turbulent burning velocities for lean premixed methane combustion," *Combustion and Flame*, vol. 153, no. 4, pp. 510–524, Jun. 2008, <https://doi.org/10.1016/j.combustflame.2008.03.014>.
- [8] A. R. Choudhuri and S. R. Gollahalli, "Effects of Isothermal Mixing Process on the Stability of Hydrogen-hydrocarbon Blend Turbulent Jet Diffusion Flames," in *International Joint Power Generation Conference*, Florida, FL, USA, Jul. 2000.
- [9] M. Karbasi and I. Wierzbna, "The effects of hydrogen addition on the stability limits of methane jet diffusion flames," *International Journal of Hydrogen Energy*, vol. 23, no. 2, pp. 123–129, Feb. 1998, [https://doi.org/10.1016/S0360-3199\(97\)00031-1](https://doi.org/10.1016/S0360-3199(97)00031-1).
- [10] P. Kumar and D. P. Mishra, "Experimental investigation of laminar LPG-H₂ jet diffusion flame," *International Journal of Hydrogen Energy*, vol. 33, no. 1, pp. 225–231, Jan. 2008, <https://doi.org/10.1016/j.ijhydene.2007.09.023>.
- [11] A. R. Choudhuri and S. R. Gollahalli, "Characteristics of hydrogen-hydrocarbon composite fuel turbulent jet flames," *International Journal of Hydrogen Energy*, vol. 28, no. 4, pp. 445–454, Apr. 2003, [https://doi.org/10.1016/S0360-3199\(02\)00063-0](https://doi.org/10.1016/S0360-3199(02)00063-0).
- [12] H. J. Burbano, A. A. Amell, and J. M. Garcia, "Effects of hydrogen addition to methane on the flame structure and CO emissions in atmospheric burners," *International Journal of Hydrogen Energy*, vol. 33, no. 13, pp. 3410–3415, Jul. 2008, <https://doi.org/10.1016/j.ijhydene.2008.04.020>.
- [13] C. J. Tseng, "Premixed combustion of hydrogen/methane mixtures in a porous medium burner," in *36th Intersociety Energy Conversion Engineering Conference*, Savannah, GE, USA, Aug. 2001.
- [14] S. Naha and S. K. Aggarwal, "Fuel effects on NO_x emissions in partially premixed flames," *Combustion and Flame*, vol. 139, no. 1, pp. 90–105, Oct. 2004, <https://doi.org/10.1016/j.combustflame.2004.07.006>.
- [15] G. J. Rortveit and J. E. Hustad, "NO_x Formations in Diluted CH₄/H₂ Counterflow Diffusion Flames," *International Journal of Energy for a Clean Environment*, vol. 4, no. 4, 2003, <https://doi.org/10.1615/InterJEnrCleanEnv.v4.i4.20>.
- [16] S. A. El-Ghafour, "Pollutant emissions control utilizing hybrid fuel and air nozzle-cascading techniques," M.S. thesis, Port Said University, Port-Said, Egypt, 2011.
- [17] J. B. Heywood and F. R. Vilchis, "Comparison of Flame Development in a Spark-Ignition Engine Fueled with Propane and Hydrogen," *Combustion Science and Technology*, vol. 38, no. 5–6, pp. 313–324, Jul. 1984, <https://doi.org/10.1080/00102208408923778>.
- [18] C. K. Law, *Combustion Physics*. Cambridge, UK: Cambridge University Press, 2006.
- [19] Y. Lafay, B. Renou, G. Cabot, and M. Boukhalfa, "Experimental and numerical investigation of the effect of H₂ enrichment on laminar methane-air flame thickness," *Combustion and Flame*, vol. 153, no. 4, pp. 540–561, Jun. 2008, <https://doi.org/10.1016/j.combustflame.2007.10.002>.
- [20] M. Fukuda, K. Korematsu, and M. Sakamoto, "On Quenching Distance of Mixtures of Methane and Hydrogen with Air," *Bulletin of JSME*, vol. 24, no. 193, pp. 1192–1197, 1981, <https://doi.org/10.1299/jsme1958.24.1192>.
- [21] Z. Dulger, "Adiabatic flame temperature and product composition for lean combustion of hydrogen-methane combination," *The American Society of Mechanical Engineers*, vol. 2, pp. 707–712, 1994.

- [22] A. C. Mangra, "Design and Numerical Analysis of a Micro Gas Turbine Combustion Chamber," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6422–6426, Dec. 2020, <https://doi.org/10.48084/etasr.3835>.
- [23] M. W. Khalid and M. Ahsan, "Computational Fluid Dynamics Analysis of Compressible Flow Through a Converging-Diverging Nozzle using the k-ε Turbulence Model," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5180–5185, Feb. 2020, <https://doi.org/10.48084/etasr.3140>.
- [24] J. B. Bell, R. K. Cheng, M. S. Day, and I. G. Shepherd, "Numerical simulation of Lewis number effects on lean premixed turbulent flames," *Proceedings of the Combustion Institute*, vol. 31, no. 1, pp. 1309–1317, Jan. 2007, <https://doi.org/10.1016/j.proci.2006.07.216>.
- [25] M. Day, J. Bell, P.-T. Bremer, V. Pascucci, V. Beckner, and M. Lijewski, "Turbulence effects on cellular burning structures in lean premixed hydrogen flames," *Combustion and Flame*, vol. 156, no. 5, pp. 1035–1045, May 2009, <https://doi.org/10.1016/j.combustflame.2008.10.029>.
- [26] G. Solero, A. Coghe, and G. Scribano, "Influence of natural gas injection procedure in a swirl Burner," in *7th Mediterranean Combustion Symposium*, Sardinia, Italy, Sep. 2011.
- [27] M. R. Lecic, A. S. Cocic, and S. M. Cantrak, "Original measuring and calibration equipment for investigation of turbulent swirling flow in circular pipe," *Experimental Techniques*, vol. 38, no. 3, pp. 54–62, May 2014, <https://doi.org/10.1111/j.1747-1567.2012.00812.x>.
- [28] B. E. Launder, *Lectures in Mathematical Models of Turbulence*. London, New York: Academic Press, 1972.
- [29] V. Yakhot and S. A. Orszag, "Renormalization group analysis of turbulence. I. Basic theory," *Journal of Scientific Computing*, vol. 1, no. 1, pp. 3–51, Mar. 1986, <https://doi.org/10.1007/BF01061452>.
- [30] F. Cozzi, A. Olivani, and A. Coghe, "Combustion of NG + H₂ fuel mixtures in a partially premixed swirlburner," presented at the Third European Combustion Meeting ECM 2007, 2007.
- [31] A. Olivani, F. Cozzi, G. Solero and A. Coghe, "Analysis of the Environmental Impact of a Swirl Burner: Hydrogen and Natural Gas Mixture Feeding," presented at the 28th Meeting of the Italian Section of The Combustion Institute, 2005
- [32] J. F. Widmann, S. R. Charagundla, and C. Presser, *Benchmark Experimental Database for Multiphase Combustion Model Input and Validation: Characterization of the Inlet Combustion Air*. Gaithersburg, MD, USA: National Institute of Standards and Technology, 1999.
- [33] M. W. Eldeen, M. A. Taemah, and W. M. El-Maghalany, "Influence of Swirl Intensity and Blended Methane-Hydrogen on Gas Turbine Combustion Characteristics," *International Journal of Engineering Research & Technology*, vol. 7, no. 7, pp. 405–412, 2018.
- [34] A. Frassoldati, S. Frigerio, E. Colombo, F. Inzoli, and T. Faravelli, "Determination of NO_x emissions from strong swirling confined flames with an integrated CFD-based procedure," *Chemical Engineering Science*, vol. 60, no. 11, pp. 2851–2869, Jun. 2005, <https://doi.org/10.1016/j.ces.2004.12.038>.

Classification of Voltage Sag Propagation under Double Line to Ground Fault on a Low Voltage Distribution Power Network

Abdul Khaliq Junejo

Department of Electrical Engineering
Quaid-e-Awam University of Engineering, Science and
Technology, Nawabshah, Pakistan
ak.junejo@quest.edu.pk

Mohsin Ali Koondhar

Department of Electrical Engineering
Quaid-e-Awam University of Engineering, Science and
Technology, Nawabshah, Pakistan
enr.mohsinkoondhar@quest.edu.pk

Munwar Ayaz Memon

Department of Electrical Engineering
Quaid-e-Awam University of Engineering, Science and
Technology, Nawabshah, Pakistan
enr.mam@quest.edu.pk

Irfan Ali Channa

Department of Automation
Beijing University of Chemical Technology
Beijing, China
irfanali@mail.buct.edu.cn

Syed Abid Ali Shah Bukhari

Department of Electrical Engineering
Quaid-e-Awam University of Engineering, Science and Technology
Larkana, Pakistan
abidshah@quest.edu.pk

Abstract-The key issue of the voltage-sag propagation on the different types of winding connections for the low voltage transformer in the distribution power network is considered in this study. The issues of voltage sag propagation depend upon the type of the asymmetrical fault and the connection of the transformer windings. In this paper, the double line to ground fault is adopted for the voltage sag characterization. The depiction of the sage propagation is an outcome of the zero sequence components. Voltage sag propagation detection can be used to improve the power quality of the system. The obtained results reveal the characteristics of the voltage drop propagation from the primary to the secondary winding of the transformer throughout different connection. The implemented method will help analyze the Power Quality (PQ) problem in terms of voltage sag of the distributed transmission networks of the power system.

Keywords-voltage sag; power quality; double line to ground fault

I. INTRODUCTION

Power Quality (PQ) is an important issue in electrical systems [1]. Due to the existence of symmetric faults in the transformer windings, the sinking of the distribution network will deviate from the waveform of the power supply voltage, which depends on the nominal sine waveform and the nominal frequency [2, 3]. Many industrial and home customers often have equipment that is sensitive to electrical interference [4].

Corresponding author: Mohsin Ali Koondhar

Therefore, it is important to understand the quality of the supplied power [5, 6]. If sufficient power supply voltage is not provided, electrical appliances may or may not be sued at all. End users have increased their awareness of voltage quality and challenge network operators to obtain appropriate information and provide them with better power supply quality [3]. Nowadays, PQ has become a major aspect of the Power Distribution System (PDS), due to reasons such as the presence of sensitive and nonlinear loads in the distribution system [7, 8]. The main problem is the quality of energy in terms of voltage constancy in the PDS. Voltage sag, defined as a sharp decrease in the amplitude of the supply voltage, is a PQ problem that can cause loss of susceptible loads in the PDS. In recent years, many defects due to low voltage, such as electrical equipment failure, production line loss, and complete equipment failure, have been studied [9-12]. The voltage sag defined by IEEE is the decrease of the RMS voltage value between 0.1 and 0.9pu [10]. Voltage sag can interfere with the operation of sensitive equipment. These circumstances may lead to huge cost-effective fatalities by disturbing the industrial manufacture process [13, 14].

II. DISCRPTION AND SYSTEM MODELING

The variation in voltage, current, or frequency may cause equipment failure/malfunction. According to IEEE 1159-1995 standard, the sag reduces the voltage between 0.1 to 0.9pu. As

shown in Figure 1, the RMS at the power frequency ranges from 0.5 to 1 minute. Figure 2 shows a one line schematic diagram power network with two buses. If a bus failure occurs, the voltage on the primary transformer would be in a fault state, and on the other hand, on the secondary of the transformer has not a zero level of the voltage. It can be said that voltage sag can propagate from primary to secondary.

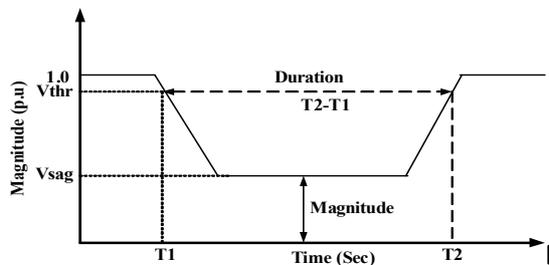


Fig. 1. Sag voltage discription.

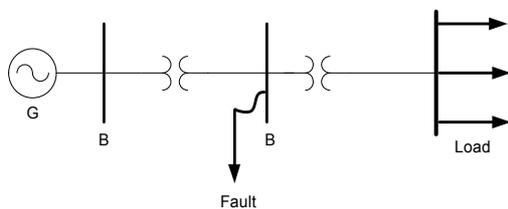


Fig. 2. One line diagram of the power system.

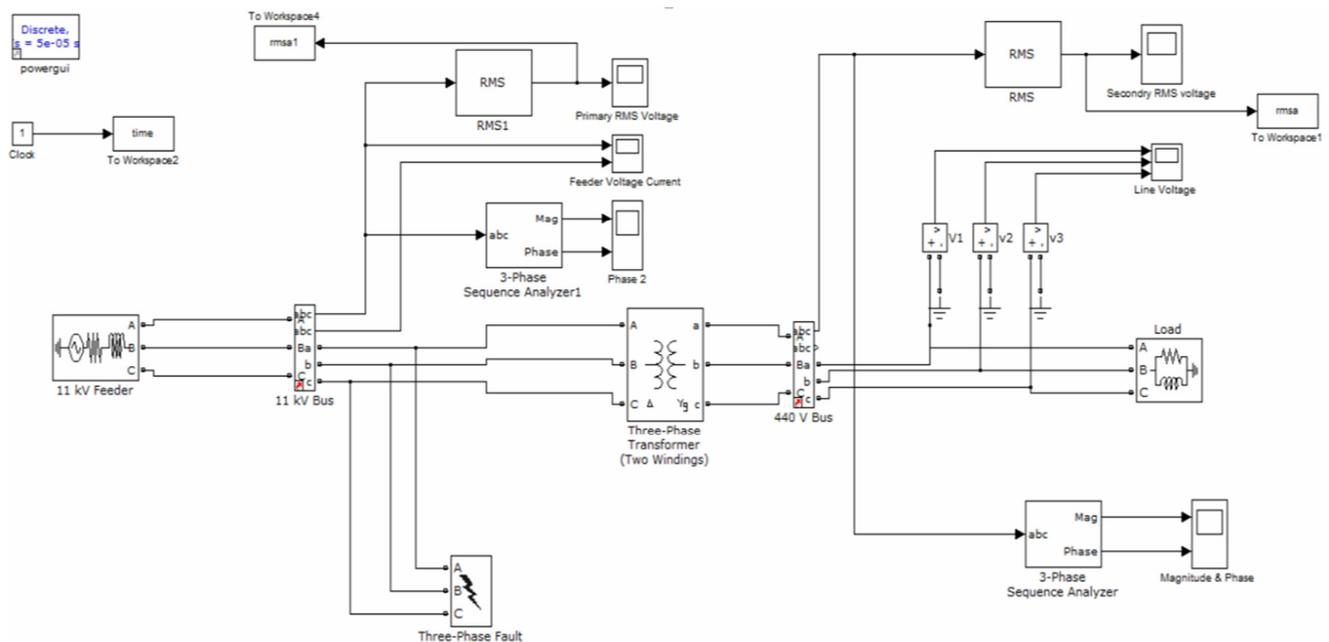


Fig. 3. The Matlab/Simulink model for the classification of voltage sag propagation of the distribution network.

4) *RMS Voltage Block*

This block measures the RMS value.

5) *Two Winding Transformer*

A. *Simulink Model of the Distribution Network for Voltage Sag Propagation*

This study aims to classify the voltage sag propagation under double line to ground fault on a low voltage distribution power network. The power system has four stages: generation, transmission, distribution, and utilization. In the model shown in Figure 3, one distribution feeder of 11kV from the feeder voltage has been sent to the load side and between them there is transformer. When fault occurs on the primary side of the transformer, the voltage will sag at the primary side and also propagate on the secondary side depending on various factors, i.e. fault type, winding connection, winding impedance, and fault impedance. Voltage sag was measured with a measurement block and calculated from the sequence analyzer on the primary and secondary side of transformer.

1) *Three Phase Source Block*

This block is formed in Simulink where three phase voltage displaced with reference degree is present. Here, 11kV voltage is selected and reference angle is selected.

2) *Voltage Current Measurement Block*

This block is formed where the three phase of voltage and current are measured and shown with a connected scope.

3) *Three Phase Fault Block*

This block is formed where the fault type (line to ground fault, double line fault, double lone to ground fault, and three phase faults) is selected.

In this block, winding connections are selected with suitable winding resistance and inductance.

6) *RLC Load Block*

In this block, active and reactive powers are selected.

7) Sequence Analyzer Block

In this block, sequence component values and phase angles are measured and shown.

The discrete time (5×10^{-5}) solver is used in the Simulink model.

B. Modeling for Sequence Analyzer's Symmetrical and Unsymmetrical Faults

In a power system, two kinds of faults occur, due to a variety of reasons, namely the symmetrical and the unsymmetrical faults. In symmetrical faults, all phases are short circuited, so all phases have the same amount of voltage, i.e. zero or maximum. In an unsymmetrical fault, two phases are short circuited with one another or to the ground: line to line, phase to line, and phase-phase to ground faults (LL, LG, and LLL).

C. Sequence Components

Fault calculation may be conducted with the sequence components. There are three sequence components (positive, negative, and zero sequence components) used in the analysis of the power system's faults and its stability. The sequence components are used to analyze the voltage sag propagation under unsymmetrical faults.

1) Positive Sequence Components

Positive sequence components have the same magnitude and displaced phase angle as the normal phase voltage or current, as shown in Figure 4.

2) Negative Sequence Components

Negative sequence components have the same magnitude but opposite phase displacement with normal phase voltage or current as shown in Figure 5.

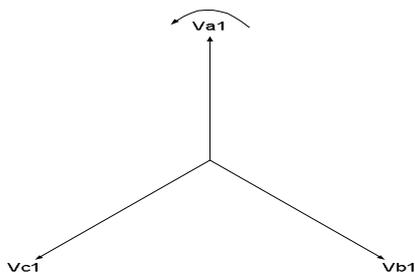


Fig. 4. Positive sequence components.

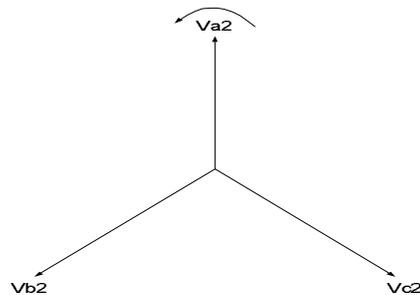


Fig. 5. Negative sequence components.

3) Zero Sequence Components

Zero sequence components have the same magnitude and angle as shown in Figure 6.

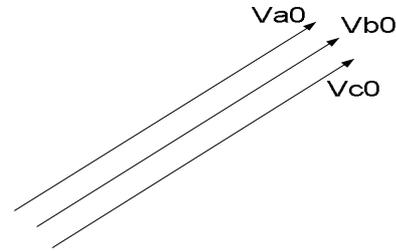


Fig. 6. Zero sequence components.

D. Phase Voltages Calculation by Sequence Components

In a three phase system V_a , V_b , and V_c are the three phase voltages displaced at 0° , 120° and 240° or -120° . Equations (1)-(3) [15] describe the system:

$$V_a = V_1 + V_2 + V_0 \quad (1)$$

$$V_b = a^2 V_1 + a V_2 + V_0 \quad (2)$$

$$V_c = a V_1 + a^2 V_2 + V_0 \quad (3)$$

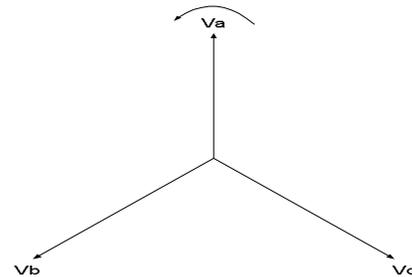


Fig. 7. A three phase vector.

III. RESULTS AND DISCUSSION

In this section, the Simulink sequence analyzer is used to display the simulation results. In this article, the different types of the connection of the three-phase transformers are considered and the voltage sag propagation on the distribution power network under double line to ground fault is analyzed. Sequence components (positive and zero components) are implemented to characterize the voltage sag. When the fault appears on the power system, then the voltage sag also occurs on each side of the transformer, but due to the transformer winding connection, the voltage sag will expand. Voltage sag analysis is performed by changing the fault resistance on the primary side of the transformer to monitor the propagation voltage on the secondary side by modifying the fault resistance, selecting double line to ground fault for voltage sag analysis, and selecting start-delta connection for comparative analysis.

A. LLG Fault with Y-Y Connection

In this section, the LLG with Y-Y connection of the transformer is used to test the voltage sag propagation on the power network. When an LLG fault takes place in the power network, voltage sag emerges on the primary and secondary sides, as shown in Figures 8 and 9. It can be clearly seen from the RMS value waveform that when a fault occurs on the primary to the secondary side of the transformer, then the voltage drop can appear as a different RMS value. The RMS values under LLG fault with different types of transformer connection were simulated in the Matlab model by using the phase sequence analyzer (Table I).

TABLE I. SAG VOLTAGE ON THE PRIMARY AND SECONDARY OF T/F WITH Y-Y CONNECTION

T/F Primary Side	T/F Secondary Side
$V_{a1}=0.333 \angle -1.3$	$V_{a1}=0.332 \angle -4.059$
$V_{a2}=0.333 \angle -121.28$	$V_{a2}=0.3325 \angle -124.09$
$V_{a0}=0.3321 \angle -120.11$	$V_{a0}=0$
$V_a = 0$	$V_a = 0.332 \angle -64$
$V_b = 0$	$V_b = 0.332 \angle -64$
$V_c = 0.99 \angle 119.17$	$V_c = 0.66 \angle 115.92$

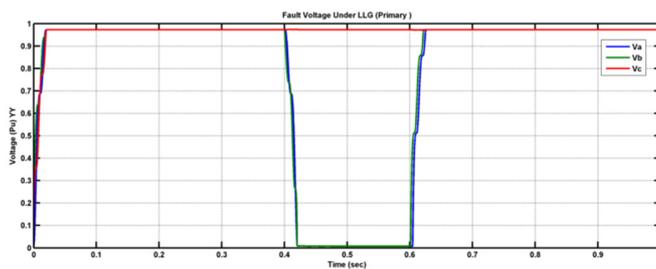


Fig. 8. Sag voltage at the primary side by using Y-Y connection.

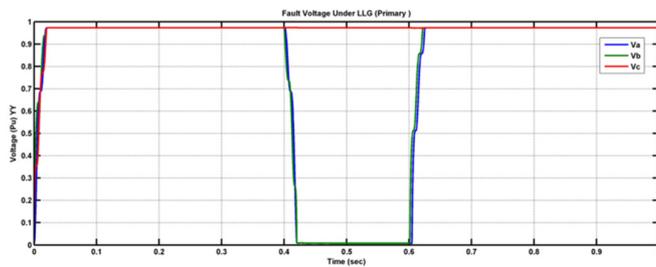


Fig. 9. Sag voltage at the secondary side by using Y-Y connection.

B. LLG Fault with Y-Delta Connection

When a LLG fault occurs in the power network, voltage sag would come out on both sides of the transformer as illustrated in Figures 10-11. The RMS value waveforms obviously state that once the primary side of the transformer fails, at different RMS values, the voltage sag of the transformer will drop to the secondary side. This is simulated in the model and can be passed. The phase sequence analyzer performs the calculations, as shown in Table II.

C. LLG Fault with YG-YG Connection

LLG fault with YG-YG connection is considered in this section. When a LLG fault takes place in the distribution

network, there will be voltage sag on the primary and secondary sides as shown in Figures 12-13.

TABLE II. PRIMARY AND SECONDARY OF T/F ON Y-D CONNECTION VOLTAGE SAG

T/F Primary Side	T/F Secondary Side
$V_{a1}=0.333 \angle -1.3$	$V_{a1}=0.332 \angle -34.05$
$V_{a2}=0.333 \angle -121.28$	$V_{a2}=0.3325 \angle -94.039$
$V_{a0}=0.3321 \angle -120.11$	$V_{a0}=0$
$V_a = 0$	$V_a = 0.57 \angle -64$
$V_b = 0$	$V_b = 0$
$V_c = 0.99 \angle 119.17$	$V_c = 0.57 \angle 115.98$

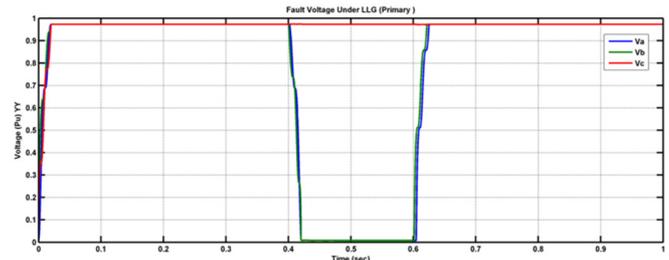


Fig. 10. Y-D connection sag voltage at the primary side.

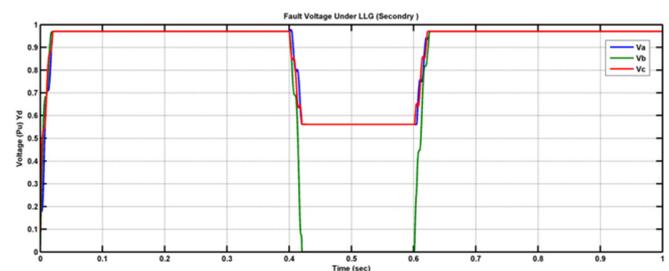


Fig. 11. Y-D connection sag voltage at the secondary side.

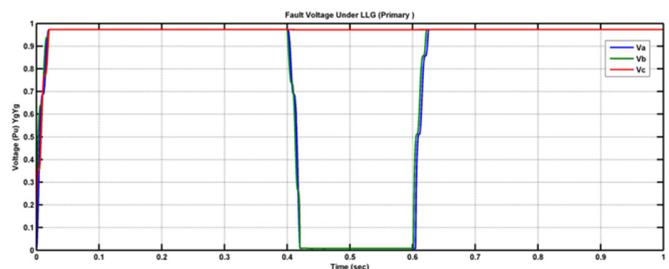


Fig. 12. YG-YG connection sag voltage at the primary side.

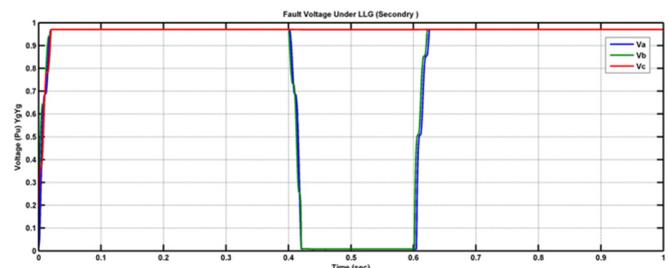


Fig. 13. YG-YG connection sag voltage at the secondary side.

From the RMS value waveform, it is observed that when the primary of the transformer fails, at different RMS value, the

voltage sag of the transformer will also drop to the secondary side. The phase sequence analyzer performance is shown in Table III.

TABLE III. PRIMARY AND SECONDARY OF T/F ON YG-YG CONNECTION VOLTAGE SAG

T/F Primary Side	T/F Secondary Side
$V_{a1} = 0.333 \angle -1.3$	$V_{a1} = 0.332 \angle -4.25$
$V_{a2} = 0.333 \angle -121.28$	$V_{a2} = 0.332 \angle -124.23$
$V_{a0} = 0.3321 \angle 120.11$	$V_{a0} = 0.332 \angle -117.15$
$V_a = 0$	$V_a = 0$
$V_b = 0$	$V_b = 0$
$V_c = 0.99 \angle 119.17$	$V_c = 0.99 \angle 116.2$

D. LLG with YG-Delta Connection

Following, the LLG with YG-Delta connection has been tested for voltage sag propagation. When an LLG fault occurs, the voltage sag on the primary and secondary would appear on the transformer. It can be observed from the waveform of the RMS values in Figures 14-15 that when the primary of the transformer is under fault condition, then the voltage sags would be dropped to the secondary side of the transformer at different RMS values. Table IV represents the voltage sags on both sides of the transformer under phase sequence analyzer performance.

TABLE IV. PRIMARY AND SECONDARY OF T/F ON YG-D CONNECTION VOLTAGE SAG

T/F Primary Side	T/F Secondary Side
$V_{a1} = 0.333 \angle -1.3$	$V_{a1} = 0.332 \angle -34.05$
$V_{a2} = 0.333 \angle -121.28$	$V_{a2} = 0.3325 \angle -94.039$
$V_{a0} = 0.3321 \angle 120.11$	$V_{a0} = 0$
$V_a = 0$	$V_a = 0.57 \angle -67$
$V_b = 0$	$V_b = 0$
$V_c = 0.99 \angle 119.17$	$V_c = 0.57 \angle 115.98$

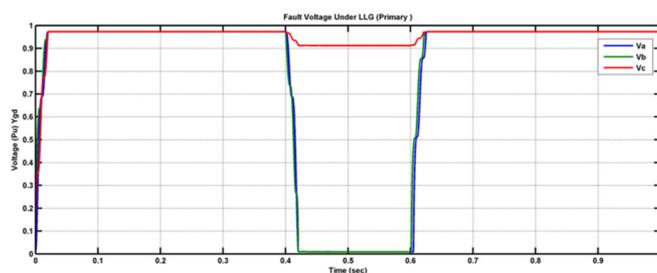


Fig. 14. YG-delta connection sag voltage at the primary side.

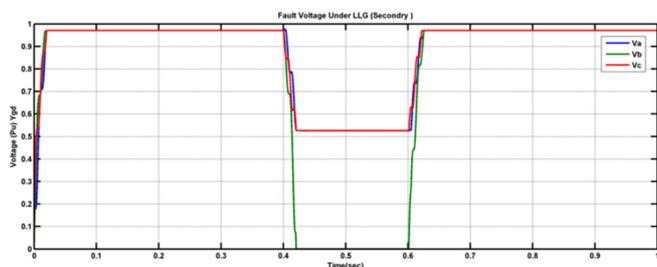


Fig. 15. YG-delta connection sag voltage at the secondary side.

E. Double Line to Ground Fault with Delta-Star connection

The RMS values of an LLG fault are shown in Figures 16-17. The voltage sags on the primary and secondary sides are again calculated with the sequence analyzer in the Matlab model. The calculated RMS values are shown in Table V.

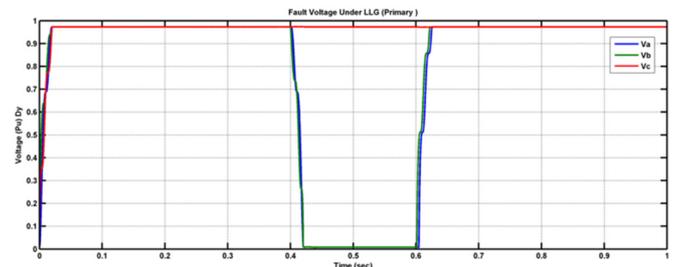


Fig. 16. D-Y connection sag voltage at the primary side.

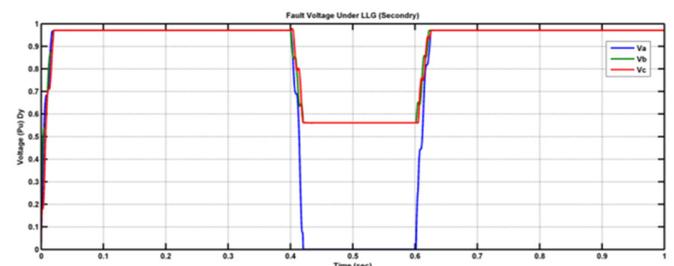


Fig. 17. D-Y connection sag voltage at the secondary side.

TABLE V. PRIMARY AND SECONDARY OF T/F ON D-Y CONNECTION VOLTAGE SAG

T/F Primary Side	T/F Secondary Side
$V_{a1} = 0.333 \angle -1.3$	$V_{a1} = 0.332 \angle 25.945$
$V_{a2} = 0.333 \angle -121.28$	$V_{a2} = 0.3325 \angle -154$
$V_{a0} = 0.3321 \angle 120.11$	$V_{a0} = 0$
$V_a = 0$	$V_a = 0$
$V_b = 0$	$V_b = 0.57 \angle -64$
$V_c = 0.99 \angle 119.17$	$V_c = 0.57 \angle 115.98$

F. Double Line to Ground Fault with Delta-Delta Connection

The RMS values of an LLG fault are shown in Figures 18-19. The voltage sags on the primary and secondary sides are again calculated with the sequence analyzer in the Matlab model. The calculated RMS values are shown in Table VI.

TABLE VI. PRIMARY AND SECONDARY OF T/F ON D-D CONNECTION COLTAGE SAG

T/F Primary Side	T/F Secondary Side
$V_{a1} = 0.333 \angle -1.3$	$V_{a1} = 0.332 \angle -4.059$
$V_{a2} = 0.333 \angle -121.28$	$V_{a2} = 0.3325 \angle -124.09$
$V_{a0} = 0.3321 \angle 120.11$	$V_{a0} = 0$
$V_a = 0$	$V_a = 0.332 \angle -64$
$V_b = 0$	$V_b = 0.332 \angle -64$
$V_c = 0.99 \angle 199.17$	$V_c = 0.66 \angle 115.92$

G. LLG Fault with D-YG Connection

The RMS values of an LLG fault are shown in Figures 20-21. The voltage sags on the primary and secondary sides are

again calculated with the sequence analyzer in the Matlab model. The calculated RMS values are shown in Table VII.

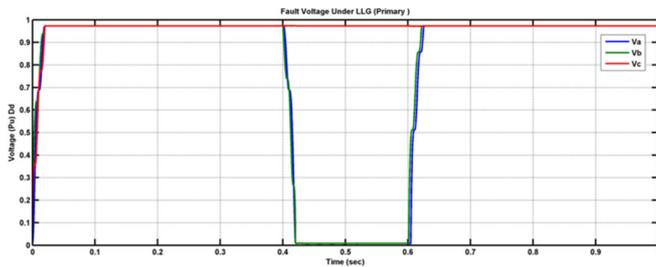


Fig. 18. D-D connection sag voltage at the primary side.

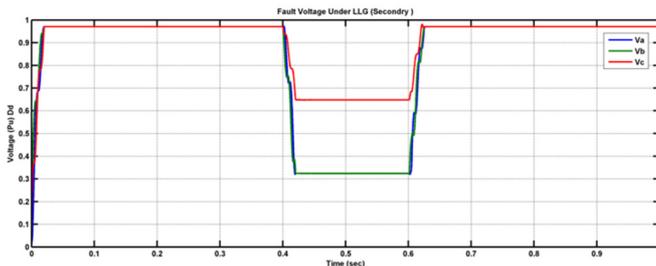


Fig. 19. D-D connection sag voltage at the secondary side.

TABLE VII. PRIMARY AND SECONDARY OF T/F ON D-YG CONNECTION VOLTAGE SAG

T/F Primary Side	T/F Secondary Side
$V_{a1}=0.333 \angle -1.3$	$V_{a1}=0.332 \angle 25.945$
$V_{a2}=0.333 \angle -121.28$	$V_{a2}=0.3325 \angle -154$
$V_{a0}=0.3321 \angle 120.11$	$V_{a0}=0$
$V_b = 0$	$V_b = 0.57 \angle -64$
$V_c = 0.99 \angle 119.17$	$V_c = 0.57 \angle 115.98$

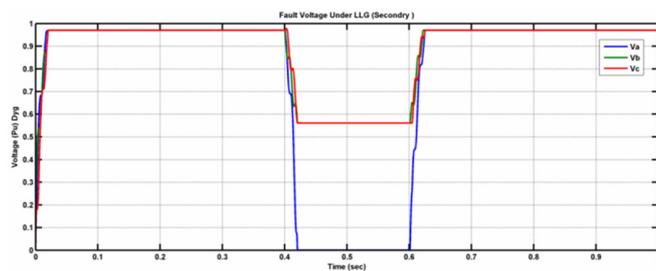


Fig. 20. D-YG connection sag voltage at the primary side.

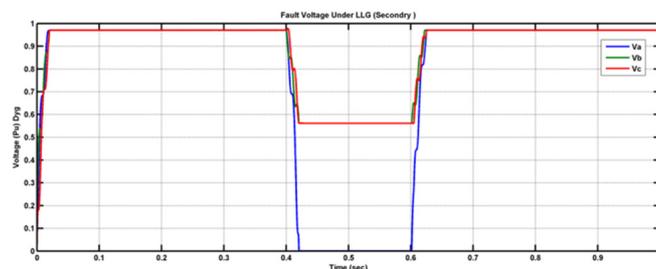


Fig. 21. D-YG connection sag voltage at the secondary side.

IV. CONCLUSION

This paper conducts in-depth research on the voltage sag propagation under double line to ground fault (LLG) with different transformer connections, which are used for low-voltage in distribution power networks. The results have been acquired with the use of the sequence analyzer. The obtained results show the voltage sag classification, which propagates from the primary to the secondary of the transformer. The voltage sag characterizations are important in order to check the power quality under fault condition of the distribution network. This study's results may be found helpful when designing proper protection schemes and ways to improve the voltage sag under fault conditions at various levels of the power system. It is worth noting that when a zero-sequence component is included, then voltage sag propagation would be different on the secondary side of the transformer. If the zero sequence components are not included, the same voltage sag would propagate from the primary to the secondary of the transformer. In addition, it can be observed that a phase shift would appear when a fault occurs on the secondary side of the transformer.

REFERENCES

- [1] A. H. Soomro, A. S. Larik, M. A. Mahar, A. A. Sahito, and I. A. Sohu, "Simulation-based Analysis of a Dynamic Voltage Restorer under Different Voltage Sags with the Utilization of a PI Controller," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5889–5895, Aug. 2020, <https://doi.org/10.48084/etasr.3524>.
- [2] A. K. Junejo, G. M. Bhutto, A. S. Saand, and M. U. Keerio, "Analysis of Unsymmetrical Voltage Sag Propagation Through Distribution Transformer," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 13, no. 5, pp. 403–408–408, 2016, <https://doi.org/10.19026/rjaset.13.2958>.
- [3] P. S. Kiran, B. Nookesh, M. N. Bhavani, B. A. Kumar, and G. Mani, "Voltage Sag Mitigation Using Pulse Width Modulation Switched Autotransformer through Matlab Simulation," *International Journal of Engineering Research and Applications*, vol. 4, no. 4, pp. 33–36, Apr. 2014, <https://doi.org/10.19026/rjaset.13.2958>.
- [4] F. Jandan, S. Khokhar, Z. A. Memon, and S. a. A. Shah, "Wavelet Based Simulation and Analysis of Single and Multiple Power Quality Disturbances," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3909–3914, Apr. 2019, <https://doi.org/10.48084/etasr.2409>.
- [5] A. A. Malak, G. M. Bhutto, E. A. Buriro, and M. A. Koondhar, "Mathematical Analysis of Voltage Dip propagation in CIGRE (International council of large electric system) Low Voltage Distribution Test Network," *Journal of Applied and Emerging Sciences*, vol. 10, no. 2, pp. 109–116, Dec. 2020, <https://doi.org/10.36785/jaes.v10i2.405>.
- [6] E. A. Al-Ammar, A. Ul-Haq, A. Iqbal, M. Jalal, and A. Anjum, "SRF based versatile control technique for DVR to mitigate voltage sag problem in distribution system," *Ain Shams Engineering Journal*, vol. 11, no. 1, pp. 99–108, Mar. 2020, <https://doi.org/10.1016/j.asej.2019.09.001>.
- [7] M. Ramachandran, A. M. C. Mary, M. Muthukumaran, J. Ganesan, A. Krishnaveni, and D. E. Selvaraj, "A Review on Basic Concepts and Important Standards of Power Quality in Power System," *International Journal of Science and Engineering Applications*, vol. 4, no. 5, pp. 299–303, Oct. 2015, <https://doi.org/10.7753/IJSEA0405.1013>.
- [8] M. M. Ghahderijani, M. Castilla, L. G. de Vicuña, A. Camacho, and J. T. Martínez, "Voltage sag mitigation in a PV-based industrial microgrid during grid faults," in *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, Edinburgh, UK, Jun. 2017, pp. 186–191, <https://doi.org/10.1109/ISIE.2017.8001245>.

- [9] S. Thirukkovai, J. Venkatesan, and D. S. M. Girirajkumar, "Voltage Sag/Swell Mitigation Using UPQC," *International Journal of Scientific & Engineering Research*, vol. 5, no. 4, pp. 161–167, Apr. 2014.
- [10] S. I. A. J. and K. K. V. Kumar, "Sag/Swell Migration Using Multi Converter Unified Power Quality Conditioner," *International Journal of Advances in Engineering & Technology*, vol. 1, no. 5, pp. 437–440, Nov. 2011.
- [11] J. Chen, L. Qi, G. Jin, and Y. Xue, "Analysis of the effect on propagation of voltage sag considering different winding connection modes of transformer," presented at the 2016 International Conference on Advanced Electronic Science and Technology (AEST 2016), Nov. 2016, pp. 547–556, <https://doi.org/10.2991/aest-16.2016.74>.
- [12] O. H. Eyüboğlu, B. Dindar, and Ö. Gül, "Series Resonance Type Fault Current Limiter for Fault Current Limitation and Voltage Sag Mitigation in Electrical Distribution Network," in *2020 2nd Global Power, Energy and Communication Conference (GPECOM)*, Izmir, Turkey, Oct. 2020, pp. 256–261, <https://doi.org/10.1109/GPECOM49333.2020.9247878>.
- [13] "IEEE Recommended Practice for Monitoring Electric Power Quality," *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009)*, pp. 1–98, Aug. 2019, <https://doi.org/10.1109/IEEESTD.2019.8796486>.
- [14] S. S. Rao, *Switchgear Protection and Power Systems*, 13th ed. Delhi, India: Khanna Publishers, 2008.



Irfan Ali Channa was born at Nawabshah, Pakistan. He received his B.E and M.E from the Electrical Engineering Department of the Quaid-e-Awam University of Engineering, Science and Technology in 2011 and 2015 respectively. He is currently enrolled as a Ph.D. student in Beijing University of Chemical Technology, China. His fields of interest are power systems, power quality, signal processing, and AI.



Syed Abid Ali Shah was born in 1983 in Sukkur Sindh-Pakistan. He received the B.E in Electrical Engineering from Quaid-e-Awam University of Engineering, Science and Technology in 2009, his M.E in Electrical Power Engineering from Mehran University, Pakistan in 2015, and his Ph.D. from Aston University, UK in 2019. He has successfully completed the Associate Fellowship from the Higher Education Academy, UK. He is an Assistant Professor at the Quaid-e-Awam University of Engineering. His research interests include the design and optimization of electrical machines and drives and analyzing of special machines. He is a registered engineer of Pakistan Engineering Council and a member of IEEE and IET of UK and Ireland section.

AUTHOR PROFILES



Abdul Khaliq Junejo was born in Larkana, Sindh, Pakistan in 1989. He received his B.E. and M.E. from the Electrical Engineering department of the Quaid-e-Awam University of Engineering, Science and Technology, Pakistan in 2011 and 2015 respectively. He is currently working towards his PhD in the School of Electrical and Electronics Engineering at the State Key Laboratory of Advanced Electromagnetic Engineering in Huazhong University of Science and Technology,

Wuhan, China. He is currently employed as an Assistant Professor in the Quaid-e-Awam University of Engineering, Science and Technology. He has served as a reviewer for various IEEE journals. His research interests include sliding mode control, direct torque control, model predictive control, and sensorless control methods for permanent magnet synchronous machines, induction machines, and linear induction machines and drives.



Mohsin Ali Koondhar was born in Nawabshah, Sindh, Pakistan in 1985. He received his B.E. and M.E. from the Electrical Engineering department of the Quaid-e-Awam University of Engineering, Science and Technology in 2007 and 2016 respectively. He is currently employed as a Lecturer in the Quaid-e-Awam University of Engineering, Science and Technology. He has served as a reviewer for IET Generation, Transmission & Distribution Journal. His research interests include

the control of DC and AC machines, renewable energy, and programmable logic controllers.



Munwar Ayaz Memon was born at Bhiria, Pakistan in 1987. He received his B.E and M.E from Quaid-e-Awam University of Engineering Science and Technology and nowadays is working as an Assistant Professor at the Department of Electrical Engineering at the Quaid-e-Awam University of Engineering, Science and Technology. He is currently pursuing his Ph.D. in Electrical Engineering from the same University. His fields of interest are power systems, power

quality, and renewable energy systems. He is an author/co-author of 8 publications. He is a member of Pakistan's Engineering Council.

A Techno-Economic Comparative Study of a Grid-Connected Residential Rooftop PV Panel

The Case Study of Nahr El-Bared, Lebanon

Hüseyin Çamur

Department of Mechanical Engineering
Engineering Faculty
Near East University
Nicosia, Cyprus
huseyin.camur@neu.edu.tr

Youssef Kassem

Department of Mechanical Engineering
Engineering Faculty
Near East University
Nicosia, Cyprus
yousseuf.kassem@neu.edu.tr

Ebaa Alessi

Department of Mechanical Engineering
Engineering Faculty
Near East University
Nicosia Cyprus
ibaa_a88@hotmail.com

Abstract—Lebanon suffers from daily electricity shortages. The country has paid much attention to renewable energy sources, particularly solar, to gradually replace conventional energy. Installing a photovoltaic (PV) system becomes increasingly attractive for residential consumers due to the rising electricity tariff rates while it reduces the dependency on domestic power generators. No known study has dealt with the investigation of potential grid-connected rooftop PV systems with various sun-tracking modes and PV technologies in Nahr El-Bared, Lebanon. Consequently, the main objective of the current paper is to investigate the feasibility of a 5kW grid-connected PV system of various technologies (mono-crystalline silicon and poly-crystalline silicon) and sun-tracking modes including fixed tilt and 2-axis systems for rooftop households in Lebanon. The Nahr El-Bared camp was the case study was of the paper. RETScreen Expert software was used to evaluate the techno-economic performance of the proposed systems. The results show that the annual electrical energy from a fixed 5kW PV panel tilted at an optimal angle ranged from 8564.47kWh to 8776.81kWh, while the annual electrical energy from the PV tracking system was within the range of 11511.67-12100.92kWh. This amount of energy output would contribute significantly to reduce the energy shortage in the country. A typical household was selected to establish a load profile and load supply during both grid availability and outage periods. The highest energy consumption that can be covered by the PV systems was recorded during the spring and summer seasons. Also, the average energy production cost ranged from 0.0239 to 0.0243\$/kWh for all the proposed systems. It was concluded that a 5kW grid-connected rooftop PV system could be economically justifiable. Finally, this study tried to increase the awareness about utilizing PV sun-tracking systems and the feasibility of small-scale grid-connected rooftop PV systems in the selected regions. The results of this research can help investors in the energy and building sectors.

Keywords—Lebanon; Nahr El-Bared camp; rooftop PV system; grid-connected; RETScreen

I. INTRODUCTION

Energy is an indispensable need, which, until quite recently, was met with the consumption of conventional fossil fuels. However, the use of fossil fuels has polluted the environment and led to climate change mainly due to greenhouse gas emissions [1]. The environmental problems resulting from the increasing consumption of fossil fuels have encouraged scientific researchers to search for alternative sources of energy that would affect less the environment [2, 3]. Many researchers found that renewable energy sources, like solar and wind, play a significant role in reducing Greenhouse Gas (GHG) emissions. For instance, authors in [4] concluded that the utilizing of solar thermal energy provided an important step towards sustainable zero-emission production in the industry. Authors in [5] found that renewable energy sources have a high potential for providing an increasing share of future energy growth without increasing GHG emissions. Authors in [6] reported that the production of electricity with the use of solar systems can protect the environment. Recently, solar power has been considered as a potential, economically viable, and environment friendly energy source. Authors in [7] evaluated the solar energy potential in Pakistan based on data from 58 meteorological stations covering the whole country. The results indicated that in an area of 100m², 45 MW to 83 MW power per month may be generated in the southern Punjab, Sindh, and Balochistan regions. Authors in [8] compared the potential of utilizing wind energy and solar energy as power sources for a small household in three urban regions in Northern Cyprus. The results showed that the selected regions have huge solar

Corresponding author: Youssef Kassem

potential compared to the wind energy potential. Authors in [9] studied the feasibility of a 12MW grid-connected wind/PV project in two regions of Northern Cyprus. The results indicated that the selected regions have a high potential compared to wind energy. Authors in [10] evaluated a techno-economic analysis of a 50MW wind/PV system in various locations in Libya. The results showed that a solar PV plant project was more economical than a wind farm. Authors in [11] reported that the production of electricity using solar systems would have a great material impact on the lives of low-income households as compared to their high-income counterparts in United States. Authors in [12] investigated the feasibility of a 10MW grid-connected PV system at different locations in Libya. The results indicated that the proposed system offered the best solution for generating electricity in Libya.

Lebanon suffers from problems in the electricity sector. Generally, electricity cuts range from 3 to 13 hours per day, depending on the area. As a result, families are forced to pay two electricity bills: one to the Electricity Provider of Lebanon (Electricite du Liban-EDL) for the energy it provides, and another for generators, in which many families have to invest in order to ensure continuous access to electricity. Moreover, the electric rationing, has returned at a high rate due. Presently, the demand is estimated to be more than 3458MW, while the current capacity is around 2000MW. Besides, the cost of generating electricity is high. With peak time electricity demand of about 3458MW, which is probably less than its real size, the EDL only provides a part of this demand, and the remaining demand is met through private diesel generators, illegally connected to low voltage distribution networks. Most of the households depend on this source despite its very high prices and its low capacity. The potential for utilizing wind and solar energy in Lebanon has been investigated in [13-22]. Authors in [13] investigated the feasibility of 100MW grid-connected wind/solar systems in the Rayak region. The results indicated that wind power was more efficient than solar power in the selected region. Authors in [17] investigated the potential of wind energy at 3 coastal regions of Lebanon. The results showed that small-scale wind turbines offered high wind potential when compared to other types of wind turbines. Similar results were found in [18, 19]. Furthermore, according to the solar atlas map, the Global Horizontal Irradiation (GHI) and Direct Normal Irradiation (DNI) values vary from 1600kWh/m² to 2000kWh/m² and 1700W/m² to 2200W/m², respectively. Based on the values of GHI and DNI, it is concluded that the country has high solar resources, being categorized as good, excellent, or outstanding according to solar potential classes [23]. The coastal regions have lesser values of GHI and DNI compared to other regions based on the solar atlas map. According to the wind atlas map, the mean wind speed and wind power density are within the range of 2.5-3m/s and 25-75W/m², respectively at a height of 10m in coastal regions. Therefore, it can be concluded that the country has a huge solar energy potential compared to wind energy. The conclusions after reviewing the previous studies are:

- Lebanon has a huge solar energy potential when compared to wind energy.

- Solar power systems can reduce fossil fuel consumption and CO₂ emissions.
- No scientific studies proposed small-scale grid-connected rooftop PV solar power systems to meet the electricity demand of households in Nahr El-Bared.
- To the best of our knowledge, no studies investigated the economic feasibility of small-scale PV systems with various sun-tracking systems (fixed-tilt and two-axis tracking systems) in Lebanon.

Thus, there is a literature discussing the performance of rooftop PV systems in Lebanon. Also, according to [13], citizens rely on domestic power generators or small home generators during the periods of power outage, adding financial burdens to the citizens. Furthermore, none of the reviewed studies addresses the case of Nahr El-Bared in Lebanon. According to United Nations Relief and Works Agency in 2018, the total energy generated from the generators in the Nahr El-Bared is found to be about 80MWh/day. Therefore, grid-connected PV systems can be an attractive solution to reduce electricity consumption, dependence on utility power, and increase electricity generation from renewable energy sources for residential electricity users. Furthermore, household rooftop PV systems will create significant savings in the cost of electricity bills. Additionally, such systems can help the EDL to reduce fossil fuel consumption and make electricity available throughout the day. Consequently, the study aims to investigate the techno-economic and environmental sustainability of rooftop PV systems in Nahr El-Bared. To achieve this, the NASA's database has been utilized as a source of meteorological information. Sun-tracking systems, including fixed tilt and 2-axis systems, are investigated in this study. Moreover, numerous economic indices including net present value, payback period, annual life cycle savings, internal rate of return, and the Levelized Cost of Energy have been taken into consideration as measure of performance indicators for the proposed projects. To this aim, RETScreen software was used in the present study. The proposed system can provide valuable inputs for the development of new policies and innovative solutions for the PV market growth in the country.

II. MATERIALS AND METHODS

The solar energy potential in the selected region in the Northern part of Lebanon is discussed based on NASA's average monthly global solar radiation and air temperature. Besides, the grid-connected rooftop PV system is investigated as a solution for the electricity crisis with the use of the RETScreen software. Figure 1 shows the methodology used in this study.

A. Study Area and Data

Nahr El-Bared camp is one of the largest Palestinian camps in the northern part of Lebanon, located near Tripoli and near the Mediterranean Sea at 34°30'28.19"N latitude and 35°57'23.99"E longitude. From a topographic perspective, the highest and lowest average elevation reach 27m and 12m above the sea level respectively. The area of the old camp is estimated to be about 2km². According to United Nations Relief and Works Agency (2018), Nahr El-Bared camp is divided in two

camp, the old camp with an average apartment area of 78m² and total built residential area of 360000m² and the new camp with an average household area of 150m² and total built area of 300000m². Several studies evaluated the solar energy potential in different locations using the NASA database [24, 25].

Moreover, the NASA database showed good agreement with the measured data of global solar irradiation of previous studies [25-27]. Therefore, the monthly NASA database is utilized in the current study to investigate the potential of solar energy in the selected region.

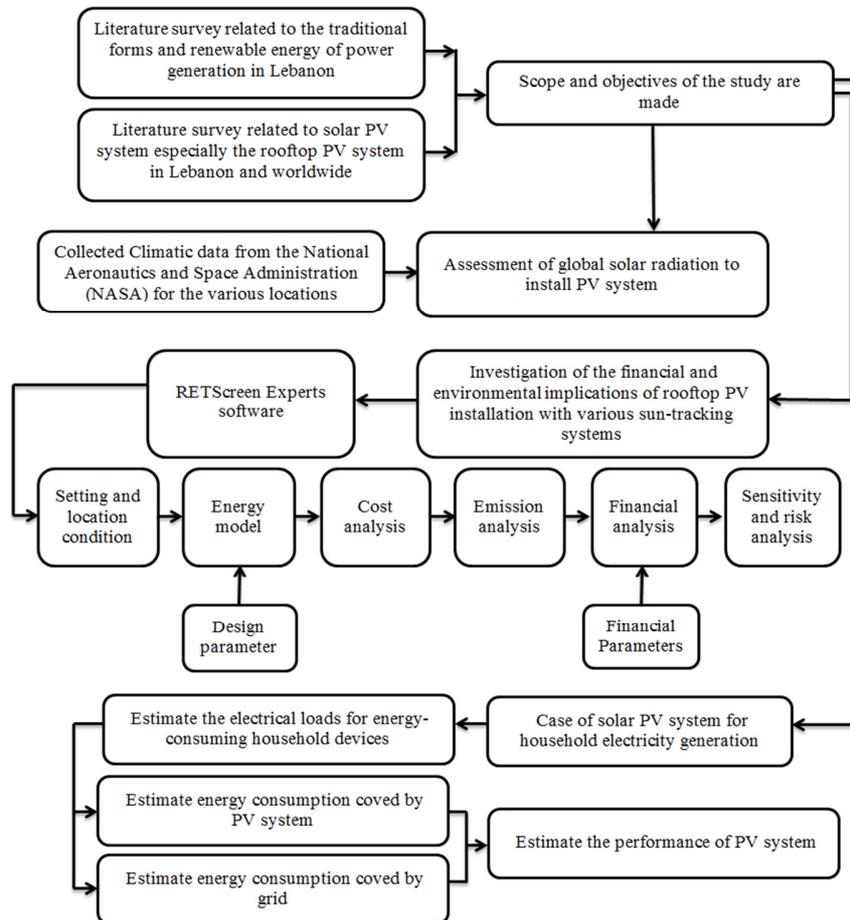


Fig. 1. The flowchart of the analysis steps of the current study.

B. Design of the PV Power System

For the proposed small-scale grid-connected PV system, mono-Si-CS6X-300M and poly-Si - CS6X-310P were selected which are manufactured by Canadian Solar. Table I summarizes the specifications of the selected modules. The total number of modules and the area required for the proposed system are 17 modules and 33m². The description method of designing a solar PV system is available in [10, 13, 25].

Power generating factor (PGF):

$$PGF = \frac{SI \times SH}{STCI} \quad (1)$$

Energy demand (ED):

$$ED = ECAL \quad (2)$$

Solar PV Energy Required (SPVER):

$$SPCER = PER \times ELS \quad (3)$$

PV Module Sizing:

$$TWPR = \frac{SPVER}{PGF} \quad (4)$$

$$PVMS = \frac{TWPR}{PVOPR} \quad (5)$$

Inverter Sizing (IS:)

$$IS = PER \times FS \quad (6)$$

where *SI* is solar irradiance, *SH* is the sunshine hours, *STCI* is standard test condition irradiance, *ECAL* is the energy consumption of all loads, *TWPR* is total Watt peak rating, *PGF* is panel generation factor, *PVMS* is PV module size, *PVOPR* is PV output power rating, *FS* is the factor of safety, *PER* is peak energy requirements, *ELS* is energy lost in the system and the factor of safety is 1.3. Moreover, one inverter with 6kW capacity and 97.9% efficiency is used to convert the DC into AC and feed it directly to the grid (Table II).

TABLE I. PV MODULE SPECIFICATION

Item	Specification	
Module technology	Mono-Si	Poly-Si
Manufacturer	Canadian Solar	
Model	mono-Si - CS6X-300M	poly-Si - CS6X-310P
Nominal power [W]	300	310
Open-circuit voltage [V]	45	44.9
Short-circuit current [A]	8.74	9.08
Voltage at point of maximum power [V]	36.5	36.4
Current at point of maximum power [A]	8.22	8.52
Module area [m ²]	1.919	1.918
Efficiency [%]	15.63	16.16
Warranty [year]	25	25
Cost [USD/Wdc]	0.83	0.8

TABLE II. SPECIFICATIONS OF THE SELECTED INVERTOR

Input DC parameters	Value
MPP voltage range V_{DC} (at 25°C / at 50°C)	850V-1425V/1275V
Min. input voltage $V_{DC,min}$ / start voltage $V_{DC,start}$	778V/878V
Max. input voltage $V_{DC,max}$	1500V
Max. input current $I_{DC,max}$ (at 25°C / at 50°C)	3000A / 2700A
Max. short-circuit current rating	4300A
Output DC parameters	Value
Nominal AC power at $\cos\phi = 1$ (at 25°C / at 40°C / at 50°C)	2500kVA/2350kVA/2250kVA
Nominal AC power at $\cos\phi = 0.8$ (at 25°C / at 40°C / at 50°C)	2000kW/1880kW/1800kW
Nominal AC IAC, nom = Max. output current $I_{AC,max}$	2624A

C. Economic Analysis

Software like RETScreen and HOMER are used to estimate the economic indicators for financial analysis. Several researchers have utilized RETScreen to investigate the feasibility of a grid-connected PV system [8, 10]. In the present study, 6 economic indicators, namely NPV, LCOE, IRR, SP, EP, and ALCS are estimated using RETScreen. Also, the GHG emission reduction, energy production, and the Capacity Factor (CF) for the proposed systems are determined using RETScreen.

Net Present Value (NPV):

$$NPV = \sum_{n=0}^N \frac{C_n}{(1+r)^n} \quad (7)$$

Levelized Cost Of Energy (LCOE):

$$LCOE = \frac{\text{sum of cost over lifetime}}{\text{s of electricity generated over the lifetime}} \quad (8)$$

Internal Rate of Return (IRR):

$$0 = \sum_{n=0}^N \frac{C_n}{(1+IRR)^n} \quad (9)$$

Simple Payback (SP):

$$SP = \frac{C-IG}{(C_{ener}+C_{capa}+C_{RE}+C_{GHG})-(C_{o\&M}+C_{fuel})} \quad (10)$$

Equity Payback (EP):

$$EP = \sum_{n=0}^N C_n \quad (11)$$

Annual Life Cycle Savings (ALCS):

$$ALCS = \frac{NPV}{\frac{1}{r}\left(1-\frac{1}{(1+r)^N}\right)} \quad (12)$$

GHG Reduction Cost (GRC):

$$GRC = \frac{ALCS}{\Delta_{GHG}} \quad (13)$$

Benefit-Cost ratio (B-C):

$$B - C = \frac{NPV+(1-f_d)C}{(1-f_d)^{FDC}} \quad (14)$$

Capacity Factor (CF):

$$CF = \frac{P_{out}}{P \times 8760} \quad (15)$$

where P_{out} is the energy generated per year, P is the installed capacity, N is the project life in years, C_n is the after-tax cash flow in year n , r is the discount rate, C is the total initial cost of the project, f_d is the debt ratio, B is the total benefit of the project, IG is the incentives and grants, C_{enter} is the annual energy savings or income, C_{appa} is the annual capacity savings or income, C_{RE} is the annual Renewable Energy (RE) production credit income, C_{GHG} is the GHG reduction income, $C_{o\&M}$ is the yearly operation and maintenance costs incurred by the clean energy project, C_{fuel} is the annual cost of fuel, which is zero for renewable projects, and Δ_{GHG} is the annual GHG emission reduction.

III. RESULTS AND DISCUSSION

A. Monthly Solar Radiation

It is well known that the performance of PV modules is influenced by several climate parameters including the meteorological factors of the specific location, and solar irradiance is one of the most essential among them. Several studies utilized various simulation tools to assess the performance simulation of small-scale grid-connected rooftop PV systems. Authors in [28] assessed the feasibility of a 6.4kW grid-connected rooftop PV system in Ujjain, India by using 4 simulation tools. The results indicated that PV*SOL is an easy, fast, and reliable software tool for the simulation of the solar PV system in the selected region. Authors in [29] investigated the performance of a 1MW grid-connected rooftop PV system in an educational institute in India using 3 simulation tools and the actual data and the PVGIS has the lowest average mean bias error and average normalized mean bias error compared to the other two tools (PVSyst and PVWatt). In this paper, the performance of the developed PV systems is estimated by 2 simulation tools, PVGIS and RETScreen. With the PVGIS (Figure 2), the highest solar radiation was recorded in July with a value of 246.892kWh/m² while the minimum was obtained in December with a value of 76.211kWh/m². In the RETScreen simulation tool, the highest value of solar radiation was recorded in July (250.48kWh/m²) as shown in Figure 2. Moreover, it was found the annual solar radiation is 1941 kWh/m² and 2010kWh/m², given by PVGIS and RETScreen

respectively. It is noticed that the PVGIS and RETScreen almost give the same annual value of solar radiation. Authors in [25] found that the estimated values of solar radiation, obtained from PVGIS and RETScreen, are close to the actual data. Based on the value of solar radiation at the selected location, it was found that the solar source of Nahr El-Bared is categorized as excellent (class 5) according to [23]. Therefore, this region is suitable for installing a PV system.

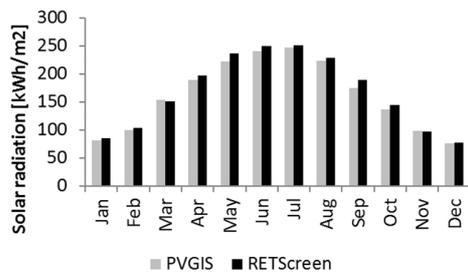


Fig. 2. Average monthly solar radiation at the studied location.

B. Electricity Generation and Capacity Factor

The optimum slope angle and azimuth angle for the fixed-tilt system are estimated with the Photovoltaic Geographical Information System (PVGIS) simulation tool. Several scientific studies have used PVGIS to find the slope angle and azimuth angle for PV systems [30, 31]. Generally, the PVGIS provides the optimum slope and azimuth angles that give the maximum annual global solar radiation for a specific location. The slope and azimuth angles for the selected study were estimated to be 30° and -3° respectively. The economic feasibility of the developed system is investigated with RETScreen. According to [32, 33], solar radiation and the number of clear sunny days are important factors that affect the annual energy exported to the grid by the panel and the Capacity Factor (CF). Table III shows the monthly and annual Electricity Generation (EG) and CF of the proposed systems. For the fixed-tilt system, it is found that the total annual value of EG is 8564.47kWh for mono-Si and 8776.81kWh for poly-Si. The total annual EG for the two-axis tracking system is within the range of 11511.67kWh (mono-Si) to 12100.92kWh (poly-Si). The maximum value of EG is recorded in July for all the proposed systems. It can be concluded that the amount of output power could be increased from 34% to 37% by a two-axis tracking system. Furthermore, it is found that the CF values of both systems vary from 19.01% to 26.43% for the fixed-tilt and two-axis sun-tracking systems. These observations can be supported by the findings of other researchers who analyzed the feasibility of grid-connected PV systems. For instance, authors in [34] found that the CF of the proposed PV system in Oman was within the range of 16-23%. Authors in [35] found that the value of CF of grid-connected PV systems with different sun-tracking modes was within the range of 17.54-27.42%. Authors in [35, 36] concluded that the use of the two-axis instead of the fixed-tilt option significantly increases the generated electricity. Therefore, it can be concluded that the value gotten from the present study for the selected location is compatible with the acceptable values. Consequently, it is technically sustainable to build a grid-connected rooftop PV system in Nahr El-Bared.

TABLE III. ELECTRICITY GENERATION IN KWH AND CAPACITY FACTOR FOR THE DEVELOPED SYSTEMS

Month	Fixed-tilt axis		Two-axis	
	Mono-Si	Poly-Si	Mono-Si	Poly-Si
Jan.	514.14	526.89	663.81	680.27
Feb.	549.70	563.33	663.81	700.90
Mar.	726.28	744.29	663.81	963.47
Apr.	778.32	797.61	1015.02	1040.19
May	861.40	882.76	1261.22	1292.49
Jun.	853.38	874.54	1293.29	1325.35
Jul.	864.42	885.85	1296.47	1328.61
Aug.	852.87	874.02	1246.95	1277.86
Sep.	805.77	825.75	1083.05	1109.90
Oct.	718.62	736.43	970.07	994.12
Nov.	562.21	576.15	730.92	749.05
Dec.	477.36	489.19	623.25	638.70
Annual [kWh]	8564.47	8776.81	11511.67	12100.92
CF [%]	19.17	19.01	26.43	26.21

C. Performance of the Proposed Systems

The performance of the aforementioned sun-tracking systems was evaluated by the estimation of the economic and environmental factors for each system. In this study, the financial parameters (Table IV) inflation rate, discount rate, reinvestment rate, debt ratio, debt interest rate, were considered as input variables for the estimated economic indicators assumed based on [10, 13, 25, 35, 37-40]. In the present study, the system cost is around \$5261 and \$5243 for mono-Si and poly-Si fixed-tilt and \$7365 and \$7347 for mono-Si and poly-Si two-axis tracking systems which are estimated based on recent market data in the country and are consistent with the cost prices available in the literature. In general, the net present value of any investment project can be defined as the difference between the present value of cash inflows and outflows of the project [41]. Therefore, it is clear that in order to calculate the net present value, there must be a discount rate based on which the cash flows associated with the investment are deducted [42, 43]. Therefore, it can be concluded that the difficulty in determining the discount rate that is used as a basis for calculating the net present value has an impact on the investment decision. Additionally, the average credit interest rate can be used as the discount rate, or the interest rate on project loans can be considered as the discount rate [44]. There is an inverse relationship between the discount rate and the current value of the project, which means that choosing a discount rate is very important [45]. The debt ratio is the financial ratio used to assess and measure the leverage of an entity over the relationship between total debt (long-term debt and short-term debt) and total assets [46]. If the ratio is higher than 1, this means that the total liabilities are higher than the total assets, which means that the facility's leverage is high and it faces more financial risks. If the ratio is less than 1, then this means that the total liabilities are less than the assets, which means that the facility is financially sound. Besides, high-interest rates lead to an increase in the future value of the project as the period of investment leads to a future value increase. The results regarding the economic performance of the 5kW grid-connected rooftop PV system for all the developed PV systems are summarized in Table V. The obtained results show that the value of NPV for the proposed systems is positive and makes the project to be financially and

economically feasible [24, 35]. Also, it is found that the proposed projects in the selected location are economically acceptable based on the internal rate of return, which is a measure of a project’s profitability [24, 47]. It is observed that the developed PV project has the longest value of EP of 3.2 years for mono-Si and the lowest one for poly-Si system with a value of 3.1 years. For the fixed-tilt system, the SP values are 6.14 years and 5.97 years for mono-Si and poly-Si respectively. For the two-axis tracking modes, the SP values are 6.24 and 6.07 years respectively for mono-Si and poly-Si. The results indicate that the PV projects in the region make financial sense.

TABLE IV. FINANCIAL PARAMETERS

Factor	Unit	Value
Inflation rate	%	2.5
Discount rate	%	0
Reinvestment rate	%	9
Project life	year	25
Debt ratio	%	50
Debt interest rate	%	0
Debt term	year	20
Electricity export escalation rate	%	5

TABLE V. ECONOMIC PERFORMANCE –CURRENT STUDY

Parameters	Fixed-tilt axis		Two-axis	
	Mono-Si	Poly-Si	Mono-Si	Poly-Si
IRR [%]	15.99	16.14	15.92	16.05
SP [year]	6.14	5.97	6.24	6.07
EP [year]	3.21	3.12	3.26	3.17
NPV [\$]	37658.51	38740.63	51809.73	53294.87
ALCS p\$/year	1506.34	1549.63	2072.39	2131.79
NPV [\$]	-248.79	-249.75	-248.26	-249.20
LCOE [\$/kWh]	0.0246	0.0239	0.0249	0.0243
Performance [%]	82.63	81.90	91.62	93.16

TABLE VI. ECONOMIC PERFORMANCE COMPARISON

Ref.	Country	Size/ load	Tracking mode	LCOE [\$/kWh]
[8]a	Northern Cyprus	6.4kW	Fixed Tilt	0.0194-0.0199
[10]a	Libya	5kW	Fixed Tilt	0.0307-0.0270
[38]a	Norway	2.07kW	Fixed Tilt	0.149-11.632
[48]b	India	6kW	Fixed Tilt	2.944
			Two-axis adjustment tracking	3.165
[49]c	Nigeria	1kW	Two-axis continuous tracking	0.299
			Fixed Tilt	0.097
[50]a	Jordan	7.98kW	Two-axis continuous tracking	0.122
			Fixed Tilt	0.226-0.287
[51]b	India	5.19 kWh/d	Two-axis continuous tracking	0.523-0.732
			Fixed Tilt	0.075
[52]a	Spain	1kW	Two-axis continuous tracking	0.109
			Fixed Tilt	0.09
[53]c	Iraq	4kW	Fixed Tilt	0.120-0.390
[54]a	Europe	4kW	Fixed Tilt	0.09
[55]a	Indonesia	3kWp	Fixed Tilt	0.078
[56]a	Iraq	5kW	Fixed Tilt	0.0239-0.0246
			Two-axis continuous tracking	0.0243-0.0246

CS: Current Study, a: grid-connected PV system, b: stand-alone PV with micro WT hybrid power system, c: Stand-alone PV system

The lowest value of LCOE is found for poly-Si systems as shown in Table V. The LCOE value of the proposed projects is compared with the existing value of small-scale PV systems in the literature. Table VI summarizes the economic analysis of solar projects with various tracking modes in different countries. It is found that the LCOE values of the proposed systems are within the range of the maximum (3.165\$/kWh) and minimum (0.0199\$/kWh) of LCOE values obtained from the literature. It is noticed that the value of LCOE is slightly increased by 4.23% when the two-axis system is used. This increment of LCOE is primarily due to the high cost of the two-axis system. The developed systems provide a very good insight into the economic viability of the project for all regions. Additionally, the obtained results demonstrated that the development of the proposed 5kW PV power system is economically acceptable due to the obtained favorable economic results.

D. The Case of a Solar PV System for Household Electricity Generation

This section aims to evaluate the techno-economic performance of 5kW grid-connected PV systems for rooftop PV systems. Designing the electrical load is an essential part of this section. The energy demand (E_{load}) of the considering household can be estimated by [57]:

$$E_{load} = \sum_{j=1}^{N_{category}} P_j n_j T_j \quad (16)$$

where P_j is the rated power of the j -th kind of household appliance (kW), n_j is the number of the j -th kind of household appliances, T_j is the used hours per day of the j -th kind of household appliance (h/day) and $N_{category}$ is the category number of household appliances.

According to [58, 59], lighting, TV, air conditioning, refrigerator, electric cooker, washing machine, water heater, and small power appliances are the most important energy-consuming household devices. In this case, the chosen house is equipped with efficient appliances. The characteristics of this house are:

- Room number: three rooms.
- Lighting: rooms, kitchen, toilet, bathroom, courtyard, and corridor.
- Appliances: refrigerator, TV, water heater, computer, laptop, washing machine, air conditioner, and fan.

The total energy consumption for the chosen house is estimated as 20kWh/day. As mentioned previously, the total annual energy generating from both PV systems is varied from 8564.47kWh to 8776.81kWh and 11511.67kWh to 12100.92kWh for fixed-tilt and two-axis tracking systems, respectively.

Figure 3 shows the energy consumption covered by the proposed PV systems and grid for each month. It is observed that around 13% out of the total energy consumption is covered by the grid and the remaining is supplied by a fixed-tilt PV system, which is about 87% for the winter season. This is due to a heavy electrical load connected in the system like a water heater. For the rest of the year, the PV systems could cover all

the energy consumption. Also, due to the length of the day that depends on sun altitude, the geographical latitude of the location, declination angle of the sun, and hour angle, the amount of energy production from the PV system will be increased.

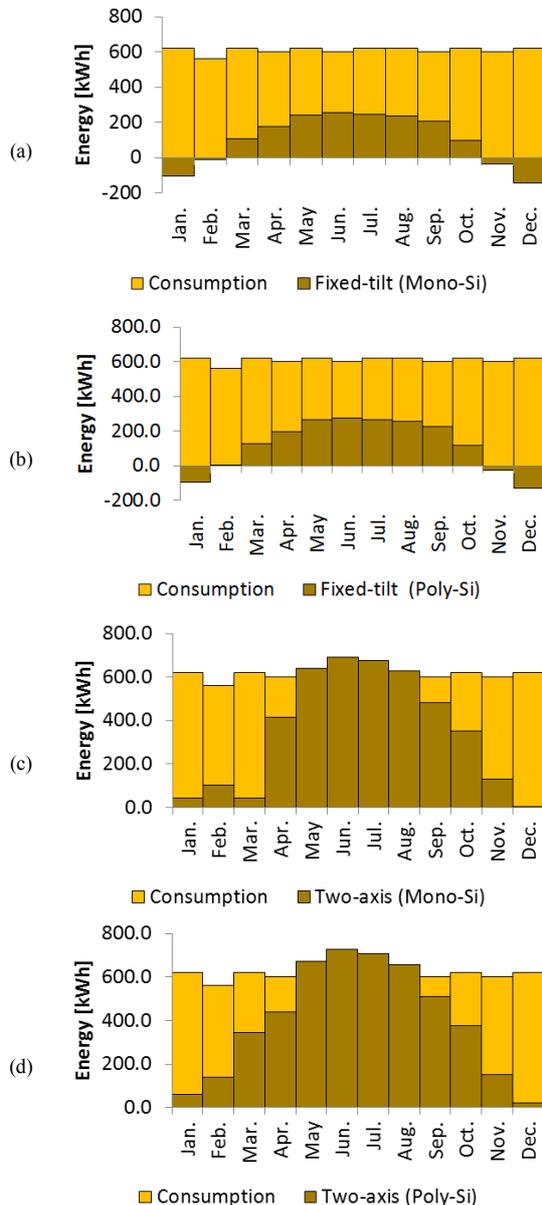


Fig. 3. Monthly variation of energy generation and energy consumption.

