

Implementation of Lean Manufacturing in a Cement Industry

Majed Masmali

Industrial Engineering Department
College of Engineering
Northern Border University
Arar, Saudi Arabia
majed.masmali@nbu.edu.sa

Abstract- The lean manufacturing concept is a systematic minimization of waste and non-value activities in production processes introduced by the Toyota production system. In this research, lean manufacturing is implemented in a cement production line. Value Stream Mapping (VSM) is applied to give a clear picture of the value chain in cement production processes and to highlight the non-value-added in the shop floor. To begin, the existing VSM is constructed based on the information and data gathered during visiting and observing the manufacturing process in the firm. As a result, the excess inventory between workstations was identified as a major waste generation, hence, the proposed VSM conducts further improvement and makes action plans to alleviate the unwanted activities. Then, the takt time to ensure smooth material flow and to avoid any occurring delay or bottleneck in the production line was figured out. The supermarket pull-based production control is suggested to be adopted in the future map. Two pull production strategies are selected in this case. The first is applying the Kanban system to control the level of inventory between workstations. The other is the CONWIP approach to control the amount of work in process to the entire production line. The outcome of the proposed models indicates a decrease of the none-value time from 23 days in the current state to about 4 and 2 days in Kanban and CONWIP systems respectively, so the CONWIP was suggested as most efficient. Some suggestions for further research are also mentioned.

Keywords- cement industry; CONWIP; Kanban; pull systems; value stream mapping; lean manufacturing

I. INTRODUCTION

The cement industry is the most prominent building material industry, contributing greatly to the global and national economy and growth. Since cement is a primary construction material, cement economics are closely related to the construction sector. The history of the cement industry goes back to 1824, when the Portland cement was invented [1-3]. Thereafter, the cement industry has developed and the volume of production has increased steadily. The raw materials include clay limestone transform into clinker. The clinker is transformed into cement by adding gypsum and other materials. The cement production process passes through numerous workstations including crushers, milling, rotary kiln, and milling finishing cements. Between these manufacturing

processes there are stocks and silos between workstations. The cement is made by mixing lime, clay, and iron materials in certain proportions to give the coated materials calcium, silica, corundum, and iron oxide. These raw materials are then grinded well to facilitate the burning process, which takes place at a temperature of 1500°C. This material after burning is called clinker. The lime materials for clinker make up about 90% of the cement mass, as gypsum material is added to it in proportions ranging from 3 to 5% to control the hardening and cohesion of the cement, which varies in terms of the added amount according to the type produced [1-6].

This series of activities from the suppliers to the customers is known as the value chain. The customer order is considered as the input of flow information, while the output is the order fulfillment. The aim of the value chain is to make the enterprise competitive [7, 8]. In order to ensure the sustainability of production and competitiveness with a view to improve the cement production line, production management principles should be implemented in this industry. The purpose of these strategies is to make decisions to develop and improve the shop floor operations in cement factories. One of these aspects is reducing the work in process, since the excess inventory could contribute to the holding cost. In addition, no value is added to the clinker and cement which are stored in the factory. Therefore, the lean manufacturing concept is chosen in this research to be implemented in the production processes. It is a systematic minimization of non-value activities in the manufacturing systems [9, 10].

The lean manufacturing philosophy has been introduced in the Toyota production system. The principle of this technique is to eliminate the waste that occurs in the production process. The waste in lean manufacturing comes from over production, waiting, transportation, over processing, movement, inventory, and defects [9-11]. Liker [12] extended the definition by considering the non-utilized talents as a waste. When the lean manufacturing is implemented a sort of tools must be used to reduce and eliminate the added non-value, tools such as pull system, value stream map, Heijunka, SMED, Total Productive Maintenance (TPM), Poka-yoke 5'S, and Kaizen [9-11].

Lean manufacturing has been implemented to different manufacturing aspects. This part will focus on the studies

related to the cement industry. Authors in [13] implemented the lean philosophy to improve the cement production line performance. Maximum efficiency and minimum lead time and production cost can be achieved if the right training is provided to the employees and a motivation system is introduced to the workplace environment. Authors in [14] developed an approach for the application of lean manufacturing in the cement industry. A numerical model was built to compare the current and future state VSM. The output of the proposed model outlined that there was an improvement in throughput and reduction in lead time. In addition, the Kanban was applied to study the effects of the inventory in cement manufacturing. Authors in [15] implemented TPM in the cement production line. Authors in [16] applied VSM in order to explore the various non-value added activities in the supply chain of a cement factory. Authors in [17] determined the essential points for implementing the management of Just in Time philosophy on the cement industry in Pakistan. Inventory management, production design, production plan, supply chain integration, and total quality control have been selected as factors in order to eliminate the waste of resources. Authors in [18] demonstrated a numerical approach to study the cement manufacturing lines taking into account the supply chain management and its effect on Just in Time production. Authors in [19] proposed a simulation model to study the supply chain management in cement industry. Authors in [20] introduced a matrix for the waste relationship, a questionnaire of waste assessment, and a value stream analysis approach to eliminate the waste in cement production. In the previous study, the approach consists of four stages which are identification of waste, value stream analysis, fishbone diagram, and failure mode and effect analysis. The literature review shows significant contributions that should be further researched in terms of the application of lean production in the cement industry.

In the current article, the VSM tool is applied in a cement production line. The main contribution of this article is the implementation of lean manufacturing in a cement factory in order to control and minimize the level of work in process. There is an opportunity to compare the outcomes of the proposed pull production systems which are Kanban and CONWIP strategies among the push production system.

II. RESEARCH METHODOLOGY

In this study, VSM is applied to implement lean manufacturing towards a cement production line. It is a significant lean tool that identifies the potential waste and forms a clear picture of the production process in order to build a road map of improving the production process. The four main steps of implementing VSM [11] are given in Figure 1.



Fig. 1. Value stream mapping process.

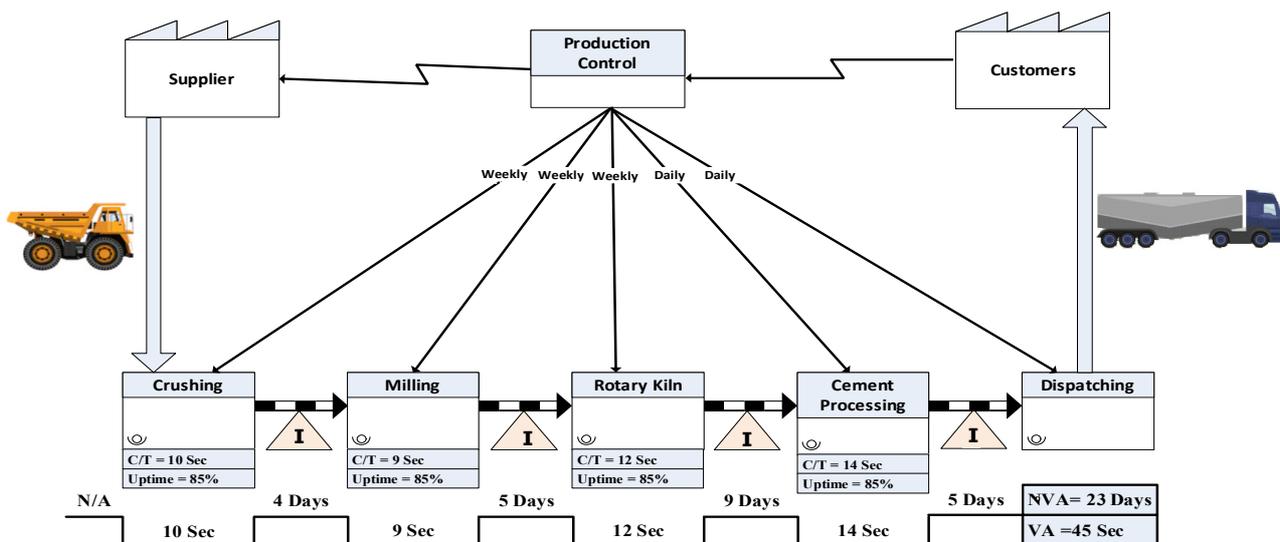


Fig. 2. Current value stream mapping.

The research study is conducted at a cement factory. The design capacity of this plant is 6,000 tons per day, whereas the

annual production is 1.98 million tons, based on 330 working days per year and taking into account the shutdown for

maintenance. The current VSM is constructed according to the guidelines in [11]. The relevant data and information associated with the production processes, i.e. information flow, physical flow, and time line are gathered. Information flow is related to the customer requests (from the sales department or from orders from the factory’s website) where the production-planning can be conducted. The second element is material flow. Data were taken during visiting the factory and observing the progress of operations and workflow in the shop floor, in order to understand the cement manufacturing industry processes. The cycle times are taken as the average cycle time from production department records. When it comes to the inventories, there are four stores between each workstation. The average time of materials stored in silos has been reported. The down time for each process has been assumed to be 15%. The third element is trend line. It consists of two time components. The first timeline is the non-value added time. It is also called the lead time or waiting time. It is the summation of the amount of each inventory. Another timeline is value added time. It can be computed by the summation of cycle times for each workstation process.

The current VSM is illustrated in Figure 2. It can be seen that from the trend line, the non-value time is about 23 days, whereas the value time is approximately 45 seconds.

III. THE PROPOSED FUTURE VALUE STREAM MAPPING

Based on the existing VSM outcomes, it is necessary for the proposed VSM model to eliminate the non-value added in the cement production line. First of all, the maximum time required to meet the customer demands should be identified in order to avoid any waste or delay of delivery. The takt time is computed as [11]:

$$\text{Takt Time} = \frac{\text{Available working time}}{\text{Customer demand}} \quad (1)$$

The maximum demand required by customers for the cement production line is about 6000 tons per day. As the plant runs 24 hours a day, the available working time is 86,400 seconds. Therefore, the takt time can be obtained as:

$$\text{Takt Time} = \frac{86400 \text{ seconds}}{6000 \text{ tons per day}} = 14.4 \text{ seconds per ton}$$

This means that 1 ton of cement needs to be produced in 14.4 seconds. Subsequently, the overall takt time is compared against cycle time for each workstation as shown in Figure 3. In addition, the current production strategy for the factory follows the push production control shown in Figure 4.

Each workstation pushes completed work to the next workstation. As a result, the inventory level between workstations is increased. In the proposed future map, pull production is applied to produce the right quantity of products at the right time. A common approach of pull based production

process is the Kanban. One of its benefits is the control of the level of inventory between workstations [21]. However, authors in [22-25] introduced another pull system, Constant Work In Process (CONWIP), to control the production line. Both the above pull approaches were implemented in the future state maps.

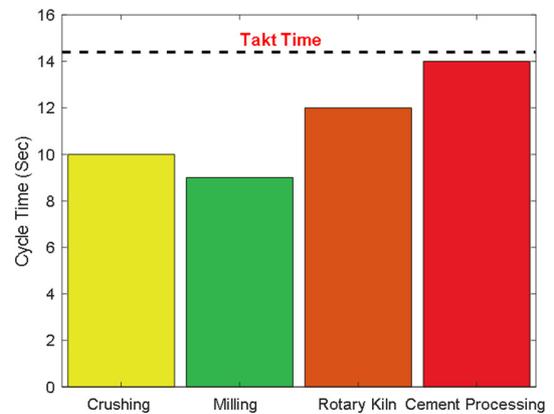


Fig. 3. Cement production line balancing.

A. Kanban

Kanban is a lean tool associated with the pull production system and Just in Time production [7-9]. To demonstrate the use of the Kanban pull system to control the excess inventory, the following assumptions are made: The cement production line is divided into four stages, as shown in Figure 5. Each phase consists of a process workstation which consists of the input buffer, processes, and the output buffer. The number of Kanbans is computed [26] as (2):

$$N = \frac{dL+S}{c} \quad (2)$$

where *d* is the average demand, *L* is the lead time, *S* is the safety stock, and *C* is the container size.

To illustrate, the Withdrawal Kanban (WK) is transferred from the succeeding workstation to the preceding workstation and placed in the output buffer of the workstation and the production for the rotary kiln workstation respectively. For further details, the WK3 and PK3 are assigned as the numbers of withdrawal and production for the rotary kiln workstation respectively. The clinker container with Withdrawl Kanban (WK3) is moved from the cement processing station into the rotary kiln station and placed in the output of this workstation. The future state maps for the Kanban approach is shown in Figure 6.

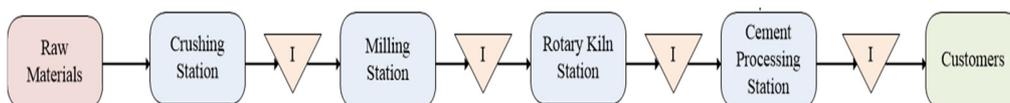


Fig. 4. Push production system.

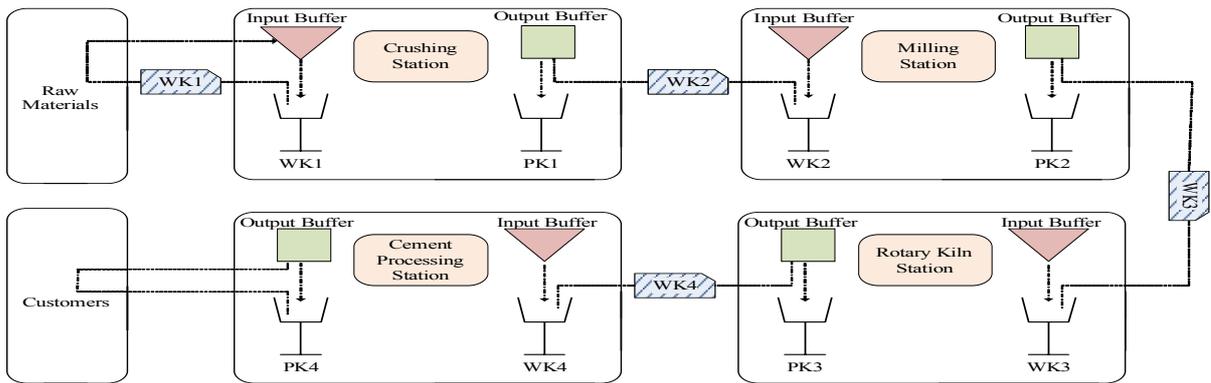


Fig. 5. Proposed Kanban in the cement production line.

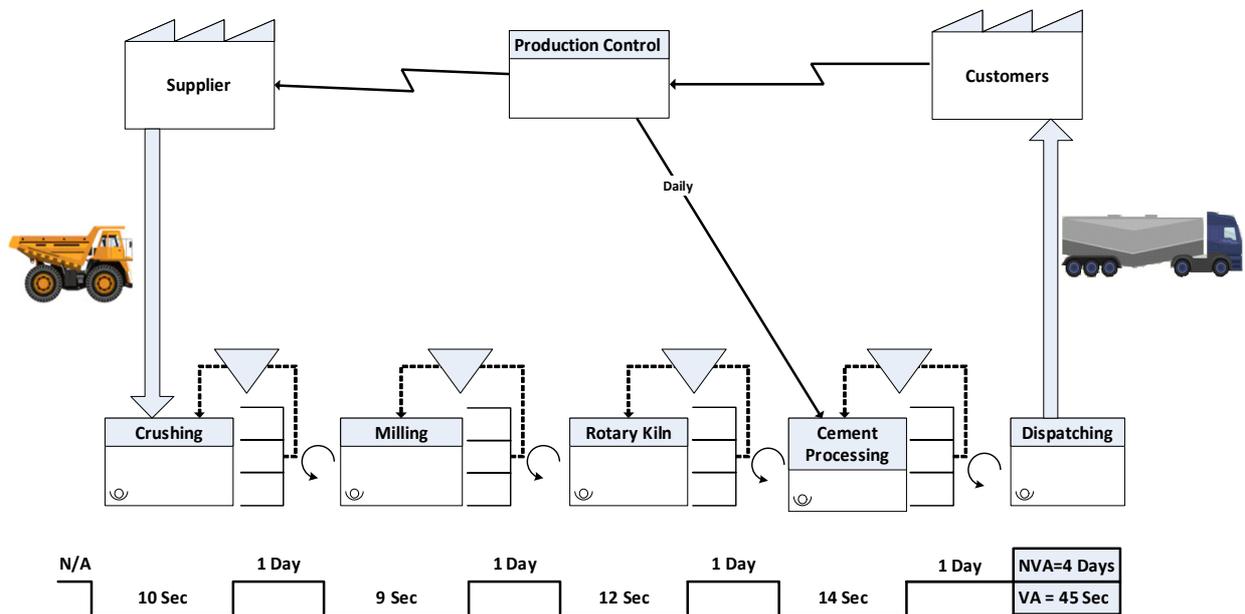


Fig. 6. Proposed VSM future state (Kanban).

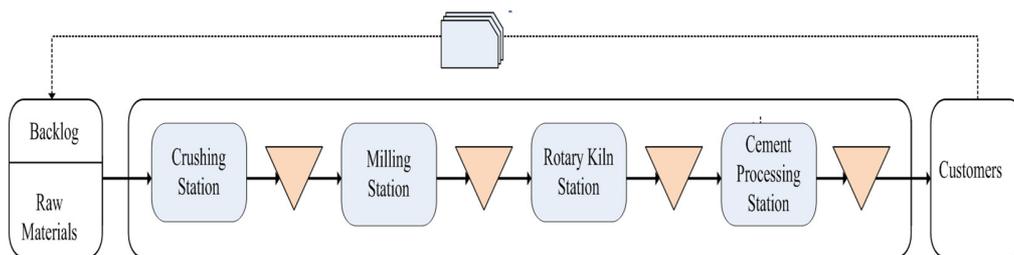


Fig. 7. The proposed CONWIP in the cement production line.

B. CONWIP

Figure 7 shows the proposed mechanism of CONWIP in cement production line. The assumptions are made by identifying the maximum WIP level in the cement production line, to ensure the level of inventory remains constant to the entire shop floor. The material flow directly enters the backlog. The backlog-sequencing rule in this proposed model is First In, First Out (FIFO). The required quantity of finished cement is dispatched to the customer. Then, the next customer orders can

be initiated. The benefit gains of reducing the work cycle time according to the Little's law are [27]:

$$WIP = CT \times TH \quad (3)$$

where TH is the throughput time that is the average output of a production process per unit time.