IV. LIMITATIONS AND CONCLUSIONS

Due to rising electricity tariff rates and the potential reduction of the dependency on domestic power generators in Lebanon, installing PV systems has become increasingly attractive for residential consumers, a trend that is supported by the reviewed studies. Before starting the main conclusions in the present study, it is essential to acknowledge the limitations

of this work. First, the financial parameters were assumed based on historical values in the literature. Second, the influence of various parameters such as dust, irradiation intensity, air temperature, and relative humidity were neglected due to limitations of the RETScreen software. Third, the cost of the proposed projects was estimated based on the existing costs in the literature. The findings from the present study showed that the annual value of solar radiation for the selected region is 2010kWh/m². Based on annual global solar irradiation, the analysis indicates that the selected region in Lebanon has a potential for the distribution of PV power systems in residential applications. Moreover, the annual energy output showed that the 5kW grid-connected PV system in Nahr El-Bared was within the range of 8564.47-8776.81kWh and 11511.67-12100.92kWh respectively when fixed tilt and two-tracking systems have been used. Based on this analysis, it was found that the highest energy consumption that can be covered by PV systems is recorded in spring and summer seasons.

The average energy production cost ranged in 0.0239-0.0249\$/kWh for all the proposed systems. The electricity price depends on the amount of energy consumption, i.e. the energy cost calculation starts from 0.025\$/kWh for energy consumption of 0-100kWh, 0.04\$/kWh for 100-300kWh, 0.0584\$/kWh for 300-400kWh, 0.0875\$/kWh for 400-500kWh, and 0.146\$/kWh over 500kWh. Therefore, the energy production cost of the proposed systems is competitive with the electricity tariffs. Furthermore, it was found that the difference between the energy production cost from a fixed-tilt system and from a two-axis tracking system is lower than 2%. It can be concluded that a 5kW grid-connected PV tracking system is economically justifiable.

The results of this paper demonstrate that a small-scale grid-connected rooftop PV system in the Nahr El-Bared has the potential to solve the electricity issue in the area, reduce the consumption of fossil fuels and environmental pollution by minimizing the emissions of CO₂. The present study tried to increase the awareness about utilizing PV sun-tracking systems and the feasibility of small-scale grid-connected rooftop PV systems in the studied region. The results of this research can help the investors in the energy and building sectors and accelerate an informed transition towards a more sustainable future.

ACKNOWLEDGMENTS

The authors would like to thank the Faculty of Engineering and especially the Mechanical Engineering Department at Near East University for their support and encouragement.

REFERENCES

[1] F. Chermat, M. Khemliche, A. E. Badoud, and S. Latreche, "Techno-Economic Feasibility Study of Investigation of Renewable Energy System for Rural Electrification in South Algeria," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3421-3426, Oct. 2018, <https://doi.org/10.48084/etasr.2253>.

[2] A. Q. Jakhriani, A. R. Jatoui, M. M. Jakhriani, A. N. Laghari, and S. A. Kamboh, "Appraisal of Climatic Conditions for Potential Solar Energy Applications in Nawabshah and Quetta Cities," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3711-3714, Feb. 2019, <https://doi.org/10.48084/etasr.2472>.

- [3] E. A. Al-Ammar, N. H. Malik, and M. Usman, "Application of using Hybrid Renewable Energy in Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 1, no. 4, pp. 84–89, Aug. 2011, <https://doi.org/10.48084/etasr.33>.
- [4] H. Schnitzer, C. Brunner, and G. Gwehenberger, "Minimizing greenhouse gas emissions through the application of solar thermal energy in industrial processes," *Journal of Cleaner Production*, vol. 15, no. 13–14, pp. 1271–1286, Sep. 2007, <https://doi.org/10.1016/j.jclepro.2006.07.023>.
- [5] A. Shahsavari and M. Akbari, "Potential of solar energy in developing countries for reducing energy-related emissions," *Renewable and Sustainable Energy Reviews*, vol. 90, pp. 275–291, Jul. 2018, <https://doi.org/10.1016/j.rser.2018.03.065>.
- [6] A. Shahsavari, F. T. Yazdi, and H. T. Yazdi, "Potential of solar energy in Iran for carbon dioxide mitigation," *International Journal of Environmental Science and Technology*, vol. 16, no. 1, pp. 507–524, Jan. 2019, <https://doi.org/10.1007/s13762-018-1779-7>.
- [7] S. Adnan, A. Hayat Khan, S. Haider, and R. Mahmood, "Solar energy potential in Pakistan," *Journal of Renewable and Sustainable Energy*, vol. 4, no. 3, May 2012, Art. no. 032701, <https://doi.org/10.1063/1.4712051>.
- [8] Y. Kassem, R. Al Zoubi, and H. Gokcekus, "The Possibility of Generating Electricity Using Small-Scale Wind Turbines and Solar Photovoltaic Systems for Households in Northern Cyprus: A Comparative Study," *Environments*, vol. 6, no. 4, Apr. 2019, Art. no. 47, <https://doi.org/10.3390/environments6040047>.
- [9] Y. Kassem, H. Gokcekus, and H. Camur, "Economic assessment of renewable power generation based on wind speed and solar radiation in urban regions," *Global Journal of Environmental Science and Management*, vol. 4, no. 4, pp. 465–482, Oct. 2018, <https://doi.org/10.22034/gjesm.2018.04.007>.
- [10] Y. Kassem, H. Camur, and R. A. F. Aateg, "Exploring Solar and Wind Energy as a Power Generation Source for Solving the Electricity Crisis in Libya," *Energies*, vol. 13, no. 14, Jan. 2020, Art. no. 3708, <https://doi.org/10.3390/en13143708>.
- [11] B. O. Sigrin and M. E. Mooney, "Rooftop Solar Technical Potential for Low-to-Moderate Income Households in the United States," National Renewable Energy Lab. (NREL), Golden, CO, United States, NREL/TP-6A20-70901, Apr. 2018, <https://doi.org/10.2172/1434891>.
- [12] Y. Kassem, H. Camur, and O. a. M. Abughinda, "Solar Energy Potential and Feasibility Study of a 10MW Grid-connected Solar Plant in Libya," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5358–5366, Aug. 2020, <https://doi.org/10.48084/etasr.3607>.
- [13] Y. Kassem, H. Gökçekuş, and W. Janbein, "Predictive model and assessment of the potential for wind and solar power in Rayak region, Lebanon," *Modeling Earth Systems and Environment*, Jul. 2020, <https://doi.org/10.1007/s40808-020-00866-y>.
- [14] K. Abdeladim, S. Bouchakour, A. H. Arab, S. O. Amrouche, and N. Yassaa, "Promotion of renewable energy in some MENA region countries," *IOP Conference Series: Earth and Environmental Science*, vol. 154, May 2018, Art. no. 012003, <https://doi.org/10.1088/1755-1315/154/1/012003>.
- [15] G. Al Zohbi, P. Hendrick, and P. Bouillard, "Wind characteristics and wind energy potential analysis in five sites in Lebanon," *International Journal of Hydrogen Energy*, vol. 40, no. 44, pp. 15311–15319, Nov. 2015, <https://doi.org/10.1016/j.ijhydene.2015.04.115>.
- [16] G. Al Zohbi, P. Hendrick, C. Renier, and P. Bouillard, "The contribution of wind-hydro pumped storage systems in meeting Lebanon's electricity demand," *International Journal of Hydrogen Energy*, vol. 41, no. 17, pp. 6996–7004, May 2016, <https://doi.org/10.1016/j.ijhydene.2016.01.028>.
- [17] Y. Kassem, H. Gokcekus, and M. Zeitoun, "Modeling of techno-economic assessment on wind energy potential at three selected coastal regions in Lebanon," *Modeling Earth Systems and Environment*, vol. 5, no. 3, pp. 1037–1049, Sep. 2019, <https://doi.org/10.1007/s40808-019-00589-9>.
- [18] H. Gokcekus, Y. Kassem, and M. Al Hassan, "Evaluation of Wind Potential at Eight Selected Locations in Northern Lebanon Using Open Source Data," *International Journal of Applied Engineering Research*, vol. 14, no. 11, pp. 2789–2794, 2019.
- [19] Y. Kassem, H. Gokcekus, M. M. Mizran, and S. M. Alsayas, "Evaluation of the Wind Energy Potential in Lebanon's Coastal Regions using Weibull Distribution Function," *International Journal of Engineering Research and Technology*, vol. 12, no. 6, pp. 784–792, 2019.
- [20] S. Tannous, R. Manneh, H. Harajli, and H. El Zakhem, "Comparative cradle-to-grave life cycle assessment of traditional grid-connected and solar stand-alone street light systems: A case study for rural areas in Lebanon," *Journal of Cleaner Production*, vol. 186, pp. 963–977, Jun. 2018, <https://doi.org/10.1016/j.jclepro.2018.03.155>.
- [21] G. Ibarra-Berastegi, A. Ulazia, J. Saenz, and S. J. Gonzalez-Roji, "Evaluation of Lebanon's Offshore-Wind-Energy Potential," *Journal of Marine Science and Engineering*, vol. 7, no. 10, Oct. 2019, Art. no. 361, <https://doi.org/10.3390/jmse7100361>.
- [22] A. E. H. Berjawi, S. Najem, G. Faour, and C. Abdallah, "Assessing Solar PV's Potential in Lebanon," *Issam Fares Institute for Public Policy and International Affairs, working paper 17*, Aug. 2017, <https://doi.org/10.13140/RG.2.2.17493.32485>.
- [23] R. Pravalie, C. Patriche, and G. Bandoc, "Spatial assessment of solar energy potential at global scale. A geographical approach," *Journal of Cleaner Production*, vol. 209, pp. 692–721, Feb. 2019, <https://doi.org/10.1016/j.jclepro.2018.10.239>.
- [24] A. B. Owlabi, B. E. K. Nsafon, J. W. Roh, D. Suh, and J.-S. Huh, "Validating the techno-economic and environmental sustainability of solar PV technology in Nigeria using RETScreen Experts to assess its viability," *Sustainable Energy Technologies and Assessments*, vol. 36, Dec. 2019, Art. no. 100542, <https://doi.org/10.1016/j.seta.2019.100542>.
- [25] Y. Kassem, H. Camur, and S. M. A. Alhuoti, "Solar Energy Technology for Northern Cyprus: Assessment, Statistical Analysis, and Feasibility Study," *Energies*, vol. 13, no. 4, Jan. 2020, Art. no. 940, <https://doi.org/10.3390/en13040940>.
- [26] K. Belkilani, A. Ben Othman, and M. Besbes, "Assessment of global solar radiation to examine the best locations to install a PV system in Tunisia," *Applied Physics A*, vol. 124, no. 2, Jan. 2018, Art. no. 122, <https://doi.org/10.1007/s00339-018-1551-3>.
- [27] K. Gairaa and Y. Bakelli, "Solar Energy Potential Assessment in the Algerian South Area: Case of Ghardaïa Region," *Journal of Renewable Energy*, vol. 2013, Mar. 2013, Art. no. 496348, <https://doi.org/10.1155/2013/496348>.
- [28] C. Dondariya *et al.*, "Performance simulation of grid-connected rooftop solar PV system for small households: A case study of Ujjain, India," *Energy Reports*, vol. 4, pp. 546–553, Nov. 2018, <https://doi.org/10.1016/j.egyr.2018.08.002>.
- [29] S. Thotakura *et al.*, "Operational performance of megawatt-scale grid integrated rooftop solar PV system in tropical wet and dry climates of India," *Case Studies in Thermal Engineering*, vol. 18, Apr. 2020, Art. no. 100602, <https://doi.org/10.1016/j.csite.2020.100602>.
- [30] R. Abdallah, E. Natsheh, A. Juaidi, S. Samara, and F. Manzano-Agugliaro, "A Multi-Level World Comprehensive Neural Network Model for Maximum Annual Solar Irradiation on a Flat Surface," *Energies*, vol. 13, no. 23, Jan. 2020, Art. no. 6422, <https://doi.org/10.3390/en13236422>.
- [31] N. Bailek, K. Bouchouicha, N. Aoun, M. EL-Shimy, B. Jamil, and A. Mostafaiepour, "Optimized fixed tilt for incident solar energy maximization on flat surfaces located in the Algerian Big South," *Sustainable Energy Technologies and Assessments*, vol. 28, pp. 96–102, Aug. 2018, <https://doi.org/10.1016/j.seta.2018.06.002>.
- [32] A. Mehmood, F. A. Shaikh, and A. Waqas, "Modeling of the solar photovoltaic systems to fulfill the energy demand of the domestic sector of Pakistan using RETSCREEN software," in *International Conference and Utility Exhibition on Green Energy for Sustainable Development*, Pattaya, Thailand, Mar. 2014, pp. 1–7.
- [33] A. Khandelwal and V. Shrivastava, "Viability of grid-connected solar PV system for a village of rajasthan," in *International Conference on Information, Communication, Instrumentation and Control*, Indore, India, Aug. 2017, pp. 1–6, <https://doi.org/10.1109/ICOMICON.2017.8279175>.
- [34] H. A. Kazem and T. Khatib, "Techno-economical assessment of grid connected photovoltaic power systems productivity in Sohar, Oman,"

- Sustainable Energy Technologies and Assessments*, vol. 3, pp. 61–65, Sep. 2013, <https://doi.org/10.1016/j.seta.2013.06.002>.
- [35] K. Mohammadi, M. Naderi, and M. Saghaifar, "Economic feasibility of developing grid-connected photovoltaic plants in the southern coast of Iran," *Energy*, vol. 156, pp. 17–31, Aug. 2018, <https://doi.org/10.1016/j.energy.2018.05.065>.
- [36] M. A. Vaziri Rad, A. Toopshekan, P. Rahdan, A. Kasaeian, and O. Mahian, "A comprehensive study of techno-economic and environmental features of different solar tracking systems for residential photovoltaic installations," *Renewable and Sustainable Energy Reviews*, vol. 129, Sep. 2020, Art. no. 109923, <https://doi.org/10.1016/j.rser.2020.109923>.
- [37] M. EL-Shimy, "Viability analysis of PV power plants in Egypt," *Renewable Energy*, vol. 34, no. 10, pp. 2187–2196, Oct. 2009, <https://doi.org/10.1016/j.renene.2009.01.010>.
- [38] M. S. Adaramola, "Techno-economic analysis of a 2.1kW rooftop photovoltaic-grid-tied system based on actual performance," *Energy Conversion and Management*, vol. 101, pp. 85–93, Sep. 2015, <https://doi.org/10.1016/j.enconman.2015.05.038>.
- [39] K. Y. Kebede, "Viability study of grid-connected solar PV system in Ethiopia," *Sustainable Energy Technologies and Assessments*, vol. 10, pp. 63–70, Jun. 2015, <https://doi.org/10.1016/j.seta.2015.02.003>.
- [40] Y. Himri, M. Merzouk, N. Kasbadji Merzouk, and S. Himri, "Potential and economic feasibility of wind energy in south West region of Algeria," *Sustainable Energy Technologies and Assessments*, vol. 38, Apr. 2020, Art. no. 100643, <https://doi.org/10.1016/j.seta.2020.100643>.
- [41] M. Hirschey, *Fundamentals of Managerial Economics*. London, UK: Cengage Learning, 2008.
- [42] A. Griffith and P. Watson, *Construction Management: Principles and Practice*. Houndmills, UK: Red Globe Press, 2003.
- [43] L. Booth, W. S. Cleary, and I. Rakita, *Introduction to Corporate Finance*, 5th edition. New York, USA: Wiley, 2020.
- [44] S. Lumby and C. Jones, *Corporate Finance: Theory and Practice*, 7th edition. London, UK: Cengage Learning EMEA, 2003.
- [45] S. Park, "Optimal Discount Rates for Government Projects," *ISRN Economics*, vol. 2012, Jul. 2012, Art. no. 982093, <https://doi.org/10.5402/2012/982093>.
- [46] I. El-Sayed Ebaid, "The impact of capital-structure choice on firm performance: empirical evidence from Egypt," *The Journal of Risk Finance*, vol. 10, no. 5, pp. 477–487, Jan. 2009, <https://doi.org/10.1108/15265940911001385>.
- [47] S. Rehman, M. A. Ahmed, M. H. Mohamed, and F. A. Al-Sulaiman, "Feasibility study of the grid connected 10MW installed capacity PV power plants in Saudi Arabia," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 319–329, Dec. 2017, <https://doi.org/10.1016/j.rser.2017.05.218>.
- [48] S. Sinha and S. S. Chandel, "Analysis of fixed tilt and sun tracking photovoltaic-micro wind based hybrid power systems," *Energy Conversion and Management*, vol. 115, pp. 265–275, May 2016, <https://doi.org/10.1016/j.enconman.2016.02.056>.
- [49] A. Bahrami, C. O. Okoye, and U. Atikol, "Technical and economic assessment of fixed, single and dual-axis tracking PV panels in low latitude countries," *Renewable Energy*, vol. 113, pp. 563–579, Dec. 2017, <https://doi.org/10.1016/j.renene.2017.05.095>.
- [50] B. Hammad, A. Al-Sardeah, M. Al-Abed, S. Nijmeh, and A. Al-Ghandoor, "Performance and economic comparison of fixed and tracking photovoltaic systems in Jordan," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 827–839, Dec. 2017, <https://doi.org/10.1016/j.rser.2017.05.241>.
- [51] S. Bhakta and V. Mukherjee, "Techno-economic viability analysis of fixed-tilt and two axis tracking stand-alone photovoltaic power system for Indian bio-climatic classification zones," *Journal of Renewable and Sustainable Energy*, vol. 9, no. 1, Jan. 2017, Art. no. 015902, <https://doi.org/10.1063/1.4976119>.
- [52] D. L. Talavera, E. Munoz-Ceron, J. P. Ferrer-Rodriguez, and P. J. Perez-Higueras, "Assessment of cost-competitiveness and profitability of fixed and tracking photovoltaic systems: The case of five specific sites," *Renewable Energy*, vol. 134, pp. 902–913, Apr. 2019, <https://doi.org/10.1016/j.renene.2018.11.091>.
- [53] R. K. Fakher Alfahed, S. Oudah, and K. Al-jabori, "Electrification of a Rural Home by Solar Photovoltaic System in Haur Al-Hammar of Iraq," *Journal of Energy Management and Technology*, vol. 3, no. 3, pp. 30–40, Sep. 2019, <https://doi.org/10.22109/jemt.2018.134845.1102>.
- [54] G. Martinopoulos, "Are rooftop photovoltaic systems a sustainable solution for Europe? A life cycle impact assessment and cost analysis," *Applied Energy*, vol. 257, p. 114035, Jan. 2020, <https://doi.org/10.1016/j.apenergy.2019.114035>.
- [55] E. Tarigan, "Rooftop PV System Policy and Implementation Study for a Household in Indonesia," *International Journal of Energy Economics and Policy*, vol. 10, no. 5, pp. 110–115, Aug. 2020.
- [56] H. H. Abada, A. N. Al-Shamani, and M. D. Faisal, "A grid connected PV system for a rural resident in Al-Najaf, Iraq," *AIP Conference Proceedings*, vol. 2290, no. 1, Dec. 2020, Art. no. 050056, <https://doi.org/10.1063/5.0027826>.
- [57] J. Vishnupriyan and P. S. Manoharan, "Prospects of hybrid photovoltaic-diesel standalone system for six different climate locations in Indian state of Tamil Nadu," *Journal of Cleaner Production*, vol. 185, pp. 309–321, Jun. 2018, <https://doi.org/10.1016/j.jclepro.2018.03.061>.
- [58] B. K. Das, N. Hoque, S. Mandal, T. K. Pal, and M. A. Raihan, "A techno-economic feasibility of a stand-alone hybrid power generation for remote area application in Bangladesh," *Energy*, vol. 134, pp. 775–788, Sep. 2017, <https://doi.org/10.1016/j.energy.2017.06.024>.
- [59] A. A. Y. Barzanchi, S. A. Kareem, and N. M. H. Bayaty, "Solar Power for Households in Iraq," *Eurasian Journal of Science & Engineering*, vol. 6, no. 2, pp. 31–38, 2020.

Performance Improvement for Medical Image Transmission Systems using Turbo-Trellis Coded Modulation (TTCM)

Menezla Fayssal

Department of Technology and Science, Nour Bachir
University Center of El-Bayadh, Algeria and
LEPO Laboratory, Djillali Liabès University, Algeria
menezla@yahoo.fr

Mahdjoub Zoubir

LEPO Laboratory
Djillali Liabès University
Sidi Bel-Abbès, Algeria
z.mahdjoub@yahoo.fr

Bendelhoum Mohammed Sofiane

Department of Technology and Science
Nour Bachir University Center of El-BayadhEl-Bayadh
El-Bayadh, Algeria
Bendelhoum_med@yahoo.fr

Mohamed Reda Lahcene

Department of Technology and Science
Nour Bachir University Center of El-BayadhEl-Bayadh
El-Bayadh, Algeria
lahceneredal@gmail.com

Abdelhalim Zerroug

Sonatrach Oil and Gas Company
Algeria
eloual@yahoo.fr

Abstract-Digital images have become an essential working tool in several areas such as the medical field, the satellite and astronomical field, film production, etc. The efficiency of a transmission system to exchange digital images is crucial to allow better and accurate reception. Generally, transmitted images are infected with noise. In the medical field, this noise makes the process of diagnosing difficult. To eliminate the transmission errors, an Error Correcting Code (ECC) can be used with the aim to guarantee excellent reception of the images and allowing a good diagnosis. In this paper, source and channel encoding/decoding functions are studied during medical image transmission (LUNG). At first, the Turbo-Code (TC) is used as ECC and subsequently the Turbo-Trellis Coded Modulation (TTCM). The simulation results represent the curves giving the Bit Error Rate (BER) as a function of the signal to noise ratio (E_b/N_0). In order to compare these two systems properly, the MSSIM (Mean Structural Similarity) parameter was used. The obtained results showed the effect and importance of ECC on the transmission of medical images using TTCM which gave better results compared to TC with regard to increasing the performance of the transmission system by eliminating transmission noise.

Keywords-channel coding; turbo-codes; TTCM; MSSIM; medical image

I. INTRODUCTION

Given their importance and their massive use, digital images have become an essential mean of communication.

Corresponding author: Menezla Fayssal

Medical images are widely used as a diagnostic tool and in biomedical research [1]. Sometimes it is necessary to transmit such images to distant doctors to have their opinions/diagnosis. The transmitted images are very sensitive to transmission errors, so it is very important to protect them against transmission noise in order to have on reception a sharp and clear image identical to the original one, allowing doctors to make a good diagnosis. As a solution to the noise problem, Error Correcting Code (ECC) can be added to the medical image transmission chain in order to eliminate the effect of noise during transmission. Such codes are: block codes, convolutional codes, LDPC codes [2], Turbo-Code (TC) and Turbo-Trellis Coded Modulation (TTCM). TC and ECC allow approaching the theoretical limit of correction predicted by Shannon more than 60 years ago [3-5]. More recently, a new family of ECCs, the TTCM [6, 7], was introduced.

In this article, we first explain the basic principles of TC and TTCM, then we will propose the model of our digital image transmission chain, in which we will transmit a medical image of a lung, using TC and TTCM in order to make a comparison between the two systems and propose the best solution for noisy transmissions. To evaluate our simulation results, additively to BER and signal to noise ratio (E_b/N_0) parameters, we will use the MSSIM (Mean Structural Similarity) parameter to see and analyze the results in a subjective and objective way. The image [8] used in this article shows infected lungs with the Corona virus.

II. TURBO-CODE

The TC corrects error produced by transmission noise. This code, invented and presented by Claude-Berrou [4], is obtained by concatenation of two or more ECCs of low complexity [5], allowing to approach the theoretical limit of correction. Their decoding uses an iterative (or turbo) process. In the principal plan of TC [9] the input binary message of length k is encoded in its natural order and in a permuted order, by two coders called C_1 and C_2 . The two elementary coders are identical, but this is not necessary. In our example, the performance of the natural encoding, without punching is $1/3$, for each source bit (d_k) three bits (x, y_1, y_2) are sent on the channel [7]. Turbo-decoding is carried out according to the principle of iterative decoding [10], based on the use of SISO (Soft-Input, Soft-Output) decoders which exchange information from each other's reliability Z_k [11]. These are called extrinsic information, through a feedback, to improve the correction over the iterations. The circuit of a turbo-decoder is constituted by cascading P modules, corresponding to the P identical decoding iterations and its structure is perfectly modular [9]. The input of the P^{th} module consists of the properly received delayed sequences $\{x_k\}_{P-1}, \{y_k\}_{P-1}$ and the sequence $\{Z_k\}_{P-1}$ of the feedback generated by the $(P-1)^{th}$ module.

III. TURBO-TRELLIS CODED MODULATION

TTCM is an ECC introduced by Robertson and Wörz in 1995 [12, 13]. This code is based on the concatenation of two Trellis Coded Modulations (TCM). The TCM was introduced in the early '80s [14] and is based on the joint optimization of error correcting coding and modulation. Coding is done directly in the signal space, so the ECC and the binary signal code of the modulation can be represented all along using a single trellis. The optimization criterion for a TCM consists in maximizing the minimum Euclidean distance between two coded sequences. In the TTCM scheme each MCT coder consists of a recursive systematic convolutional coder, or CSR coder, of efficiency $q/(q+1)$ and of a modulation without order memory $Q = 2^{q+1}$ [15]. Binary symbols from the source are grouped by q -bit symbols. These symbols are coded by the first MCT in the order they are issued by the source and by the second after interleaving.

IV. SIMULATION RESULTS

To eliminate the noise introduced by the channel and to improve the reception quality of a given communication system, an error correction coder must be used [9, 16-20]. The aim of the conducted simulation is to show the effect of the TC and TTCM on a medical image transmission chain in a Gaussian canal and to make a comparison between the performance of these two codes. In all simulations we use code of $1/2$ return and MAQ16 modulation to have the same spectral efficiency in the two transmission chains. The spectral efficiency of a transmission is given by [9, 16]:

$$\eta = R \log_2 M \quad (\text{bit/s/Hz}) \quad (1)$$

So in both cases of our transmission chain, spectral efficiency is given by:

$$\eta = 1/2 \log_2 16 = 2\text{bit/s/Hz} \quad (2)$$

The results are obtained by evaluating the signal to noise ratio (E_b/N_0) for a number of iterations. In each case the (BER) and the Structural Similarity Index (SSIM) are calculated. Two ways are generally used to measure the quality of degraded images, the subjective methods and the objective methods. In this paper, besides the evaluation criteria, we will use a new criterion, which is the SSIM. This criterion presents the similarity and compares the luminosity, the contrast, and the structure between each pair of vectors of the two images (original and received). The SSIM between two signals x and y is given by [9]:

$$SSIM(x,y) = L(x,y), C(x,y) \text{ and } S(x,y) \quad (3)$$

where $L(x,y)$ represents the comparison of the brightness between the original and the received image, $C(x,y)$ represents the comparison of the contrast between the original and the received image, and $S(x,y)$ represents the comparison of the structure between the original image and the received image.

The proposed simulation algorithm for the image transmission chain coded with TC and TTCM is presented in Figure 1.

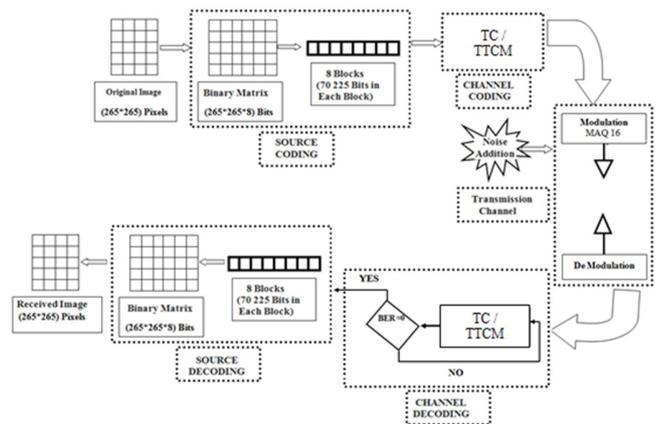


Fig. 1. Block diagram of the image transmission chain.

Figure 2 shows the effect of TC in the transmission chain of the image. The obtained simulation results represent the evaluation of the BER function of the signal-to-noise ratio in a image transmission in a Gaussian channel using TC with rate $R=1/2$ and MAQ16 modulation.

Figure 3 shows the effect of TTCM in the transmission chain. The obtained simulation results represent the evaluation of the BER function of the signal-to-noise ratio in a medical image transmission in a Gaussian Channel using TTCM with rate $R=1/2$ and MAQ16 modulation.

Figure 4 shows the performance comparison between the TC whith $\eta=2$ and TTCM whith $\eta=2$ in the transmission chain.

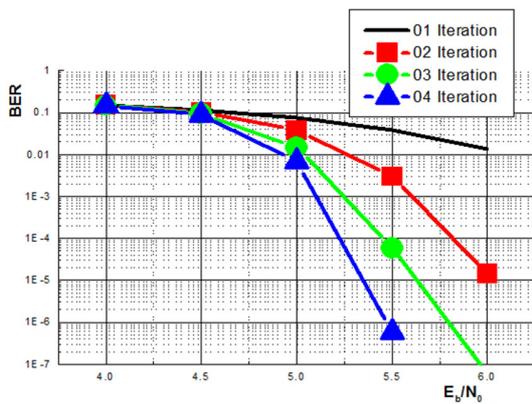


Fig. 2. The influence of TC on the transmission of a medical image in a Gaussian channel with $\eta= 2\text{bit/s/Hz}$.

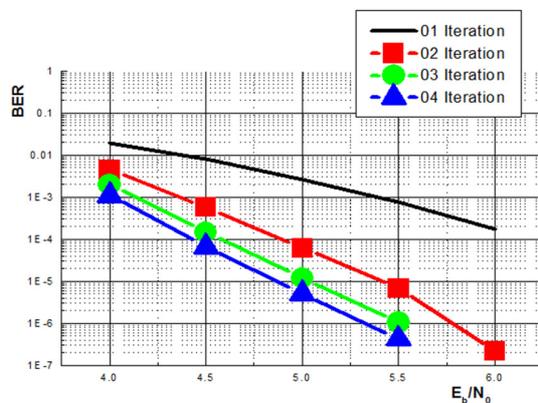


Fig. 3. The influence of TTCM on the transmission of a medical image in a Gaussian channel with $\eta= 2\text{bit/s/Hz}$.

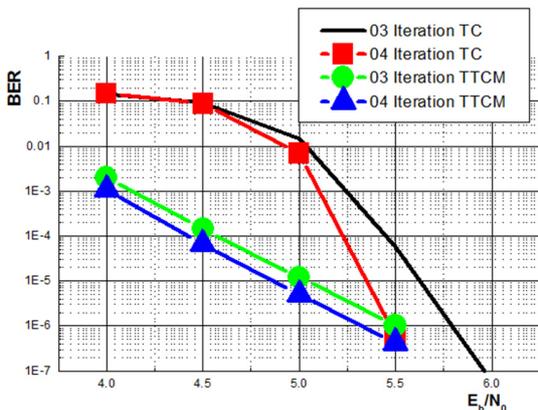


Fig. 4. Comparison between the TC and TTCM on the transmission of a medical image in a Gaussian channel with $\eta= 2\text{bit/s/Hz}$.

The BER measurement gives a numerical value on the damage, but it does not describe its type, it does not quite represent the quality perceived by human observers. For medical imaging applications, where the degraded images must eventually be examined by experts, traditional evaluation remains insufficient. For this reason, objective approaches are needed to assess the medical imaging quality [21]. A new paradigm is evaluated to estimate the quality of medical

images. For this reason and to have more details on our system the SSIM will be used as indicated above. Figures 5 and 6 show the effect of TC and TTCM in the transmission chain of a medical image. The obtained simulation results represent the evaluation of the SSIM function of the signal-to-noise ratio in a medical image transmission in a Gaussian channel using TC and TTCM with $R=1/2$ rate and MAQ16 modulation. Figure 7 shows the comparison between the performance of the TC and TTCM. The results are represented in detail in Tables I and II and the histograms in Figures 8 and 9.

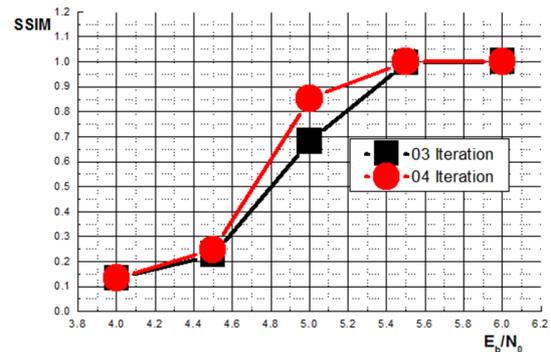


Fig. 5. The obtained SSIM using TC on the transmission of a medical image in a Gaussian channel with $\eta= 2\text{bit/s/Hz}$.

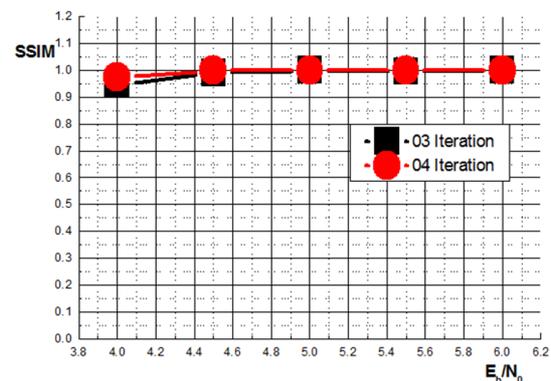


Fig. 6. The obtained SSIM using TTCM on the transmission of a medical image in a Gaussian channel with $\eta= 2\text{bit/s/Hz}$.

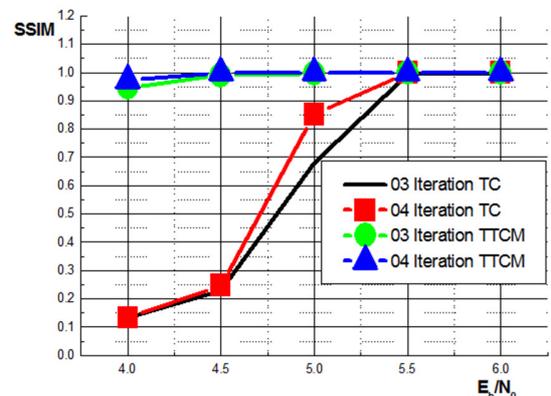


Fig. 7. Comparison of the SSIM between the TC and the TTCM on the transmission of a medical image in the Gaussian channel with $\eta= 2\text{bit/s/Hz}$.

TABLE I. OBTAINED RESULTS OF BER AS A FUNCTION OF THE SIGNAL TO NOISE RATIO AND THE NUMBER OF ITERATIONS

E_b/N_0 (dB)	Iteration No.	BER	
		TC	TTCM
4	1	0.150801	0.0196
	2	0.1494215	0.0046
	3	0.14854397	0.0020
	4	0.148309	0.0011
4.5	1	0.11651121	0.00797721609113564
	2	0.10195265	0.000585083659665361
	3	0.0946173	0.000145781416874333
	4	0.08811321	6.62157351370595 e ⁻⁰⁵
5	1	0.07558384	0.002655927376290500
	2	0.03763973	6.20327518689925 e ⁻⁰⁵
	3	0.01492702	1.19971520113920 e ⁻⁰⁵
	4	0.006852973	5.00177999288003 e ⁻⁰⁶
5.5	1	0.038408685	0.000749875400498398
	2	0.002993948	6.63937344250623 e ⁻⁰⁶
	3	0.000058739766	1.03239587041652 e ⁻⁰⁶
	4	0.	4.22997508009968 e ⁻⁰⁷
6	1	0.013770025	0.000176913492346031
	2	0.000014239943	2.13599145603418 e ⁻⁰⁷
	3	0.	0.
	4	0.	0.

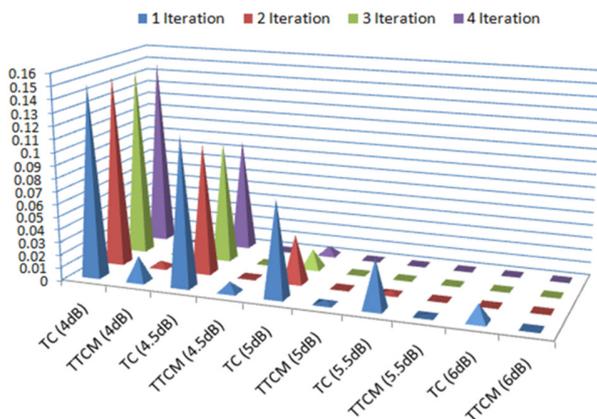


Fig. 8. Graphical representation of the data in Table I.

V. COMMENTS AND INTERPRETATION OF THE OBTAINED RESULTS

From the simulation results, it can be noticed that the role of the ECCs is very important in the transmission of images. The simulation results show that by using transmission systems with TTCM, there is a coding gain compared to transmission systems that use TC. With TTCM, the BER and SSIM results as functions of SNR are quite satisfactory. In the case of TTCM it is more adequate to remain at the 3rd iteration, allowing the receiver to have an easy structure to put in. The decoding will be simpler and the time shorter leading to optimal results. To see the importance of our work and in order to justify the reliability of our system, we have compared the obtained results with the ones obtained in [21, 22]. In these articles, the authors chose joint encoding with TCM with a spectral efficiency of $\eta = 2$ bit/s/Hz. It can be said that the obtained results of the current study have a significant coding gain and also we arrived at SSIM = 1 (100%) which is not the case in the cited articles.

TABLE II. OBTAINED RESULTS OF SSIM AS A FUNCTION OF THE SIGNAL TO NOISE RATIO AND THE NUMBER OF ITERATIONS

E_b/N_0 (dB)	Iteration No.	SSIM	
		TC	TTCM
4	3	0.132 13.2%	0.947 94.7%
	4	0.133 13.3%	0.973 97.3%
4.5	3	0.227 22.7%	0.989 98.9%
	4	0.247 24.7%	0.999 99.9%
5	3	0.681 68.1%	0.996 99.6%
	4	0.852 85.2%	1 100%
5.5	3	0.995 99.5%	1 100%
	4	1 100%	1 100%
6	3	1 100%	1 100%
	4	1 100%	1 100%

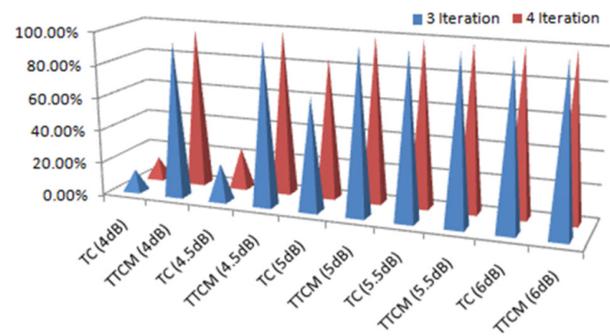


Fig. 9. Graphical representation of the data in Table II.

VI. CONCLUSION

The main objective of our work is to study the performance of TC and TTCM when transmitting medical images over a Gauss channel. The simulation results obviously depict that our proposed transmission scheme performs better for the transmission of digital images and BER is significantly decreased (BER=0). This efficiency leads to the best reception quality of the digital images, 100% close to the transmitted ones (SSIM=1). Moreover, we demonstrated that the TTCM implementation outperforms the TC. By using TTCM, it is suggested to stop the process at the 3rd iteration. This allows having a simple architecture of the decoder and optimizes the decoding process in terms of time and complexity. The obtained results make the proposed transmission scheme a strong candidate when transmitting medical images, allowing medical staff to have best quality at the reception and facilitate the diagnosis process. The obtained results are better than the results found in [21].

ACKNOWLEDGEMENT

We express our sincere thanks to the General Directorate of Scientific Research and Technological Development (DGRSDT) for their support in the development of this work.

REFERENCES

- [1] B. Gharnali and S. Alipour, "MRI Image Segmentation Using Conditional Spatial FCM Based on Kernel-Induced Distance Measure," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2985–2990, Jun. 2018, <https://doi.org/10.48084/etasr.1999>.
- [2] L. Jordanova, L. Laskov, and D. Dobrev, "Influence of BCH and LDPC Code Parameters on the BER Characteristic of Satellite DVB Channels," *Engineering, Technology & Applied Science Research*, vol. 4, no. 1, pp. 591–595, Feb. 2014, <https://doi.org/10.48084/etasr.394>.
- [3] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, Jul., Oct. 1948.
- [4] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proceedings of ICC '93 - IEEE International Conference on Communications*, Geneva, Switzerland, May 1993, vol. 2, pp. 1064–1070, <https://doi.org/10.1109/ICC.1993.397441>.
- [5] Ha-Young Yang, Suk-Hyun Yoon, and Chang-Con Kang, "Iterative decoding of serially concatenated convolutional codes applying the SOVA," in *VTC '98. 48th IEEE Vehicular Technology Conference. Pathway to Global Wireless Revolution (Cat. No.98CH36151)*, Ottawa, Canada, May 1998, vol. 1, pp. 353–357, <https://doi.org/10.1109/VETEC.1998.686594>.
- [6] C. Berrou, Ed., *Codes and turbo codes*. Paris, France: Springer-Verlag, 2010.
- [7] L. Mostari, R. Méliani, and A. Bounoua, "Turbo-code in a bit interleaved turbo-coded modulation system," in *CMT'2010*, Ecole Supérieure de Technologie, Hassan 2 University, Casablanca, Morocco, 2010.
- [8] "In pictures ... this is how the Corona virus appears in the infected lungs," *Ramallah News*, Mar. 14, 2020. <https://ramallah.news/post/151117/> (accessed Feb. 24, 2021).
- [9] M. Fayssal, "Optimisation d'une transmission d'images fixes utilisant un turbo-code," Ph.D. dissertation, University of Djillali Liabes, Sidi bel Abbes, Algeria, 2019.
- [10] P. Robertson and T. Wörz, "Coded modulation scheme employing turbo codes," *Electronics Letters*, vol. 31, no. 18, pp. 1546–1547, Aug. 1995, <https://doi.org/10.1049/el:19951064>.
- [11] P. Hoher, "New Iterative ('Turbo') Decoding Algorithms," in *Proc. International Symposium on Turbo Codes & Related Topics*, Brest, France, 1997, pp. 63–70.
- [12] J. D. Kene and K. D. Kulat, "Soft Output Decoding Algorithm for Turbo Codes Implementation in Mobile Wi-Max Environment," *Procedia Technology*, vol. 6, pp. 666–673, 2012, <https://doi.org/10.1016/j.protcy.2012.10.080>.
- [13] P. Robertson and T. Wörz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 206–218, Feb. 1998, <https://doi.org/10.1109/49.661109>.
- [14] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982, <https://doi.org/10.1109/TIT.1982.1056454>.
- [15] M. R. Lahcene, "Amélioration des performances des codeur/décodéur sur canal WIMAX," Ph.D. dissertation, University of Tahri Mohammed, Bechar, Algeria, 2018.
- [16] M. R. Lahcene, A. Bassou, M. Beladgham, and A. Taleb-Ahmed, "On the Employment of Unpunctured Turbo-Trellis Coded Modulation Channel Coding to Improve the Performance of WiMax System," *International Journal on Communications Antenna and Propagation*, vol. 8, no. 3, pp. 248–256–256, Jun. 2018, <https://doi.org/10.15866/irecap.v8i3.13609>.
- [17] F. Menezla, M. S. Bendelhoum, Z. Mahdjoub, L. M. Reda, and R. Meliani, "The Effect of Error Correcting Codes in the Chain of Transmission and Comparison between the Performances of these Codes," in *2019 6th International Conference on Image and Signal Processing and their Applications (ISPA)*, Mostaganem, Algeria, Nov. 2019, <https://doi.org/10.1109/ISPA48434.2019.8966917>.
- [18] H. Mounira, "Reception dans un système d'accès multiples a repartition par code et application aux mode FDD et TDD de l'UMTS," Ph.D. dissertation, University of Biskra, Biskra, Algeria, 2014.
- [19] K. D. Eduardo, "Les code correcteurs d'erreurs LDPC," M.S. thesis, University of Sidi bel Abbes, Sidi bel Abbes, Algeria, 2012.
- [20] F. Menezla, R. Meliani, and Z. Mahdjoub, "Comparative study on performance of the LDPC codes and turbo-codes used in the image transmission in the gaussian and rayleigh channel," *Journal of Applied Research and Technology*, vol. 17, no. 2, Oct. 2019, <https://doi.org/10.22201/icat.16656423.2019.17.2.802>.
- [21] B. Mohamed, B. Abdesselam, B. Mohammed, T.-A. Abdelmalik, and M. L. Abdelmounaim, "International Journal of Image, Graphics and Signal Processing(IJIGSP)," *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, vol. 6, no. 10, pp. 10–17, Sep. 2014.
- [22] S. Wang, C. Perrine, C. Olivier, and C. Chatellier, "Stratégie de codage conjoint pour la transmission de séquences d'images via un canal bruité," Nov. 2007, Accessed: Feb. 24, 2021. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-00348734>.

The Energy Output from the Kuching Barrage in East Malaysia

Kamran Ahmed Samo

Electrical Engineering Department
Quaid-e-Awam University of Engineering, Science and
Technology, Larkana, Sindh, Pakistan
kamransamo2@gmail.com

Muhammad Usman Keerio

Electrical Engineering Department
Quaid-e-Awam University of Engineering, Science and
Technology, Nawabshah, Sindh, Pakistan
usmankeerio@quest.edu.pk

Shahid Hussain Shaikh

Electrical Engineering Department
Quaid-e-Awam University of Engineering, Science and
Technology, Larkana, Sindh, Pakistan
shaikhshahid@quest.edu.pk

Andrew Ragai Henry Rigit

Mechanical Engineering Department
Universiti Malaysia Sarawak Kota Samarahan
Sarawak, Malaysia
arigit@unimas.my

Kishan Chand Mukwana

Energy & Environment Engineering Department
Quaid-e-Awam University of Engineering, Science and
Technology, Nawabshah, Sindh, Pakistan
kcmukwana@quest.edu.pk

Abstract—Electricity generation from the sea has many advantages in comparison with other renewable energy resources. Power can be generated from new or existing barrages. Based on previous location research, a suitable system to produce tidal range energy from a potential site was developed in this paper. The main objective of this research is to calculate the energy output of the Kuching Barrage of Sarawak State of Malaysia. The daily flushing process of Kuching Barrage is conducted during the low tide period and therefore to put up the ebb generation process is appropriate. The calculated period of power generation is determined to about 6 hours. The annual energy output is calculated based on a theoretical method, with the average daily potential energy calculated to be 5.8MW and approximately 10.23GWh/year could be harnessed. This research can be beneficial for energy generation with the use of a double basin scheme for the construction of new barrages in East Malaysia.

Keywords—generation; Kuching Barrage; energy output; potential energy; Malaysia

I. INTRODUCTION

The oceans possess a huge potential of generating electrical power [1]. The generation of electricity from oceans offers many advantages when compared with other renewable energy sources [2]. Oceans are now globally established as mainstream renewable sources of energy [3]. It is calculated that the theoretical ocean energy resources are over 30,000TWh/year and the net potential power is larger than wind and solar power combined [4]. Malaysia is a big producer of solar Photovoltaic

(PV) panels and is ranked as the world's third-largest producer of solar PV energy [5]. Although Malaysia is located at the equatorial zone and surrounded by sea, the ocean potential power has not been given full attention by the Malaysian government [6]. The country's total coastline is 4,675km with West and East Malaysia having 2,068km and 2,607 km of coastline respectively [7]. The vast area of Malaysia's coastline is a huge advantage making tidal range energy a reliable alternative energy source [8]. Tidal range energy is convertible and can be used at a big scale for sustainable electrical power generation [9]. Tidal range power is produced due to the consistent rise and fall of seawater [10, 11]. Tidal range power can be generated where the tidal range flow is available which means that it cannot be generated inland [12]. In Malaysia, the overall renewable energy creation is 32% of the yearly target. Malaysia has to restore its responsibility regarding accomplishing the target of electricity generation from renewable sources.

Worldwide, there are a few barrage plants in operation. The Sihwa dam in Korea and La Rance in France have mounting energy capability of 254MW and 240MW respectively [14, 15]. The described project supports the government's endeavors of electricity production from renewable energy resources and promotes tidal range energy. Unlike other known renewable energy resources, tidal range is an expectable phenomenon. Energy productions from a tidal range power can be assessed appropriately [13].

Corresponding author: Kamran Ahmed Samo

II. LITERATURE REVIEW AND METHODOLOGY

A. Methods of Calculating Potential Energy and Annual Energy Output

The theoretical calculation method is based upon derived equations. For tidal barrage, basin size and mean tidal range have important influence on power generation. Equation (1) shows the correlation of potential energy in mean tidal range enclosed in a basin [16]:

$$E_p = 0.5 \rho g A_b \Delta h_b^2 \quad (1)$$

where E_p is the potential energy over a tide cycle (GJ), ρ is the density of sea water (1025 kg/m^3), g is the acceleration of gravity (9.81 m/s^2), A_b is the horizontal area of the basin (km^2), and Δh_b is the monthly or annually mean tidal range in the basin (m). The factor 0.5 is its average value. The total basin area bounded by Kuching Barrage is 1430 km^2 . However, this may not be valid when used for power generation therefore Google maps was used to calculate the area.

For coastal areas where tidal range power generation is economically attractive, the tidal regime generally consists of two flood and two ebb tides, with a semi-diurnal period of 12.42h at the low tide when the potential energy is zero. Therefore, the total potential energy per day from the barrage approximates to $24/12.42 E_p$, and the corresponding mean potential power can be written as:

$$\bar{P} = 24/12.42 E_p / 86400 \quad (2)$$

where \bar{P} is the power in (GW), and E_p is the potential energy over a tide cycle (GJ) and time (s) [14].

Authors in [14] worked on the calculation of annual energy output from a tidal barrage. At the Severn Barrage, UK, researchers used a method for approximating the annual tidal energy output from a barrage. The total annual energy output from a barrage was calculated from the theoretical energy of tidal dynamics in the map of the Bristol Channel and Severn estuary with the site of the Severn Barrage [14]. It was suggested that the energy available from a barrage depends on the area of the water surface seized by the barrage and the mean tidal range h . Thus, the potential annual tidal energy output from a barrage can be considered as [17]:

$$E_{yr} = 0.987 A_b \Delta h_b^2 \eta \quad (3)$$

where E_{yr} is the potential annual tidal power output (GWh/year), A_b is the horizontal area of the basin (km^2), Δh_b is the tidal range in the basin (m), and η is the efficiency. The constant value 0.987 as mentioned in (3) is the converted value of tidal cycles per year [17]. The per cycle value is 1397 and the world mean period for tides is 12 hours and 24 minutes, hence, there are 706 tidal cycles per year, and therefore the converted value is 0.987kWh [17].

To sum up, the power output over a tide cycle can be calculated by using (1). Total potential energy per day can be calculated by (2). The relationship of (1) and (3) is that (1) can be used to calculate the daily potential energy and (3) can be used to calculate the annually power output. Authors in [18]

researched the potential energy generation on total 34 sites at Sabah and Sarawak, Malaysia and identified 18 sites as suitable for the construction of a barrage to generate energy. Moreover, these suitable sites were supposed as preliminary, as the final sites should be supposed after conducting a feasibility study. Among the 18 sites, 2 sites were selected as having the maximum potential. The maximum potential power calculated sites were at the Tanjung Manis and at the Pending. The maximum energy potential was calculated to come from Tanjung Manis and was measured between 50.7kW and 39.2kW, while the second highest power was calculated for Pending, between 33.1kW and 25.1kW.

In continuity of the above research, the calculation of potential energy and annual energy of a tidal range energy extraction barrage at the Kuching is calculated in the current research.

B. Calculation of Potential Energy and Annual Energy Output:

To calculate the potential energy over a tide cycle, (1) will be used to calculate the mean potential energy in a day. The calculation is completed by using (2) whereas (3) is used to calculate the annual energy output. The annual calculation power output was calculated with the theoretical method. The annual output power is calculated with the lower power conversion efficiency.

III. RESULTS AND DISCUSSION

A. Calculated Basin Area of Kuching Barrage

The basin area of Pending site at Kuching Barrage is calculated to get the potential energy and annual power output. The total basin area bounded by Kuching Barrage is 1430 km^2 , however, this may not be valid when used for power generation. Hence, it is assumed that the total bounded area may not be the effective area contributing to power generation. The effective basin area has been assumed as 2.94 km^2 .

B. Daily Power Output Calculation

Based on (1) and with a mean tidal range of 4.2m, Kuching Barrage potential energy value is 260GJ over a tide cycle. The daily potential energy is about 520GJ with nil potential energy during the low tide. In demand to get daily power output, the potential energy is distributed with the tidal period. The tidal period of semidiurnal tidal form is 12.42h or 44712s. The daily potential energy is $24/12.42$. Equation (2) is used to calculate the daily potential energy. Thus, Kuching Barrage expected mean daily potential energy is 0.0058GW or 5.8MW. However, the energy output may not be continuous during the 12h cycle (ebb tides occur twice a day) operation due to the changing sea level. The Kuching Barrage energy plant is calculated to produce energy output power for 6h. However, generation duration could accept dependency on tide pattern. Due to this reason the mean tidal range reduces as the downstream water increases. As the mean tidal range decreases in height the power generation also reduces.

Figure 1 shows a simple cycle of power generation during the operation of Kuching Barrage gate. Phases I, II, and III are respectively for filling, holding and generating. There is a 12h

operation during the ebb tide, a 6h generation and a 6h filing of the basin. The schedule of opening the Kuching Barrage gates to flush out the water to the sea depends on the tide level. The average schedule of opening the gates is 4-5h and sometimes 6-7h, depending on the tide level. Flush out water to downstream is possible when there is a low tide. The main aim of Kuching Barrage is to flush out river water before upstream got flooded during rainy season. The flush out of the upstream water to the downstream is possible twice a day as of low and high tides.

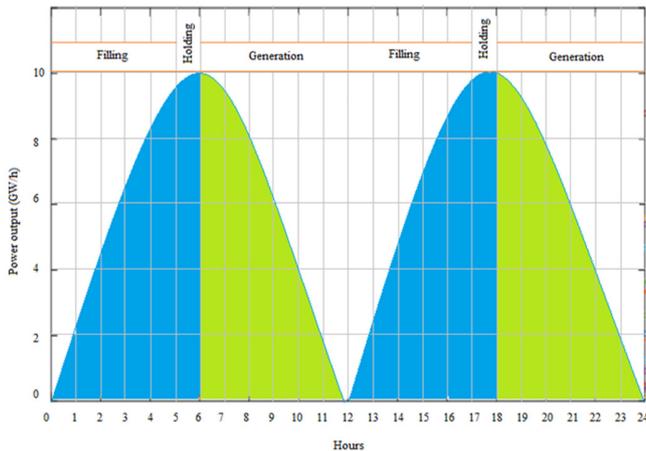


Fig. 1. A simple cycle of power generation.

C. Annual Power Output Calculation

Annual energy output is calculated by (3). At a minimum, 10.23GWh/year ($\eta_p = 20\%$) of power can be harnessed with the lowest efficiency. Table I summarizes the power output value of Kuching Barrage tidal energy plant system. In Table I, the actual and theoretical annual energy output from different tidal power plants [14] are listed. Table I documents the technical conditions of the tidal power plants, with the actual output and theoretical calculation range of the annual energy output from each plant. Table I also depicts the installed capacity, basin area, and mean tidal range of 5 main tidal power plants. Tidal range energy schemes were measured by low standards of power conversion efficiency, frequently reaching from 20% to 40%, (with an average of 33%). However, minimum power conversation efficiency (20%) is used to calculate the power output for this project. The mean tidal range is the monthly or yearly average. In Table I, the theoretical standards are tabulated and are related to a present tidal energy plant to confirm the calculation methods.

Summarizing, Kuching Barrage is an additional proposed tidal power plant. The calculated basin area of Kuching Barrage is 2.94km² with calculated mean tidal range is 4.2m. The annual power output is calculated based on the lowest power conversion efficiency of 20%. The expected power output for Kuching Barrage tidal energy plant Malaysia is calculated as 10.23GWh/year.

Table II shows the power output of the Kuching Barrage tidal energy plant. The mean daily potential energy is calculated by (2). The potential annual power output is calculated by (3) at the lower efficiency of 20%.

TABLE I. COMPARISON OF THEORETICAL AND ACTUAL POWER OUTPUT OF TIDAL POWER PLANTS

Country	Site	Installed capacity (MW)	A_b (km ²)	Δh_b (m)	E_{yr} range at $\eta = 20\%$ (GWh/yr)	Actual or expected output (GWh/yr)
France	La Lance	240	22.5	8.5	320.90	533
Canada	Annapoils	20	6.0	6.4	48.51	50
China	Jiangxia	3.9	1.37	5.1	8.88	6-7
Korea	Shiwa	254	43	5.6	266.19	553
UK	Severn	8640	570	7.5	6329.14	15600
Malaysia	*Kuching Barrage	-	2.94	4.2	10.23	10.23

*Kuching Barrage at Pending site is our proposal for a tidal power plant scheme

TABLE II. CALCULATED POWER OUTPUT OF KUCHING BARRAGE TIDAL RANGE ENERGY PLANT SCHEME

Power estimation	Value
Mean daily potential energy, \bar{P}	2.8MW
Potential annual power, E_{yr}	5TWh/year ($\eta = 20\%$)

IV. CONCLUSION

The potential energy and annual power output of the Pending site at the Kuching Barrage basin area were calculated in this paper. The basin area of Kuching Barrage has been considered as 2.94km². The daily flushing process of Kuching Barrage is for the period of low tides and therefore is appropriate to put up the ebb generation process. The calculated period of power generation is determined to about 6h. Thus, the expected mean daily potential energy of Kuching Barrage is 5.8MW. The minimum annual energy output is calculated to be 10.23GWh/year.

REFERENCES

- [1] A. S. Bahaj, "Generating electricity from the oceans," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 7, pp. 3399–3416, Sep. 2011, <https://doi.org/10.1016/j.rser.2011.04.032>.
- [2] L. Myers and A. S. Bahaj, "Simulated electrical power potential harnessed by marine current turbine arrays in the Alderney Race," *Renewable Energy*, vol. 30, no. 11, pp. 1713–1731, Sep. 2005, <https://doi.org/10.1016/j.renene.2005.02.008>.
- [3] F. M. Sim, "Global Direction on Renewables," *The Ingenieur*, vol. 67, pp. 32-37, Jul. 2016.
- [4] T. D. Corsatea and D. Magagna, *Overview of European innovation activities in marine energy technology*. Brussels, Belgium: European Commission, 2014.
- [5] "Malaysia is third largest producer of PV cells," www.thesundaily.my. <https://www.thesundaily.my/archive/2193974-XTARCH433330> (accessed Mar. 06, 2021).
- [6] O. Yaakob, Y. Ahmed, M. Mazlan, K. Jaafar, and R. Muda, "Model Testing of an Ocean Wave Energy System for Malaysian Sea," *World Applied Sciences Journal*, vol. 22, no. 5, pp. 667-671, Apr. 2013.
- [7] "Malaysia - The World Factbook," *CIA*. <https://www.cia.gov/the-world-factbook/countries/malaysia/> (accessed Mar. 06, 2021).
- [8] S. M. Shafie, T. M. I. Mahlia, H. H. Masjuki, and A. Andriyana, "Current energy usage and sustainable energy in Malaysia: A review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4370–4377, Dec. 2011, <https://doi.org/10.1016/j.rser.2011.07.113>.
- [9] G. Lalchand, "Fostering green growth for Malaysia, Energy-Wise," *Jurutera, the Monthly Bulletin of the Institution of Engineers*, vol. 2016, pp. 6–14.

-
- [10] O. Yaakob, T. Mohd, A. Bin, T. A. Rashid, M. Affi, and A. Mukti, "Prospects for Ocean Energy in Malaysia," in *International Conference on Energy and Environment 2006*, 2006, pp. 62–68.
- [11] J. Xia, R. A. Falconer, and B. Lin, "Impact of different tidal renewable energy projects on the hydrodynamic processes in the Severn Estuary, UK," *Ocean Modelling*, vol. 32, no. 1, pp. 86–104, Jan. 2010, <https://doi.org/10.1016/j.ocemod.2009.11.002>.
- [12] P. Meisen and A. Loiseau, *Ocean energy technologies for renewable energy generation*. San Diego, CA, USA: Global Energy Network Institute, 2009.
- [13] S. Waters and G. Aggidis, "Tidal range technologies and state of the art in review," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 514–529, Jun. 2016, <https://doi.org/10.1016/j.rser.2015.12.347>.
- [14] J. Xia, R. A. Falconer, B. Lin, and G. Tan, "Estimation of annual energy output from a tidal barrage using two different methods," *Applied Energy*, vol. 93, pp. 327–336, May 2012, <https://doi.org/10.1016/j.apenergy.2011.12.049>.
- [15] R. Kempener and F. Neumann, *Tidal Energy Technology Brief*. Abu Dhabi, UAE: IRENA, 2014.
- [16] H. Lamb, *Hydrodynamics*, 6th ed. New York, NY, USA: Dover Publications, 1945.
- [17] J. W. Tester, E. M. Drake, M. J. Driscoll, M. W. Golay, and W. A. Peters, *Sustainable Energy, second edition: Choosing Among Options*, 2nd ed. Cambridge, MA, USA: The MIT Press, 2012.
- [18] K. A. Samo, I. A. Samo, Z. A. Siyal, and A. R. H. Rigit, "Determination of Potential Tidal Power Sites at East Malaysia," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6047–6051, Aug. 2020, <https://doi.org/10.48084/etasr.3674>.

The Effects of Steam-Curing on the Properties of Concrete

Mustapha Boukendakdji
College of Engineering
University of Hail
Hail, Saudi Arabia
m.boukendakdji@uoh.edu.sa

Mabrouk Touahmia
College of Engineering
University of Hail
Hail, Saudi Arabia
m.touahmia@uoh.edu.sa

Belkacem Achour
College of Engineering
University of Hail
Hail, Saudi Arabia
b.achour@uoh.edu.sa

Ghazi Albaqawy
College of Engineering
University of Hail
Hail, Saudi Arabia
g.albaqawy@uoh.edu.sa

Mohamed Abdelhafez
College of Engineering
University of Hail
Hail, Saudi Arabia
mo.abdelhafez@uoh.edu.sa

Khaled Elkhayat
College of Engineering
University of Hail
Hail, Saudi Arabia
m.elkhayat@uoh.edu.sa

Emad Noaime
College of Engineering
University of Hail
Hail, Saudi Arabia
e.noaime@uoh.edu.sa

Abstract-Worldwide, concrete is the most preferred construction material. The steam curing method is favored when there is a need for accelerating strength. This paper presents the study of the compressive and flexural tensile strength of concrete subjected to eight different steam cures. In addition, the stress-strain curve and the modulus of elasticity were determined at the age of 28 days. The compressive strength test results show that after treatment, strength increases with concrete maturity. A cycle with a pre-heating period gives better results than a cycle without a pre-heating period. The longer the duration of the maximum temperature period, the lower the strength drop compared to the control concrete. The best results were obtained for concrete treated according to the following cycle: a 3-hour pre-heating period at 20°C, a 2-hour increase of temperature from 20 to 70°C, and a 3 hour of maximum temperature of 70°C.

Keywords-steam curing; compressive strength; flexural tensile strength; modulus of elasticity; maturity

I. INTRODUCTION

Concrete is extensively used as a construction material and, globally, it is the second highest consumed material [1]. Ordinary Portland Cement (OPC) is usually used as the primary binder to make concrete and is one of the most energy-intensive construction materials. The primary reason for using steam in the curing of concrete is to produce high early strength, which is very desirable to the manufacturers of precast and pre-stressed concrete units [2]. There are many advantages in steam curing such as the production of concrete with high early strength, short production cycle, and superior economic benefits [3, 4]. The steam curing process includes the following four phases: the delay period, the isothermal period, the heating period, and the cooling period [5]. Heat treatment has been used to accelerate the development of the strength of concrete products, such as bricks and small concrete elements. By increasing temperature and duration, the strength of concrete is increased [6]. Authors in [7] showed that 85°C is

the optimum temperature in order to obtain high early strength. Authors in [8] demonstrated the improved early-age strength of high performance concrete through steam curing. Authors in [9] showed that as the temperature increases, the compressive and flexural strength of concrete at early age increase. Authors in [5] reported that 65% of the strength of the control concrete at 28 days was obtained by thermal treatment within 24h after mixing. Similar results have also been reported in [6,10, 11]. The same results have been obtained for concrete with superplasticizers [12] and slag [5]. In order to obtain good final strength, it is not advisable to immediately heat the concrete. The faster the speed of temperature rising, the more the final strength of the treated concrete is weakened. Authors in [9] showed that the increase in temperature has significant impact on the compressive and flexural tensile strength of concrete. However, there is no effect on strength when hard concrete is exposed to 6-hour fire at 1000°C [13]. Authors in [5, 14] reported that the effect of heat treatment on the modulus of elasticity is the same as that on strength. Authors in [14] reported that the optimum rate of heating and preset period should be selected based on the maximum chamber temperature in use in order to secure the desired results. There was enhancement in the early compressive strength due to the effect of steam curing, however the delay period before steam curing is very essential for the early strength gain [2, 15]. Authors in [5] showed that the reduction in shrinkage of steam-cured concrete when compared to normal cured concrete is about 10%. Finally, steaming leads to a reduction of 20 to 30% of the creep compared to the same concrete, cured at room temperature and loaded with equal stress/strength ratio [16]. This reduction in creep has been confirmed by [5, 12].

II. MATERIALS AND EXPERIMENTAL DETAILS

Nine concrete mixtures were tested in this program. The concrete had an OPC content of 350kg/m³. The total aggregate/cement ratio was 5.41, the gravel/sand ratio was 1.61, and the

water/cement ratio was 0.57. For all the mixtures, the coarse and fine aggregates were weighed in dry room conditions. The control OPC concrete was designed to develop a 28-day cube compressive strength of 35MPa. All the concrete mixtures were mixed for 2min in a laboratory mixer. All molded specimens were cured for 24h under wet hessian, then were demolded and cured in water at $20\pm 2^\circ\text{C}$ for 28 days. For each mixture, twelve 100mm cubes were made to determine the compressive strength and two $300\times 150\text{mm}$ diameter cylindrical specimens were used to determine the static modulus of elasticity. The flexure tensile strength was measured on $100\times 100\times 500\text{mm}$ prisms (Figure 1). The compressive strength was measured after 1, 3, 7, and 28 days (Figure 2). The flexural tensile strength, the stress-strain curve and the modulus of elasticity were determined at the age of 28 days.



Fig. 1. The concrete specimens.



Fig. 2. The cube compressive strength test.

The studied types of concrete were: A control concrete stored in water at an ambient temperature (20°C) and concretes subjected to 8 different steam curing cycles (Table I). The curing cycles necessarily include four distinct phases:

- Preheating period
- Heating period
- Maximum temperature period
- Cooling period

The molds were placed into the steam chamber, the latter being at the processing temperature. This steam chamber is equipped with a fan allowing the circulation of ambient air. A thermocouple was placed in the concrete in the middle of the central specimen of the cubic mold. After demolding, the concrete specimens were stored in water at 20°C

TABLE I. DIFFERENT TYPES OF CURING CYCLES

Cycle	Preheating period (h)	Heating period (h)	Max. Tem. period (h)	Initial Tem. ($^\circ\text{C}$)	Max. Tem. ($^\circ\text{C}$)	Maturity ($^\circ\text{C}\cdot\text{h}$)
A	0	2	3	20	70	350
B	1.5	2	3	20	70	395
C	3	2	3	20	70	440
D	3	2	1.5	20	70	320
E	3	2	0	20	70	200
F	1.5	2	3	20	50	315
G	1.5	2	3	24	70	405
H	1.5	2.5	3	24	70	433

III. RESULTS AND DISCUSSION

A. Influence of the Duration of the Pre-Heating Period on Compressive Strength

Figure 3 shows that the strength at the end of the treatment (0 days) and after 28 days of curing is minimal for zero preheating period. The decrease in final strength (28 days) compared to control concrete kept at 20°C is around 33%. For the cycles B and C, the fall is 26% and 15% respectively. So it is not advisable to heat the concrete immediately after pouring, which is in agreement with the results in [10]. The heating of fresh concrete causes the appearance of physical and chemical phenomena [17]. The physical order phenomena are essentially linked to the expansion of the various constituents of concrete under the effect of the rise in temperature. However, these various constituents have very different linear expansion coefficients. The rise in temperature causes the material to expand, which ultimately results in the appearance of pores. These pores indeed remain after the consolidation of the concrete structure. It is well known that the strength of concrete is inversely proportional to the number of pores in it. Chemical phenomena are linked to the appearance of insoluble, compact hydrate films, which oppose the subsequent penetration of water and somehow isolate the grains of cement [17].

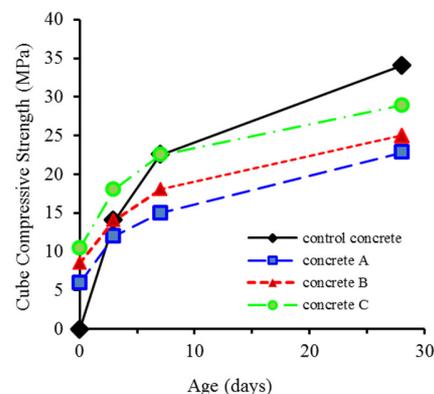


Fig. 3. The effect of the delay period on concrete strength.

B. Influence of the Speed of Temperature Rise on Compressive Strength

In order to study the influence of the speed of temperature rising, the evolution of the compressive strength for heating at 70°C with rise speeds of 18°C/h and 24.8°C/h are compared in Figure 4. It is observed that the best strength at 28 days is obtained for slowly heated concrete (18°C/h). However, the quality of treated concrete remains lower than that of control concrete. The drop in strength compared to control concrete is 13% and 15% for concrete heated to 18°C/h and 24°C/h respectively. This confirms the results of [5]. The physical cause of this alteration is known: it is the very significant expansion of the constituents of fresh concrete, water and air. A high rate of temperature rise causes disorders in the structure of concrete which can go as far as the appearance of micro-cracks.

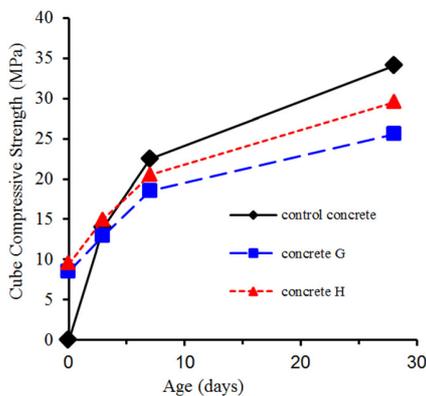


Fig. 4. The effect of heating period on concrete strength.

C. Influence of Maximum Temperature

When the temperature rises the compressive strength increases more rapidly at 0 days, whereas at 28 days, the strength of concrete treated at 50°C is greater than that of the concrete treated at 70°C (Figure 5). Authors in [7] found that the optimum maximum temperature of steam curing was near 60°C. However, higher temperatures of 70°C and 80°C show a distinct reduction in strength after 28 days [7]. A significant increase of compressive strength for fly ash mortars has been found at the early age of 3 days when curing temperature was raised [18].

D. Influence of the Maximum Temperature Period Time

In order to study this phenomenon, three cycles of heat treatments (C, D and E) were considered, so the maximum temperature period varied from 0 to 3h. The three cycles had the same preheating period (3h), heating period (2h), and maximum temperature (70°C), however the maximum temperature period was 3, 1.5, and 0h respectively. It should be noted that the longer the maximum temperature period time, the higher the compressive strength in the short and long term is (Figure 6). Concrete maintained at a 3hour level presents, at 28 days, a reduction in strength around 15% compared to control. For 1.5h and 0h concrete, the strength drop is around 24% and 28%, respectively. Therefore, the greater the duration of the maximum temperature period, the smaller the drop in

strength compared to control concrete at 28 days, which is in agreement with the results in [19].

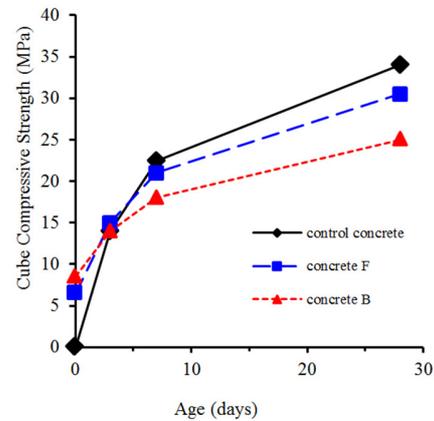


Fig. 5. The effect of maximum temperature on concrete strength.

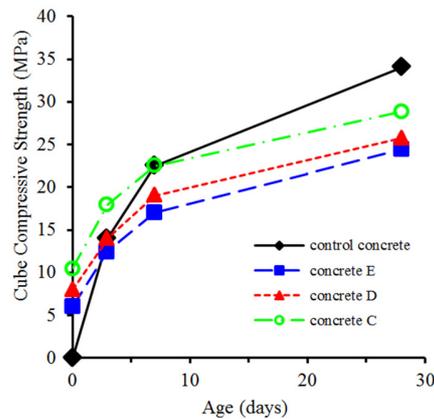


Fig. 6. Effect of maximum temperature holding time on concrete strength.

E. Maturity Factor

We can see that the compressive strength after treatment increases with maturity (Figure 7). The latter is the sum of the product of temperature and time:

$$Maturity = \sum (T + 10^\circ C) \times \Delta t \quad (1)$$

where T is the temperature in °C and Δt is the time interval (h).

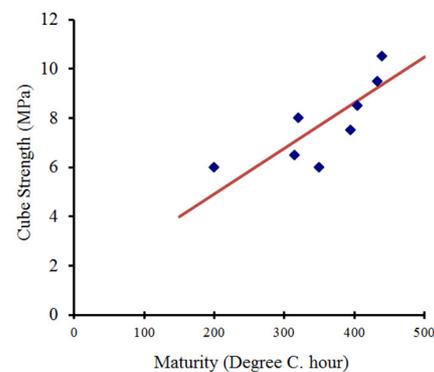


Fig. 7. Relationship between maturity and cube compressive strength.

F. Tensile Strength

The values of tensile strength are given in Table II. It can be seen that the maximum tensile strength (4.27MPa) is obtained by the control concrete. The tensile strength of all concrete mixtures was determined by the flexure test and is presented as a function of cube compressive strength in Figure 8. The European Concrete Committee (CEB) suggests that the axial tensile strength (f_t) can be related to the characteristic compressive strength (f_c) by:

$$f_t = 0.3f_c^{2/3} \quad (2)$$

The flexural tensile strength of all the concrete mixtures ranged between 2.9 and 4.27MPa. For a given compressive strength, the values for flexural tensile strength for all concretes are slightly higher than to the curve proposed by the CEB in (2). The ratio of the flexural tensile strength to the compressive strength was about 0.13. The results of [21] showed that the steam-curing provides better performance at 3 days for flexural strength and the ratio of the flexural tensile strength to the compressive strength was about 0.10.

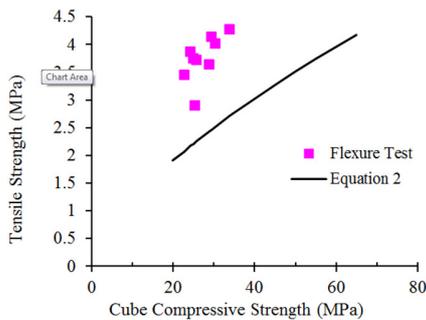


Fig. 8. Relationship between tensile and cube compressive strength.

TABLE II. FLEXURAL TENSILE STRENGTH

Cycle	Control	A	B	C	D	E	F	G	H
Tensile strength (MPa)	4.27	3.45	3.75	3.64	3.72	3.86	4.01	2.90	4.14

G. Stress-Strain Curve

The stress-strain curves in compression of the control and both B and C concretes are shown in Figure 9. It is clear that the shape of the curves is similar for all concretes. Moreover, it is worth mentioning that before failure, it is difficult to assess any value for limiting axial strain. However, comparison can be made on the basis of 95% of failure stress. The limiting axial strain for control concrete is about 1880 microstrains. However, for concretes B and C, these values are 1420 and 1596 microstrains respectively. These values are 24% and 15% lower than the control concrete's. The secant modulus of elasticity, as measured at a stress-strength ratio of 0.3 is shown in Table III. Compared with normal concrete, the values of concrete B and C are less by approximately 14% and 8% respectively. On the other hand, the value proposed by ACI [14] is 24.3GPa. This result is in agreement with the results in [20] for steam-cured slag concrete.

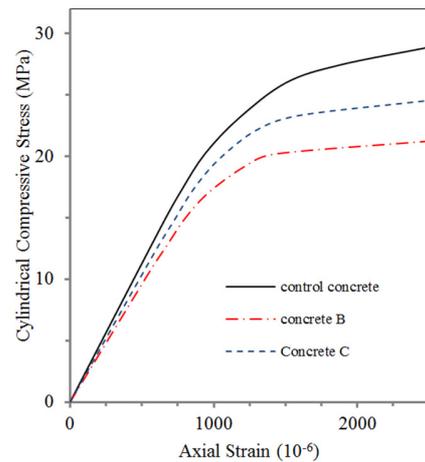


Fig. 9. Stress-strain curves for concretes in compression.

TABLE III. SECANT MODULUS OF ELASTICITY (GPA)

Age (days)	Control	B	C
28	26.5	22.8	24.4

IV. CONCLUSION

For the materials and test conditions reported in this study, it was found that:

- A decrease in the final strength of about 33% compared to that of the control concrete was observed for the cycle without pre-heating (Cycle A). However, for the cycles B and C (with pre-heating 1.5 and 3 hours), the decrease was 26% and 15% respectively.
- The greater the maximum temperature period time, the smaller the drop in compressive strength at 28 days compared to the control.
- The best result for the treated concrete is obtained according to the following cycle:
 - A 3-hour of pre-heating period at 20°C,
 - A 2-hour increase of temperature from 20°C to 70°C, and
 - A 3-hour heating at the maximum temperature of 70°C.

ACKNOWLEDGMENT

The current research was funded by the Deanship of Scientific Research at the University of Hail, Saudi Arabia, under the contract No. RG-191241. The authors would like to extend their gratitude to the Deanship of Scientific Research and to the College of Engineering at the University of Hail for providing all the required facilities.

REFERENCES

[1] A. A. Jhatial, S. Sohu, N. ul K. Bhatti, M. T. Lakhiar, and R. Oad, "Effect of steel fibres on the compressive and flexural strength of concrete," *International Journal of Advanced and Applied Sciences*, vol. 5, no. 10, pp. 16–21, Oct. 2018, <https://doi.org/10.21833/ijaas.2018.10.003>.

- [2] R. Nayaka, G. S. Diwakar, and V. Gudur, "Effect of Steam Curing on the Properties of High Early Strength Silica Fume Concrete," *IOP Conference Series: Materials Science and Engineering*, vol. 431, no. 4, Nov. 2018, Art. no. 042011, <https://doi.org/10.1088/1757-899X/431/4/042011>.
- [3] J. L. García Calvo, M. C. Alonso, L. Fernández Luco, and M. Robles Velasco, "Durability performance of sustainable self compacting concretes in precast products due to heat curing," *Construction and Building Materials*, vol. 111, pp. 379–385, May 2016, <https://doi.org/10.1016/j.conbuildmat.2016.02.097>.
- [4] A. M. Ramezani pour, Kh. Esmacili, S. A. Ghahari, and A. A. Ramezani pour, "Influence of initial steam curing and different types of mineral additives on mechanical and durability properties of self-compacting concrete," *Construction and Building Materials*, vol. 73, pp. 187–194, Dec. 2014, <https://doi.org/10.1016/j.conbuildmat.2014.09.072>.
- [5] M. Boukendakdji, J. J. Brooks, and P. J. Wainwright, "Influences of steam curing on strength, shrinkage and creep of OPC and slag concretes," in *Radical concrete technology: proceedings of the international conference held at the University of Dundee*, Dundee, Scotland, UK, Jun. 1996, <https://doi.org/10.1201/9781482294941-49>.
- [6] K. Elakkiya, N. Una, and A. Govandan, "Effect of thermal and microwave oven curing on concrete," *Indian Journal of Scientific Research*, vol. 17, no. 2, pp. 262–267, 2018.
- [7] S. Türkel and V. Alabas, "The effect of excessive steam curing on Portland composite cement concrete," *Cement and Concrete Research*, vol. 35, no. 2, pp. 405–411, Feb. 2005, <https://doi.org/10.1016/j.cemconres.2004.07.038>.
- [8] J.-S. Park, Y. J. Kim, J.-R. Cho, and S.-J. Jeon, "Early-Age Strength of Ultra-High Performance Concrete in Various Curing Conditions," *Materials (Basel, Switzerland)*, vol. 8, no. 8, pp. 5537–5553, Aug. 2015, <https://doi.org/10.3390/ma8085261>.
- [9] M. A. Panhyar, S. N. R. Shah, F. A. Shaikh, and A. A. Jhatial, "Influence Of Casting Temperature On The Structural Behavior Of Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4480–4483, Aug. 2019, <https://doi.org/10.48084/etasr.2924>.
- [10] B. Liu, Y. Xie, and J. Li, "Influence of steam curing on the compressive strength of concrete containing supplementary cementing materials," *Cement and Concrete Research*, vol. 35, no. 5, pp. 994–998, May 2005, <https://doi.org/10.1016/j.cemconres.2004.05.044>.
- [11] Y.-J. Kim, S.-W. Kim, C.-W. Park, and J.-S. Sim, "Compressive Strength Properties of Concrete Using High Early Strength Cement and Recycled Aggregate with Steam Curing Conditions," *Journal of the Korean Recycled Construction Resources Institute*, vol. 4, no. 1, pp. 76–81, 2016, <https://doi.org/10.14190/JRCR.2016.4.1.076>.
- [12] A. Makani, T. Vidal, G. Pons, and G. Escadeillas, "Time-dependent behaviour of high performance concrete: influence of coarse aggregate characteristics," *EPJ Web of Conferences*, vol. 6, p. 03002, 2010, <https://doi.org/10.1051/epjconf/20100603002>.
- [13] A. H. Buller, M. Oad, and B. A. Memon, "Flexural Strength of Reinforced Concrete RAC Beams Exposed to 6-hour Fire – Part 2: Rich Mix," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3814–3817, Feb. 2019, <https://doi.org/10.48084/etasr.2494>.
- [14] A. Hanif, Y. Kim, K. Lee, C. Park, and J. Sim, "Influence of cement and aggregate type on steam-cured concrete – an experimental study," *Magazine of Concrete Research*, vol. 69, no. 13, pp. 694–702, Apr. 2017, <https://doi.org/10.1680/jmacr.17.00015>.
- [15] S.-D. Hwang, R. Khatib, H. K. Lee, S.-H. Lee, and K. H. Khayat, "Optimization of steam-curing regime for high-strength, self-consolidating concrete for precast, prestressed concrete applications," *PCI Journal*, vol. 57, no. 3, pp. 48–62, Jun. 2012, <https://doi.org/10.15554/pcij.06012012.48.62>.
- [16] *318-08: Building Code Requirements for Structural Concrete and Commentary*. Farmington Hills, MI, USA: American Concrete Institute, 2008.
- [17] A. M. Neville, *Properties of Concrete*, 5th ed. Harlow, England, UK; New York, NY, USA: Trans-Atlantic Publications, 2012.
- [18] J. Payá, J. Monzó, M. V. Borrachero, E. Peris-Mora, and F. Amahjour, "Mechanical treatment of fly ashes: Part IV. Strength development of ground fly ash-cement mortars cured at different temperatures," *Cement and Concrete Research*, vol. 30, no. 4, pp. 543–551, Apr. 2000, [https://doi.org/10.1016/S0008-8846\(00\)00218-0](https://doi.org/10.1016/S0008-8846(00)00218-0).
- [19] M. Ba, C. Qian, X. Guo, and X. Han, "Effects of steam curing on strength and porous structure of concrete with low water/binder ratio," *Construction and Building Materials*, vol. 25, no. 1, pp. 123–128, Jan. 2011, <https://doi.org/10.1016/j.conbuildmat.2010.06.049>.
- [20] M. Boukendakdji, "Mechanical properties and long-term deformation of slag cement concrete," Ph.D. dissertation, University of Leeds, Leeds, UK, 1989.
- [21] Y. H. Ahmed and R. Azhari, "Effect of atmospheric steam curing parameters on concrete strength and durability," in *Second Conference of Civil Engineering in Sudan*, Khartoum, Sudan, Dec. 2018, pp. 48–54.

DyTE: An Effective Routing Protocol for VANET in Urban Scenarios

Abdul Karim Kazi
NED University of Engineering and Technology
Karachi, Pakistan
karimkazi@neduet.edu.pk

Shariq Mahmood Khan
NED University of Engineering and Technology
Karachi, Pakistan
shariq@neduet.edu.pk

Abstract—A Vehicular Ad-hoc Network (VANET) is a subclass of wireless ad-hoc networks, widely used in on-road vehicles and roadside equipment, having applications in various areas including passenger safety, smart traffic solutions, and connectivity on vehicles. The VANET is the backbone of the Intelligent Transport System (ITS) that establishes connectivity between vehicles through a wireless medium. When it comes to the communication between high-speed vehicles there is the challenge of dynamic mobility. In order to provide a higher Packet Delivery Ratio (PDR) and increase the throughput, a new routing protocol called Dynamic Trilateral Enrolment (DyTE) is introduced which chooses a dynamic trilateral zone to find the destination vehicle by allowing only relevant nodes to participate in the communication process using the location coordinates of the source and destination nodes. The proposed routing protocol is compared with Ad-hoc On-demand Distance Vector (AODV), Ad-hoc On-demand Multipath Distance Vector (AOMDV), and Dynamic Source Routing (DSR), and the results show a remarkable improvement in reducing the Network Routing Load (NRL) and in increasing the PDR and throughput of the network. DyTE has performed more efficiently in terms of PDR (23% approximately), throughput (26% approximately) and drastically minimized the NRL by a factor of almost 3.

Keywords—vehicular ad-hoc networks; location aided routing; trilateral zone; broadcast storm

I. INTRODUCTION

Vehicular Ad-hoc Networks (VANETs) are becoming quite popular because nowadays fast moving vehicles are able to communicate with each other and on-road equipment [1]. The network topology in a VANET changes rapidly as it consists of fast moving nodes. A VANET is considered to be an opposite option for the development of ITS [2]. A VANET is a specialized type of Mobile Ad-hoc Network (MANET), through which vehicular communication can take place in urban and highway scenarios. Many applications such as security, information services, accident alarm, road safety and traffic managements are associated with VANETs. The speed of the vehicle is the major factor which affects the communication process due to which rapid changes in network topology occur, so the selection of the intermediate vehicle is very crucial. A typical VANET scenario is depicted in Figure 1 where communication can take place between vehicle-to-vehicle (V-2-V), vehicle-to-infrastructure (V-2-I) and hybrid (either V-2-V or V-2-I) [3]. In a VANET, routing protocols are generalized in five categories [4]: Ad-hoc based, cluster-based,

position-based, broadcast-based, and geocast-based routing. Position-based routing protocols use geographical coordinates from the Global Positioning System (GPS). They also use digital maps to obtain road information. Low overhead and its effectiveness for rapidly moving nodes are the key advantages of location-based routing protocols. Geocast-based routing protocols basically limit the area of transmission in a specific portion by identifying the next hop using the node's position. Information related to the position and speed of the node help to decrease overhead packets and to increase PDR [5].

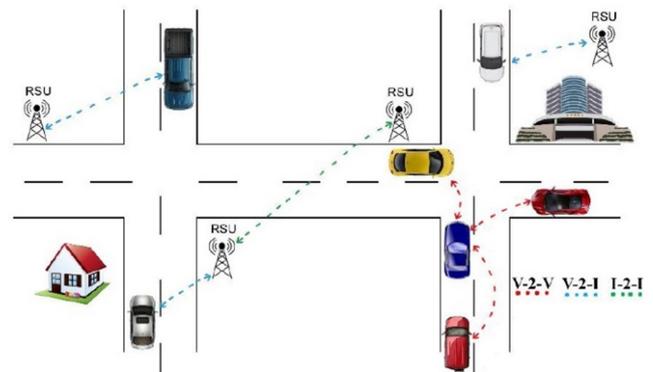


Fig. 1. A VANET scenario.

Location Aided Routing (LAR) is based on the exact GPS location of the destination node [6, 7]. In [8], an advancement of LAR that works on the principle of choosing next-hop with minimum angle difference to the straight line between source and destination node to reduce the packet overhead is reported. In [9], two strategies are opted for the transmission of packets while selecting the next-hop, greedy forwarding strategy chooses the closest node with respect to the destination and perimeter forwarding strategy is opted whenever the distance between the forwarding and the target node is closer than the distance to the destination of its neighbors. In [10], the lifespan of the node is calculated using the speed and position in order to get the time for the node's availability within the transmission range. Greedy Perimeter Coordinator Routing (GPCR) is also a position-based routing protocol which differs from the GPSR because it does not rely on the static global map [11].

Corresponding author: Abdul Karim Kazi

In DSR [12], a complete list of nodes in sequence is attached in the packet header therefore network overhead is increased. AODV [13] is a reactive and loop-free routing protocol which is still considered relevant [14-16]. An optimization in AODV is proposed in [17] by formulating route weight. The Improved AODV (IAODV) is proposed in [18] for dense networks and focuses on forwarding timely and accurate information to nodes by keeping a list of all one hop nodes as a backup route. Multiple routing path is accumulated during the route discovery phase by AOMDV [19] from source to destination node to get an uninterrupted communication in case of failure of the primary path. A variant of the AOMDV protocol is the Speed Direction AODV (SD-AODV) [20], in which the parameters of speed and direction are included in the hop count field during the process of route establishment. In [21], data transfer rate is optimized using Signal-to-Noise Ratio (SNR), improving link quality and minimizing energy consumption. A reactive routing protocol is proposed in [22] that minimizes the routing load by selecting reliable routes during the route discovery phase. An upgraded version of TORA was introduced in [23] where only highly reliable discovered links are used. In ERRRP [24], the disconnection of routes is minimized by the Reliability Factor (RF) which is based on Route Expiration Time (RET) and hop count. Higher RF value routes are selected as reliable paths. Roadside Service Discovery Protocol (RSDP) [25] proposed an application layer beaconing-based protocol to address the problem of service discovery upon request. Velocity-aided routing protocol [26] works on the principle of predicting the moving behavior of mobile nodes. The possible trajectory of the destination node is predicted on the basis of its position and velocity information and determines the packet forwarding region. The incorporation of mobility characteristics in the routing enhances its performance and makes it more adaptive.

Ant Colony Optimization (ACO) algorithms [27, 28] extract the optimal routing information by selecting the best available path among the available routes or by finding an alternate route in case of route failure. ACO algorithms successfully reduce link failures by applying the above strategies. An Improved Genetic Algorithm-based Route Optimization Technique (IGAROT) [29] is implemented by considering road anomalies resulting in prompt notification by road maintenance agencies of persistent road conditions through V2I communication. Using an iterative K-Means algorithm, IGAROT generates clusters into two non-overlapping groups and then updates the size of the successful cluster using likelihood metric in the initial population size. Overhead reduction was achieved in [30] by forwarding the location of the destination to every generated RREQ packet. Efficient Routing Protocol (ERP) [31] minimizes the unnecessary broadcast in the network by finding the minimum set of vehicles for backbone communication that act as forwarders and only those nodes will be having the responsibility to broadcast. Using the GPS system, the selection process of the next forwarding node is based on the lesser distance to destination node in GeOpps [32]. Due to network topology dependence, not every node is needed to calculate the optimal path.

A large number of RREQs leads to frequent change in topology, and thus impede the transmission of information packets. Also excessive requests might cause congestion and ultimately packet loss. In [33], efficiency is achieved by selecting a carrier group of vehicles to relay the data packet to the destination. This group maintains the vehicle information only in those streets that are indispensable in reaching the destination.

It is observed that the problem of broadcast storm during the route finding phase still needs to be addressed. Therefore, in order to address that issue a routing protocol is proposed in this paper which not only minimizes the routing overhead but also increases the efficiency of the overall network. The proposed routing protocol is influenced by the above mentioned protocols. It uses the vehicles' positions and limits the transmission in a specific zone. Extensive simulations have been carried out in the NS2 simulator. Node mobility has been generated by the Simulation of Urban MObility (SUMO) whereas real environment of the traffic was generated by OpenStreetMap.

II. THE PROPOSED ROUTING PROTOCOL

Much research has been being carried out to improve the efficiency of routing protocols in VANETs. Densely populated networks with excessive number of nodes and their scalability issue have drawn much attention. A novel approach is proposed in this work to increase throughput and PDR and to decrease the overhead packet count. The route discovery is initiated from the source by sending route request packets to all neighbors. The neighbor nodes subsequently forward the route request to their respective neighbors. The process continues until the packet reaches its destination. The route reply packet confirms the route establishment. In the proposed protocol, initially the source node generates a trilateral zone using the destination vehicle's last known position where the destination vehicle could reside and makes a list of nodes which lie inside that zone. Information about the list of nodes is added to the RREQ packet header which is sent to the neighbors. Each neighbor who receives the RREQ packet checks the list of nodes. If its information related to the trilateral membership is present in that list, then it can participate, otherwise the packet will be dropped. The receiving node checks whether the packet targets its own vehicle. If not, it then calculates a new trilateral zone using the updated position of the destination vehicle and adds that information in its new RREQ packet. The process continues until the destination node is reached. The proposed protocol avoids unnecessary RREQs, thus efficient usage of bandwidth is achieved during route discovery.

A. DyTE – The Algorithm

The proposed protocol (DyTE) combines the power of position-based routing and Geocast-based routing protocol because it uses the coordinates of the vehicles and limits the packet request in a limited trilateral zone. In Figure 2, an area is shown where the trilateral zone is extracted as a zone of interest and the remaining area (red colored) is excluded.

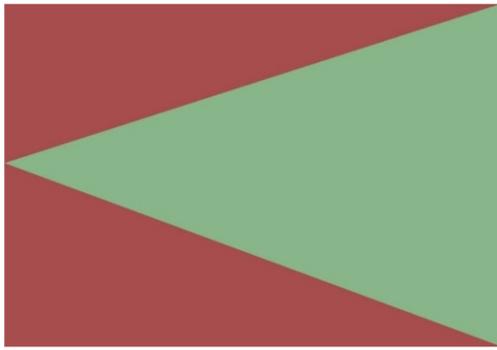


Fig. 2. A trilateral zone example.

The sequence of steps to be taken is (Figure 3):

- Get source and destination location through GPS.
- Calculate the trilateral zone where the broadcast of packets is going to be accepted.
- Add the list of nodes to the RREQ packet.
- Each node that receives the RREQ packet will examine the list to identify its trilateral zone's membership.
- Only if the receiving vehicle (destination/intermediate) lies within the trilateral zone, it can process it further. Else, the packet will be dropped.
- If the request received by a vehicle lies within the trilateral zone, it checks whether that packet was destined to it. If it does not, it will create a new list of trilateral zone's membership to forward the packet but with a newly calculated list of members.
- The above steps are repeated until the destination node is reached.

B. Creation of Trilateral Zone

When an intermediate node receives the request, it checks whether the packet is generated by its own vehicle in order to avoid looping. The request can be seen initially by every neighbor node that lies within the wireless transmission range but the RREQ packet header contains the special array field that consists of a list of nodes which are allowed to take part in the RREQ process. To create a trilateral zone, we used source and destination vehicle's GPS coordinates to find the slope (1) of the straight line between the source (S_x, S_y) and destination (D_x, D_y) vehicle.

$$m = \frac{\Delta y}{\Delta x} = \frac{D_y - S_y}{D_x - S_x} \quad (1)$$

The distance between source and destination nodes is calculated by (2):

$$\text{dist} = \sqrt{(D_x - S_x)^2 + (D_y - S_y)^2} \quad (2)$$

After calculating the slope and distance, the perpendicular slope is calculated by (3):

$$m_{\text{perpendicular}} = -\frac{1}{m} \quad (3)$$

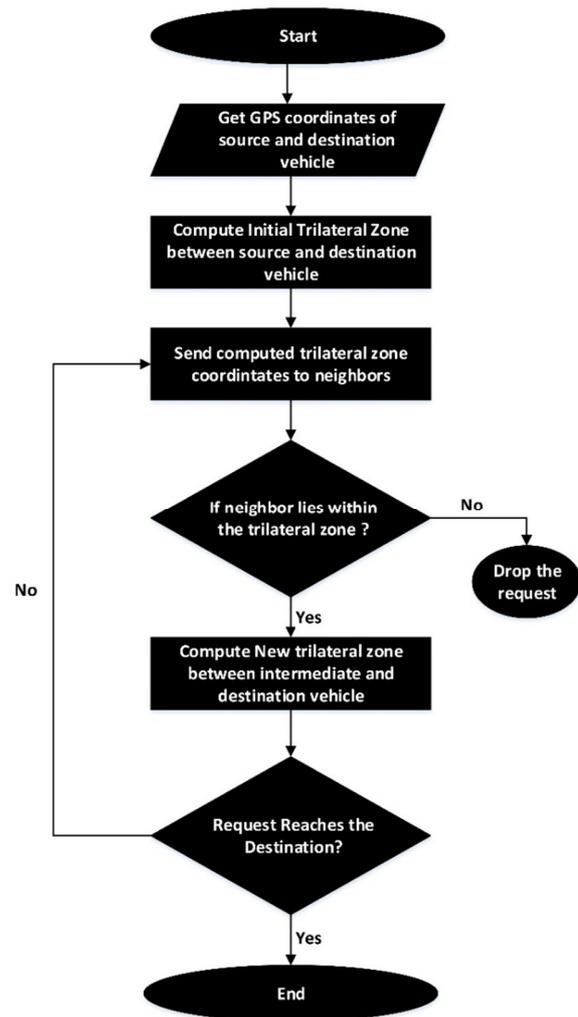


Fig. 3. Flowchart of the algorithm.

C. Trilateral Enrolment Identification Process

The next step is to identify whether the receiving node exists within the trilateral zone. To find out that a given node can participate in RREQ flooding, consider the trilateral zone, as shown in Figure 4, where the coordinates of A, B, C and an arbitrary node P are given as (A_x, A_y), (B_x, B_y), (C_x, C_y) and (P_x, P_y) respectively. In order to allow the node to get the enrolment for the RREQ forwarding process, DyTE goes through calculating the area and sub areas for each node.

The area of the trilateral zone (Triangle ABC) as shown in Figure 4 is calculated by:

$$\text{Area}_{\text{total}} = \frac{|A_x \times (B_y - C_y) + B_x \times (C_y - A_y) + C_x \times (A_y - B_y)|}{2} \quad (4)$$

The sub area #1 of the trilateral zone (triangle PAB) is calculated by:

$$\text{Sub}_{\text{Area1}} = \frac{|A_x \times (B_y - P_y) + B_x \times (P_y - A_y) + P_x \times (A_y - B_y)|}{2} \quad (5)$$

The sub area #2 of the trilateral zone (triangle PBC) is calculated by:

$$Sub_{Area2} = \frac{|P_x \times (B_y - C_y) + B_x \times (C_y - P_y) + C_x \times (P_y - B_y)|}{2} \quad (6)$$

The sub area #3 of the trilateral zone (triangle PAC) is calculated by:

$$Sub_{Area3} = \frac{|A_x \times (P_y - C_y) + P_x \times (C_y - A_y) + C_x \times (A_y - P_y)|}{2} \quad (7)$$

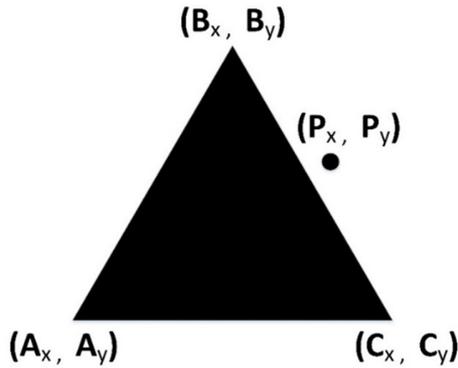


Fig. 4. Coordinates of a trilateral zone of an arbitrary point P.

D. Forwarding Path and Request / Reply Process in DyTE

Each participating node during the phase of RREQ discovery sets up a special entry called reverse route which contains the IP address of the source node, a sequence number along with hop counts to the source node, and the IP address of the neighbor from which the RREQ was obtained. Since no information should be kept in forever in the routing table, a timer variable called lifetime is associated with the reverse route so that if the route entry is not used for the specified allocated time, then the information will be wiped off from the table to minimize the routing overhead. Routing loops are handled by the sequence numbers with source node information in order to check the freshness of the RREQs of the unicasted response of the RREP back to the node. Broadcasting of a route request within a trilateral zone is permitted again only when an intermediate node gets out of contact, because disconnection of the intermediate node will make the destination node inaccessible. The RREP packet contains the information of the source and the IP address of the destination. If the destination node is able to reply, then it resets the hop count information and the sequence number and lifetime are also updated in the RREP packet. In case the intermediate node is able to reply, then the hop count will be updated as per the intermediate node's distance to the destination node along with the destination sequence number information and the time validity of routes is calculated within the routing table.

A forwarding path is stored in the routing table when the RREP packet is received by an intermediate node. This path entrance contains the neighbor's IP address from which the RREP is received, the IP address of the destination and the hop counts to the destination. To obtain the distance of the

destination, the receiving node increases the value in the hop count by one and also associates this entry with the lifetime. The associated lifetime gets updated whenever the route is used and is deleted when the route is not used within the specified time limit. An RREP can be received from multiple neighbors for a similar destination. In that case the RREP which is received first will be considered and the alternative RREPs are only used if they contain information with smaller hop count or if the destination sequence number is greater to ensure quick and updated routing. In other words, after receiving the first RREP, the source node will begin the data transmission process and if it discovers, at a later stage, a better route, then the routing information will be updated and a different path will be taken for data transmission.

III. SIMULATION AND RESULTS

Network simulation was performed on NS-2 simulator [34]. For mobility simulation within urban environment, SUMO [35] was used. Figure 5(a) shows the simulation plot which is an original image containing a map portion of Karachi extracted from OpenStreetMap [36] whereas Figure 5(b) is the network animation within the NS-2 simulator and Figure 5(c) is the converted map using SUMO. The performance of the proposed protocol is compared with the performance of other popular routing protocols (AODV, AOMDV, and DSR). The metrics considered for comparison are PDR, network routing load (overhead), and throughput. To avoid the possibility of fluctuations in simulations, the presented results are the average of multiple simulations. The network is considered from moderate to highly-dense and random distribution of nodes was incorporated. The simulation parameters are given in Table I.

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Mobility	Manhattan
Channel type	Wireless
Antenna	Omni directional
Model	Two ray ground
No. of vehicles	50, 100, 150, 200, 250
Simulation area	2500×2000m ²
Simulation time	100s
Routing protocols	AODV, AOMDV, DSR, DyTE
Speed range	50–100Km/h
Transmission zone	250m
Connection type	UDP
Traffic type	CBR
Packet rate	4 per second

In Figure 6, the NRL is depicted which can be defined as the proportion between the number of packets generated by each node and the number of packets successfully transmitted to the destination nodes. Since only relevant packets are allowed to get transmitted over the network, improvement is achieved by the proposed DyTE protocol. Figure 7 represents the PDR which is defined as the ratio of data packets that are successfully received at the destination with the number of data packets generated at the source node. The nodes enrolled in the trilateral zone are the most favorable to transmit the request to the destination, therefore the results are improved. Figure 8 depicts the throughput of the network which is the cumulative

number of packets transmitted successfully to the destination over the entire simulation. The comparison shows a direct relation between the number of nodes and the number of delivered packets to the destination node, reflecting a higher throughput.

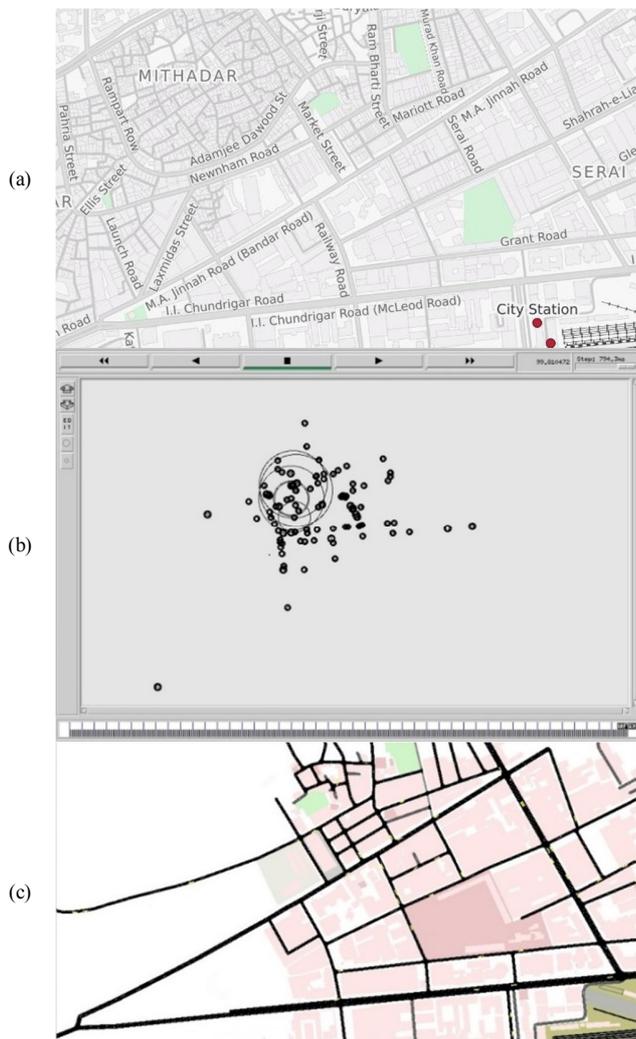


Fig. 5. (a) The map of a part of Karachi, (b) Network animation in NS-2, (c) SUMO map conversion.

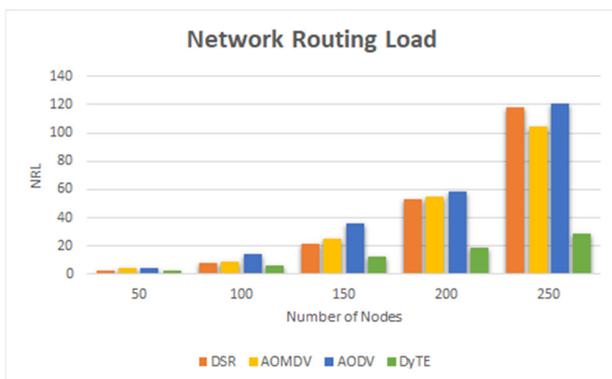


Fig. 6. NRL against node density.

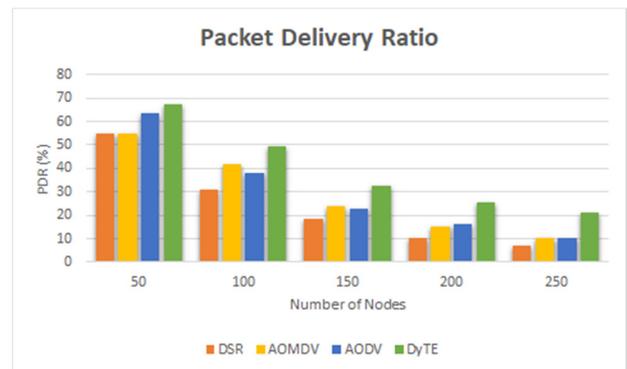


Fig. 7. PDR against node density.

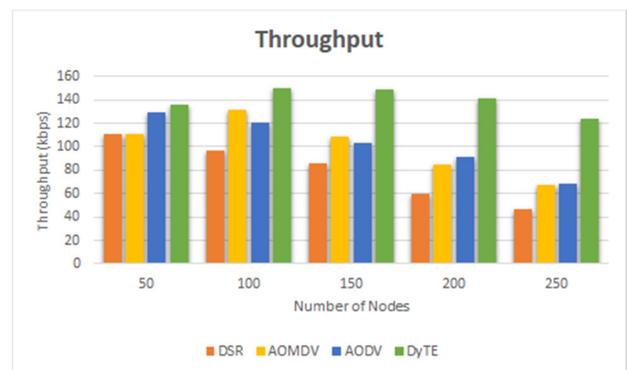


Fig. 8. Throughput against node density.

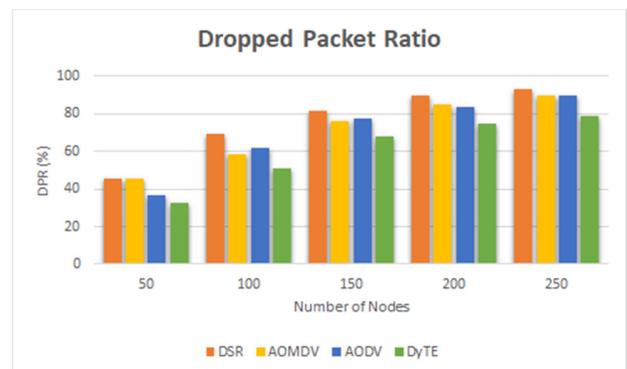


Fig. 9. Dropped packet ratio against node density.

Figure 9 shows the dropped packet ratio which is a reciprocal of PDR, since PDR is increased therefore it is quite understandable that the dropped packet ratio is decreased. The request packets are controlled before sending the actual data to minimize the unnecessary overhead. It is conclusive that minimizing the unnecessary route request within the network actually lowers the burden on the overall network. Since only expected and reachable nodes can participate in the route request phase, NRL is minimized, the chance of reaching the packet to the destination is increased, and therefore PDR is increased.

IV. CONCLUSION

A novel way of limiting the request zone for the effective communication is presented in this paper with focus given in

essential parameters such as NRL, throughput, and PDR. The main challenge in VANETs is designing routing protocols for dynamic topology along with the consideration of mobility. The proposed routing protocol DyTE uses the location of the vehicles and limits the area of communication. The results have shown that DyTE is not only efficient in increasing the PDR and throughput but also greatly minimized the NRL. DyTE has been found to have better performance than AODV, AOMDV, and DSR in VANETs.

ACKNOWLEDGEMENT

The research reported in this study was conducted at the NED University of Engineering and Technology (NEDUET) and is funded by MoST (Ministry of Science and Technology).

REFERENCES

- [1] Y. Kim and I. Kim, "Security issues in vehicular networks," in *The International Conference on Information Networking*, Bangkok, Thailand, Jan. 2013, pp. 468–472, <https://doi.org/10.1109/ICIN.2013.6496424>.
- [2] C. Wan and J. Zhang, "Efficient identity-based data transmission for VANET," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 6, pp. 1861–1871, Nov. 2018, <https://doi.org/10.1007/s12652-017-0650-x>.
- [3] F. Cunha *et al.*, "Data communication in VANETs: Protocols, applications and challenges," *Ad Hoc Networks*, vol. 44, pp. 90–103, Jul. 2016, <https://doi.org/10.1016/j.adhoc.2016.02.017>.
- [4] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey," *IEEE Vehicular Technology Magazine*, vol. 2, no. 2, pp. 12–22, Jun. 2007, <https://doi.org/10.1109/MVT.2007.912927>.
- [5] A. K. Kazi and S. M. Khan, "Working of various routing protocols in Vehicular Ad-hoc Network: A Survey," *University of Sindh Journal of Information and Communication Technology*, vol. 4, no. 4, pp. 278–286, Dec. 2020.
- [6] Y. Ko and N. H. Vaidya, "Location-Aided Routing (LAR) in mobile ad hoc networks," *Wireless Networks*, vol. 6, no. 4, pp. 307–321, Sep. 2000, <https://doi.org/10.1023/A:1019106118419>.
- [7] G. G. Finn, "Routing and Addressing Problems in Large Metropolitan-Scale Internetworks," University Of Southern California Marina Del Rey Information Sciences Institute, ADA180187, Mar. 1987. Accessed: Mar. 10, 2021. [Online]. Available: <https://apps.dtic.mil/sti/citations/ADA180187>.
- [8] R. S. Raw, D. K. Lobiyal, S. Das, and S. Kumar, "Analytical Evaluation of Directional-Location Aided Routing Protocol for VANETs," *Wireless Personal Communications*, vol. 82, no. 3, pp. 1877–1891, Jun. 2015, <https://doi.org/10.1007/s11277-015-2320-7>.
- [9] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *6th annual international conference on Mobile computing and networking*, New York, NY, USA, Aug. 2000, pp. 243–254, <https://doi.org/10.1145/345910.345953>.
- [10] S. A. Rao, M. Pai, M. Boussejra, and J. Mouzna, "GPSR-L: Greedy perimeter stateless routing with lifetime for VANETS," in *8th International Conference on ITS Telecommunications*, Phuket, Thailand, Oct. 2008, pp. 299–304, <https://doi.org/10.1109/ITST.2008.4740275>.
- [11] C. Lochert, M. Mauve, H. Fubler, and H. Hartenstein, "Geographic routing in city scenarios," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 9, no. 1, pp. 69–72, Jan. 2005, <https://doi.org/10.1145/1055959.1055970>.
- [12] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," in *Mobile Computing*, T. Imielinski and H. F. Korth, Eds. Boston, MA, USA: Springer, 1996, pp. 153–181.
- [13] C. E. Perkins, E. M. Belding-Royer, and S. R. Das, *Ad hoc On-Demand Distance Vector (AODV) Routing*. The Internet Society, 2003.
- [14] A. S. Azman, M. Y. Lee, S. K. Subramaniam, and F. S. Feroz, "Performance Evaluation of Routing Protocols for Wireless Mesh Network using Grid Topology Suitable for Downstream Oil and Gas Pipeline," in *International Conference on Computing and Information Technology*, Tabuk, Saudi Arabia, Sep. 2020, <https://doi.org/10.1109/ICCIT-144147971.2020.9213718>.
- [15] S. K. Subramaniam, S. M. Khan, and R. Nilavalan, "Static Network Performance Optimisation Using Dual Interleave Routing Algorithm," *IJUM Engineering Journal*, vol. 19, no. 1, pp. 129–143, Jun. 2018, <https://doi.org/10.31436/iiumej.v19i1.841>.
- [16] M. Y. Lee, A. S. Azman, S. K. Subramaniam, and F. S. Feroz, "Performance Analysis of Linear Topology Wireless Sensor Network in Oil and Gas Industry," *IOP Conference Series: Materials Science and Engineering*, vol. 765, Mar. 2020, Art. no. 012070, <https://doi.org/10.1088/1757-899X/765/1/012070>.
- [17] H. Guo, F. B. A. Thani, W. C. Wong, and Y. Wu, "An optimized routing protocol for vehicular ad hoc networks," in *TENCON 2010 - 2010 IEEE Region 10 Conference*, Fukuoka, Japan, Nov. 2010, pp. 245–250, <https://doi.org/10.1109/TENCON.2010.5685996>.
- [18] D. Sutariya and S. Pradhan, "An improved AODV routing protocol for VANETs in city scenarios," in *International Conference On Advances In Engineering, Science And Management*, Nagapattinam, India, Mar. 2012, pp. 575–581.
- [19] M. K. Marina and S. R. Das, "Ad hoc on-demand multipath distance vector routing," *Wireless Communications and Mobile Computing*, vol. 6, no. 7, pp. 969–988, 2006, <https://doi.org/10.1002/wcm.432>.
- [20] B. T. Sharef, R. A. Alsaqour, and M. Ismail, "Vehicular communication ad hoc routing protocols: A survey," *Journal of Network and Computer Applications*, vol. 40, pp. 363–396, Apr. 2014, <https://doi.org/10.1016/j.jnca.2013.09.008>.
- [21] M. A. Soomro, M. I. Channa, Z. Hussain, and M. Ahmed, "EPLAODV: Energy Priority and Link Aware Ad-hoc On-demand Distance Vector Routing Protocol for Post-disaster Communication Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5148–5152, Feb. 2020, <https://doi.org/10.48084/etasr.3223>.
- [22] S. M. Khan, R. Nilavalan, and A. F. Sallama, "A Novel Approach for Reliable Route Discovery in Mobile Ad-Hoc Network," *Wireless Personal Communications*, vol. 83, no. 2, pp. 1519–1529, Jul. 2015, <https://doi.org/10.1007/s11277-015-2461-8>.
- [23] S. J. Gudakahriz, S. Jamali, M. V. Khiavi, and A. Soleimany, "A Stable TORA Based for Routing in Mobile Ad Hoc Networks," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2532–2536, Feb. 2018, <https://doi.org/10.48084/etasr.1720>.
- [24] S. M. Khan, M. M. Khan, N. A. Khan, and Waseemullah, "Efficient and Reliable Reactive Routing Protocol for Mobile AdHoc Network," *International Journal of Computer Science and Network Security*, vol. 17, no. 4, pp. 238–244, 2017.
- [25] K. Ullah, A. Sayyed, M. Aziz, and E. Moreira, "A beaconing-based roadside services discovery protocol for vehicular ad hoc networks," *Turkish Journal of Electrical Engineering and Computer Science*, vol. 27, no. 3, pp. 2036–2051, Jun. 2019.
- [26] K. Feng, C. Hsu, and T. Lu, "Velocity-Assisted Predictive Mobility and Location-Aware Routing Protocols for Mobile Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 448–464, Jan. 2008, <https://doi.org/10.1109/TVT.2007.901897>.
- [27] S. L. O. B. Correia, J. Celestino, and O. Cherkaoui, "Mobility-aware Ant Colony Optimization routing for vehicular ad hoc networks," in *IEEE Wireless Communications and Networking Conference*, Cancun, Mexico, Mar. 2011, pp. 1125–1130, <https://doi.org/10.1109/WCNC.2011.5779289>.
- [28] H. Rana, P. Thulasiraman, and R. K. Thulasiram, "MAZACORNET: Mobility aware zone based ant colony optimization routing for VANET," in *IEEE Congress on Evolutionary Computation*, Cancun, Mexico, Jun. 2013, pp. 2948–2955, <https://doi.org/10.1109/CEC.2013.6557928>.
- [29] H. Bello-Salau, A. M. Aibinu, Z. Wang, A. J. Onumanyi, E. N. Onwuka, and J. J. Dukiya, "An optimized routing algorithm for vehicle ad-hoc networks," *Engineering Science and Technology, an International Journal*, vol. 22, no. 3, pp. 754–766, Jun. 2019, <https://doi.org/10.1016/j.jestch.2019.01.016>.

- [30] B. Moussaoui, S. Djahel, H. Khelifi, and S. Merniz, "Towards enhanced reactive routing in urban Vehicular Ad hoc Networks," in *International Conference on Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems*, Paris, France, Jul. 2015, pp. 1–6, <https://doi.org/10.1109/NOTERE.2015.7293506>.
- [31] T. Sivakumar and R. Manoharan, "ERP: An efficient reactive routing protocol for dense vehicular ad hoc networks," *Turkish Journal of Electrical Engineering and Computer Science*, vol. 25, no. 3, pp. 1762–1772, Jun. 2017.
- [32] R. Gandhi and S. Parthasarathy, "Distributed algorithms for connected domination in wireless networks," *Journal of Parallel and Distributed Computing*, vol. 67, no. 7, pp. 848–862, Jul. 2007, <https://doi.org/10.1016/j.jpdc.2007.04.003>.
- [33] S. Arianmehr and M. A. Jabraeil Jamali, "HybTGR: a hybrid routing protocol based on topological and geographical information in vehicular ad hoc networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 4, pp. 1683–1695, Apr. 2020, <https://doi.org/10.1007/s12652-019-01332-z>.
- [34] T. Issariyakul and E. Hossain, *Introduction to Network Simulator NS2*. Boston, MA, USA: Springer, 2008.
- [35] S. R. Santana, J. J. Sanchez-Medina, and E. Rubio-Royo, "How to Simulate Traffic with SUMO," in *Computer Aided Systems Theory – EUROCAST 2015*, Canary Islands, Spain, Feb. 2015, pp. 773–778, https://doi.org/10.1007/978-3-319-27340-2_95.
- [36] J. Thebault-Spieker, B. Hecht, and L. Terveen, "Geographic Biases are 'Born, not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap," in *Proceedings of the 2018 ACM Conference on Supporting Groupwork*, New York, NY, USA, Jan. 2018, pp. 71–82, <https://doi.org/10.1145/3148330.3148350>.

Improving the Recognition Performance of Lip Reading Using the Concatenated Three Sequence Keyframe Image Technique

Lap Poomhiran

Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok
Bangkok, Thailand
lap.p@email.kmutnb.ac.th

Phayung Meesad

Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok
Bangkok, Thailand
pym@kmutnb.ac.th

Sumitra Nuanmeesri

Faculty of Science and Technology
Suan Sunandha Rajabhat University
Bangkok, Thailand
sumitra.nu@ssru.ac.th

Abstract-This paper proposes a lip reading method based on convolutional neural networks applied to Concatenated Three Sequence Keyframe Image (C3-SKI), consisting of (a) the Start-Lip Image (SLI), (b) the Middle-Lip Image (MLI), and (c) the End-Lip Image (ELI) which is the end of the pronunciation of that syllable. The lip area's image dimensions were reduced to 32×32 pixels per image frame and three keyframes concatenate together were used to represent one syllable with a dimension of 96×32 pixels for visual speech recognition. Every three concatenated keyframes representing any syllable are selected based on the relative maximum and relative minimum related to the open lip's width and height. The evaluation results of the model's effectiveness, showed accuracy, validation accuracy, loss, and validation loss values at 95.06%, 86.03%, 4.61%, and 9.04% respectively, for the THDigits dataset. The C3-SKI technique was also applied to the AVDigits dataset, showing 85.62% accuracy. In conclusion, the C3-SKI technique could be applied to perform lip reading recognition.

Keywords-concatenated frame images; convolutional neural network; keyframe reduction; keyframe sequence; lip reading

I. INTRODUCTION

Deep learning applications, especially Convolutional Neural Network (CNN) applications, have recently achieved impressive success in diverse object detection and recognition tasks [1]. However, CNNs face some challenges, in particular in video recognition. A video may be incomplete and sound may be lacking during certain parts. If the audio at the crucial moment is missing, it may result in the video's contents being misunderstood [2]. These videos will be more useful if they were edited and the missing words or messages could be found. Most of the proposed solutions rely on the lip reading technique to help transcription by reading and observing the moving lips, including tongue and face to get the right words.

Corresponding author: Lap Poomhiran

Moreover, the process of transcribing or translating the speech obtained by lip reading is a skill that requires learning and practice until becoming proficient at recognizing the lip movement or lip pattern related to the pronunciation of each syllable.

In general, there are two popular multimodals or methods of supervised learning for lip reading: Visual Speech Recognition (VSR) and Audio-Visual Speech Recognition (AVSR). VSR uses a method to teach the machine with visual-only information from a video without using speech or audio for training [3]. On the other hand, AVSR trains the machine by applied images combined with audio data from a video to achieve greater accuracy [4]. Authors in [5] found that the use of Visual-Only (VO) data has better classification accuracy than that of Audio Visual (AV) and Audio Only (AO) data. There are currently two groups of research studies on lip reading recognition, the first group uses images in VO [2, 3, 5-14] while the other group uses both audio and video together [4, 15-25]. Both these groups have a model for extracting the features with a technique used in combination. Nevertheless, the latter group differs in that they have to combine audio feature data with visual features when it comes to machine learning. AV speech recognition is used commercially on various software or systems, but the recognition quality is reduced by environmental noise. This situation is not the same as using data from a quiet image, making lip reading a pivotal role in Automatic Speech Recognition (ASR) in harsh audio environments [6].

Neural networks are commonly used to help extract features and recognize lip reading patterns in machine learning, such as in CNNs [6, 11, 24], Long Short-Term Memory (LSTM) [3, 10, 21], or a combination of CNN with LSTM [2,

12, 13]. As mentioned above, in machine learning for lip reading recognition, visual training is required. The images of the face or lips are an essential factor as the primary input. This makes the number of images arranged for use as a large input. In order to sort the images of the video frame, the lip movement is determined by the number of frames, for example, 40 frames per word [7], or the number of frames in seconds such as 0.2 seconds [8, 16] or 1 second, or maybe determined using every frame in the video wherein one video will have only one word and a short duration. The methods mentioned above also use multiple frames in machine learning processing, which consumes more resources and processing time. There are also real world problems, e.g. where a speaker speaks at a slowly speaking speed, resulting in a longer video file length while the words are spread out. Therefore, limitations on the number of frames or the duration may differ from the words or messages conveyed. However, if there is a representation of the images or keyframes, it will reduce the number of images and cost used in machine learning.

Most lip reading studies are applied to English datasets of digits, alphabets, words, phrases, and sentences. The AVDigits [26] is a popular dataset for testing and improving the performance of a model. There is a total of 540 videos with a resolution of 1920×1080 pixels (px). It consists of six speakers who speak numbers from 0 to 9 in English and each repeats the numbers 9 times. Greek [12], Myanmar [14], Spanish [27], and Czech [28] have been also studied, but there are no Thai language datasets created for lip reading. It is difficult for the Thai language to be in some words with similar lip patterns but with different meanings. Since the vowels are pronounced similarly, the patterns of lip movements are similar. There is a pattern of intonation in Thai language resulting from a combination of tones, resulting in intonation in five tonal sounds. Therefore, this research aims to create a dataset in the Thai language and reduce the image dimensions and the number of frames by finding keyframes to replace syllables or words for Thai lip reading recognition using CNNs.

II. RESEARCH METHODOLOGY

The research methodology for the improving performance recognition of lip reading using C3-SKI consists of 1) dataset preparation, 2) face detection and lip localization, 3) C3-SKI creation, 4) model development, and 5) model effectiveness evaluation.

A. Dataset Preparation

The Thai digit dataset is called THDigits. It was created as a Thai video file containing numbers from 0 to 9, which are repeated three times with three different speaking speeds (slow, regular, and fast) by using 100 mixed-gender speakers. A total of 3,000 video files with length between 1 and 4 seconds were constructed. These videos have four different resolutions: 1920×1080 , 1280×720 , 960×540 , and 720×404 px and were recorded with smartphones, regardless of model and brand. The Thai numbering is shown in Table I.

B. Face Detection and Lip Localization

Before processing face detection and lip locating, each video prepared in the Thai digit dataset was separated into

individual frames, ordered from the first to the last frame. After that, the face was detected in each frame using the Viola-Jones [29, 30] technique based on the Haar-like feature. The hypothesis T represents any distinguishing characteristic, h is the distinguishing characteristic, and β represents the percentage of error classification. The characteristic $C(x)$ is given as (1) [29]:

$$C(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where $\alpha_t = \log \frac{1}{\beta_t}$.

The difference of pixel sums or features on an image is compared using the Haar-like feature with the black and white filter that assigned I and P values to represent N by N images as in (2) [30]:

$$\sum_{1 \leq i \leq N} \sum_{1 \leq j \leq N} I(i, j) 1_{P(i, j)=white} - \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq N} I(i, j) 1_{P(i, j)=black} \quad (2)$$

After face detection comes the lip positioning based on 68 facial landmarks [31]. Lip localization determines the lip area or Region Of Interest (ROI) to leave only the desired feature area by cropping in each frame to the lip area only, illustrated in Figure 1.

TABLE I. THAI NUMBER PRONUNCIATION IN THDIGITS DATASET

Numbers	Thai pronunciation	Meaning
0	Šūny	Zero
1	Hñung	One
2	Šxng	Two
3	Šām	Three
4	Šī	Four
5	Hā	Five
6	Hk	Six
7	Cēd	Seven
8	Pæd	Eight
9	Kēā	Nine

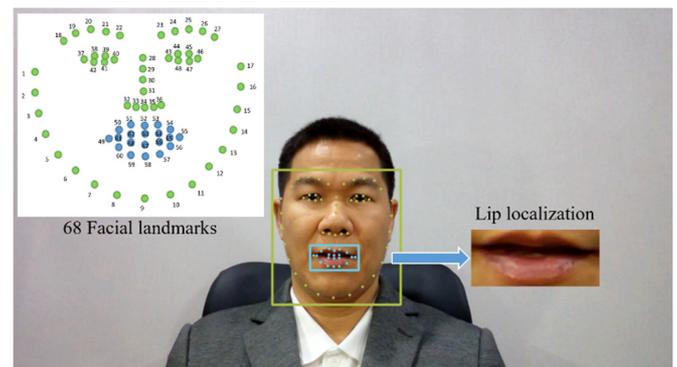


Fig. 1. Face detection and lip localization on each image frame.

C. The Concatenated Three Sequence Keyframe Image

Recent research favorably replaces any word or syllable by concatenating frames as a single image before training. For example, some research uses frames together as a single image

in 5 rows with 5 frames per row, in a total of 25 frames per image [6]. As mentioned above, if the speaker speaks slowly or extensively, the number of frames will be increased, but only one word or syllable will be conveyed. In this study, only the keyframes with pronounced lip movement were selected, relying on measuring the outer lip dimensional from the mouth's height (h) and width (w) for each frame x as shown in Figure 2.



Fig. 2. Lip width (w) and lip height (h) for the current and the next frame.

Lip dimensions were calculated by (3) on each frame. With the help of the increasing function, decreasing function, relative maximum, and relative minimum of a graph related to lip dimensions, all keyframes were extracted and selected. The relative maximum and relative minimum are lying on the slope at the x interval. If the c is some x in a graph, where c is the relative maximum or relative minimum, then the slope at c would be zero, represented in (4).

$$f(x) = w_x + h_x \quad (3)$$

$$f'(c) = 0 \quad (4)$$

$f'(c)$ is the function that has a relative maximum or minimum at c , $f(x) \leq f(c)$ for the relative maximum, and $f(x) \geq f(c)$ for the relative minimum.

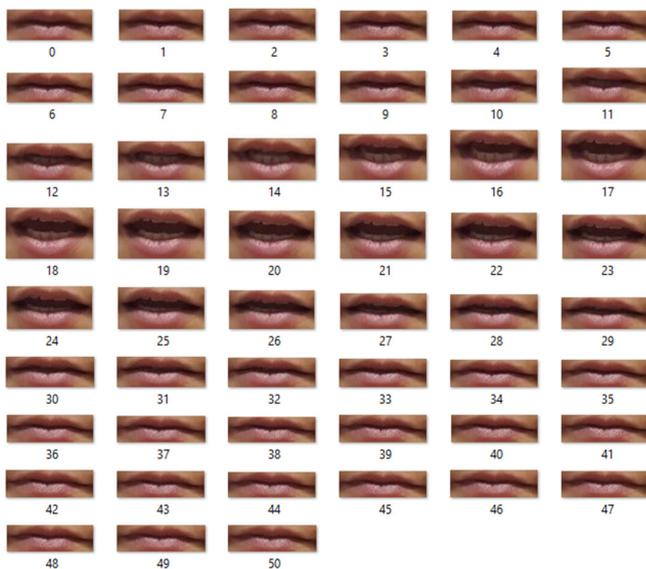


Fig. 3. Image frame sequence from a sample video of the dataset.

Each video is possible to produce several relative maxima and relative minima as keyframes. In this experiment, there are three keyframes assigned as a single syllable. The first keyframe will represent the frame where the mouth starts to

pronounce a syllable. This is the SLI. The second keyframe (MLI) is the frame where the mouth is at the maximum opening limit or movement of the syllable. The last keyframe (ELI) is at the end of the pronunciation of that syllable. For example, the word 'one' in English is pronounced as 'nueng' (หนึ่ง) in Thai language. Fifty one frames were split from a video in the dataset and the area of the lip was cropped as shown in Figure 3. The frame sequence numbers 9, 18, and 32 are representing the three keyframes based on relative maximum and relative minimum, and are shown in Figure 4.

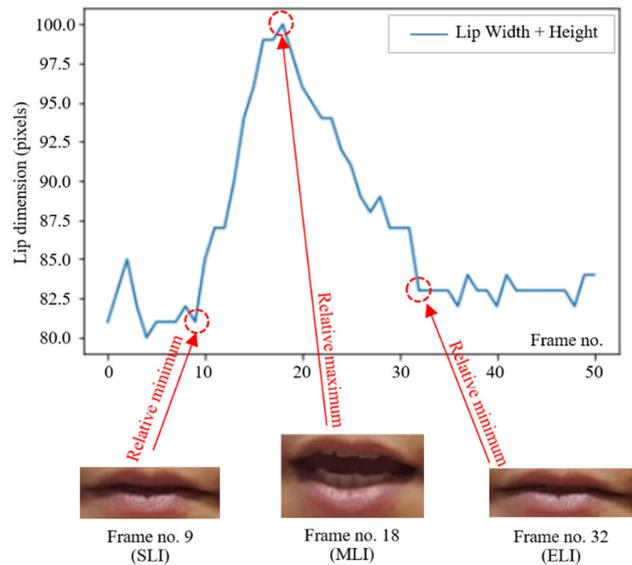


Fig. 4. The three keyframes representing a single syllable.

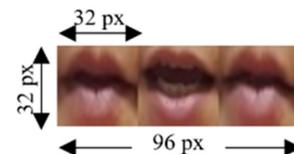


Fig. 5. The image dimensions for training in the current study.

After the three keyframes were selected, each keyframe image was scaled down to 32×32 pixels and merged into a single RGB color image with a dimension of 96×32 pixels, as shown in Figure 5, while other studies use image dimensions of 80×60 or 224×224 pixels ([7] and [6] respectively). Thus, each input color image frame as the dataset for building the model has a pixel density ratio of 3,072 pixels, which is less than the one used in other studies [6, 7].

D. Model Development

The model was designed based on the CNNs with a total of 13 layers. It consists of 1 normalization layer, 6 convolutional (Conv) layers with or Rectified Linear Unit (ReLU) activation function for each convolutional layer, 3 max-pooling layers, and 3 Fully-Connected (FC) layers, which include a flatten layer and two dense layers. There are a total of 846,890 parameters for training. The model architecture is shown in Figure 6.

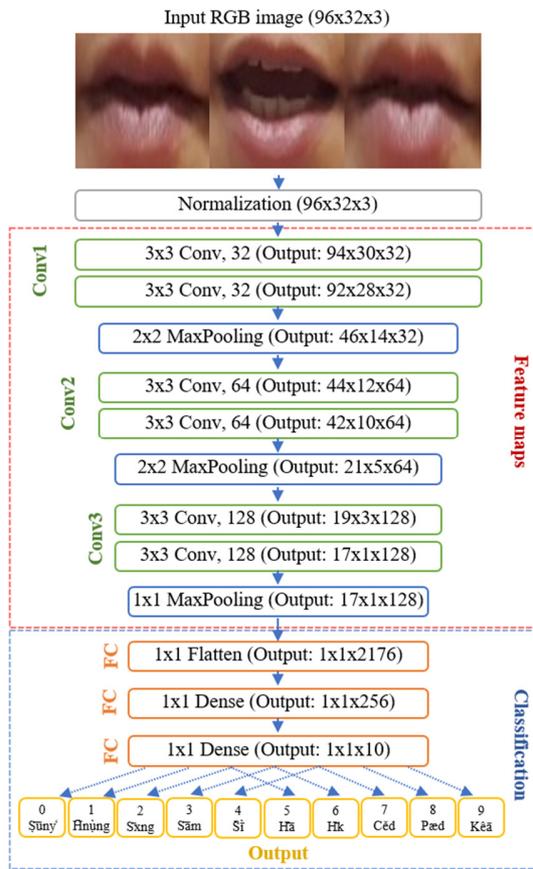


Fig. 6. Model architecture for lip reading applied to C3-SKI.

All images of the Thai digit dataset were used in the training and validation of the designed model. The training and validation images were the 80% and 20% of the dataset respectively. A total of 200 epochs of training was set with a minibatch gradient descent size of 32. The model was built using an Intel Core i7-7700HQ PC at 2.80GHz, 16GB of memory, 512GB of Samsung Solid State Drive, and a 3GB NVIDIA GeForce GTX 1060. This system was running through Python version 3.8.3 on Windows 10 x64 architecture.

E. Model's Effectiveness Evaluation

The developed model has validated by the accuracy rate and loss value. The accuracy was calculated by (5), and the loss value or loss function, also known as cross-entropy which is the favorite function for classification is defined in (6) [32-38]:

$$Accuracy = \frac{C}{N} \quad (5)$$

where C refers to the total number of samples recognized correctly, N refers to the total number of all samples.

$$Cross - entropy = - \sum_{i=1}^N \sum_{j=1}^M y_{i,j} \log(p_{i,j}) \quad (6)$$

where N refers to the total number of all samples, M refers to the total number of classes, $y_{i,j}$ refers to the true value when the

sample i belongs to the class j , $p_{i,j}$ refers to the probability predicted by the sample model of i to belong to the class j .

The model will stop training automatically using the 'EarlyStopping' feature supported by Keras library with the callback function. The patience value was 5, and monitoring to the validation loss value. The model stopped training when there was no improvement in the validation loss for 5 consecutive epochs. The minimum validation loss was reached at 37 epochs of training.

III. RESULTS

According to the experiments on the developed model, the accuracy and loss value were determined while building the model. After the 37 epochs of training, the model gave 95.06% accuracy and 4.61% loss value. The model stopped training after 47 epochs. Considering the model's validation, the model gave validation accuracy value of 86.03% and validation loss of 9.04%. The accuracy and loss for both training and validation are shown in Figures 7-9.

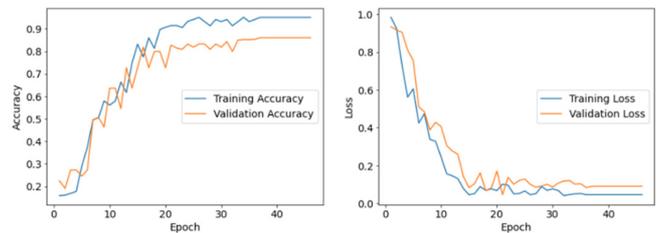


Fig. 7. The accuracy value of training and validation (left) and the loss value of training and validation (right).

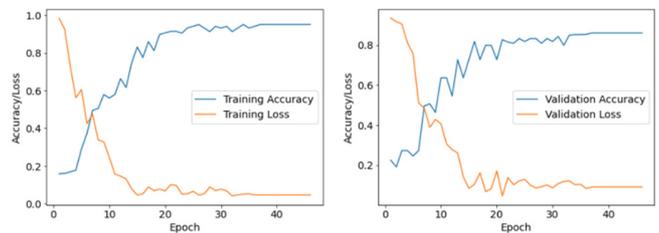


Fig. 8. The training accuracy and loss (left), and the validation accuracy and loss (right).

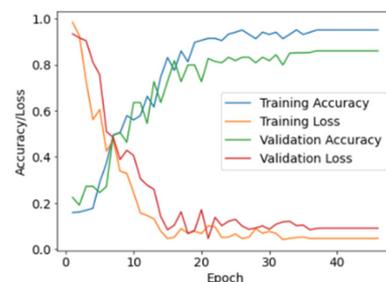


Fig. 9. The accuracy and loss of training and validation of the model.

Moreover, this experiment compares the accuracy of the technique for input image reduction with other studies using the AVDigits dataset. It was found that the lip reading models built by MDBN [39], MDAE [40], RTMRBM [26], am-LSTM

[41], and alm-GRU [20] gave accuracy of 55.00%, 66.74%, 71.77%, 85.23%, and 85.53%, respectively while this work gave 85.62%. The model performances are shown in Table II and Figure 10.

TABLE II. THE MODEL'S EFFECTIVENESS EVALUATION

Dataset	Method	Modality	Accuracy (%)
AVDigits	MDBN [39]	Visual-only	55.00
AVDigits	MDAE [40]	Audiovisual	66.74
AVDigits	RTMRBM [26]	Audiovisual	71.77
AVDigits	am-LSTM [41]	Audiovisual	85.23
AVDigits	alm-GRU [20]	Visual-only	85.53
AVDigits	C3-SKI	Visual-only	85.62
THDigits	C3-SKI	Visual-only	86.03

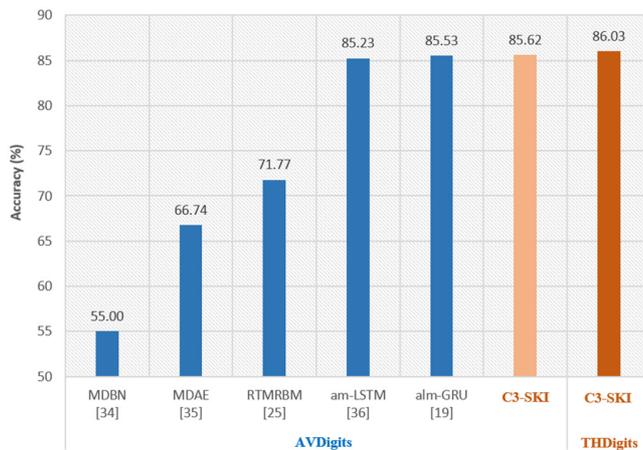


Fig. 10. Comparison with the state-of-the-art methods.

IV. CONCLUSION AND DISCUSSION

This paper proposes the application of the C3-SKI to a CNN for lip reading. The C3-SKI consisting of SLI, MLI, and ELI was tested in lip reading recognition on THDigits and AVDigits datasets. Its primary input is images applied to machine learning according to deep learning techniques. Reducing the number of images can be used by finding the image keyframes based on the relative maximum and relative minimum values. In this paper, 32×32 pixels of 3 sequence keyframes were assigned to represent any syllable of a digit between zero and nine, and these keyframes were concatenated to produce a new 96×32 px image as an input to the neural network. These new input images make the resulting image dimensions and density of pixels less than most state-of-the-art methods. Thus, there were 3,000 keyframe input images with 10 classes that were divided to 80% for training and 20% for validation. The developed CNN model included 1 normalization layer, 6 convolutional layers, 3 max-pooling layers, and 3 fully-connected layers. Training took a total of 47 epochs and was finalized when the validation loss value reached its minimum and the model did not improve any further. As a result, the model had accuracy, validation accuracy, loss, and validation loss values of 95.06%, 86.03%, 4.61%, and 9.04% respectively.

The model's accuracy was 85.62% when it was applied to the AVDigits dataset. Therefore, the reduced number of

keyframes with relative maximum and relative minimum can be applied in conjunction with CNN for lip reading. The results of this study conform to the conclusions of [6, 7] in which the CNN method with reduced concatenated frame images was applied to lip reading with a high level of effectiveness. These concatenated frame images have smaller dimension than the traditional images which are usually used to image classification in CNNs (224×224 px).

In future work, the researchers plan to create a dataset for Thai sentences and test the C3-SKI technique. New keyframes could be scaled down and compared to the number of keyframes used to represent each syllable for continuous speech or sentences, such as 5, 7, or 9 keyframe images sequentially. This technique will develop a function or equation to find the lip shift's common point from one syllable to the following syllable. Besides, word prediction methods from the corpus would be used for prediction combined with image processing to increase lip reading accuracy on deep learning in real-time processing.

ACKNOWLEDGMENT

The authors are grateful to the Graduate College, Faculty of Information Technology and Digital Innovation at King Mongkut's University of Technology North Bangkok, and the Faculty of Science and Technology, Suan Sunandha Rajabhat University, for supporting this research.

REFERENCES

- [1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings IEEE Computer Visualization and Pattern Recognit*, Las Vegas, NV, USA, 2016, pp. 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
- [2] S. Fenghour, D. Chen, and P. Xiao, "Decoder-encoder LSTM for lip reading," in *Proceedings of the 2019 8th International Conference on Software and Information Engineering*, Cairo, Egypt, Apr. 9-12, 2019, pp. 162–166, <http://doi.org/10.1145/3328833.3328845>.
- [3] S. Petridis, Z. Li, and M. Pantic, "End-to-end visual speech recognition with LSTMs," in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing*, New Orleans, LA, USA, Mar. 5-9, 2017, pp. 2592–2596, <https://doi.org/10.1109/ICASSP.2017.7952625>.
- [4] S. Chung, J. S. Chung, and H. Kang, "Perfect match: Improved cross-modal embeddings for audio-visual synchronisation," in *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Brighton, UK, May 12-17, 2019, pp. 3965–3969, <https://doi.org/10.1109/ICASSP.2019.8682524>.
- [5] R. Bi and M. Swerts, "A perceptual study of how rapidly and accurately audiovisual cues to utterance-final boundaries can be interpreted in Chinese and English," *Speech Communication*, vol. 95, pp. 68–77, 2017, <https://doi.org/10.1016/j.specom.2017.07.002>.
- [6] D. Jang, H. Kim, C. Je, R. Park, and H. Park, "Lip reading using committee networks with two different types of concatenated frame images," *IEEE Access*, vol. 7, pp. 90125–90131, 2019, <https://doi.org/10.1109/ACCESS.2019.2927166>.
- [7] A. Mesbah, A. Berrahou, H. Hammouchi, H. Berbia, H. Qjidaa, and M. Daoudi, "Lip reading with Hahn convolutional neural networks," *Image and Vision Computing*, vol. 88, pp. 76–83, 2019, <http://doi.org/10.1016/j.imavis.2019.04.010>.
- [8] J. S. Chung and A. Zisserman, "Learning to lip read words by watching videos," *Computer Vision and Image Understanding*, vol. 173, pp. 76–85, 2018, <http://doi.org/10.1016/j.cviu.2018.02.001>.
- [9] Z. Thabet, A. Nabih, K. Azmi, Y. Samy, G. Khoriba, and M. Elshehaly, "Lipreading using a comparative machine learning approach," in *Proceedings of the 2018 First International Workshop on Deep and*

- Representation Learning, Cairo, Egypt, 2018, pp. 19–25, <https://doi.org/10.1109/IWDRL.2018.8358210>.
- [10] S. Petridis, J. Shen, D. Cetin, and M. Pantic, "Visual-only recognition of normal, whispered and silent speech," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Calgary, AB, Canada, 2018, pp. 6219–6223, <https://doi.org/10.1109/ICASSP.2018.8461596>.
- [11] A. Koumparoulis and G. Potamianos, "Deep View2View mapping for view-invariant lipreading," in *Proceedings of the 2018 IEEE Spoken Language Technology Workshop (SLT)*, Athens, Greece, December 18–21, 2018, pp. 588–594, <https://doi.org/10.1109/SLT.2018.8639698>.
- [12] J. Wei, F. Yang, J. Zhang, R. Yu, M. Yu, and J. Wang, "Three-dimensional joint geometric-physiologic feature for lip-reading," in *Proceedings of the 2018 IEEE 30th International Conference on Tools with Artificial Intelligence*, Greece, 2018, pp. 1007–1012, <https://doi.org/10.1109/ICTAI.2018.00155>.
- [13] I. Fung and B. K. Mak, "End-to-end low-resource lip-reading with Maxout CNN and LSTM," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Calgary, AB, Canada, 2018, pp. 2511–2515, <https://doi.org/10.1109/ICASSP.2018.8462280>.
- [14] T. Thein and K. M. San, "Lip localization technique towards an automatic lip reading approach for Myanmar consonants recognition," in *Proceedings of the 2018 International Conference on Information and Computer Technologies*, IL, USA, 2018, pp. 123–127, <https://doi.org/10.1109/INFOCT.2018.8356854>.
- [15] S. Yang *et al.*, "LRW-1000: A naturally-distributed large-scale benchmark for lip reading in the wild," in *Proceedings of the 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition*, Lille, France, 2019, pp. 1–8, <https://doi.org/10.1109/FG.2019.8756582>.
- [16] J. S. Chung and A. Zisserman, "Lip reading in profile," in *Proceedings of the 28th British Machine Vision Conference*, London, UK, 2017, <https://doi.org/10.5244/C.31.155>.
- [17] P. P. Filntisis, A. Katsamanis, P. Tsiakoulis, and P. Maragos, "Video-realistic expressive audio-visual speech synthesis for the Greek language," *Speech Communication*, vol. 95, pp. 137–152, 2017, <https://doi.org/10.1016/j.specom.2017.08.011>.
- [18] T. Afouras, J. S. Chung, A. Senior, O. Vinyals, and A. Zisserman, "Deep audio-visual speech recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–11, 2018, <https://doi.org/10.1109/TPAMI.2018.2889052>.
- [19] S. Petridis, T. Stafylakis, P. Ma, F. Cai, G. Tzimiropoulos, and M. Pantic, "End-to-end audiovisual speech recognition," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2018, pp. 6548–6552, <https://doi.org/10.1109/ICASSP.2018.8461326>.
- [20] Y. Yuan, C. Tian, and X. Lu, "Auxiliary loss multimodal GRU model in audio-visual speech recognition," *IEEE Access*, vol. 6, pp. 5573–5583, 2018, <https://doi.org/10.1109/ACCESS.2018.2796118>.
- [21] S. Petridis, T. Stafylakis, P. Ma, G. Tzimiropoulos, and M. Pantic, "Audio-visual speech recognition with a Hybrid CTC/Attention architecture," in *Proceedings of the 2018 IEEE Spoken Language Technology Workshop*, Athens, Greece, 2018, pp. 513–520, <https://doi.org/10.1109/SLT.2018.8639643>.
- [22] W. J. Ma, X. Zhou, L. A. Ross, J. J. Foxe, and L. C. Parra, "Lip-reading aids word recognition most in moderate noise: A bayesian explanation using high-dimensional feature space," *PLoS ONE*, vol. 4, no. 3, 2009, Art. no. e4638, <https://doi.org/10.1371/journal.pone.0004638>.
- [23] M. Wand, J. Koutnik, and J. Schmidhuber, "Lipreading with long short-term memory," in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Shanghai, China, 2016, pp. 6115–6119, <https://doi.org/10.1109/ICASSP.2016.7472852>.
- [24] A. Gabbay, A. Shamir, and S. Peleg, "Visual speech enhancement," in *Proceedings of Interspeech 2018, 19th Annual Conference of the International Speech Communication Association*, India, Sep. 2–6, 2018, pp. 1170–1174.
- [25] M. Wand, J. Schmidhuber, and N. T. Vu, "Investigations on end-to-end audiovisual fusion," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Calgary, AB, Canada, 2018, pp. 3041–3045, <https://doi.org/10.1109/ICASSP.2018.8461900>.
- [26] D. Hu, X. Li, and X. Lu, "Temporal multimodal learning in audiovisual speech recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 2016, pp. 3574–3582, <https://doi.org/10.1109/CVPR.2016.389>.
- [27] A. Fernandez-Lopez and F. M. Sukno, "Automatic viseme vocabulary construction to enhance continuous lip-reading," in *Proceedings of the 12th International Conference on Computer Vision Theory and Applications*, Porto, Portugal, Feb. 27–Mar. 1, 2017, pp. 52–63.
- [28] K. Paleček, "Experimenting with lipreading for large vocabulary continuous speech recognition," *Journal on Multimodal User Interfaces*, vol. 12, no. 4, pp. 309–318, 2018, <https://doi.org/10.1007/s12193-018-0266-2>.
- [29] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004, <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>.
- [30] Y.-Q. Wang, "An analysis of the Viola-Jones face detection algorithm," *Image Processing On Line*, vol. 4, pp. 128–148, 2014, <https://doi.org/10.5201/ipol.2014.104>.
- [31] J. M. Saragih, S. Lucey, and J. F. Cohn, "Deformable model fitting by regularized landmark mean-shift," *International Journal of Computer Vision*, vol. 91, pp. 200–215, 2011, <https://doi.org/10.1007/s11263-010-0380-4>.
- [32] K. Janocha and W. M. Czarnecki, "On loss functions for deep neural networks in classification," *Schedae Informaticae*, vol. 25, pp. 49–59, 2016, <https://doi.org/10.4467/20838476SI.16.004.6185>.
- [33] Z. Zhang and M. R. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," in *Proceedings of the 32nd Conference on Neural Information Processing Systems*, Montréal, Canada, Dec. 2–8, 2018.
- [34] Q. Zhu, Z. He, T. Zhang, and W. Cui, "Improving classification performance of softmax loss function based on scalable batch-normalization," *Applied Sciences*, vol. 10, no. 8, pp. 29–50, 2020, <https://doi.org/10.3390/app10082950>.
- [35] N. Srivastava and R. Salakhutdinov, "Learning representations for multimodal data with deep belief nets," presented at *the 29th International Conference on Machine Learning Workshop*, Edinburgh, UK, Jun. 26–Jul. 1, 2012.
- [36] M. B. Ayed, "Balanced communication-avoiding support vector machine when detecting epilepsy based on EGG signals," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6462–6468, 2020, <https://doi.org/10.48084/etasr.3878>.
- [37] S. Nuanmeesri, "Mobile application for the purpose of marketing, product distribution and location-based logistics for elderly farmers," *Applied Computing and Informatics*, 2019, <https://doi.org/10.1016/j.aci.2019.11.001>.
- [38] A. N. Saeed, "A machine learning based approach for segmenting retinal nerve images using artificial neural networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5986–5991, 2020, <https://doi.org/10.48084/etasr.3666>.
- [39] A. U. Ruby, P. Theerthagiri, I. J. Jacob, and Y. Vamsidhar, "Binary cross entropy with deep learning technique for image classification," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 5393–5397, 2020, <https://doi.org/10.30534/ijatcse/2020/175942020>.
- [40] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proceedings of the 28th International Conference on Machine Learning*, Washington, USA, 2011, pp. 689–696.
- [41] C. Tian, and W. Ji, "Auxiliary multimodal LSTM for audio-visual speech recognition and lipreading," 2017, arXiv preprint arXiv:1701.04224v2.