The future state map for the CONWIP approach is illustrated in Figure 8.

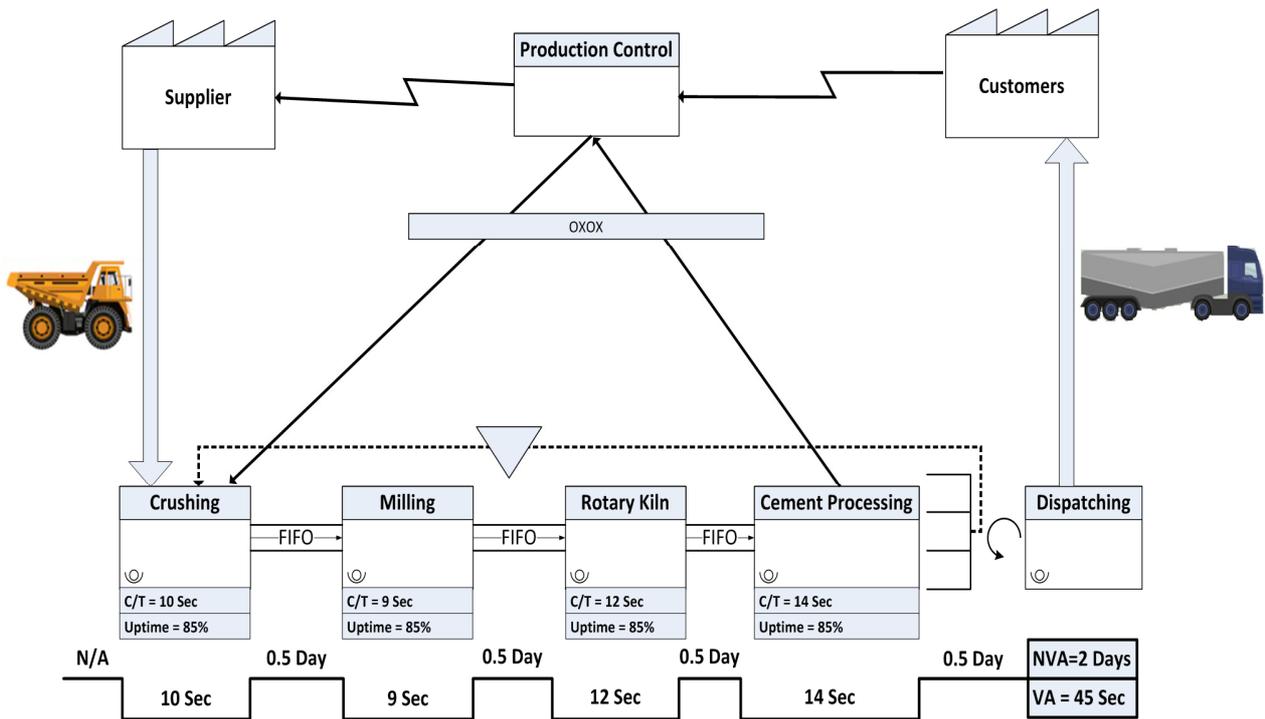


Fig. 8. Proposed VSM future state (CONWIP).

IV. RESULTS AND DISCUSSION

The comparison between the current value stream map and the proposed future maps is presented in Figure 9.

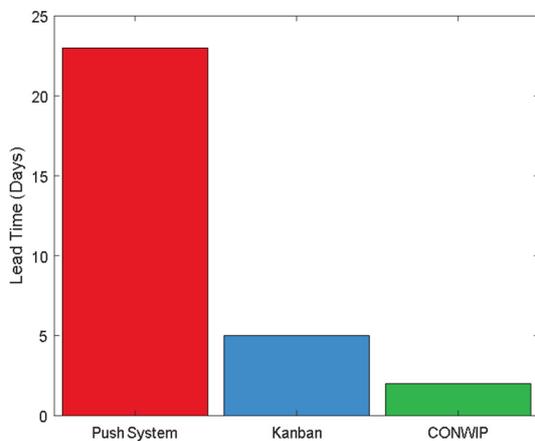


Fig. 9. Comparison of the push system and the pull systems.

It can be seen that the longest value of lead-time is about 23 days in the push production system, while in the proposed maps they are about 4 and 2 days for Kanban and CONWIP approaches respectively. The comparative results of non-value times in the cement production line for the proposed future maps is decreased by 50% in the scenario of CONWIP approach. The reason behind that is that there is no stock left between workstations in CONWIP. To illustrate, the finished cements are pulled out from the cement processing station which makes the production flows smoothly in the production

line. In the case of push production system, the lead time is fixed as the material requirements are scheduled. The relationship between the level of work in progress, flow time and throughput rate is given in [27]. Due to the increase in the WIP level, flow time increases. Consequently, the increase in throughput hinders and delays the production process and as a result, the production may get congested. Therefore, the recommendation is to implement the CONWIP approach to control the level of inventory in the cement production line.

A significant contribution of this paper is the study of the impact of the pull-based production system on the cement production line. It was shown that Kanban and CONWIP are tools capable of improving the shop floor process. They are able to control the excess inventories of clinker and cement and to eliminate the non-value activities in the cement-manufacturing environment. The outcome of this research is consistent with the foregoing findings. Authors in [21-25] reported that CONWIP is an appropriate approach to control the level of inventory. In addition, the findings could extend the methodology of [14] to implement lean production in cement industry. The future prospect is to extend this research further in order to implement a plan to construct the proposed VSM towards a cement production line.

V. CONCLUSION

The major issue in cement production process is the large inventories of clinker and cement. To cater and eliminate this issue, manufacturing strategies are applied. The VSM is a significant tool to evaluate the current state of manufacturing processes. The pull production systems (Kanban and CONWIP) produce the right quantities in the right time by

controlling the inventory level. Both the proposed models are capable of improving the environment of cement manufacturing processes.

There are further possibilities for future research. The design of experiments to run simulation models should be considered. In addition, the possibility of applying the lean manufacturing model in cement Industry 4.0 should be explored. It is important to understand the dynamic manufacturing environment of the cement industry. The scheduling for dispatching products to customers based on the demand or the identification of time waste and bottlenecks during the cement production processes could be arranged by the integrated lean production with value stream tools to represent the real time and dynamic value chain of information and materials flows.

ACKNOWLEDGMENT

The author wishes to acknowledge the support of this research study by the grant No. ENG-2017-1-7-F-7137 from the Deanship of Scientific Research of Northern Border University, Arar, Kingdom of Saudi Arabia.

REFERENCES

- [1] G. A. Rankin, "Portland cement," *Journal of the Franklin Institute*, vol. 181, no. 6, pp. 747–784, Jun. 1916, [https://doi.org/10.1016/S0016-0032\(16\)90148-0](https://doi.org/10.1016/S0016-0032(16)90148-0).
- [2] J. F. Ryan, "The story of Portland cement," *Journal of Chemical Education*, vol. 6, no. 11, pp. 1854–1868, Nov. 1929, <https://doi.org/10.1021/ed0006p1854>.
- [3] P. E. Halstead, "The Early History of Portland Cement," *Transactions of the Newcomen Society*, vol. 34, no. 1, pp. 37–54, Jan. 1961, <https://doi.org/10.1179/tns.1961.003>.
- [4] V. A. Kaiser, "Computer control in the cement industry," *Proceedings of the IEEE*, vol. 58, no. 1, pp. 70–77, Jan. 1970, <https://doi.org/10.1109/PROC.1970.7541>.
- [5] D. Le Sar, "The production of sulphate-resisting and marine cements from local raw materials," Ph.D. dissertation, Department of Chemistry, The University of Cape Town, Cape Town, South Africa, 1964.
- [6] A. Sicakova, E. Kardosova, and M. Spak, "Perlite Application and Performance Comparison to Conventional Additives in Blended Cement," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5613–5618, Jun. 2020, <https://doi.org/10.48084/etasr.3487>.
- [7] M. E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*. New York, NY, USA: Free Press, 1998.
- [8] M. T. Fernandes, "Value Analysis: Going into a Further Dimension", *Engineering, Technology & Applied Science Research*, Vol. 5, No. 2, pp. 781-789, 2015, <https://doi.org/10.48084/etasr.535>.
- [9] T. Ohno and N. Bodek, *Toyota Production System: Beyond Large-Scale Production*, 1st ed. Cambridge, MA, USA: Productivity Press, 1988.
- [10] J. P. Womack, D. Jones, D. Roos, *The Machine That Changed the World*. New York, NY, USA: Rawson Associates, 1990.
- [11] M. Rother, J. Shook, J. Womack, and D. Jones, *Learning to See: Value Stream Mapping to Add Value and Eliminate MUDA*, 1st ed. Cambridge, MA, USA: Lean Enterprise Institute, 1999.
- [12] J. K. Liker, *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*, New York, NY, USA: McGraw-Hill, 2004.
- [13] T. Tourki, "Implementation of Lean Within the Cement Industry," Ph.D. dissertation, De Montfort University, Leicester, UK, 2010.
- [14] S. Deo, "Methodology to implement lean manufacturing in cement industry," M.S. thesis, Lamar University, Beaumont, TX, USA, 2013.
- [15] M. Graisa and A. Al-Habaibeh, "An investigation into current production challenges facing the Libyan cement industry and the need for innovative total productive maintenance (TPM) strategy," *Journal of Manufacturing Technology Management*, vol. 22, no. 4, pp. 541–558, Jan. 2011, <https://doi.org/10.1108/17410381111126445>.
- [16] A. M. Saifuddoha, M. A. Habib, S. Y. Sumi, S. Jennurine, and M. S. Islam, "Minimization of waste by applying value stream mapping in the supply chain of cement industry," *IOSR Journal of Business and Management*, vol. 9, no. 3, pp. 79–84, 2013, <https://doi.org/10.9790/487X-0937984>.
- [17] M. I. Qureshi, M. Iftikhar, M. N. Bhatti, T. Shams, and K. Zaman, "Critical elements in implementations of just-in-time management: empirical study of cement industry in Pakistan," *SpringerPlus*, vol. 2, no. 1, Dec. 2013, Art. no. 645, <https://doi.org/10.1186/2193-1801-2-645>.
- [18] E. Alghadafi and M. Latif, "Simulation of a Libyan Cement Factory," in *Proceedings of the World Congress on Engineering 2010 Vol III*, London, UK, Jul. 2010.
- [19] B. Noche and T. Elhasia, "Approach to Innovative Supply Chain Strategies in Cement Industry; Analysis and Model Simulation," *Procedia - Social and Behavioral Sciences*, vol. 75, pp. 359–369, Apr. 2013, <https://doi.org/10.1016/j.sbspro.2013.04.041>.
- [20] E. Amrina and A. A. A. Lubis, "Minimizing Waste Using Lean Manufacturing: A Case in Cement Production," in *4th International Conference on Industrial Engineering and Applications (ICIEA)*, Nagoya, Japan, Apr. 2017, <https://doi.org/10.1109/IEA.2017.7939181>.
- [21] R. P. Marek, D. A. Elkins, and D. R. Smith, "Understanding the fundamentals of Kanban and CONWIP pull systems using simulation," in *Proceeding of the 2001 Winter Simulation Conference*, Arlington, VA, USA, Dec. 2001, vol. 2, pp. 921–929, <https://doi.org/10.1109/WSC.2001.977394>.
- [22] M. L. Spearman, D. L. Woodruff, and W. J. Hopp, "CONWIP: A pull alternative to kanban," *International Journal of Production Research*, vol. 28, no. 5, pp. 879–894, 1990, <https://doi.org/10.1080/00207549008942761>.
- [23] M. Bertolini, M. Braglia, G. Romagnoli, and F. Zammori, "Extending value stream mapping: The synchro-MRP case," *International Journal of Production Research*, vol. 51, no. 18, pp. 5499–5519, 2013, <https://doi.org/10.1080/00207543.2013.784415>.
- [24] M. L. Spearman and M. A. Zazanis, "Push and pull production systems. Issues and comparisons," *Operations Research*, vol. 40, no. 3, pp. 521–532, 1992, <https://doi.org/10.1287/opre.40.3.521>.
- [25] W. J. Hopp and M. L. Spearman, *Factory Physics*, 3rd ed. New York, NY, USA: McGraw-Hill, 2008.
- [26] L. J. Krajewski and L.P. Ritzman, *Operations Management: Strategy and Analysis*, 6th ed. New Jersey, NJ, USA: Prentice Hall, 2001.
- [27] J. D. C. Little, "A Proof for the Queuing Formula: $L = \lambda W$," *Operations Research*, vol. 9, no. 3, pp. 383–387, Jun. 1961, <https://doi.org/10.1287/opre.9.3.383>.

The Usability Testing of SSAAT, a Bioinformatic Web Application for DNA Analysis at a Nucleotide Level

Victor Mero

School of Computational and Communication Science and Engineering
Nelson Mandela African Institution of Science and Technology (NM-AIST)
Arusha, Tanzania
merov@nm-aist.ac.tz

Dina Machuve

School of Computational and Communication Science and Engineering
Nelson Mandela African Institution of Science and Technology (NM-AIST)
Arusha, Tanzania
dina.machuve@nm-aist.ac.tz

Abstract—Sanger sequencing remains the cornerstone method for Deoxyribonucleic Acid (DNA) sequencing due to its high accuracy in targeting smaller genomic regions in a larger number of samples. The analysis of Sanger sequence DNA data requires powerful and intelligent software tools. Most of the preferred tools are proprietary licensed tools that offer a user-friendly interface and have many features, however, their affordability, especially to individual scientists or students, is limited. On the other hand, a few free and open-source licensed tools are available but have limited features. This study focuses on the usability testing of the developed Sanger Sequence Automatic Analysis Tool (SSAAT), a free and open-source web tool for Sanger sequence analysis. Usability tests were conducted with potential users and the results demonstrate that the participants were able to use the tool easily and accomplish the test tasks at the given time. Moreover, the participants were excited with the easy-to-use interface and agreed that most users could use the tool with no need for technical assistance. However, the participants also identified some issues that require more development effort.

Keywords—Sanger sequence; usability; bioinformatics; web tool

I. INTRODUCTION

A. Background

Sanger sequencing technique is one of the most famous methods used for determining nucleotide sequences in DNA [1], due to its high sequencing accuracy compared to the Next Generation Sequencing (NGS) technologies and its efficiency in sequencing short fragments of DNA, ranging from 200 base pairs (bp) to around 1,000 bp. Sanger sequencing is extensively used to the fields of functional and comparative genomics, evolutionary genetics, and complex disease research. Particularly, the method was employed in sequencing the first human genome in 2000 [2]. The Sanger sequencing process is composed of a pipeline from the DNA extraction to the generation of a chromatogram which is stored as a file called AB1. This process can be seen in [3].

The Sanger sequencing quality relies on the "base calling quality", i.e. the relative certainty with which the nucleobases are determined [2]. Assessing the base-calling accuracy is usually performed using the visual inspection of the sequence trace chromatogram. Most often, proprietary software like CLC Genomics Workbench (Qiagen), SeqMan (DNASTAR), etc. are preferred due to their user-friendly interfaces and the features they provide. Nevertheless, some free open-source software tools for Sanger sequence analysis exist. Phred was among the earliest base-calling software tools reported to have less error rate than the ABI machine software [4-6]. However, Phred was developed as an open-source resource but it is not freely available [7]. Tracy on the other hand is a free open-source tool for Sanger sequence analysis that performs base-calling and other tasks including sequence alignment, assembly, and deconvolution of Sanger chromatogram trace files, all in a command-line interface [8]. Moreover, SangerseqR [9], Automated Sanger Analysis Pipeline (ASAP) [10], and SeqTrace [7] are also reported open-source tools for performing Sanger sequence data analysis but have limited graphical user interface and cross-platform capabilities. Some web tools are also available including Indelligent, CHILD, and Mixed Sequence Reader they are limited to a single feature usage [11-13]. The Sanger Sequence Automatic Analysis Tool (SSAAT), unlike the aforementioned tools, was developed as a web-based tool to eliminate the cross-platform issues while providing an easy-to-use interface and more DNA analysis features.

B. Usability

Usability refers to how easily a user of a specific product or design can use it to accomplish the intended goals effectively, efficiently, and acceptably [14]. In the field of human-computer interaction, usability is defined as a way to remove all possible frustrations that users may experience when using a product or design. On the other hand, usability evaluation refers to a method used in the central design to assess a product or design by testing it with a group of representative users [15], and a platform for users to give direct feedback and

recommendations [16]. Usability is a result of the basic quality components which are learnability, efficiency, memorability, error tolerance, and satisfaction [17]. There are several methods for testing products and among them, usability testing and heuristic evaluation have been the most appropriate ones [18]. Heuristic evaluation is mostly done by professionals who use generally accepted guidelines to evaluate the usability of the product through demos and report issues. In contrast, usability testing recruits users to evaluate a particular product's usability through their feedback after interacting with it [18]. The current study aims to explore the usability of SSAAT, which was developed as a user-friendly web tool for analyzing Sanger sequence data at the nucleotide level. The study focused on the task completion rates, the meantime to complete tasks, and system usability.

II. MATERIALS AND METHODS

A. SSAAT Tool Description

SSAAT was created to provide an easy-to-use interface for molecular biologists to perform Sanger sequence DNA analysis. SSAAT reads the DNA input files, performs base-calling, and provides chromatogram visualization, DNA sequence alignment, and polymorphism detection, and delivers a structured report of the analysis. The flow diagram of SSAAT (Figure 1) illustrates the working flow of the tool.

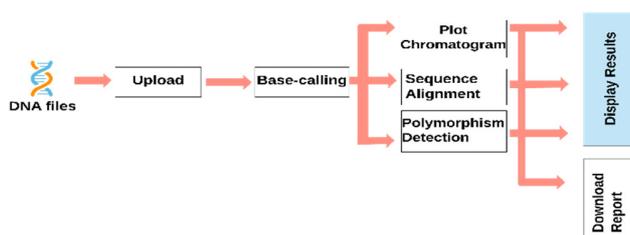


Fig. 1. Flow diagram of SSAAT.