AUTHORS PROFILE



Lap Poomhiran is currently a Ph.D. student in the Department of Information Technology, Faculty of Information Technology and Digital Innovation, King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand. His research interests include web and mobile programming, augmented reality (AR) and virtual reality (VR) development, image processing, data mining, machine learning, deep learning, and the internet of things (IoT).



Phayung Meesad is an associate professor at the King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand. His research interests include computational intelligence, artificial intelligence, machine learning, deep learning, data analytics, big data analysis, data science, data mining, digital signal processing, image processing, business intelligence, time series analysis, and natural language processing.



Sumitra Nuanmeesri is an assistant professor in the Department of Information Technology, Faculty of Science and Technology at Suan Sunandha Rajabhat University, Thailand. Her research interests include speech recognition, data mining, deep learning, image processing, mobile application, supply chain management systems, internet of things, robotics, augmented reality, and virtual reality.

Factors Affecting the Adoption of E-Learning Technology by Students during the COVID-19 Quarantine Period: The Application of the UTAUT Model

Houcine Chatti

Dar Al Uloom University
Riyadh, Saudi Arabia
houcine@dau.edu.sa

Slim Hadoussa

Brest Business School
Brest, France
slim.hadoussa@brest-bs.com

Abstract-The adoption of e-learning technology has become a major challenge for many academic institutions during the Covid-19 pandemic. More and more institutions are questioning the success of adopting this technology and are seeking to understand their adoption process. The objective of the current research work is to study the factors affecting the intent of Saudi university students to adopt online technologies during the Covid-19 pandemic. Based on the Information Technology Adoption Model, UTAUT, a research model was designed and validated by combining the factorial analysis method with simple linear regression analysis. The study revealed four key factors that had significant and positive effects on users' intent to use online technology, including the perceived usefulness, perceived ease of use, teacher influence, university management commitment, and availability of student technical assistance.

Keywords-Covid-19; UTAUT; e-learning

I. INTRODUCTION

The Covid-19 pandemic has disrupted education in a unprecedented scale. As of March 16, 2020, 777 million pupils and students had been forced to leave their school or university in a total of 100 countries, with 85 governments closing schools nationally and 15 others imposing school closures at local levels [1]. In Saudi Arabia, the decision to close schools and universities was taken early, on March 9, 2020 almost a week after the appearance of the first case of COVID-19 in the kingdom. The Saudi government, through its Ministry of Education (MoE), ensured that educational activities were not interrupted, adopting the concept of distant learning very effectively. In fact, the MoE provides a 5-way integrated system (ranging from electronic portals to YouTube channels) that takes into account the contrasting technical capabilities of the students and which can be adapted to their daily schedule. Universities have also adopted distant learning with their students, with each university using its respective channel/portal, which has resulted in undisrupted teaching cycle and final exam schedules. Thus, the use of new technologies has become essential to mitigate the impact of closure of academic institutions. To this end, distance learning

is offered as an alternative that guarantees pedagogical continuity.

Education and classroom teaching have changed radically with the rise of e-learning. The crisis of Covid-19 pushed educational institutions to make greater use of online technology or start using e-learning systems for course continuity. Institutions use many technologies such as teaching and learning programs or even social media. In Saudi Arabia, most universities turn to online education through various applications such as Zoom, Blackboard, LMS and Teams [2]. E-learning generally refers to the use of computer network technology and electronic media to deliver and transfer knowledge for educational purposes [3]. In addition, e-learning assists students to get access to different online courses, scientific databases, and learning tools [4]. E-learning is generally considered as an important educational source for improving the effectiveness and efficiency of learning services, because of its availability, low cost, ease of use, and interactive nature [5]. Furthermore, e-learning systems help to capitalize knowledge and coach students. E-learning systems have several functionalities that can be useful in the learning process. They allow SMS sending and support the learning activity on a laptop or mobile device. Students can easily integrate the learning content into their mobile devices since they can easily connect to mobile or Wi-Fi networks. E-learning is thus seen as the process of using electronic technologies to access learning material in a non-traditional way [3, 6]. In other words, e-learning refers to the provision, organization and management of educational activities, such as student registration, exams, assignments, course descriptions, lesson plans, messages, syllabus, core course material, etc. within a remote system. [7]. Since the success of the e-learning system depends on the willingness and acceptance of students to use it [3, 8], rejection of the use of this online technology may hinder the realization of the expected benefits [9]. Non-acceptance makes the implementation or adoption of this technology a failure and a waste of money for academic institutions [10]. The research on this topic is still in its infancy, as students' views are not well studied [11]. Exploring the

Corresponding author: Houcine Chatti

adoption of e-learning technology can lead universities to understand better the needs of their students and manage more efficiently such educational systems that are widely diffused due to Covid-19 crisis [12].

To date, little work has been attempted in analyzing the use of e-learning systems during the Covid-19 pandemic. This research aims to investigate the main factors affecting the use of e-learning systems during the Covid-19 pandemic in Saudi Arabia. The process of adopting online technology involves two major activities that are carried out sequentially: initiation and implementation [13]. During the initiation stage, the academic institution becomes aware of the e-learning system, forms an attitude towards it, and evaluates the innovation. Initiation includes the activities of researching information, conceptualizing, and planning for adoption of this technology [14]. Initiation includes, the stages of awareness, consideration, and intention. At the end of this stage, a decision on whether to adopt or reject the technology is made. Authors in [15] adhere to this two-stage vision by considering that the adoption process includes adoption and implementation. An organization adopts an innovation when it decides to allocate resources to it. The implementation stage includes the development and installation activities that take place when the organization begins to use the innovation [16]. However, the process of setting up the e-learning system begins without a real awareness of the different stakeholders in the system: students, teachers, and administration. In other words, the initiation stage is only cursory. As a result, the question of acceptance of this innovation arises, particularly for students. The UTAUT technology acceptance model is, a framework for studying the question under research.

II. THEORETICAL FRAMEWORK

A. Theories of Technology Acceptance

Behavioral theories about people's intentions have provided the theoretical basis for the development of models of technology adoption and acceptance [17]. These include the theory of reasoned action, the theory of planned behavior, and the Technological Acceptance Model (TAM). The theory of reasoned action [18] aims to predict and understand the behavior of individuals. It claims that the behavior (adoption of a technology) of any individual is reflected and directly influenced by his or her intention to achieve it, the latter being determined by two variables: attitudes and subjective norms. These two determinants of behavioral intention are in turn influenced respectively by beliefs about the consequences of the behavior and normative beliefs about a given behavior. The theory of planned behavior [19] is an extension of the theory of reasoned action. This theory adds a third determinant to the two determinants of the theory of reasoned action, namely the perceived control of the behavior or what is also called the perceived ability to carry out the behavior. It is based on the idea that individuals make rational use of available information to construct intention when faced with a behavioral decision [19]. In addition, the concept of perceived control of behavior stems from external and internal factors that facilitate the achievement of a given behavior, as well as the individual's perception of his or her personal effectiveness in achieving that behavior. In addition, two other variables explain this concept.

These are control beliefs and facilitating conditions, which refer to an individual's assessment of the importance and availability of resources or an environment conducive to the achievement of the behavior. Built from the theories of social psychology, particularly the theory of reasoned action, the TAM is specifically concerned with the acceptability behavior of ICT. The TAM offers a perspective for understanding problems related not only to the adoption and use of information systems, but also to the appropriation of information systems by users. The main objective of this model is to predict the acceptance of an information system and, above all, to evaluate the impact of various external factors on internal beliefs, attitudes, and intentions of its users. The AMT postulates that the actual use of a technology depends on the intention to use it and considers that this intention is influenced by the perceived usefulness and ease of use. The perceived usefulness is defined as the degree to which an individual believes that the use of a particular system could improve job performance [20]. It is therefore a function of the degree to which a technology is seen as advantageous and beneficial by its users. Perceived ease of use, on the other hand, refers to the degree to which an individual believes that the use of a particular system can reduce effort [20]. It refers to the degree to which an individual believes that the use of a system will not require "much" cognitive effort. The specificity of AMT in relation to ICT, the accuracy of the determinants it proposes, and the precision of the measures [21], make it the most widely used model in the areas of ICT adoption and acceptance [17].

The listed models focus on individual ICT acceptance by predicting usage intent for ICTs that are already implemented. This study focuses on a context in which the technology is already implemented and used by individuals. Thus, to answer the research question, the Unified Model Theory of Individual Technology Acceptance (UTAUT), which is considered to be the most recent and unifying theory, will be adopted in this study. UTAUT uses the essential elements of previously established models [21]. It is tested and empirically validated by longitudinal and cross-sectional studies on technology adoption in different contexts [22-26]. The UTAUT model [21] is a synthesized and comprehensive theory, which takes up the theories of acceptance of technologies in use, the variables whose validity and predictive power have been shown to be significant. Compared to previously established acceptance models, UTAUT is considered [24] the best account for technology adoption, use, and acceptance. UTAUT postulates that the actual use of a technology is a function of the intention to use it, which in turn is influenced by determinants such as the expected performance, expected effort, social influence, and enabling conditions. Moreover, unlike previous models, this model integrates new categories of moderating variables that vary the influence of the determining variables on the intention to use. These are: gender, age, experience of use, and whether the use is compulsory or voluntary.

B. Research Model Design

The main objective of this work is to determine the factors that influence the acceptance of e-learning technology in the Saudi academic community. The selection of the determinants of acceptance used in this research is heavily influenced by the

UTAUT model [21, 27]. UTAUT has the advantages of (1) being a general model belonging to the set of theoretical models that have been developed to explain human adoption behavior and (2) having a high number of constructs, which gives it great explanatory power for the intention to use a technology. Four independent variables define the model of [21], namely the perceived ease of use, the perceived utility, social influences, and facilitating conditions. Four moderating variables (age, gender, experience, and mandatory or voluntary use), moderate the influence of the explanatory variables in the intention to adopt the technology. Some authors consider that when the technology integrates with the individual's old tasks, and is therefore compatible with existing values, it can be recognized as useful and can influence the intention to use. This is why we have also considered, in this modeling effort the construct "compatibility with existing values" [28], also called task-technology compatibility [26], as an (independent) explanatory variable for the intention to adopt e-learning technology. Finally, the variables explaining the intention to use and actual use of the UTAUT were grouped into a single variable called intention to adopt e-learning technology.

For the present study we will not retain the moderator variables. Thus, we have: a (dependent) explanatory variable: intention to adopt, and the (independent) explanatory variables: perceived usefulness, ease of use, social influences and facilitating conditions (Figure 1).

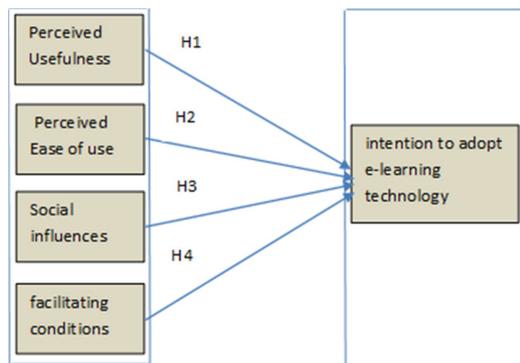


Fig. 1. The research model.

1) Perceived Usefulness

Perceived usefulness is the degree to which an individual believes that using a system would help him or her achieve gains in job performance [21]. It has been shown by several studies to be a very significant determinant in explaining the intention to adopt a technology [21, 29]. Indeed, e-learning technology can only be adopted if the students perceive gains in terms of efficiency, speed, and performance in the execution of tasks. The variable "perceived usefulness" is therefore understood as the perception of the usefulness of e-learning technology. Theoretical models of technology adoption (TAM, UTAUT) and the work of several researchers have confirmed that the perceived usefulness of a technology promotes its adoption and use. Based on the empirical research in [21, 29] we assume that:

H1: The perceived usefulness of e-learning technologies by students in performing their tasks positively influences their intention to adopt them.

2) Perceived Ease-of-Use

The perceived ease of use is seen as a direct determinant of the intent to adopt [26]. This variable is defined as the intensity with which an individual believes that the use of a particular system will occur without difficulty or additional effort [21]. Thus, the likelihood that there is an intention to use the e-learning technology is all the greater as long as the appropriation and learning process is easy to understand and use [26]. In other words, this variable explains the extent to which an individual believes that the organizational and technical infrastructure exists to support the use of the system [30]. The following hypothesis is therefore adopted:

H2: The perceived ease of use of e-learning technologies by students positively influences their intention to adopt them.

3) Social Influences

Social influences are defined as an individual's perception that most people who are important to him or her think he or she should or should not engage in the behavior in question [18]. In other words, they include the role of people who are important to the individual and who exert some influence on his/her behavior [25]. Prior to UTAUT, several technology acceptance [28, 31, 32] models showed that this "social influence" variable - also called subjective norms or social factors - exerted a significant influence on technology use. In our study, this social influence variable is measured separately by three constructs drawn from research using different models such as the GAM [31], the GAM2 [3]: influence of colleagues, influence of the hierarchical superior, and expected professional valorization. Thus, we postulate that:

H3.1: Colleagues' use of e-learning technologies influences students' adoption intentions.

H3.2: The commitment of the line manager (teacher) to the use of e-learning technologies for teaching influences the positively the students' adoption intention.

4) Facilitating Conditions

Enabling conditions refer to the extent to which an individual believes that an organizational and technical infrastructure exists to support the use of the system [21]. This variable is generally assessed through management involvement and the availability of technical assistance [33]. Organizational support or the involvement of leaders, which have been described as agents of change [34], is presented as a determining factor in the process of adopting a technology within an organization [21]. We suggest the following hypotheses:

H4.1: The involvement of university management positively influences students' intention to adopt online technologies.

H4.2: The availability of technical assistance within the university positively influences the intention of students to adopt online technologies.

III. METHODOLOGICAL FRAMEWORK

A. Study Population

The data were collected through a distributed by email questionnaire. Our study population consisted of students from a Saudi Business School. Out of the 120 contacted students, 87 (55.3% male and 44.7% female) answered the questionnaire. They were asked about the way they used to attend online classes, which indicated that 95% of the students used cell phones and 5% laptops.

B. Questionnaire Construction

The questionnaire consisted of two parts, demographic data and a 20 item questionnaire on a 5-point Likert scale ranging from 1 for strongly disagree to 5 for strongly agree. SPSS software was used to test the reliability and validity of the scales and to apply multiple regression analysis on the

collected data. For its construction, the guidelines of [25, 35-38] on the determinants of e-learning adoption and on the determinants of pedagogical integration of ICT by university teachers were followed. The items selected for each variable are summarized in Table I.

C. Methods of Analysis

For the processing of the collected data during the survey, two types of analysis were opted. At first, a principal component factor analysis was performed to test the validity and reliability of the measurement scales. Then simple linear regression was used to test the research hypotheses. This method is widely used to explain or predict a variable to be explained based on one or more explanatory variables [39]. These analyses and statistical treatments were performed using the SPSS version 25.0.

TABLE I. MEASURED ITEMS IN THE RESEARCH CONSTRUCTS

Variables		Items
Perceived usefulness	PUSE1	The use of e-learning systems helps me to perform my learning tasks in a better way.
	PUSE2	The use of online learning systems helps me to better follow my courses.
	PUSE3	Using e-learning systems can improve my level of interaction with my teachers and classmates.
	PUSE4	Using e-learning systems can improve the effectiveness of my learning.
Perceived ease-of-use	PEAS1	It would be easy for me to learn how to use the e-learning system recommended by the university.
	PEAS2	I find the e-learning system recommended by the university easy to use.
	PEAS3	I may have difficulties in using the e-learning system recommended by the university.
	PEAS4	It would be easy for me to become competent in the use of the learning system.
Classmate influence	CINF1	My classmates help me use the e-learning system.
	CINF2	I interact with my classmates on the use of e-learning systems.
Teacher influence	TINF1	My teacher explicitly supports my use of e-learning systems.
	TINF2	My professor is convinced of the advantages of the e-learning systems.
Top management	TMAN1	The university management is particularly interested in the use of the e-learning system.
	TMAN2	The use of the e-learning system is strongly encouraged by the university management.
	TMAN3	The management of the university is actively involved in the implementation of the e-learning system.
Technical assistance	TASS1	The members of the IT department are always available to help those who need to
	TASS2	The university has an IT department with technicians ready to help me in case of need.
	TASS3	The necessary instructions for a better use of the e-learning system are available for me.
Intention to adopt	BINT1	I have decided to adopt the e-learning system recommended by the university.
	BINT2	I decided to attend the online course
	BINT3	During this pandemic, I think it is good to use the e-learning system in courses.

IV. RESULT PRESENTATION AND DISCUSSION

A. Purification of Measurement Scales

There is a question of verifying the unidimensionality of the constructs and the internal coherence of the different scales of measurement. To do so, principal components factorial analysis was performed and the coefficient of Cronbach's alpha was calculated [40]. The analysis of the results of the one-dimensionality test of the measurement scales shows that the vast majority of our constructions present a Cronbach's alpha greater than 0.8, which is widely acceptable in the literature. These values show and attest to the existence of internal consistency between the scale items. The constructs are therefore reliable for testing the research hypotheses.

B. Testing the Validity of the Model Hypotheses

The obtained results, presented in Table II, show that the perceived usefulness and the perceived ease of use positively influence the intent to adopt. Thus the hypotheses H1 ($\beta=0.93$,

$p<0.01$), H2 ($\beta=0.95$, $p<0.01$) are verified. In addition, the hypothesis H3.1 ($\beta=-0.06$, $p>0.1$) is rejected while H3.2 ($\beta=0.32$, $p<0.01$) is accepted. Finally, the two constructs of the facilitation condition variable, management involvement and availability of technical assistance, both positively influence the intention to adopt the technology. The hypotheses H4.1 ($\beta=0.33$, $p<0.01$) and H4.2 ($\beta=0.59$, $p<0.01$) are therefore verified.

C. Discussion

1) The impact of Perceived Usefulness on the Intent to Adopt Online Learning Technology

The results of the analysis show that perceived usefulness has a significant influence on adoption intention. These results are consistent with those of [20, 41]. The significance of this perceived utility variable in our study shows that for students, the online learning system supports the teacher-student relationship and promotes the learning activity. Students appear to believe that a learning system improves the effectiveness of

their learning, allowing them to visualize and follow their course and to interact more easily with teachers. Usefulness is perceived by the fact that an e-learning system would be a real platform for students to continue to attend their course and complete their assignments without any constraints. The

students consider these platforms as real tools for communication, content publication (knowledge sharing), technology monitoring, etc. The respondents are therefore aware of the contributions of e-learning systems to learning activities.

TABLE II. VARIABLE PURIFICATION IN THE RESEARCH MODEL

Construits	Items	Quality of representation	Eigenvalues	Explained variance	Cumulative variance	Cronbach's alpha
Perceived usefulness	PUSE 1	0.788	2.076	62.28%	62.28%	0.872
	PUSE 2	0.789				
	PUSE 3	0.921				
	PUSE 4	0.834				
Perceived ease-of-use	PEAS 1	0.701	3.150	65.11%	65.11%	0.801
	PEAS 2	0.698				
	PEAS 3	0.744				
	PEAS 4	0.786				
Classmate influence	CINF 1	0.651	2.795	55.01%	55.01%	0.891
	CINF 2	0.569				
Teacher influence	TINF 1	0.734	1.918	71.11%	71.11%	0.863
	TINF 2	0.781				
Top management	TMAN 1	0.901	2.823	74.16%	74.16%	0.857
	TMAN 2	0.878				
	TMAN 3	0.865				
Technical assistance	TASS 1	0.689	2.350	59.19%	59.19%	0.901
	TASS 2	0.707				
	TASS 3	0.819				
Intention to adopt	BINT 1	0.780	2.090	70.35%	70.35%	0.874
	BINT 2	0.865				
	BINT 3	0.897				

TABLE III. TESTING THE VALIDITY OF THE HYPOTHESES

Construits	Items	Hypotheses	R2	β	F
Perceived usefulness, intention to adopt	PUSE, INTA	H1	0.91	0.93,***	4.125,***
Perceived ease-of-use, intention to adopt	PEAS, BINT	H2	0.93	0.95,***	0.97,***
Classmate influences, intention to adopt	CINF, BINT	H3.1	0.79	-0.06	-1.00
Teacher influence, intention to adopt	TINF, BINT	H3.2	0.81	0.32,***	7.61,***
Top management, intention to adopt	TMAN, BINT	H4.1	0.83	0.33,***	0.59,***
Technical assistance, intention to adopt	TASS, BINT	H4.2	0.73	0.59,***	7.28,***

*p<0.10, **p<0.05, ***p<0.01

2) The Impact of Perceived Ease of Use on the Intention to Adopt an E-Learning System

The results showed that there was a significant and direct relationship between these two variables. The influence of the perceived ease of use on the intent to adopt e-learning systems was found to be very significant and positive in this research. These findings are consistent with the results in [20, 21, 41], where authors assume that the more the student sees the use of e-learning systems in the learning activity as effortless, the greater the intention to adopt will be. Perceived ease of use, a logical consequence of personal experience in using such a technology, which itself has created a kind of attachment to e-learning systems, favors the willingness to use it. Thus, the results obtained can be explained by:

- The fact that, for respondents, this is not a new and unknown technology, but rather a technology that they are used to practice in a personal and sometimes playful way (social media). Almost 80% of the respondents used social media to communicate with each other. The majority of them were fairly accustomed to these platforms.

- The availability of training material on the use of online learning systems and web 2.0 platforms at the university and on the internet for free.

3) The Impact of Social Influences on the Intention to Adopt the E-Learning System

In most studies of technology adoption in workplaces [21, 25], the influence of classmates and of the teacher has been considered the as a "social influence". These studies have shown that each of these constructs has a significant influence on technology adoption. Our results showed that only one of the starting constructs of the social influence variable exerts an influence on the adoption of e-learning systems. Only the teacher influence construct that emerged during the treatment was found to significantly influence the intention to adopt e-learning systems. These results call for several comments:

- We note the absence of a significant impact of peer influence on the intention to adopt e-learning systems. Based on the results, we conclude that in the context of our survey, the students to whom our respondents belonged did not have "early" adopters [41] who could encourage them

to use e-learning systems for their learning practices. We could also assume that in a period of confinement, contact is virtually absent and the use of these platforms does not require great skill or ability to motivate students to seek help from their colleagues.

- Our study showed the presence of a causal link between the variable influence of teachers and the intention to adopt e-learning systems. This suggests that according to the surveyed students, the adoption of e-learning systems is a function of teacher encouragement. Teachers encourage their students to adopt e-learning systems because they are convinced of the benefits that the learning activity can derive from their implementation. Teachers help students to solve many technical problems and answer many of their questions regarding the use of these platforms. Teachers post all course material on these platforms and even conduct online exams. They do not give students a choice about whether or not to use these platforms.

4) *The Impact of Facilitation Conditions on the Intention to Adopt the E-Learning System*

Both constructs of the facilitation condition variable of management involvement and the availability of technical assistance were found to significantly influence the intention of adoption of e-learning systems by the sampled students. These results are consistent with several previous studies [29, 33]. The involvement of university management is necessary to improve student commitment and motivation to adopt such systems. Senior management commitment refers to a concept well identified, through work on information systems, as a determinant of successful IT acceptance [42]. Under the impetus of a policy of presence and support from management, students are obliged to align themselves with management's guidelines for distance education and to become more involved in the use of these platforms. We can therefore consider that the respondents are aware that despite the significant benefits of e-learning systems, management support is necessary for the successful adoption of these platforms. The meaning of the availability of technical assistance is simply explained by the fact that despite a certain willingness on the part of the students, if there is no technical infrastructure, or if it is not adequate and there is no high speed internet connection, there can be no claim of adopting e-learning systems. Thus, the presence of computer equipment, internet connection, and technical assistance, as well as the establishment of a training policy, if necessary, to accompany the adoption of e-learning systems is essential for the respondents.

V. CONCLUSION

Within COVID-19 context, online education and e-learning has emerged at many universities as a modern educational model and changed radically the previous traditional learning concept. To avoid the spread of the virus and enhance the social distancing among students, many Saudi universities and high schools implement and adopt different online educational systems to ensure course continuity. This research focuses on identifying the factors that determine e-learning system adoption by students at a Saudi Business School. The study adopts the UTAUT theory [21] to develop a conceptual model

which was tested through a quantitative methodology by an online survey (87 students). The most important outcome of this research was the confirmation of the importance of factors determining the acceptance and the adoption of an e-learning system. The results show that the main determining factors are: the perceived usefulness, the perceived ease of use, teacher influence, involvement of university management, and the availability of technical assistance. This research demonstrates that if students perceive the usefulness of e-learning systems with their learning tasks, they will be more inclined to adopt such a system. This result can be useful to higher education managers that focus on enhancing awareness-raising actions regarding the benefits and opportunities that e-learning systems can provide. Such policies would encourage the e-learning system adoption by students and would enhance the success of the online educational transition. This awareness can be done, for example, through offers of continuing education and/or short awareness meetings.

Moreover, the results suggest that early awareness rising among students by the School management is vital for the adoption of e-learning systems. The use of e-learning systems is relatively new at many higher educational institutions. Therefore, like any technological innovation, it faces difficulties in its launching. The e-learning system makes learners struggle with their habits, which sometimes leads to a reaction of rejection or at best, mistrust. This is apparent both in the appreciation of the usefulness of the system and its ease of use. This feeling could be also increased by the anxiety expressed by many students due to the general negative atmosphere that exists due to the lockdown. The role of management is essential in this case to facilitate the development of a positive attitude towards e-learning systems. Furthermore, the role of the academic staff is crucial to enhance the adoption and the creation of an attractive atmosphere that encourages students to not only virtually assist but also to interact during the online courses' sessions. The availability of technical assistance was also found to be a very significant influence on the intent to adopt e-learning systems. Since a technology can only be used if the technological infrastructure is available, it is essential to strengthen the IT infrastructure in general and the broadband internet connection. Many students have the importance of being technically assisted, especially when they encounter a technical problem during the course's online session.

The insights obtained by this study could be helpful for both academics and practitioners interested on e-learning adoption at the educational sector. Thus, this research helps educational managers to establish effective policies towards e-learning system use that may ensure the education continuity during the Covid-19 pandemic and would enhance the knowledge transfer quality [3, 43]. However, this research has some limitations. Although the results are valid, the tested sample is limited to only one Saudi Business School. It would be interesting in a future study to consider other higher institutions specialized in different fields with a larger number of students. Also, the sampling method could be enriched by taking consideration of the field' differences that exist between students coming from different backgrounds. A future study could also compare the performances of the public and the

private educational sector. Moreover, the current study did not consider the emotional situation, stress, and anxiety caused by the Covid-19 atmosphere that was expressed by many students during the survey and that could possibly impact the adoption of the e-learning system. In addition, the current study is limited to the students' point of view that could be different from the academic staffs. These different additional studies and data could be very valuable for exploring and understanding some causal relationships in the conceptual model and to enhance the results interpretation.

This study attempts to offer a positive and optimistic perspective regarding the online educational student experience in the context of Saudi higher education system. The digital transformation of the higher education sector will continue to change educational practices and learning methods. Future studies could be conducted in this sense and explore the impact of e-learning systems on knowledge sharing and capitalization.

ACKNOWLEDGEMENT

The authors wish to thank the Deanship of Graduate Studies and Scientific Research at Dar Al Uloom University for the support and funding of this study.

REFERENCES

- [1] "COVID-19 Educational Disruption and Response," *UNESCO*, Mar. 24, 2020. <https://en.unesco.org/news/covid-19-educational-disruption-and-response> (accessed Mar. 17, 2021). 2 Kingdom of Saudi Arabia - Ministry of Education. <https://www.moe.gov.sa/en/Pages/default.aspx> (accessed Mar. 17, 2021).
- [2] Kingdom of Saudi Arabia - Ministry of Education. <https://www.moe.gov.sa/en/Pages/default.aspx> (accessed Mar. 17, 2021).
- [3] S. Hadoussa, "Evaluation of e-learning system on higher education institutions in KSA: a survey at Saudi Electronic University," *International Journal of Technology Enhanced Learning*, vol. 12, no. 2, pp. 180–199, Jan. 2020, <https://doi.org/10.1504/IJTEL.2020.106285>.
- [4] H. Slim and M. Hafedh, "Social media impact on language learning for specific purpose: a study in english for business administration," *Teaching English with Technology*, vol. 19, no. 1, pp. 56–71, 2019.
- [5] F. A. A. Idris and Y. B. Osman, "Challenges Facing the Implementation of e-Learning at University of Gezira According to View of Staff Members," in *Fifth International Conference on e-Learning*, Manama, Bahrain, Oct. 2015, pp. 336–348, <https://doi.org/10.1109/ECONF.2015.51>.
- [6] H. E. Fazazi, M. Elgarej, M. Qbadou, and K. Mansouri, "Design of an Adaptive e-Learning System based on Multi-Agent Approach and Reinforcement Learning," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6637–6644, Feb. 2021, <https://doi.org/10.48084/etasr.3905>.
- [7] M. Haghshenas, "A Model for Utilizing Social Softwares in Learning Management System of E-Learning," *Quarterly of Iranian Distance Education Journal*, vol. 1, no. 4, pp. 25–38, Apr. 2019, <https://doi.org/10.30473/idej.2019.6124>.
- [8] M. A. Almaiah and O. A. Alismaiel, "Examination of factors influencing the use of mobile learning system: An empirical study," *Education and Information Technologies*, vol. 24, no. 1, pp. 885–909, Jan. 2019, <https://doi.org/10.1007/s10639-018-9810-7>.
- [9] M. A. Almaiah and I. Y. Alyoussef, "Analysis of the Effect of Course Design, Course Content Support, Course Assessment and Instructor Characteristics on the Actual Use of E-Learning System," *IEEE Access*, vol. 7, pp. 171907–171922, 2019, <https://doi.org/10.1109/ACCESS.2019.2956349>.
- [10] Q. N. Naveed, M. R. N. Qureshi, A. O. Alsayed, A. Muhammad, S. Sanober, and A. Shah, "Prioritizing barriers of E-Learning for effective teaching-learning using fuzzy analytic hierarchy process (FAHP)," in *4th IEEE International Conference on Engineering Technologies and Applied Sciences*, Salmabad, Bahrain, Dec. 2017, pp. 1–8, <https://doi.org/10.1109/ICETAS.2017.8277855>.
- [11] A. Tarhini, R. Masa'deh, K. A. Al-Busaidi, A. B. Mohammed, and M. Maqableh, "Factors influencing students' adoption of e-learning: a structural equation modeling approach," *Journal of International Education in Business*, vol. 10, no. 2, pp. 164–182, Jan. 2017, <https://doi.org/10.1108/JIEB-09-2016-0032>.
- [12] M. Alksasbeh, M. Abuhelaleh, M. Almaiah, M. AL-jaafreh, and A. A. Karaka, "Towards a Model of Quality Features for Mobile Social Networks Apps in Learning Environments: An Extended Information System Success Model," *International Journal of Interactive Mobile Technologies*, vol. 13, no. 5, pp. 75–93, May 2019.
- [13] S. Gopalakrishnan and F. Damanpour, "A review of innovation research in economics, sociology and technology management," *Omega*, vol. 25, no. 1, pp. 15–28, Feb. 1997, [https://doi.org/10.1016/S0305-0483\(96\)00043-6](https://doi.org/10.1016/S0305-0483(96)00043-6).
- [14] E. M. Rogers, *Diffusion of Innovations, Fourth Edition*, 4th edition. New York, NY, USA: Free Press, 1995.
- [15] V. Grover and M. Goslar, "Toward an empirical taxonomy and model of evolution for telecommunications technologies," *Journal of Information Technology*, vol. 8, no. 3, pp. 167–176, Sep. 1993, <https://doi.org/10.1057/jit.1993.23>.
- [16] M. B. Prescott and C. V. Slyke, "The Internet as an Innovation," in *Americas Conference on Information Systems*, Aug. 1996.
- [17] V. Venkatesh and F. D. Davis, "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science*, vol. 46, no. 2, pp. 186–204, Feb. 2000, <https://doi.org/10.1287/mnsc.46.2.186.11926>.
- [18] M. Fishbein and I. Ajzen, *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Boston, MA, USA: Addison-Wesley, 1975.
- [19] I. Ajzen and M. Fishbein, *Understanding Attitudes and Predicting Social Behavior*, 1st edition. Englewood Cliffs, NJ, USA: Pearson, 1980.
- [20] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, Sep. 1989, <https://doi.org/10.2307/249008>.
- [21] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quarterly: Management Information Systems*, vol. 27, no. 3, pp. 425–478, Sep. 2003, <https://doi.org/10.2307/30036540>.
- [22] J. E. Anderson and P. H. Schwager, "SME adoption of wireless LAN technology: Applying the UTAUT model," in *7th Annual Conference of the Southern Association for Information Systems*, 2004, pp. 39–43.
- [23] C. A. Lin, "Webcasting Adoption: Technology Fluidity, User Innovativeness, and Media Substitution," *Journal of Broadcasting & Electronic Media*, vol. 48, no. 3, pp. 446–465, 2004.
- [24] P. Rosen and D. Klumper, "The Impact of the Big Five Personality Traits on the Acceptance of Social Networking Website," in *AMCIS 2008 Proceedings*, Jan. 2008.
- [25] T. Lassoued, "Les déterminants de l'adoption de l'e-learning : etude empirique au sein de l'entreprise tunisienne," Ph.D. dissertation, Jean Moulin University, Lyon, France, 2010.
- [26] K. S. Kouakou, "Les déterminants de l'adoption des réseaux sociaux numériques en situation professionnelle : étude empirique au sein des bibliothèques des universités ivoiriennes," *Frantice.net*, no. 9, Dec. 2014.
- [27] R. A. Ali and M. R. M. Arshad, "Perspectives of Students' Behavior Towards Mobile Learning (M-learning) in Egypt: an Extension of the UTAUT Model," *Engineering, Technology & Applied Science Research*, vol. 6, no. 4, pp. 1109–1114, Aug. 2016, <https://doi.org/10.48084/etasr.710>.
- [28] G. C. Moore and I. Benbasat, "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation," *Information Systems Research*, vol. 2, no. 3, pp. 192–222, Sep. 1991, <https://doi.org/10.1287/isre.2.3.192>.

- [29] G. Ibanescu, *Facteurs d'acceptation et d'utilisation des technologies d'information: une étude empirique sur l'usage du logiciel « rational suite » par les employés d'une grande compagnie de services informatiques*. Montreal, Quebec, Canada: Université du Québec à Montreal, 2011.
- [30] M. Silic and A. Back, "Organizational Culture Impact on Acceptance and Use of Unified Communications & Collaboration Technology in Organizations," in *26th Bled eConference eInnovations: Challenges and Impacts for Individuals, Organizations and Society*, Bled, Slovenia, Jun. 2013, pp. 275–286.
- [31] V. Venkatesh and F. D. Davis, "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science*, vol. 46, no. 2, pp. 186–204, Feb. 2000, <https://doi.org/10.1287/mnsc.46.2.186.11926>.
- [32] R. Frambach, N. Schillewaert, M. Ahearne, and R. Moenaert, "The Acceptance of Information Technology in the Sales Force," e-Business Research Center, University Park, PA, USA, e-Business Research Center Working Paper 7–2000, 2000.
- [33] A. Singh and M. M. Shoura, "A life cycle evaluation of change in an engineering organization: A case study," *International Journal of Project Management*, vol. 24, no. 4, pp. 337–348, May 2006, <https://doi.org/10.1016/j.ijproman.2005.11.001>.
- [34] M. T. Dishaw and D. M. Strong, "Extending the technology acceptance model with task–technology fit constructs," *Information & Management*, vol. 36, no. 1, pp. 9–21, Jul. 1999, [https://doi.org/10.1016/S0378-7206\(98\)00101-3](https://doi.org/10.1016/S0378-7206(98)00101-3).
- [35] V. Venkatesh and X. Zhang, "Unified Theory of Acceptance and Use of Technology: U.S. Vs. China," *Journal of Global Information Technology Management*, vol. 13, no. 1, pp. 5–27, Jan. 2010, <https://doi.org/10.1080/1097198X.2010.10856507>.
- [36] H. Khechine, S. Lakhal, D. Pascot, and A. Bytha, "UTAUT Model for Blended Learning: The Role of Gender and Age in the Intention to Use Webinars," *Interdisciplinary Journal of E-Learning and Learning Objects*, vol. 10, pp. 33–52, 2014.
- [37] C. -M. Chao, "Factors Determining the Behavioral Intention to Use Mobile Learning: An Application and Extension of the UTAUT Model," *Frontiers in Psychology*, vol. 10, 2019, Art. no. 1652, <https://doi.org/10.3389/fpsyg.2019.01652>.
- [38] V. Acharya, S. O. Junare, and D. D. Gadhavi, "E-Payment: Buzz Word or Reality," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3S2, pp. 397–404, 2019.
- [39] J. Neter, W. Wasserman, and M. H. Kutner, *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*, 3rd edition. Homewood, IL, USA: CRC Press, 1990.
- [40] Y. Evrard, E. Roux, and B. Pras, *Market: Etudes et recherches en marketing*, 3rd Edition. Paris, France: Dunod, 2003.
- [41] M. Y. Yi, J. D. Jackson, J. S. Park, and J. C. Probst, "Understanding information technology acceptance by individual professionals: Toward an integrative view," *Information & Management*, vol. 43, no. 3, pp. 350–363, Apr. 2006, <https://doi.org/10.1016/j.im.2005.08.006>.
- [42] E. W. T. Ngai, C. C. H. Law, and F. K. T. Wat, "Examining the critical success factors in the adoption of enterprise resource planning," *Computers in Industry*, vol. 59, no. 6, pp. 548–564, Aug. 2008, <https://doi.org/10.1016/j.compind.2007.12.001>.
- [43] H. Louati and S. Hadoussa, "Study of Social media impacts on social capital and employee performance – evidence from Tunisia Telecom," *Journal of Decision Systems*, Feb. 2021, <https://doi.org/10.1080/12460125.2021.1872142>.

Analysis of Text Feature Extractors using Deep Learning on Fake News

Bilal Ahmed

Department of Electrical Engineering
Sukkur IBA University
Sukkur, Pakistan
bilal_ahmed@iba-suk.edu.pk

Gulsher Ali

Department of Electrical Engineering
Sukkur IBA University
Sukkur, Pakistan
gulsher@iba-suk.edu.pk

Arif Hussain

Department of Electrical Engineering
Sukkur IBA University
Sukkur, Pakistan
arif.hussain@iba-suk.edu.pk

Abdul Baseer Buriro

Department of Electrical Engineering
Sukkur IBA University
Sukkur, Pakistan
abdul.baseer@iba-suk.edu.pk

Junaid Ahmed

Department of Electrical Engineering
Sukkur IBA University
Sukkur, Pakistan
j.bhatti@iba-suk.edu.pk

Abstract-Social media and easy internet access have allowed the instant sharing of news, ideas, and information on a global scale. However, rapid spread and instant access to information/news can also enable rumors or fake news to spread very easily and rapidly. In order to monitor and minimize the spread of fake news in the digital community, fake news detection using Natural Language Processing (NLP) has attracted significant attention. In NLP, different text feature extractors and word embeddings are used to process the text data. The aim of this paper is to analyze the performance of a fake news detection model based on neural networks using 3 feature extractors: TD-IDF vectorizer, Glove embeddings, and BERT embeddings. For the evaluation, multiple metrics, namely accuracy, precision, F1, recall, AUC ROC, and AUC PR were computed for each feature extractor. All the transformation techniques were fed to the deep learning model. It was found that BERT embeddings for text transformation delivered the best performance. TD-IDF has been performed far better than Glove and competed the BERT as well at some stages.

Keywords-fake news; natural language processing; feature extractors; deep learning

I. INTRODUCTION

Due to the easy and excessive use of the internet and the incremented use of social media, the probability of fake news circulation is increased. This has impacted the trust on news from the media nowadays specially since the 2016 US elections. Authors in [1] conducted a survey in which they found that global trust on the news from the media of different countries ranged from 23% to 62%. The big challenge for the

researchers is to encounter this problem and provide a feasible solution. The main aspects of this problem are the fact that the same sources may provide both fake and real news, the language used can be deceiving, and biasness of the dataset and machine learning models may occur. Also, due to advancement of Artificial Intelligence applications in natural language generation has brought considerable negative impact when used in generating fake news.

In [2], news sources were under consideration instead of single articles. The idea is based on the frequency of the news sources providing fake news. The more the frequency of fake news from a source, the more the chances are that that source will provide more fake news in the future. The motive of the authors was to build an algorithm that can identify the fake news in its source before it spread. In their paper they targeted multiple sources such as URLs, twitter accounts, Wikipedia pages and articles, and found articles to be more real than the other kind of sources. Automatic text generation using Artificial Intelligence is also a popular way to spread fake news at a fast pace [3].

Fake news is one of the most difficult and sensitive topics in the field of NLP. When dealing with fake news one must keep tricky things in mind like the source of the news, the language of the news, and its pattern. Different types of text transformation techniques, machine learning algorithms and state-of-the-art methods have been introduced and applied to address this problem. TD-IDF is one of the simplest text transformation techniques which transforms each word to a

Corresponding author: Bilal Ahmed

float number as its weight according to the frequency of words in documents. This simple technique is very useful when working on a simple task. But when encountering a tricky topic like fake news detection, TD-IDF can be less effective. To encounter the complexity of text-based data in machine learning, text transformation techniques, also called as embeddings, are developed which deal with the phenomenon of the relation of words with respect to their meanings. However, TD-IDF was used before the contextual embeddings came out and became common. Authors in [4] used TD-IDF vectorizer to extract features from news articles. Authors in [5] proposed a tool for fake news detection in which they used bag of words, bigram frequency, and TD-IDF vectorizer to extract features from news articles which were tested with probabilistic classification and linear classification. Authors in [6] analyzed different machine learning models including Naïve Bayes, Support Vector Machine (SVM), Logistic Regression and Recurrent Neural Networks (RNN) on a fake news dataset from Twitter. They concluded that the Naïve Bayes and SVM are the classifiers with the best performance. Ensemble or combination of machine learning models is a technique used by researchers to deal with complex machine learning tasks. Authors in [7] analyzed different ensembles of different machine learning models and finally came with the an ensemble of Decision Tree, Logistic Regression, Bagging Classifier used with hard-voting ensemble technique which gave accuracy of 88%. For other NLP tasks, ensemble techniques give nice results. Authors in [8] used the ensemble technique to improve the translation quality from English to Hindi and used 6 different machine translation engines.

For NLP problems, neural networks provide a great range of algorithms to process and learn on sequence-based or textual data. Dense, RNN, LSTM, 1D-Convolutional (Conv1d) and GRU layers are being used in processing of textual or sequence-based data. Authors in [9] used convolutional neural networks to build a model that can classify the Arabic text. Authors in [10] used convolutional layers and bidirectional RNNs on large movie review and sentiment treebank datasets. Embeddings do not consist a feasible solution for every text-based problem, as there are two kinds of texts, static and context-based and so there should be two kinds of embeddings. Thus, the contextual embedding has been developed which works well on the text data having context with respect to every word. This paper discusses and analyzes different word transformation algorithms on fake news datasets using deep learning for each word transformation technique. It covers some well-known word transformation techniques such as BERT [11], Glove [12] and TD-IDF [13]. The main objective of this paper is to analyze whether contextual embeddings outperform the static word transformation techniques or not.

II. THE DATASETS

Two balanced datasets openly available on Kaggle were used. Both datasets contain both true and false news about western politics and international issues. The first dataset [14] contains 6335 articles about different topics mostly on politics and has 3164 fake and 3171 true articles and the second [15] contains 20,800 articles of the same nature with 10,387 fake articles and 10,413 valid articles (Figure 1). In both datasets,

article text was only analyzed and the other columns were dropped to avoid probability of false pattern learning and to reduce computational cost while training the model. For both datasets, train/test ratio was 80/20 and from the train data, 10% was kept for validation to monitor the model behavior during training.

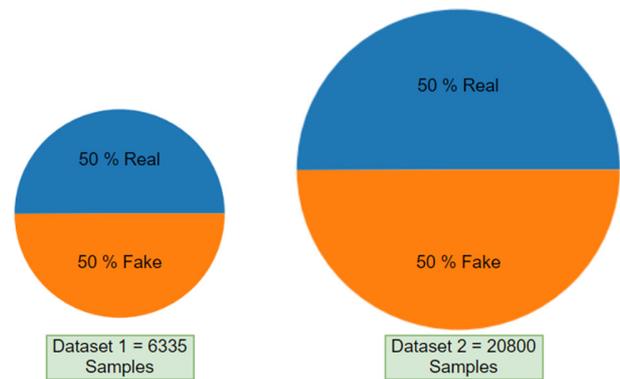


Fig. 1. Pictorial presentation of the used datasets

III. METHODOLOGY AND FEATURE EXTRACTORS

The objective of this study is to analyze the performance of 3 feature extractors, BERT embeddings, Glove embeddings and TD-IDF vectorizer using ANNs on two fake news datasets. The ANN model contains two dense hidden layers and an output layer with 8, 16 and 1 neuron(s) respectively. The feature extractors were chosen because they cover all broader classes of text feature extractors. The old fashioned TD-IDF vectorizer computes the word count (frequency) of a word, Glove is a static embedding context-independent method which works on the principle of computing the similarity between words according to their semantics, and BERT is a contextual embedding model which does not just compute the similarity between words but takes care of the context in which a specific word is used, since the word's meaning may be varied according to the context. ANNs were used as classifiers due to their ability to handle efficiently large datasets in comparison with other machine learning models. Other ANN types (LSTM, RNN, etc.) were not used, since, in this paper, the analysis of feature extractor has not been conducted on the basis of different ANN models. Also, a simple ANN requires less computational time. The BERT model was itself quite heavy. For extracting features of our dataset from Glove and TD-IDF, a personal I7 laptop with 8GB RAM was used. Feature extraction and model training took hardly 2-3 minutes for TD-IDF and 4-5 minutes for Glove. However, for BERT the required computational power surpassed the available resources, as feature extraction requires very high computational power. For BERT, the cloud-based Kaggle platform was utilized, which offers free GPU and TPU usage for higher computational tasks. For this task 15 GB GPU and 16 GB RAM were used to extract the features and pass to the ANN classification model which took about 28 minutes in total for each dataset.

This study was carried out using Python language and its available tools. The datasets were loaded from Kaggle, then

they were split into train and test groups with 80/20 ratio. The feature extractors were applied on each dataset and the output was fed to the ANN model. We used Sklearn library to implement the TD-IDF vectorizer and Keras for the implementation of the ANN model having TensorFlow at the backend. We used built-in functions in Sklearn for the evaluation of the models trained on each feature extractor for both datasets. Accuracy, precision, recall, F1, AUC ROC and AUC PR score were computed for both datasets using all three feature extractors to compare their performances. The flowchart of the process from dataset loading to the final evaluation is given in Figure 2.

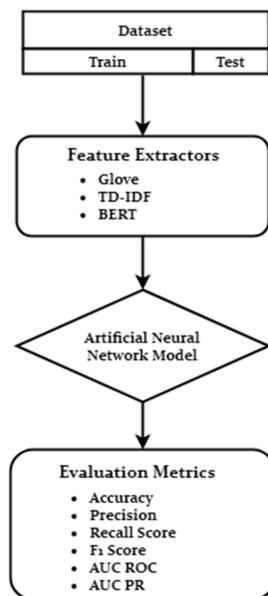


Fig. 2. Flowchart of ANN model analysis for 3 feature extractors.

A. TD-IDF Vectorizer

TD-IDF vectorizer is one of the simplest techniques to transform text into numerical values which can be fed to a machine learning model for processing. It statistically computes and finds the relevancy of a word from a document in other documents. It is computed by multiplication of two metrics: how often a word appears in a document and how rare it appears in other documents. TD-IDF vectorizer has nothing to do with the similarity between words because it is not an embedding. It is very commonly used as feature extractor for various NLP tasks [16, 17].

B. Glove Embeddings

Glove embedding is an unsupervised model for word representation in the form of vectors. These embeddings are achieved by mapping words to a meaningful space in which the distance between words is related to their semantic meaning. Cosine similarity and Euclidean distance are used in Glove embeddings to compute the distance between words. Glove comes with the advantage that it does not just depend on local context (surrounding) information of words but on the global co-occurrence of words in a given corpus by creating a co-occurrence matrix of words in a given corpus unlike Word2Vec

which relies on local contextual (surrounding) word information. Glove embeddings has been used in many text problems [18, 19]. Embedding comes in some versions with respect to the size of tokens used. We have used a pre-trained Glove model with 6 billion tokens, each of 300-dimensional vector size.

C. BERT Embeddings

Contextual embeddings [20, 21] differ from static embeddings like Word2Vec [22] and Glove. These embeddings do not just compute the similarity between words which similarity in their semantics, but also they compare the context as well in which the words are used. They are more efficient on contextual problems like sentiment analysis, sentence classification, text summarization, etc.. An interesting thing about BERT is that it does not just compute words or token embeddings, but also sentence embeddings to differentiate the sentences and positional embedding of the word in a given sequence. These combined embeddings can clearly help context each word in a given corpus. Another interesting thing in the development of the BERT model is the use of the concept of masked language modeling. This means that they hid 15% of words and used their positional embeddings to address or infer them to make the learning more effective. As a result, the BERT model outperformed all state-of-the-art existing language models even before its convergence. Authors in [23] used a BERT model with Bayesian Network to classify text data of people's livelihood governance. The BERT model comes in various versions with respect to the number of layers, heads, hidden units, cased, uncased and for languages other than English. We used the BERT Base Uncased model with 12 layers with 768 hidden units, 12 attention heads which has 110 million learning parameters.

IV. CLASSIFICATION MODEL AND EXPERIMENTS

Features from the split data into train and test sets were fed to a simple classifier 2-layer feed forward ANN model which evaluated each feature extractor (TD-IDF, Glove and BERT). The same ANN model was used for all 3 feature extractors' outputs to balance and rationalize the experimental results.

A. The Artificial Neural Network

A simple ANN was selected to classify the news in two classes (fake and real) after getting the features from all 3 feature extractors. The ANN contains two hidden dense layers having 8 and 16 neurons respectively with a final output layer with a single neuron as this is a binary classification problem. Relu activation function was used in hidden layers and sigmoid in the output layer with Adam optimizer to update the leaning weights while training.

B. Experimental Results

In this section, the results obtained from the 3 feature extractors are analyzed and compared. Accuracy is commonly used as an evaluation metric to analyze performance [24]. Figures 3, 4 illustrate the obtained results including accuracy, precision, and AUC ROC for both datasets. Outperforming the TD-IDF and Glove, BERT achieved 96% accuracy on the first dataset and 99% accuracy on the second dataset. TD-IDF achieved 93% on the first dataset and 96% on the second.

In the first dataset, BERT outperformed Glove and TD-IDF by 12% and 3% in accuracy. The precisions of BERT and TD-IDF do not differ much but achieved better by 13% than Glove. The AUC ROC scores of TD-IDF and BERT are similar and better than Glove. In the second dataset, Glove performed better. This time BERT outperformed TD-IDF by 3% as was in first dataset and Glove by 7% in accuracy. In precision, BERT outperformed Glove and TD-IDF by 7% and 3% respectively, and on the AUC ROC score, both BERT and TD-IDF achieved 3% higher than Glove. The reason because Glove performed a bit better in the second dataset (but not better than BERT and TD-IDF), is that the second dataset is 3 times the size of the first dataset as we implemented an ANN which requires large datasets to perform better. All feature extractors' performances are better on the second dataset. The detailed comparisons of all discussed evaluation metrics are given in Tables II and II.

In both datasets, BERT performed better than the other feature extractors. TD-IDF also performed well, but Glove embeddings did not because the problem at hand is a contextual one and Glove computes static embeddings. Surprisingly, TD-IDF performed far better than Glove and reached BERT's performance at some stage in the first dataset. There are 3 explanations for this. The first reason for this surprise performance of TD-IDF is that the word length of the document for Glove and BERT was fixed. TD-IDF uses all words in given documents and normally we do not fix the size of TD-IDF word length in its implementation, but for embeddings a size must be defined. The second reason is that Glove and BERT might have over fit the data as their vocabulary size is too large. The third reason may be that signal from embeddings is noisy as they have complex architecture in their implementation which can cause the model to learn false information from the given data during training.

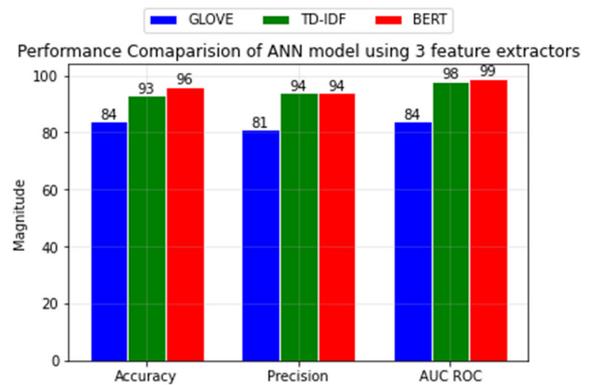


Fig. 3. Analysis of feature extractors' performance on the first dataset.

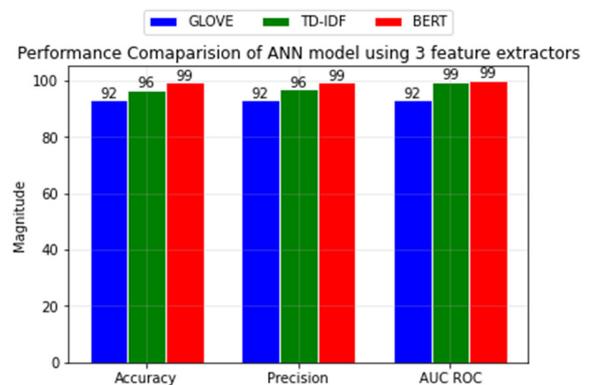


Fig. 4. Analysis of feature extractors' performance on the second dataset.

TABLE I. PERFORMANCE EVALUATION OF FEATURE EXTRACTORS ON DATASET 1

Feature extractor	Accuracy	Precision	Recall	F1 Score	AUC ROC score	AUC PR score
BERT	96.37	94.71	98.27	96.46	99.08	98.71
TD-IDF	93.68	94.85	92.47	93.65	98.70	98.77
Glove	84.21	81.80	87.31	84.47	84.26	87.67

TABLE II. PERFORMANCE EVALUATION OF FEATURE EXTRACTORS ON DATASET 2

Feature extractor	Accuracy	Precision	Recall	F1 Score	AUC ROC score	AUC PR score
BERT	99.23	99.14	99.33	99.24	99.97	99.98
TD-IDF	96.61	96.71	96.57	96.64	99.46	99.48
Glove	92.95	92.99	92.94	92.97	92.95	94.73

V. CONCLUSION

The easy access of social media to everyone has obvious advantages but also it has some disadvantages, such as the rapid quick spread of fake news. It is a very tedious job to check every news item manually, so, in order to overcome this problem, researchers are developing algorithms to detect fake news automatically. Fake news identification is a contextual problem in which the meaning of the same words may be different depending on the context. Various feature extractors have been built to efficiently solve this problem. In this paper, we analyzed two publicly available fake news datasets using

three different feature extractors: TD-IDF vectorizer, Glove static embeddings, and BERT contextual embeddings on the fake news datasets and the outputs were fed to an ANN model for classification. It was found experimentally that the BERT model outperformed the TD-IDF and Glove in both datasets. TD-IDF outperformed Glove for both datasets and competed well with BERT.

REFERENCES

[1] T. Lima-Quintanilha, M. Torres-da-Silva, and T. Lapa, "Fake news and its impact on trust in the news. Using the Portuguese case to establish lines of differentiation," *Communication & Society*, vol. 32, no. 3, pp. 17–32, Apr. 2019, <https://doi.org/10.15581/003.32.3.17-32>.

- [2] R. Baly, G. Karadzhev, D. Alexandrov, J. Glass, and P. Nakov, "Predicting Factuality of Reporting and Bias of News Media Sources," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, Belgium, Oct. 2018, pp. 3528–3539, <https://doi.org/10.18653/v1/D18-1389>.
- [3] R. Zellers *et al.*, "Defending Against Neural Fake News," *arXiv:1905.12616 [cs]*, Dec. 2020, Accessed: Mar. 19, 2021. [Online]. Available: <http://arxiv.org/abs/1905.12616>.
- [4] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," in *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, Vancouver, Canada, Oct. 2017, pp. 127–138, https://doi.org/10.1007/978-3-319-69155-8_9.
- [5] B. A. Asaad and M. Erascu, "A Tool for Fake News Detection," in *2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, Timisoara, Romania, Sep. 2018, pp. 379–386, <https://doi.org/10.1109/SYNASC.2018.00064>.
- [6] Abdullah-Ali-Tanvir, E. M. Mahir, S. Akhter, and M. R. Huq, "Detecting Fake News using Machine Learning and Deep Learning Algorithms," in *2019 7th International Conference on Smart Computing Communications (ICSCC)*, Sarawak, Malaysia, Jun. 2019, pp. 1–5, <https://doi.org/10.1109/ICSCC.2019.8843612>.
- [7] S. Sangamerkar, R. Srinivasan, M. R. Christuraj, and R. Sukumaran, "An Ensemble Technique to Detect Fabricated News Article Using Machine Learning and Natural Language Processing Techniques," in *2020 International Conference for Emerging Technology (INCET)*, Belgaum, India, Jun. 2020, <https://doi.org/10.1109/INCET49848.2020.9154053>.
- [8] D. Chopra, N. Joshi, and I. Mathur, "Improving Translation Quality By Using Ensemble Approach," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3512–3514, Dec. 2018, <https://doi.org/10.48084/etasr.2269>.
- [9] M. Biniz, S. Boukil, F. Adnani, L. Cherrat, and A. Moutaouakkil, "Arabic Text Classification Using Deep Learning Technics," *International Journal of Grid and Distributed Computing*, vol. 11, no. 9, pp. 103–114, Sep. 2018, <https://doi.org/10.14257/ijgcd.2018.11.9.09>.
- [10] A. Hassan and A. Mahmood, "Efficient Deep Learning Model for Text Classification Based on Recurrent and Convolutional Layers," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Cancun, Mexico, Dec. 2017, pp. 1108–1113, <https://doi.org/10.1109/ICMLA.2017.00009>.
- [11] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv:1810.04805 [cs]*, May 2019, Accessed: Mar. 19, 2021. [Online]. Available: <http://arxiv.org/abs/1810.04805>.
- [12] J. Pennington, R. Socher, and C. Manning, "GloVe: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, Oct. 2014, pp. 1532–1543, <https://doi.org/10.3115/v1/D14-1162>.
- [13] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, Nov. 2011.
- [14] "Fake News: Balanced dataset for fake news analysis," *Kaggle*. <https://kaggle.com/hassanamin/textdb3> (accessed Mar. 19, 2021).
- [15] "Fake news: Fake News Classifier Using Bidirectional LSTM," *Kaggle*. <https://kaggle.com/saratchendra/fake-news> (accessed Mar. 19, 2021).
- [16] H. Christian, M. P. Agus, and D. Suhartono, "Single Document Automatic Text Summarization using Term Frequency-Inverse Document Frequency (TF-IDF)," *ComTech: Computer, Mathematics and Engineering Applications*, vol. 7, no. 4, pp. 285–294, Dec. 2016, <https://doi.org/10.21512/comtech.v7i4.3746>.
- [17] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based Framework for Text Categorization," *Procedia Engineering*, vol. 69, pp. 1356–1364, Jan. 2014, <https://doi.org/10.1016/j.proeng.2014.03.129>.
- [18] W. K. Sari, D. P. Rini, and R. F. Malik, "Text Classification Using Long Short-Term Memory With GloVe Features," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 5, no. 2, pp. 85–100, Dec. 2019, <https://doi.org/10.26555/jiteki.v5i2.15021>.
- [19] U. Khan, K. Khan, F. Hassan, A. Siddiqui, and M. Afaq, "Towards Achieving Machine Comprehension Using Deep Learning on Non-GPU Machines," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4423–4427, Aug. 2019, <https://doi.org/10.48084/etasr.2734>.
- [20] T. Wolf *et al.*, "Transformers: State-of-the-Art Natural Language Processing," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Online, Oct. 2020, pp. 38–45, <https://doi.org/10.18653/v1/2020.emnlp-demos.6>.
- [21] M. Zaheer *et al.*, "Big Bird: Transformers for Longer Sequences," *arXiv:2007.14062 [cs, stat]*, Jan. 2021, Accessed: Mar. 19, 2021. [Online]. Available: <http://arxiv.org/abs/2007.14062>.
- [22] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," *arXiv:1301.3781 [cs]*, Sep. 2013, Accessed: Mar. 19, 2021. [Online]. Available: <http://arxiv.org/abs/1301.3781>.
- [23] S. Liu, H. Tao, and S. Feng, "Text Classification Research Based on Bert Model and Bayesian Network," in *2019 Chinese Automation Congress (CAC)*, Hangzhou, China, Nov. 2019, pp. 5842–5846, <https://doi.org/10.1109/CAC48633.2019.8996183>.
- [24] A. Hussain, G. Ali, F. Akhtar, Z. H. Khand, and A. Ali, "Design and Analysis of News Category Predictor," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6380–6385, Oct. 2020, <https://doi.org/10.48084/etasr.3825>.

Development of an Automatic Contactless Thermometer Alert System Based on GPS and Population Density

Naif K. Al-Shammari

Department of Mechanical Engineering
College of Engineering
University of Ha'il
Hail, Saudi Arabia
naif.alshammari@uoh.edu.sa

Husam B. Almansour

Department of Health Management
College of Public Health and Health Informatics
University of Ha'il
Hail, Saudi Arabia
h.almansour@uoh.edu.sa

Muzamil Basha Syed

School of Computing and Information Technology
REVA University
Bangaluru, Karnataka, India
muzamilbasha.s@reva.edu.in

Abstract-In today's out-breaking Covid-19 circumstance, treatments are preferred to be contactless. Social distancing has become a mandate in order to prevent disease spreading. In such a scenario, checking the body temperature is preferable to be made contactless because it helps the doctors and social workers to stay away from the symptomatic patients. Infrared (IR) contactless thermometers are employed in measuring the temperature while preventing direct contact with the body. Improved functionalities in the contactless thermometer can provide accurate precision in measurements and calculations. Technological advancement in pharmacy has cohesively improved over time. Coupling Machine Learning (CML) will revolutionize the process of testing. The demand for automated temperature test equipment is likely to grow at a significant pace, with the continuous advancements in technology and the adoption of ATE (Automated Test Equipment). The Global Positioning System (GPS) easy tracking and navigation can be used for easy tracking. Population density can be used to calculate the amount of population in a particular area. The proposed automatic contact-less thermometer system has the potential to replace the traditional temperature measuring techniques and safeguard from human-to-human transmission diseases.

Keywords-Covid-19; IR technology; ATE; GPS; population density

I. INTRODUCTION

The Global Positioning System (GPS) is used to determine locations using a satellite navigation process. It is well suited for different kinds of tracking applications [1] and is used for safety and security purposes. The main components of the GPS system are the space segment, control segment, and user segment. These segments collaborate with each other and

process the signals received from the satellites and use them for further information processing like tracking, security in banking applications, automatic toll services, location identification, and navigation [2]. In our example, the idea is to develop an automatic contactless thermometer system based on GPS location (latitude-longitude) and population density. Population density is used to calculate the amount of population in a particular area. It is calculated as [3]:

$$Pd = \frac{Pn}{Ln} \quad (1)$$

where Pd is the population density, Pn is the total population, and Ln the total land area covered by the population. The calculation of population density varies depending on environmental (climate, resources) and human (economic and societal) factors [4]. Globally, millions of Covid-19 and temperature tests are conducted every day. Testing temperature for every person is very crucial: most enterprises, governments, and organizations are following very strict protocols in order to minimize the impact. Educational organizations are struggling to reopen schools, colleges, and research centers, mainly due to the lack of testing centers. Testing is time-consuming, and is conducted with low accuracy. The automatic contactless thermometer provides a solution. The contactless process is efficient and provides easy scalability among large groups. Automatic temperature test-centers will be the future regarding this matter. Public places, and the transportation sector need such technology during these hours.

In Figure 1, the details of the product development stages are presented. The first and second stages of the project are already completed, with stage three being our future plan. The increased adoption of smartphones brings a great opportunity to

Corresponding author: Naif Khalaf Al-Shammari

locate people affected with covid-19. Smartphones enable the governments to track and locate affected people and virtually fencing people is now possible. Technological advancement in pharmacy has improved with times. Coupling Machine Learning (CML) will revolutionize the way we test. The demand for automated temperature test equipment is likely to grow at a significant pace along with the continuous advancement in technology and the adoption of ATE (Automated Test Equipment). The increased adoption of smartphones brings great opportunity to locate Covid-19 affected people. Focusing on minimizing error in detection is a very important task. Using an automated contact-less thermometer and geolocation feature from a smartphone will help monitoring on a large scale. However, technology nowadays has become more sophisticated. For example, the number of available infrared (IR) detectors has highly increased, and thanks to their selective filtering capabilities, these detectors can promise more efficiency and speed.

thermometer system to work in such an environment where the system can be deployed in identifying critically ill people. The proposed work is automated to initiate the functioning of the thermometer system by gathering the location identification of the densely populated area with the help of the GPS tracking system, and the IR signals of the contactless thermometer that provide an accurate measurement of the body temperature. As an outcome, the proposed system will help in identifying red alert areas, in which the population is affected with fever in order to isolate and prevent the general public from entering.

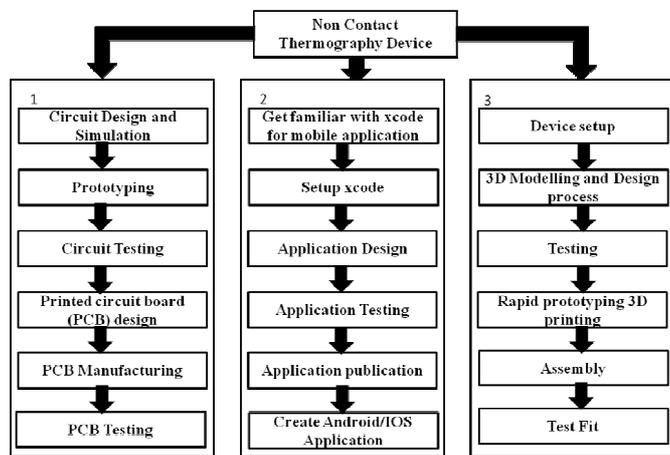


Fig. 1. Design and development steps of the complete product.

Minimizing error in detecting is a very important task. Public transportation is highly dynamic, testing the public in a fast phase and with high accuracy is essential. Non-contact temperature measurement is the preferred technique for monitoring [5]. During the Covid-19 pandemic, the bolometer IR sensor is commonly used, since it provides an IR based solution to measure the temperature and it is efficient with ~4% error rate. GPS provides satellite tracking that is very useful in many ways. The main objective of this research work is to develop and deploy an automatic alert system based on the IR thermometer. The two main parameters that are taken into consideration to fully automate the body temperature measurements are the population density and its accurate location identification. The hardware circuit diagram of the traditional non-contact IR thermometer is presented in Figure 2. In the proposed work, focus is given on developing the contact-less thermometer to work in a situation where there is a necessity to calculate the body temperature in mass gathering places.

The GPS tracking system in collaboration with the population density helps in identifying and locating densely populated areas. The aim is to develop an automatic contactless

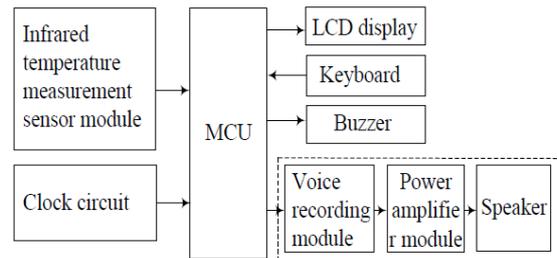


Fig. 2. The hardware circuit diagram of the non-contact IR thermometer.

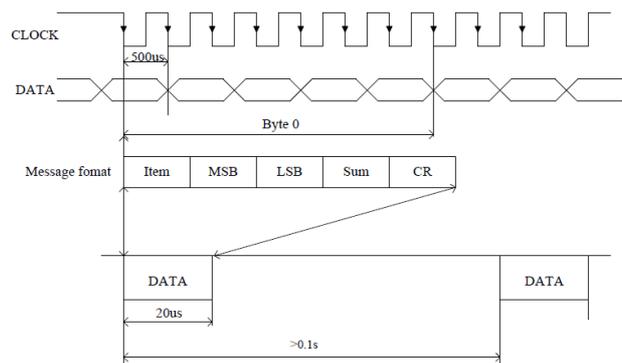


Fig. 3. The timing diagram of the IR temperature measurement module.

II. LITERATURE REVIEW

Globally, health centers have uploaded Covid-19 strains long ago and researchers and scientists are reverse-engineering the strains to find the solution [6]. Research institutions collect data regarding temperature, location, test results, and personal details to predict the future consequences using Machine Learning algorithms.

A. Smart IR Thermometers

Companies are focused on developing smart IoT thermometers. Real-time trackability is very crucial. To provide compliance with non-contact temperature measurements, it is necessary to calculate the effective efficiency of the black body and the difference in temperature between the workplace and the point or object in which the contact temperature is set [7]. Effective recreation can be calculated by applying the theory of evolution or by adjusting the temperature. The latter method seems to be difficult to use, as long as the black hole is cleaned.

B. Health-Internet of Things (H-IoT)

Considering the impact of the 2019 pandemic, it is natural that many contact tracing and self-assessment mobile applications have already been developed, such as the Aarogya

Setu App1 of the Indian Government. Health organizations like Apollo have developed chatbots for analyzing the symptoms. Health IoT devices typically belong to two broad types, personal H-IoT devices and clinical H-IoT devices [8, 9].

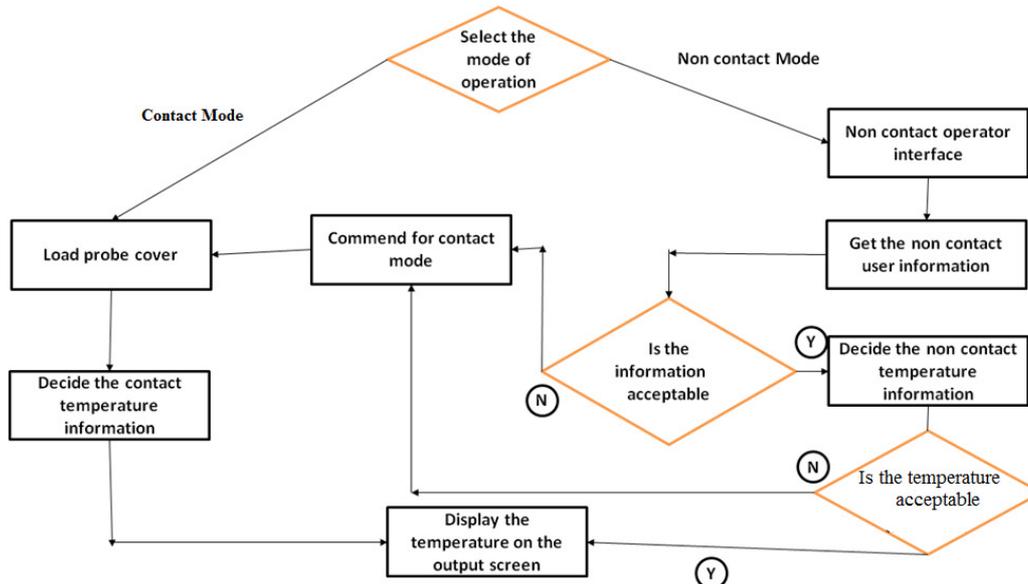


Fig. 4. Flow chart of the automatic test equipment.

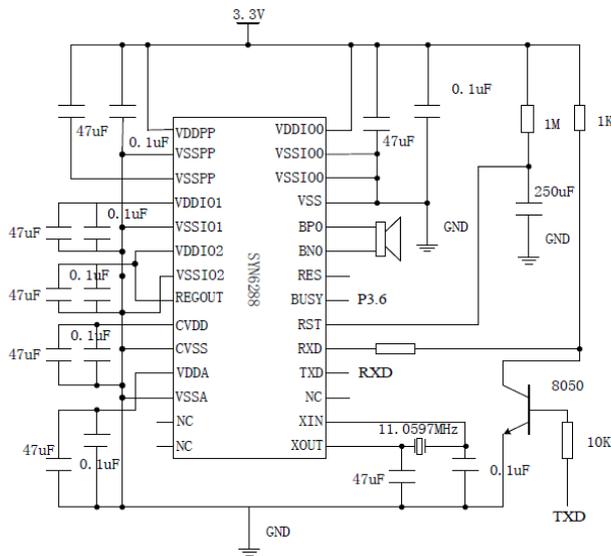


Fig. 5. The voice broadcast circuit.