SSAAT utilizes a web-based interface that allows the molecular biologist to perform DNA analysis in a user-friendly graphical interface and also enables working in cross-platform environments. One important element in accomplishing the goals of SSAAT is to ensure that its design satisfies the needs of the users and is flexible to handle the continued evolution of this emerging area.

B. Usability Testing Methodology

In this study, both qualitative and quantitative methods were used to capture user interactivity with the web tool. Qualitative data were collected through a Likert scale questionnaire while quantitative data, such as total users who were able to complete all the tasks, total complete tasks, complete task time, etc., were also collected. The testing session was designed in a way that each participant performed all the tasks and summative assessment was done to examine and evaluate participants' insights. The study aimed at capturing the indicating factors of usability such as learnability, efficiency, usefulness, and satisfaction through the conducted test session.

1) Participants and Duration

A total of 15 individuals participated in the usability test sessions. Based on previous studies, 5-20 participants are a valid sample for usability testing [19]. The first 3 participants were scheduled on the first day and were regarded as a pilot for the next sessions. The testing sessions were conducted for 5 days (3 participants per day). The duration for each session was 60 minutes and after every session, a break period of 60 minutes was given. The session duration was based on previous studies on usability which suggest that 60-90 minutes time is valid for test sessions [19]. During the testing sessions, the moderator provided a brief overview of the test session and requested the participants to fill in a pre-test questionnaire in order to collect the general data. The participants then read the task instructions and began to perform the tasks on the tool using a web browser. As soon as the participants completed all the tasks, the moderator requested the participants to rate the web tool (SSAAT) using the Likert scale questionnaire. This was done as a post-test session in order to find out more information about the overall usability of the web tool.

2) Tasks

Table I presents the results of the tasks that were obtained during the test sessions. Each participant was required to attempt the tasks and the moderator observed the time of completion and participant's behavior while attempting the tasks.

TABLE I. SUMMARY OF USABILITY TEST RESULTS BY TASK

Code	Tasks	Baseline time (min)/estimated time (min)	Mean time (min)	Completion rate (%)
Task 1	Identify the use of the web tool.	3/5	0.56	100
Task 2	File upload, view the sequence quality, and download the extracted sequence as a FASTA file.	3/5	4	87
Task 3	Navigate to Chromatogram Viewer, trim the 5' end 50 base and trim 3' end 100 base and download the chromatogram as a PDF file.	7/10	7.3	73
Task 4	Upload a reference sequence and calculate the global/local sequence alignment with the previously uploaded file as a primary sequence.	7/10	9	60
Task 5	Generate a report with sequence detail, chromatogram quality score plot, and sequence alignment results.	3/5	3.1	73

III. RESULTS AND DISCUSSION

A. Participant Characteristics

Among the 15 participants, there were 7 masters and three PhD students in life sciences, 3 molecular biologists, and 2 bioinformaticians. Eight were women and 7 were men, aged between 21 to 47 years old. All the participants claimed to use the internet daily. None of the participants received formal training or had the chance to review a user guide before participating in the usability test.

B. Task Completion Rate Results

All participants successfully completed Task 1 marking a 100% completion rate. Task 2 was completed by 87% of the participants. Among the long duration tasks, Task 3 scored 73% completion rate, while Task 4 scored 60% completion rate, and Task 5 was successfully completed by the 73% of the participants. The task definitions, completion rate, and mean time are illustrated in Table I.

C. Mean Time to Complete Task Results

The moderator recorded the task execution time for each participant. The allocated time for each task ranged from 5 to 10 minutes where simple tasks were allocated with less time and lengthy tasks with more. Task 1 had the shortest completion time with a mean time of 0.56 minutes. This was followed by Task 5 and Task 2, with times of 3.1 minutes and 4 minutes respectively. Task 3 and Task 4 were the longest to complete with mean times of 7.3 minutes and 9 minutes respectively. The overall completion time ranged from 0.56 to 9 minutes, with a commonly recorded time of less than 5 minutes for the majority of tasks.

D. System Usability Survey Results

During the post-testing session, the participants were asked to rate the web tool to capture the general usability aspects of the SSAAT. The detailed results are shown in Table II. The measurements which were captured from participants' post-test questionnaires included:

- Its ease of use
- If the users would prefer to use the web tool
- Its learnability
- If assistance from technical personnel was required
- System functionality integrations
- If the participants would recommend the tool to a colleague

The majority of the participants (86.67%) agreed that the web tool was easy to use. Additionally, most participants (93%) reported that they would prefer to work with the web tool often. Regardless of the higher scores of participants agreeing that the tool was easy to use, 40% of them agreed that technical assistance was needed to operate the tool effectively. More than half of the participants agreed that the integrated features were functioning well. Lastly, the majority of participants reported that they would recommend the tool to a colleague.

E. Discussion

The usability of SSAAT a bioinformatic tool for DNA analysis at the nucleotide level was assessed in this study. Our findings suggest that SSAAT is easy and learnable, and even new users may be able to use the tool without prior exposure or technical assistance and accomplish the required tasks at a given time. The participants identified a number of possible improvements to the tool such as the addition of batch processing capabilities, trace file editing, and connection to remote DNA databases. However, the mentioned suggestions would require more development effort and time, therefore we plan to work on them in the future versions. Other modifications such as the suggested lighter interface background colors and the creation of a user guide with some visual illustrations were easier and more straightforward to implement.

TABLE II. SYSTEM USABILITY SURVEY RESULTS

Statement	Response	Frequency
"The web tool is user friendly"	Strongly agree	5 (33.33%)
	Agree	8 (53.33%)
	Neutral	2 (13.33%)
"I would like to use web tool often"	Strongly agree	4 (26.67%)
	Agree	10 (66.67%)
	Neutral	1 (6.67%)
"I think most of the users will be able to use the web tool fast"	Strongly agree	4 (26.67%)
	Agree	8 (53.33%)
	Neutral	1 (6.67%)
	Disagree	2 (13.33%)
"I will not need technical assistance to be able to use the web tool"	Agree	7 (46.67%)
	Neutral	2 (13.33%)
	Disagree	4 (26.67%)
	Strongly disagree	2 (13.33%)
"I think the web tool units/parts are well integrated"	Strongly agree	4 (26.67%)
	Agree	6 (40%)
	Neutral	4 (26.67%)
	Disagree	1 (6.67%)
"I will recommend this web tool to my colleagues"	Strongly agree	4 (26.67%)
	Agree	9 (60%)
	Neutral	2 (13.33%)

IV. CONCLUSION

The primary focus of this study was to examine the usability of the developed SSAAT tool. This is an attempt towards eliminating the barriers to the availability of free and user-friendly software for Sanger sequence DNA analysis. The usability assessment results suggest that most of the users will be able to use the web tool without assistance. This is a good indicator that the tool is easy to use and hence most of the users are likely to often use it in their work and would probably recommend it to their colleagues.

The participants in this usability study encountered several minor usability issues while using the prototype of this web tool. To ensure the effective use of SSAAT, the issues identified during the usability testing sessions will be addressed in the upcoming version of the tool and the analysis results will affect the future development. Usability evaluations are invaluable to the success of technology in an emerging area, especially in a complex domain such as genetics.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support from the Bioinformatics Unit at the Beca-ILRI hub in Nairobi Kenya and CoCSE Laboratory at NM-AIST in Arusha Tanzania.

REFERENCES

- [1] J. Shendure *et al.*, "DNA sequencing at 40: past, present and future," *Nature*, vol. 550, no. 7676, pp. 345–353, Oct. 2017, <https://doi.org/10.1038/nature24286>.
- [2] B. Ewing, L. Hillier, M. C. Wendl, and P. Green, "Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment," *Genome Research*, vol. 8, no. 3, pp. 175–185, Mar. 1998, <https://doi.org/10.1101/gr.8.3.175>.
- [3] <https://commons.wikimedia.org/wiki/File:Sanger-sequencing.svg>
- [4] "Applied Biosystems Genetic Analysis Data File Format." Applied Biosystems, Jul. 2006.
- [5] B. Ewing and P. Green, "Base-calling of automated sequencer traces using phred. II. Error probabilities," *Genome Research*, vol. 8, no. 3, pp. 186–194, Mar. 1998.
- [6] M. Machado *et al.*, "Phred-Phrap package to analyses tools: a pipeline to facilitate population genetics re-sequencing studies," *Investigative Genetics*, vol. 2, no. 1, Feb. 2011, Art. no. 3, <https://doi.org/10.1186/2041-2223-2-3>.
- [7] B. J. Stucky, "SeqTrace: A Graphical Tool for Rapidly Processing DNA Sequencing Chromatograms," *Journal of Biomolecular Techniques: JBT*, vol. 23, no. 3, pp. 90–93, Sep. 2012, <https://doi.org/10.7171/jbt.12-2303-004>.
- [8] T. Rausch, M. H.-Y. Fritz, A. Untergasser, and V. Benes, "Tracy: basecalling, alignment, assembly and deconvolution of sanger chromatogram trace files," *BMC Genomics*, vol. 21, no. 1, Mar. 2020, Art. no. 230, <https://doi.org/10.1186/s12864-020-6635-8>.
- [9] J. T. Hill and B. Demarest, *sangerseqR: Tools for Sanger Sequencing Data in R*. Bioconductor version: Release (3.12), 2021.
- [10] A. Singh and P. Bhatia, "Automated Sanger Analysis Pipeline (ASAP): A Tool for Rapidly Analyzing Sanger Sequencing Data with Minimum User Interference," *Journal of Biomolecular Techniques: JBT*, vol. 27, no. 4, pp. 129–131, Dec. 2016, <https://doi.org/10.7171/jbt.16-2704-005>.
- [11] C.-T. Chang *et al.*, "Mixed Sequence Reader: A Program for Analyzing DNA Sequences with Heterozygous Base Calling," *The Scientific World Journal*, vol. 2012, Jun. 2012, Art. no. e365104, <https://doi.org/10.1100/2012/365104>.
- [12] D. A. Dmitriev and R. A. Rakitov, "Decoding of Superimposed Traces Produced by Direct Sequencing of Heterozygous Indels," *PLOS Computational Biology*, vol. 4, no. 7, 2008, Art. no. e1000113, <https://doi.org/10.1371/journal.pcbi.1000113>.
- [13] I. Zhidkov, R. Cohen, N. Geifman, D. Mishmar, and E. Rubin, "CHILD: a new tool for detecting low-abundance insertions and deletions in standard sequence traces," *Nucleic Acids Research*, vol. 39, no. 7, Apr. 2011, Art. no. e47, <https://doi.org/10.1093/nar/gkq1354>.
- [14] M. Shitkova, J. Holler, T. Heide, N. Clever, and J. Becker, "Towards Usability Guidelines for Mobile Websites and Applications," in *Wirtschaftsinformatik Proceedings*, Mar. 2015.
- [15] P. T. Koziokas, N. D. Tselikas, and G. S. Tselikis, "Usability Testing of Mobile Applications: Web vs. Hybrid Apps," in *Proceedings of the 21st Pan-Hellenic Conference on Informatics*, New York, NY, USA, Sep. 2017, Art. no. 55, <https://doi.org/10.1145/3139367.3139410>.
- [16] J. R. Bergstrom and A. Schall, Eds., *Eye Tracking in User Experience Design*, 1st ed. Amsterdam, Netherlands; Boston, MA, USA: Morgan Kaufmann, 2014.
- [17] E. Folmer and J. Bosch, "Architecting for usability: a survey," *Journal of Systems and Software*, vol. 70, no. 1, pp. 61–78, Feb. 2004, [https://doi.org/10.1016/S0164-1212\(02\)00159-0](https://doi.org/10.1016/S0164-1212(02)00159-0).
- [18] A. Fernandez, E. Insfran, and S. Abrahão, "Usability evaluation methods for the web: A systematic mapping study," *Information and Software Technology*, vol. 53, no. 8, pp. 789–817, Aug. 2011, <https://doi.org/10.1016/j.infsof.2011.02.007>.
- [19] J. Nielsen, "Usability for the MassesJUS," *Journal of Usability Studies*, vol. 1, no. 1, pp. 2–3, Nov. 2005.

Evaluation of the Variation in Dynamic Load Factor Throughout a Highly Skewed Steel I-Girder Bridge

Yousif Almoosi

Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
yousif@udel.edu

Jennifer McConnell

Department of Civil and Environmental Engineering
College of Engineering
University of Delaware
Delaware, USA
righman@udel.edu

Nazar Oukaili

Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
nazar.oukaili@coeng.uobaghdad.edu.iq

Abstract—The Dynamic Load Factor (DLF) is defined as the ratio between the maximum dynamic and static responses in terms of stress, strain, deflection, reaction, etc. DLF adopted by different design codes is based on parameters such as bridge span length, traffic load models, and bridge natural frequency. During the last decades, a lot of researches have been made to study the DLF of simply supported bridges due to vehicle loading. On the other hand, fewer works have been reported on continuous bridges especially with skew supports. This paper focuses on the investigation of the DLF for a highly skewed steel I-girder bridge, namely the US13 Bridge in Delaware State, USA. Field testing under various load passes of a weighed load vehicle was used to validate full-scale three-dimensional finite element models and to evaluate the dynamic response of the bridge more thoroughly. The results are presented as a function of the static and dynamic tensile and compressive stresses and are compared to DLF code provisions. The result shows that most codes of practice are conservative in the regions of the girder that would govern the flexural design. However, the DLF sometimes exceeds the code-recommended values in the vicinity of skewed supports. The discrepancy of the DLF determined based on the stress analysis of the present study, exceeds by 13% and 16% the values determined according to AASHTO (2002) for tension and compression stresses respectively, while, in comparison to BS5400, the differences reach 6% and 8% respectively.

Keywords—dynamic load factor; steel bridges; skewed bridges; codes of practice; field test; finite element analysis

I. INTRODUCTION

Skewed supports occur when the supporting abutments for the girders are not normal to the girder lines. This may be required due to the characteristics of the intersecting roadways or due to the geological terrain. Since skew angles increase the interaction between the steel girders and the cross-frames or

diaphragms, the behavior of bridges with skewed supports becomes more complicated than that in bridges with normal supports. Differential girder deflections, obtuse and acute angle corners, unbalanced construction loads, and the possibility for unequal girder lengths are among other characteristics of skew bridges that subject the girders to significant torsional and lateral moments induced by moving truck loads, which may affect the dynamic responses. More accurate knowledge of the dynamic responses will lead to safer and more economical designs of new bridges and to more effective evaluation of existing ones. DLF is defined as the ratio between the maximum dynamic and static responses [1]:

$$\text{Dynamic Load Allowance (DLA)} = \frac{D_{dyn} - D_{stat}}{D_{stat}} \quad (1)$$

$$\text{Dynamic Load Factor (DLF)} = 1 + \text{DLA} = \frac{D_{dyn}}{D_{stat}} \quad (2)$$

where D_{dyn} and D_{stat} are the dynamic and static responses of the bridge girders, respectively (e.g. deflections, strains, stress, or reactions). In the present study, the responses refer to the maximum tensile and compressive stresses respectively.

Today, design codes and specifications recommend the use of a dynamic load factor as a function of span length, natural frequency, and traffic load models, regardless of the type of bridge. Studies have also shown that DLF is influenced by parameters such as the road surface condition, vehicle weight, speed, the number of axles, and bridge materials and type. The American Association of State Highway and Transportation Officials (AASHTO (2002)) [2] specified the Dynamic Load Allowance (DLA) as a function of the bridge span length:

$$\text{DLA (IM)} = \frac{15.24}{L_{(\text{in meters})} + 38.10} \leq 30\% \quad (3a)$$

Corresponding author: Yousif Almoosi

$$DLA (IM) = \frac{50}{L(\ln feet)^{+125}} \leq 30\% \quad (3b)$$

where L is the clear span length measured from center to center of the bearing supports. This empirical equation has been in effect since 1944. Similar to the expression given by AASHTO, the Japan Road Association also defines the DLA as a function of the bridge span length [3]. However, in the new AASHTO LRFD [4], for strength considerations, the DLA adopted is equal to 0.33 of the truck effects.

The 1983 Ontario Highway Bridge Design Code [5], and the 1992 Australia's Bridge Design Code [6], both specify the DLF as a function of the first longitudinal natural frequency of the bridge. The DLA values fall between 0.2 and 0.4, with the higher values corresponding to frequencies between 2.5 and 4.5Hz. In some codes such as the 2003 EUROCODE [7], and the 2006 BS5400-2 [8], two types of loading for highway bridges are considered, namely the HA and HB loading for normal and abnormal traffic loads respectively. For both loading types, a constant DLA of 0.25 is used. The review of the aforementioned codes also reveals what factors are not considered in DLF. These include different bridge types, skews, curvatures, and cross-section properties (except for how these variables may influence the natural frequency in the Ontario and Australian specifications) and the fact that a single DLF is used for the entire bridge versus various points in the bridge having different DLFs. There is a large volume of research data on impact factors and bridge-vehicle interaction. The most relevant ones to this study concern the verification of impact factor relations introduced in design codes, and comparison with empirical and numerical results. Authors in [9] implemented a study on the Dynamic Amplification Factor (DAF) by evaluating continuous beam bridges and conducting vehicle bridge interaction analysis. The result showed that when resonance condition was approached, the DAF increased, and therefore, vehicle speed influenced the DAF in the studied girder bridges.

Authors in [10], proposed a method to evaluate the impact factors for fatigue design purposes taking into account the effect of the deterioration of the road surface condition, vehicle velocity, and bridge span length of steel I-girder bridges. Maximum stress range instead of the maximum stress is used to calculate the conventional impact factor. The new impact factor can be used to correlate fatigue-induced damage. A non-linear dynamic simulation was conducted in [1] to determine the dynamic impact for composite steel bridges. The modeling involved rigid bodies (vehicles) connected to nonlinear suspensions. The composite steel girder bridge assembly consists of 3D thin-walled beam elements. The parametric study includes vehicle velocity, span length, elastomeric bearings stiffness, vehicle mass, and eccentricity with respect to deck centerline. The conclusion is that current code-specified values are adequate for normal design situations. Authors in [11] investigated the impact factor for concrete-steel composite I-girder bridges. FEA was conducted using ABAQUS for 120 different bridges, considering parameters such as span length, girders number, and the number of traffic lanes in order to model the traffic loads. The vehicle-bridge interaction was disregarded. The outcome of this research was

that the impact factor for composite bridges based on the AASHTO formula is over-estimated for moment and deflection and is under-estimated for support reaction. Authors in [12] evaluated the impact factor for simple and continuous beams subjected to moving vehicles. The parameters that were taken into account are the frequency ratio of vehicle-bridge interaction, bridge damping, and road pavement roughness. The study revealed that the impact factor outputs of the bridge responses (e.g. moment, support reactions, and deflections) are different and suggested new formulas for the impact factor. The DLFs according to different graphic codes can be seen graphically in [1].

Figure 1 shows the DLF of the bridge that is the subject of the present research compared to the DLFs determined according to the above-mentioned codes as well as the Italian, French, and West German ones. This shows that the span length for the bridge of interest results in the highest variability among the different codes.

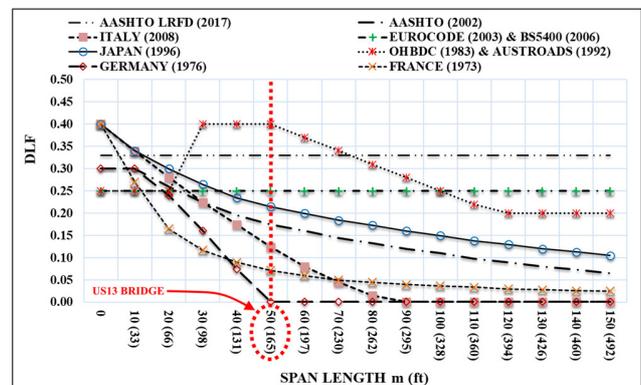


Fig. 1. Dynamic load factors for US13 Bridge according to different practice codes. Data taken from [1].

The current study aims to evaluate the DLF of a highly skewed steel I-girder bridge with perpendicular cross-frames and the variation in DLF along the length of a representative girder. This is accomplished using a bridge previously field-tested to validate a finite element analysis and then using the validated model to simulate various traffic loadings. A case study including a 3D FEA model for the same bridge specifications and dimensions but with staggered cross-frames configuration is implemented to investigate the effect of different arrangements of bracing systems on the bridge dynamic response and thus on DLFs values. This is an extension of the prior field test that suggested that the DLF was highly variable at a small number of points where it was possible to install field instrumentation. Thus, the presentation of results focuses on the computation of DLF at key points along the length of the girders.

II. BRIDGE FIELD TESTING

A. Bridge General Description

The bridge of interest, labeled as the US13 Bridge, is a 65° heavily skewed steel I-girder bridge in Delaware State, USA (Figure 2(a)).

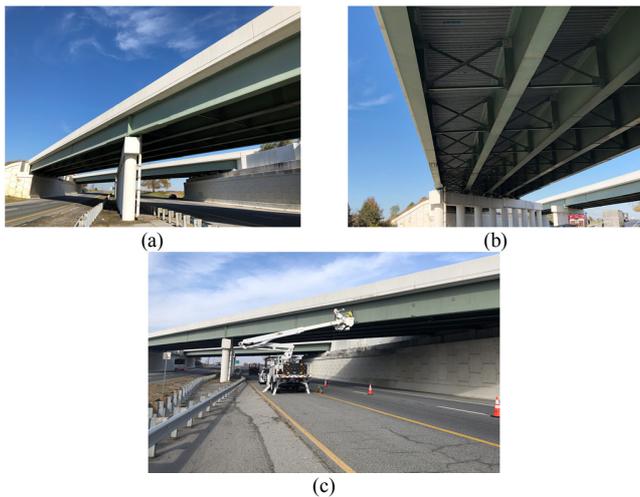


Fig. 2. US13 Bridge in Delaware State, USA, (a) general view, (b) perpendicular cross-frame configuration, (c) gauge installing and lane closing.

Twin spans carry the north- and south-bound lanes. The bridge consists of two continuous spans of equal lengths of 50m (165ft). There are 5 girders spaced 2.9m (9.5ft) on center with exterior girders spaced 0.86m (2.83ft) and 1.16m (3.83ft) away from the outer edge of the bridge concrete guard wall on the west and east sides respectively. Therefore, the total width of the bridge is 13.37m (44.67ft), carrying two 3.65m (12ft) lanes, a 3.65m (12ft) shoulder on the west side, and a 1.82m (6ft) shoulder on the east side. The concrete guard wall located on each side has a width of 0.4m (1.34ft) and a height of 0.86m (2.83ft). The bridge contains inline X-shaped cross-frames (between girders and at pier location) oriented perpendicular to the girders, as shown in Figure 3, which are connected to the girders using full-depth connection plates. Also, inline K-shaped cross-frames are used in the vicinity of the abutment supports.

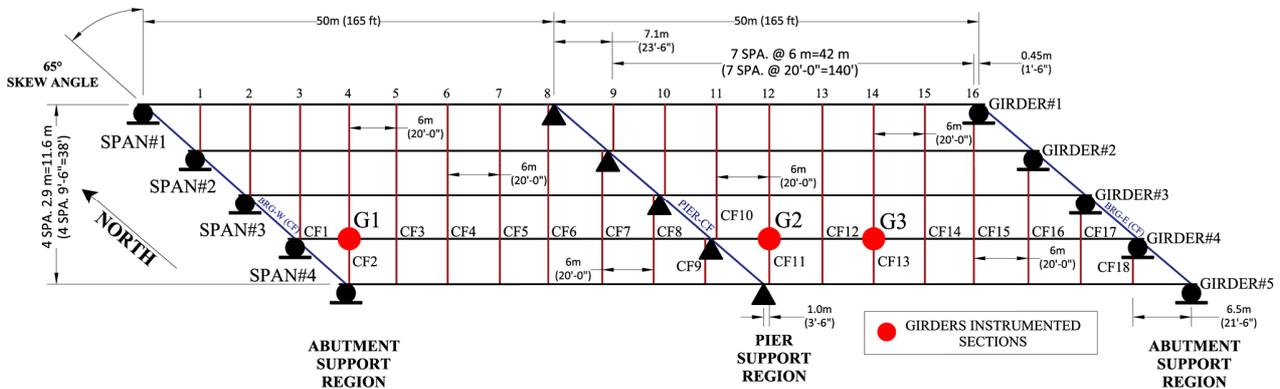


Fig. 3. Framing plan and instrumentation locations for US13 Bridge.

B. Instrumentation Layout and Loading Pattern

Bridge Diagnostics Inc. ST-350 strain gauges (BDI gauges) [13] and their associated data-acquisition system were used in the field test. Specifically, 12 strain gauges were installed on girder #4 at 3 different cross-sections labeled G1, G2, and G3 (Figure 3). Each of these cross-sections was instrumented with 4 strain gauges (Figure 4). One pair of them was placed 5cm (2in) away from the outer edges of the bottom surface of the bottom flange. These positions are labeled as BF-1 and BF-2. The other pair was placed on the opposite sides of the web at approximately mid-height of the web and are labeled W-1 and W-2. One additional pair was used only for the numerical FEA located at the bottom of the web and labeled as WB-1 and WB-2. Three different truck passes, with 24km/hr (15mph) speed, were conducted for the load test (Figure 5). Pass #1 had the loaded truck travel down the center of the left lane. This position was intended to maximize the stress and induce differential deflection in Girder #4. Pass #2 was designed to produce a high level of stress in both Girder #3 and Girder #4, while Pass #3 had the truck travel with the left side wheels aligned with the centerline of the two lanes, intending to

maximize differential deflections between the instrumented girder and the adjacent one [14]. Three additional passes are implemented using FEA and are labeled as Passes #4–#6 which mimic Passes #1–#3 respectively, but with different travel speed of 104km/hr (65mph).

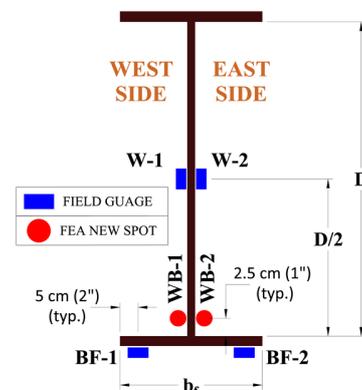


Fig. 4. Instrumentation locations across the cross-section of the US13 Bridge during the field test and FEA.

III. FINITE ELEMENT ANALYSIS

A. Geometry, Meshing, Elements, Materials, and Boundary Conditions

The geometry of the bridge for both static and dynamic FEA models was created according to the structural plans provided by the bridge owner (the Delaware Department of

Transportation). Software like AutoCAD-3D, FEMAP, NX Nastran, and ABAQUS/CAE were used to perform the bridge final finite element model [15] (Figure 6). Over one million four-node reduced-integration shell elements [16], were used for modeling all girders, cross-frames, and stay-in-place profiled metal deck forms in each of the static and dynamic models.

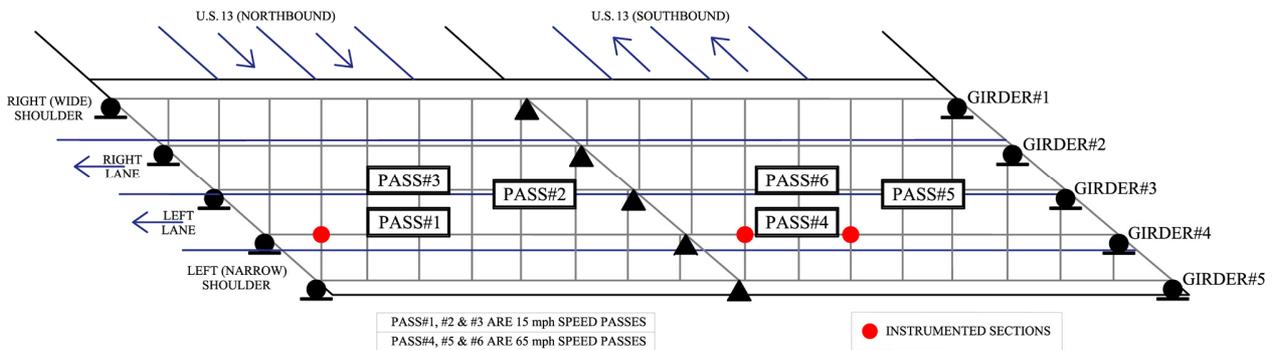


Fig. 5. Truck load passes on the US13 Bridge.

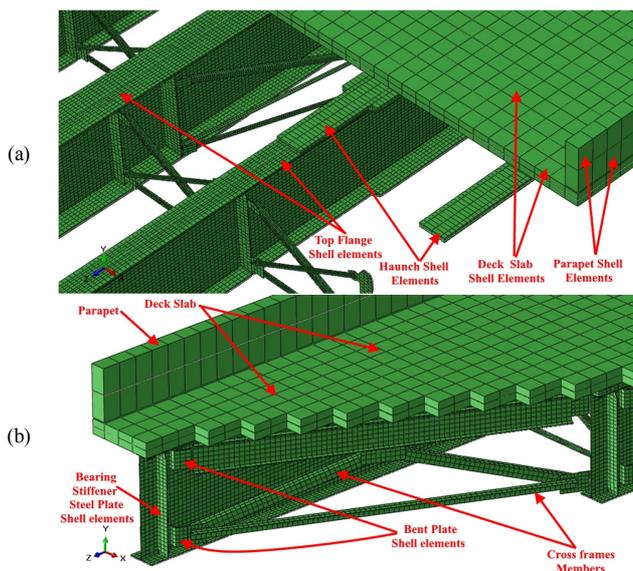


Fig. 6. Abaqus finite element model (static/shell model) for the US13 Bridge: (a) Pier support region, (b) abutment support region.

For modeling the concrete deck, haunch, and concrete guard wall, four-node reduced-integration shell elements (S4R) were used in the static models. For comparison, the dynamic models use three-dimensional eight-node reduced-integration brick elements (C3D8R) [16-18]. Reinforcement in the concrete is defined by Abaqus' rebar option using the actual geometry of the reinforcement and its spacing in both the static and dynamic models. In general, the mesh size of both the concrete deck slab and the metal forms was 30cm×30cm (1ft×1ft), while 8 or 12 elements were used across the width of the girder's top and bottom flanges and 28 elements through the height of each web. Linear isotropic elastic material properties were used for the FEA because the applied loads' result does not cause the proportional limit of the materials to

be exceeded. Expansion bearings, at abutments, were modeled with translation constraints in both vertical and transverse directions at the center node of the bottom flange cross-section of each girder, and only vertical direction constraint for the remaining bottom flange nodes of the abutment cross-sections. Fixed bearings, at the pier, were modeled similar to the expansion bearing constraints except that the center node of the bottom flange cross-sections was also restrained in the longitudinal direction.

B. Loading (Vehicle Modeling)

Two different approaches were used for modeling the static versus dynamic load passes. In the static models, the loading truck was modeled using six-point loads (one to represent each physical wheel of the loading vehicle). A Visual Basic for Applications (VBA) programming routine was used to model the variable nodal positions of the load as the truck traveled across the bridge for each load case. In this approach, only the load being transferred to the bridge due to the vehicle's weight is considered and the inertia of the vehicle is ignored. Previous studies classified moving load modeling (for the dynamic models) into three main types based on the technique used to model the vehicle. In one-dimensional (1D) models [19, 20] the vehicle is modeled as a spring-mass of one or two degrees of freedom, a planar model with multiple axles is considered in two-dimensional (2D) models [21, 22], and there are the 3D complete vehicle models [23, 24]. In this study, the one-dimensional (1D) dynamic model was adopted, where each wheel of the truck was modeled as a moving mass by defining a load versus time history for each axle, which is constant, and a displacement history for the position versus time. Consequently, both load and inertia are used to evaluate the dynamic effects of the moving load. Surface-based contact with a rigid pressure-overclosure relationship and frictionless tangential behavior was used to model the interaction between the bridge and moving masses.

C. Interaction Mechanisms and Analysis

Tie constraints were used to simulate all connections between the steel components of the bridge (e.g. between cross-frame members and vertical connection plates on the girders) for both the static and the dynamic models. The metal decking and the top flange were connected via merged nodes with all degrees of freedom constrained. Timoshenko (shear flexible) beam elements [16] with circular cross-sections and six degrees of freedom (three translational and three rotational) at each node to represent the shear studs and an isotropic friction model with a coefficient of friction of 0.4 at the steel-concrete interface were used to model the steel-concrete interaction mechanism in the dynamic models, while in the static ones, surface-to-surface tie constraints were used to model the connection between the haunch and the slab, and node-to-surface tie constraints were used to model the connection between the top flange and the haunch. The analysis was performed using an Expert Subroutine System programmed to extract key information from the Abaqus output result file [15]. This was implemented using the Caviness High-Performance Computing (HPC) cluster at the University of Delaware, USA. The static analysis was performed using Abaqus standard implicit static analysis while Abaqus explicit dynamic analysis was used for the dynamic models.

IV. FINITE ELEMENT VALIDATION

Figures 7 and 8 illustrate the maximum tensile and compressive stresses respectively, recorded from the field test and the FEA data for all girder strain gauges.

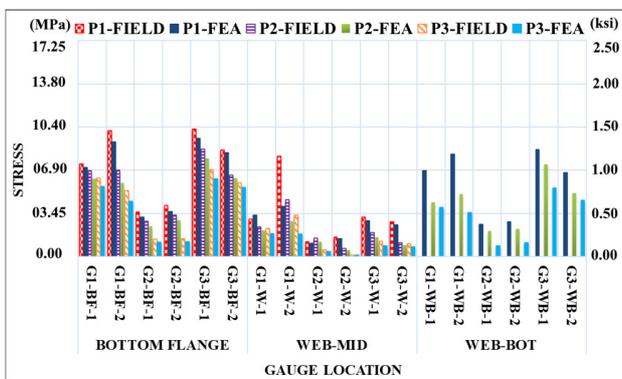


Fig. 7. Maximum tensile stress recorded from field test vs FEA data for all gauges located in the instrumented sections (G1, G2, and G3) due to the three passes.

Generally, the FEA accurately reproduces the general trends observed in the field testing and a favorable quantitative comparison is obtained in most cases, especially for the gauges located in the bottom flanges of the instrumented sections, and even more so when these gauges are in tension during Passes #1 and #2 (for Pass #1, there is 7.7% difference between the FEA and field test for tensile and 10.6% for compressive stresses, for Pass #2, 10.5% for tensile and 14.8% for compressive stresses, while for Pass #3, 13.4% for tensile and 18.9% for compressive stresses). A weaker correlation was

expected and achieved when comparing the compressive stress results since a linear elastic material with infinite tensile strength was used to model the concrete in tension. For the webs, the best overall correlation between these results was achieved during Pass #1 at the G2 gauge location during tension and compression, where the error is 8% or less. The overall average percentage errors are: for Pass #1, 12.43% for tension and 15.97% for compression, for Pass #2, 21.16% for tension and 23.49% for compression, and for Pass #3, 27.36% for tension and 31.28% for compression.

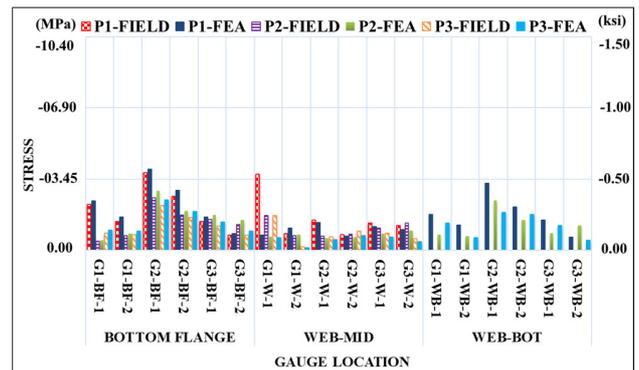


Fig. 8. Maximum compressive stress recorded from field test vs FEA data for all gauges located in the instrumented sections (G1, G2, and G3) due to the three passes.