III. METHODOLOGY

In Figure 4, the data flow between the modules of the proposed technique is presented. In the first step, the non-contact mode of operation is selected. In step two, we get the information from the source using sensors. If the reading received from the sensor is within the acceptable range, then the reading temperature is processed for future decision making. Else, the system is switched to contact mode to get the temperature, thus always making sure that acceptable accurate

temperature is displayed on the display unit. Switching between the possible modes is only made based on the format of the data received from the sensor and the acceptable range of the received data. The circuit diagram is presented in Figure 5 with complete details.

IV. EXPERIMENTAL SETUP AND RESULTS

The circuit unit is simulated on the Proteus platform, where the PCB board is designed and built after simulation. The MCU download system and other components are mounted on the PCB board [9]. The temperature of the object is usually simultaneously measured by a mercury thermometer and with a non-contact IR thermometer, time is measured by the length of the clock, temperature is set at 25°C to 48°C. The results of the measurements are: Temperature error is less than 0.1°C, measurement time is about 5s with an IR contact thermometer. The measurements are accompanied with a voice message transmission, and LCD display of temperature and time. In general, the measured data are very accurate. A screenshot of the experimental work is presented in Figure 6. The display unit IR with the VBC is presented in Figure 7. When a temperature of an object below 25°C is detected with the IR contact thermometer, a buzzer alarm and a voice transmission take place. Object temperature at 25°C ~ 48°C is accompanied with voice transmission and LCD display of the temperature and time. Temperature above 48°C is accompanied with a buzzer alarm, and a voice transmission. The data from the sensors are fed into the circuit and the outcome is obtained from the display component shown in Figure 7. To estimate the performance of the experimental setup, the collected data (temperature, GPS attributes) from the sensors are stored in an

Excel sheet. The corresponding output label (message to be displayed) for each input is included to establish the train data with the help of human radiologists for diagnosis. In order to test the experimental setup, new data instances are used (test data) and the output is stored in the same sheet. The Decision Tree model is constructed using the trained dataset [10]. The evaluation parameters used to estimate the performance of the model constructed are shown in Table I.

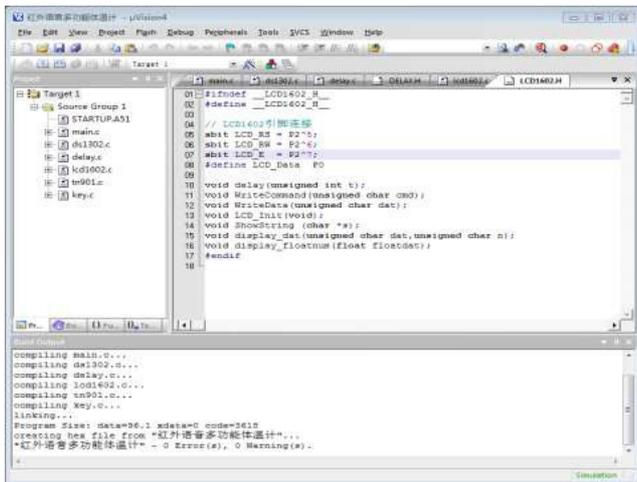


Fig. 6. Matlab experimental setup.

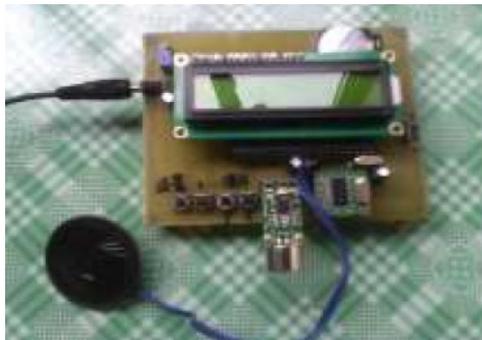


Fig. 7. Display unit interface.

TABLE I. EVALUATION PARAMETERS

Parameters	Formula
Accuracy (A)	$A = \frac{TP+TN}{Total}$
Misclassification Rate (MCR)	$MCR = \frac{FP+FN}{Total}$
Recall (R)	$R = \frac{TP}{Actual True}$
Precision (P)	$P = \frac{TP}{Predicted True}$
Prevalence (PV)	$PV = \frac{Actual True}{Total}$
F Score (FS)	$FS = 2 \times \frac{R \times P}{R + P}$

TP: true positive, TN: true negative, FP: false positive, FN: false negative

In Figure 8, the x-axis represents the evaluation metric used in estimating the performance of the experiment carried and in the y-axis the performance of the experiment is shown on a scale of 0 to 1. The experiment carried out in the present research work is reliable and proved to be robust with classification accuracy $A=0.96$. The present experimental setup and its performance are compared with the results of the similar research carried out in [12], in which, the recorded temperature is $\pm 0.2^{\circ}C$ accurate. The reason for achieving such accuracy is because of the metal cap attached to the thermopile sensor. The present circuit is designed to have ± 0.5 trueness (P), which is better than the design of [11]. The provision of switching between contactless and contact modes allows much better precision in the current design. The GPS and population density values are used in identifying the crowded areas. The aim of the proposed work is to keep the average temperature $36.5-37.5^{\circ}C$ in the tested locations. The latitude and longitude of the locations help to give alert messages to people in problematic areas in order to maintain their social distance.

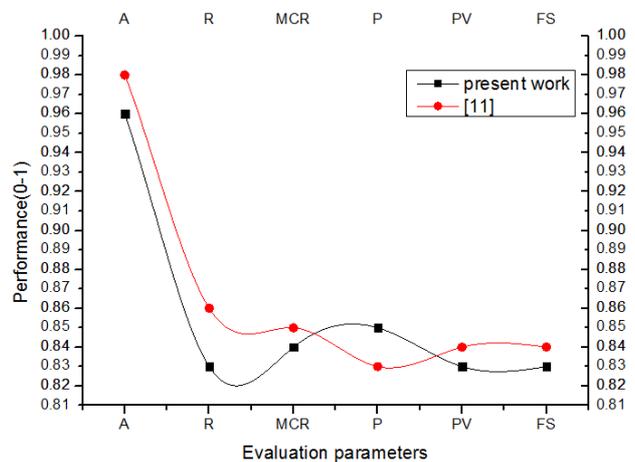


Fig. 8. Accuracy plot of the experimental work.

V. CONCLUSIONS AND FUTURE WORK

Advancement in technology and infrastructure will give us the cure to Covid-19 and its fellow strains. Automated testing and GPS-based geofencing will provide trackability on large scale. The GPS tracking system in collaboration with the population density help in identifying and locating the densely populated areas. The proposed system is aimed to develop an automatic contact-less thermometer system to work in such an environment where the system can be deployed in identifying critically ill people. The advantage of the proposed model is that it uses GPS coordinates and population density values towards identifying the most crowded areas. The aim of the proposed work is to keep the average temperature between 36.5 and $37.5^{\circ}C$ in the tested locations. The latitude and longitude details of the locations allow sending alert messages in order to maintain social distance. The limitation of the product designed in the present research work is that its contactless mode has not been implemented yet. Regarding future work, we would like to extend our work in adding a metal cap to the thermopile sensor towards achieving better accuracy.

TABLE II. CHARACTERISTICS OF THE DATASET

Name of the area (India)	Latitude	Longitude	Density level	Population	Area (m ²)	Density value	Average temperature of the area (°C)	Alert message (text)
Coimbatore	11.0168	76.9558	Low	190	7516	0.0252	36.6	Temperature is normal
Coimbatore	11.0168	76.9558	Low	239	13,200	0.10176	36.6	Temperature is normal
Bangalore	12.9716	77.5946	Low	670	32,666	0.0205	36.7	Temperature is normal
Bangalore	12.9716	77.5946	Low	189	90,100	0.0012	37.4	Temperature is normal
Coimbatore	11.0168	76.9558	High	165	5501	0.03	37.4	Temperature is unusually high
Coimbatore	11.0168	76.9558	High	208	6890	0.0303	37.9	Temperature is unusually high
Coimbatore	11.0168	76.9558	High	260	9634	0.0304	38.2	Temperature is unusually high
Bangalore	12.9716	77.5946	High	259	6890	0.0303	38.3	Temperature is unusually high

ACKNOWLEDGMENT

This research has been funded by the Scientific Research Deanship of the University of Ha'il, project number COVID-1908.

REFERENCES

- [1] R. Maddison and C. Ni Mhurchu, "Global positioning system: a new opportunity in physical activity measurement," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 6, no. 1, Nov. 2009, Art. no. 73, <https://doi.org/10.1186/1479-5868-6-73>.
- [2] F. Abulude, A. Akinnusotu, and A. Adeyemi, "Global Positioning System and It's Wide Applications," *Continental Journal of Information Technology*, vol. 9, no. 1, pp. 22–32, Oct. 2015, <https://doi.org/10.5707/cjit.2015.9.1.22.32>.
- [3] "Population Density: Definition, Formula & Examples - Video & Lesson Transcript," *Study.com*. <https://study.com/academy/lesson/population-density-definition-formula-examples.html> (accessed Mar. 21, 2021).
- [4] J. F. McDonald, "Econometric studies of urban population density: A survey," *Journal of Urban Economics*, vol. 26, no. 3, pp. 361–385, Nov. 1989, [https://doi.org/10.1016/0094-1190\(89\)90009-0](https://doi.org/10.1016/0094-1190(89)90009-0).
- [5] J.-H. Min, N. Gelo, and H. Jo, "Non-contact and Real-time Dynamic Displacement Monitoring using Smartphone Technologies," *Life Cycle Reliability and Safety Engineering*, vol. 4, no. 2, pp. 40–51, May 2015.
- [6] T. M. Fernández-Caramés and P. Fraga-Lamas, "Towards The Internet of Smart Clothing: A Review on IoT Wearables and Garments for Creating Intelligent Connected E-Textiles," *Electronics*, vol. 7, no. 12, p. 405, Dec. 2018, <https://doi.org/10.3390/electronics7120405>.
- [7] P. F. Paradis and W. Rhim, "Non-Contact Measurements of Thermophysical Properties of Titanium at High Temperature," *Journal of Chemical Thermodynamics*, vol. 32, no. 1, pp. 123–133, Mar. 1999.
- [8] V. Siva Brahmaiah, B. Rajkiran, and S. Pradeep, "An Intellectual Dual Axis Efficient Solar Tracking System by Using IoT Integrated Controller," *SSRN*, <https://dx.doi.org/10.2139/ssrn.3554005>.
- [9] S. K. Kumaravel *et al.*, "Investigation on the impacts of COVID-19 quarantine on society and environment: Preventive measures and supportive technologies," *3 Biotech*, vol. 10, no. 9, Sep. 2020, Art. no. 393, <https://doi.org/10.1007/s13205-020-02382-3>.
- [10] S. M. Basha and D. S. Rajput, "Chapter 9 - Survey on Evaluating the Performance of Machine Learning Algorithms: Past Contributions and Future Roadmap," in *Deep Learning and Parallel Computing Environment for Bioengineering Systems*, A. K. Sangaiah, Ed. Academic Press, 2019, pp. 153–164, <https://doi.org/10.1016/B978-0-12-816718-2.00016-6>.
- [11] E. Moisello, M. Vaiana, M. E. Castagna, G. Bruno, P. Malcovati, and E. Bonizzoni, "An Integrated Micromachined Thermopile Sensor With a Chopper Interface Circuit for Contact-Less Temperature Measurements," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 9, pp. 3402–3413, Sep. 2019, <https://doi.org/10.1109/TCSI.2019.2928717>.
- [12] B. Trstenjak, D. Donko, and Z. Avdagic, "Adaptable Web Prediction Framework for Disease Prediction Based on the Hybrid Case Based Reasoning Model," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1212–1216, Dec. 2016, <https://doi.org/10.48084/etasr.753>.

AUTHORS PROFILE

Naif Khalaf Al-Shammari Academic Qualifications: 1993 B.S., MEng, KSU, rating (4.25/5), 1999 M.S., MEng, KSU, rating (4.56), 2011 PhD, BEng, Birmingham University, UK. Professional Experience: 1994-2000 Saudi Electric Company, 2001-2010 Sultan B.A Al Saud Foundation. 2010-to date: University of Hail Research Interests Rehabilitation Engineering Robotics & Control. Scientific areas: Mechanisms design, biomechanical systems and bio/nano-robotics, dual-arm & flexible manipulator dynamics, smart materials, and structure mechatronics.

Husam Almansour is a distinguished doctor in Nursing Education and Leadership and Management programs at the University of Hail. He worked as the Head of the Health Management Department and as a vice-dean for Quality and Development of the Faculty of Public Health & Health Informatics. He has been inclined with research focusing on practice and education.

Syed Muzamil Basha completed his PhD at Vit University, Vellore, India in 2020. His current research interests include BigData Analytics, BlockChain Management, and Internet of Things. He has more than 30 publications, which include 21 research papers, and 7 conference papers, and 6 book chapters. He is an editor of 4 textbooks. He conducted guest lectures in various Engineering Colleges in India. He received an award for his scientific contribution from VIT in 2018.

Comparative Study between Fuzzy Logic and Interval Type-2 Fuzzy Logic Controllers for the Trajectory Planning of a Mobile Robot

Boucetta Kasmi

Department of Electronics
Laboratory of Intelligent Systems
University Ferhat Abbas Setif 1
Setif, Algeria
boucetta.kasmi@univ-setif.dz

Abdelouaheb Hassam

Department of Electronics
Laboratory of Intelligent Systems
University Ferhat Abbas Setif 1
Setif, Algeria
abdelhassam@univ-setif.dz

Abstract—In this study, Fuzzy Logic (FL) and Interval Type-2 FL (IT-2FL) controllers were applied to a mobile robot in order to determine which method facilitates navigation and enables the robot to overcome real-world uncertainties and track an optimal trajectory in a very short time. The robot under consideration is a non-holonomic unicycle mobile robot, represented by a kinematic model, evolving in two different environments. The first environment is barrier-free, and moving the robot from an initial to a target position requires the introduction of a single action module. Subsequently, the same problem was approached in an environment closer to reality, with objects hindering the robot's movement. This case requires another controller, called obstacle avoidance. This system allows the robot to reach autonomously a well-defined target by avoiding collision with obstacles. The robustness of the structures of the defined controllers is tested in Matlab simulations of the studied controllers. The results show that the IT-2FL controller performs better than the FL controller.

Keywords—Interval Type-2 Fuzzy Logic (IT2-FL); Fuzzy Logic (FL); mobile robot; non-holonomic; obstacle avoidance; trajectory planning

I. INTRODUCTION

A mobile robot is generally equipped with perception and decision-making capabilities and actions that allow it to navigate safely and successfully in a given environment and to follow a desired trajectory. This objective is achieved without or with a reduced human intervention [1]. Today, independent mobile robots are increasingly used in many applications, such as service tasks, agriculture, handling of nuclear waste, and in space industry [2]. The developments of sensors, microprocessors, and control technology have enabled mobile robots to perform very complex tasks. Today, the main challenge regarding mobile robots is the development of intelligent navigation systems [3]. Navigation is of great importance, since virtually every task requires the robot to travel between different positions by tracking a desired trajectory while being able to localize itself and plan its future movements without human assistance [4] in order to accomplish the defined task [5]. Therefore, once the trajectory

of the robot in a complex environment is determined, it must be capable of following it [6]. The generated trajectory must take into account the environmental and kinematic constraints of the moving object [7-8]. To tackle these difficulties and to enable a large number of parameters to be managed, a new control strategy is proposed, which is broken down into weakly coupled units whose interactions are limited and perfectly controlled in a distributed manner. The purpose is to develop robots that can move safely in unstructured environments, despite any unforeseen changes.

The main problem of navigation of mobile robots can be broken down into three sub-problems: reaching the target, avoiding obstacles, and tracking an optimal trajectory in a very short time. To reach the target, we can indicate the trajectory to the robot by employing stable techniques based on the search for optimal trajectories. However, since the environments are rarely predictable, it makes little sense to provide the robot with a planned trajectory. Obstacle avoidance has often been solved by using local information, perceived by the robot via its sensors. Initially, research has focused on solving the obstacle avoidance problem by presenting it as a high-level control component of hierarchical robotic systems. Thus, the problem arose as a trajectory planning where the controller at the low level leads the robot towards its final destination while avoiding the surrounding obstacles using the trajectory found at the high level.

The main objective of this paper is the tracking of an optimal trajectory in a very short time, which is the third sub-problem. A comparative study of Fuzzy Logic (FL) and Interval Type-2 FL (IT-2FL) controllers was conducted in order to assess their performance. The robot under consideration is a non-holonomic unicycle mobile robot, represented by a kinematic model.

II. MODELING OF A MOBILE ROBOT

The differential mobile robot is a platform with two motorized wheels (Figure 1), mounted on the same axis and controlled independently while having in addition a free front

Corresponding author: Boucetta Kasmi

wheel [7]. The simplified hypotheses considered for modeling are:

- The ground wheel contact is a contact point.
- The rolling of each wheel is done without slipping.

The generalized coordinates of the system are given by $q = [x, y, \theta, \varphi_r, \varphi_l]^T$ where $[x, y]$ are the Cartesian coordinates of the mobile robot, θ is its orientation measured from the x-axis and φ_r, φ_l are the angular positions of the right and left wheel respectively [2, 3].

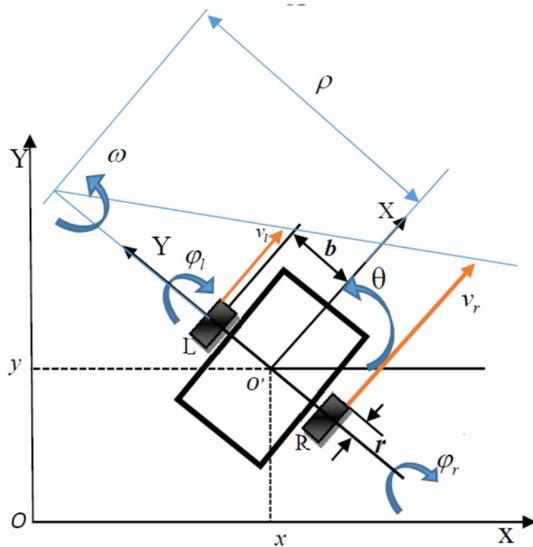


Fig. 1. The unicycle-type mobile robot under study.

III. KINEMATIC MODEL OF THE MOBILE ROBOT

The kinematic model of the mobile robot can be written as [3]:

$$\dot{q} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\varphi}_r \\ \dot{\varphi}_l \end{bmatrix} = \begin{bmatrix} \frac{r \cos(\theta)}{2} & \frac{r \cos(\theta)}{2} \\ \frac{r \sin(\theta)}{2} & \frac{r \sin(\theta)}{2} \\ \frac{r}{2b} & \frac{-r}{2b} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\varphi}_r \\ \dot{\varphi}_l \end{bmatrix} = \begin{bmatrix} \dot{\varphi}_r \\ \dot{\varphi}_l \end{bmatrix} \quad (1)$$

By introducing the following control inputs:

$$\begin{bmatrix} \dot{\varphi}_r \\ \dot{\varphi}_l \end{bmatrix} = \frac{b}{r} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}$$

Equation (1) may be written as:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{bmatrix} v + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \omega \quad (2)$$

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} \frac{r}{2} & \frac{r}{2} \\ \frac{r}{2b} & -\frac{r}{2b} \end{bmatrix} \begin{bmatrix} \dot{\varphi}_r \\ \dot{\varphi}_l \end{bmatrix} \quad (3)$$

where v and ω are respectively the linear and angular velocities of the mobile robot and $2b$ and r represent respectively the radius of the wheels and the distance between them.

The non-holonomic constraint is represented in a simple mathematical form [3]:

$$\dot{x} \cos(\theta) - \dot{y} \sin(\theta) = 0 \quad (4)$$

Equation (4) implies that a perfect tracking is achievable only if the reference trajectories are feasible for the physical platform.

IV. LOCALIZATION OF A MOBILE ROBOT

One of the fundamental problems of autonomous mobile robotics is the locationing of the robot during its movement. In fact, to locate a mobile robot is to determine, in a given work reference, its position and its orientation, in order to accomplish the control structure that is based on these data.

A. Presentation of the Odometry

The odometry allows determining the position and the orientation of a mobile robot navigating on a plane ground, with respect to the reference mark, which is the robot's mark in its initial configuration. This technique is based on the integration of the elementary motions of the wheels measured by means of incremental encoders.

B. The Odometry for the Localization of a Mobile Robot

This locomotion system is very popular for indoor robots because of its maneuverability and ease of operation. In this case, the displacement ΔD and the elementary rotation $\Delta \theta$ of the robot model in the plane can be expressed as a function of the elementary displacements of the right and left wheels respectively Δd_r and Δd_l , by [13, 14]:

$$\Delta D = \frac{\Delta d_r + \Delta d_l}{2} \quad (5)$$

$$\Delta \theta = \frac{\Delta d_r - \Delta d_l}{2b} \quad (6)$$

where (x_k, y_k, θ_k) is the configuration of the robot at the instant k , and $(\Delta D_k, \Delta \theta_k)$ the components of the elementary displacement measured between instants k and $k+1$. The elementary rotation at time $k+1$ is:

$$\theta_{k+1} = \theta_k + \Delta \theta_k \quad (7)$$

These very simple formulas are obtained by considering that the robot moves in a straight line ΔD_k in the direction defined by θ_k , and then makes a rotation on site of $\Delta \theta_k$:

$$x_{k+1} = x_k + \Delta D_k \cos(\theta_k + \frac{\Delta \theta_k}{2}) \quad (8)$$

$$y_{k+1} = y_k + \Delta D_k \sin(\theta_k + \frac{\Delta\theta_k}{2}) \quad (9)$$

This robot control was applied in both FL and IT-2FL strategies.

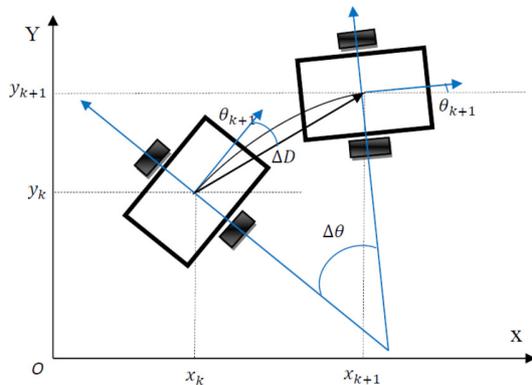


Fig. 2. The odometry applied to the mobile robot.

V. STRUCTURE OF THE FL CONTROLLER

A classic fuzzy controller consists of a fuzzification interface, a rule base, an inference system, and a defuzzification interface [15-21]. The structure of an FL system is illustrated in Figure 3.

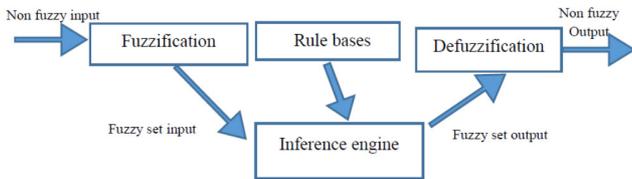


Fig. 3. Structure of an FL system.

In our work, the fuzzy controller (FL and IT-2FL) has 4 triangular-shaped membership functions for the robot-target distance, 7 membership functions for the variation of the robot-target angle and an interval-type of fuzzy sets for the linear velocity and angular velocity output for IT-2FLC. The output variables for FLC are 4 triangular-shaped membership functions for linear velocity and 7 triangular-shaped membership functions for the angular velocity.

VI. THE PROPOSED NAVIGATION SYSTEM

The movement of the unicycle robot is carried out on a flat ground and the position of the robot can be expressed at every moment according to its kinematic model (x, y, θ) . When meeting obstacles or walls, the relevant decision is made by two controllers. Figure 4 shows the structure of the system, consisting of a free navigation controller and an obstacle avoidance controller [8-10].

A. Implementation of the Free Navigation Controller

If we take a mobile robot operating in a non-binding environment, then the optimal path from an initial configuration to a final situation would naturally be a straight line joining the two situations (Figure 5) [8, 11, 12].

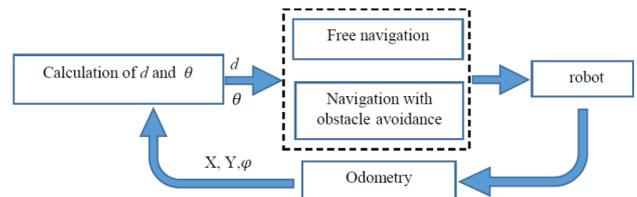


Fig. 4. Structure of the proposed navigation system.

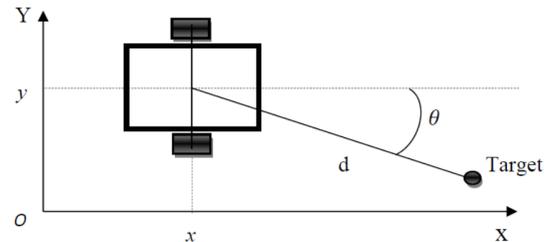


Fig. 5. The mobile robot in a free environment.

1) Input Variables

For the robot-target distance d (m), we have chosen 4 membership functions (Figure 6): very near (VN), near (NR), far (F), and very far (VF) distributed over the discourse universe $[0, 30]$. For the entry θ (rad), the robot-target angle has 7 membership functions that are associated (Figure 7): negative big (NB), negative medium (NM), negative small (NS), zero (Z), positive small (PS), positive medium (PM), and positive big (PB) distributed over the discourse universe $[-3, 3]$.

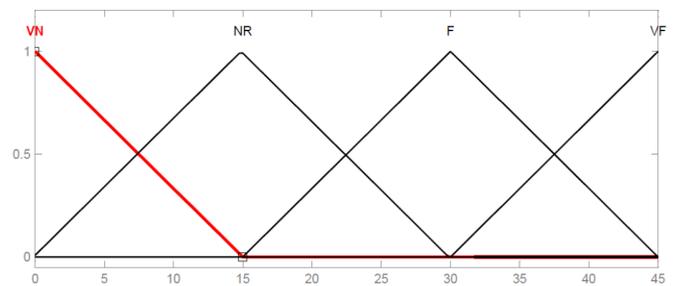


Fig. 6. The membership functions of the input variable distance d .

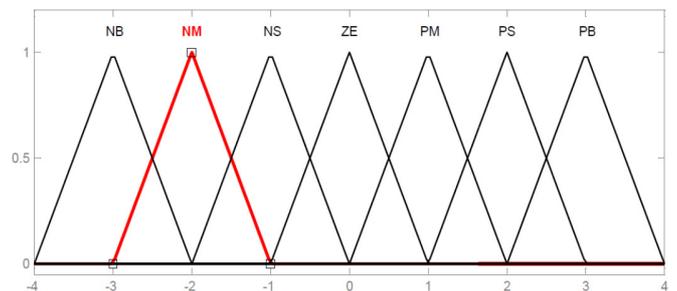


Fig. 7. The membership functions of the input variable θ .

2) Output Variables

For the linear velocity v (m/s), 4 intervals were chosen: very slow (VS), slow (S), fast (F), and very fast (VF) distributed over the discourse universe $[0, 0.2]$.

The angular velocity ω (rad) has seven (7) intervals: negative big (NB), negative medium (NM), negative small (NS), zero (Z), positive small (PS), positive medium (PM) and positive big (PB), distributed over the discourse universe [-0.8, 0.8].

3) Representation of Input Variables

The membership functions of the triangular input variables are shown in Figures 8, 9.

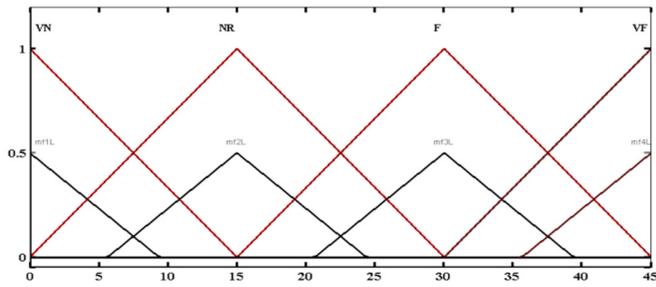


Fig. 8. The membership functions of the input variable d in Matlab Toolbox.

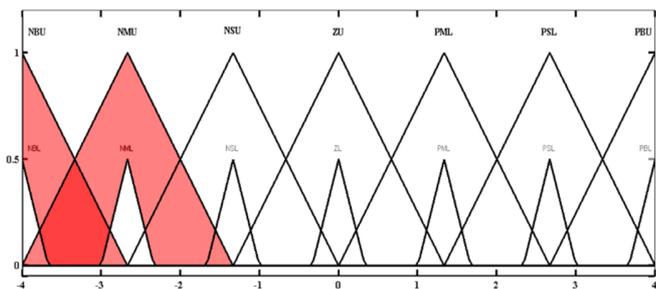


Fig. 9. The membership functions of the input variable θ in Matlab Toolbox.

4) The Rule of the Free Navigation Controller

TABLE I. THE RULE BASE OF THE FREE NAVIGATION CONTROLLER

Orientation θ	Distance d			
	VN	N	F	VF
NB	VS, PB	VS, PB	S, PB	S, PM
NM	VS, PM	VS, PS	S, PS	F, PS
NS	VS, S	S, PS	F, PS	F, PS
Z	VS, Z	S, Z	F, Z	V, FZ
PS	VS, NS	S, NS	F, NS	VF, NS
PM	VS, NM	VS, NM	F, NM	F, NS
PB	VS, NB	VS, NB	S, NB	S, NB

The rule base of the free navigation controller is the following:

- Rule 1: IF d is VN AND θ is NB
THEN v is VS AND ω is PB
- Rule 2: IF d is VN AND θ is NM
THEN v is VS AND ω is PM
-

Rule 28: IF d is VF AND θ is PB

THEN v is S AND ω is NM

We applied these rules of the free navigation controller in both FL and IT-2FL strategies.

B. Implementation of the Obstacle Avoidance Controller

In the case where the robot moves close to an obstacle, another fuzzy controller is used to avoid the obstacle and steer the robot away from d_{ob} using the navigation controller. Figure 10 shows the configuration of the robot in the obstacle area [12, 14, 20].

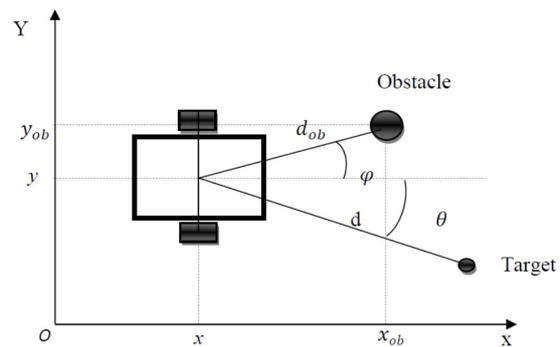


Fig. 10. The mobile robot in an environment with an obstacle.

The obstacle avoidance controller that we used has two input variables: robot-obstacle distance and robot-obstacle angle (d_{ob} and ϕ) respectively and two output variables: robot linear velocity v and angular velocity ω .

1) The Membership Functions of the Input Variables for FLC

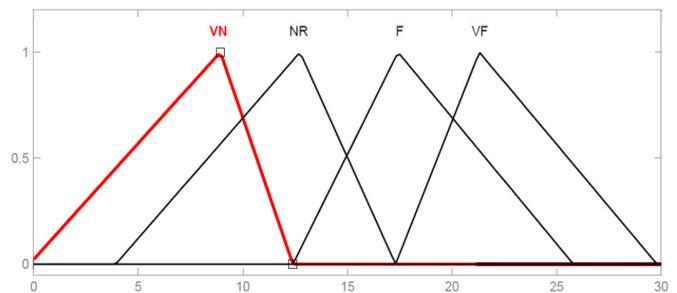


Fig. 11. The membership functions of the input variable d_{ob} .

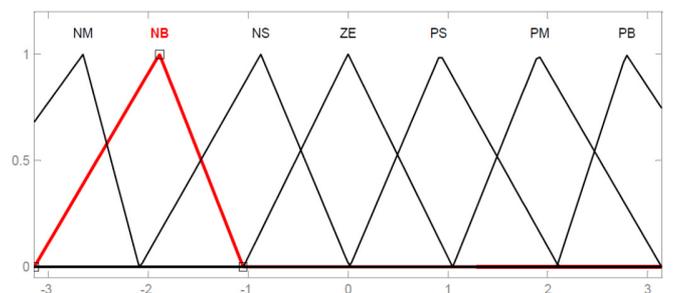


Fig. 12. The membership functions of the input variable ϕ .

2) The Membership Functions of the Input Variables for IT-2FLC

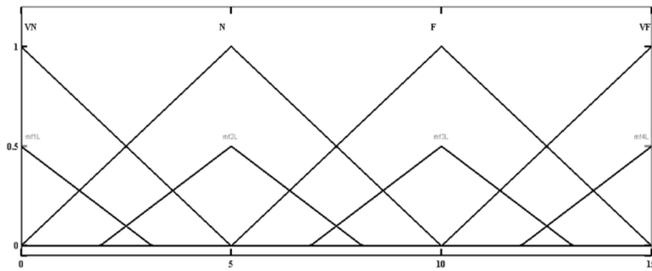


Fig. 13. The membership functions of the input variable d_{ob} in Matlab Toolbox.

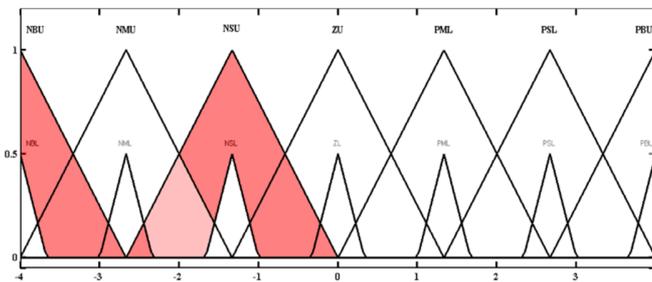


Fig. 14. The membership functions of the input variable ϕ in Matlab Toolbox.

3) The Rule Base of the Obstacle avoidance controller

TABLE II. THE RULE BASE OF THE OBSTACLE AVOIDANCE CONTROLLER

Orientation ϕ	The obstacle distance d_{ob}			
	VN	N	F	VF
NB	VS, PM	VS, PM	S, PM	S, PM
NM	VS, PM	VS, PS	S, PS	F, PS
NS	VS, PS	S, PS	F, PS	F, PS
Z	VS, PM	S, NS	F, PS	VF, PS
PS	VS, NM	S, NS	F, NS	VF, NS
PM	VS, NM	VS, NM	F, NS	F, NS
PB	VS, NB	VS, NB	S, NB	S, NB

The rule base of the obstacle avoidance controller is the following:

Rule 1: IF d_{ob} is VN AND ϕ is NB
THEN v is VS AND ω is PM

Rule 2: IF d_{ob} is VN AND ϕ is NM
THEN v is VS AND ω is PM

.....
Rule 28: IF d_{ob} is VF AND ϕ is PB
THEN v is S AND ω is NB

We applied these rules of the obstacle avoidance controller in both the FL and IT-2FL strategies.

VII. SIMULATION RESULTS

To compare the control and planning performances of FL and IT2-FL controllers, simulations were conducted and analyzed with MATLAB Fuzzy Logic Toolbox R2014a.

A. Trajectory without Obstacles

From Figure 15, we notice that from its initial position, ($x_i=2.5$ and $y_i=4.5$, $\theta=45^\circ$) the mobile robot could reach the target whose coordinates are ($x_f=7$, $y_f=-1$). In Figures 16 and 17 we can see the velocity and the angular velocity of the mobile robot. From the obtained results, we notice that the mobile robot adopts the following behavior: when the robot-target angle is large, the angular velocity is high, whereas the linear velocity is small. Once the robot-target angle becomes zero, the linear velocity reaches its maximum. The latter gradually decreases by canceling once the target is reached.

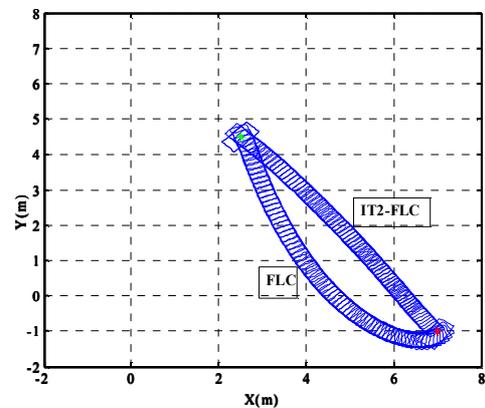


Fig. 15. Attraction to the target: barrier-free environment.

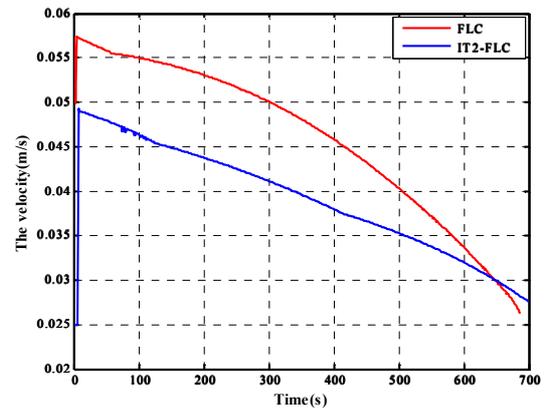


Fig. 16. The velocity of the mobile robot.

B. Obstacle Avoidance

From Figure 18 we can see how the mobile robot reached the target ($x_f=6$, $y_f=7$) from its initial position ($x_i=3$ and $y_i=8$, $\theta=-45^\circ$).

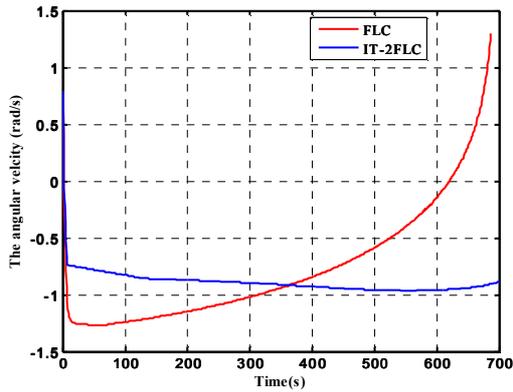


Fig. 17. The angular velocity of the mobile robot.

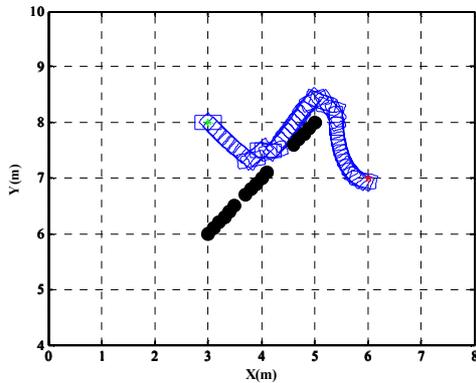


Fig. 18. Convergence towards the target by FLC in the presence of obstacles.

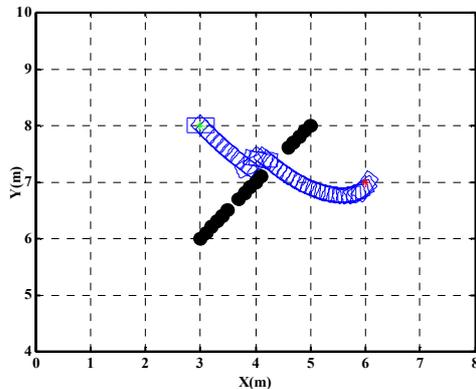


Fig. 19. Convergence towards the target by IT-2FLC in the presence of obstacles.

We can highlight from the above results that the IT-2FL sets can be quite useful when considering the control of a mobile robot. It was shown in depth that the proposed IT2-FL controller is more efficient in terms of saving time, smooth trajectory, and optimal distance than its FL counterpart.

VIII. CONCLUSION

In this paper, movement control methods of a wheeled mobile robot were studied. To achieve this control target, we

have designed the FL and IT-2FL controllers and simulated the mobile robot movement from an initial to a desired position in different environment configurations, with and without obstacles. The proposed fuzzy control exploits the interactive variables between the mobile robot and the unknown environment to generate the robot's velocity and steering, which makes it possible to bring the mobile robot towards the target while avoiding any obstacles present in this environment.

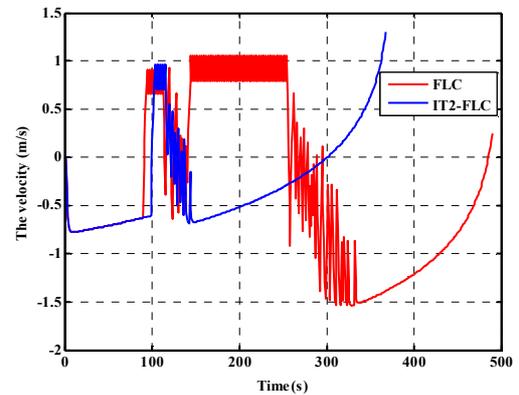


Fig. 20. The velocity of the mobile robot in the presence of obstacles.

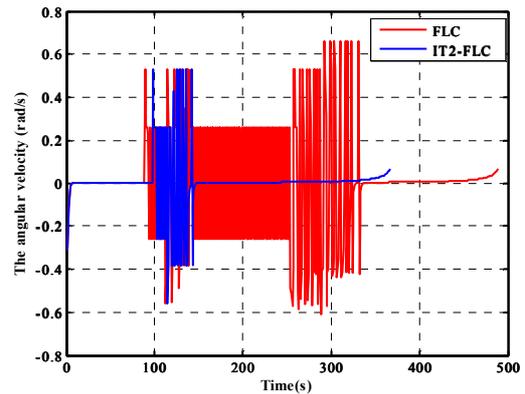


Fig. 21. The angular velocity of the mobile robot in the presence of obstacles.

In order to test the applicability of the IT-2FLC system, we compared its performance with that of an FL controller. The IT-2FLC offers better results than its FL counterpart in environments with obstacles. The main characteristic of the IT-2FLC sets is its ability to handle uncertainties more efficiently than FLC. This is made possible because a larger number of parameters and more freedom degrees are available in the IT-2FLC sets.

REFERENCES

[1] P. Gil, Y. Mezouar, M. Vincze, and J. A. Corrales, "Robotic Perception of the Sight and Touch to Interact with Environments," *Journal of Sensors*, vol. 2016, Dec. 2016, Art. no. e1751205, <https://doi.org/10.1155/2016/1751205>.
 [2] U. Libal and J. Plaskonka, "Noise sensitivity of selected kinematic path following controllers for a unicycle," *Bulletin of the Polish Academy of Sciences, Technical Sciences*, vol. 62, no. 1, pp. 3-13, Mar. 2014, <https://doi.org/10.2478/bpasts-2014-0001>.

- [3] G. Abdelhakim and H. Abdelouahab, "A New Approach for Controlling a Trajectory Tracking Using Intelligent Methods," *Journal of Electrical Engineering & Technology*, vol. 14, no. 3, pp. 1347–1356, May 2019, <https://doi.org/10.1007/s42835-019-00112-1>.
- [4] B. Hua, E. Rama, G. Capi, and M. Jindai, "A human-like robot intelligent navigation in narrow indoor environments," *International Journal of Information and Electronics Engineering*, vol. 6, no. 5, pp. 308–312, Sep. 2016, <https://doi.org/10.18178/IJIEE.2016.6.5.644>.
- [5] K. Chen, F. Yang, and X. Chen, "Planning with task-oriented knowledge acquisition for a service robot," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, New York, NY, USA, Jul. 2016, pp. 812–818.
- [6] Y. Gigras and K. Gupta, "Artificial Intelligence in Robot Path Planning," *International Journal of Soft Computing and Engineering*, vol. 2, no. 2, pp. 471–474, May 2012.
- [7] B. Damas and J. Santos-Victor, "Avoiding moving obstacles: the forbidden velocity map," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, St. Louis, MO, USA, Oct. 2009, pp. 4393–4398, <https://doi.org/10.1109/IROS.2009.5354210>.
- [8] W. Yu, J. Peng, X. Zhang, and K.-C. Lin, "A Cooperative Path Planning Algorithm for a Multiple Mobile Robot System in a Dynamic Environment," *International Journal of Advanced Robotic Systems*, vol. 11, no. 8, Aug. 2014, Art. no. 136, <https://doi.org/10.5772/58832>.
- [9] Xiaoyu Yang, M. Moallem, and R. V. Patel, "A layered goal-oriented fuzzy motion planning strategy for mobile robot navigation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 6, pp. 1214–1224, Dec. 2005, <https://doi.org/10.1109/TSMCB.2005.850177>.
- [10] D. J. Huh, J. H. Park, U. Y. Huh, and H. I. Kim, "Path planning and navigation for autonomous mobile robot," in *IEEE 2002 28th Annual Conference of the Industrial Electronics Society. IECON 02*, Seville, Spain, Nov. 2002, vol. 2, pp. 1538–1542, <https://doi.org/10.1109/IECON.2002.1185508>.
- [11] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet, "A generic framework for topological navigation of urban vehicle," presented at the ICRA09 - Workshop on Safe navigation in open and dynamic environments Application to autonomous vehicles, May 2009.
- [12] T. Belker and D. Schulz, "Local action planning for mobile robot collision avoidance," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Lausanne, Switzerland, Sep. 2002, vol. 1, pp. 601–606, <https://doi.org/10.1109/IRDS.2002.1041457>.
- [13] M. Defoort, J. Palos, A. Kokosy, T. Floquet, W. Perruquetti, and D. Boulinguez, "Experimental Motion Planning and Control for an Autonomous Nonholonomic Mobile Robot," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, Rome, Italy, Apr. 2007, pp. 2221–2226, <https://doi.org/10.1109/ROBOT.2007.363650>.
- [14] E.-H. Guechi, J. Lauber, and M. Dambrine, "Suivi de trajectoire d'un robot mobile non holonome en présence de retards sur les mesures," in *2010 IEEE CIFA*, Nancy, France, 2010.
- [15] N. N. Karnik, J. M. Mendel, and Qilian Liang, "Type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 6, pp. 643–658, Dec. 1999, <https://doi.org/10.1109/91.811231>.
- [16] O. Kahouli, B. Ashammari, K. Sebaa, M. Djebali, and H. H. Abdallah, "Type-2 Fuzzy Logic Controller Based PSS for Large Scale Power Systems Stability," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3380–3386, Oct. 2018, <https://doi.org/10.48084/etasr.2234>.
- [17] O. Castillo and P. Melin, *Type-2 Fuzzy Logic: Theory and Applications*. Berlin Heidelberg, Germany: Springer-Verlag, 2008.
- [18] B. Bouchon-Meunier and C. Marsala, *Logique floue. Principes, aide à la décision*. Paris, France: Hermes Science Publications, 2002.
- [19] Qilian Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: theory and design," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 5, pp. 535–550, Oct. 2000, <https://doi.org/10.1109/91.873577>.
- [20] N. N. Karnik and J. M. Mendel, "Type-2 fuzzy logic systems: type-reduction," in *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics*, San Diego, CA, USA, Oct. 1998, vol. 2, pp. 2046–2051, <https://doi.org/10.1109/ICSMC.1998.728199>.
- [21] H. A. Hagra, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 524–539, Aug. 2004, <https://doi.org/10.1109/TFUZZ.2004.832538>.

A Novel Drone-based Search and Rescue System using Bluetooth Low Energy Technology

Anas M. Hashmi

Department of Electrical and Electronic Engineering
University of Jeddah
Jeddah, Saudi Arabia
ahashmi@uj.edu.sa

Abstract-Drones are widely known for their mobility and ease of use and represent a significant technological breakthrough with various applications. In this paper, a novel inexpensive Search and Rescue (SAR)-based approach for application in indoor environments is presented. The usage of Bluetooth Low Energy (BLE) has been evaluated with respect to other technologies and a conceptual view of the complete setup has been presented. Besides the cost of the drone and the locator devices, the other hardware is relatively inexpensive costing only a fraction of a US dollar. The system is believed to cover a wide area in a small-time frame ranging a few minutes, for instance, a 3600m² surface area could be scanned in less than 5 minutes. The system is tested by attaching a BLE device in the payload to evaluate the presence of target beacons. Potential upgrades in the system are also proposed, including design modifications for outdoor use and the application in locating missing objects. This system can confidently replace search parties dealing with missing children in public places or venues, with minimal human interaction while bearing the potential for complete automation.

Keywords-BLE; wearable electronics; UAVs; localization

I. INTRODUCTION

During the last decade, the usage of Unmanned Aerial Vehicles (UAVs) has risen exponentially [1]. It is estimated that the global market share of multi-rotor drones is expected to reach \$36.9 billion by 2022 from \$6.8 billion in 2016 [2]. Drones are directly or indirectly used as an effective tool in agriculture, industrial maintenance, logistics, military, entertainment, and traffic control [3, 4]. In most of the aforementioned applications, drones have proven to be an effective replacement of the existing technologies, mainly in terms of cost maintenance, time, training, and automation. One of the most significant applications of UAVs lies in the domain of Search and Rescue (SAR)-based missions. Usually, such operations are performed in conventional matters, which are time consuming and costly [5]. For example, a successful 19-day rescue operation in July 2020 to locate two hikers lost in New Zealand cost the New Zealand authorities over 80k NZD (\$57.8k) [6]. Using an autonomous UAV would potentially cut the cost and time of a SAR operation by 70% in comparison with classic operations [7]. Typically, drone-based SAR operations are carried out manually by continuously monitoring live transmitted broadcast imagery through technologies, such

as the First-Person View (FPV). Some advances in the automation of the search process have been implemented which involve AI-based protocols to identify pre-trained objects or shapes, or face recognition. In confined places, such as events held at parks, theatres, or concerts, parents or guardians that lose their children usually seek the assistance of a dedicated lost child office to locate them. The proposed UAV system can potentially help in such cases or in similar cases regarding elder patients (e.g. patients with Alzheimer's or autism). The search process in these cases can often be costly and time consuming or may involve the risk of a child/patient leaving the dedicated event area [8].

This paper discusses a novel inexpensive approach of utilizing multirotor drones as a SAR tool whereby it scans passive pre-defined passive receivers. This device overcomes the characteristic challenges of events being held in large venues, allowing object/person finding applications. Potential future upgrades for transforming the system to form a completely autonomous decision support system tool for event management may even be possible. Moreover, these developments may be uploaded to upgrade the system for open space usage with certain limitations.

II. EXISTING TECHNOLOGY

There are various existing drone-based SAR applications and these attempts have been featured to be utilized in various types of technologies as described in the existing literature. Table I enlists some common technologies used in a typical SAR operation.

As shown in Table I, most of the technologies are focused on searching wireless radio activities emitted by target mobile phones. However, usually minors and individuals with special needs are not expected to carry mobile phones, therefore technology dependent on mobile phones may be difficult to implement in such cases. Other technologies achieve this task by different means which include live streaming footage, QR codes, or audio signals. Cellular network-based systems received enormous attention when it comes to SAR application due to the increased network coverage and enhanced communication capabilities, including the rapid expansion of 5G networks [16, 17].

Corresponding author: Anas M. Hashmi

TABLE I. SELECTED EXISTING SAR DRONE-BASED APPLICATIONS

Product	Technology	Features	Drawbacks	Ref.
AltiGator	IR thermal imaging	Manual inspection of live footage.	Manual inspection requires additional human power, thermal cameras can be expensive.	[9]
Intelligent drone navigation for SAR operations	Computer vision	QR codes are used for indoor localization.	Uses GPS for localization, difficult to operate indoors	[10]
SARDO	Mobile network and IR cameras	Locates a mobile phone in 3 minutes in an open field.	Mobile network-based solution, requires a phone carried by the target	[11]
SLAM	3D map generation from stitching still images	Compares movement between two frames of the images.	Images are difficult to be evaluated in crowded environments	[12]
MOBNET	Mobile phones	Searches for mobile phones, signal of trapped people.	Mobile network-based solution, requires a phone carried by the target	[13]
GuideLoc	Wireless signals from mobile phones	Localization system uses target guiding technology based on region division.	Mobile network-based solution, requires a phone carried by the target	[14]
Audio-based SAR with a drone	Drone-embedded sound source localization	Estimates the Time Difference Of Arrival (TDOA) of sound waves in a microphone pair in addition to the angle of arrival.	Difficult to operate in crowded environments due to various sources of sound	[15]

The possibility of utilizing electronic ID tags, such as Bluetooth Low Energy (BLE) devices has been explored. We believe that featuring such inexpensive tags will make the task less complex and more achievable comparing to the methods proposed in Table I.

III. CONCEPT AND METHODOLOGY

The concept of the design is to re-design the drone payload in order to be able to carry the required electronics and means of communication for accurate identification of the ID tags. In order to design the system appropriately, various elements of the design have been investigated.

A. ID Tags

The ID tags are an important element of the process of development since they act as the primary point of communication. Tags should communicate with the host once close proximity is detected, and may be operated actively or passively. They should feature reliable operation, rapid response, inexpensive price tag, reasonable energy consumption, lightweight arrangement, and acceptable communication range. If the total weight of those tags is low, they can be manufactured as wearable electronics. The recent advances in the development of short to medium range wireless technologies provided multiple options for the UAV designed in this study. There are various communication protocols meeting the aforementioned criteria, which are enlisted in Table II [18, 19]. Referring to Table II, it can be seen that the BLE technology features have low cost, low weight, moderate communication range, and relatively low energy consumption. Conventional Bluetooth has stable connection, reasonable price, and moderate IPv6 communication energy overhead with a slow data transfer rate. As the name suggests, BLE is designed to optimize power consumption compared to the classic Bluetooth, while maintaining a similar protocol stack profile. Usually BLE devices are powered by a single coin battery that lasts up to two years. It was found that BLE consumes very little energy compared to competing low-energy communication technology, such as ZigBee [20]. It uses a 2.4GHz ISM Band with 2MHz channel spacing, and is governed by IEEE 802.11 and IEEE 802.15 standard protocols

[21]. The low energy operation of BLE is achieved by the potential trade off of the following features:

- Communication channels: Instead of using multiple channels for discovery or connections, only 3 broadcasting channels would be used.
- Memory: Only a few bytes are featured with BLE, while additional memory would require additional hardware, increasing the chances of current leakage.
- Data packets: BLE sends shorter data packets which results in fewer calibration iterations drawing less leakage current.
- Sleep time: Classic Bluetooth takes over 100ms for waking from sleep, transmitting a data packet, and then going to sleep mode, while BLE performs this process 15 times faster. This will reduce scanning time for BLE saving more battery, especially if small data segments are broadcast.

TABLE II. EXISTING SAR DRONE-BASED TECHNOLOGIES

Technology	Cost	Weight	Range	Energy consumption
Radio Frequency Identification (RFID)	Very low	Low	Read range of 100m, unidirectional	Passive
Near Field Communication (NFC)	Very low	Low	10cm, bidirectional	Very low or passive
Bluetooth	Medium	Low	100m (class dependent), bidirectional	Medium (class dependent)
BLE	Low	Low	10-50m, bidirectional	Very low
Wifi	High	High	100m	High
ZigBee	Moderate	Low	10-20m	Low

BLE operates in star network topology where the nodes can be configured as slaves connected to a single master, or a master connected to multiple slaves. The master scans the designated channels to search for slave devices. Once a device is discovered, synchronous connection is established between them resulting in the change of state of activity mode from 'sleep' to 'wake up'. Another option would be to set the slave devices to sleep mode by default to save energy [22].

B. Data Transmission and Positioning

Data are transmitted in BLE using two service profiles, namely Find Me Profile (FMP) and Proximity Profile (PXP), which take into consideration the slave/master (target/locator) positions. As the name suggests, FMP makes the localization of devices a significantly straightforward process. The target seeks the search command broadcast by the locator. Once the message is delivered, the target device is triggered, usually in the form of an audio signal or other alert mechanisms. On the other hand, PXP detects any drop in the connection with the target or initiates a connection with a different locator. This profile is effective only if a connection is initiated. Other available protocols found in the domain of consumer-based electronics include the iBeacon which operates between pre-programmed manufacturer's devices.

In terms of internal hardware arrangements in BLE devices, one point to consider is the directionality of their chip antennas, which are characterized by non-isotropic radiation. This is significantly challenging to mitigate due to the size of the additional antenna. There are various fundamental concepts when considering the positioning of BLE devices. The simplest approach of positioning is determining if both the target and the locator are within the radio coverage range, while providing a binary decision about their presence within that range [23]. In order to determine the distance between devices in free space, the following equation that relates distance (d) and the Received Signal Strength Indicator (RSSI) was utilized.

$$d = 20(\text{RSSI} + \alpha) \quad (1)$$

where α is an offset based on the maximum RSSI of the locator. The equation could also be used to determine the lateration based on multiple locators. Unlike outdoors, the indoor environment usually exhibits multipath effect due to the presence of obstacles that may vary in form and arrangement. This significantly worsens the approximation taken into consideration in (1). However, the utilization of the Time-of-Flight (ToF) approach has proven to produce more accurate results in indoor locations. Unfortunately, ToF estimation requires additional electronics (e.g. oscillators) which increase power consumption and cost. It is also important to note that since the locator is fixed in a moving station (i.e. drone), the angulation process may be ignored. In the same context, it is highly challenging to utilize fingerprinting which identifies the pre-recorded target RSSI values of both the target and locator, since they tend to exhibit continuous change in their locations.

C. Locator Movement and Beacon Analysis

One of the most challenging issues when attempting to locate the target in indoor environments is the multipath propagation effect. This issue can worsen if several devices are available in a confined space and moving objects are present. Usually, no accurate RSSI model in case of indoor environments is available and a rather simplistic way of estimation involves the use of ToF values. Therefore, a worst-case scenario should be taken into consideration when planning such a system. In some commercial applications, such as the iBeacon, it is recommended that the advertising signal intervals should be set below 100ms, when the tag is operated by a coin cell battery for maximizing the battery life to a few months. In

addition, when the intervals are set just below 1000ms the battery may last up to 3 years [23].

BLE devices continuously broadcast their data packets with a unique identifier known as the Universally Unique Identifier (UUID). This UUID acts as a network MAC address for the manufactured chipset and consists of 128-bits, with a 48-bit mandatory advertiser address. For BLE target finder, an open-source program, Raspberry Pi 3 BLE sniffer was used [24]. This code allows the detection of various compatible devices from different manufacturers, with all data being analyzed in a dedicated dashboard. While initiating the software, a signal search interval of 200ms was set. The advertising signal interval was set to 500ms, so that the target could operate confidently using a coin cell battery for one full year. In Figure 1, we see the Raspberry Pi (locator) attached to a drone payload for locating advertising channels.

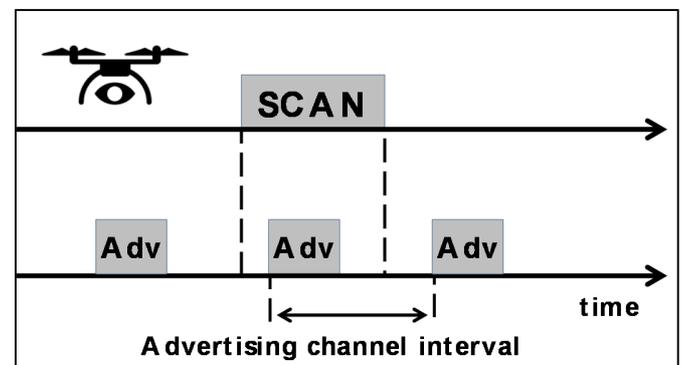


Fig. 1. Basic arrangement of target localization with scan and synchronization with the advertising beacons.

IV. EXPERIMENTAL SETUP AND ROUTE PLANNER

For route estimation purposes, the maximum Bluetooth coverage radius was assumed to be 10m. To ensure that all targets in the designated search zone are discoverable and that their discoverability is enhanced, the detection range was reduced to 5m. Moreover, the drone should fly in intervals whereby the movement was paused for 1000ms for each 5m flown. This interval of inactivity ensures that the locator exhibited at least 4 iterations of advertising signals broadcast by the target. For doubling the chances of discovery, each drone's vertical/horizontal straight parallel route should have a separation of <5m. Raw experimental data are shown in Table III.

TABLE III. RAW EXPERIMENTAL DATA

AGL	Flight duration		Wind speed
	Route length	Stop mode	
5m	Vertical movement: 12×55m = 660m Horizontal lines: 11×5m = 55m Total route: 715m	1000ms every 5m flown	0m/s (indoor)

For demonstration purposes, the exhibition hall Dhahran Expo, located in Dhahran, Saudi Arabia and covering a surface

area of 3600m^2 ($60\text{m}\times 60\text{m}$), was chosen [25]. To estimate the planned route, an open-source code named Reinforcement Learning for Autonomous navigation for UAVs [26] was used. This code was designed to implement reinforcement learning algorithms for autonomous navigation of drones in indoor environments. Figure 2 shows the exhibition hall and the planned route using predefined hall edges. The total horizontal route of the drone was $60\text{m}\times 12\text{m}=720\text{m}^2$, with an additional $5\text{m}\times 11\text{m}=55\text{m}^2$ surface area of horizontal connecting lines, which makes the total route approximately around 775m^2 . Considering the average speed of the drone to be 6.7ms^{-1} (15mph), the total area can be covered in around 115s, with an additional 1s per 5m taken into consideration. It would take a total of 227s to completely evaluate the whole area of 3600m^2 .

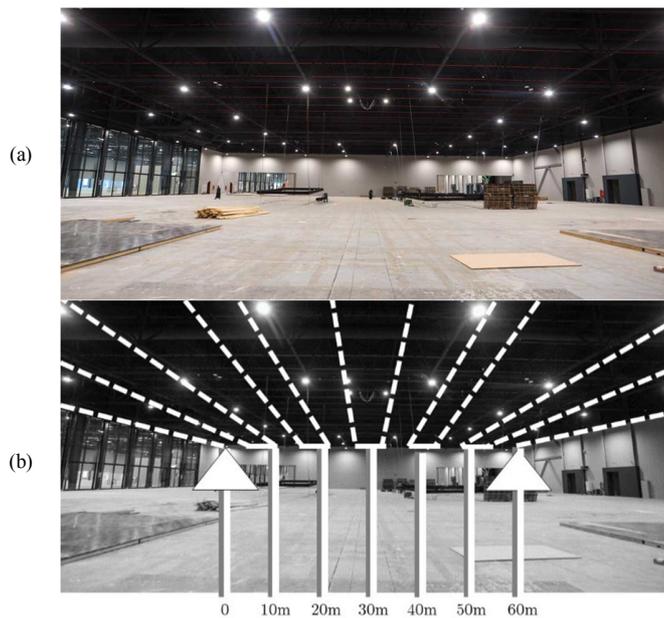


Fig. 2. (a) A raw image of hall #3 in Dhahran Expo, Dhahran, Saudi Arabia. (b) Raw image processed using planned routing with the aid of open-source code named Reinforcement Learning for Autonomous navigation for UAVs. The triangle on the left-hand side demonstrates the drone's starting point, while the triangle on the right is suggestive of the end point of the flight. The dashed lines demonstrate the main path of vertical movement, and the solid lines represent the path of the horizontal flight.

A. Accuracy and Operationality

In order to ensure the maximum accuracy of the system, repeated movements may be performed. Once the target is located, the movement of the drone is locked around this area, while it continuously performs search operations in its proximity. If the connection drops, the drone should potentially move in a spiral path to reconnect with the target. The RSSI value could be taken into consideration when determining the static positioning of the drone. The proposed model was validated by using a BLE device as a target moving backward and forward from the host. The signal loss was evaluated to ensure scan signal being continuously broadcast within the expected Bluetooth range.

B. Potential Applications

The technological development reported in this study may be used in applications involved in locating objects or persons in halls or exhibitions. This development is easy to implement, since the BLE target devices require inexpensive hardware. BLE devices could come in different forms, such as wristbands, keychains, tags, pins, each costing less than one US dollar. Therefore, the tool shows promising potential for commercialization. One potential upgrade of this application would be its outdoors potential use. In that case, uploading a suggested route plan with details of streets or landmarks may be pertinent. A swarm drone technology may be applied to cover a wider area.

V. CONCLUSIONS

A BLE-based SAR system with indoor applications was proposed which potentially could be applied to locate lost individuals. The proposed BLE-based, cell-battery powered device has a year-long operational durability, and bears relatively inexpensive electronic tags in convenient packaging such as a wristband or a keychain. The drone is able to scan an indoor area of 3600m^2 with high accuracy in just below 5 minutes. This time can be further minimized with tradeoff with battery lifetime. This system can be used for additional purposes such as event management and enhanced localization capability based on beacons from target or high-density spots based on attendees' mobile signals. Dedicated AI-based software was utilized in order to estimate the UAV route. This system could be upgraded for outdoor operations using swarm technology.

REFERENCES

- [1] S. Maharana, "Commercial Drones," *International Journal of Advances in Science Engineering and Technology*, vol. 5, no. 1 Suppl. Is. 3, pp. 96–101, Mar. 2017.
- [2] D. Gonzalez-Aguilera and P. Rodriguez-Gonzalvez, "Editorial: Drones—An Open Access Journal," *Drones*, vol. 1, no. 1, Dec. 2017, <https://doi.org/10.3390/drones1010001>.
- [3] Z. Liu, Z. Li, B. Liu, X. Fu, I. Raptis, and K. Ren, "Rise of Mini-Drones: Applications and Issues," in *Proceedings of the 2015 Workshop on Privacy-Aware Mobile Computing*, New York, NY, USA, Jun. 2015, pp. 7–12, <https://doi.org/10.1145/2757302.2757303>.
- [4] M. Almutiry, "UAV Tomographic Synthetic Aperture Radar for Landmine Detection," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5933–5939, Aug. 2020, <https://doi.org/10.48084/etasr.3611>.
- [5] G. Remy, S. M. Senouci, F. Jan, and Y. Gourhant, "SAR drones: drones for advanced search and rescue missions," *Journées Nationales des Communications dans les Transports*, vol. 1, pp. 1–3, 2013.
- [6] "Cost of missing tramper rescue operation \$80,000 and counting," *Stuff*, Jul. 21, 2020. <https://www.stuff.co.nz/national/122198869/cost-of-missing-tramper-rescue-operation-80000-and-counting> (accessed Mar. 24, 2021).
- [7] W. Stecz and K. Gromada, "UAV Mission Planning with SAR Application," *Sensors*, vol. 20, no. 4, Jan. 2020, Art. no. 1080, <https://doi.org/10.3390/s20041080>.
- [8] S. A. Riggins and D. P. Sharp, "Missing child reporting, tracking and recovery method and system," US20100262367A1, Oct. 14, 2010.
- [9] "Drones for search & rescue missions," *AltiGator Drone & UAV Technologies*, Sep. 17, 2014. <https://altigator.com/en/drones-for-search-rescue-missions/> (accessed Mar. 24, 2021).
- [10] D. Yavuz, H. Akbıyık, and E. Bostancı, "Intelligent drone navigation for search and rescue operations," in *2016 24th Signal Processing and*

- Communication Application Conference (SIU)*, Zonguldak, Turkey, May 2016, pp. 565–568, <https://doi.org/10.1109/SIU.2016.7495803>.
- [11] A. Albanese, V. Sciancalepore, and X. Costa-Perez, "SARDO: An Automated Search-and-Rescue Drone-based Solution for Victims Localization," *IEEE Transactions on Mobile Computing*, 2021, <https://doi.org/10.1109/TMC.2021.3051273>.
- [12] F. R. Heukels, "Simultaneous Localization and Mapping (SLAM): towards an autonomous search and rescue aiding drone," M. S. thesis, University of Twente, Enschede, Netherlands, 2015.
- [13] M. Półka, S. Ptak, and L. Kuziora, "The Use of UAV's for Search and Rescue Operations," *Procedia Engineering*, vol. 192, pp. 748–752, 2017, <https://doi.org/10.1016/j.proeng.2017.06.129>.
- [14] A. Wang *et al.*, "GuideLoc: UAV-Assisted Multitarget Localization System for Disaster Rescue," *Mobile Information Systems*, vol. 2017, Mar. 2017, Art. no. e1267608, <https://doi.org/10.1155/2017/1267608>.
- [15] A. Deleforge, D. D. Carlo, M. Strauss, R. Serizel, and L. Marcenaro, "Audio-Based Search and Rescue With a Drone: Highlights From the IEEE Signal Processing Cup 2019 Student Competition [SP Competitions]," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 138–144, Sep. 2019, <https://doi.org/10.1109/MSP.2019.2924687>.
- [16] S. K. Khan, A. Al-Hourani, and K. G. Chavez, "Performance Evaluation of Amplify-and-Forward UAV Relay in Millimeter-Wave," in *2020 27th International Conference on Telecommunications (ICT)*, Bali, Indonesia, Oct. 2020, <https://doi.org/10.1109/ICT49546.2020.9239435>.
- [17] S. K. Khan, M. Farasat, U. Naseem, and F. Ali, "Performance Evaluation of Next-Generation Wireless (5G) UAV Relay," *Wireless Personal Communications*, vol. 113, no. 2, pp. 945–960, Jul. 2020, <https://doi.org/10.1007/s11277-020-07261-x>.
- [18] S. J. Danbatta and A. Varol, "Comparison of Zigbee, Z-Wave, Wi-Fi, and Bluetooth Wireless Technologies Used in Home Automation," in *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, Barcelos, Portugal, Jun. 2019, pp. 1–5, <https://doi.org/10.1109/ISDFS.2019.8757472>.
- [19] N. Chhabra, "Comparative Analysis of Different Wireless Technologies," *International Journal of Scientific Research in Network Security and Communication*, vol. 1, no. 5, pp. 13–17, Dec. 2013.
- [20] M. Siekkinen, M. Hienkari, J. K. Nurminen, and J. Nieminen, "How low energy is bluetooth low energy? Comparative measurements with ZigBee/802.15.4," in *2012 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Paris, France, Apr. 2012, pp. 232–237, <https://doi.org/10.1109/WCNCW.2012.6215496>.
- [21] R. Heydon, *Bluetooth Low Energy: The Developer's Handbook*, 1st ed. Pearson, 2012.
- [22] C. Gomez, J. Oller, and J. Paradells, "Overview and Evaluation of Bluetooth Low Energy: An Emerging Low-Power Wireless Technology," *Sensors*, vol. 12, no. 9, pp. 11734–11753, Sep. 2012, <https://doi.org/10.3390/s120911734>.
- [23] A. a. M. K. Abuelgasim and K. M. Yusof, "High Speed Mobility Management Performance in a Real LTE Scenario," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5175–5179, Feb. 2020, <https://doi.org/10.48084/etasr.3245>.
- [24] "diyActive: Make a Raspberry Pi Hub," *ReelyActive*. <https://reelyactive.github.io/make-a-pi-hub.html> (accessed Mar. 24, 2021).
- [25] "Exhibition Halls," *Dhahran Expo*. <https://www.dhahranexpo.com.sa/venue-facilities-exhibition-halls.aspx> (accessed Mar. 24, 2021).
- [26] M. Liaq and Y. Byun, "Autonomous UAV Navigation Using Reinforcement Learning," *International Journal of Machine Learning and Computing*, vol. 9, no. 6, pp. 756–761, Dec. 2019, <https://doi.org/10.18178/ijmlc.2019.9.6.869>.

The Development of a Thermal-Hydraulic and Nonlinear Dynamic System for Molten Salt Reactors

Asif Ali

College of Nuclear Science and
Technology
Harbin Engineering University
Harbin, China
aliasif@hrbeu.edu.cn

Yong Kuo Liu

College of Nuclear Science and
Technology
Harbin Engineering University
Harbin, China
lyk08@126.com

Waleed Raza

College of Underwater Acoustic
Engineering
Harbin Engineering University
Harbin, China
waleed@hrbeu.edu.cn

Muhammad Nazim

College of Materials Science and
Chemical Engineering
Harbin Engineering University
Harbin, China
nazimlakhan@gmail.com

Amir Ali

College of Underwater Acoustic
Engineering
Harbin Engineering University
Harbin, China
amir@hrbeu.edu.cn

Sajid Hussain

College of Nuclear Science and
Technology
Harbin Engineering University
Harbin, China
sajidhussainmailto@gmail.com

Abstract-The Molten Salt Reactor (MSR) is the most important system suggested by Generation IV for the future direction in the nuclear reactor field. For more development of the MSR reactor, the core system inside the tube is proposed by naturally circulating molten fuel salt. The nonlinear kinetic equations form a linearized function and are obtained in state-space form. Reactivity feedback and delayed neutrons are extremely important for reactor control. In this paper, a thermal-hydraulic system for the commercial computation dynamic model is proposed. Currently, there is no commercial software to simulate the natural circulation flow. The proposed method can be easily employed to detect faults and can provide a feasible overall system performance.

Keywords-molten salt reactor; thermal-hydraulic; state-space equation; delayed neutron precursor

I. INTRODUCTION

There are six reactors in Generation-IV. Among them, Molten Salt Reactor (MSR) has a unique concept that distinguishes it from the other reactor types such as light water reactors, pressure water reactors, boiling water reactors, etc. The Generation-IV reactors do not have solid nature but utilize a homogenous liquid. The benefits of the MSR and its excellent characteristics and sustainability eliminate the fuel element fabrication accurately. Therefore, high-temperature and high power density are reached in the MSR core. In the present MSR method, the circulated fuel itself is mixed with the fuel salt. Correspondingly, the reactor control mechanism or trip primary circuit may allow or produce fuel salt in the primary circuit, and circulation of the delayed-neutron precursor is generated in the core. Hence, the decay core is less important than the outer of the core [1]. The MSR is normally operated in the liquid phase because its composition tasks are affiliated with dynamic reactors. Furthermore, in the MSR operation in

steady-state, resultantly the loss of delayed neutrons affects the reactor.

In terms of modeling, MSRs are associated with solid fuel which is used in traditional reactors. Nuclear Power Plants (NPPs) are widely sophisticated systems regarding the transit and duration of modeling time. The fuel salt is always held on the outer side of the MSR core. The multiplier term for the neutron kinetics equations is added with the delayed neutron precursor [2, 3]. The MSRs are completely different from thermal-hydraulic reactor systems. In the processes of the liquid fuel failure stability of materials can be easily resolved [3]. However, a graphite zone and a thermal-hydraulic system are developed in the core [4]. In [5, 6], the fuel region is modeled as a double lasting region, which is very important in order to improve the state space equation for the outlet core temperature. In previous studies, one or two lumped parameters were applied for the thermal-hydraulic system with two lumped regions. Resultantly, the outlet core temperature is measured more accurately. The core outlet temperature is the same as the average temperature of the upper lumped region [7]. The state-action space equation is used for the reactor, thermal-hydraulic modeling, energy balance, and controls all the volume of the core and the heat exchanger (HX) unit [8]. In MSRs, when the core temperature increases, the measurement of the core outlet temperature will be inaccurate [8].

In this paper, we explore and propose a novel thermal-hydraulic method which give us more accurate and confident results without the ambiguity of the previous models.

II. THE THERMAL-HYDRAULIC METHOD

A novel thermal-hydraulic method is proposed in which a new approach has been taken to detect and resolve the problems and faults from the core outlet. In the baseline

system, an energy equilibrium equation is obtained from fuel temperature which is used without any problems in the assembled system. A maintenance equation must be described considering the core outlet temperature. In the thermal-hydraulic method, a balance equation is formed by using the control volume of the pipeline at the core outlet. Due to this distinct temperature and the insulation, the pipe section is very small (relatively to the core region). During this process, the present temperature of the core can be calculated more accurately.

The motion of delayed neutron precursors is calculated and analysis of the fuel cycle along the state action space equation is conducted. The core neutronic as well as the extended neutronic solver 1-D thermal feedbacks are acquired by the thermal-hydraulic and neutronic calculations. The thermal neutronic features of the MSR are resolved. Before the MSR fuel studies, the thermal response was known to have high neutronic in thermal-hydraulic characteristics. Hence, the obtained power distribution from the fission salt, as corresponding fuel density distribution, and the fuel salt temperature distribution throughout the core flow loop were described by the thermal-hydraulic measurement. The schematic view of the conceptual MSR is presented in Figure 1.

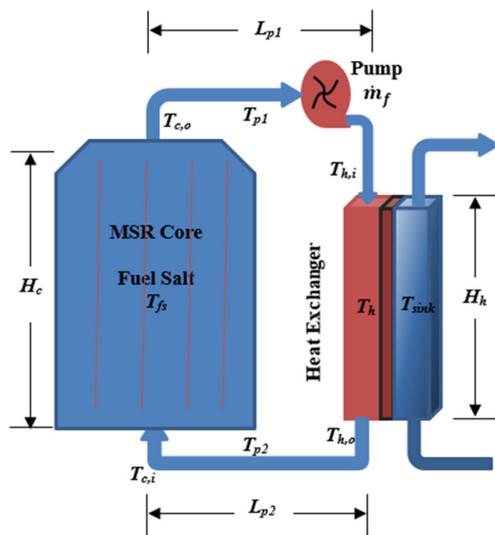


Fig. 1. Schematic representation of the MSR primary circuit.

The structure of the fast-spectrum thermal core is made by the graphite moderator elements, each of them characterized by the fast-thermal spectrum model [5, 9, 10]. The thermal spectrum liquid fuel reactor core is associated with fluoride salt and the thermal system generated a one-reserved fuel region. One circulating pump indicates the top-bottom side of the HX unit. The heat transfer mechanism is simulated with the statement of a "cooler" which operates to stabilize temperature. The heat transfer between the hot and cold sides is calculated by using the average temperature. The heat transfer coefficient is assumed to be constant [11-13].

III. MATHEMATICAL DERIVATION AND REACTOR NEUTRONS OF THE MSR

In this work, the flow effect on the liquid fuel salt and the traditional point of view of the kinetic model of the MSR core are studied and the behavior of delayed neutrons is analyzed. The acute response mechanism has been changed by the fuel density and temperature. The aggregate method has been used for fuel salt in both the reactor core and the HX unit. When more fuel salt is circulating inside the core then fission power is generated inside and outside of the core, due to the very fast delayed neutrons. The space kinetic equations derived for these regions are:

$$\frac{dp(t)}{dt} = \frac{\rho_{net}(t) - \beta_{eff}}{\Lambda} p(t) + \lambda_{eff} C(t) \quad (1)$$

$$\frac{dC(t)}{dt} = \frac{\beta_{eff}}{\Lambda} p(t) - \lambda_{eff} C(t) - \frac{1}{\tau_{core}} C(t) + \frac{e^{-\lambda_{eff}\tau_{loop}}}{\tau_{core}} C(t - \tau_{loop}) \quad (2)$$

where P and C represent the precursor concentration of the reactor power and delayed neutron precursor respectively. The terms ρ_{net} , β_{eff} , λ_{eff} , and Λ depict net reactivity and the fraction delayed neutrons emitted energy from the core. However, the various prompt neutrons per fission and the numerous delayed neutrons of precursor per fission affect the velocity of the delayed neutron decay constants of the precursor, and generation of neutron time respectively. τ shows the transit time [12]. The fuel salt passes through the core, pipe section #1, heat exchanger, and pipe section #2, and then it enters the core again as shown in Figure 1. τ_{loop} denotes the transit time outlet of the core region and it is expressed as: $\tau_{loop} = \tau_{p1} + \tau_{p2} + \tau_h$, where subscripts $p1, p2$, and h represent pipe section #1, pipe section #2, and HX respectively.

A. Neutronic Calculation Model Development

The thermodynamic characteristics of the core region are represented by the thermal power which is provided by the fuel lumped region. The average temperature of the lumped region is represented by the core fuel salt temperature. In this study, the energy balance equations are determined via an one dimension (1-D) model. The specific heat capacity is assumed to be constant and it is computed at the average temperature of the fuel. The balanced equation for the core region can be expressed as:

$$m_f c_{p,f} \frac{dT_f(t)}{dt} = p(t) - m_f c_{p,f} [T_{c,o}(t) - T_{c,i}(t)] \quad (3)$$

where m_f and \dot{m}_f show the fuel mass in the core and the fuel mass flow rate respectively, $c_{p,f}$ represents the average heat capacity of the fuel in the core, T_f , $T_{c,i}$, and $T_{c,o}$ show the core lumped fuel temperature of the moderator surface, the core inlet temperature, and the core outlet temperature [11, 14, 16].

The balanced equation for the fuel side of the HX unit is:

$$m_h c_{p,h} \frac{dT_h(t)}{dt} = m_h c_{p,h} [T_{h,i}(t) - T_{h,o}(t)] - (UA)_h [T_h(t) - T_{sink}(t)] \quad (4)$$

where m_h and \dot{m}_h represent the fuel mass in the HX and the fuel mass flow rate ($\dot{m}_h = \dot{m}_f$) respectively, $c_{p,h}$ illustrates the average heat capacity of the fuel in the HX, T_h , $T_{h,i}$, $T_{h,o}$, T_{sink}

represent the lumped fuel temperature in the HX, the inlet and the outlet temperature of the HX, and the average temperature of the heat sink. U and A_h show the overall heat transfer coefficient of the HX unit and the heat transfer area of the HX. The overall heat transfer coefficient is assumed to be constant [15, 16]. The balanced equation for pipe section #1 can be expressed as:

$$m_{p1}c_{p,p1} \frac{dT_{p1}(t)}{dt} = m_{p1}c_{p,p1} [T_{c,o}(t) - T_{h,i}(t)] - (UA)_{p1} [T_{p1}(t) - T_{\infty}] \quad (5)$$

where subscript $p1$ represents the pipe section #1, T_{p1} and T_{∞} represent the lumped fuel temperature in pipe section #1 and the ambient temperature. It is assumed that the overall heat transfer coefficient is also constant for pipe section #1.

TABLE I. MSR PARAMETERS AND RANGE OF VALUES IN THE LITERATURE

Parameters	Symbol	Unit	Range
Reactor power	P	MWth	400 – 4000
Prompt neutron generation time	Λ	S	$10^{-6} - 10^{-7}$
Initial balancing reactivity	ρ_0	Pcm	$\beta_{eff} \left[\frac{1 - e^{-\lambda_{eff}\tau_{loop}}}{\lambda_{eff}\tau_{core} + 1 - e^{-\lambda_{eff}\tau_{loop}}} \right]$
Steady-state external reactivity	ρ_{ext}	Pcm	$\rho_{ext} = -\rho_0$
Effectively delayed neutron fraction	β_{eff}	-	~ 0.00033
Effectively delayed neutron decay constant	λ_{eff}	s ⁻¹	~ 0.006
Core transit time	τ_{core}	S	$\sim m_f / \dot{m}_f$
Loop transit time	τ_{loop}	S	$\tau_h + \tau_{p1} + \tau_{p2}$
Pipe section #1 transit time	τ_{p1}	S	$\sim m_{p1} / \dot{m}_f$
Pipe section #2 transit time	τ_{p2}	S	$\sim m_{p2} / \dot{m}_f$
Heat exchanger transit time	τ_h	S	$\sim m_h / \dot{m}_f$
Fuel temperature coefficient of reactivity	α_f	K ⁻¹	$\sim 3.04 \times 10^{-5}$
Fuel mass flow rate	\dot{m}_f	kg.s ⁻¹	Depends on the power
Fuel mass in the core	m_f	Kg	$d_f @ T_f V_{core}$
Fuel mass in the HX	m_h	Kg	$d_f @ T_h V_h$
Fuel mass in the pipe section #1	m_{p1}	Kg	$d_f @ T_{p1} V_{p1}$
Fuel mass in the pipe section #2	m_{p2}	Kg	$d_f @ T_{p2} V_{p2}$
Specific heat capacity of the fuel salt	c_p	J.kg ⁻¹ K ⁻¹	$-1111 + 2.782T_f$
Heat exchanger heat transfer coefficient	$(UA)_h$	W.K ⁻¹	$P = N_{HX}(UA)_h \Delta T_{lm}^{(**)}$
Average fuel temperature (in the core)	T_f	K	~ 975
Average fuel temperature (in the HX)	T_h	K	~ 975
The average temperature of the heat sink	T_{sink}	K	750-900
Ambient temperature	T_{∞}	K	~ 3000
The density of the fuel salt	d_f	kg.m ⁻³	$49983.56 - 0.8982T_f$
The volume of the core	V_{core}	m ³	Depends on the power
The volume of the HX	V_h	m ³	Depends on the HX capacity
The volume of pipelines #1, #2	$V_{p1,p2}$	m ³	Depends on the design

B. Thermal Feedback Mechanism and State-Action Space

The net reactivity of the MSR can be expressed as:

$$\rho_{net}(t) = \rho_0 + \rho_{ext}(t) - \alpha_f [T_f(t) - T_{f0}] \quad (6)$$

where ρ_0 illustrates the initial balancing reactivity, ρ_{ext} represents the external reactivity which can be added with bubbles, and α_f represents the core fuel salt temperature which is changed by the graphite moderator and the reactivity coefficient. Similarly, the inlet and outlet flow rates are the same, and T_{f0} represents the steady-state fuel temperature in the core.

There is only one way to write in linear form the nonlinear differential equations considering the small perturbation around the steady-state operating point. These equations can be transformed into state-space form [16]. Around the steady-state power (P_0 and C_0), (1) and (2) can be written as:

$$\delta p(t) \frac{p_0}{dt} \delta \rho(t) + \frac{\delta \rho(t) \delta p(t) \beta_{eff}}{\Lambda} \delta p(t) + \lambda_{eff} \delta C(t) \quad (7)$$

$$\delta C(t) \frac{\beta_{eff}}{\Lambda} \delta \rho(t) - (\lambda_{eff} + \frac{1}{\tau_{core}}) \delta C(t) + \frac{e^{-\lambda_{eff} \tau_{loop}}}{\tau_{core}} \delta C(t - \tau_{loop}) \quad (8)$$

where $P(t) = P_0 + \delta P(t)$, $C(t) = C_0 + \delta C(t)$, and $\rho(t) = \delta \rho(t) = \alpha_f [T_f(t) - T_{f0}]$ due to the nature of the reactivity. The reactor has to be critical at steady-state operation, so the reference reactivity ($\rho_0 + \rho_{ext}$) will be zero at this point.

The nonlinear term, $\delta \rho(t) \delta P(t)$ still exists in (7). It can be inferred that the point kinetics equations are linear only for constant reactivity conditions. To perform linear analysis, this second-order term is neglected in this study. Lumped fuel temperatures for the core and the HX regions are calculated as:

$$T_f(t) = \frac{1}{2} [T_{c,o}(t) + T_{c,i}(t)] \quad (9)$$

$$T_h(t) = \frac{1}{2} [T_{c,o}(t) - T_{c,i}(t)] \quad (10)$$

By taking into consideration the transit times at pipe sections #1 and #2, the following substitutions:

$$T_{h,i}(t) \approx T_{c,o}(t - \tau_{p1}) \quad (11)$$

$$T_{c,i}(t) \approx T_{h,o}(t - \tau_{p2}) \quad (12)$$

can be made to state the relations between the temperatures. The temperature decrease of the fuel salt in the pipeline is neglected due to the insulation. By using (9)-(12), equations (3)-(5) can be written in the form of (13)-(15):

$$m_f c_{p,f} \dot{T}_f(t) = p_o + \delta p(t) - \dot{m}_f c_{p,h} [T_{c,o}(t) + (\tau_{p1} - \tau_{p2}) - 2T_h(t - \tau_{p2})] \quad (13)$$

$$m_h c_{p,h} \dot{T}_h(t) = 2\dot{m}_h c_{p,h} [T_{c,o}(t - \tau_{p1}) - T_h(t)] - (UA)_h [T_h(t) - T_{sink}] \quad (14)$$

$$m_{p1} c_{p,p1} \dot{T}_{c,o}(t) = \dot{m}_{p1} c_{p,p1} [T_{c,o}(t) - T_{c,o}(t - \tau_{p1})] - (UA)_{p1} [T_{c,o}(t) + T_{c,o}(t - \tau_{p1}) - 2T_\infty]/2 \quad (15)$$

where:

$$T_{p1}(t) = 1/2 [T_{c,o}(t) + T_{c,o}(t - \tau_{p1})].$$

It is assumed that $\dot{T}_{p1}(t) \approx \dot{T}_{c,o}(t)$ and that T_{sink} is constant. Finally, (7), (8), (13)-(15) take the state space form as follows:

$$\dot{\delta x} = A_1 x + A_2 x * \delta(t - \tau_{loop}) + A_3 x * \delta(t - \tau_{p1}) + A_4 x * \delta(t - \tau_{p1} - \tau_{p2}) + B \mu \quad (16)$$

where:

$$x = [\delta p(t) \delta C(t) T_f(t) T_h(t) T_{c,o}(t)]^T \quad (17)$$

$A_1 =$

$$\begin{bmatrix} -\frac{\beta_{eff}}{\Lambda} & \lambda_{eff} & \frac{\alpha_f p_o}{\Lambda} & 0 & 0 \\ \frac{\beta_{eff}}{\Lambda} & -\lambda_{eff} - \frac{1}{\tau_{core}} & 0 & 0 & 0 \\ \frac{1}{m_f c_{p,f}} & 0 & 0 & \frac{2\dot{m}_f}{m_f} & -\frac{\dot{m}_f}{m_f} \\ 0 & 0 & 0 & -\frac{2\dot{m}_f}{m_f} - \frac{(AU)_h}{m_f c_{p,h}} & 0 \\ 0 & 0 & 0 & 0 & -\frac{\dot{m}_f (UA)_{p1}}{m_f 2m_f c_{p,p1}} \end{bmatrix} \quad (18)$$

$$A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -e^{\lambda_{eff} \tau_{loop}} / \tau_{core} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (19)$$

$$A_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2\dot{m}_f}{m_f} & 0 \\ 0 & 0 & 0 & -\frac{\dot{m}_f}{m_f} - \frac{(UA)_{p1}}{2m_f c_{p,p1}} & 0 \end{bmatrix} \quad (20)$$

$$A_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{\dot{m}_f}{m_f} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (21)$$

$$\beta = \begin{bmatrix} -\frac{\alpha_f p_o}{\Lambda} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{m_f c_{p,f}} & 0 & 0 \\ 0 & 0 & 0 & \frac{(UA)_h}{m_f c_{p,h}} & 0 \\ 0 & 0 & 0 & 0 & \frac{(UA)_{p1}}{m_f c_{p,p1}} \end{bmatrix}^T \quad (22)$$

$$u = [T_{f0} \ P_0 \ T_{sink} \ T_\infty]^T \quad (23)$$

It is important to note that $\dot{m}_f = \dot{m}_h = \dot{m}_{p1} = \dot{m}_{p2}$ due to the closed-loop operation (the variation of the expansion tank level is neglected). All physical properties were taken as constant and were calculated at the average temperature of the lumped regions. Calculated transit time can be divided into to the length of system (or height) by the velocity of the fluid.

IV. RESULTS AND DISCUSSION

In this section, the results of the numerical analysis for the proposed system are presented. The primary loop of MSR is simulated under the steady state in Matlab/Simulink. When the fuel pump rotation decreases in the start of transient, the primary flow will reduce within a few seconds. Furthermore, the reactor core fraction of the delayed neutron precursors will increase because they have the tendency to improve the reactivity and the reactivity loss due to the fuel lumped temperature. Therefore, heat is produced by fuel lumped temperature in the reactor core, so that the temperature difference between core inlet and core outlet is increased in the dynamic model as shown in Figure 2.

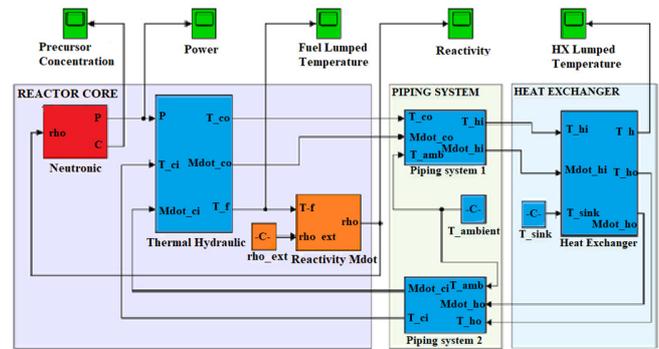


Fig. 2. Simulink model.

The core dynamic model is associated with the neutronic and thermal-hydraulic coupled system. Coupling is obtained by the reactivity feedback mechanism. In this device, the input components are outer reactivity, ambient temperature, and heat sink temperature. The output values are the reactor power and the core outlet temperature. The simulation parameters are given in Table II.

Several different MSR concepts have been presented. Therefore, there is a wide variety of values for both input and output parameters. The module "1" presents the MSR (fast or thermal) parameters and their range value can be used for the developed model and tool.

TABLE II. MSR PARAMETERS USED IN THE SIMULATION

Parameter	Value
Reactor power	4000MWth
Average core temperature	700°C
Core transit time	2s
Loop transit time	2s
Transit time for each pipe	0.1s
Effectively delayed neutron fraction	0.0033
Effectively delayed neutron decay constant	0.0611s ⁻¹
Prompt neutron generation time	0.95e-6s
Fuel temperature coefficient of reactivity	3.4e-5K ⁻¹
Coolant bulk temperature	570°C
Ambient temperature	40°C

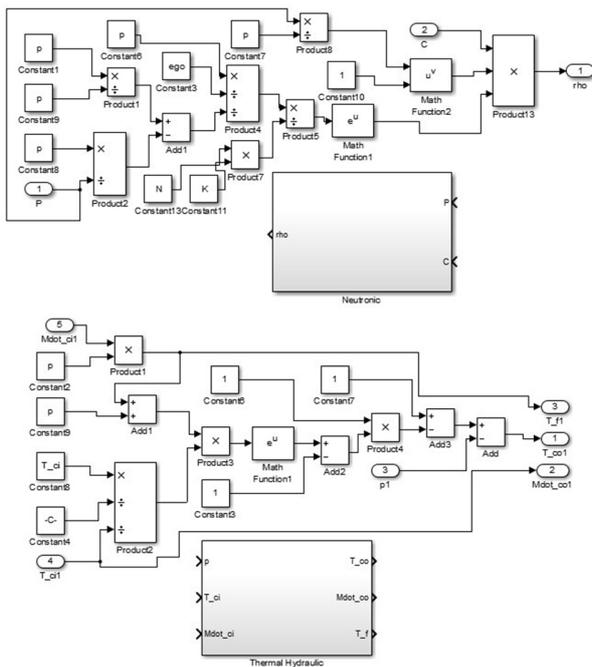


Fig. 3. The sub model of neutronic and thermal hydraulic.

Three sub models of the Simulink model are used to describe the Simulink model shown in Figure 2 representing three individual parts of the Simulink model. Figure 3 shows the neutronic and thermal-hydraulic sub model. It is used to diagnose faults from the core and give information about the fuel salt temperature distribution from the core. Figure 4 shows the reactivity model and the piping system #1 of the MSR. Therefore it shows the lumped fuel temperature and it represents all constant transfer temperatures in the piping system, while fuel salt passes through piping system #1. Figure 5 represents the piping system #2 and the HX of the MSR. It shows the heat capacity inlet and outlet temperature of the HX. Figure 6 shows the thermal reactor which reached the 3.02GW equilibrium by using the parameters illustrated in Table I. The reactor reached the steady-state equation level at 100s. The fuel temperature changes between the HX region and the core region, because temperature is reduced due to the loss of the pipeline of the heat by 1°K. The fuel temperature increased by 100°K in the reactor transverse core. These features can be used in the perturbation system.

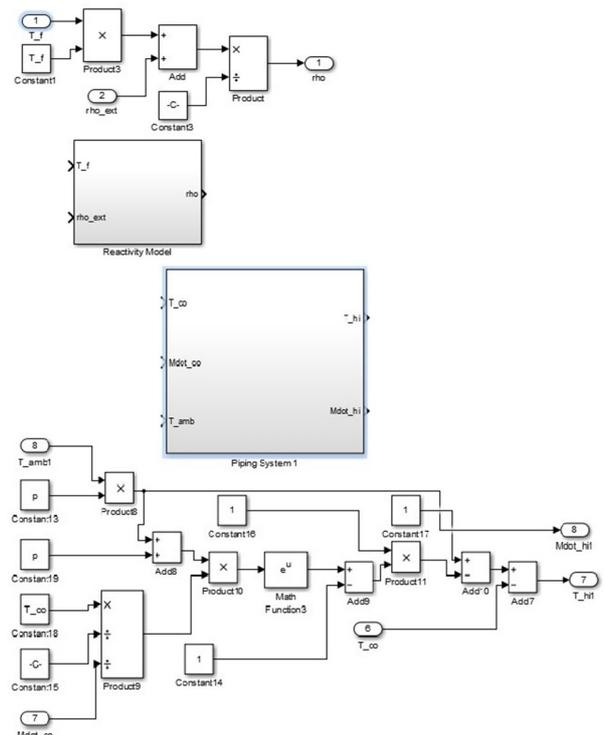


Fig. 4. The sub model of the reactivity model and piping system #1.

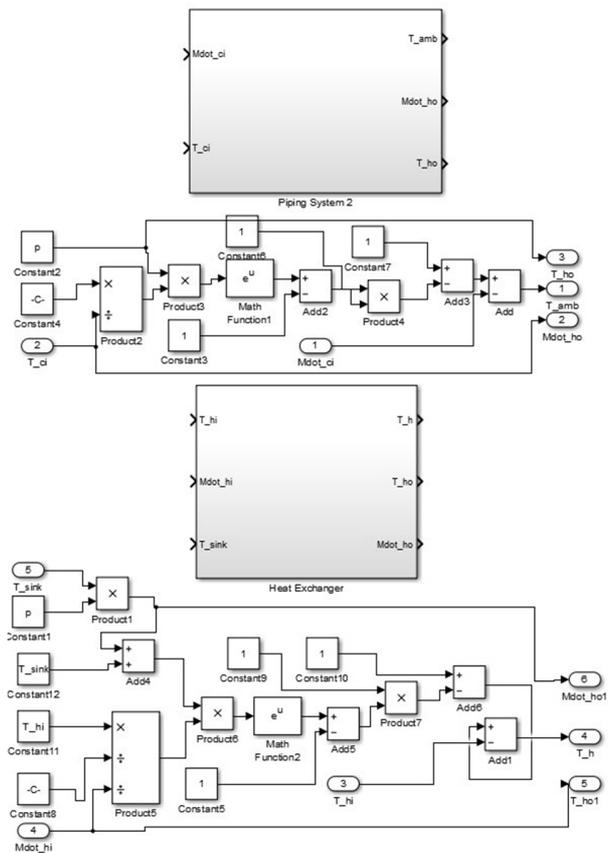


Fig. 5. The sub model of the piping system #2 and the HX.

The proposed model's derived state-space equations can be used for any specific fast MSR. The equations are written for generic parameters. The fuel temperature variance between the core region and the HX region is very small, due to the 1°K temperature decrease because of the losses of heat at the pipelines. In this type of reactor, temperature of 100°K is developed transversely in the core. The MSRs can be simulated with the proposed model. This feature can also be applied into perturbation cases. The components of the reactor system are adopted from the literature.

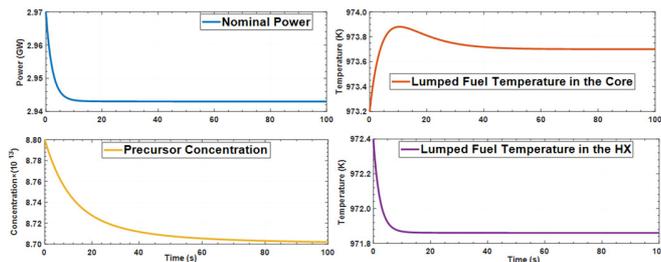


Fig. 6. Reactor parameters at steady-state operation.

V. CONCLUSION

Many dynamic methods considering the thermal MSR have been presented in previous works. In this paper, we used two or more lumped regions at the reactor core for the thermal-hydraulic modeling. It was assumed that the outlet core temperature is the same as the average temperature of the higher lumped region. The outlet core temperature is measured inaccurately due to the high temperature at the core reactor. For all types of MSR, measuring the fast spectrum or thermal fuel types is very important. There are no distinct dynamic methods for the fuel position in the core. The methods regarding other positions of the thermal reactor have significant differences that are caused by the moderator region. Considering the feedback effects, the moderator temperature of the graphite region has been modeled in the thermal MSRs, even though there is no concept of the moderator region in the fast MSR.

A new kinetic method has been proposed for fast MSRs in this paper. There is only one lumped region that has been assumed for the reactor core, which has only fuel salt. Therefore, energy has been balanced for the outlet core temperature that is described using pipeline temperature. The steady-state operation of MSR, the nonlinear kinetic equation, the Matlab/Simulink system, and the state-space model have been presented. The new model was tested in simulations and the conclusion is that the conceptual reactor can be used in real applications.

ACKNOWLEDGMENT

The authors would like to thank the Natural Science Foundation of Heilongjiang Province, China (Grant No. A2016002), the Foundation of Science and Technology on Reactor System Design Technology Laboratory (HT-KFKT-14-2017003), the technical support project for Suzhou Nuclear Power Research Institute (SNPI) (No.029-GN-B-2018- C45-P.0.99-000 03), and the Research Institute of Nuclear Power

Operation (No. RIN180149-SCCG) for funding this research project.

REFERENCES

- [1] C. Yu *et al.*, "Th-U breeding performance in a Channel-type molten salt Fast reactor with different starting fuels," *Annals of Nuclear Energy*, vol. 122, pp. 91–100, Dec. 2018, <https://doi.org/10.1016/j.anucene.2018.08.003>.
- [2] V. Singh, "Study of the Dynamic Behavior of Molten-Salt Reactors," M. S. thesis, University of Tennessee, Knoxville, TN, USA, 2019.
- [3] D. Suescún-Díaz and G. Espinosa-Paredes, "On the numerical solution of the point reactor kinetics equations," *Nuclear Engineering and Technology*, vol. 52, no. 6, pp. 1340–1346, Jun. 2020, <https://doi.org/10.1016/j.net.2019.11.034>.
- [4] V. V. Ignatiev *et al.*, "Molten-salt reactors: new possibilities, problems and solutions," *Atomic Energy*, vol. 112, no. 3, pp. 157–165, Jul. 2012, <https://doi.org/10.1007/s10512-012-9537-2>.
- [5] V. Singh, M. R. Lish, A. M. Wheeler, O. Chvála, and B. R. Upadhyaya, "Dynamic Modeling and Performance Analysis of a Two-Fluid Molten-Salt Breeder Reactor System," *Nuclear Technology*, vol. 202, no. 1, pp. 15–38, Apr. 2018, <https://doi.org/10.1080/00295450.2017.1416879>.
- [6] A. Asif *et al.*, "Advancement of Integral Fast Reactor," *Journal of Applied and Emerging Sciences*, vol. 10, no. 2, pp. 149–154, Dec. 2020, <https://doi.org/10.36785/jaes.v10i2.453>.
- [7] J. Stewart, "Gen-Foam Multiphysics Model Development for Molten Salt Reactors," Ph.D. dissertation, University of Nevada, Las Vegas, NV, USA, 2020.
- [8] A. M. Leandro, F. Heidet, R. Hu, and N. R. Brown, "Thermal hydraulic model of the molten salt reactor experiment with the NEAMS system analysis module," *Annals of Nuclear Energy*, vol. 126, pp. 59–67, Apr. 2019, <https://doi.org/10.1016/j.anucene.2018.10.060>.
- [9] B. Mignacca and G. Locatelli, "Economics and finance of Molten Salt Reactors," *Progress in Nuclear Energy*, vol. 129, p. 103503, Nov. 2020, <https://doi.org/10.1016/j.pnucene.2020.103503>.
- [10] Y. Kassem, H. Camur, and O. a. M. Abughinda, "Solar Energy Potential and Feasibility Study of a 10MW Grid-connected Solar Plant in Libya," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5358–5366, Aug. 2020, <https://doi.org/10.48084/etasr.3607>.
- [11] C. Shi, M. Cheng, and G. Liu, "Development and application of a system analysis code for liquid fueled molten salt reactors based on RELAP5 code," *Nuclear Engineering and Design*, vol. 305, pp. 378–388, Aug. 2016, <https://doi.org/10.1016/j.nucengdes.2016.05.034>.
- [12] Z. Zou, C. Shen, X. Zhang, S. Wang, and H. Chen, "3D thermal hydraulic characteristics analysis of pool-type upper plenum for lead-cooled fast reactor with multi-scale coupling program," *Nuclear Engineering and Design*, vol. 370, Dec. 2020, Art. no. 110892, <https://doi.org/10.1016/j.nucengdes.2020.110892>.
- [13] A. Asif *et al.*, "Recent Advances in Nuclear Power Plant for Fault Detection and Diagnosis - A Review," *Journal of Critical Reviews*, vol. 7, Dec. 2020, Art. no. 10, <https://doi.org/10.31838/jcr.07.18.535>.
- [14] P. Zhao, Z. Liu, T. Yu, J. Xie, Z. Chen, and C. Shen, "Code development on steady-state thermal-hydraulic for small modular natural circulation lead-based fast reactor," *Nuclear Engineering and Technology*, vol. 52, no. 12, pp. 2789–2802, Dec. 2020, <https://doi.org/10.1016/j.net.2020.05.023>.
- [15] B. Doğan and T. Ölmez, "A novel state space representation for the solution of 2D-HP protein folding problem using reinforcement learning methods," *Applied Soft Computing*, vol. 26, pp. 213–223, Jan. 2015, <https://doi.org/10.1016/j.asoc.2014.09.047>.
- [16] R. Yoshioka, Y. Shimazu, and K. Mitachi, "Guidelines for MSR Accident Analysis," presented at the Thorium Energy Conference, Shanghai, China, 2013.

A Microservice-Based System for Industrial Internet of Things in Fog-Cloud Assisted Network

Fida Hussain Khoso

Department of Basic Science
Dawood University of Engineering & Technology
Karachi, Pakistan
fidahussain.khoso@duet.edu.pk

Abdullah Lakhani

Computer Science and IT Department
Benazir Bhutto Shaheed University
Karachi, Pakistan
abdullahrazalakhani@gmail.com

Aijaz Ahmed Arain

Department of Computer Science
Quaid-e-Awam University of Engineering, Science and
Technology
Nawabshah, Pakistan
aijaz@quest.edu.pk

M. Ali Soomro

Department of Computer Systems Engineering
Quaid-e-Awam University of Engineering, Science and
Technology
Nawabshah, Pakistan
kalakot2@gmail.com

Shah Zaman Nizamani

Department of Information Technology
Quaid-e-Awam University of Engineering, Science and
Technology
Nawabshah, Pakistan
hahzaman@quest.edu.pk

Kelash Kanwar

Department of Electronic Engineering
Quaid-e-Awam University of Engineering Science and
Technology
Nawabshah, Pakistan
kelashkanwar@quest.edu.pk

Abstract-Nowadays, the usage of the Industrial Internet of Things (IIoT) in practical applications has increased. The primary utilization is a fog cloud network, which offers different services, such as network and remote edges, at different places. Existing studies implemented the Service-Oriented Architecture (SOA) based on the fog-cloud network to run IIoT applications, such as e-healthcare, e-agriculture, renewable energy, etc. However, due to the applications' monolithic property, issues like failures, security, and cost factors occur, e.g. the failure of one service in SOA affects monolithic applications' performance in the system. With this motivation, this study suggests a microservice-based system to deal with the cost, security, and failure risks of IIoT applications in the fog-cloud system. The study improves the existing SOA systems for e-healthcare, e-agriculture, and renewable energy and minimizes the applications' overall cost. The performance evaluation shows that the devised systems outperform the existing SOA system in terms of failure, cost, and the deadline for all applications.

Keywords-IIoT; security; e-healthcare; SOA; microservice

I. INTRODUCTION

Real world applications of the digital transformation of technologies such as the Industrial Internet of Things (IIoT) possess many advantages [1]. Such applications belong to the e-agriculture, e-healthcare, e-transport, e-servicing, and augmented fields [2]. Cloud computing is an incipient archetype which offers services based on Service-Oriented

Architecture (SOA) [3]. Generally, SOA comprises of different services which are monolithically managed by one centralized component [4]. SOA architecture capabilities extend to the fog computing, edge computing, and cloudlet data center to efficiently execute time and latency-sensitive applications [5]. However, fault tolerance, security, and cost of services in SOA architecture depend directly upon the centralized controller. For instance, if any specific service fails or is attacked, then the impact and cost will affect the entire system [6]. The microservice-based system is an innovative solution which offers autonomous service to run IIoT applications. Each microservice executes its application independently, and service failure has no impact on the other services during the process. All microservices are communicating with each other via an REST API. The generalization is a standard interface that exploits the REST API to create cross-platform for all microservices to run one application. Each application requires many microservices to run its tasks. For instance in e-health applications, blood pressure, heartbeat, and ECG tasks each require one microservice for execution [7]. However, a microservice-based system for IIoT has not been developed yet.

The current study extends the existing IIoT work [7] with new suggestions that improve the handling of cost, security, and failure issues of the system in fog-cloud networks. The current work is a survey paper. However, it proposes a solution

Corresponding author: Abdullah Lakhani

for those systems to improve applications' cost and security mechanisms. The paper makes the following contributions to the existing systems: It discusses the general IIoT process based on the microservices system. It also discusses industry automation based on microservices with cost and security constraints and e-healthcare services based on microservices connected with the distributed fog-cloud network. Generally, all systems exploit microservice systems in a fog-cloud network with deadline, cost, security, and fault-tolerant constraints. The study solves the offloading and scheduling problems for the considered systems.

II. RELATED WORK

Authors in [1] suggested the SOA aware framework to solve the IIoT mechanism issue of e-agriculture applications in a fog cloud network. The goal was to minimize the end to end latency of the applications. However, they did not consider the security and fault-tolerance of the applications. Fault-tolerant aware SOA is discussed in [2]. The goal was to minimize the failure risks and exploit the primary backup method to handle any SOA architecture failure. However, it consumed much more resources and had high cost during the process. The security-aware framework [3] handles application-level security in SOA architecture. Applications can offload tasks with secure data in the distributed network, however the authors did not consider failures and network and cloud security mechanisms. A hybrid security-aware system based on SOA architecture was suggested in [4]. The goal was to protect data from denial of service and brute force attacks to IIoT applications. The security and failure aware centralized SOA systems for monolithic and coarse-grained applications were studied in [5]. The objective was to improve resource utilization and minimize the application cost. However, the centralized control could fail anytime, and this would affect the entire system in the network. Cost-efficient offloading for healthcare applications based on SOA architecture by considering security and deadlines was investigated in [6]. All considered applications are monolithic and fine-grained and run their tasks distinctly under their deadline. SOA architecture offers services based on an RPC model where users are treated as thin clients and servers as thick clients. However, still the centralized failure of any service affects the entire system. An e-agriculture and a body area sensor aware network based on SOA architecture was suggested and studied in [7, 8]. The goal was to optimize the service process with minimum end to end latency of the applications while offering 24/7 services with the robot and faultless services. However, there are many issues in the system. The SOA architecture selects the services from a published pool for consumer and producer and a hacker can publish services and quickly become part of SOA monolithic architecture. The container and function aware IIoT solutions were suggested in [9, 10] to handle the offloading and scheduling problem's transient failure and security issues. These studies achieved better results as compared to SOA architecture. However, the proposed systems have not matured yet for all kinds of IIoT applications.

All reviewed studies proposed systems for IIoT based on the SOA and container function mechanisms. Many of them considered different constraints such as security, failure, cost,

and latency. However, failure in service due to hacking or resource-imbalance will impact the overall system performance. The suggested process was very costly and consumed a lot of resources. With the motivations mentioned earlier, this paper suggests microservice aware different IIoT system to minimize failure risk, cost, and security and meet the deadline of applications.

III. THE PROPOSED SOLUTION

The study proposes a different system based on microservice architecture instead of SOA in order to improve cost, failure, and security risk of the applications.

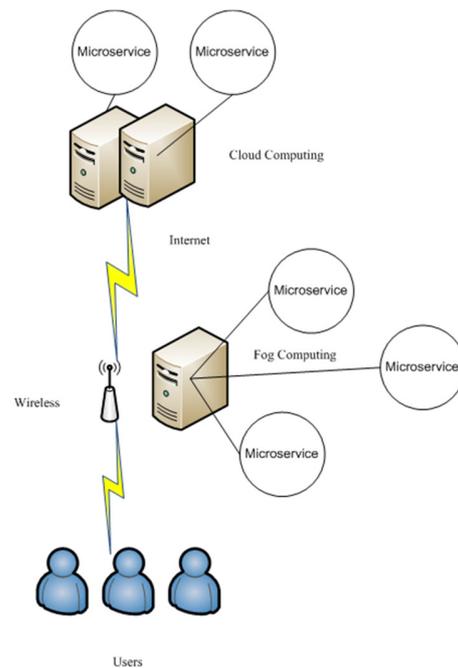


Fig. 1. Microservice based fog-cloud system.

Cloud computing is a distributed service provider which offers services at different rates to run any application on the network. Fog computing is an extension of cloud computing which brings cloud resources at the end of a wireless network. The main advantage of fog computing is the reduction of the end to end latency of an application. This study complements the existing architecture [15-24] which is based on SOA. The work creates a system which consists of three layers such as user-end, fog end and cloud end as shown in Figure 1. The user-end is connected with different computer systems that offload their data to the fog server via communication channels for further processing. The fog servers are located at the edge of the network and have short distance and minimum end to end latency during submission or offloading. The basic mechanism design is based on the remote process call which is a popular method allowing the communication channels between users and serves via stub and skeleton objects. The renewable IIoT aware network performs different tasks such as monitoring temperature, humidity, weather, etc. sensors which require microservices to run and save their data in the cloud as shown in Figure 1. Generally, these are real-time tasks,

working 24/7 and generating big data to the fog system. Due to the time sensitive and security sensitive data, the initial process will be done on the fog servers. After that, for persistent saving, the fog node offloads all data to the cloud for further big data analytics. The high volume of data with various types and speed will be easily managed by the rich resource cloud computing. A body-area network is connected with different sensors for blood pressure, heartbeat, ECG, and online monitoring tasks at the user level. All tasks offload their data to the fog node which applies computation by using a unique microservice to run each task under its deadline. Big data are generated by the tasks to be offloaded to the remote cloud for further analyzing.

The main advantage of the system is that the failure of one microservice will not affect the entire system because they all are working isolated to each other. Another advantage is that microservices have execution charges. If one microservice fails it will not charge for anything until and unless a task is finished without completing its execution. The fault-isolation, security, and autonomous execution make this system more efficient and intelligent in comparison with all the existing SOA architectures. This system selects microservices from the published pool which is publically available with different charges and time-slots. However, the considered system is smart, if one microservice fails, the system will select another without any violation of task deadlines.

IV. PERFORMANCE EVALUATION

The performance evaluation was conducted by implementing the Edgex Foundry microservice and SOA architecture to run all IIoT applications. We conducted the experiment in different cases as shown in Tables I-III. Table I summarizes the security methods at the user level only and the execution cost of the applications. Table II summarizes the study of the failure mechanisms in the case only one system failure and determines the execution cost of the applications. Table III shows the results of the comparison based on deadline methods at the application level only and the execution cost of the applications.

The overall service cost of security of the monolithic applications based on SOA architecture became high during the process due to the centralized handling of all components. However, microservices had low cost because they are run inside a container and charged for their execution instead of the resource provisioning of SOA which charged for all applications. The primary backup and check pointing always required a lot of resources in SOA architecture. If one node fails, it may transfer a task from a service to another node. This improves the overall performance of the applications, but with high cost. The bottom line is that users pay for the failure costs for their applications.

Generally, fog-cloud applications co-operatively work together to achieve the goals of the applications. However, due to the monolithic architecture, if a centralized component fails or is attacked, it needs a lot of time to restart. In this way, the deadline of and the overall performance of the applications will be degraded. The deadline of applications directly depends upon execution time and the availability of resources. The cost

of the applications by combining both failure and security becomes high when IIoT uses the SOA architecture (Table IV). The monolithic applications need a lot of resources for their execution. However, the isolated microservice-based system works efficiently for all IIoT in the fog-cloud network.

TABLE I. SECURITY MECHANISMS

Reference	Security	Cost	Application
[1, 2]	RSA	10\$	Monolithic
[3, 4]	RSA	20\$	Monolithic
[5, 6]	MD5	17\$	Monolithic
[7, 8]	SHA	23\$	Monolithic
[9, 10]	CRC32	27\$	Coarse-grained
Proposed	RSA	2\$	Fine-grained

TABLE II. FAILURE MECHANISMS

Reference	Fail	Cost	Application
[1, 2]	Primary backup	10\$	Monolithic
[3, 4]	Primary backup	20\$	Monolithic
[5, 6]	Check pointing	17\$	Monolithic
[7, 8]	Check pointing	23\$	Monolithic
[9, 10]	Check pointing	27\$	Coarse-grained
Proposed	Isolated	2\$	Fine-grained

TABLE III. DEADLINE

Reference	Deadline	Cost	Application
[1, 2]	Missed	40\$	Monolithic
[3, 4]	Missed	70\$	Monolithic
[5, 6]	Average	88\$	Monolithic
[7, 8]	Reached	75\$	Monolithic
[9, 10]	Reached	99\$	Coarse-grained
Proposed	Reached	23\$	Fine-grained

TABLE IV. OVERALL SOA AND MICROSERVICE SYSTEM PERFORMANCE

Ref.	Security	Cost	Fail	Applications
[1, 2]	No	100\$	No	E-agriculture
[3, 4]	Yes	200\$	No	E-agriculture
[5, 6]	No	150\$	Yes	E-health
[7, 8]	Yes	300\$	Yes	E-health
[9, 10]	Yes	500\$	Yes	Industry automation
Proposed	Yes	30, 50, 70\$	Yes	ALL

V. CONCLUSION

This study suggests the microservice based system for IIoT applications instead of the SOA architecture in order to minimize cost, failure effects, and risk. Simulation results show that the proposed microservice based system can enhance the performance of the applications in comparison with the monolithic SOA architecture. In future work, the Internet of Vehicle Things and mobility aware microservices in distributed fog-cloud networks will be considered. The study will design a security system based on blockchain-enable network in order to avoid any kind of attack on the system.

REFERENCES

[1] S. Latif, Z. Idrees, J. Ahmad, L. Zheng, and Z. Zou, "A blockchain-based architecture for secure and trustworthy operations in the industrial Internet of Things," *Journal of Industrial Information Integration*, vol. 21, Mar. 2021, Art. no. 100190, <https://doi.org/10.1016/j.jii.2020.100190>.

- [2] Q. Hao, S. Nazir, X. Gao, L. Ma, and M. Ilyas, "A Review on Multicriteria Decision Support System and Industrial Internet of Things for Source Code Transformation," *Scientific Programming*, vol. 2021, Jan. 2021, Art. no. e6661272, <https://doi.org/10.1155/2021/6661272>.
- [3] S. Kunal, A. Saha, and R. Amin, "An overview of cloud-fog computing: Architectures, applications with security challenges," *Security and Privacy*, vol. 2, no. 4, 2019, Art. no. e72, <https://doi.org/10.1002/spy2.72>.
- [4] A. Bamhdi, "Requirements capture and comparative analysis of open source versus proprietary service oriented architecture," *Computer Standards & Interfaces*, vol. 74, Feb. 2021, Art. no. 103468, <https://doi.org/10.1016/j.csi.2020.103468>.
- [5] L. Li, H. Huang, X. Zou, F. Zhao, G. Li, and Z. Liu, "An energy-efficient service-oriented energy supplying system and control for multi-machine in the production line," *Applied Energy*, vol. 286, Mar. 2021, Art. no. 116483, <https://doi.org/10.1016/j.apenergy.2021.116483>.
- [6] A. Lakhan and X. Li, "Transient fault aware application partitioning computational offloading algorithm in microservices based mobile cloudlet networks," *Computing*, vol. 102, no. 1, pp. 105–139, Jan. 2020, <https://doi.org/10.1007/s00607-019-00733-4>.
- [7] A. Lakhan and X. Li, "Mobility and Fault Aware Adaptive Task Offloading in Heterogeneous Mobile Cloud Environments," *EAI Endorsed Transactions on Mobile Communications and Applications*, vol. 5, no. 16, Jan. 2019, Art. no. e4, <https://doi.org/10.4108/eaic.3-9-2019.159947>.
- [8] D. K. Sajjani, A. R. Mahesar, A. Lakhan, and I. A. Jamali, "Latency Aware and Service Delay with Task Scheduling in Mobile Edge Computing," *Communications and Network*, vol. 10, no. 4, pp. 127–141, Oct. 2018, <https://doi.org/10.4236/cn.2018.104011>.
- [9] Q.-u.-A. Mastoi, A. Lakhan, F. A. Khan, and Q. H. Abbasi, "Dynamic Content and Failure Aware Task Offloading in Heterogeneous Mobile Cloud Networks," in *2019 International Conference on Advances in the Emerging Computing Technologies (AECT)*, Al Madinah Al Munawwarah, Saudi Arabia, Feb. 2020, pp. 1–6, <https://doi.org/10.1109/AECT47998.2020.9194161>.
- [10] S. F. Issawi, A. A. Halees, and M. Radi, "An Efficient Adaptive Load Balancing Algorithm for Cloud Computing Under Bursty Workloads," *Engineering, Technology & Applied Science Research*, vol. 5, no. 3, pp. 795–800, Jun. 2015, <https://doi.org/10.48084/etasr.554>.
- [11] J. Uma, V. Ramasamy, and P. Vivekanandan, "Load Balancing Algorithms in Cloud Computing Environment - A Methodical Comparison," *International Journal of Engineering Research*, vol. 3, no. 2, pp. 79–82, Feb. 2014.
- [12] A. N. Saeed, "A Machine Learning based Approach for Segmenting Retinal Nerve Images using Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5986–5991, Aug. 2020, <https://doi.org/10.48084/etasr.3666>.
- [13] Y. L. Ng, X. Jiang, Y. Zhang, S. B. Shin, and R. Ning, "Automated Activity Recognition with Gait Positions Using Machine Learning Algorithms," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4554–4560, Aug. 2019, <https://doi.org/10.48084/etasr.2952>.
- [14] L. Bittencourt *et al.*, "The Internet of Things, Fog and Cloud continuum: Integration and challenges," *Internet of Things*, vol. 3–4, pp. 134–155, Oct. 2018, <https://doi.org/10.1016/j.iot.2018.09.005>.
- [15] M. Taneja, N. Jalodia, J. Byabazaire, A. Davy, and C. Olariu, "SmartHerd management: A microservices-based fog computing-assisted IoT platform towards data-driven smart dairy farming," *Software: Practice and Experience*, vol. 49, no. 7, pp. 1055–1078, 2019, <https://doi.org/10.1002/spe.2704>.
- [16] C. Puliafito, E. Mingozzi, F. Longo, A. Puliafito, and O. Rana, "Fog Computing for the Internet of Things: A Survey," *ACM Transactions on Internet Technology*, vol. 19, no. 2, Apr. 2019, Art. no. 18, <https://doi.org/10.1145/3301443>.
- [17] H. Chegini, R. K. Naha, A. Mahanti, and P. Thulasiraman, "Process Automation in an IoT-Fog-Cloud Ecosystem: A Survey and Taxonomy," *IoT*, vol. 2, no. 1, pp. 92–118, Mar. 2021, <https://doi.org/10.3390/iot2010006>.
- [18] A. Kallel, M. Rekek, and M. Khemakhem, "IoT-fog-cloud based architecture for smart systems: Prototypes of autism and COVID-19 monitoring systems," *Software: Practice and Experience*, vol. 51, no. 1, pp. 91–116, 2021, <https://doi.org/10.1002/spe.2924>.
- [19] S. Guo, K. Wang, G. Pau, and A. Rayes, "Edge Intelligence for the Industrial Internet of Things," *IEEE Network*, vol. 33, no. 5, pp. 8–10, Sep. 2019, <https://doi.org/10.1109/MNET.2019.8863719>.
- [20] K. Janjua, M. A. Shah, A. Almogren, H. A. Khattak, C. Maple, and I. U. Din, "Proactive Forensics in IoT: Privacy-Aware Log-Preservation Architecture in Fog-Enabled-Cloud Using Holochain and Containerization Technologies," *Electronics*, vol. 9, no. 7, Jul. 2020, Art. no. 1172, <https://doi.org/10.3390/electronics9071172>.
- [21] R. K. Naha *et al.*, "Fog Computing: Survey of Trends, Architectures, Requirements, and Research Directions," *IEEE Access*, vol. 6, pp. 47980–48009, 2018, <https://doi.org/10.1109/ACCESS.2018.2866491>.
- [22] B. K. Mohanta, D. Jena, U. Satapathy, and S. Patnaik, "Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology," *Internet of Things*, vol. 11, p. 100227, Sep. 2020, <https://doi.org/10.1016/j.iot.2020.100227>.
- [23] L. Lu, L. Xu, B. Xu, G. Li, and H. Cai, "Fog Computing Approach for Music Cognition System Based on Machine Learning Algorithm," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 1142–1151, Dec. 2018, <https://doi.org/10.1109/TCSS.2018.2871694>.
- [24] F. Yang, Y. Zhang, B. Lv, and W. Dai, "A Task-Oriented Automatic Microservice Deployment Method For Industrial Edge Applications," in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, Oct. 2020, pp. 2149–2154, <https://doi.org/10.1109/IECON43393.2020.9254447>.

Application of the Dependency Graph Method in the Analysis of Automatic Transmission Gearboxes

Adam Deptuła

Department of Management and Production Engineering

Faculty of Production Engineering and Logistics

Opole University of Technology

Opole, Poland

a.deptula@po.edu.pl

Abstract-The current article discusses the use of the dependency graph method to design and analyze automatic transmissions. Different goals may be served by modeling an automatic transmission using graphs. The most important of them are: determining the gear ratios for gears and the analysis of speed and acceleration of the rotating elements. Game tree-structure methods can be used to analyze the functional schemes of the selected gears. This paper describes the method of generating a system of kinematics equations for signal dependency graphs. This allows the generalization and extension of the algorithmic approach, and also further analyses and syntheses, such as checking the isomorphism of the proposed solutions and determining the validity of the structure and the operational parameters of the analyzed gears.

Keywords-computational model; computer aided design; optimization; graph theory; mechanical engineering

I. INTRODUCTION

A machine as a system is a set of objects (blocks, elements), each of which is described by an appropriate mathematical model with indications of all connections existing between objects. System objects are usually individual devices: heat exchangers, extractors, compressors, etc. Nowadays, the use of artificial intelligence, expert systems, or neural networks is increasingly utilized to solve many technical and engineering problems [1]. Engineering practice requires a correct assessment of the mathematical model describing a given system by means of variables. Models describe a given system with different accuracy. Graphs and structural numbers play a role as models of mechanical systems [2-5] and are still systematically developed [6-11]. In addition, there are special stream graphs, e.g. in chemical and process engineering. Many optimization methods are used in technical solutions to improve the efficiency of internal combustion and electric vehicles [12, 13]. Unlike graphs, dendrite-tree structures do not have cycles, but there may be a different number of initial vertices. Therefore, such structures are applicable to variant searching and optimizing the solutions of the designed system [6]. Some of the functional requirements are transformed into the structural characteristics and incorporated in the generator as rules of enumeration. The generator enumerates all possible solutions using graph theory and combinatorial analysis. The

remaining functional requirements are incorporated in the evaluator as evaluation criteria for the selection of concepts [14]. This results in a class of feasible mechanisms. Finally, the most promising candidate is chosen for the product design.

The process may be iterated several times until the final product is achieved. This methodology has been successfully applied in the structure synthesis of epicyclic gear trains, automotive transmission mechanisms, variable-stroke engine mechanisms, robotic wrist mechanisms, etc. Generally, if a functional requirement can be written in a mathematical form, it should be included in the generator [2]. The current article presents the method of dependence graphs in the analysis of automatic transmission. The present work describes the algorithm of searching and generating parametric trees as outline graphs. The proposed method allows determining the optimal values of teeth, using only the relationships linking transmission elements with their presentation in the form of a graph. This allows a more accurate calculation of the number of teeth and an algorithmic approach to the problem. Methods using graphs and structural numbers have long been used in mechanics. Many previous works considered the theory of graphs as tools for studying system dynamics, both in the analysis and in the synthesis of complex mechanical systems. In particular, authors in [15, 16] focused on determining the power flow through closed kinematic chains. The circulating power dependency is used to analyze toothed gears. A graph describes the scalar product of the coupled graphs. The transmission system is described using kinematic chains, and the coupled toothed transmission system is described using block diagrams, which enables such systems to be analyzed in n degrees. In addition, from the diagrams (blocks) describing the coupled transmission systems, the authors constructed coupled graphs and pseudo-graphs. By varying the associated values, optimization is performed to minimize the load on a given kinematic chain of the transmission system. This is a different way of using dependency graphs than that described in the present work. The present methodology is based on input and output information (values). The graphs described in the cited works are not subject to decomposition or distribution over decision trees. In seeking optimum decisions (paths) it is desirable to define the properties of the structure in such a way that the nodes represent the sought parameter, and the leaves

Corresponding author: Adam Deptuła

determine the output decision that is optimum at a given moment, as a response of the system to a particular change. Decisions on the tree structure simulate a certain fictional "game" (mutual interaction) between the decision-maker (the decision taken) and the transmission system (the response to the decision).

II. GRAPHIC TRANSMISSION MODELS

The objectives of gear modeling with graphs vary: dynamic analysis, kinematic analysis [4], synthesis [17], structure analysis [18], enumeration [19], optimization of gear sequences, automatic design, etc. are based on the so-called graph grammars. The advantage of modeling gears with graphs is that issues considered using graph models can be solved in an algorithmic manner, which allows the use of computer programs and widely understood integrated decision-making systems in a simple manner. In graph theory, a graph is associated with many other algebraic structures as arrays, matroids, structural numbers, and linear spaces of cutoffs. These objects enable the coding of the transmission structure, which allows the use of advanced artificial intelligence (evolutionary, genetic, or immunological) algorithms. An important advantage of modeling mechanical systems with graphs is that some considerations can be carried out in parallel in the field of mechanics and graph theory [20]. The relevance of the results is based on the transformation of knowledge between these two areas. At present, there is a considerable interest in graphical methods in optimization, and especially in modeling of gears, hydraulic systems, mechanisms, trusses and frames. Among the methods of analysis of planetary gears, one can distinguish among others [5, 20-22]. In the case of Hsu rules, the graph is built according to the following rules: geometrical dimensions are omitted and the considered kinematic pairs are: rotary, "planet-yoke", and meshing. The contour graph method was discussed in [6, 7, 22, 23] for the analysis of mechanical systems. It is particularly useful for considering mechanisms of various types (so-called planar, crossheads, etc.). In particular, it can be used in the analysis of planetary gears. In the analysis of planetary gears, the contour graph method is used. This method is based on distinguishing a number of subsequent rigid links of mechanisms that form a closed loop, the so-called contour. In particular, [21] concerns only outline graphs and their applications.

Unlike graphs, dendrite-tree structures do not have cycles, but there may be a different number of initial vertices. Therefore, a different approach has been developed as a translation of a directed dependence graph, among others for game-tree structures. This approach is different from the previous literature on parametric automation machines and their applications, related to control systems, operating systems, importance analysis of construction and/or operating parameters, and gear analysis previously modeled using other graph types. Various methodologies have been developed for the systematic enumeration of kinematic structures. A thorough understanding of the structural characteristics of a given type of mechanism is critical for the development of an efficient algorithm. During the '80s, some approaches for structural synthesis of kinematic chains were presented. Authors in [24] presented a systematic approach for the kinematic analysis of

multi degree of freedom robotic bevel-gear trains. In the '00s, authors in [25, 26] presented the structure synthesis of mechanisms. Authors in [7] showed the kinematic, power flow, and efficiency analysis of epicyclic gear trains. Authors in [27] presented a survey of works connected with the problem of gear modeling by means of versatile graph theory models. More recently, authors in [28] studied the kinematic, static force, power flow, and efficiency analysis of gear trains. Other authors researched the topological synthesis of gear trains. The review analysis of the graph application issues in the transmission analysis can include kinematic analysis, torque analysis, power flow analysis, and mechanical efficiency analysis.

The general algorithm for gear modeling with graphs can be described as:

- Selecting a problem for kinematic analysis or synthesis and considering the elements related to it (abstracting).
- Determining the relationships for the specified elements (e.g. listing kinematic rotational pairs and meshing).
- Listing the selected subgraphs based on the basic transmission graph and saving the codes.
- Generating of the equations describing the gearing on the basis of codes – which is an orderly (algorithmic) way and solving the obtained system of equations.

III. ANALYSIS OF THE AUTOMATIC TRANSMISSION GEARBOXES WITH GAME-TREE STRUCTURES

A. Automatic Transmission Gearboxes

Modern automatic transmissions are complex devices, consisting of several hundred elements of a mechanical-hydraulic system and another several hundred in the electronic module. Automatic planetary gearboxes belong to a special type of gearbox. There are several of them, each of which consists of a sun gear with external meshing, satellites, and a crown gear with internal meshing. The specified gear ratio is implemented by braking or connecting a sun wheel or satellites. Brakes or clutches of individual planetary gears are controlled by hydraulic valves (transmission oil flow). The gearboxes are characterized by high durability and the ability to instantly change gears [29].

The current work is based on [6] concerning various possible developments of applications of graph methods in the analysis of gears. In particular, in [6] the concepts for describing the model of machine systems data using dependency graphs and signal flow were formulated. Among other things, the authors in [30] developed a method for finding the optimal number of teeth without transforming the graph. The algorithm for generating equations and determining the optimal number of teeth is calculated directly from the dependency graph considered as a signal flow graph. This approach is different from other works on the application of the graph method in gear analysis. For example, in [31], the gear analysis uses the Davie's equations and the interdependencies of cylindrical gear alignment. In turn, authors in [32] used hypergraphs to determine the relationship between elements and generate kinematic equations. The use of graph grammars

for the description of graph chains in the description of the gearboxes was used in [28, 33]. This is a different approach from that used in the works discussed above. The construction of contours in a graph begins in the supporting system and passes through successive rotating gear elements, each of which transfers rotary motion to the next. The contour is closed when we return in the analysis to the supporting system. To create a complete graph, all contours must be analyzed. The main goal of kinematic analysis of a toothed gear is to determine the transmission ratio and to identify any surplus gears.

The analysis of automatic gearboxes is similar to the analysis of a single planetary gear. During the process of kinematic structure enumeration using graph theory, isomorphism identification of graphs is an important and complicated problem. There are many approaches to detect isomorphic graphs, and these approaches have been largely algorithmic. This path is formed by the corresponding edges of the gear graph. Input and output are marked additionally. This path allows the analysis of the sequence of transmission of rotational motion by the subsequent elements of the transmission. In addition, it allows the detection of the so-called redundant elements for the given gear which is currently under consideration. The main purpose of the kinematic analysis of the planetary gear is to determine the kinematic ratio and possibly detect oversized gears. One of the methods of analysis, based on the use of the Willis formula for any complex planetary gears, is described in detail in the [34], in which the theory of contour graphs was used for kinematic analysis. This method allowed not only to determine the kinematic ratio, i.e. the angular velocities of all wheels and gear yokes, but also to determine the angular accelerations of rotating gear elements. In addition, redundant (oversized) wheels could be detected when plotting contour graphs. Individual gears were implemented using brakes, clutches and backstops (free wheels) as shown in Table I [34].

TABLE I. FUNCTIONAL MATRIX OF THE A4LD GEARBOX

Position	DE (1 2 3 4)	2 nd gear	1 th gear	Reverse
MT cl	(- - X X)	-	-	-
cl1	(- - - -)	X	X	X
b1	(- - - X)	-	-	-
Fw1	(X X X -)	X	X	X
cl2	(- - X X)	-	-	X
cl3	(X X X X)	X	X	-
b2	(- X - -)	X	-	-
b3	(- - - -)	-	X	X
Fw2	(X - - -)	-	-	-

The object of the analysis is a complex four-speed automatic transmission with overdrive type A4LD [29, 34].

B. Dependency Graph and Game-Tree Structures

A graph is an ordered pair $G = (V, E)$, in which V is a fine set of elements called the vertices of the graph, and E is a set of $(v_i, v_j) (v_i, v_j \in V)$ pairs called the edges of the graph. To fully specify the graph, one must also specify the relationship P

formed by the individual elements of the set of vertices $V (G)$ and edges $E (G)$. Then the graph can be called an ordered three:

$$G = (V, E, P) \quad (1)$$

where V is the set of graph vertices, E the set of edges (graph branches), and P is the three-member relationship which meets the following condition: There is such a pair of vertices $x, y \in V$ for each branch e , such that $\langle x, e, y \rangle \in P$. If for branch e exist $\langle x, e, y \rangle \in P$ and $\langle w, e, z \rangle \in P$ then either $x=w$ and $y=z$ or $x=y=z$. A dependency graph is an ordered pair $G=(X, R)$, in which X is a finite set of elements called vertices of the graph, and R is a set of pairs $(x_i, x_j) (x_i, x_j \in X)$ called the edges of the graph. In the case of parametric graphs, the notation introduced in [34] defines the signs: $G=(Q, Z)$, where Z is a set of pairs $(z_i, z_j) (z_i, z_j \in Z)$. The oriented dependency (game) graph shown in Figure 1 is composed of a set of vertices Q :

$$Q = \{q_1, q_2, q_3, q_4, q_5\} \quad (2)$$

and of set of edges Z :

$$Z = \{z_1, z_2, z_3, z_4, z_5, z_6, z_7\} \quad (3)$$

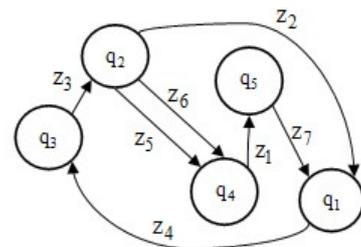


Fig. 1. An example of a dependency graph.

The path in the $G=(Q, Z)$ is the edge sequence $(z_{i_1}, z_{i_2}), (z_{i_2}, z_{i_3}), \dots, (z_{i_{k-1}}, z_{i_k})$ in which each $j \in \{2, 3, \dots, k\} (z_{i_{j-1}}, z_{i_j}) \in R$ and vertices $q_{i_1}, q_{i_2}, \dots, q_{i_k}$ are different pairs. Vertex q_{i_1} is called the beginning of the pat, and the top q_{i_k} the end of the road. As a result of a graph distribution from the chosen vertex, a tree structure with cycles is obtained in the first step and then, a general game tree structure is obtained. Each of them has an appropriate analytical formulation G_i^+ and G_i^{++} . A game tree structures is a part of the systematic searching method. The algorithm of graph distribution of dependence on parametric structures has been presented, among others, in [35]. The expression describing the degree of inferiority of a given component graph is marked with a pair of parentheses, inside which the expression concerning the given component graph is written. When starting to build an analytical expression representing a graph (Figure 1), first the starting vertex of the graph should be determined. In the case of starting the analysis from the initial vertex q_1 , we get the following expressions:

$$G(1)_{q1}^+ = ({}^0q_1({}^1z_2q_3({}^2z_3q_2({}^3z_4q_1, z_5q_4({}^4z_1q_5({}^5z_7q_1)^5)^4, ({}^6z_6q_4)^3)^2)^1)^0, \quad (4)$$

$$G(1)_{q1}^{++} = ({}^0q_1({}^1z_2q_3({}^2z_3q_2({}^3z_4q_1, z_5q_4({}^4z_1q_5({}^5z_7q_1)^5)^4, ({}^6z_6q_4({}^4z_1q_5({}^5z_7q_1)^5)^3)^2)^1)^0, \quad (5)$$

IV. KINEMATIC ANALYSIS OF THE GEAR WITH THE USE OF DEPENDENCY GRAPHS

Power is transmitted through the torque converter, h_1 yoke, free wheel Fw_1 , Cl_3 and Cl_2 clutches, gears 6, 5, 4 and yoke h_2 and in parallel through wheels 7, 8, 9 on the output shaft II (thanks to the Fw_2 backstop activated), as can be seen in Figure 2. The signal flow dependence graph of gears is shown in Figure 3.

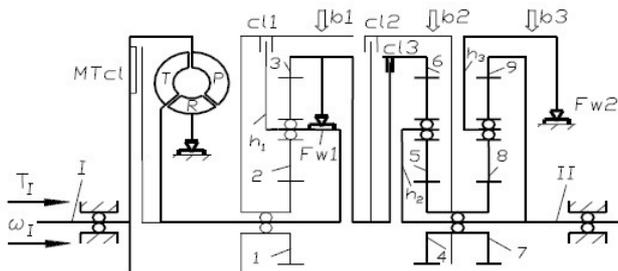


Fig. 2. Power efficiency diagram for DE1 gear.

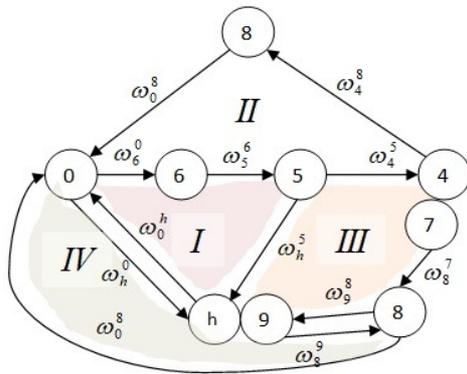


Fig. 3. Signal dependency graph for DE.

For the graph model, a system of contour equations of velocity $\omega_{i,i-1}$, peripheral velocities $\omega_{i,i-1} \times r_{Ai}$, angular $\varepsilon_{i,i-1}$, tangential $\varepsilon_{i,i-1} \times r_{Ai}$, and centripetal accelerations $\omega_i^2 \cdot r_{Ai,Ai+1}$ was generated and can be seen in (6- 9):

$$\begin{cases} \sum_{(i)} \omega_{i,i-1} = 0 \\ \omega_{6,0} + \omega_{5,6} + \omega_{h2,5} + \omega_{0,h2} = 0 \\ \omega_{6,0} + \omega_{5,6} + \omega_{4,5} + \omega_{8,7} + \omega_{0,8} = 0 \\ \omega_{4,5} + \omega_{8,7} + \omega_{9,8} + \omega_{5,h2} = 0 \\ \omega_{6,0} + \omega_{5,6} + \omega_{h2,5} + \omega_{8,9} + \omega_{0,8} = 0 \\ \sum_{(i)} r_{Ai} \times \omega_{i,i-1} = 0 \end{cases} \quad (6)$$

$$\begin{cases} \sum_{(i)} r_{Ai} \times \omega_{i,i-1} = 0 \\ r_6 \times \omega_{5,6} + (r_4 + r_5) \times \omega_{h2,5} = 0 \\ r_6 \times \omega_{5,6} + r_4 \times \omega_{4,5} + r_7 \times \omega_{8,7} + r_9 \times \omega_{9,8} = 0 \\ r_4 \times \omega_{4,5} + r_7 \times \omega_{8,7} + r_9 \times \omega_{9,8} + (r_4 + r_5) \times \omega_{5,h2} = 0 \\ r_6 \times \omega_{5,6} + (r_4 + r_5) \times \omega_{h2,5} + r_9 \times \omega_{8,9} = 0 \end{cases} \quad (7)$$

$$\begin{cases} \sum_{(i)} \varepsilon_{i,i-1} = 0 \\ \varepsilon_{6,0} + \varepsilon_{5,6} + \varepsilon_{h2,5} + \varepsilon_{0,h2} = 0 \\ \varepsilon_{6,0} + \varepsilon_{5,6} + \varepsilon_{4,5} + \varepsilon_{8,7} + \varepsilon_{0,8} = 0 \\ \varepsilon_{4,5} + \varepsilon_{8,7} + \varepsilon_{9,8} + \varepsilon_{5,h2} = 0 \\ \varepsilon_{6,0} + \varepsilon_{5,6} + \varepsilon_{h2,5} + \varepsilon_{8,9} + \varepsilon_{0,8} = 0 \end{cases} \quad (8)$$

$$\begin{cases} \sum_{(i)} r_{Ai} \times \varepsilon_{i,j-1} - \omega_i^2 \cdot r_{Ai,Ai+1} = 0 \\ r_6 \times \varepsilon_{5,6} + (r_4 + r_5) \times \varepsilon_{h2,5} = 0 \\ r_6 \times \varepsilon_{5,6} + r_4 \times \varepsilon_{4,5} + r_7 \times \varepsilon_{8,7} + r_9 \times \varepsilon_{9,8} = 0 \\ r_4 \times \varepsilon_{4,5} + r_7 \times \varepsilon_{8,7} + r_9 \times \varepsilon_{9,8} + (r_4 + r_5) \times \varepsilon_{5,h2} = 0 \\ \dot{\omega}_{6,9} = \omega_{6,0} / \omega_{9,0} = \\ r_6 \times \varepsilon_{5,6} + (r_4 + r_5) \times \varepsilon_{h2,5} + r_9 \times \varepsilon_{8,9} = 0 \end{cases} \quad (9)$$

where $\omega_{i,i-1}$ is the relative angular velocity vector of the element i relatively to the previous element $i-1$, $\omega_{i,0}$ is the vector of the absolute angular velocity of the element (relative to the fixed base), $r_{Ai} = r_{OAi}$ is the radius of the vector of point A_i (point A on element i), and $r_{Ai,Ai+1} = r_{Ai+1} - r_{Ai}$, $\varepsilon_{i,j-1}$ the vector of relative angular acceleration of the element relative to the previous element $i-1$.

V. APPLICATION OF THE SIGNAL FLOW GRAPH METHOD IN CALCULATING THE NUMBER OF GEAR TEETH

A new approach for the dependency graph can be obtained by using relationships that bind systems with their representation in the form of a dependency graph (in particular in relation to contour graphs). As mentioned above, a graph is an ordered pair $G=(V, E)$, in which V represents its vertices and E its edges. The attached \hat{G} search graph to a regular signal flow graph G is defined by replacing all G graph branches with the attached equivalents (Figure 4).

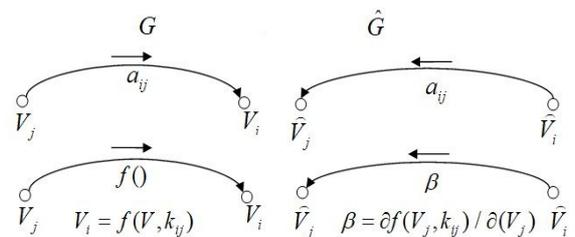


Fig. 4. Signal dependency graph for DE1.

A search graph \hat{G} for linear branches is created by reversing the direction of the signal flow. For the non-linear branch described by the function $f(V_j, k_{ij})$ (k_{ij} -number coefficient) in addition to reversing the direction of the branches, there is also a change in its description from $f(V_j, k_{ij})$ to $\partial f(V_j, k_{ij}) / \partial(V_j)$, where the derivative is determined at the point V_j corresponding to the solution of the graph G , i.e. searching the entire vertex j . This graph definition, which is the equivalent of the search circuit, allows to directly use the search graph method to calculate the sensitivity. The sensitivity of the V_o output signal relative to the gain of the linear branch or nonlinear branch stick factor is determined on the basis of the solution of graphs G and \hat{G} in the form of:

$$\begin{aligned} \frac{\partial V_o}{\partial a_{ij}} &= \hat{V}_i V_j \\ \frac{\partial V_o}{\partial k_{ij}} &= \hat{V}_i \frac{\partial f(V_j, k_{ij})}{\partial k_{ij}} \end{aligned} \quad (10)$$

The sensitivity determination method applies to both coupled and unidirectional graphs. By minimally changing only the excitation value (e.g. the number of teeth) in the search graph \hat{G} , the method can be adapted to determine the full gradient vector of the objective function [35]. The objective function will be expressed as the sum of squares of differences between the current values of the output signals V_{oi} and the set values d_i :

$$E(W) = \frac{1}{2} \sum_{i=1}^M (V_{oi} - d_i)^2 \quad (11)$$

where:

$$W = [W_1, W_2, \dots, W_N]^T \quad (12)$$

The gradient ∇E is obtained as:

$$\nabla E = \begin{pmatrix} \frac{\partial E}{\partial W_1} \\ \frac{\partial E}{\partial W_2} \\ \dots \\ \frac{\partial E}{\partial W_N} \end{pmatrix} \quad (13)$$

where:

$$\frac{\partial E}{\partial W_j} = \sum_{i=1}^M (V_{oi} - d_i) \frac{\partial V_{oi}}{\partial W_j} \quad (14)$$

The calculation of the j -th vector component of the gradient consists in the calculation of the derivative output signals V_{oi} and multiplying them by the appropriate difference: $V_{oi} - d_i$. The back propagation algorithm has been developed for use in unidirectional networks. Its interpretation and generalization in the form of graphs applies to both directed graphs and recursive

networks containing feedback cycles, e.g. a drawing with trees. It should be noted that in the case of unidirectional networks, it is easy to extract successive layers of neurons and search the signal flow between them, while in the dependence graphs this task is more complicated due to the existing cycles (feedback and contours). Figure 5 shows an example of the implementation of the search graph. Figure 6 shows the signal response for the selected graph vertices.

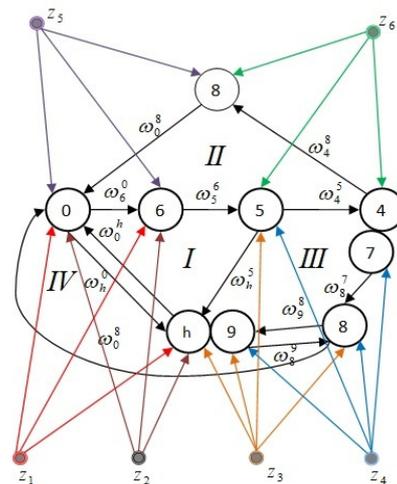


Fig. 5. Implementation of a search graph for a gear.

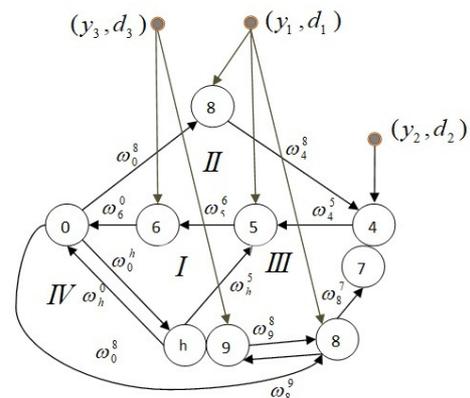


Fig. 6. The signal response for selected graph vertices.