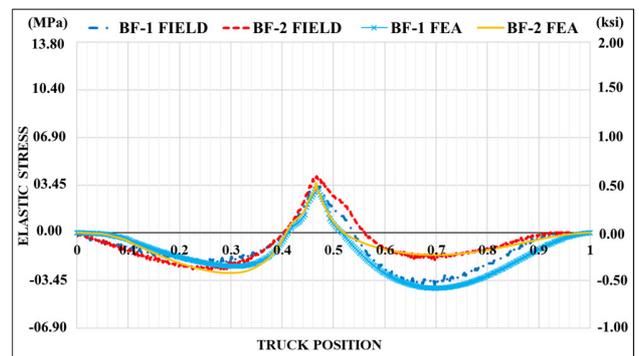


Fig. 9. Field vs FEA stresses influence line for the pair of gauges located in the bottom flange of the instrumented section G2.

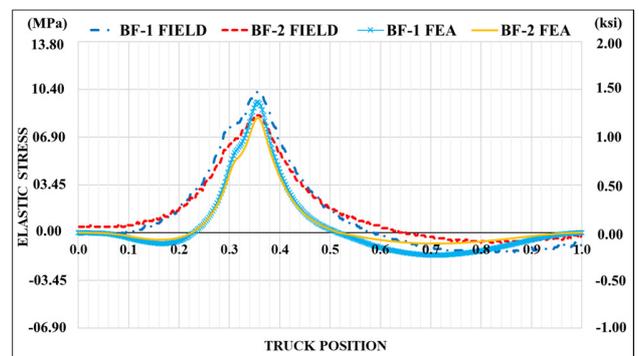


Fig. 10. Field vs FEA stresses influence line for the pair of gauges located in the bottom flange of the instrumented section G3.

Figures 9, 10 show the field results versus the FEA data for the bottom flange gauge positions G2 and G3 respectively in terms of stress versus truck position as it travels across the bridge. These Figures reveal that the FEA results match the expected trends in stress versus time as the truck travels across the bridge, including that the FEA also accurately captures the load locations that causing peak stress. It is worth mentioning that in Figures 9, 10, the x-axis represents the position of the truck as it travels over the two spans of the bridge. The value 0 of the x-axis declares that the truck is over the left support, the value 0.5 indicates that the truck reaches the intermediate support, while the value 1 shows that the truck is over the right support.

V. DYNAMIC LOAD FACTOR RESULTS

A. Dynamic Load Factors for G1, G2, and G3 Sections

Tables I and II illustrate the FEA maximum tensile and compressive stresses for different locations across the height of the cross-sections G1, G2, and G3 during the six Passes.

TABLE I. FEA MAXIMUM TENSILE STRESSES FOR INSTRUMENTED SECTIONS G1, G2, AND G3 DURING THE SIX PASSES

FEA MAXIMUM TENSILE STRESS							
PASSES	Gauge Location	SECTION G1		SECTION G2		SECTION G3	
		BF-1 MPa (ksi)	BF-2 MPa (ksi)	BF-1 MPa (ksi)	BF-2 MPa (ksi)	BF-1 MPa (ksi)	BF-2 MPa (ksi)
PASS#1	Bottom Flange	7.122	9.136	3.130	3.599	9.439	8.288
		(1.033)	(1.325)	(0.454)	(0.522)	(1.369)	(1.202)
		6.226	5.895	2.420	2.889	7.853	6.247
		(0.903)	(0.855)	(0.351)	(0.419)	(1.139)	(0.906)
		5.585	4.420	1.158	1.186	6.219	5.536
		(0.810)	(0.641)	(0.168)	(0.172)	(0.902)	(0.803)
PASS#2	Bottom Flange	9.485	11.935	3.772	4.546	10.251	9.470
		(1.376)	(1.731)	(0.547)	(0.659)	(1.487)	(1.373)
		8.103	7.487	2.859	3.607	8.436	7.067
		(1.175)	(1.086)	(0.415)	(0.523)	(1.224)	(1.025)
		7.111	5.565	1.359	1.464	6.627	6.229
		(1.031)	(0.807)	(0.197)	(0.212)	(0.961)	(0.903)
PASS#3	Bottom Flange	3.309	3.999	1.062	1.420	2.841	2.544
		(0.480)	(0.580)	(0.154)	(0.206)	(0.412)	(0.369)
		2.096	2.827	1.186	0.517	1.538	0.883
		(0.304)	(0.410)	(0.172)	(0.075)	(0.223)	(0.128)
		1.827	1.800	0.393	0.103	0.869	0.772
		(0.265)	(0.261)	(0.057)	(0.015)	(0.126)	(0.112)
PASS#4	Bottom Flange	3.718	4.819	1.172	1.639	3.000	2.740
		(0.539)	(0.699)	(0.170)	(0.238)	(0.435)	(0.397)
		2.326	3.347	1.283	0.585	1.613	0.935
		(0.337)	(0.485)	(0.186)	(0.085)	(0.234)	(0.136)
		1.994	2.099	0.423	0.116	0.899	0.815
		(0.289)	(0.304)	(0.061)	(0.017)	(0.130)	(0.118)
PASS#5	Bottom Flange	6.853	8.184	2.576	2.792	8.523	6.722
		(0.994)	(1.187)	(0.374)	(0.405)	(1.236)	(0.975)
		4.343	4.978	2.033	2.196	7.342	5.060
		(0.630)	(0.722)	(0.295)	(0.318)	(1.065)	(0.734)
		3.934	3.502	0.869	1.091	5.473	4.477
		(0.571)	(0.508)	(0.126)	(0.158)	(0.794)	(0.649)
PASS#6	Bottom Flange	8.722	9.972	3.050	3.257	9.131	7.628
		(1.265)	(1.446)	(0.442)	(0.472)	(1.324)	(1.106)
		5.393	5.985	2.374	2.716	7.773	5.638
		(0.782)	(0.868)	(0.344)	(0.394)	(1.127)	(0.818)
		4.848	4.166	1.012	1.342	5.746	4.961
		(0.703)	(0.604)	(0.147)	(0.195)	(0.833)	(0.720)
PASS#1	Web Mid	3.309	3.999	1.062	1.420	2.841	2.544
		(0.480)	(0.580)	(0.154)	(0.206)	(0.412)	(0.369)
		2.096	2.827	1.186	0.517	1.538	0.883
		(0.304)	(0.410)	(0.172)	(0.075)	(0.223)	(0.128)
		1.827	1.800	0.393	0.103	0.869	0.772
		(0.265)	(0.261)	(0.057)	(0.015)	(0.126)	(0.112)
PASS#2	Web Mid	3.718	4.819	1.172	1.639	3.000	2.740
		(0.539)	(0.699)	(0.170)	(0.238)	(0.435)	(0.397)
		2.326	3.347	1.283	0.585	1.613	0.935
		(0.337)	(0.485)	(0.186)	(0.085)	(0.234)	(0.136)
		1.994	2.099	0.423	0.116	0.899	0.815
		(0.289)	(0.304)	(0.061)	(0.017)	(0.130)	(0.118)
PASS#3	Web Mid	6.853	8.184	2.576	2.792	8.523	6.722
		(0.994)	(1.187)	(0.374)	(0.405)	(1.236)	(0.975)
		4.343	4.978	2.033	2.196	7.342	5.060
		(0.630)	(0.722)	(0.295)	(0.318)	(1.065)	(0.734)
		3.934	3.502	0.869	1.091	5.473	4.477
		(0.571)	(0.508)	(0.126)	(0.158)	(0.794)	(0.649)
PASS#4	Web Mid	8.722	9.972	3.050	3.257	9.131	7.628
		(1.265)	(1.446)	(0.442)	(0.472)	(1.324)	(1.106)
		5.393	5.985	2.374	2.716	7.773	5.638
		(0.782)	(0.868)	(0.344)	(0.394)	(1.127)	(0.818)
		4.848	4.166	1.012	1.342	5.746	4.961
		(0.703)	(0.604)	(0.147)	(0.195)	(0.833)	(0.720)
PASS#5	Web Mid	3.309	3.999	1.062	1.420	2.841	2.544
		(0.480)	(0.580)	(0.154)	(0.206)	(0.412)	(0.369)
		2.096	2.827	1.186	0.517	1.538	0.883
		(0.304)	(0.410)	(0.172)	(0.075)	(0.223)	(0.128)
		1.827	1.800	0.393	0.103	0.869	0.772
		(0.265)	(0.261)	(0.057)	(0.015)	(0.126)	(0.112)
PASS#6	Web Mid	3.718	4.819	1.172	1.639	3.000	2.740
		(0.539)	(0.699)	(0.170)	(0.238)	(0.435)	(0.397)
		2.326	3.347	1.283	0.585	1.613	0.935
		(0.337)	(0.485)	(0.186)	(0.085)	(0.234)	(0.136)
		1.994	2.099	0.423	0.116	0.899	0.815
		(0.289)	(0.304)	(0.061)	(0.017)	(0.130)	(0.118)
PASS#1	Web Bot	6.853	8.184	2.576	2.792	8.523	6.722
		(0.994)	(1.187)	(0.374)	(0.405)	(1.236)	(0.975)
		4.343	4.978	2.033	2.196	7.342	5.060
		(0.630)	(0.722)	(0.295)	(0.318)	(1.065)	(0.734)
		3.934	3.502	0.869	1.091	5.473	4.477
		(0.571)	(0.508)	(0.126)	(0.158)	(0.794)	(0.649)
PASS#2	Web Bot	8.722	9.972	3.050	3.257	9.131	7.628
		(1.265)	(1.446)	(0.442)	(0.472)	(1.324)	(1.106)
		5.393	5.985	2.374	2.716	7.773	5.638
		(0.782)	(0.868)	(0.344)	(0.394)	(1.127)	(0.818)
		4.848	4.166	1.012	1.342	5.746	4.961
		(0.703)	(0.604)	(0.147)	(0.195)	(0.833)	(0.720)

Figures 11 and 12 illustrate the DLF results for all gauges installed on G1, G2, and G3 sections, (see Figure 4), due to the static and dynamic passes shown in Figure 5 in terms of maximum tensile and compressive stresses respectively. The goal of this preliminary analysis is to evaluate the passes that produced higher DLF values.

TABLE II. FEA MAXIMUM COMPRESSIVE STRESSES FOR INSTRUMENTED SECTIONS G1, G2, AND G3 DURING THE SIX PASSES

FEA MAXIMUM COMPRESSIVE STRESS							
PASSES	Gauge Location	SECTION G1		SECTION G2		SECTION G3	
		BF-1 MPa (ksi)	BF-2 MPa (ksi)	BF-1 MPa (ksi)	BF-2 MPa (ksi)	BF-1 MPa (ksi)	BF-2 MPa (ksi)
PASS#1	Bottom Flange	-2.386	-1.613	-3.937	-2.903	-1.613	-0.793
		(-0.346)	(-0.234)	(-0.571)	(-0.421)	(-0.234)	(-0.115)
		-0.496	-0.793	-2.882	-1.910	-1.710	-1.462
		(-0.072)	(-0.115)	(-0.418)	(-0.277)	(-0.248)	(-0.212)
		-0.972	-0.931	-2.434	-1.882	-1.358	-0.924
		(-0.141)	(-0.135)	(-0.353)	(-0.273)	(-0.197)	(-0.134)
PASS#2	Bottom Flange	-3.216	-2.207	-4.696	-3.653	-1.797	-0.911
		(-0.466)	(-0.320)	(-0.681)	(-0.530)	(-0.261)	(-0.132)
		-0.656	-1.065	-3.374	-2.379	-1.870	-1.645
		(-0.095)	(-0.155)	(-0.489)	(-0.345)	(-0.271)	(-0.239)
		-1.278	-1.243	-2.795	-2.300	-1.461	-1.028
		(-0.185)	(-0.180)	(-0.405)	(-0.334)	(-0.212)	(-0.149)
PASS#3	Bottom Flange	3.309	-1.069	-1.351	-0.683	-1.145	-1.007
		(-0.106)	(-0.155)	(-0.196)	(-0.099)	(-0.166)	(-0.146)
		-0.662	-0.752	-0.572	-0.627	-0.814	-0.958
		(-0.096)	(-0.109)	(-0.083)	(-0.091)	(-0.118)	(-0.139)
		-0.614	-0.131	-0.510	-0.696	-0.641	-0.407
		(-0.089)	(-0.019)	(-0.074)	(-0.101)	(-0.093)	(-0.059)
PASS#4	Bottom Flange	-0.857	-1.314	-1.535	-0.813	-1.207	-1.134
		(-0.124)	(-0.191)	(-0.223)	(-0.118)	(-0.175)	(-0.165)
		-0.771	-0.917	-0.642	-0.742	-0.851	-1.076
		(-0.112)	(-0.133)	(-0.093)	(-0.108)	(-0.123)	(-0.156)
		-0.702	-0.157	-0.562	-0.816	-0.665	-0.455
		(-0.102)	(-0.023)	(-0.081)	(-0.118)	(-0.096)	(-0.066)
PASS#5	Bottom Flange	6.853	-1.210	-3.240	-2.090	-1.457	-0.643
		(-0.251)	(-0.176)	(-0.470)	(-0.303)	(-0.211)	(-0.093)
		-0.766	-0.670	-2.421	-1.451	-0.816	-1.184
		(-0.111)	(-0.097)	(-0.351)	(-0.211)	(-0.118)	(-0.172)
		-1.311	-0.618	-1.825	-1.732	-1.195	-0.497
		(-0.190)	(-0.090)	(-0.265)	(-0.251)	(-0.173)	(-0.072)
PASS#6	Bottom Flange	-2.235	-1.594	-3.845	-2.531	-1.507	-0.727
		(-0.324)	(-0.231)	(-0.558)	(-0.367)	(-0.219)	(-0.105)
		-0.968	-0.871	-2.807	-1.726	-0.820	-1.305
		(-0.140)	(-0.126)	(-0.407)	(-0.250)	(-0.119)	(-0.189)
		-1.637	-0.791	-2.090	-2.006	-1.166	-0.540
		(-0.237)	(-0.115)	(-0.303)	(-0.291)	(-0.169)	(-0.078)

In general, both Figures 11 and 12 revealed that the ratio between Pass#4 (dynamic) and Pass#1 (static), generates the higher dynamic responses compared to the other passes in terms of DLF. This suggests that the ideal use of the model is when the load is placed as close to the girder of interest as possible (instrumented section), which also has the benefit of producing the highest stresses in the member of interest. The average DLF value for the spots investigated in the web's middle region is 1.10, while 1.17 and 1.21 are the DLFs for the web's bottom region and the spots at the bottom flange cross-section respectively. In Figures 11-12, the data series titled "AASHTO LRFD" refer to the value of 1.33 (DLA=0.33 ⇒ DLF=1+DLA=1.33), used by AASHTO LRFD specifications.

DLF of 1.25 and 1.4 was used for the data series "BS5400" and "OHBD" respectively. The series titled "AASHTO (US13)" refers to the output of (3) using US13 Bridge span Length.

B. Dynamic Load Factors for Girder Cross-Sections Adjacent to Cross-Frames Located at Girder#4

The preliminary investigation revealed that Passes #1 and #4 result in higher values of DLF especially for the gauges located at the bottom flange section of the instrumented girder (Girder #4).

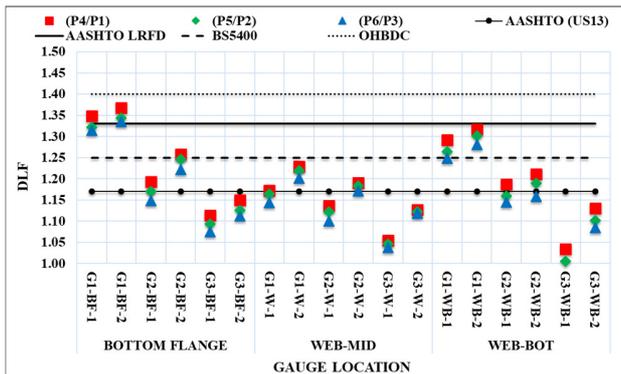


Fig. 11. FEA dynamic load factor values for instrumented sections G1, G2, and G3 in terms of tensile stresses.

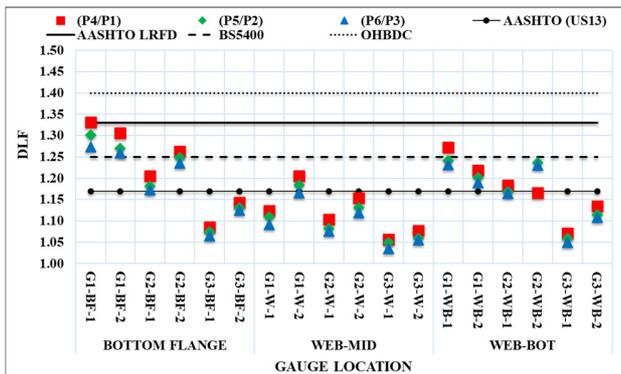


Fig. 12. FEA dynamic load factor values for instrumented sections G1, G2, and G3 in terms of compressive stresses.

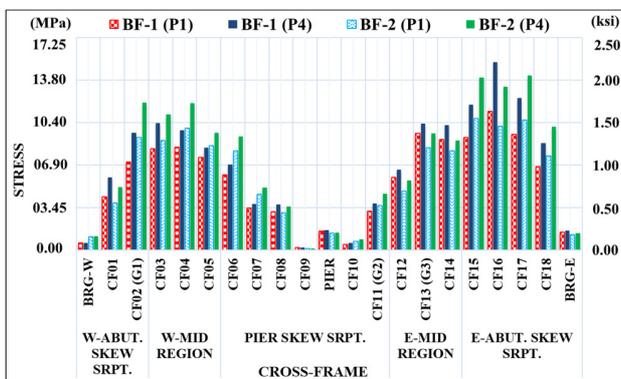


Fig. 13. FEA maximum tensile stresses for BF-1 and BF-2 locations at girder cross-sections adjacent to cross-frames connected to Girder #4 due to truck Pass#1 and Pass#4.

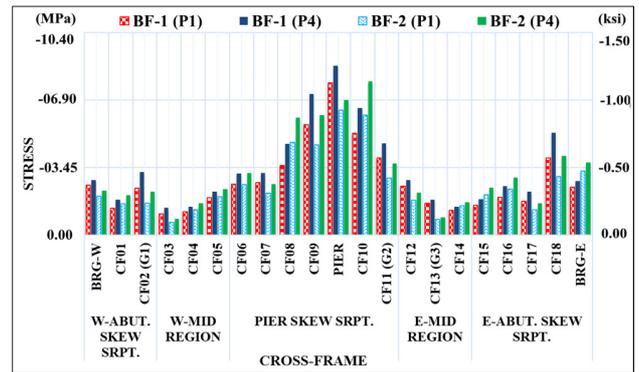


Fig. 14. FEA maximum compressive stresses for BF-1 and BF-2 locations at girder cross-sections adjacent to cross-frames connected to Girder #4 due to truck Pass#1 and Pass#4.