The variables $(z_1, z_2, z_3, z_4, z_5, z_6)$ in Figure 5 are components of the input vector Z . The output of the graph consists of neurons marked by their output signals (y_1, y_2, y_3) and their corresponding set signals d_1, d_2, d_3 . When creating an attached \hat{G} (Figure 6), the output nodes of the normal graph are extracted, which are now available as input nodes for the \hat{G} graph. Input signal values can now be expressed as $\varepsilon_1 = y_1 - d_1$, $\varepsilon_2 = y_2 - d_2$, $\varepsilon_3 = y_3 - d_3$. For searching, the range of values for the number of teeth searched are assumed as: $z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9$: $z_1 \in (0, 1, 2, \dots, 100)$, $z_2 \in (0, 1, 2, \dots, 100)$, $\dots, z_9 \in (0, 1, 2, \dots, 100)$. After steps T , the solution $z_1(T), z_2(T), z_3(T) \dots$ is obtained, which can be considered as the solution to the search graph. In the case of

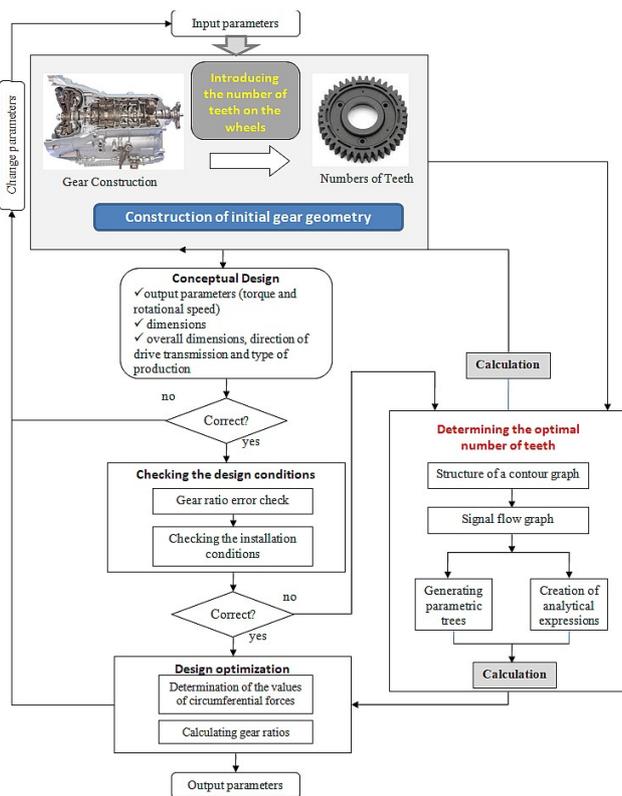


Fig. 10. Flowchart of the general optimization algorithm of a transmission system with selection of the appropriate number of teeth.

VI. CONCLUSION

The graph-based methods of analysis and synthesis of planetary gears provide an alternative for the accomplishing of the tasks in question. The automatic transmission presented in this article was modeled with the use of signal flow graphs. Due to the use of the new modeling method, there is a need to analyze the technical risk of such a solution. An element necessary in the risk assessment process is to take into account the requirements of the engineering design methodology. In this regard, it is worth choosing a design methodology that meets the following criteria:

- Completeness of technical criteria.
- Adequacy of the set of parameters describing the designed object.

The criteria formulated for assessment should be synthetic because this increases the objectivity of assessment of a given solution. Designing facilities involves increasing efficiency (to some extent) while reducing the efficiency of another team. Determining effectiveness consists in comparing the effects of a given measure with its expenditure. Particularly in this respect, it should be ensured that the risk of the adopted boundary standards in the scope of e.g. costs is not exceeded. The current article presents the application of dependency graphs to determine the optimal number of gear teeth for an automatic transmission. Each gearbox (automatic or manual planetary gear) can be additionally modeled with Hsu graphs [36].

REFERENCES

- [1] K. S. Belkhir, "Intelligent Maximum Power Tracking Control For a Wind Energy System Based on Magnetic Gear Generator," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4329–4333, Aug. 2019, <https://doi.org/10.48084/etasr.2851>.
- [2] F. Buchsbaum and F. Freudenstein, "Synthesis of kinematic structure of geared kinematic chains and other mechanisms," *Journal of Mechanisms*, vol. 5, no. 3, pp. 357–392, Sep. 1970, [https://doi.org/10.1016/0022-2569\(70\)90068-6](https://doi.org/10.1016/0022-2569(70)90068-6).
- [3] W. J. Sohn and F. Freudenstein, "An Application of Dual Graphs to the Automatic Generation of the Kinematic Structures of Mechanisms," *Journal of Mechanisms, Transmissions, and Automation in Design*, vol. 108, no. 3, pp. 392–398, Sep. 1986, <https://doi.org/10.1115/1.3258745>.
- [4] F. R. E. Crossley, "The Permutations of Kinematic Chains of Eight Member or Less from the Graph-Theoretic Viewpoint," *Developments in Theoretical and Applied Mechanics*, vol. 2, pp. 467–486, 1965.
- [5] F. Freudenstein, "An Application of Boolean Algebra to the Motion of Epicyclic Drives," *Journal of Engineering for Industry*, vol. 93, no. 1, pp. 176–182, Feb. 1971, <https://doi.org/10.1115/1.3427871>.
- [6] A. Deptuła and M. A. Partyka, "Application of dependence graphs and game trees for decision decomposition for machine systems," *Journal of Automation Mobile Robotics and Intelligent Systems*, vol. 5, no. 4, pp. 17–26, 2011.
- [7] C. A. Nelson and R. J. Cipra, "Simplified Kinematic Analysis of Bevel Epicyclic Gear Trains With Application to Power-Flow and Efficiency Analyses," *Journal of Mechanical Design*, vol. 127, no. 2, pp. 278–286, Mar. 2005, <https://doi.org/10.1115/1.1814392>.
- [8] X. Yu, J. Zhang, C. Fan, and S. Chen, "Stability analysis of governor-turbine-hydraulic system by state space method and graph theory," *Energy*, vol. 114, pp. 613–622, Nov. 2016, <https://doi.org/10.1016/j.energy.2016.07.164>.
- [9] N. F. Timerbaev, A. R. Sadrtidinov, and R. G. Safin, "Software Systems Application for Shafts Strength Analysis in Mechanical Engineering," *Procedia Engineering*, vol. 206, pp. 1376–1381, Jan. 2017, <https://doi.org/10.1016/j.proeng.2017.10.648>.
- [10] J. Xu *et al.*, "The Construction of Design Science Knowledge Graphs - Based on National Natural Science Foundation of China," *Proceedings of the Design Society: International Conference on Engineering Design*, vol. 1, no. 1, pp. 2607–2616, Jul. 2019, <https://doi.org/10.1017/dsi.2019.267>.
- [11] G. Chatterjee and L.-W. Tsai, "Enumeration of Epicyclic-Type Automatic Transmission Gear Trains," *SAE Transactions*, vol. 103, pp. 1415–1426, 1994.
- [12] G. Gururaj and K. S. Gowri, "Performance Optimization of an eCAR by Parametric Analysis," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4968–4973, Dec. 2019, <https://doi.org/10.48084/etasr.3139>.
- [13] A. Y. Al-Rawashdeh, "Investigation of an Induction Wound Rotor Motor to Work as a Synchronous Generator," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 4071–4074, Apr. 2019, <https://doi.org/10.48084/etasr.2606>.
- [14] G. Kron, "Generalized Theory of Electrical Machinery," *Transactions of the American Institute of Electrical Engineers*, vol. 49, no. 2, pp. 666–683, Apr. 1930, <https://doi.org/10.1109/T-AIEE.1930.5055554>.
- [15] J. Wojnarowski, "The graph method of determining the loads in complex gear trains," *Mechanism and Machine Theory*, vol. 11, no. 2, pp. 103–121, Jan. 1976, [https://doi.org/10.1016/0094-114X\(76\)90003-3](https://doi.org/10.1016/0094-114X(76)90003-3).
- [16] D. J. Sanger, "The determination of power flow in multiple-path transmission systems," *Mechanism and Machine Theory*, vol. 7, no. 1, pp. 103–109, Mar. 1972, [https://doi.org/10.1016/0094-114X\(72\)90020-1](https://doi.org/10.1016/0094-114X(72)90020-1).
- [17] F. Freudenstein and L. Dobrjanskyj, "On a theory for the type synthesis of mechanisms," in *Applied Mechanics*, H. Gortler, Ed. Berlin, Heidelberg: Springer, 1966, pp. 420–428.
- [18] L.-W. Tsai and C.-C. Lin, "The Creation of Nonfractionated, Two-Degree-of-Freedom Epicyclic Gear Trains," *Journal of Mechanisms*,

- Transmissions, and Automation in Design*, vol. 111, no. 4, pp. 524–529, Dec. 1989, <https://doi.org/10.1115/1.3259033>.
- [19] C. -H. Hsu, "Graph notation and kinematic equations of motion of planetary differential gear trains," *International Journal of Vehicle Design*, vol. 13, no. 3, pp. 233–241, Jan. 1992, <https://doi.org/10.1504/IJVD.1992.061725>.
- [20] H.-S. Yan and Y.-T. Chiu, "An algorithm for the construction of generalized kinematic chains," *Mechanism and Machine Theory*, vol. 62, pp. 75–98, Apr. 2013, <https://doi.org/10.1016/j.mechmachtheory.2012.11.005>.
- [21] D. B. Marghitu, *Kinematic Chains and Machine Components Design*, 1st edition. Amsterdam, Boston: Academic Press, 2005.
- [22] A. Deptula, J. Drewniak, and M. A. Partyka, "Application of Dependence Graphs and Game Trees in Analysis of a Planetary Gear Modeled With a Contour Graph," *Machine Dynamics Research*, vol. 41, no. 3, pp. 105–114, 2017.
- [23] D. Gibson and S. Kramer, "Symbolic Notation and Kinematic Equations of Motion of the Twenty-Two Basic Spur Planetary Gear Trains," *Journal of Mechanisms, Transmissions, and Automation in Design*, vol. 106, no. 3, pp. 333–340, Sep. 1984, <https://doi.org/10.1115/1.3267416>.
- [24] J. M. del Castillo, "Enumeration of 1-DOF Planetary Gear Train Graphs Based on Functional Constraints," *Journal of Mechanical Design*, vol. 124, no. 4, pp. 723–732, Nov. 2002, <https://doi.org/10.1115/1.1514663>.
- [25] A. C. Rao and P. B. Deshmukh, "Computer aided structural synthesis of planar kinematic chains obviating the test for isomorphism," *Mechanism and Machine Theory*, vol. 36, no. 4, pp. 489–506, Apr. 2001, [https://doi.org/10.1016/S0094-114X\(00\)00030-6](https://doi.org/10.1016/S0094-114X(00)00030-6).
- [26] H. Ding, Z. Huang, and Y. Cao, "The topological graphs creation automatically of kinematic chains and atlas database establishment," *Journal of Mechanical Engineering*, vol. 42, no. 4, pp. 32–36, 2006.
- [27] J. Wojnarowski, J. Kopec, and S. Zawislak, "Gears and graphs," *Journal of Theoretical and Applied Mechanics*, vol. 44, no. 1, pp. 139–162, 2006.
- [28] M.-C. Tsai, C.-C. Huang, and B.-J. Lin, "Kinematic Analysis of Planetary Gear Systems Using Block Diagrams," *Journal of Mechanical Design*, vol. 132, May 2010, Art. no. 065001, <https://doi.org/10.1115/1.4001598>.
- [29] J. Drewniak, S. Zawislak, and A. Wieczorek, "Analysis method of automatic planetary transmission kinematics," *Scientific Journal of Silesian University of Technology-Series Transport*, vol. 83, pp. 65–71, 2014.
- [30] A. Deptula, J. Drewniak, and M. A. Partyka, "Selection of the Optimal Number of Teeth for a Biplanetary Bevel Gear Based on Discrete Optimization," *Machine Dynamics Research*, vol. 42, no. 3, pp. 55–66, 2018.
- [31] L. P. Laus, H. Simas, and D. Martins, "Efficiency of gear trains determined using graph and screw theories," *Mechanism and Machine Theory*, vol. 52, pp. 296–325, Jun. 2012, <https://doi.org/10.1016/j.mechmachtheory.2012.01.011>.
- [32] J. R. Goma Ayats, U. Diego-Ayala, J. Minguella Canela, F. Fenollosa, and J. Vivancos, "Hypergraphs for the analysis of complex mechanisms comprising planetary gear trains and other variable or fixed transmissions," *Mechanism and Machine Theory*, vol. 51, pp. 217–229, May 2012, <https://doi.org/10.1016/j.mechmachtheory.2011.07.011>.
- [33] W. Yang, H. Ding, B. Zi, and D. Zhang, "New Graph Representation for Planetary Gear Trains," *Journal of Mechanical Design*, vol. 140, Nov. 2017, Art. no. 012303, <https://doi.org/10.1115/1.4038303>.
- [34] A. Deptula, *Application of graphic decision structures in the design and management methodology*. Opole, Poland: Publishing House of the Opole University of Technology, 2020.
- [35] M. S. Pishvaei and M. Rabbani, "A graph theoretic-based heuristic algorithm for responsive supply chain network design with direct and indirect shipment," *Advances in Engineering Software*, vol. 42, no. 3, pp. 57–63, Mar. 2011, <https://doi.org/10.1016/j.advengsoft.2010.11.001>.
- [36] C.-H. Hsu, "An Analytic Methodology for the Kinematic Synthesis of Epicyclic Gear Mechanisms," *Journal of Mechanical Design*, vol. 124, no. 3, pp. 574–576, Aug. 2002, <https://doi.org/10.1115/1.1485094>.

The Effect of Sisal Juice Extract Admixture on Compressive and Flexural Strength of Cement Concrete

Michael O. Eloget
SMARTEC, Jomo Kenyatta University of Agriculture
and Technology
Nairobi, Kenya
miklate2@gmail.com

Silvester O. Abuodha
Department of Civil Engineering
University of Nairobi
Nairobi, Kenya
sochieng@yahoo.com

Mathew M. O. Winja
Department of Civil Engineering
Jomo Kenyatta University of Agriculture and Technology
Nairobi, Kenya
mathewwinja@yahoo.com

Abstract-The characteristics of concrete are influenced by the ratio of water to cementitious materials (w/c) used in the mixture. An increase in paste quality will yield higher compressive and flexural strength, lower permeability, increased resistance to weathering, improve the bond between concrete and reinforcement, reduced volume change from drying and wetting, and reduced shrinkage cracking tendencies. Admixtures are used to improve the properties of concrete or mortar. The current study investigates the effect of Sisal Juice Extract (SJE) as an admixture on concrete durability. SJE contains unrefined minerals which can be used as organic retarders to increase the rate of strength development at an early age. A total of 84 concrete cubes were produced in 7 sets of 12 samples each. One set was the control mix which had zero SJE content. The remaining sets had varying dosages of SJ namely 5%, 10%, 15%, 20%, 25%, and 30%. Twelve beam specimens were also cast and subjected to the three-point flexural test. To establish the effect on strength of concrete, compressive strength was tested at 7, 14, 28, and 56 days while flexural strength was tested at 28 days. The highest compressive strength was achieved at 5% dosage beyond which a decrease in strength occurred for all the higher dosages.

Keywords-sisal juice extract; compressive strength; flexural strength

I. INTRODUCTION

Concrete is one of the most commonly used construction materials [1]. The production of concrete is achieved by mixing predetermined proportions of cement and aggregates with water to form a plastic material which transforms to a hard strong and load bearing material when dried [2]. The hardening of concrete is attained through hydration which is an exothermic reaction between cement and water. The addition of water to concrete during mixing enables hydration to take place

while the excess quantity forms a slurry with the cement which coats the aggregates in the mix, lubricating them and allowing easy movement during compaction [3]. The first stage of hydration that leads to stiffening is caused by the reaction between water and a chemical constituent of cement called tricalcium silicate (C_3S) [4]. The time taken for concrete to harden after the addition of water to the cement-aggregate mix is referred to as setting time. The last stage occurs over a longer period, called final setting time, and is caused by the reaction of water with calcium silicate (C_2S). It is during this stage that hardening occurs, as a result of the exothermic nature of hydration. The higher the rate of hydration is, the higher the rate of both heat generation and strength development [5-7].

The expansion and contraction of the hardened concrete are associated with the generated heat which causes stresses that lead to microscopic cracks within the concrete structure. This phenomenon causes problems when constructing large concrete structural members such as piers, dams and bridges and generally when concreting structures in hot climates [8, 9]. In order to slow both the rate of hydration and the generation of heat and strength development, admixture materials called retarders may be used as they have the effect of delaying the initial setting time of concrete thereby providing enough time for concrete operations [10, 11]. Retarding admixtures reduce the rate of hydration by blocking cement from reacting with water [12-15]. This happens because the retarding admixtures are mainly based on materials having lignosulfonic acids and their salts, hydroxyl-carboxylic acids and their salts, sugars including their derivatives, inorganic salts such as borates phosphates, and zinc and lead salts. [7, 16-19]. Admixtures come in various forms and are usually imported and expensive to acquire, hence cheaper admixture materials need to be established from locally available sources [20, 21]. The ability

Corresponding author: Michael O. Eloget

of retarders to reduce the rate of heat generation and strength development when concreting large structural members also helps to prevent the occurrence of capillary cracks that compromise the durability of concrete structures [22].

In [23], the effects of sugar on the setting time of cement and the compressive strength of concrete were investigated. In all curing periods the compressive strength was found minimum for concrete with 0.3% sugar and maximum for concrete with 0.08% sugar. Only in 0.3% sugar, the compressive strength was found less than the strength found for concrete with 0% sugar. The compressive strengths of all the other sugar percentages were found increased in comparison with plain concrete's. The authors concluded that the concrete compressive strength was maximum for 0.08% sugar in all curing periods among the sugar percentages used. Authors in [24] evaluated the effect of sugar on the compressive strength of concrete. Ordinary Portland Cement-OPC (ACC cement) of 43 grades was used for the cube casting. The cubes having sugar dosages of 0%, 0.06%, 0.08% by weight of cement were cast and the compressive strength of cubes for 7, 21, and 28 days were tested. It was observed that the setting time and workability of concrete increased with the increase in sugar dosage and that there was an increase in the compressive strength of about 16.02% at 28 days compared to ordinary concrete [24]. Authors in [25] studied the effect of sugar on strength and durability of HVFAC concrete with similar results. It was observed that the addition of sugar increased the workability and compressive strength of concrete.

Waste materials can be used to produce new products or can be used as admixtures so that natural resources are used more efficiently and the environment is protected from waste deposits [26]. Kenya, like many other countries, grows sisal in large quantities for extraction of fiber but the remains after extraction are not fully utilized. The objective of this paper is to study the potential of utilizing SJE that is generated from sisal waste as an economical admixture. Statistical assessment on compressive strength, flexural strength, and setting time has been conducted in order to investigate the effectiveness of SJE as an admixture. In addition, the utilization of the sisal waste will address the environmental pollution by making good use of sisal fiber waste material

II. MATERIALS AND METHODS

A. Materials

1) Cement

OPC class 53 conforming to KEBS SRN 103 was used in the production of the concrete for all mixes.

2) Fine and Coarse Aggregates

Natural river sand obtained from local suppliers was used as fine aggregates. Coarse aggregates originated from crushed stone type of nominal size of 10/20 mm. Their properties can be seen in Table I.

3) Sisal Juice

The SJ used for the experiment was obtained from sisal leaves harvested from a local single plantation to ensure the use of a single species. The juice was extracted by squeezing the

leaves in a cleaned improvised machine that allowed collection in a container.

TABLE I. PROPERTIES OF FINE AND COARSE AGGREGATES

Properties	River sand	Coarse aggregates
Fineness modulus	2.70	4.82
Silt content	0.5	-
Specific gravity	2.48	2.50
Water absorption (%)	2.1	3.10

4) Water

Water conforming to the requirements for concreting and curing as per IS: 456-20009 was used.

B. Test Methods

The preparation of specimens was done by measuring the required proportions of fine and coarse aggregates and cement and thereafter mixing was done in dry state before the addition of water and SJ. Water was mixed with different dosages of SJ before addition to the concrete mix. A homogenous mix was achieved through thorough mixing of the concrete ingredients. Two concrete mixes were produced. Mix "A" was produced without any addition of SJE while mix "B" had different dosages of SJE. Batched concrete mixes were prepared by replacement of 5%, 10%, 15%, 20%, 25% and 30% water with SJE in six sets as shown in Table II.

TABLE II. DETAILS OF MIX PROPORTIONS CLASS 25 CONCRETE

Mix Set	Mix A (OPC)	Mix B (OPC + SJ)					
	1	2	3	4	5	6	7
Dosage (%)	0	5	10	15	20	25	30

TABLE III. CONCRETE SPECIMEN DETAILS

	No. of cubes (specimen size: 150×150×150mm)	No. of beams (specimen size: 150×150×530mm)
Mix A	12	2
Mix B	72	10

1) Sisal Juice PH Test

The PH value of the fresh SJ was tested in the laboratory using the universal PH scale. SJE was subjected to chemical analysis using XRF technology (X-Ray Florescence spectra analysis).

2) Compressive Strength Test

Concrete specimens were compacted into a detachable mould with dimensions of 150×150×150mm. The specimens were demoulded after 24 hours and were immersed in water to cure. A total of 84 cubes were cast, and they were cured for 7, 14, 28, and 56 days. Compressive strength tests were conducted after the respective curing days. A compression machine was used in the laboratory, in accordance with BS EN 12390.

3) Flexural Strength Test

The test was carried out to determine the strength of concrete according to BS EN: 12390-5: 2009. The specimens

for various SJ dosages were prepared and cast in moulds measuring 150×150×530mm and were cured for 28 days. A total of 12 beams were tested using the three point flexural test method. The beam specimens were placed between two supports in the universal testing machine and load was applied at the center until failure occurred.

4) *Setting Time of Cement Mortar*

The setting time of cement mortar made with replacement of water with SJ by 0%, 5%, 10%, 15%, 20%, 25% and 30% were obtained in accordance with BS EN 196-3:1995.

III. RESULTS AND DISCUSSION

A. *Chemical Analysis of Sisal Juice Extract*

The fresh SJE was subjected to a litmus test giving a PH of 5, which is acidic. The decrease in compressive strength of concrete cubes with increase in the quantity of SJE was caused by the fact that the components of concrete break down during contact with acid [18]. From the results in Table IV, it can be noted that the content of calcium oxide (CaO) is the most predominant element at 35% followed closely by potassium oxide (K₂O). These two ingredients are also found in gypsum which is a mineral used for retarding the setting of cement during manufacture. Other elements found in both gypsum and SJE include magnesium oxide (MgO) and sulphur (S) [27, 28].

TABLE IV. CHEMICAL COMPOSITION OF SISAL JUICE

Element	Percentage by Weight (%)
Calcium oxide (CaO)	35.88
Potassium oxide (K ₂ O)	32.40
Magnesium oxide (MgO)	13.74
Aluminum oxide (AL ₂ O ₃)	4.19
Silicon dioxide (SiO ₂)	4.08
Iron (Fe)	3.55
Sulphur (S)	2.95
Phosphorous pentoxide (P ₂ O ₅)	1.48
Chlorine (Cl)	1.00
Zinc (Zn)	0.62
Manganese (Mn)	0.38
Titanium (Ti)	0.23
Strontium(Sr)	0.11

B. *Effects of SJE on Compressive Strength*

TABLE V. EFFECT OF SISAL JUICE ON COMPRESSIVE STRENGTH

		Mean strength (N/mm)						
		Percentage						
		0%	5%	10%	15%	20%	25%	30%
Days	7	17.02	20.18	16.74	15.63	11.86	12.43	9.76
	14	22.10	24.87	20.85	19.55	17.29	14.97	11.92
	28	25.00	26.08	21.97	22.79	20.27	18.07	17.96
	56	28.90	30.47	25.03	23.38	21.21	19.53	18.69

The test results in Table V indicate an increase in strength between control and cement with 5% of SJ but beyond 5%, the strength decreased steadily. The steady decrease in compressive strength can be attributed to the fact that SJE is acidic and an increase of SJE breaks down the components of concrete during contact as shown in [18]. The results in Figure 1 point out a strength gain trend with 5% dosage attracting the

highest compressive strength while 30% SJE gained the least strength at all ages. In Figure 2, the compressive strength difference between the control mix and 5% SJ dosage is 4.32%. The results show that there is a low rate of decrease in strength between 10% and 15% of SJ dosage but beyond 15% to 25% SJ, the rate increases. This result is in line with the results demonstrated in [23].

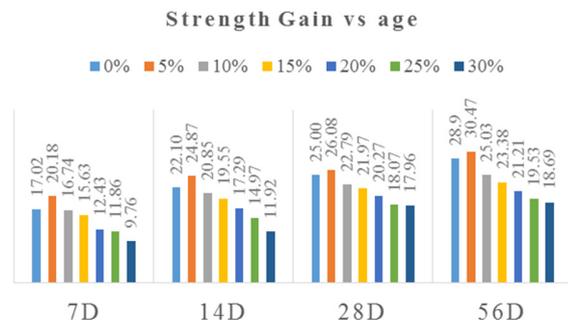


Fig. 1. Strength of concrete at different ages.

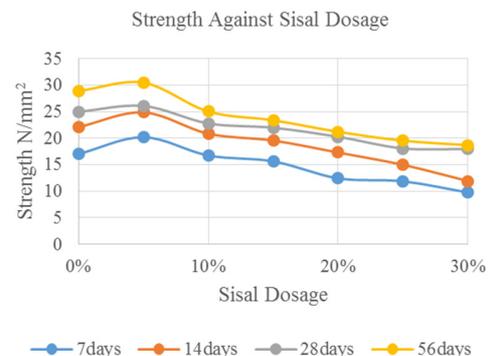


Fig. 2. Effects of SJ dosage on concrete's strength.

C. *Flexural Strength Test*

The test was conducted in accordance with BS EN 12390:2009 by using the three point load system. Beams, with a size of 150×150×550mm were used (Figure 3).

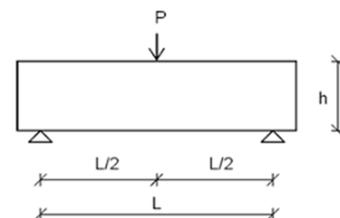


Fig. 3. Three point load system.

The results for flexural strength of SJE concrete are plotted against various percentages of SJE dosage. In Figure 4, it can be seen that when SJE dosage is increased, a significant increase in flexural strength is obtained. The predominant increase in flexural strength is found at 20% SJE dosage. The percentage of increase in flexural strength is 18% compared to control concrete. There was a slight decrease in flexural strength at 25% and 30% dosage.

TABLE VI. FLEXURAL STRENGTH TEST RESULTS

SJ dosage	S/No	Load (N)	Flexural strength (N/mm ²)	Avg.
0%	Specimen 1	32.21	4.29	4.3
	Specimen 2	32.40	4.32	
5%	Specimen 1	33.00	4.40	4.4
	Specimen 2	33.00	4.40	
10%	Specimen 1	33.75	4.50	4.6
	Specimen 2	34.88	4.65	
15%	Specimen 1	36.00	4.80	4.9
	Specimen 2	36.75	4.90	
20%	Specimen 1	39.98	5.33	5.1
	Specimen 2	36.98	4.93	
25%	Specimen 1	37.95	5.06	4.9
	Specimen 2	36.00	4.80	
30%	Specimen 1	31.50	4.20	4.5
	Specimen 2	36.00	4.80	

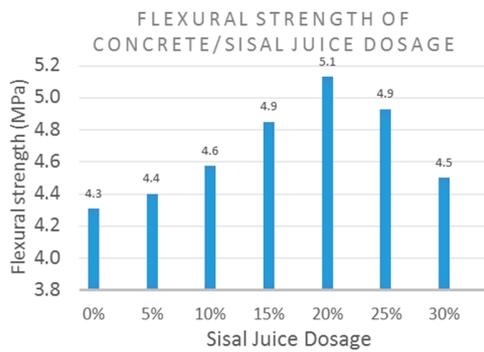


Fig. 4. Flexural strength of concrete at different SJ dosages.

The outcomes demonstrated that beam flexural performance increased rapidly by the addition of SJ, up to 20% and then it decreased slightly which is in line with results presented in [29]. Furthermore, diagonal cracks were observed on all the beams, including the control. Ultimate collapse occurred by concrete crushing within the compression zone for all the beams.



Fig. 5. Ultimate collapse of a concrete beam.

D. Comparison between Flexural and Compressive Strength

Table VI shows the relationship between the flexural and the compressive strength of concrete with SJE.

TABLE VII. FLEXURAL AND COMPRESSIVE STRENGTH COMPARISON AT 28 DAYS

Sisal dosage (%)	Flexural strength (N/mm ²)	Compressive strength (N/mm ²)
0	4.3	25.00
5	4.4	26.08
10	4.6	21.97
15	4.9	22.79
20	5.1	20.27
25	4.9	18.07
30	4.5	17.96

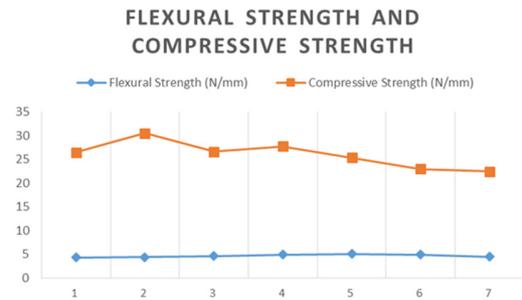


Fig. 6. Comparison between flexural and compressive strength at different sisal dosages.

Flexural strength must be between 10 to 20% of compressive strength depending on the type, size and volume of coarse aggregates used, see Table VIII.

TABLE VIII. ACCEPTABLE RANGE OF FLEXURAL STRENGTH

Sisal dosage (%)	Flexural strength (N/mm ²)	Compressive strength (N/mm ²)	Acceptable range of flexural strength	Remarks on the range of flexural strength
0	4.3	25.00	2.2 to 4.4	Acceptable
5	4.4	26.08	2.6 to 5.2	Acceptable
10	4.6	21.97	2.3 to 4.4	Unacceptable
15	4.9	22.79	2.3 to 4.6	Unacceptable
20	5.1	20.27	2.0 to 4.1	Unacceptable
25	4.9	18.07	1.8 to 3.6	Unacceptable
30	4.5	17.96	1.8 to 3.6	Unacceptable

High flexural strength is essential for stress-bearing restorations, when high pressure/stress is exerted on the material or restoration. From Table VIII, an increase of about 4.32% in compressive strength is witnessed and a significant improvement in flexural toughness is observed in the case of replacing water by 5% SJE. The flexural strength of 4.4N/mm² falls within the required range of 10% to 20% of compressive strength. The appropriate SJ dosage for the production of SJE concrete was found to be 5%. The water replacement by SJE enhanced the flexural strength compared to control mix.

E. Effects of SJE on Setting Time

Figure 7 shows that both the initial setting time and the final setting time increased at an average of 14% and 23% respectively with increase in SJ dosage. According to Table IX, the initial setting time of the cement paste increased with increase in SJ dosage when compared to the reference value. The final setting time also increased exponentially with increase in replacement of water by SJE. This can be attributed to the adsorption of the retarding SJ on the surface of cement

particles, forming a protective skin which slows down the process of hydration hence the reason why there is a steady rise of setting time. The layer of retarding admixture around the cement particles acts as a diffusion barrier. Due to this diffusion barrier, it becomes difficult for the water molecules to reach the surface of the un-hydrated cement grains and hence the hydration slows down, and the dormant period (period of relatively inactivity) is lengthened and thus the paste remains plastic for a longer time.

TABLE IX. INITIAL AND FINAL SETTING TIMES OF CEMENT MORTAR WITH VARYING DOSAGES OF SISAL JUICE EXTRACT

SJE dosage (%)	Initial setting time (min)	Percentage increase (%)	Final setting time (tin)	Percentage increase (%)
0	179		334	
5	214	20	467	40
10	234	31	521	56
15	237	32	620	86
20	242	35	893	167
25	265	48	1767	429
30	372	107	1840	451

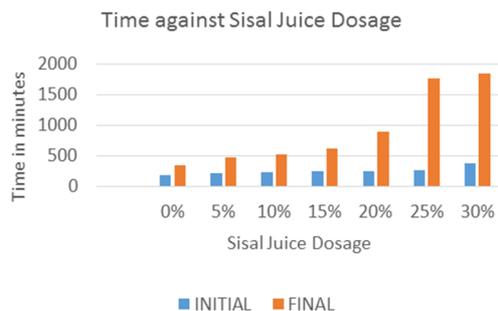


Fig. 7. Effects of SJE on initial and final setting time.

IV. CONCLUSIONS

The following conclusions can be drawn from the experimental study:

- SJE is acidic in nature and contain elements like calcium oxide (CaO) and potassium oxide (K₂O) which are predominant in retarders.
- Early compressive strength development is achieved at all ages with replacement of water by 5% SJE.
- The gradual replacement of water by SJE dosage enhanced flexural strength.
- The optimum water replacement by SJE that achieved high compressive strength and appropriate flexural strength is 5%.
- The replacement of water by SJE retarded the setting time of concrete.

V. RECOMMENDATIONS

The 22.79N/mm² concrete strength achieved at 28 days for 15% dosage can be used for low loading retaining wall with low surcharge. The 5% replacement of water by SJE concrete can be used to improve site capacity by increasing formwork rotation and also fasten construction time.

It is hereby suggested that further work on SJE should be carried out at controlled conditions. It is also recommended that prior to the acceptance of SJE as a retarder, the properties of different species of sisal leaves need to be studied.

REFERENCES

- [1] M. Uysal and V. Akyuncu, "Durability performance of concrete incorporating Class F and Class C fly ashes," *Construction and Building Materials*, vol. 34, pp. 170–178, Sep. 2012, <https://doi.org/10.1016/j.conbuildmat.2012.02.075>.
- [2] N. Bheel, R. A. Abbasi, S. Sohu, S. A. Abbasi, A. W. Abro, and Z. H. Shaikh, "Effect of Tile Powder Used as a Cementitious Material on the Mechanical Properties of Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4596–4599, Oct. 2019, <https://doi.org/10.48084/etasr.2994>.
- [3] N. Bheel, K. A. Kalhor, T. A. Memon, Z. U. Z. Lashari, M. A. Soomro, and U. A. Memon, "Use of Marble Powder and Tile Powder as Cementitious Materials in Concrete," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5448–5451, Apr. 2020, <https://doi.org/10.48084/etasr.3378>.
- [4] L. Zhang, L. J. J. Catalan, R. J. Balec, A. C. Larsen, H. H. Esmacili, and S. D. Kinrade, "Effects of Saccharide Set Retarders on the Hydration of Ordinary Portland Cement and Pure Tricalcium Silicate," *Journal of the American Ceramic Society*, vol. 93, no. 1, pp. 279–287, 2010, <https://doi.org/10.1111/j.1551-2916.2009.03378.x>.
- [5] R. Ilangovana, N. Mahendrana, and K. Nagamanib, "Strength and durability properties of concrete containing quarry rock dust as fine aggregates," *ARN Journal of Engineering and Applied Science*, vol. 3, no. 5, pp. 20–26, 2008.
- [6] R. Nagalakshmi, "Experimental study on strength characteristics on M25 concrete with partial replacement of cement with fly ash and coarse aggregate with coconut shell," *International Journal of Scientific & Engineering Research*, vol. 4, no. 1, 2013.
- [7] K. S. Pann, T. Yen, C.-W. Tang, and T. D. Lin, "New Strength Model Based on Water-Cement Ratio and Capillary Porosity," *Materials Journal*, vol. 100, no. 4, pp. 311–318, Jul. 2003, <https://doi.org/10.14359/12669>.
- [8] A. M. Neville, *Properties of Concrete*, 5th ed. Harlow, UK; New York, NY, USA: Trans-Atlantic Publications, Inc., 2012.
- [9] Z. Li, *Advanced Concrete Technology*. New York, NY, USA: John Wiley & Sons, 2011.
- [10] M. Jammal, J. Nurdiana, and A. M. Ghulam, "The effect of the combination additive admixtures to increase setting time and compressive strength at the concrete casting in warm tropical temperature," *International Journal of Civil Engineering and Technology*, vol. 11, no. 4, pp. 61–72, 2020.
- [11] S. A. Chandio, B. A. Memon, M. Oad, F. A. Chandio, and M. U. Memon, "Effect of Fly Ash on the Compressive Strength of Green Concrete," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5728–5731, Jun. 2020, <https://doi.org/10.48084/etasr.3499>.
- [12] S. B. Pathan and V. V. Singh, "Using Molasses in Concrete As A Time Retarding Admixture," *International Journal of Engineering Research & Technology*, vol. 6, no. 11, pp. 509–513, Nov. 2017.
- [13] R. Rixom and N. Mailvaganam, *Chemical Admixtures for Concrete*, 3rd ed. London, UK: E & FN SPON, 1999.
- [14] S. Vadivelan, K. Abirami, R. Agasthiya, D. Karthikayeni, and V. Meena, "Influence of Cactus Milk and Blood in Natural Soil Concrete For Improving Compressive Strength of Concrete," *International Journal of Engineering Research & Technology*, vol. 6, no. 4, pp. 1–5, Apr. 2018.
- [15] T. K. Akchurin, A. V. Tukhareli, and T. F. Cherednichenko, "Effective Concrete Modified by Complex Additive Based on Waste Products of Construction Acrylic Paints," *Procedia Engineering*, vol. 150, pp. 1468–1473, Jan. 2016, <https://doi.org/10.1016/j.proeng.2016.07.083>.
- [16] T. Y. Erdoğan, *Admixtures for concrete*. Ankara, Turkey: Middle East Technical University Press, 1997.

- [17] B. Khan and B. Baradan, "The Effect of Sugar on setting-time of various types of cements," *Quarterly Science Vision*, vol. 8, no. 1, pp. 71–78, 2002.
- [18] A. W. Otunyo, S. C. Onwusiri, and N. Nwaiwu, "Effect of Sugar Cane Juice on Slump Values, Setting Times and Strength of Concrete," *Nigerian Journal of Technology*, vol. 34, no. 2, pp. 254–258, Mar. 2015, <https://doi.org/10.4314/njt.v34i2.6>.
- [19] A. K. Parande, K. Stalin, R. K. Thangarajan, and M. S. Karthikeyan, "Utilization of Agroresidual Waste in Effective Blending in Portland Cement," *ISRN Civil Engineering*, vol. 2011, May 2011, Art. no. 701862, <https://doi.org/10.5402/2011/701862>.
- [20] E. Mello, C. Ribellato, and E. Mohamedelhasan, "Improving Concrete Properties with Fibers Addition," *International Journal of Civil and Environmental Engineering*, vol. 8, no. 3, pp. 249–254, 2014.
- [21] N. Tripathi, C. D. Hills, R. S. Singh, and C. J. Atkinson, "Biomass waste utilisation in low-carbon products: harnessing a major potential resource," *npj Climate and Atmospheric Science*, vol. 2, no. 1, pp. 1–10, Oct. 2019, <https://doi.org/10.1038/s41612-019-0093-5>.
- [22] B. Suhendro, "Toward Green Concrete for Better Sustainable Environment," *Procedia Engineering*, vol. 95, pp. 305–320, Jan. 2014, <https://doi.org/10.1016/j.proeng.2014.12.190>.
- [23] M. A. Azad, S. Rahman, and R. Chowdhury, "Effect of Sugar on Setting Time of Cement and Compressive Strength of Concrete," in *2nd International Conference on Research and Innovation in Civil Engineering (ICRICE 2020)*, Chittagong, Bangladesh, Jan. 2020.
- [24] J. Chand and S. Dhyani, "Effect of sugar on the compressive strength of concrete," *International Journal of Advanced Technology & Engineering Research*, vol. 5, no. 4, pp. 5–7, 2015.
- [25] T. R. Nikhil, H. S. Naveen Kumar, and T. P. Lakshmi, "Effect of sugar on Strength and Durability Characteristics of HVFAC Pavements," *International Journal of Engineering Research & Technology*, vol. 8, no. 11, pp. 288–291, Nov. 2019.
- [26] A. Baikerikar, "A Review on Green Concrete," *Journal of Emerging Technologies and Innovative Research*, vol. 1, no. 6, pp. 472–474, 2014.
- [27] Y. Chen and I. Odler, "On the origin of portland cement setting," *Cement and Concrete Research*, vol. 22, no. 6, pp. 1130–1140, Nov. 1992, [https://doi.org/10.1016/0008-8846\(92\)90042-T](https://doi.org/10.1016/0008-8846(92)90042-T).
- [28] J. H. Wilder, "Gypsum: its Occurrence, Origin, Technology and Uses. Iowa Geological Survey Volume XXVIII Annual Reports 1917 and 1918," *Iowa Geological Survey Annual Report*, vol. 28, pp. 47–245, Jan. 1918, <https://doi.org/10.17077/2160-5270.1205>.
- [29] M. T. Lakhari, S. Sohu, I. A. Bhatti, N. Bhatti, S. A. Abbasi, and M. Tarique, "Flexural Performance of Concrete Reinforced by Plastic Fibers," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 3041–3043, Jun. 2018, <https://doi.org/10.48084/etasr.2084>.

Individual Tree Crown Detection Using UAV Orthomosaic

Khairul Nizam Tahar

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA
Shah Alam, Malaysia
khairul0127@uitm.edu.my

Mimie Asmida Asmadin

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA
Shah Alam, Malaysia
advancedphotogrammetry@gmail.com

Saiful Aman Hj Sulaiman

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA
Shah Alam, Malaysia
saifulaman@uitm.edu.my

Nafisah Khalid

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA
Shah Alam, Malaysia
nafisahkhalid@uitm.edu.my

Ahmad Nurhisyam Idris

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA
Shah Alam, Malaysia
ahmad_nurhisyam@uitm.edu.my

Mohammad Hezri Razali

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA
Shah Alam, Malaysia
hezrirazali@uitm.edu.my

Abstract—Unmanned Aerial Vehicles (UAVs) are increasingly used in forestry as they are economical and flexible. This study aims to present the advantages of the drone photogrammetry method in collecting individual tree crowns, as individual tree crown detection could deliver essential ecological and economic information. The referred accuracy for individual tree crown extraction is 79.2%. Only crowns that were clearly visible were selected and manually delineated on the image because the distribution of the true crown size is significantly different from the segmented crowns. The aim of this study is to investigate UAVs orthomosaics in individual tree crown detection. The objectives of this study are to produce the orthomosaic of tree crown extraction mapping using the Pix4D software and analyze the tree crowns using tree crown delineation and the OBIA algorithm. Data processing involves the processing of aerial images using Pix4Dmapper. Automatic tree crown detection involves a tree crown delineation algorithm and OBIA operations to process the tree crown extraction. The crown delineation algorithm and OBIA algorithm operation will be compared to the actual tree crown measurement in terms of diameter and area. The tree crown delineation method obtained a 0.347m mean diameter difference from the actual tree crown diameter, while the OBIA approach obtained 4.98m. The tree crown delineation method obtained 97.26% of the actual tree crown area, while OBIA obtained 91.74%.

Keywords—orthomosaic; unmanned aerial vehicle; digital surface model; digital terrain model; tree crown

I. INTRODUCTION

Individual tree crown extraction parameters deliver essential ecological and economic information to decision makers [1]. The accurate extraction of the tree crown is a

necessary precondition to derive tree position and dimensions such as height, width, and volume. A tree crown is a distinct object with a strong spectral variability [5, 6]. Tree crown can be defined as the top part of the tree which features branches that grow out from the main trunk and support the whole leaves. Tree crown is essential as it strongly links with other tree parameters that are used to estimate the growing stock [2]. Tree crowns can be divided into four classes, namely dominant, codominant, intermediate, and overtopped. Dominant crowns extend above the general level of the canopy. They are the largest and fullest crowns in the stand as they receive full light from above and some light from the sides [3]. Codominant crowns make up the general level of the canopy [4]. They are shorter than the dominant crowns and they receive direct light from above but little or no light from the sides. Intermediate crowns occupy a subordinate position in the canopy. They are narrow in position and shorter than dominant and codominant crowns. They receive some direct light from above but no light from the sides. Overtopped crowns are below the general level of the canopy.

This study presents the advantages of the drone photogrammetry method in collecting individual tree crowns. Unmanned Aerial Vehicles (UAVs) have been increasingly used in forestry as they are economical and flexible. A UAV can be defined as an aircraft that can navigate without a human pilot on board or can fly autonomously through software-controlled flight plans. To take measurements from photographs, a camera is mounted on the UAV and pointed vertically towards the ground. The output of UAV photogrammetry is typically a map or 3D model of a real-world object [7]. Many products can be extracted by aerial imagery,

Corresponding author: Khairul Nizam Tahar

including surface models, orthophotos, 3D building models, contour maps, and volumetric surveys [8]. A drone-based technique for topographic analysis has better accuracy than satellite remote sensing, which has limited resolution and quality depending on various factors, resulting in discontinuities in obtaining correct land information. UAV-derived photogrammetric data are relatively more sensitive, cheaper, and more productive than data from other platforms such as airborne LIDAR systems [9].

Orthomosaic mapping has been used in many commercial sectors such as surveying, mining, construction, and agriculture for its precise measurement and calculations [10, 11]. Orthomosaics are very useful for analyzing the features of a large plot or area of land. An orthomosaic can be divided into two parts, i.e. ortho and mosaic. Ortho is defined as straightening and mosaic is defined as piecing things together [12, 13]. The orthomosaic allows the drones to take many photos from above and stitch these images together into a single scene. To stitch these images together, every photograph must be ortho- and geo-rectified in terms of lens distortion, camera tilt perspective, etc. In other words, an orthomosaic can be defined as an aerial image of an area that is composed of multiple photographs stitched together using photogrammetry, which has been scaled and geographically corrected for accuracy [14, 15]. Generally, orthomosaic derives a much bigger resolution and detail compared to satellite imagery. The automatic individual tree crown delineation in a highly diverse tropical forest using WorldView22 satellite is presented in [16, 17]. The method works in successive steps that involve pre-processing, selection of forested pixels, enhancement of borders, detection of pixels in the crown's borders, correction of shade in large trees, and finally, segmentation of the tree crowns. These methods are done in hierarchical stages to produce accurate individual tree crown delineation. The reported accuracy in individual tree crown extraction is 79.2%. Only crowns that are clearly visible are selected and manually delineated on the image. Authors in [18, 19] identified false positives and false negatives of tree crown extraction. False positives identified a crown where there was no crown, and false negatives did not identify a crown where there was a true crown. This situation happened as the true crown size distribution was significantly different from the segmented crowns. This approach overestimated the size of small crowns, when it was less than 5m.

In the current study, the UAV-based photogrammetric approach is used in individual tree crown extraction as aerial photos have proved useful in the efficient conservation and management of forest inventory. A drone-based approach can provide a more accurate result compared to satellite images. The image quality in terms of spectral and radiometric resolution can help produce tree crown inspection more accurately. The aim of this study is to investigate UAV orthomosaic in individual tree crown detection.

II. MATERIALS AND METHODS

The objectives of this study are to detect tree crowns using a crown delineation algorithm and analyze the individual tree crown result with the actual tree crown. The preliminary study

is the initial exploration of tree crown extraction that involves review and research. A preliminary study involves designing the problem statement, aims, and objectives. Data collection involves the fieldwork observation to extract the tree crown. After data collection, the data will be processed using the necessary software, then, the results are analyzed to determine the accuracy of the outcome. Figure 1 illustrates the workflow of this study in detail.

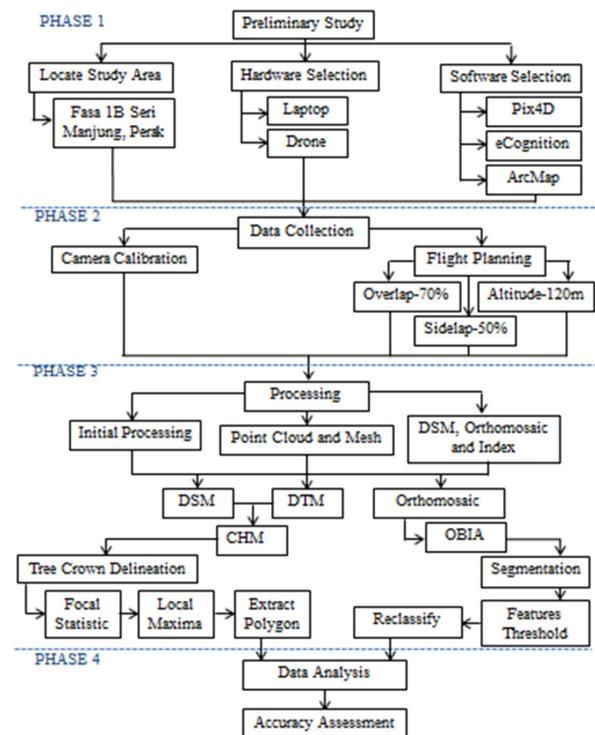


Fig. 1. Research methodology.

The study area is located in Fasa 1b Seri Manjung, Perak, Malaysia. The coordinates of the survey area are $4^{\circ}12'12.14''N$ and $100^{\circ}40'21.41''E$. The study area covers 54.6243 acres. The selection of the study area focused on the area that contains trees. The types of trees in the study area were not a concern as tree crown detection can be done on any kind of tree.

A. Flight Planning



Fig. 2. Flight route planning in Altizure.

For flight planning, altitude, overlap, and side lap should be considered. The altitude in this flight planning was 120m. The overlap was 70%, with a side lap of approximately 50%. The software used in designing the flight plan during data collection is Altizure. The application automatically generates the fly paths over the study area. The capturing of data is designed for generating a 2D orthomosaic as it captures vertical photographs (Figure 2).

B. Image Processing

In Pix4Dmapper, the initial processing includes computing keypoints on the images. The keypoints will then be used to find matches between the images. This initial processing also allows the software to run an automatic aerial triangulation and bundle block adjustment. After that, point cloud and mesh take place. This processing option allows the setting of parameters for point cloud densification. The image scale was set to half image size, and multiscale was checked to compute additional 3D points on multiple image scales. The point density chosen was "optimal". For the minimum number of matches, the default value "3" was chosen to allow each 3D point to correctly re-project at least three images. The final process is the Digital Surface Model (DSM), the orthomosaic, and the index operation. The processing options allow changes of the desired outputs for DSM and orthomosaic generation to be made. For this step, DSM and orthomosaics are automatically generated. Canopy Height Model (CHM) represents the height between the ground and the top of the object above the ground. The CHM is used for automatic tree crown detection. The CHM is generated by subtracting the DTM from the DSM.

C. Tree Crown Delineation Algorithm

The tree crown delineation algorithm was implemented on ArcMap software. The tools involved in the automatic tree crown detection processing are the focal statistics tool, the raster calculator, the slope tool, and the export polygon operation. To generate local maxima, the focal statistics tool should be used first. The focal statistics tool performs a neighborhood process that calculates an output raster where the value for each output cell is a function of the values of all the input cells that are in a specified neighborhood around that location. In the window, CHM was the input raster. The neighborhood chosen had a "rectangle" shape, with a neighborhood setting of 3×3 (height×width). The option selected for statistic type was "mean". Local maxima are individual pixels that represent high surfaces. To compute local maxima, focal statistics were needed. The local maxima operation was operated using a raster calculator. The slope operation allows the creation of a tree crown outline. This operation involves CHM as the input raster. The extracted tree crown will then be converted to polygon.

D. Object-Based Image Analysis Algorithm Technique

Object-Based Image Analysis (OBIA) is the most accurate relevant technique. OBIA produces higher accuracy than pixel-to-pixel analysis. The OBIA technique used in this research is multiresolution segmentation. The segmented orthomosaic was then used for thresholding features indices and classification in order to detect tree crowns. This algorithm allowed the identification of one pixel size for single image objects. It also

merged them with neighboring pixels based on relative homogeneity. For this research, multiresolution segmentation was performed against red, green, blue, and near-infra-red bands. The scale parameter was set to 90 and the image layer weight for near-infra-red was set to 2. For shape and compactness, the value was set at 0.3 and 0.7 respectively. The thresholding features indices used were NDVI and NPV. NDVI threshold was performed for tree detection, while NPV threshold was performed for non-tree detection. The arithmetic threshold for NDVI involved the near-infra-red and red bands, while NPV involved green and blue bands. Reclassification of tree and non-tree indices was performed next. Classification was performed to discriminate tree from non-tree features. The classification was performed using the assign class tool. For tree classification, the condition of thresholding was set at $NDVI > 0.08$ for tree and $NPV > 0$ and for non-tree.

III. RESULTS AND DISCUSSION

Aerial image processing was conducted to generate the orthomosaic. The orthomosaic generation was conducted with the Pix4D software. There were three main processing stages, namely initial processing, point cloud and mesh, and DSM, orthomosaic and index. The orthomosaic was generated for individual tree crown detection analysis. Figure 3(a) shows the orthomosaic mapping of Fasa 1B Seri Manjung, Perak. Orthomosaic products are free from any distortion and can be used for true measurement [20]. Figure 3(b) shows the DSM and Figure 3(c) shows the Digital Terrain Model (DTM) of the study area. The DSM includes all objects like buildings, trees, and cars to generate the earth's surface. DTM, on the other hand, excludes all objects on the earth's surface and represents the bare earth.

A. Tree Crown Delineation (TCD)

The CHM represents the height between the ground and the top of the objects above it. Figure 4(a) shows the CHM model on the orthomosaic. The CHM can be generated by subtracting the DTM from the DSM. In Figure 4(a), the white cell indicates trees in the CHM model that are located along the road. Figure 4(b) shows the output of the focal statistics operation which was used for generating the local maxima. The focal statistics tool performs a neighborhood process that calculates an output raster where the value for each output cell is a function of the values of all the input cells that are in a specified neighborhood around that location. In Figure 4(b), the white cell shows tree features along the road. Local maxima are individual pixels that represent high surfaces. Local maxima were computed to find the tree tops. The output of the local maxima was then converted into points. Figure 4(c) shows the local maxima output and converted points. Figure 5(a) shows the tree crown outline after the extraction process using tree crown delineation algorithm. The results were obtained from the CHM model as input raster using the slope tool. The tree crown is outlined in red. It can be seen that the outlined tree crown is along the road side. In Figure 5(b), the left image shows the result of tree crown extraction using slope operation. The right image shows the tree crown outline overlays with the orthomosaic at 40% transparency. Figure 5(c) shows the final result for this approach on the orthomosaic. The tree crowns can be seen in red line.

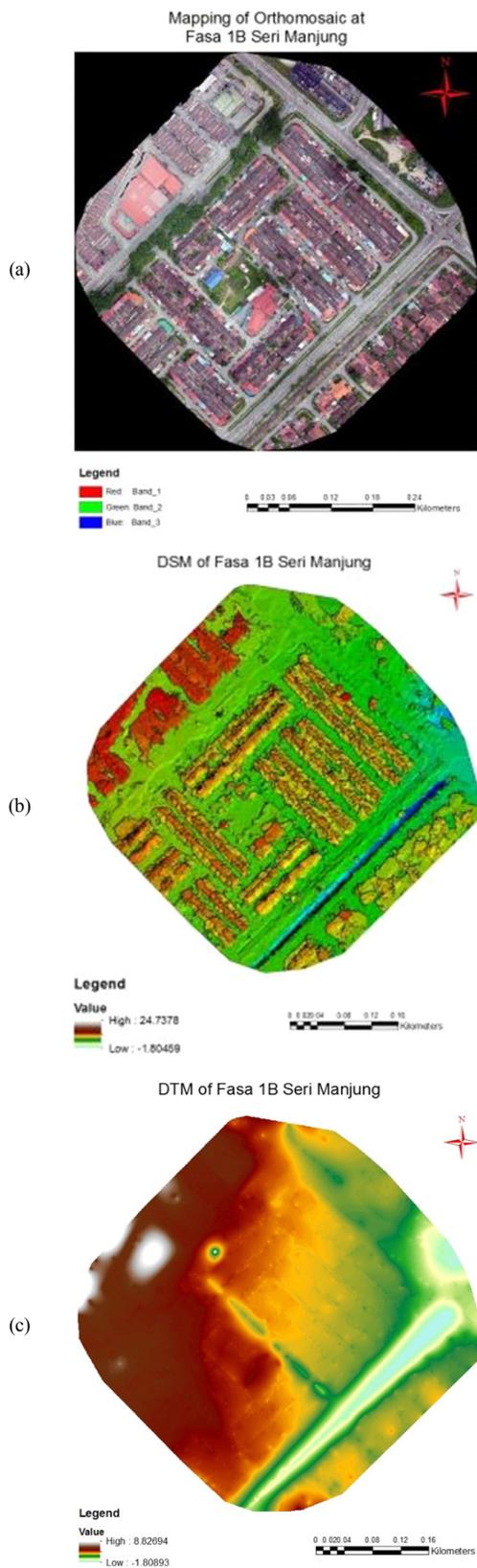


Fig. 3. Results. (a) Orthomosaic, (b) DSM, (c) DTM.

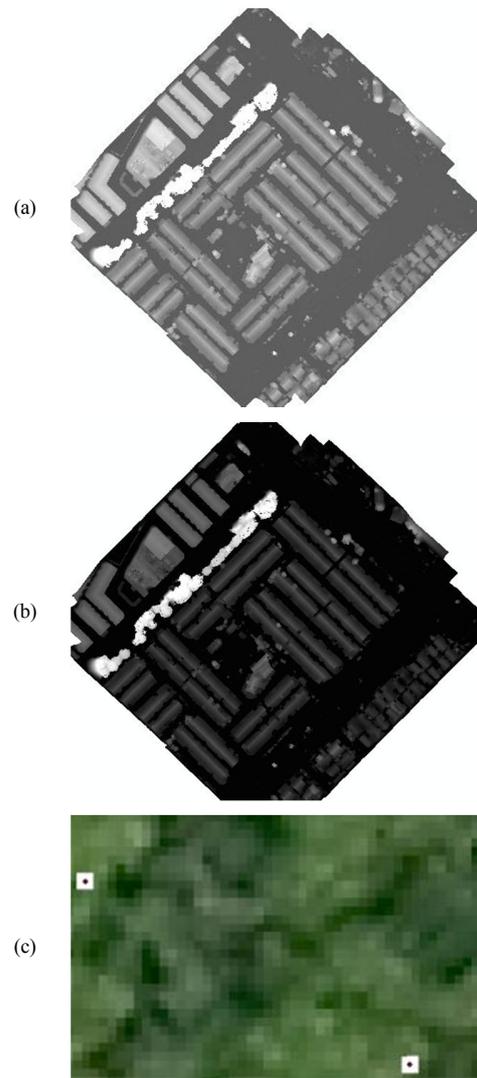


Fig. 4. TCD results. (a) CHM model, (b) focal statistic, (c) converted local maxima in points (right).

B. Object-Based Image Analysis Algorithm

Multiresolution segmentation operation could turn each pixel segmented into single image objects. Figure 6 shows the result of the multiresolution operation. The operation approached the scale parameter at 90 with shape and compactness of 0.3 and 0.7 respectively. The thresholding feature indices involved two features indices, which are NDVI and NPV (Figure 7(a)). NDVI indices are used for tree detection while NPV indices are used for non-tree detection. Both indices are computed to discriminate trees from non-trees. The threshold of NDVI was set at >0.08 for tree detection, and the threshold of NPV was set at >0 for non-tree detection. In Figure 7(b), the detected trees can be seen as green lines, and detected non-trees as red lines. From the final result, it can be said that NDVI indices are not suitable for this process as they detect other green areas like grass despite extracting only trees.

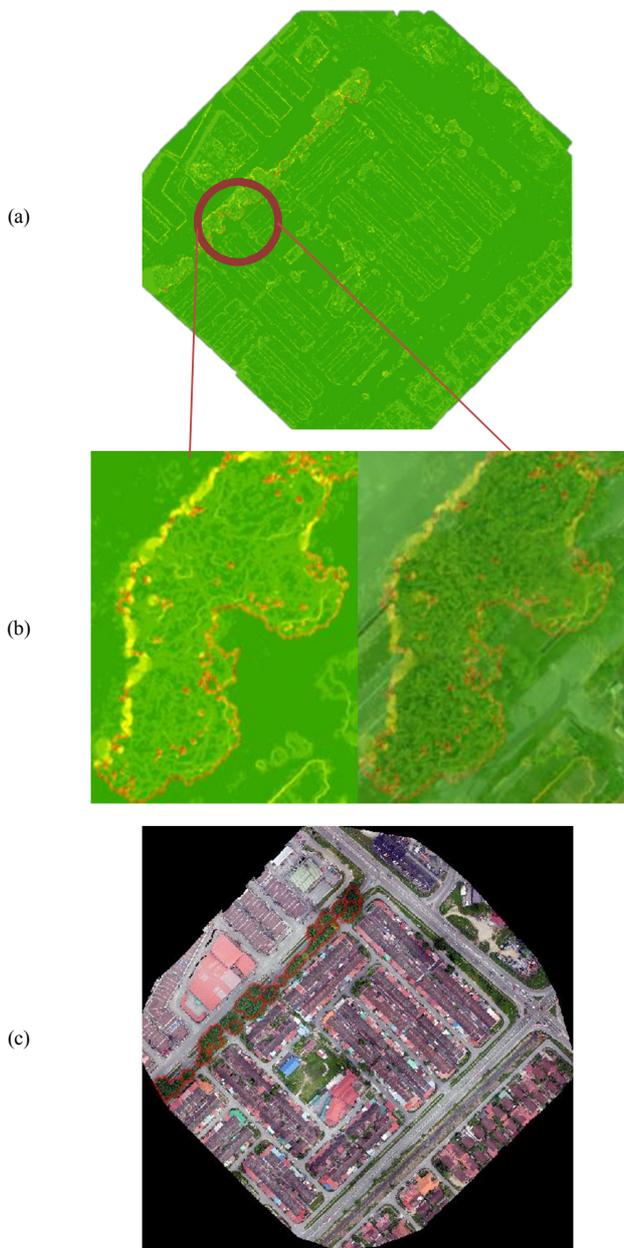


Fig. 5. Crown results. (a) Tree crown outline, (b) tree crown outline (left) and tree crown outline overlay with orthomosaic (right), (c) final tree crown delineation overlay with the orthomosaic.

The final result of the OBIA method is affected by a few conditions. The classification of tree and non-tree is affected by the thresholding features indices, which are NDVI and NPV (Figure 7(c)). Both thresholds are suitable as both detect features into their own classes, which are tree and non-tree. The scale parameter during segmentation operation also affects the final result. The scale parameter should be set below 90 as the scale parameter value affects the object classification.

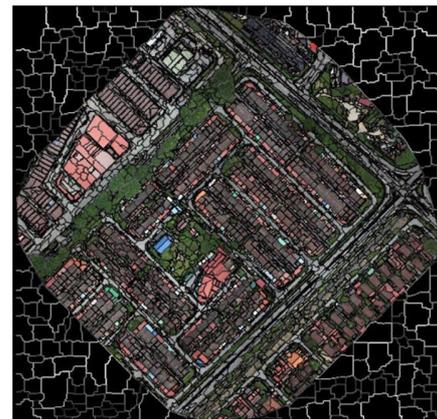


Fig. 6. The result of multiresolution segmentation.

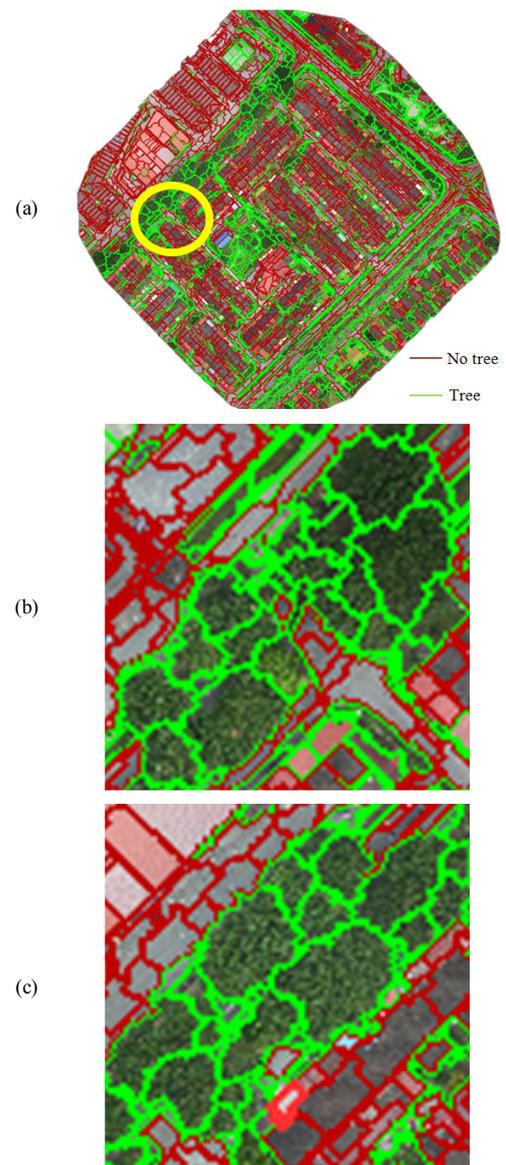
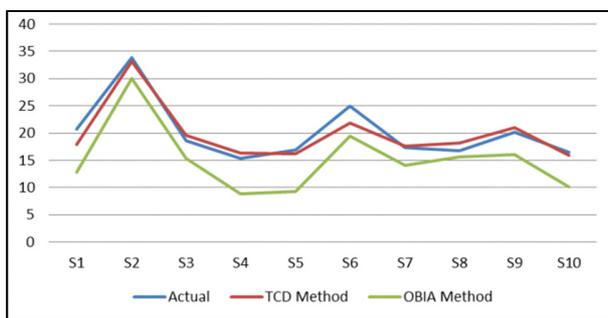


Fig. 7. OBIA result. (a) Thresholding, (b) reclassification, (c) final result.

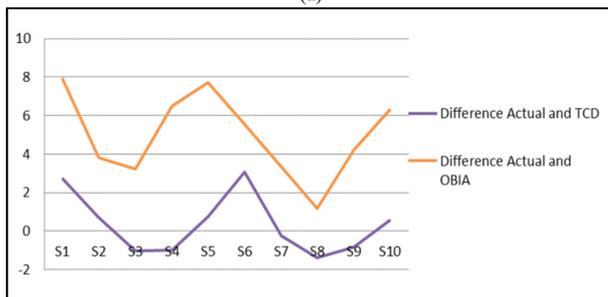
C. Accuracy Assessment

1) Diameter Assessment

Accuracy assessment operation involves analyzing the comparison between the actual tree crown, the tree crown delineation algorithm approach, and the OBIA algorithm approach in terms of diameter and area. The comparison was conducted based on ten samples. The comparison results were then converted into a line graph. In Figure 8(a), the tree crown delineation method shows small differences of actual tree crown diameter compared to the OBIA approach, which shows big differences from the actual tree crown diameter. All the 10 samples of the TCD method in red lines are closer to the actual tree crown diameter measurement in blue lines compared to the OBIA method in green lines.



(a)



(b)

Fig. 8. Tree crown diameter. (a) Measurements of actual, tree crown delineation, and OBIA, (b) difference between actual and tree crown delineation and actual and OBIA.

Figure 8(b) shows the tree crown diameter comparison between the actual tree crown with TCD method and the actual tree crown with the OBIA method. The difference in diameter between the actual tree crown and the TCD method had a mean value of 0.347m over 10 tree samples, while the OBIA method obtained a mean value of 4.98m. The difference between the actual tree crown and TCD in purple lines is not big. The difference between the actual crown and the OBIA method is slightly bigger.

2) Area Assessment

Figure 9(a) shows the small differences from the actual tree crown area compared to the OBIA approach. It can be seen that the TCD method is close to the actual tree crown. Figure 9(b) shows the comparison between the actual tree crown with the TCD method and between the actual tree crown with the OBIA

method. The TCD and the OBIA methods obtained the 97.25% and the 91.74% of the tree crown area respectively. The difference between the actual crown and the TCD is not big. The difference between the actual crown and the OBIA method is bigger.

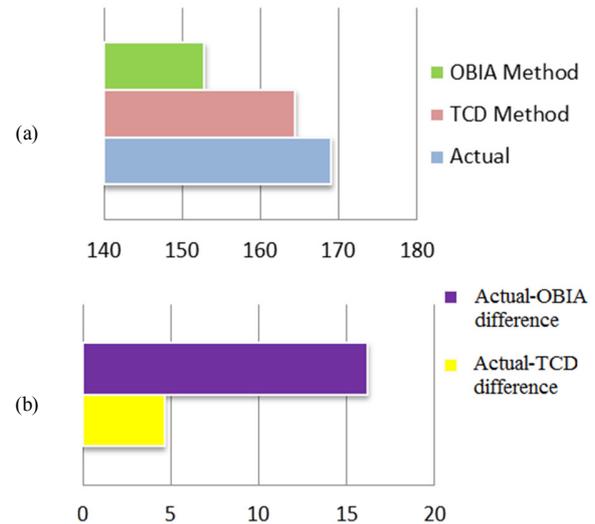


Fig. 9. Tree crown area. (a) Measurements of actual, tree crown delineation, and OBIA, (b) comparison.

This study has almost similar results with previous relevant studies. Authors in [21] correctly classified 95% of individual trees, authors in [22] also successfully detect 90% of individual trees using fuzzy system, and authors in [23] used the deep learning approach to detect individual trees and the result of tree detection was about 90%. Therefore, the results achieved in this study have an improvement of about 2.25% by using the TCD method.

IV. CONCLUSIONS

The differences between the actual crown and the TCD and OBIA results are affected due to a few conditions in terms of diameter and area measurements. The tree crown measurement differences may be due to the number of overlapping images. Based on the aerial image processing, the trees along the road are located under less than five overlapping images. This may affect the results of the tree crown diameter and area measurements between the actual crown, the TCD method, and the OBIA method. The process of an automatic tree crown that uses two different software may also affect the final result of both methods compared to the actual tree crown measurement. The OBIA method result may be due to the scale parameter value during segmentation operation.

ACKNOWLEDGMENT

Faculty of Architecture, Planning, and Surveying Universiti Teknologi MARA (UiTM), Research Management Institute (RMi) and Ministry of Higher Education (MOHE) are greatly acknowledged for providing the fund BESTARI 600-RMC/DANA 5/3/BESTARI (TD) (006/2019).

REFERENCES

- [1] S. Krause, T. Sanders, J.-P. Mund, and K. Greve, "UAV-Based Photogrammetric Tree Height Measurement for Intensive Forest Monitoring," *Remote Sensing*, vol. 11, no. 7, Mar. 2019, Art. no. 758, <https://doi.org/10.3390/rs11070758>.
- [2] M. Immitzer, F. Vuolo, and C. Atzberger, "First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe," *Remote Sensing*, vol. 8, no. 3, Mar. 2016, Art. no. 166, <https://doi.org/10.3390/rs8030166>.
- [3] A. Rajput, A. Hussain, F. Akhtar, Z. H. Khand, and H. Magsi, "A Versatile Decentralized 3D Volumetric Fusion for On-line Reconstruction," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6584–6588, Dec. 2020, <https://doi.org/10.48084/etasr.3838>.
- [4] S. Baena, J. Moat, O. Whaley, and D. S. Boyd, "Identifying species from the air: UAVs and the very high resolution challenge for plant conservation," *PLOS ONE*, vol. 12, no. 11, 2017, Art. no. e0188714, <https://doi.org/10.1371/journal.pone.0188714>.
- [5] J. Maschler, C. Atzberger, and M. Immitzer, "Individual Tree Crown Segmentation and Classification of 13 Tree Species Using Airborne Hyperspectral Data," *Remote Sensing*, vol. 10, no. 8, Aug. 2018, Art. no. 1218, <https://doi.org/10.3390/rs10081218>.
- [6] C. Torresan *et al.*, "Forestry applications of UAVs in Europe: a review," *International Journal of Remote Sensing*, vol. 38, no. 8–10, pp. 2427–2447, May 2017, <https://doi.org/10.1080/01431161.2016.1252477>.
- [7] S. S. Abdullah, K. N. Tahar, M. F. A. Rashid, and M. A. Osman, "Capabilities of UAV-Based Watershed Segmentation Method for Estimating Tree Crown: A Case Study of Oil Palm Tree," *IOP Conference Series: Earth and Environmental Science*, vol. 385, Nov. 2019, Art. no. 012015, <https://doi.org/10.1088/1755-1315/385/1/012015>.
- [8] N. Liba and J. Berg-Jürgens, "Accuracy of Orthomosaic Generated by Different Methods in Example of UAV Platform MUST Q," *IOP Conference Series: Materials Science and Engineering*, vol. 96, Nov. 2015, Art. no. 012041, <https://doi.org/10.1088/1757-899X/96/1/012041>.
- [9] A. Grznárová, M. Mokroš, P. Surový, M. Slavík, M. Pondelík, and J. Merganič, "The Crown Diameter Estimation from Fixed Wing Type of Uav Imagery," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4213, pp. 337–341, Jun. 2019, <https://doi.org/10.5194/isprs-archives-XLII-2-W13-337-2019>.
- [10] F. Bayat, H. Arefi, and F. Alidoost, "Individual Tree Detection and Determination of Tree Parameters Using Uav-Based LIDAR Data," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4218, pp. 179–182, Oct. 2019, <https://doi.org/10.5194/isprs-archives-XLII-4-W18-179-2019>.
- [11] H. Yao, R. Qin, and X. Chen, "Unmanned Aerial Vehicle for Remote Sensing Applications—A Review," *Remote Sensing*, vol. 11, no. 12, p. 1443, Jan. 2019, <https://doi.org/10.3390/rs11121443>.
- [12] C. Suwanprasit, J. Strobl, and J. Adamczyk, "Extraction of Complex Plantations from VHR Imagery using OBIA Techniques," *International Journal of Geoinformatics*, vol. 11, no. 1, pp. 73–83, Mar. 2015.
- [13] S. H. Adil, M. Ebrahim, S. S. A. Ali, and K. Raza, "Performance Analysis of Duplicate Record Detection Techniques," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4755–4758, Oct. 2019, <https://doi.org/10.48084/etasr.3036>.
- [14] M. Kamal, S. Phinn, and K. Johansen, "Object-Based Approach for Multi-Scale Mangrove Composition Mapping Using Multi-Resolution Image Datasets," *Remote Sensing*, vol. 7, no. 4, pp. 4753–4783, Apr. 2015, <https://doi.org/10.3390/rs70404753>.
- [15] Z. Zhang, A. Kazakova, L. M. Moskal, and D. M. Styers, "Object-Based Tree Species Classification in Urban Ecosystems Using LiDAR and Hyperspectral Data," *Forests*, vol. 7, no. 6, Jun. 2016, Art. no. 122, <https://doi.org/10.3390/f7060122>.
- [16] M. Brach, J. Chan, and P. Szymański, "Accuracy assessment of different photogrammetric software for processing data from low-cost UAV platforms in forest conditions," *iForest - Biogeosciences and Forestry*, vol. 12, no. 5, pp. 435–441, Sep. 2019, <https://doi.org/10.3832/ifor2986-012>.
- [17] F. H. Wagner *et al.*, "Individual tree crown delineation in a highly diverse tropical forest using very high resolution satellite images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 362–377, Nov. 2018, <https://doi.org/10.1016/j.isprsjprs.2018.09.013>.
- [18] K. Johansen, T. Raharjo, and M. McCabe, "Using Multi-Spectral UAV Imagery to Extract Tree Crop Structural Properties and Assess Pruning Effects," *Remote Sensing*, vol. 10, no. 6, Jun. 2018, Art. no. 854, <https://doi.org/10.3390/rs10060854>.
- [19] Z. Zhen, L. Quackenbush, and L. Zhang, "Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data," *Remote Sensing*, vol. 8, no. 4, Apr. 2016, Art. no. 333, <https://doi.org/10.3390/rs8040333>.
- [20] N. Che Mat and K. N. Tahar, "Surf Zone Mapping Using Multirotor Unmanned Aerial Vehicle Imagery," *International Journal of Optics*, vol. 2019, May 2019, Art. no. e2987301, <https://doi.org/10.1155/2019/2987301>.
- [21] J. Holmgren and Å. Persson, "Identifying species of individual trees using airborne laser scanner," *Remote Sensing of Environment*, vol. 90, no. 4, pp. 415–423, Apr. 2004, [https://doi.org/10.1016/S0034-4257\(03\)00140-8](https://doi.org/10.1016/S0034-4257(03)00140-8).
- [22] M. A. A. Khameneh, "Tree Detection and Species Identification using LiDAR Data," M. S. thesis, Royal Institute of Technology, Stockholm, Sweden, 2013.
- [23] A.-I. Pleșoianu, M.-S. Stupariu, I. Șandric, I. Pătru-Stupariu, and L. Drăgăuț, "Individual Tree-Crown Detection and Species Classification in Very High-Resolution Remote Sensing Imagery Using a Deep Learning Ensemble Model," *Remote Sensing*, vol. 12, no. 15, Jan. 2020, Art. no. 2426, <https://doi.org/10.3390/rs12152426>.