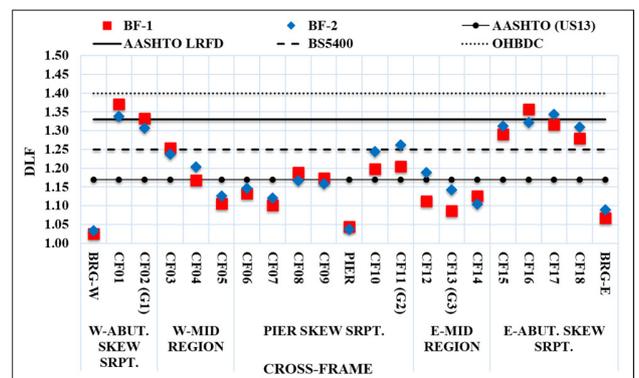


Fig. 15. FEA dynamic load factor variation along the bottom flange of Girder #4 in terms of maximum tensile stresses.

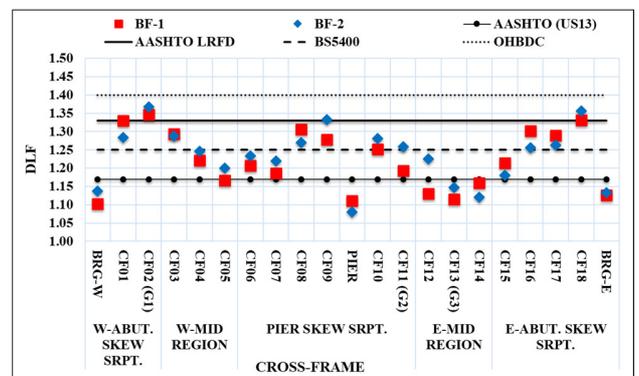


Fig. 16. FEA dynamic load factor variation along the bottom flange of Girder #4 in terms of maximum compressive stresses.

In Figures 13-16, the symbols BRG-W, and BRG-E refer to the west and east bridge expansion bearing supports, and PIER, refer to the pier fixed bearing support, while CF#, refer to the cross-frame adjacent to the girder#4 cross-section (see Figure 3). Figures 15 and 16 show the DLF variation along the bottom flange of Girder #4. The higher DLF values are concentrated near the skew support regions (abutments). This is hypothesized as being due to the skew bridge support regions, which are subjected to significant torsional and lateral moments due to the interaction effect of both the bracing system and the moving loads. It is hypothesized that the dynamic responses of

the bridge exacerbate these effects. The average DLF values for the abutments skew support regions are 1.25 in tension and 1.27 in compression, and for the pier skew support region 1.16 in tension and 1.22 in compression, while for the non-skew middle region they are 1.15 in tension and 1.18 in compression.

VI. CASE STUDY

This case study is implemented in order to investigate the effect of different arrangements of bracing systems on the

bridge dynamic response and thus on DLF values. Figure 17 represents the US13 Bridge but with staggered cross-frames configuration keeping the other components of the bridge the same as in the real bridge. Sections G1, G2, and G3 in addition to cross-frames CF1, CF8, CF15, CF16, and CF17 are selected to compare the DLF results for perpendicular vs staggered cross-frames configuration of the US13 Bridge. These sections (bottom flange sections) are selected since they are matching in locations in perpendicular and staggered bracing systems of US13 Bridge as shown in Figures 5 and 17 respectively.

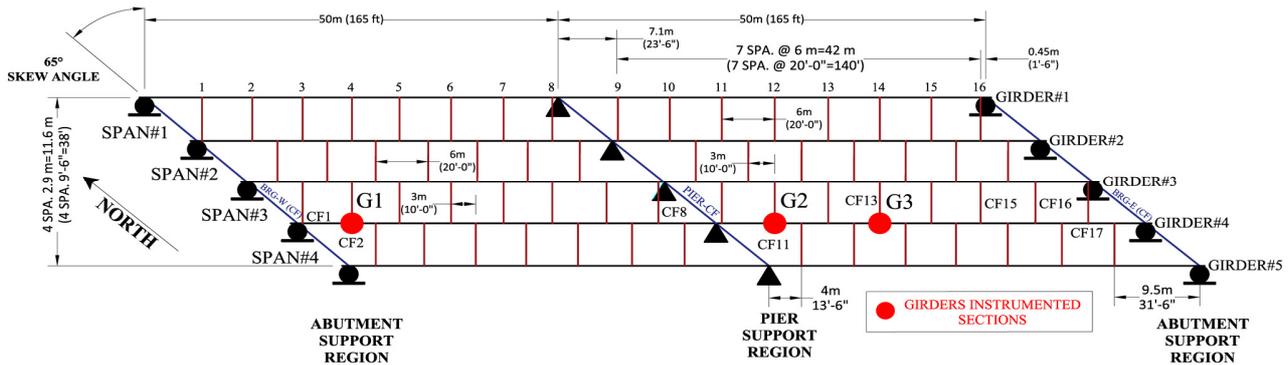


Fig. 17. US13 Bridge with staggered cross-frames configuration.

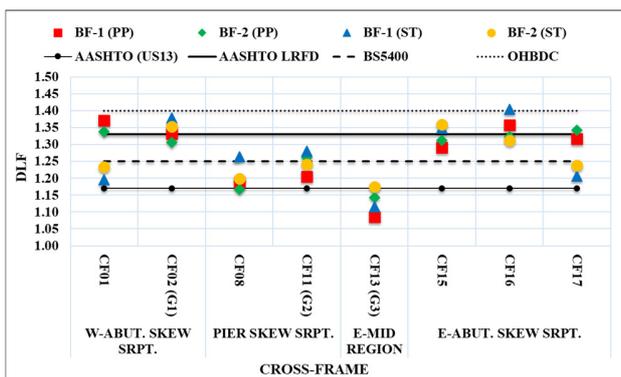


Fig. 18. FEA perpendicular vs staggered DLFs for selected sections along the bottom flange of Girder #4 in terms of maximum tensile stresses.

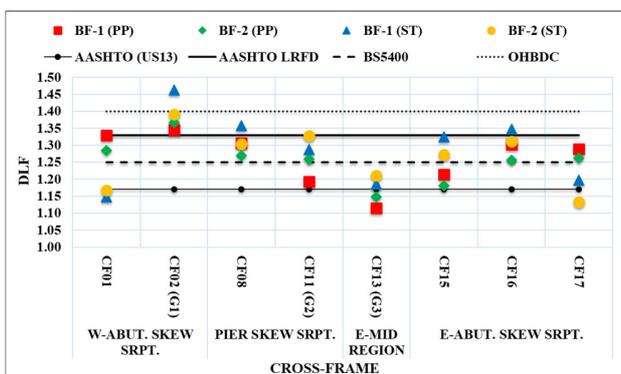


Fig. 19. FEA perpendicular vs staggered DLFs for selected sections along the bottom flange of Girder #4 in terms of maximum compressive stresses.

Figures 18 and 19 show the comparison of the DLF results for both cases. The comparisons indicate a slight difference between the DLFs results of both perpendicular and staggered cross-frames configurations within this range of bridge length and skew angle. For the selected cross-frames, the average DLF values in tensions are 1.25 for perpendicular bracing and 1.26 for staggered bracing, while in compression the average values are 1.26 and 1.28 for perpendicular and staggered bracing respectively.

VII. CONCLUSION

Field testing data of different passes of a weighed load vehicle were used to validate a full-scale 3D FEA model of a highly skewed steel girder bridge, created with the FEA software ABAQUS/CAE. The results include the maximum tension and compression stresses resulting from static and dynamic travel speeds along the length of the most-heavily loaded girder. These are used to calculate DLF and are compared with the code provisions for DLF. The results of this study lead to the following conclusions:

- The DLF is highly variable throughout the length of the girder and the results indicate that the DLF from the compressive stresses of the skewed continuous bridge is larger than those for tensile stresses. The average value for all cross-frames DLF in tension is 1.19 and in compression is 1.23.
- The DLF used in the current practice codes is intended for the design of new bridges and the case of in-service bridges. This shows that most codes of practice are appropriately conservative in the regions of the girder that would govern the flexural design. However, the DLF sometimes exceeds the code-recommended values in the

vicinity of skewed supports. The discrepancy of the DLF, in the west abutment skew support region, exceeds by 13% and 16% the values determined by [2] for tension and compression stresses respectively, while in comparison to [8] the differences reach 6% and 8%, respectively.

According to the review of the widely used codes of practice, it can be concluded that the specifications related to DLF vary significantly, indicating that there is no unanimity for the evaluation of DLF. Also, only a few parameters are adopted in the expression of DLF, such as bridge span length in [2-4], bridge natural frequency [5, 6], and traffic load models in [7, 8]. The outputs of this study were based on the numerical models on a skew continuous girder bridge with specific span length and skew. More diverse numerical investigations or field testing of bridges with a wider range of both span lengths and skewness are suggested to draw a more comprehensive perspective on the general dynamic behavior of the skewed continuous and simple span bridges. Also, more studies are needed in order to evaluate the effect of skewness and bracing system configurations (e.g. perpendicular, parallel, and staggered) on the bridge dynamic responses.

REFERENCES

- [1] H. Moghimi and H. Ronagh, "Impact factors for a composite steel bridge using non-linear dynamic simulation," *International Journal of Impact Engineering*, vol. 35, no. 11, pp. 1228–1243, Nov. 2008, <https://doi.org/10.1016/j.ijimpeng.2007.07.003>.
- [2] *Standard Specifications for Highway Bridges: 2002*, 17th edition. Washington, DC, USA: AASHTO, 2003.
- [3] *Specifications for highway bridges. Part 1: Common Specifications*. Tokyo, Japan: Japan Road Association, 1996.
- [4] *Load and Resistance Factor Design Specifications*, 8th ed. Washington DC, USA: AASHTO, 2017.
- [5] *Ontario Highway Bridge Design Code*, 2nd ed. Toronto, Canada: Ontario Ministry of Transportation and Communications, 1983.
- [6] *Bridge Design Code*. Sydney, Australia: Australia's National Road Authority, 1992.
- [7] *Eurocode 1: Actions on structures - Part 2: Traffic loads on bridges*. Brussels, Belgium: CEN, 2003.
- [8] *BSS400-2: Steel, Concrete and Composite Bridges. Part 2: Specification for Loads*. London, UK: BSI, 2006.
- [9] L. Ma, W. Zhang, W. S. Han, and J. X. Liu, "Determining the dynamic amplification factor of multi-span continuous box girder bridges in highways using vehicle-bridge interaction analyses," *Engineering Structures*, vol. 181, pp. 47–59, Feb. 2019, <https://doi.org/10.1016/j.engstruct.2018.11.059>.
- [10] W. Wang, L. Deng, and X. Shao, "Fatigue Design of Steel Bridges Considering the Effect of Dynamic Vehicle Loading and Overloaded Trucks," *Journal of Bridge Engineering*, vol. 21, no. 9, Sep. 2016, Art. no. 04016048, [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0000914](https://doi.org/10.1061/(ASCE)BE.1943-5592.0000914).
- [11] X. Zhang, K. Sennah, and J. B. Kennedy, "Evaluation of impact factors for composite concrete-steel cellular straight bridges," *Engineering Structures*, vol. 25, no. 3, pp. 313–321, Feb. 2003, [https://doi.org/10.1016/S0141-0296\(02\)00160-8](https://doi.org/10.1016/S0141-0296(02)00160-8).
- [12] Y.-B. Yang, S.-S. Liao, and B.-H. Lin, "Impact Formulas for Vehicles Moving over Simple and Continuous Beams," *Journal of Structural Engineering*, vol. 121, no. 11, pp. 1644–1650, Nov. 1995, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1995\)121:11\(1644\)](https://doi.org/10.1061/(ASCE)0733-9445(1995)121:11(1644)).
- [13] "BDI strain transducer ST-350 specifications sheet." BDI, 2006.
- [14] J. R. McConnell, M. Radovic, and K. Ambrose, "Field Evaluation of Cross-Frame and Girder Live-Load Response in Skewed Steel I-Girder Bridges," *Journal of Bridge Engineering*, vol. 21, no. 3, Mar. 2016, Art. no. 04015062, [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0000846](https://doi.org/10.1061/(ASCE)BE.1943-5592.0000846).
- [15] Y. Almoosi, J. McConnell, and N. Oukaili, "Structural Modeling of Cross-Frame Behavior in Steel Girder Bridges," in *2019 12th International Conference on Developments in eSystems Engineering (DeSE)*, Kazan, Russia, Oct. 2019, pp. 620–625, <https://doi.org/10.1109/DeSE.2019.00117>.
- [16] *ABAQUS 6.14: ABAQUS/CAE user's guide*. Simulia, 2019.
- [17] M. a. J. Hassan and A. F. Izzet, "Experimental and Numerical Comparison of Reinforced Concrete Gable Roof Beams with Openings of Different Configurations," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 5066–5073, Dec. 2019, <https://doi.org/10.48084/etasr.3188>.
- [18] H. M. Hekmet and A. F. Izzet, "Numerical Analysis of Segmental Post Tensioned Concrete Beams Exposed to High Fire Temperature," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4759–4768, Oct. 2019, <https://doi.org/10.48084/etasr.3059>.
- [19] D. Chang and H.-H. Lee, "Impact Factors for Simple-Span Highway Girder Bridges," *Journal of Structural Engineering*, vol. 120, no. 3, pp. 704–715, Mar. 1994, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1994\)120:3\(704\)](https://doi.org/10.1061/(ASCE)0733-9445(1994)120:3(704)).
- [20] Y.-B. Yang, C. W. Lin, and J. D. Yau, "Extracting bridge frequencies from the dynamic response of a passing vehicle," *Journal of Sound and Vibration*, vol. 272, no. 3, pp. 471–493, May 2004, [https://doi.org/10.1016/S0022-460X\(03\)00378-X](https://doi.org/10.1016/S0022-460X(03)00378-X).
- [21] P. K. Chatterjee, T. K. Datta, and C. S. Surana, "Vibration of Suspension Bridges under Vehicular Movement," *Journal of Structural Engineering*, vol. 120, no. 3, pp. 681–703, Mar. 1994, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1994\)120:3\(681\)](https://doi.org/10.1061/(ASCE)0733-9445(1994)120:3(681)).
- [22] K. Chompooming and M. Yener, "The influence of roadway surface irregularities and vehicle deceleration on bridge dynamics using the method of lines," *Journal of Sound and Vibration*, vol. 183, no. 4, pp. 567–589, Jun. 1995, <https://doi.org/10.1006/jsvi.1995.0273>.
- [23] D. Huang, T.-L. Wang, and M. Shahawy, "Impact Analysis of Continuous Multigirder Bridges due to Moving Vehicles," *Journal of Structural Engineering*, vol. 118, no. 12, pp. 3427–3443, Dec. 1992, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1992\)118:12\(3427\)](https://doi.org/10.1061/(ASCE)0733-9445(1992)118:12(3427)).
- [24] E. J. OBrien, D. Cantero, B. Enright, and A. González, "Characteristic Dynamic Increment for extreme traffic loading events on short and medium span highway bridges," *Engineering Structures*, vol. 32, no. 12, pp. 3827–3835, Dec. 2010, <https://doi.org/10.1016/j.engstruct.2010.08.018>.

Recurrent Neural Network-based Path Planning for an Excavator Arm under Varying Environment

Nga Thi-Thuy Vu
School of Electrical Engineering
Hanoi University of Science and Technology
Hanoi, Vietnam
nga.vuthithuy@hust.edu.vn

Nam Phuong Tran
School of Electrical Engineering
Hanoi University of Science and Technology
Hanoi, Vietnam
tranphuongnam3098@gmail.com

Nam Hoai Nguyen
School of Electrical Engineering
Hanoi University of Science and Technology
Hanoi, Vietnam
nam.nguyenhoai@hust.edu.vn

Abstract-This paper proposes an algorithm to generate the reference trajectory based on recurrent neural networks for an excavator arm working in a dynamic environment. Firstly, the dynamic of the plant which includes the tracking controller, the arm, and the pile is appropriated by a recurrent neural network. Next, the recurrent neural network combined with a Model Reference Adaptive Controller (MRAC) is used to calculate the reference trajectory for the system. In this paper, the generated trajectory is changed depending on the variation of the pile to maximize the dug weight. This algorithm is simple but effective because it only needs information about the weight at each duty cycle of the excavator. The efficiency of the overall system is verified through simulations. The results show that the proposed scheme gives a good performance, i.e. the dug weight always remains at the desired value (nominal load) as the pile changes its shape during working time.

Keywords-adaptive controller; excavator arm; neural network; path planning; uncertainties

I. INTRODUCTION

The automatic control of an excavator system is a major issue in the field of excavator research [1, 2]. By unmanned operation, the excavator systems not only keep the workers safe but also increase efficiency. However, in order to finish the task without an operator, the trajectory should be designed carefully. Thus, many researches on creating the trajectory for the general manipulator and the excavator have been published. In [3, 4], the 3D trajectory is built by using the information from sensors, cameras, and scanners. The advantage of these types of feedback signals is that they can track the change of the working environment. However, if the working environment lacks light or is dusty, the reliability of the obtained images can be reduced. In [5], a laser scanner is used to get the shape of the pile. From this information, the model of the pile is built and divided into small layers. The local path is designed according to these layers before creating

the global path. In [6], the integrated physics-based model is presented for a mobile excavator. In this work, the current position of the excavator arm is returned to the control system to predict the trajectory for the next cycle. In [7, 8], a neural network is used to calculate the optimal trajectory for the excavator arm. However, these trajectories only work well in static environments. In [9], excavation trajectories are generated using the velocity and the acceleration of each hydraulic cylinder. The generated trajectories are optimal and stable but the velocity and the acceleration are difficult to measure. In [10], with the purpose of optimizing the efficiency for a semi-automated or fully automated excavator system, the trajectory is classified into 4 categories based on the location and the angle of the bucket. From these 4 trajectory types, the operator or the automatic controller will make a suitable decision as the environment changes. The problem of optimal working time and torque motion of the excavator in consideration with the boundary of the actuator's ability is the role of path planning [11]. In this scheme, the trajectory is planned based on B-spline technique using information about soil parameters and system dynamics.

Neural networks are known as a good way of dealing with path planning problems. In [12-14], neural networks are used to generate collision-free trajectories for robots. In [12], the robot works in a dynamic environment with U-shaped and varying obstacles. The reference trajectory of the robot is generated by using a topologically organized neural network. In this network, the dynamic of each neuron is characterized by a shunting equation. The same neural network topology is used in [13] for a multirobot system with moving obstacles. The trajectory planning for the manipulator robot based on a neural network model of the harmonic function is introduced in [14]. Trajectories are built based on neural networks to optimize the jerk in [15] or the working time in [16]. Neural networks are used in the field of path planning with many different purposes [17-19].

Corresponding author: Nga Thi-Thuy Vu

In this paper, an algorithm based on neural networks is proposed for an excavator arm working in a dynamic environment. The dynamic model of the inner loop which includes the tracking controller, the excavator arm, and the pile is approximated by a neural network. A second, Recurrent Neural Network (RNN) combined with the Model Reference Adaptive Controller (MRAC) algorithm is used to calculate the reference trajectory for the system. By this combination, the calculated trajectory can be adjusted after each duty cycle to adapt with the change of the pile and the dug weight remains around the nominal value despite the reduction of the pile. The effectiveness of the overall system is verified through simulations. The contributions of the proposed algorithm are concluded as:

- It can work in a dynamic environment, something that is restricted in [8].
- The dug weight remains in an acceptable range during the digging process although the material is reducing.

The proposed scheme uses only the feedback signal from the weight sensor to generate the path. This is more reliable than cameras or scanners [3-5] and it is easier to measure than velocity and acceleration [9].

II. PROBLEM DESCRIPTION

In order to maximize the efficiency of the excavator, one of the requirements is that during the digging process, the dug mass at each time must be maintained in a given acceptable range. However, the shape of the pile is changing while the excavator is digging up the material. Therefore, if the trajectory of the excavator remains the same and the dug stack declines over each period of the process, the requirement will not be guaranteed. Hence, the excavator driver has to observe the excavated weight and trajectories in the previous digging to choose an appropriate trajectory next time with an expectation that the mass in the next period will be acceptable. A trajectory generator based on this structure is proposed to replace the excavator driver, which will completely automate the digging process. Before going to solve the described problem, the following assumptions are made:

- The pile lies above the ground and has a triangular shape as shown in Figure 1.
- After a digging cycle, the material on the top of the pile will lie down and fill in the space that was taken, so the pile will maintain the triangular shape with a different slope.
- The dug weight in each period is limited by the volume of the bucket.
- The trajectory in each period is represented by a set of parameters.

The task of the trajectory generator is to observe the weight and the trajectory's parameters in the previous period in order to adjust suitably the trajectory parameters in the next one. In order to execute the algorithm, the excavator's trajectory has a parabolic shape, however, in fact, the excavator's trajectory often has a more complex shape due to many practical conditions. The parabolic shape is described by:

$$y - K = \frac{(x - h)^2}{4p} \quad (1)$$

where K , h and p are scalars. In this work, we have just change the parameter h , which corresponds to the coordinate of the vertex on the horizontal axis. Therefore, each trajectory will be represented by a single value h .

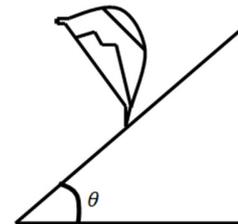


Fig. 1. The shape of the hypothetical pile model.

The block diagram of the overall system is illustrated in Figure 2. It can be seen that the desired trajectory which is generated by the module is denoted as q^* . The desired trajectory is a reference for the excavator controller, and the output of the excavator is q , which affects the dug weight. If the tracking controller has a good performance, the equivalent model of the controller and the excavator can be viewed as a dynamic model which has a static gain of one. In this paper, the equivalent model is considered as a first order system as follows:

$$G(s) = \frac{1}{5s + 1} \quad (2)$$

This means that the trajectory error reduces exponentially over time.

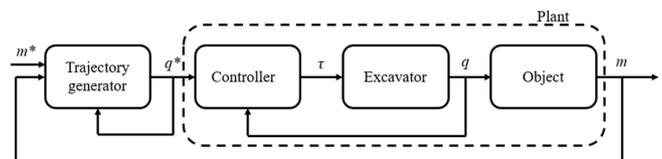


Fig. 2. Block diagram of the overall system.

III. ALGORITHM FOR GENERATING THE REFERENCE TRAJECTORY

In fact, one does not have complete information about the dynamics of the pile. Therefore, a decision on the appropriate trajectory can only be done in practice by observing the pile's response in accordance with the trajectories that the excavator made. Hence, the pile can be considered as a black box, and designing a controller for a black box leads to the idea of the neural network-based MRAC. The control structure of the neural-network based MRAC can be seen in Figure 3. In this Figure, the trajectory generator block can be seen as a controller of the outer loop. This controller is responsible for keeping the dug weight within the specific range as the shape of the pile changes.

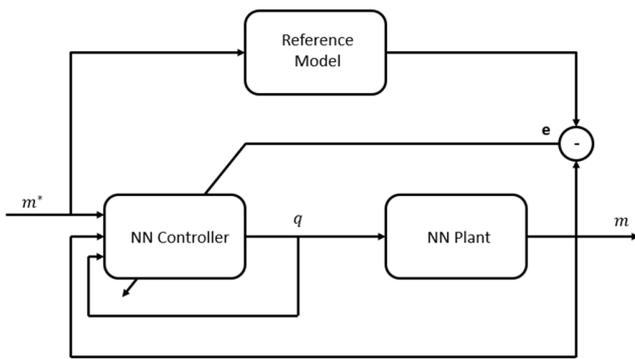


Fig. 3. Control structure of neural network-based MRAC.

This is an offline process, the first step of the design is identifying the plant with the usage of neural network, and the next step is creating a neural network-based controller according to the identified plant in the previous step. Note that, the plant now contains the pile, the excavator, and the tracking controller. The input of the plant is the desired trajectory q^* , whereas the practical trajectory created by the excavator is q because of the control error which should affect the dug weight. Then, the robot excavates the pile with the trajectory q and obtains a weight m . In the identification problem, because the user concerns only about the dug weight after each time the excavator finishes, the object can be viewed as a discrete-time system. A discrete-time system can be described as:

$$m_k = f(q_k^*, \dots, q_{k-n_u}^*, m_{k-1}, \dots, m_{k-n_y}) \quad (3)$$

Discrete-time system identification can be carried out by using a nonlinear autoregressive network with exogenous inputs denoted as NARX. The NARX's architecture is illustrated in Figure 4, where the input and feedback output of the network pass through the Tapped-Delay-Lines (TDLs) which make up the dynamics of the network.

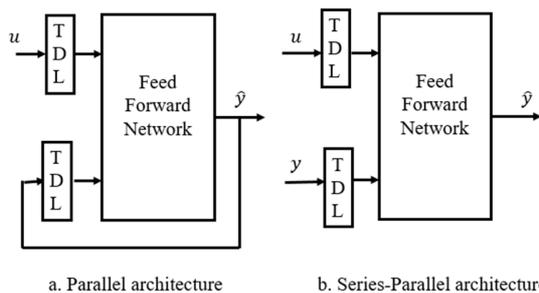


Fig. 4. Parallel and series-parallel architectures of the NARX network.

A feature that makes the NARX become regularly usable in identification problems is that it has only a feedback from the network's output. Therefore, in the training process, one can remove this feedback and consider the feedback output as a second input to the network. The opened architecture, referred as a series-parallel architecture (Figure 4), is used in one-step-ahead prediction. The training process of the opened network is much simpler than that of the closed network, because one can utilize the traditional back-propagation algorithm for the

opened network training. After the opened-network training process is completed, the network can predict one-step ahead, however the objective is to identify the system, which means that the network can make a multi-step-ahead prediction. Therefore, the performance of the closed network is often not good enough. Hence, one should continue training the closed network based on the opened network. However, training a closed network can still confront the problem of gradient vanishing, which is a well-known problem in the training of RNNs. Thus, the original training set is divided into subsets which have a smaller size. The network is trained with these subsets. When the training process with these subsets is over, the size of each subset is increased by a small number, and the training restarts with the new bigger subset. These steps are repeated until the size of the subset becomes equal with the size of the original training set. The algorithm is shown in Figure 5, which begins with the size of each subset equal to the maximum number of delays (MD) in the TDL of the network plus one that is equivalent to the training opened network. In Figure 5, MD is the Maximum number of Delays of both input's TDL and feedback output's TDL. Q, c, n are the length of the original batch, the length of the mini batch, and the number of subsets' respectively, \underline{x} is the network's parameters, e_{ik} represents the error which corresponds with the k^{th} element of the i^{th} subset, $l_i(\underline{x})$ is the squared error which corresponds to the i^{th} subset, $L(\underline{x})$ is the loss function which is the Mean Square Error of the original batch.

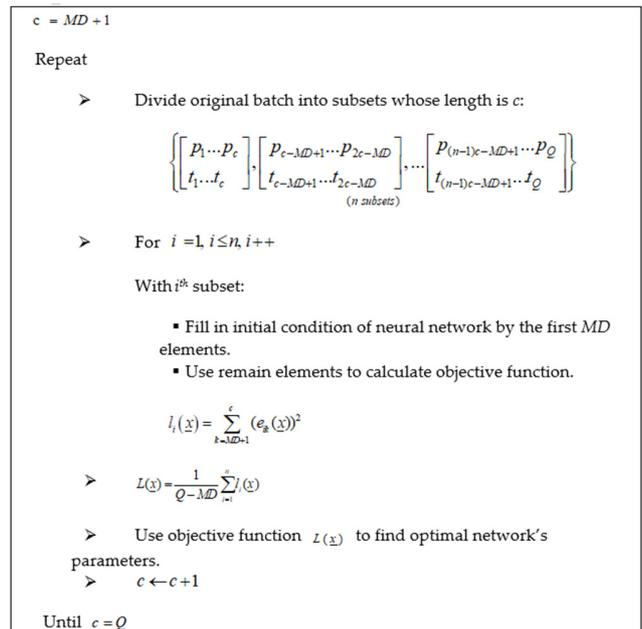


Fig. 5. Training algorithm of the closed network.

Note that each adjacent subset has some similar elements, with size equivalent to MD in the TDL. These elements are used as the initial conditions for the network. To guarantee that the objective function includes all elements in the original training batch, a subset is created to overlap the adjacent one. Once the closed network has finished training, we will create the controller network based on the object network. To train the

controller network, the whole structure that contains the plant network and the controller network must be considered as a major network. Then, this network is used to approximate the given reference model. However, during the training process, only the controller network's parameters are adjusted in order to optimize the objective function, while the object network's parameters are kept unchanged. The well-known back-propagation algorithm is used to train the major network, the error between reference model's output and this network's output back-propagates through layers of the network, and the controller network's parameters are updated.

IV. NETWORK TRAINING AND SIMULATION

A. Plant Network

Neural networks are known as a powerful technique to model a system with a mathematical model unknown or difficult to build [20, 21]. In this section, a neural network is used to model the plant which consists of the pile, the excavator, and the tracking controller. The first step of training the plant network is collecting data from the object. The performance of the network depends significantly on the way the data cover and describe the operating range of the object. In system identification problems, the data are often obtained by generating the input signal in the form of a sequence of step functions that have random durations and amplitudes. In this case, the reference input signal is produced with a random amplitude within the intervals [0.3, 0.5] and [0.5, 0.8] in the first 100 and the next 100 steps respectively. The duration of each step function is $d \in [1, 10]$. The main reason is that the collected data need to cover the operation range of the object. In the first 100 steps, because the size of the pile is still large, one should use a trajectory with small h . In contrast, in the next 100 steps, because a large amount of the pile is already taken, one needs to shift the range of the trajectory deep inside the pile with the expectation that the excavated weight is still acceptable. This signal is applied to the plant, then the plant's output is collected. The obtained data are shown in Figure 6.

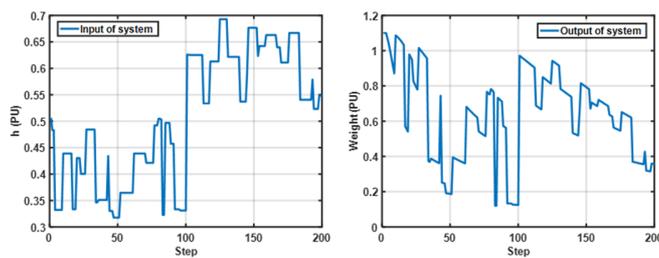


Fig. 6. A mini batch of the training data for the plant network.

Note that there are 40 mini batches whose length is 200 samples, however Figure 6 only shows an instance of the mini batches. Before training, the input data are normalized in the range [-1, 1]. Note that h is the representation of the desired trajectory, the control error in the controlling robot affects the practical trajectory, so the dug weight could not be the same as expected. In this case, the collected data include this control error. When the training data are available, the plant network can be designed. This network is illustrated in Figure 7. The

input of the network passes through a 0:1 TDL. The TDL of the input has a current element (zero delay) due to the pile feature. When one adjusts the excavator's trajectory, the dug weight immediately changes depending on the shift of the trajectory, therefore the current component is added to demonstrate this feature. The TDL of network's feedback output is 1:5. The Bayesian regularization algorithm is used for training. The training result is shown in Figure 8. Then, the network is tested with the testing data. The test result is illustrated in Figure 9.

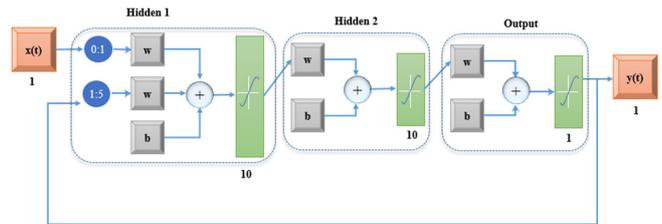


Fig. 7. Plant network's architecture.

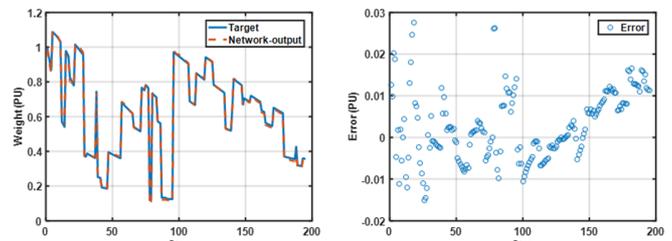


Fig. 8. Training result of the plant network.

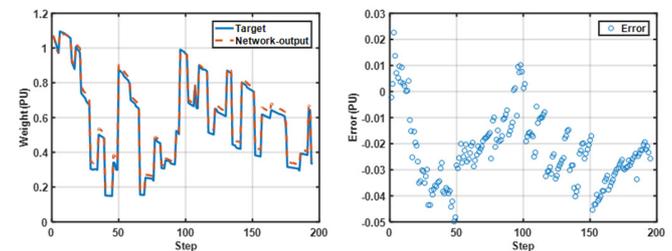


Fig. 9. Testing result of the plant network.

B. Controller Network

First of all, one must choose an appropriate reference model. Note that the plant network is of the 5th order, thus, a fifth-order dead-beat system is utilized as a reference model. The transfer function of the chosen reference model is:

$$Gm(z) = \frac{5z^{-1} + 4z^{-2} + 3z^{-3} + 2z^{-4} + z^{-5}}{15} \quad (4)$$

The next step is collecting data from the reference model for the controller training set. The training data for the controller is demonstrated in Figure 10. The left side in Figure 10 shows the input data of the training set. The right side shows the response of the reference model. Note that the initial value of the reference model's output is the limited dug weight, this means that the dug weight can reach the maximum value in the first period. It can be seen that although the controller only

operates with unit set-point, the input has different values that range from 0.4 to 1. The reason is that the controller has to work with different states of the object, so this type of input helps training the controller with practical states of the object, e.g. the controller can be trained to give a suitable trajectory that will help the dug weight increase from 0.5, which is its current value, to the desired value of 1. The training data have 8 mini batches, each mini batch having a length of 90 samples. Figure 10 illustrates an example of the total training set. The major neural network's architecture, which includes the plant network and the controller network, is shown in Figure 11. The transfer function of the output layer of the controller is the *tansig* function, which means that the output of the controller is within the range from -1 to 1 which is equal to the normalized input range of the plant network.

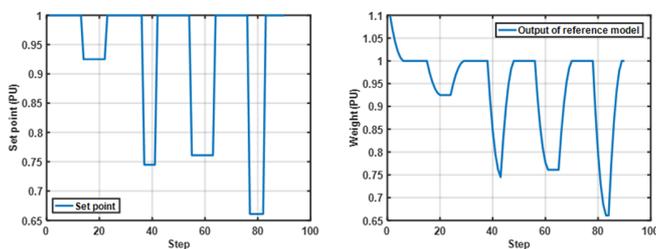


Fig. 10. Training data of the controller network.

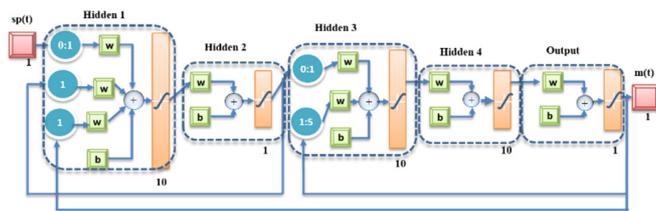


Fig. 11. Major network's architecture.

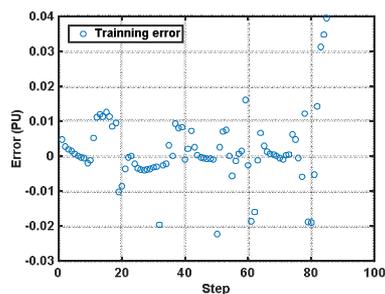


Fig. 12. Training result of the major network.