AUTHORS PROFILE

Mimie Asmida Asmadin is a graduate student in the Centre of Studies for Surveying Science and Geomatics, Universiti Teknologi MARA (UiTM), Malaysia. His research interests include Unmanned Aerial Vehicle (UAV) for mapping, photogrammetry and modeling.

Khairul Nizam Tahar is the Head of the Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, UiTM, Malaysia. His major research interests include photogrammetry and close range photogrammetry especially on UAV applications. He has published 1 book, 1 book chapter, and over 70 articles in journals and conference proceedings. In the last 5 years, he received a number of research grants from the University and the Ministry. He has been appointed as the chief editor and an editorial board member of several international journals. He is also a keynote and invited speaker at national and international conferences and short courses.

Saiful Aman Hj Sulaiman is a Senior Lecturer at the Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, UiTM, Malaysia. His major research interests include GNSS, geodesy, cadastra, and surveying.

Nafisah Khalid is a Senior Lecturer at the Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, UiTM, Malaysia. His major research interests include LiDAR, remote sensing, and geospatial.

Ahmad Norhisyam Idris is a Lecturer at the Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, UiTM, Malaysia. His major research interests include GNSS, geodesy, and surveying.

Mohammad Hezri Razali is a Lecturer at the Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, UiTM, Malaysia. His major research interests include cadastra, surveying, and ground penetrating radars.

A Novel Modeling and Control Design of the Current-Fed Dual Active Bridge Converter under DPDPS Modulation

Phuong Vu

School of Electrical Engineering
Hanoi University of Science and
Technology
Hanoi, Vietnam
phuong.vuhoang@hust.edu.vn

Do Tuan Anh

School of Electrical Engineering
Hanoi University of Science and
Technology
Hanoi, Vietnam
tuananhdo.hust@gmail.com

Hoang Duc Chinh

School of Electrical Engineering
Hanoi University of Science and
Technology
Hanoi, Vietnam
chinh.hoangduc@hust.edu.vn

Abstract—This paper proposes a novel control design for a Current-Fed Dual Active Bridge (CFDAB) converter in boost mode. The Double PWM plus Double Phase Shifted (DPDPS) modulation is applied to the converter due to its considerable merits. A small-signal model is developed to control the output voltage stably in boost mode. Simulations of the control design for the CFDAB converter were conducted to verify the proposed model. The results show that the system can achieve high performance, not only in the dynamic response but also in the steady-state.

Keywords—Current-Fed Dual Active Bridge (CFDAB) converter; small-signal model; Double PWM plus Double Phase Shifted (DPDPS) modulation

I. INTRODUCTION

Nowadays, the growth in renewable power systems urges researchers and engineers to solve the problems of managing integrated storages and exchange powers to enhance overall system efficiency. In order to connect the battery at low voltage to the DC link at high voltage in a storage system, the utilized DC/DC converter needs to have a high gain voltage factor and an ability of bidirectional transferring power. The DC/DC converters can be classified as isolated and non-isolated. Between these two, the isolated DC/DC converter is highly recommended for its high reliability. It is able to eliminate current leakage in the system. This undesired current is the cause of EMI, additional loss, and unsafe installation and operation [1]. Another advantage of the isolated DC/DC converter is that the flexible gain ratio can be obtained by using a high frequency transformer [2, 3]. The Dual Active Bridge (DAB) converter belongs to the isolated DC/DC type and is well-known for its advantages of high frequency operating ability, inherit zero voltage switching, and bidirectional power flow [4-6]. The structure of the DAB can have modified flexibly with different power sources [7-8]. The DAB converter is divided into two categories: Voltage-Fed (VFDAB) and Current-Fed (CFDAB). When using the VFDAB structure in a variety of applications [9-11], the voltage source is directly supplied to the converter, which produces high current ripple

and becomes large and costly. Besides, the VFDAB needs a bulky capacitor in series-connection with the primary coil to avoid the flux saturation of the transformer [11]. Therefore, the CFDAB structure is used to deal with these problems, the input interleaved boost inductors in CFDAB help to decrease the current ripple significantly, and ZVS can be achieved for all switches over a wide range of load [9-12].

On the other hand, a small-signal model needs to be built for the accurately controlling problem. Naturally clamped CFDAB modeling structure has been introduced in [13, 14]. Besides, a state-space model under SPSPWM modulation for interleaved boost CFDAB structure is mentioned in [15]. In the current research, DPDPS modulation technique in comparison with SPSPWM modulation [9] is applied, and a small-signal model to regulate the output voltage is built up in boost mode. The simulation results show the effectiveness of the proposed method.

II. RESEARCH METHOD

A. Operation Principles and Control Structure

1) Operation Principles

The applied structure of CFDAB converter in this research is presented in Figure 1, which consists of 2 main parts: the interleaved boost part and the dual active full-bridge part. In the interleaved boost part, two DC inductors are considered as two current sources. The inductor L_{dc1} is combined with the left leg containing the switch pair Q_1, Q_{1a} to create the first boost converter, while the right leg containing the Q_2, Q_{2a} switch pair is combined with the inductor L_{dc2} to create the second boost converter. These boost converters are 180° phase-shifted in order to compose an interleaved boost part, in which the boost voltage is kept by the clamp capacitor C_c .

On the other hand, the dual active full-bridge part consists of two H-bridge modules in two sides of an isolated high frequency transformer with the turn ratio of $N: 1$.

Corresponding author: Phuong Vu

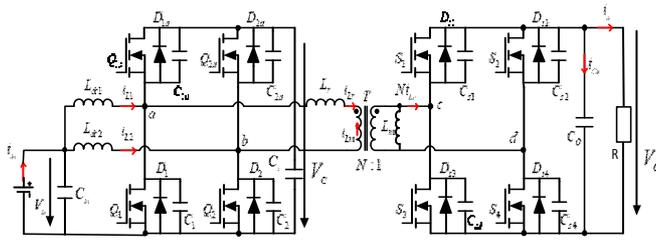


Fig. 1. Structure of the CFDAB converter.

The voltages of the clamp capacitor and the output capacitor are assumed as the Low-Voltage Side (LVS) and High-Voltage Side (HVS). The bidirectional transfer power between the LVS, the HVS, and the power flow are determined by the phase shifted angle. The AC inductor L_r , which is the sum of the primary-referred transformer leakage inductor, which can be considered as a power link between the two sides of the converter.

2) Modulation Strategy

Due to the high amount of switch devices and inductors, the modulation strategy for the converter needs to be considered before designing the control loop. To handle the voltage ratio variation problems and to minimize the conduction loss in power transfer stages, the PWM plus Phase Shifted (PPS) is introduced while the duty cycle for HVS switches is equal to 50% and the duty ratio for the main LVS switches Q_1, Q_2 is a variable D . However, in the non-power transfer stage of PPS, the circulation loss is significant with high leak current spike. An additional phase shifted in Double PWM plus Double Phase Shifted (DPDPS) method [9], which equals to $(2D-1) \times T_s / 2$, is applied to eliminate the leak current spike. Figure 2 illustrates the modulation rule of the CFDAB converter, where the phase shifted between $Q_1 - Q_2, Q_2 - S_1$ and $S_1 - S_4$ are $180^\circ, \varphi$ and φ_s respectively. The value of φ_E determines the power flow and direction, where $\varphi_E > 0$ represents the operation boost mode with power flow from LVS to HVS and vice versa. By utilizing DPDPS modulation, 8 operation modes in one switching period along with leakage current, the switching LVS voltage and the transformer voltage waveform v_{cd} are displayed in Figure 3 [9].

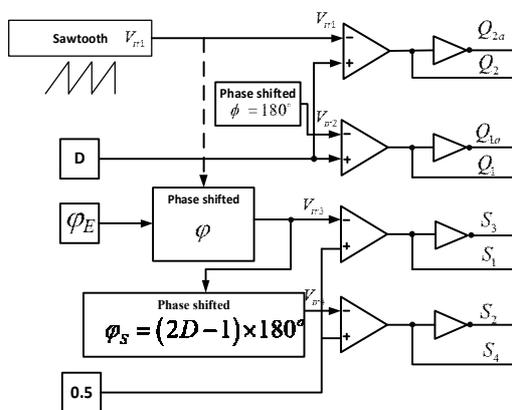


Fig. 2. DPDPS modulation technique for the CFDAB converter.

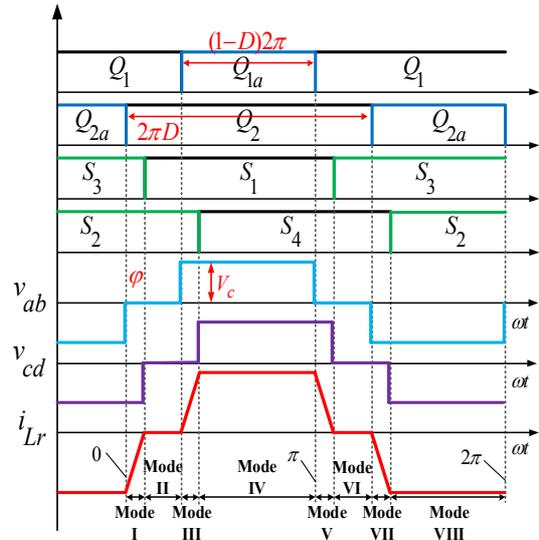


Fig. 3. Pulse patterns, transformer voltage, and leakage current waveforms.

3) Control Structure

For CFDAB converter structure, the transfer power is related to the clamp voltage V_c and the phase shifted φ_E between the two H-bridges. Therefore, the control structure in boost mode has to control the voltage of the clamp capacitor and the output capacitor simultaneously. There are two pairs of control variables: $V_c - d$ and $V_o - \varphi$, as presented in Figure 4.

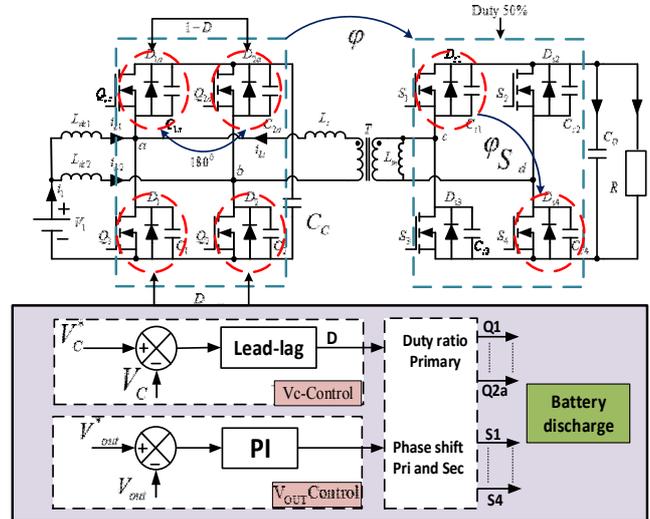


Fig. 4. Control structure in boost mode.

B. Small Signal Modeling

It is required to determine the transfer functions of the two pairs of control variables mentioned above so that the control scheme in Figure 4 can be implemented. The transfer function perturbations between clamp voltage \hat{v}_c and duty ratio \hat{d} is depicted in (1) [9]:

$$G_{v_c-d} = \frac{\hat{v}_c}{\hat{d}} = \frac{2V_c\phi - \omega L_r(I_{L1} + I_{L2})}{\omega L_r C_d s + 2\phi(1-D) - \frac{\phi^2}{2\pi}} \quad (1)$$

In this paper, the small signal model along with the closed loop controller is proposed to regulate the output voltage of the converter under DPDPS modulation. In the secondary side, the output capacitor is charged and discharged in power transfer and non-power transfer stage respectively. From the switching sequence in Figure 3, the time interval of each switching mode is given by:

$$\begin{cases} d_1 = d_3 = d_5 = d_7 = \frac{\phi}{2\pi} \\ d_4 = d_8 = 1 - d, \\ d_2 = d_6 = \frac{2d-1}{2} - \frac{\phi}{2\pi} \end{cases} \quad (2)$$

The relationship between small signals of variables is presented in Table I for 8 modes in one DPDPS switching period. The inductors and capacitors are considered as ideal, i.e. the internal resistors are neglected. The continuous-time equation of output voltage is defined as:

$$\begin{aligned} C_o \frac{dv_o}{dt} &= N(d_4 - d_8) \frac{\langle v_c \rangle}{\omega L_r} \phi \\ &+ (-d_1 + d_2 + d_3 - d_4 - d_5 + d_6 + d_7 - d_8) \frac{v_o}{R} \quad (3) \\ &+ Nd_5 \frac{\langle v_c \rangle}{\omega L_r} (\pi - \theta_5 + \phi) - Nd_1 \frac{\langle v_c \rangle}{\omega L_r} (\theta_1 - \phi) \end{aligned}$$

Using the average model:

$$\theta_7 = 2d\pi + \frac{\phi}{2}, \langle v_c \rangle = v_c, \theta_3 = (2d-1)\pi + \frac{\phi}{2}$$

we can obtain:

$$\begin{aligned} C_o \frac{dv_o}{dt} &= N(d_4 + d_8) \frac{\langle v_c \rangle}{\omega L_r} \phi + N(d_1 + d_5) \frac{\langle v_c \rangle}{\omega L_r} \frac{\phi}{2} \quad (4) \\ &+ (-d_1 + d_2 + d_3 - d_4 - d_5 + d_6 + d_7 - d_8) \frac{v_o}{R} \end{aligned}$$

Replacing (2) to (4) we get:

$$C_o \frac{dv_o}{dt} = (4d-3) \frac{v_o}{R} + N \left(\frac{-\phi^2}{2\pi} + 2\phi(1-d) \right) \frac{v_c}{\omega L_r} \quad (5)$$

Using the perturbations of duty ratio, phase-shifted angle output voltage, and clamp voltage we have:

$$\begin{aligned} C_o \frac{d(V_o + \hat{v}_o)}{dt} &= (4(D + \hat{d}) - 3) \frac{(V_o + \hat{v}_o)}{R} + \\ &+ N \left(\frac{-(\Phi + \hat{\phi})^2}{2\pi} + 2(\Phi + \hat{\phi})(1 - (D + \hat{d})) \right) \frac{V_c + \hat{v}_c}{\omega L_r} \end{aligned}$$

Discarding the steady-state value and the second order perturbations and assuming that $\hat{d} = \hat{v}_{cc} = 0$ we get:

$$C_o \frac{dv_o}{dt} = (4d-3) \frac{v_o}{R} + N \left(\frac{-\phi^2}{2\pi} + 2\phi(1-d) \right) \frac{v_c}{\omega L_r} \quad (7)$$

By Laplace transformation, the transfer function from \hat{v}_o to $\hat{\phi}$ in the frequency domain can be expressed as:

$$G_{v_o-\phi} = \frac{NRV_c(2-2D-\frac{\Phi}{\pi})}{\omega RL_r C_o s - \omega L_r(4D-3)} = \frac{A}{Bs+C} \quad (8)$$

TABLE I. SMALL SIGNAL MODEL IN EACH SWITCHING MODE

Mode	Continuous-time equations
Mode I [0, φ]	$C_o \frac{dv_o}{dt} = -\frac{v_o}{R} - Ni_{Lr}, i_{Lr} = \frac{\langle v_c \rangle_{Ts}}{\omega L_r} (\theta - \phi)$
Mode II [φ, (2d-1)π]	$C_o \frac{dv_o}{dt} = \frac{v_o}{R}, i_{Lr} = 0$
Mode III [(2d-1)π, (2d-1)π + φ]	$C_o \frac{dv_o}{dt} = \frac{v_o}{R}, i_{Lr} = \frac{\langle v_c \rangle_{Ts}}{\omega L_r} [\theta - (2d-1)\pi]$
Mode IV [(2d-1)π + φ, π]	$C_o \frac{dv_o}{dt} = -\frac{v_o}{R} + Ni_{Lr}, i_{Lr} = \frac{\langle v_c \rangle_{Ts}}{\omega L_r} \phi$
Mode V [π, π + φ]	$C_o \frac{dv_o}{dt} = -\frac{v_o}{R} + Ni_{Lr}, i_{Lr} = \frac{\langle v_c \rangle_{Ts}}{\omega L_r} (\pi - \theta + \phi)$
Mode VI [π + φ, 2dπ]	$C_o \frac{dv_o}{dt} = \frac{v_o}{R}, i_{Lr} = 0$
Mode VII [2πd, 2πd + φ]	$C_o \frac{dv_o}{dt} = \frac{v_o}{R}, i_{Lr} = -\frac{\langle v_c \rangle_{Ts}}{\omega L_r} [\theta - 2d\pi]$
Mode VIII [2πd + φ, 2π]	$C_o \frac{dv_o}{dt} = -\frac{v_o}{R} - Ni_{Lr}, i_{Lr} = -\frac{\langle v_c \rangle_{Ts}}{\omega L_r} \phi$

C. Simulation Verification

The control design with the proposed method is firstly verified by simulation in Matlab/Simulink (Figure 5). The specifications of the converter are given in Table II and the transfer function in frequency domain between \hat{v}_o and $\hat{\phi}$ is turned into:

$$G_{v_o-\phi} = \frac{\hat{v}_o}{\hat{\phi}} = \frac{A}{Bs+C} = \frac{1.643 \times 10^4}{0.3213s + 4.7} \quad (9)$$

A PI controller is applied to regulate the output voltage:

$$G_c = K_p \left(1 + \frac{1}{T_i s} \right) \quad (10)$$

where T_i is equal to the pole of $G_{v_o-\phi}$: $T_i = \frac{B}{C} = 0.068$.

The output voltage close loop transfer can be written as:

$$G_k = \frac{1}{T_k s + 1} \quad (11)$$

with $T_k = \frac{B}{A * K_p}$. By choosing $T_k = 5 * T_s = 10^{-4}$, $K_p = 0.276$.

The simulation scenario consists of 4 steps: the clamp capacitor is charged from 0 to 0.2s, at 0.2s the V_c controller starts and the LVS switches are enabled, at 0.25s the V_o controller starts and

in the final step, the load changes suddenly to the nominal value. Figures 6, 7 show the simulation results of clamp and output voltage respectively. As can be observed, in the start-up period, the voltages of the two capacitors are charged and controlled to increase gradually as well as track the reference values with negligible error and no over-shoot.

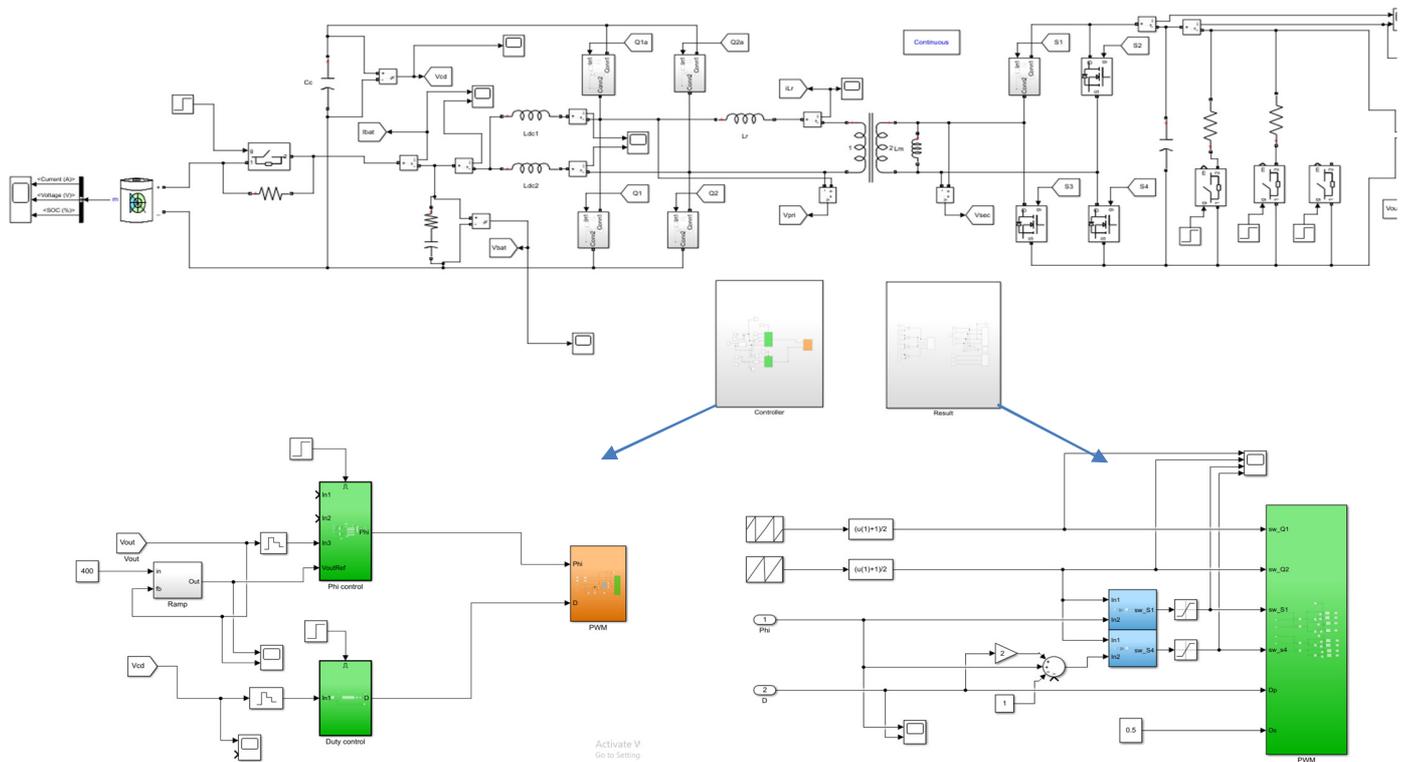


Fig. 5. Simulink model for the control of the CFDAB converter in boost mode.

TABLE II. SIMULATION PARAMETERS

Specification	Symbol	Value
Nominal power	P_n	2500W
Input voltage	V_{in}	144V
Output voltage	V_{out}	400V
Duty ratio	D	0.64
Clamp voltage	V_c	391V
Switching frequency	f_s	50kHz
Nominal duty cycle	D	0.64
Nominal phase-shifted	Φ	0.241

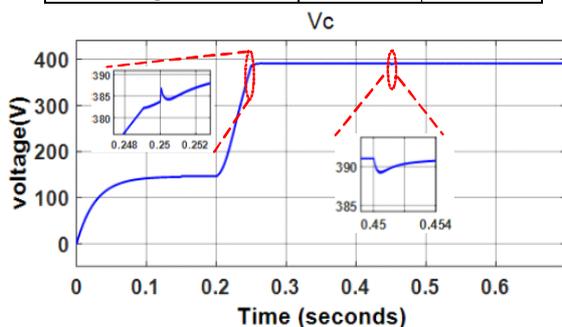


Fig. 6. Voltage response on the clamp capacitor.

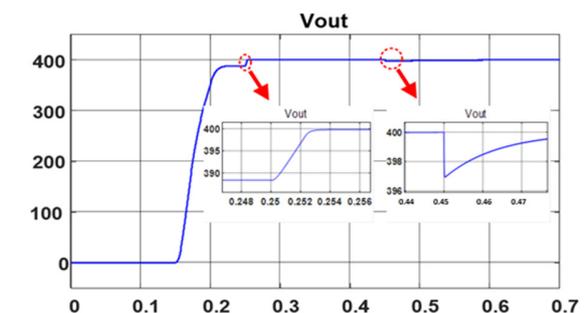


Fig. 7. Voltage response on the output capacitor.

Furthermore, the currents through the boost inductor and the leakage inductor with little spike in the start interval which ensure the feasibility of the control design in the experiment, are shown in Figures 8 and 9. The clamp voltage and the out voltage decrease slightly, about 3V, at 0.45s when the nominal load is changed. After that, they quickly return to the reference value within 0.01s and 0.03s. In steady-state operation, the voltages on both the LVS and HVS side of the transformer along with leakage current are shown in Figure 9. Due to the

small difference between the reference of clamp voltage and the output voltage, the bias current in the leakage current helps HVS switches achieve the ZVS easier [16]. In addition, the switches voltage and current are presented for S_1 and Q_2 in Figures 10 and 11.

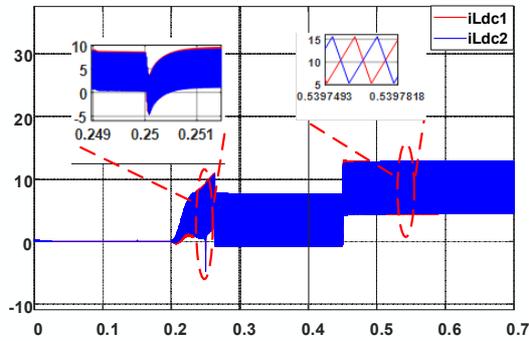


Fig. 8. Interleaved-boost current in simulation.

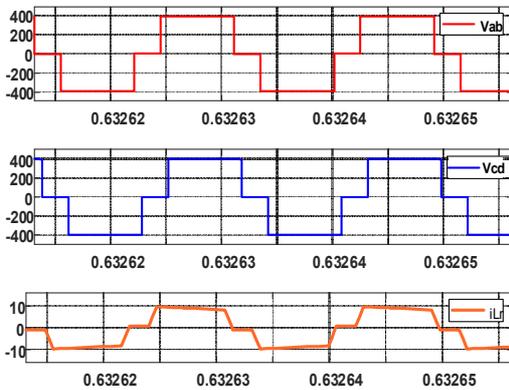


Fig. 9. Voltage on the two sides of the transformer and leakage current.

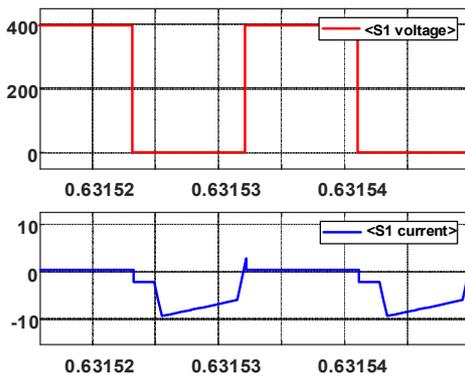


Fig. 10. Voltage and current on switch device S_1 .

III. CONCLUSION

The current paper presents the control structure along with the modeling of the CFDAB converter in boost mode. A small signal model is built up with the chosen modulation technique to control the output voltage of the converter. Simulations were

carried out to verify both the proposed model and the control design. The simulation results prove the good performance of the control design not only in the steady-state but also in the dynamic responses.

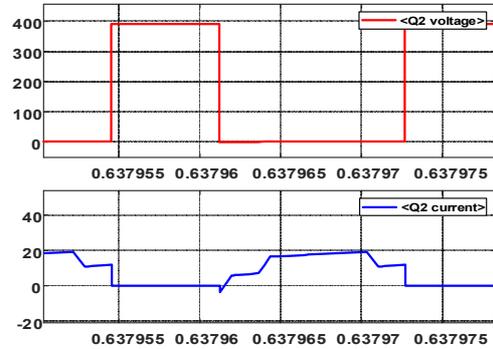


Fig. 11. Voltage and current on switch device Q_2 .

ACKNOWLEDGMENT

This research is funded by the Vietnamese Ministry of Education and Training under the project number CT2020.02.BKA.004.

REFERENCES

- [1] W. Chen, X. Yang, W. Zhang, and X. Song, "Leakage Current Calculation for PV Inverter System Based on a Parasitic Capacitor Model," *IEEE Transactions on Power Electronics*, vol. 31, no. 12, pp. 8205–8217, Dec. 2016, <https://doi.org/10.1109/TPEL.2016.2517740>.
- [2] Y. Almalag and M. Matin, "Two-Switch High Gain Non-Isolated Cuk Converter," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6362–6367, Oct. 2020, <https://doi.org/10.48084/etasr.3826>.
- [3] K. Jayaswal and D. K. Palwalia, "Performance Analysis of Non-Isolated DC-DC Buck Converter Using Resonant Approach," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3350–3354, Oct. 2018, <https://doi.org/10.48084/etasr.2242>.
- [4] B. Zhao, Q. Song, W. Liu, and Y. Sun, "Overview of Dual-Active-Bridge Isolated Bidirectional DC-DC Converter for High-Frequency-Link Power-Conversion System," *IEEE Transactions on Power Electronics*, vol. 29, no. 8, pp. 4091–4106, Aug. 2014, <https://doi.org/10.1109/TPEL.2013.2289913>.
- [5] B. J. Byen, B.-H. Jeong, and G.-H. Choe, "Single pulse-width-modulation strategy for dual-active bridge converters," *Journal of Power Electronics*, vol. 18, no. 1, pp. 137–146, Jan. 2018, <https://doi.org/10.6113/JPE.2018.18.1.137>.
- [6] B. J. Byen, C.-H. Ban, Y.-B. Lim, and G.-H. Choe, "An efficiency-optimized modulation strategy for dual-active-bridge DC-DC converters using dual-pulse-width-modulation in the low power region," *Journal of Power Electronics*, vol. 17, no. 6, pp. 1413–1421, Nov. 2017, <https://doi.org/10.6113/JPE.2017.17.6.1413>.
- [7] M. L. Mendola, M. di Benedetto, A. Lidozzi, L. Solero, and S. Bifaretti, "Four-Port Bidirectional Dual Active Bridge Converter for EVs Fast Charging," in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, Baltimore, MD, USA, Sep. 2019, pp. 1341–1347, <https://doi.org/10.1109/ECCE.2019.8912252>.
- [8] N. Hou and Y. W. Li, "A Tunable Power Sharing Control Scheme for the Output-Series DAB DC-DC System With Independent or Common Input Terminals," *IEEE Transactions on Power Electronics*, vol. 34, no. 10, pp. 9386–9391, Oct. 2019, <https://doi.org/10.1109/TPEL.2019.2911059>.
- [9] D. Sha and G. Xu, *High-Frequency Isolated Bidirectional Dual Active Bridge DC-DC Converters with Wide Voltage Gain*, 1st ed. Berlin Heidelberg, Germany: Springer, 2018.

- [10] A. K. Rathore and U. Prasanna, "Comparison of soft-switching voltage-fed and current-fed bi-directional isolated Dc/Dc converters for fuel cell vehicles," in *2012 IEEE International Symposium on Industrial Electronics*, Hangzhou, China, May 2012, pp. 252–257, <https://doi.org/10.1109/ISIE.2012.6237093>.
- [11] A. Pressman, K. Billings, and T. Morey, *Switching Power Supply Design*, 3rd ed. New York, NY, USA: McGraw-Hill Education, 2009.
- [12] Y. Shi, R. Li, Y. Xue, and H. Li, "Optimized Operation of Current-Fed Dual Active Bridge DC–DC Converter for PV Applications," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 11, pp. 6986–6995, Nov. 2015, <https://doi.org/10.1109/TIE.2015.2432093>.
- [13] P. Xuewei and A. K. Rathore, "Small signal modeling of naturally clamped soft-switching current-fed dual active bridge dc/dc converter and control design using Cypress PSoC," in *2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Mumbai, India, Dec. 2014, <https://doi.org/10.1109/PEDES.2014.7041961>.
- [14] A. Pal and S. Kapat, "Discrete-time Modeling of a Naturally Commutated Current-Fed Dual Active Bridge DC-DC Converter," in *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*, Anaheim, CA, USA, Mar. 2019, pp. 2877–2881, <https://doi.org/10.1109/APEC.2019.8722235>.
- [15] A. Pal and S. Kapat, "Accurate Discrete-Time Modeling of an Interleaved Current-Fed Dual Active Bridge DC-DC Converter," in *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*, Anaheim, CA, USA, Mar. 2019, pp. 1616–1621, <https://doi.org/10.1109/APEC.2019.8721870>.
- [16] D. Sha, X. Wang, K. Liu, and C. Chen, "A Current-Fed Dual-Active-Bridge DC–DC Converter Using Extended Duty Cycle Control and Magnetic-Integrated Inductors With Optimized Voltage Mismatching Control," *IEEE Transactions on Power Electronics*, vol. 34, no. 1, pp. 462–473, Jan. 2019, <https://doi.org/10.1109/TPEL.2018.2825991>.

Backstepping Terminal Sliding Mode MPPT Controller for Photovoltaic Systems

Khalissa Behih

Department of Electrical Engineering
LSI Laboratory
Ferhat Abbas Sétif 1 University
Sétif, Algeria
khalissabehih@univ-setif.dz

Hadjira Attoui

Department of Electrical Engineering
QUERE Laboratory
Ferhat Abbas Sétif 1 University
Sétif, Algeria
attoui_hadjira@univ-setif.dz

Abstract-In this paper, a new Maximum Power Point Tracking (MPPT) control for a Photovoltaic (PV) system is developed based on both backstepping and terminal sliding mode approaches. This system is composed of a solar array, a DC/DC boost converter, an MPPT controller, and an output load. The Backstepping Terminal Sliding Mode Controller (BTSMC) is used via a DC-DC boost converter to achieve maximum power output. The stability of the closed-loop system is guaranteed using the Lyapunov method. This novel approach provides good transient response, low tracking error, and very fast reaction against solar radiation and PV cell temperature variations. Furthermore, chattering, which constitutes the main disadvantage of the classic sliding mode technique is eliminated. To show the effectiveness and robustness of the proposed control, different simulations under different atmospheric conditions are conducted in Matlab/Simulink.

Keywords-backstepping; terminal sliding mode control; Lyapunov stability; maximum power point tracking; photovoltaic system

I. INTRODUCTION

Renewable energy sources are nowadays an important part of power generation. Photovoltaic (PV) generation is one of the most promising renewable sources since it exhibits many merits such as availability, cleanness, little maintenance, and no noise pollution. However, all PV systems have two problems: very low electric-power generation efficiency, especially under low-irradiation states and the interdependence of the amount of the electric power generated by solar arrays and the weather conditions. Load mismatch can occur under these weather varying conditions and maximum power may not be extracted and delivered to the load. This issue constitutes the so-called Maximum Power Point Tracking (MPPT) problem [1-4]. Many methods have been developed to determine the Maximum Power Point (MPP) under varying conditions [5-7]. Some of them are based on the well-known principle of perturb and observe (P&O) [8], others are based on sliding mode control [9-12], on artificial neural networks, or on fuzzy logic algorithms [1, 2, 7]. In [12-14], Maximum Power Voltage (MPV) based approaches are developed using a two-loop MPPT control scheme. The first loop is to determine the MPV reference of the PV array and the second loop is to regulate the

PV array voltage to the reference voltage. The procedure repeats the MPV reference searching and the PV voltage tracking until maximum power is reached. To track MPP more efficiently a hybrid method consisting of two loops is proposed in [15]. In the first loop, the MPP is estimated using an incremental conductance method, and in the second loop a terminal sliding mode controller is developed to drive the system to the searched reference MPP. Authors in [16] proposed the use of backstepping sliding mode control for the second loop. The backstepping sliding mode control law is based on the asymptotic stability analysis whilst the system trajectories evolve to a specified attractor reaching the equilibrium in an infinite time. Many authors have proposed an alternative way to get a finite time convergence based on terminal attractor techniques [15-18], providing high-precision performance besides disturbance attenuation. Moreover, chattering in BSMC remains the problem that has to be overcome.

In order to avoid chattering, the major disadvantage in the sliding mode methodology, various methods have been proposed. One of them consists in replacing the sign function by a continuous approximation in the vicinity of the sliding surface. Saturation function or sigmoid function was used in fuzzy logic to build the transition band [9, 21, 22]. The authors of [22] proposed to vary the sliding gain using a fuzzy system which adjusts the distance between the system and the sliding surface. Thus, its value decreases as the system state approaches the sliding surface. However, the ultimate accuracy and robustness of the sliding mode are partially lost.

In this paper, a Backstepping Terminal Sliding Mode Controller (BTSMC) is developed for MPPT. By using a DC/DC boost converter in the power control circuit, the BTSMC is proposed to drive the system to the MPV reference in the second loop. Using this approach, finite time convergence of the error is guaranteed and the chattering effect is eliminated without losing robustness.

II. PV SYSTEM MODELING

A DC-DC boost converter constituting the heart of the MPPT is inserted between the PV module and its load to achieve optimum power transfer, as can be seen in Figure 1.

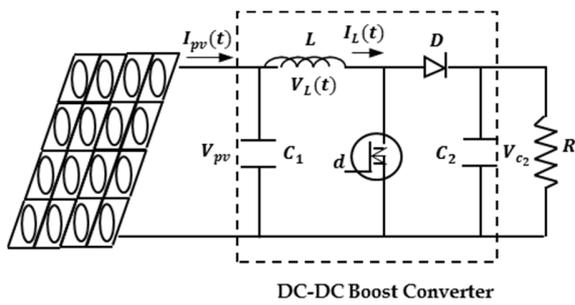


Fig. 1. Structure of the PV system.

The converter is used to regulate the PV module output voltage V_{pv} in order to extract as much power as possible from the PV module. Referring to [15], the dynamics of the boost converter are given by:

$$\left\{ \begin{aligned} \frac{dV_{pv}}{dt} &= \frac{1}{C_1} (I_{pv} - I_L) \\ \frac{dI_L}{dt} &= \frac{1}{L} V_{pv} - \frac{R_c(1-d)}{L(1+\frac{R_c}{R})} I_L \\ &\quad + \frac{(1-d)}{L} \left(\frac{R_c}{R_c+R} - 1 \right) V_{C_2} - \frac{V_D(1-d)}{L} \\ \frac{dV_{C_2}}{dt} &= -\frac{(1-d)}{C_2(1+\frac{R_c}{R})} I_L - \frac{1}{C_2(R_c+R)} V_{C_2} \end{aligned} \right. \quad (1)$$

where the three states variables V_{pv} , I_L and V_{C_2} are respectively the output voltage of the PV module, the inductor current and the voltage of the capacitor C_2 (i.e. the voltage across the load). V_D is the forward voltage of the power diode, d is the duty ratio of the PWM control input signal, and R is the resistive load.

By taking $x(t) = [V_{pv}(t) \ I_L(t) \ V_{C_2}(t)]^T$, the set of equations in (1) can be written in the following form [15]:

$$\left\{ \begin{aligned} \frac{dV_{pv}}{dt} &= \frac{1}{C_1} (I_{pv} - I_L) \\ \frac{dI_L}{dt} &= f_1(x) + g_1(x)d \\ \frac{dV_{C_2}}{dt} &= f_2(x) + g_2(x)d \end{aligned} \right. \quad (2)$$

where:

$$f_1(x) = \frac{1}{L} V_{pv} - \frac{R_c}{L(1+\frac{R_c}{R})} I_L + \frac{1}{L} \left(\frac{R_c}{R_c+R} - 1 \right) V_{C_2} - \frac{V_D}{L}$$

$$f_2(x) = \frac{1}{C_2(1+\frac{R_c}{R})} I_L - \frac{1}{C_2(R_c+R)} V_{C_2}$$

$$g_1(x) = \frac{R_c}{L(1+\frac{R_c}{R})} I_L - \frac{1}{L} \left(\frac{R_c}{R_c+R} - 1 \right) V_{C_2} + \frac{V_D}{L}$$

$$g_2(x) = -\frac{1}{L} \frac{1}{C_2(1+\frac{R_c}{R})} I_L$$

III. DESIGN OF THE BACKSTEPPING TERMINAL SLIDING MODE MPPT CONTROLLER

The overall control structure is illustrated in Figure 2. Here, i_{pv} and V_{pv} are measured from the PV array and transmitted to the MPP searching algorithm, which generates the reference maximum power voltage V_{ref} . Then, the reference voltage is given to the MPV based BTSM controller for maximum power tracking.

A. MPP Searching Algorithm

To seek the MPP voltage V_{ref} , we use an incremental conductance method [1, 15]. The power slope dP_{pv}/dV_{pv} can be expressed as:

$$\frac{dP_{pv}}{dV_{pv}} = I_{pv} + V_{pv} \frac{dI_{pv}}{dV_{pv}} \quad (3)$$

When the power slope $\frac{dP_{pv}}{dV_{pv}} = 0$, i.e. $\frac{dI_{pv}}{dV_{pv}} = -\frac{I_{pv}}{V_{pv}}$, the PV system operates at maximum power generation. Therefore, the update law for V_{ref} is given by the following rules [1, 15]:

$$\left\{ \begin{aligned} V_{ref} &= V_{ref}(k-1) + \Delta V, \quad \text{for } \frac{dI_{pv}}{dV_{pv}} > -\frac{I_{pv}}{V_{pv}} \\ V_{ref} &= V_{ref}(k-1) - \Delta V, \quad \text{for } \frac{dI_{pv}}{dV_{pv}} < -\frac{I_{pv}}{V_{pv}} \end{aligned} \right. \quad (4)$$

B. Backstepping Terminal Sliding Model Controller

The backstepping terminal sliding mode controller is designed to extract maximum power from a PV panel. The objective of the controller is to let the panel PV voltage V_{pv} track the reference maximum power voltage V_{ref} by acting on the duty cycle $d(t)$ of the boost converter switch. The recursive nature of the proposed control design is similar to the standard backstepping methodology. However, the proposed control design uses backstepping to design controllers with a terminal sliding surface at the last step [22-23]. The design proceeds as follows:

For the first step we consider a zero-order sliding surface:

$$e_1 = x_1 - V_{ref} \quad (5)$$

Considering an auxiliary tracking error variable:

$$e_2 = \dot{e}_1 + \alpha_1 \quad (6)$$

Let the first Lyapunov function candidate is:

$$V_1 = \frac{1}{2} e_1^2 \quad (7)$$

The time derivation of (7) is given by:

$$\begin{aligned} \dot{V}_1 &= e_1 \dot{e}_1 \\ &= e_1 (e_2 - \alpha_1) \quad (8) \\ &= -\lambda_1 e_1^2 + e_1 e_2 \end{aligned}$$

The stabilization of e_1 can be obtained by introducing a new virtual control α_1 , such that:

$$\alpha_1 = \lambda_1 e_1, \quad \lambda_1 > 0 \quad (9)$$

where λ_1 is a positive feedback gain, such that α_1 can be chosen in order to eliminate the nonlinearity and getting $\dot{V}_1(x) < 0$.

Equation (8) shows that, if the designed control law makes e_2 converge, then $\dot{V}_1(x) < 0$ which guarantees global stability. In order to make e_2 converge in finite time and improve the convergence rate and the steady-state tracking accuracy of the

system, the higher-order non-singular terminal sliding mode surface is designed for e_2 as follows:

$$s = e_2 + \gamma \dot{e}_2^{\frac{p}{q}}, \quad 1 < p/q < 2, \gamma > 0 \quad (10)$$

where $\gamma > 0$ is a positive constant that contributes to force the error e_2 to converge to zero while p and q are positive impair constants such that $1 < p/q < 2$. This last inequality must be enforced to avoid the inherent singularity in such types of control. The surface selected in (10) is called Nonsingular Terminal Sliding Mode (NTSM). It is used to realize a second-order SMC, enabling to have $s = \dot{s} = \ddot{s} = 0$, often called higher-order sliding mode. If the system converges on t_r , s converges to zero, which means $s(t) = 0, t \geq t_r$, then from (10), it can be shown that e_2 and \dot{e}_2 will converge to zero in finite time, given by:

$$t_s = t_r + \gamma^{p/q} \frac{p}{(p-q)} |e_2(t_r)|^{\frac{p-q}{p}} \quad (11)$$

The system will remain on the second-order sliding mode ($e_2 = \dot{e}_2 = 0$), according to (11), and by selecting the parameters p, q , and γ one can adjust the convergence speed of e_2 .

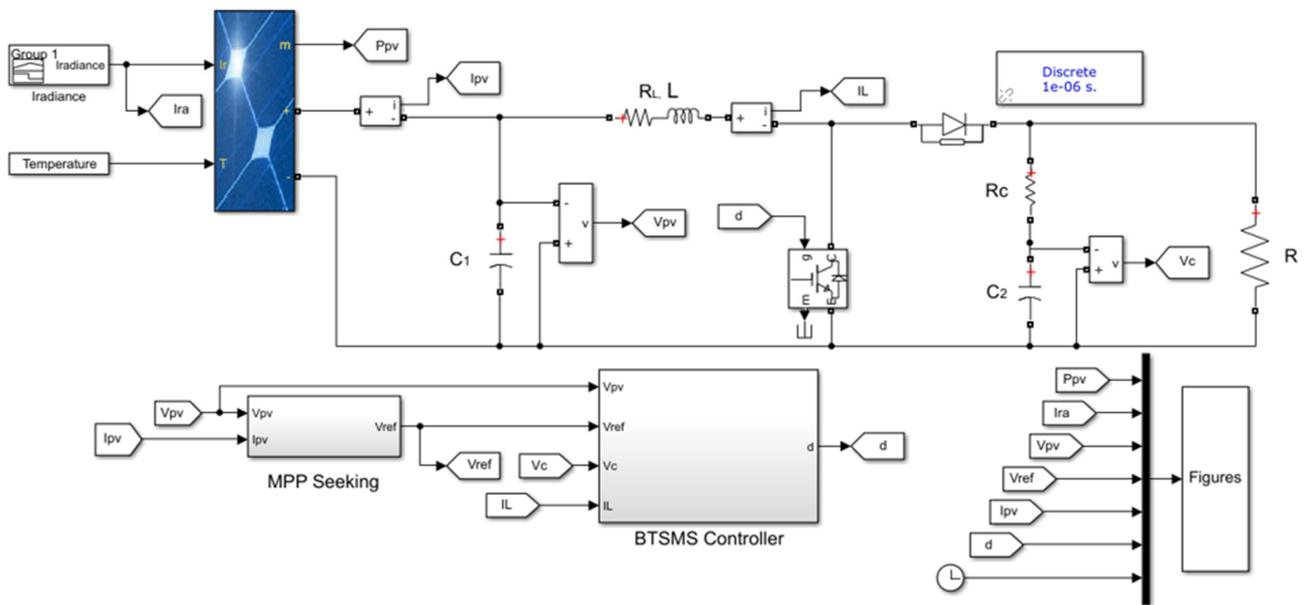


Fig. 2. Simulink PV system model.

C. Stability and Robustness Analysis

The stability and robustness issues of the controller are addressed here by using Lyapunov stability theory as follows:

The augmented Lyapunov function is given by:

$$V_2 = \frac{1}{2} s^2 \quad (12)$$

The time derivative of $V_2(s)$ is then:

$$\dot{V}_2 = s \dot{s} \quad (13)$$

$$\begin{aligned} \dot{V}_2 &= s \left[\dot{e}_2 + \gamma \frac{p}{q} \dot{e}_2^{\left(\frac{p-1}{q}\right)} \dot{e}_2 \right] \\ &= s \gamma \frac{p}{q} \dot{e}_2^{\left(\frac{p-1}{q}\right)} \left[\ddot{e}_2 + \frac{1}{\gamma} \frac{q}{p} \dot{e}_2^{\left(2-\frac{p}{q}\right)} \right] \end{aligned} \quad (14)$$

with:

$$\begin{aligned} \dot{e}_2 &= \ddot{e}_1 + \dot{\alpha}_1 \\ &= \ddot{V}_{pv} - \ddot{V}_{ref} + \dot{\alpha}_1 \end{aligned} \quad (15)$$

$$\begin{aligned} \dot{V}_2 &= s \left[\dot{e}_2 + \gamma \frac{p}{q} \dot{e}_2^{\left(\frac{p-1}{q}\right)} \dot{e}_2 \right] \\ &= s \gamma \frac{p}{q} \dot{e}_2^{\left(\frac{p-1}{q}\right)} \left[\ddot{e}_2 + \frac{1}{\gamma} \frac{q}{p} \dot{e}_2^{\left(2-\frac{p}{q}\right)} \right] \end{aligned} \quad (16)$$

To satisfy the Lyapunov stability condition, the input signal control is designed as in (17):

$$d(t) = \frac{1}{g_1(x)} [d_1(t) + d_2(t)], \quad g_1(x) \neq 0 \quad (17)$$

where

$$d_1(t) = -f_1(x) + \dot{V}_{pv} - C_1 \ddot{V}_{ref} + C_1 \dot{\alpha}_1 \quad (18)$$

and

$$d_2(t) = C_1 \left[\int_0^t \left[\frac{1}{\gamma} \frac{q}{p} \dot{e}_2^{\left(2-\frac{p}{q}\right)} + \eta \operatorname{sign}(s) \right] dt \right] \quad (19)$$

Equation (14) can therefore be developed to give:

$$\dot{V}_2 = -\gamma \frac{p}{q} \dot{e}_2^{\left(\frac{p-1}{q}\right)} \eta |s| \quad (20)$$

If $s \neq 0$, as p and q are odd and $1 < p/q < 2$, then $\dot{e}_2^{p/q-1} \geq 0$, which means $\dot{V}_2 \leq 0$. If $s = 0$, we will consider the two following conditions:

- If $z_n \neq 0$, then $\dot{V}_2 = -\gamma \frac{p}{q} \dot{e}_2^{\left(\frac{p-1}{q}\right)} \eta |s| < 0$ and
- If $\dot{z}_n = 0, z_n \neq 0$ then $\dot{V}_2 = 0$ cannot be kept continuously.

Thus, according to Lyapunov stability theory, the system can reach the manifold in finite time, for e_2 converges to zero in finite time.

IV. SIMULATION RESULTS

The Simulink PV system model shown in Figure 2 was selected to assess the performance and the effectiveness of the proposed BTSMC controller. Fixed-step type and ode4 (Runge Kutta) solver were the Simulink solver selection parameters. The specification parameters for PV power generation are given in Table I. The parameters of the boost converter used are: $L=12\text{mH}$, $R_L=0.15\Omega$, $R_C=39.6\Omega$, $C_2=1000\mu\text{F}$, $R=25\Omega$, and $V_D=0.82\text{V}$. The controller parameters used are: $\gamma=50$, $\eta=0.01$, $p=7$, $q=5$. In this section, we present the simulation results when applying the Backstepping Terminal Sliding Mode control law.

TABLE I. ELECTRICAL SPECIFICATIONS OF THE KC200GH-2P PV MODULE

Parameter	Value
Maximum power P_{mpp}	200W
Short circuit current I_{scr}	8.21A
Open circuit voltage V_{OC}	32.9V
Voltage at maximum power point V_{mpp}	26.3V
Current at maximum power point I_{mpp}	7.61A
P-N junction characteristic factor A	1.8

A. Simulation Results with Standard Operating Conditions

The simulation results with standard operating conditions ($S=1000\text{W/m}^2$, $T=25^\circ\text{C}$) are shown in Figures 3–6. From the obtained results, it is clearly noted that V_{pv} reaches quickly the desired set point V_{ref} .

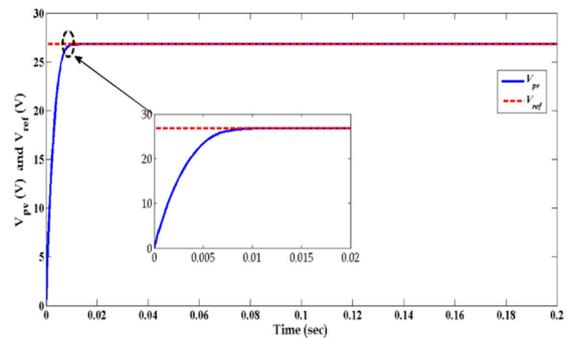


Fig. 3. Evolution of V_{pv} and V_{ref} .

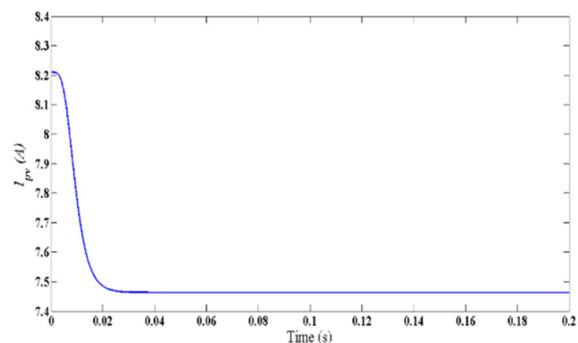


Fig. 4. Evolution of I_{pv} .

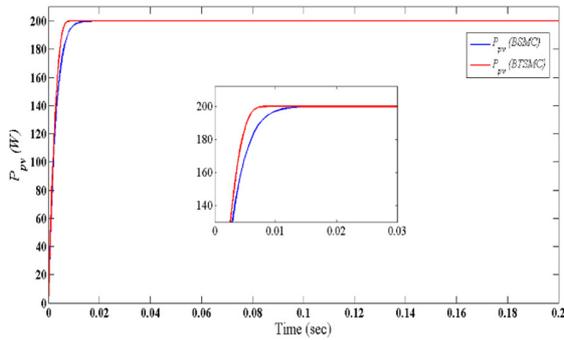


Fig. 5. Evolution of P_{pv} .

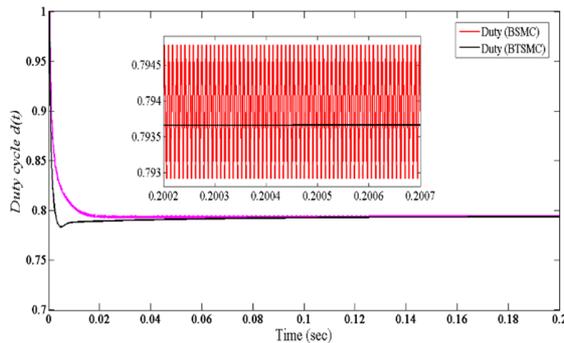


Fig. 6. Evolution of duty cycle $d(t)$.

The P_{pv} achieves maximum power in a settling time of 0.008s. Moreover, it can be deduced that the proposed approach presents good transient response, low tracking error, and a very fast system reaction against set point change. Figures 5-6 present a comparative study between the proposed controller and the BSMC. When comparing these results, it is obvious that the proposed method provides faster responses while eliminating the phenomenon of chattering.

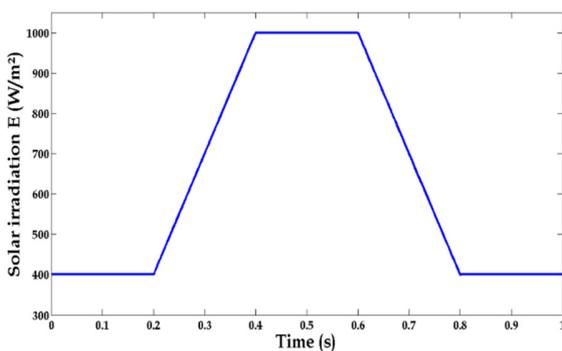


Fig. 7. Solar irradiation variation.

B. Simulation Results under Solar Irradiation Variation

It is known that meteorological parameters, especially the temperature, do not remain constant all day long, but change significantly. It is then worth investigating the influence of the daily average temperature variation on the performance of the optimized system. In order to evaluate the effect of changing

irradiation conditions, an irradiation ramp change was used as shown in Figure 7. The temperature is considered constant with a value of 25°C during the simulation tests. The performances including V_{pv} , I_{pv} , P_{pv} , and duty cycle are given in Figures 8-11.

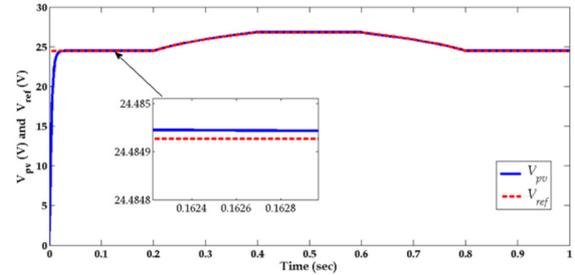


Fig. 8. Evolution of V_{pv} and V_{ref} .

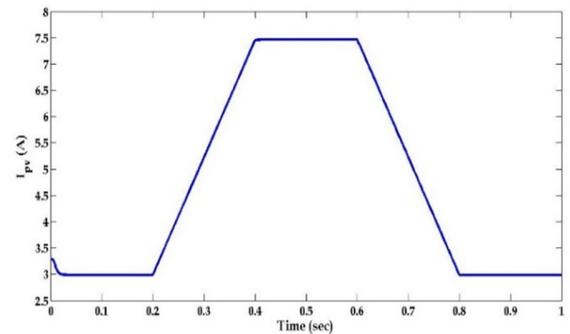


Fig. 9. Evolution of I_{pv} .

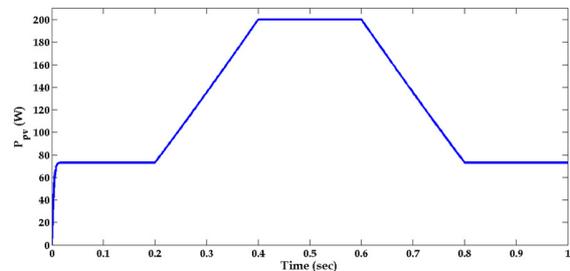


Fig. 10. Evolution of P_{pv} .

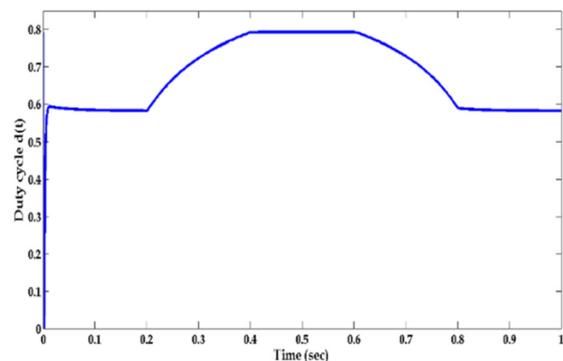


Fig. 11. Evolution of duty cycle $d(t)$.

As shown in Figures 8-11, when the irradiance level changes, the MPPT controller can quickly track the MPP. It can be seen from the results that the proposed MPPT controller follows the direction of the MPP perfectly.

C. Simulation Results under Temperature Variations

Simulations were carried out to evaluate the effect of changing temperature (Figure 12). The solar irradiation is considered constant with a value of 1000W/m². The performances including V_{pv} , I_{pv} , P_{pv} , and duty cycle are given in Figures 13-16.

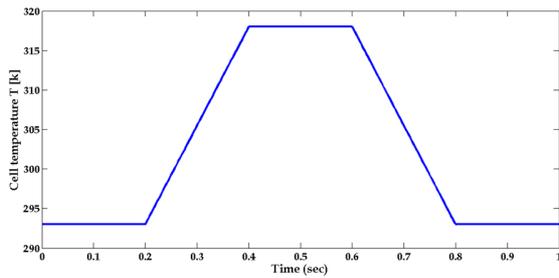


Fig. 12. Temperature variation.

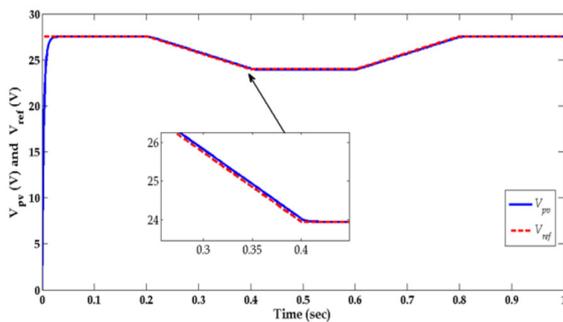


Fig. 13. Evolution of V_{pv} and V_{ref} .

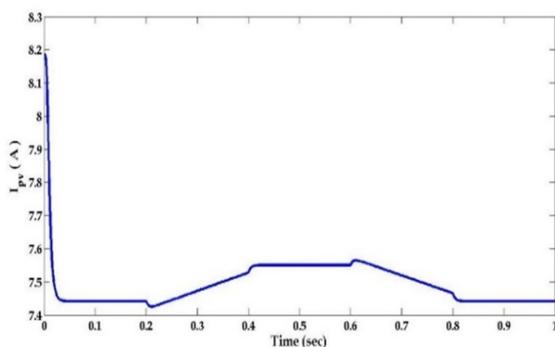


Fig. 14. Evolution of I_{pv} .

It can be seen in Figures 13-16 that when the temperature changes, the MPPT controller can quickly track the MPP. The proposed MPPT controller provides a good performance in all simulation tests. Figure 17 shows that the trace of the operating point stays close to the MPPs during the simulation.

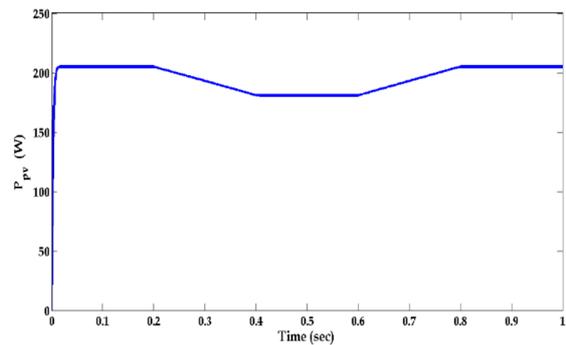


Fig. 15. Evolution of P_{pv} .

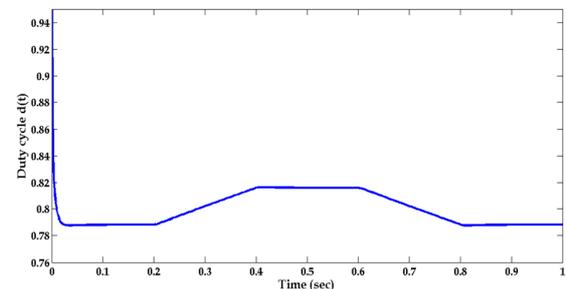


Fig. 16. Evolution of duty cycle $d(t)$.

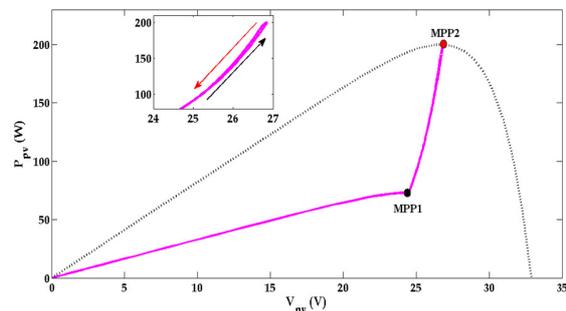


Fig. 17. $V_{pv} - P_{pv}$ characteristics under solar irradiation variation.

D. Simulation Results under Load Variation

Considering load change from 25Ω to 80Ω under the same irradiance and temperature, the corresponding results are shown in Figures 18-20. It can be easily concluded that the proposed controller achieves strong robustness and has satisfactory response under these types of disturbance.

In all the simulation results, the proposed BTSMC provides a good performance. The proposed MPPT controller follows the direction of the MPP perfectly and in finite time. Furthermore, the elimination of the chattering problem permits the smoothness of the control law.

V. CONCLUSION

In this study, a backstepping sliding mode control scheme for MPPT of a PV system has been developed, which integrates the backstepping methodology and the terminal sliding mode

control strategy. The combined strategies proved to have the advantages of both the terminal sliding mode and the backstepping approach. The parameters of the proposed controller are adjusted using adaptation laws, based on the Lyapunov synthesis in order to ensure, in the presence of external disturbances, the stability of the closed-loop system and the tracking performance is achieved in finite time.

The objective of the study was accomplished. Low tracking error, very fast system reaction against solar radiation change, and good transition response without overshoot was obtained. Moreover, it has been shown that the chattering phenomenon appearing in the classical backstepping sliding mode control was reduced. Besides, the proposed algorithm provides a shorter rise-time than the backstepping sliding mode control. Future work will include the implementation of the algorithm on a PV system using dSPACE1104 including a partial shading influence study.

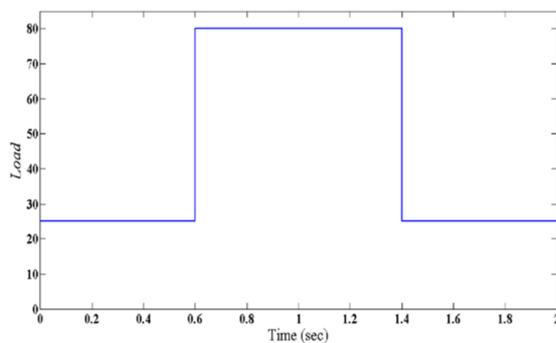


Fig. 18. Charge variation.

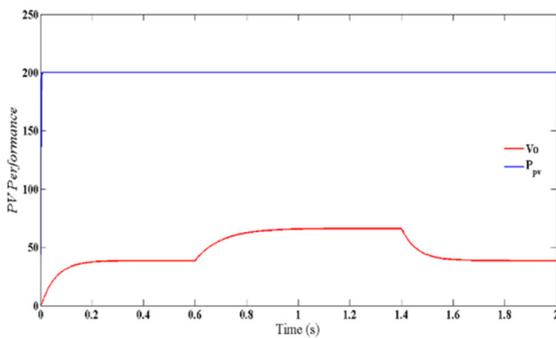


Fig. 19. PV performance.

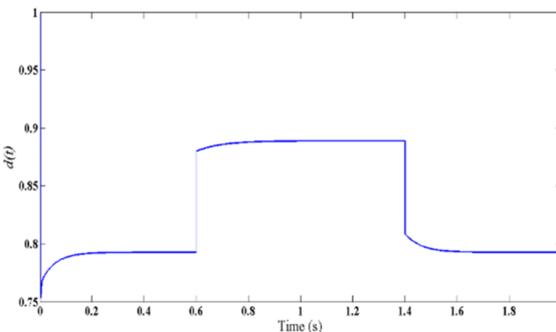


Fig. 20. Evolution of duty cycle.

REFERENCES

- [1] M. Y. Allani, D. Mezghani, F. Tadeo, and A. Mami, "FPGA Implementation of a Robust MPPT of a Photovoltaic System Using a Fuzzy Logic Controller Based on Incremental and Conductance Algorithm," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4322–4328, Aug. 2019, <https://doi.org/10.48084/etasr.2771>.
- [2] A. Elgharbi, D. Mezghani, and A. Mami, "Intelligent Control of a Photovoltaic Pumping System," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4689–4694, Oct. 2019, <https://doi.org/10.48084/etasr.2982>.
- [3] M. Aly and H. Rezk, "A Differential Evolution-Based Optimized Fuzzy Logic MPPT Method for Enhancing the Maximum Power Extraction of Proton Exchange Membrane Fuel Cells," *IEEE Access*, vol. 8, pp. 172219–172232, 2020, <https://doi.org/10.1109/ACCESS.2020.3025222>.
- [4] J. E. Hernández-Díez, C. F. Méndez-Barrios, S. I. Niculescu, and E. Bárcenas-Bárcenas, "A current sensorless delay-based control scheme for MPPT-boost converters in photovoltaic systems," *IEEE Access*, vol. 8, pp. 174449–174462, 2020, <https://doi.org/10.1109/ACCESS.2020.3024566>.
- [5] J. Wang, C. Wang, S. jia, S. Xiang, and X. Shan, "Photovoltaic power MPPT controller based on fuzzy-backstepping method," in *2019 8th International Symposium on Next Generation Electronics (ISNE)*, Zhengzhou, China, Oct. 2019, pp. 1–3, <https://doi.org/10.1109/ISNE.2019.8896680>.
- [6] O. Abdel-Rahim and H. Wang, "A new high gain DC-DC converter with model-predictive-control based MPPT technique for photovoltaic systems," *CPSS Transactions on Power Electronics and Applications*, vol. 5, no. 2, pp. 191–200, Jun. 2020, <https://doi.org/10.24295/CPSSPEA.2020.00016>.
- [7] K. Y. Yap, C. R. Sarimuthu, and J. M.-Y. Lim, "Artificial Intelligence Based MPPT Techniques for Solar Power System: A review," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1043–1059, Nov. 2020, <https://doi.org/10.35833/MPCE.2020.000159>.
- [8] Z. R. Labidi, H. Schulte, and A. Mami, "A Systematic Controller Design for a Photovoltaic Generator with Boost Converter Using Integral State Feedback Control," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 4030–4036, Apr. 2019, <https://doi.org/10.48084/etasr.2687>.
- [9] Y. Zhang, Q. Zhang, J. Zhang, and Y. Wang, "Sliding Mode Control for Fuzzy Singular Systems With Time Delay Based on Vector Integral Sliding Mode Surface," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 4, pp. 768–782, Apr. 2020, <https://doi.org/10.1109/TFUZZ.2019.2916049>.
- [10] A. Bag, B. Subudhi, and P. K. Ray, "A combined reinforcement learning and sliding mode control scheme for grid integration of a PV system," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 4, pp. 498–506, Dec. 2019, <https://doi.org/10.17775/CSEEJPES.2017.01000>.
- [11] Z. Meng, W. Shao, J. Tang, and H. Zhou, "Sliding-mode control based on index control law for MPPT in photovoltaic systems," *CES Transactions on Electrical Machines and Systems*, vol. 2, no. 3, pp. 303–311, Sep. 2018, <https://doi.org/10.30941/CESTEMS.2018.00038>.
- [12] I.-S. Kim, "Robust maximum power point tracker using sliding mode controller for the three-phase grid-connected photovoltaic system," *Solar Energy*, vol. 81, no. 3, pp. 405–414, Mar. 2007, <https://doi.org/10.1016/j.solener.2006.04.005>.
- [13] E. Koutroulis, K. Kalaitzakis, and N. C. Voulgaris, "Development of a microcontroller-based, photovoltaic maximum power point tracking control system," *IEEE Transactions on Power Electronics*, vol. 16, no. 1, pp. 46–54, Jan. 2001, <https://doi.org/10.1109/63.903988>.
- [14] M. Veerachary, T. Senjyu, and K. Uezato, "Neural-network-based maximum-power-point tracking of coupled-inductor interleaved-boost-converter-supplied PV system using fuzzy controller," *IEEE Transactions on Industrial Electronics*, vol. 50, no. 4, pp. 749–758, Aug. 2003, <https://doi.org/10.1109/TIE.2003.814762>.
- [15] C.-S. Chiu, Y.-L. Ouyang, and C.-Y. Ku, "Terminal sliding mode control for maximum power point tracking of photovoltaic power generation systems," *Solar Energy*, vol. 86, no. 10, pp. 2986–2995, Oct. 2012, <https://doi.org/10.1016/j.solener.2012.07.008>.

- [16] K. Dahech, M. Allouche, T. Damak, and F. Tadeo, "Backstepping sliding mode control for maximum power point tracking of a photovoltaic system," *Electric Power Systems Research*, vol. 143, pp. 182–188, Feb. 2017, <https://doi.org/10.1016/j.epsr.2016.10.043>.
- [17] N. Zerroug, M. N. Harmas, S. Benaggoune, Z. Bouchama, and K. Zehar, "DSP-based implementation of fast terminal synergetic control for a DC–DC Buck converter," *Journal of the Franklin Institute*, vol. 355, no. 5, pp. 2329–2343, Mar. 2018, <https://doi.org/10.1016/j.jfranklin.2018.01.004>.
- [18] R. Ma, H. Zhang, M. Yuan, B. Liang, Y. Li, and Y. Huangfu, "Chattering Suppression Fast Terminal Sliding Mode Control for Aircraft EMA Braking System," *IEEE Transactions on Transportation Electrification*, 2021, <https://doi.org/10.1109/TTE.2021.3054510>.
- [19] K. Behih, K. Benmahammed, Z. Bouchama, and M. N. Harmas, "Real-Time Investigation of an Adaptive Fuzzy Synergetic Controller for a DC-DC Buck Converter," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4984–4989, Dec. 2019, <https://doi.org/10.48084/etasr.3172>.
- [20] Z. B. Duranay, H. Guldemir, and S. Tuncer, "Fuzzy Sliding Mode Control of DC-DC Boost Converter," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 3054–3059, Jun. 2018, <https://doi.org/10.48084/etasr.2116>.
- [21] S. Latreche and S. Benaggoune, "Robust Wheel Slip for Vehicle Anti-lock Braking System with Fuzzy Sliding Mode Controller (FSMC)," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6368–6373, Oct. 2020, <https://doi.org/10.48084/etasr.3830>.
- [22] T. L. Nguyen, T. H. Vo, and N. D. Le, "Backstepping Control for Induction Motors with Input and Output Constrains," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5998–6003, Aug. 2020, <https://doi.org/10.48084/etasr.3680>.
- [23] B. K. Oubbati, M. Boutoubat, A. Rabhi, and M. Belkheiri, "Experiential Integral Backstepping Sliding Mode Controller to achieve the Maximum Power Point of a PV system," *Control Engineering Practice*, vol. 102, Sep. 2020, Art. no. 104570, <https://doi.org/10.1016/j.conengprac.2020.104570>.

Erratum and Addendum: "Mandible Bone Osteoporosis Detection using Cone-beam Computed Tomography"

Rasha F. Abu Marar
Department of Computer Science
Faculty of Information Technology
Middle East University
Amman, Jordan
ramarar82@gmail.com

Diaa M. Uliyan
Department of Information Computer Science
College of Computer Science and Engineering
University of Hai'l
Hai'l, Saudi Arabia
d.ulyan@uoh.edu.sa

Hamza A. Al-Sewadi
Department of Computer Science
Faculty of Information Technology
Middle East University
Amman, Jordan
hsewadi@meu.edu.jo

The authors would like to clarify that [1] is the original source of a Figure that was included in our previously published manuscript as Figure 7 [2]. Although a reference to [1] was provided in [2], a direct reference was not inserted in the image caption. The image is reshown here as Figure 1 for convenience. We would like to thank Dr E. Klintström, the copyright holder, for agreeing to the reuse of the Figure.

Engineering, Technology & Applied Science Research, vol. 10, no. 4, pp. 6027–6033, Aug. 2020, <https://doi.org/10.48084/etasr.3637>.

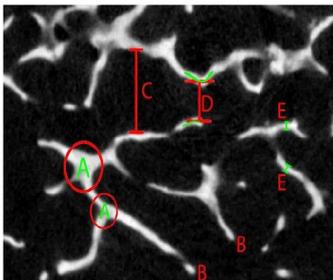


Fig. 1. The above figure was included as Fig. 7 in [2] with the following caption: "The figure Bone structure parameters. (A) Trabecular nodes, (B) trabecular termini, (C) trabecular separation, (D) trabecular spacing, and (E) trabecular thickness." The authors of [2] wish to clarify that the source of this Figure is [1].

REFERENCES

- [1] E. Klintström, "Image Analysis for Trabecular Bone Properties on Cone-Beam CT Data," Linköping University Medical Dissertations No. 1594, Department of Medical and Health Sciences, Linköping University, Sweden, Linköping, 2017.
- [2] R. F. A. Marar, D. M. Uliyan, and H. A. Al-Sewadi, "Mandible Bone Osteoporosis Detection using Cone-beam Computed Tomography,"