The used training algorithm is the Bayesian regularization. Figure 12 describes the training result of the major network. It can be seen that the training error is very small, therefore the controller can be tested on the practical object.

C. Stability Analysis

The system contains a plant neural network and a controller neural network. It is a five-layer discrete-time RNN. To analyze the stability of the RNN, there are several stability criteria [22-25]. In this case, the stability criteria from [22]

were applied. To do so, the recurrent neural network will be put in the standard form as:

$$x(k+1) = f(W^1x(k) + W^2x(k+1) + b) \quad (5)$$

where $x(k)$ is defined as a vector of layers' output in the past (such as $a^1(k-1), a^2(k-1), a^2(k-2), a^2(k-3) \dots$, where a^i is the output of layer $i, i=1,2,\dots,5$), W^1 and W^2 are the weight matrices of layer 1 and 2. The size of the state vector is 41 in this case. We did obtain the matrices W^1 and W^2, b and f , but the upper bound and lower bound matrices for the function vector $f(41 \times 1)$ must be found as the requirement of the stability criteria in [20]. This leads to the problem that the equilibrium point of the system in (5) must be determined. It is not easy to find the equilibrium point in this case because its size is 41 and the function vector includes nonlinear functions such as the *tansig*. Thus, in this work the stability of the system is verified through simulations.

D. Simulation Results

In order to verify the effectiveness of the proposed scheme, simulations are done in Matlab/Simulink. In the simulations, the initial path of the arm is set so that the dug mass is 0.4. The trajectory generator should calculate the reference trajectory for the plant to meet the requirement. The dug weight desired value is tracked during working time despite the change of the pile. Figure 13(a) shows the change of the pile during working time. Initially, the slope of the pile is about 62 degrees. Because of the digging, this slope will decrease gradually and reach the value of 46 degrees at the 42th step. Figure 13(b) shows the values of h for each step. The initial value of h is set at -0.9 and after 50 steps its value is 0.8. Corresponding with these 50 values of h , 50 trajectories will be generated to adapt to the change of the pile.

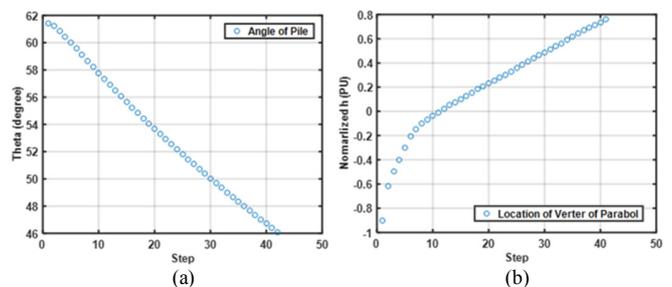


Fig. 13. The change of system during working time. (a) θ , (b) h .

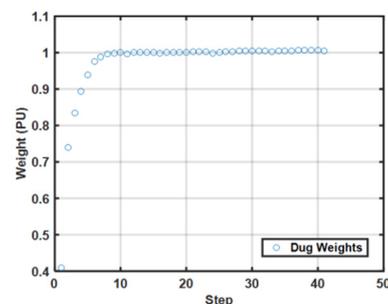


Fig. 14. The dug weight response during working time.

Figure 14 illustrates the dug weight response of the system during working time. In the first step, the excavator works with the given trajectory and the dug weight is 0.4. The information about the trajectory and the weight will be given to the controller to calculate the reference trajectory for the next step. After 5 steps, the dug weight accomplishes the desired value and remains at this state during the working time despite the change of the pile.

V. CONCLUSION

A simple but efficient algorithm has been proposed to design the reference trajectory for an excavator arm in a dynamic environment. The model of the tracking controller, the arm, and the pile is approximated, at first by an RNN. A second RNN combined with the MRAC algorithm is used to calculate the reference trajectory of the system. By this combination, the calculated trajectory can be adjusted after each duty cycle to adapt with the change of the pile so that the dug weight remains around the nominal value despite the reduction of the pile. The effectiveness of the overall system was verified through simulations. The results show that the proposed scheme gives a good performance, i.e. the dug weight always tracks the nominal value as the pile changes its shape during working time.

REFERENCES

- [1] H. Feng *et al.*, "Robotic excavator trajectory control using an improved GA based PID controller," *Mechanical Systems and Signal Processing*, vol. 105, pp. 153–168, May 2018, <https://doi.org/10.1016/j.ymsp.2017.12.014>.
- [2] R. Ding, B. Xu, J. Zhang, and M. Cheng, "Self-tuning pressure-feedback control by pole placement for vibration reduction of excavator with independent metering fluid power system," *Mechanical Systems and Signal Processing*, vol. 92, pp. 86–106, Aug. 2017, <https://doi.org/10.1016/j.ymsp.2017.01.012>.
- [3] H. Shao, H. Yamamoto, Y. Sakaida, T. Yamaguchi, Y. Yanagisawa, and A. Nozue, "Automatic Excavation Planning of Hydraulic Excavator," in *Intelligent Robotics and Applications*, Berlin, Heidelberg, 2008, pp. 1201–1211, https://doi.org/10.1007/978-3-540-88518-4_128.
- [4] A. Stentz, J. Bares, S. Singh, and P. Rowe, "A robotic excavator for autonomous truck loading," in *Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovations in Theory, Practice and Applications*, Victoria, BC, Canada, Oct. 1998, vol. 3, pp. 1885–1893, <https://doi.org/10.1109/IROS.1998.724871>.
- [5] J. Seo, S. Lee, J. Kim, and S.-K. Kim, "Task planner design for an automated excavation system," *Automation in Construction*, vol. 20, no. 7, pp. 954–966, Nov. 2011, <https://doi.org/10.1016/j.autcon.2011.03.013>.
- [6] Y. H. Zweiri, L. D. Seneviratne, and K. Althoefer, "Model-based automation for heavy duty mobile excavator," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Lausanne, Switzerland, Sep. 2002, vol. 3, pp. 2967–2972, <https://doi.org/10.1109/IRDS.2002.1041723>.
- [7] S. Lee, D. Hong, H. Park, and J. Bae, "Optimal path generation for excavator with neural networks based soil models," in *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Seoul, Korea, Aug. 2008, pp. 632–637, <https://doi.org/10.1109/MFI.2008.4648015>.
- [8] N. T.-T. Vu, N. P. Tran, and N. H. Nguyen, "Adaptive Neuro-Fuzzy Inference System Based Path Planning for Excavator Arm," *Journal of Robotics*, vol. 2018, Dec. 2018, Art. no. e2571243, <https://doi.org/10.1155/2018/2571243>.
- [9] Z. Li, X. Li, S. Liu, and L. Jin, "A study on trajectory planning of hydraulic robotic excavator based on movement stability," in *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, Xi'an, China, Aug. 2016, pp. 582–586, <https://doi.org/10.1109/URAI.2016.7625784>.
- [10] R. Tiwari, J. Knowles, and G. Danko, "Bucket trajectory classification of mining excavators," *Automation in Construction*, vol. 31, pp. 128–139, May 2013, <https://doi.org/10.1016/j.autcon.2012.11.006>.
- [11] Y. B. Kim, J. Ha, H. Kang, P. Y. Kim, J. Park, and F. C. Park, "Dynamically optimal trajectories for earthmoving excavators," *Automation in Construction*, vol. 35, pp. 568–578, Nov. 2013, <https://doi.org/10.1016/j.autcon.2013.01.007>.
- [12] S. X. Yang and M. Meng, "Neural network approaches to dynamic collision-free trajectory generation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 31, no. 3, pp. 302–318, Jun. 2001, <https://doi.org/10.1109/3477.931512>.
- [13] Howard Li, S. X. Yang, and Y. Biletskiy, "Neural network based path planning for a multi-robot system with moving obstacles," in *2008 IEEE International Conference on Automation Science and Engineering*, Arlington, VA, USA, Aug. 2008, pp. 163–168, <https://doi.org/10.1109/COASE.2008.4626446>.
- [14] A. Pashkevich, M. Kazheunikau, and A. E. Ruano, "Neural network approach to collision free path-planning for robotic manipulators," *International Journal of Systems Science*, vol. 37, no. 8, pp. 555–564, Jun. 2006, <https://doi.org/10.1080/00207720600783884>.
- [15] D. Simon, "The application of neural networks to optimal robot trajectory planning," *Robotics and Autonomous Systems*, vol. 11, no. 1, pp. 23–34, May 1993, [https://doi.org/10.1016/0921-8890\(93\)90005-W](https://doi.org/10.1016/0921-8890(93)90005-W).
- [16] J. Giri, P. Giri, and R. Chadge, "Neural Network Based Modelling of Time Optimal Interpolation of Non-linear Trajectory," *Materials Today: Proceedings*, vol. 5, no. 2, Part 2, pp. 7981–7990, Jan. 2018, <https://doi.org/10.1016/j.matpr.2017.11.482>.
- [17] G. Atmeh and K. Subbarao, "A Dynamic Neural Network with Feedback for Trajectory Generation," *IFAC-PapersOnLine*, vol. 49, no. 1, pp. 367–372, Jan. 2016, <https://doi.org/10.1016/j.ifacol.2016.03.081>.
- [18] Y. Li, R. Cui, Z. Li, and D. Xu, "Neural Network Approximation Based Near-Optimal Motion Planning With Kinodynamic Constraints Using RRT," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 11, pp. 8718–8729, Nov. 2018, <https://doi.org/10.1109/TIE.2018.2816000>.
- [19] P. Zhang, C. Xiong, W. Li, X. Du, and C. Zhao, "Path planning for mobile robot based on modified rapidly exploring random tree method and neural network," *International Journal of Advanced Robotic Systems*, vol. 15, no. 3, May 2018, Art. no. 1729881418784221, <https://doi.org/10.1177/1729881418784221>.
- [20] A. S. Kote and D. V. Wadkar, "Modeling of Chlorine and Coagulant Dose in a Water Treatment Plant by Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4176–4181, Jun. 2019, <https://doi.org/10.48084/etasr.2725>.
- [21] L. B. Salah and F. Fourati, "Systems Modeling Using Deep Elman Neural Network," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3881–3886, Apr. 2019, <https://doi.org/10.48084/etasr.2455>.
- [22] N. H. Nguyen and M. Hagan, "Stability analysis of layered digital dynamic networks using dissipativity theory," in *The 2011 International Joint Conference on Neural Networks*, San Jose, CA, USA, Jul. 2011, pp. 1692–1699, <https://doi.org/10.1109/IJCNN.2011.6033428>.
- [23] N. E. Barabanov and D. V. Prokhorov, "Stability analysis of discrete-time recurrent neural networks," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 292–303, Mar. 2002, <https://doi.org/10.1109/72.991416>.
- [24] N. E. Barabanov and D. V. Prokhorov, "A new method for stability analysis of nonlinear discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 2250–2255, Dec. 2003, <https://doi.org/10.1109/TAC.2003.820158>.
- [25] M. Liu, "Delayed Standard Neural Network Models for Control Systems," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1376–1391, Sep. 2007, <https://doi.org/10.1109/TNN.2007.894084>.

Modified Numerical Modeling of Axially Loaded Concrete-Filled Steel Circular-Tube Columns

Phu-Cuong Nguyen

Advanced Structural Engineering Laboratory
Faculty of Civil Engineering
Ho Chi Minh City Open University
Ho Chi Minh City, Vietnam
cuong.pn@ou.edu.vn, henycuong@gmail.com

Duc-Duy Pham

Advanced Structural Engineering Laboratory
Faculty of Civil Engineering
Ho Chi Minh City Open University
Ho Chi Minh City, Vietnam
duy.pd@ou.edu.vn

Thanh-Tuan Tran

Institute of Offshore Wind Energy
Kunsan National University
Kunsan City, Cheollabuk-Do, Republic of Korea
tranthanhtuan@kunsan.ac.kr

Trong Nghia-Nguyen

Faculty of Civil Engineering
Ho Chi Minh City Open University
Ho Chi Minh City, Vietnam
nghia.nt@ou.edu.vn

Abstract—Predicting the behavior of concrete in a Concrete-Filled Steel Tubular (CFST) column is challenging due to the sensitivity to input parameters such as the size of the cross-section, the material modeling, and the boundary conditions. The present paper proposes a new modified finite element model to predict the behavior and strength of a CFST subjected to axial compression. The development is based on the concrete damaged plasticity model, with its stress-strain relationship revised from the available model. The predicted accuracy of the modified model is verified via a wide range of experimental tests. The proposed model has more accuracy than the available models in predicting the ultimate compression strength. The results show good agreement with the test data, allowing its use in modeling CFST columns.

Keywords—CFST columns; axial compression; finite element modeling; stress-strain relationship; steel; concrete

I. INTRODUCTION

Concrete-Filled Steel Tubular (CFST) columns are widely used in modern constructions. Many studies have been conducted in order to understand the behavior of these composite columns, either experimental [1, 2] or analytical [3, 4]. Based on the outcomes from these investigations, different guidelines and codes have been published for the design procedure [5, 6]. CFST columns combined with steel-framed structures [7, 8] form very powerful structures. Authors in [9] used fiber reinforced polymer tubes filled with recycled materials and concrete for pile foundations. Using Finite Element (FE) models has become a popular technique to model CFST via software such as ANSYS and ABAQUS. This approach can model the interaction between concrete core and steel tube accurately. In this study, the FE software ABAQUS is used to FE model CFST. Recently, authors in [10] studied the influence of geometric and material properties on the behavior of the Concrete-Filled Double Skin Steel Tube

(CFDST) member with M16 studs under bending. The material model is an input parameter that is mainly affected by the behavior of CFST. Several concrete and steel stress-strain $\sigma - \varepsilon$ models have been developed [11-16]. Almost all models are developed from experimental data based on regression analysis. Authors in [17, 18] proposed FE modeling for analyzing box and circular CFST columns. Authors in [16] developed a simplified model to calculate the capacity of the cross-section and the axial force via the load-deformation $N - \varepsilon$ curve based on 50 test data of Self-Consolidating Concrete (SCC) filled Hollow Structural Steel (HSS) stub columns. Similarly, authors in [14, 15] proposed a simple and effective model for CFST subjected to extreme loadings. In [11], the steel $\sigma - \varepsilon$ relationship was revised based on the available models by [19]. In the present study, the concrete $\sigma - \varepsilon$ curve developed in [12] is revised via the modification of the plateau branch to represent the actual behavior of the material.

This paper aims to develop an FE model of circular CFST columns under axial compression loading. A modified $\sigma - \varepsilon$ model is proposed for the concrete material. The purpose of this study is to improve the accuracy of the FE model using ABAQUS. The accuracy and effectiveness of the new model are verified against the test data from [3, 20]. In the future, focus may be given to the behavior of CFST with steel fiber reinforcement [21], concrete using fly ash [22], or concrete with cement paste including silica oxide nanoparticles.

II. DEVELOPMENT OF THE MATERIAL MODELS

A. Steel Material Model

In this study, the material model proposed in [19] is adopted for the characteristics of the stress-strain curves, as illustrated in Figure 1. The relationship of $\sigma - \varepsilon$ can be expressed as follows:

Corresponding author: Phu-Cuong Nguyen

$$\sigma = \begin{cases} E_s \varepsilon & 0 \leq \varepsilon \leq \varepsilon_y \\ f_y & \varepsilon_y \leq \varepsilon \leq \varepsilon_p \\ f_u - (f_u - f_y) \times \left(\frac{\varepsilon_u - \varepsilon}{\varepsilon_u - \varepsilon_p}\right)^p & \varepsilon_p \leq \varepsilon \leq \varepsilon_u \\ f_u & \varepsilon \geq \varepsilon_u \end{cases} \quad (1)$$

where E_s is Young's modulus of steel, f_y is the yield strength corresponding to the yield strain $\varepsilon_y = \frac{f_y}{E_s}$, f_u is the ultimate strength corresponding to the ultimate strain ε_u , p is the hardening exponent that can be calculated as:

$$p = E_p \times \left(\frac{\varepsilon_u - \varepsilon_p}{f_u - f_y}\right) \quad (2)$$

where E_p is the initial elastic modulus at the onset of strain hardening, and can be taken as $0.02E_s$. Yield strain ε_p and ultimate strain ε_u can be expressed based on the range of yield strength f_y which is 200MPa-800MPa:

$$\varepsilon_p = \begin{cases} 15 \times \varepsilon_y & f_y \leq 300 \\ [15 - 0.018(f_y - 300)] \times \varepsilon_y & f_y \geq 300 \end{cases} \quad (3)$$

$$\varepsilon_u = \begin{cases} 100 \times \varepsilon_y & f_y \leq 300 \\ [100 - 0.15(f_y - 300)] \times \varepsilon_y & f_y \geq 300 \end{cases} \quad (4)$$

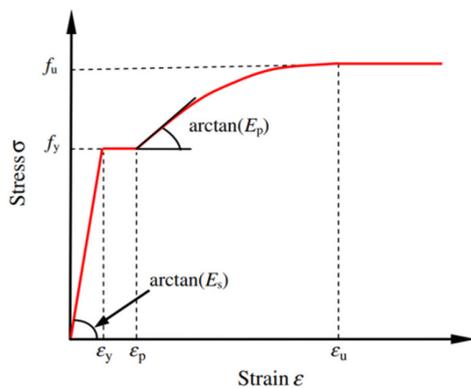


Fig. 1. Constitutive model of steel.

B. Concrete Material Model

Characteristics of stress-strain curves for concrete have been studied in [1, 24]. Most researchers focused on the development of a $\sigma - \varepsilon$ model based on the experimental results. For example, in [25], $\sigma - \varepsilon$ curves for the fiber model of concrete have been developed to predict the monotonic and cyclic force-deformation behavior of CFST columns. Similar studies have been performed in [2, 26]. In these works, empirical models were proposed for the strain-softening response for CFST. In this paper, the damaged plasticity model of concrete material is utilized. This model is available in ABAQUS and can be easily used by defining some key material parameters. These factors are calibrated and verified with the extensive test results of [27-29] and they are briefly introduced as follows:

Young's modulus of concrete is predicted based on the concrete strength (f'_c) [30]:

$$E_c = 4700 \sqrt{f'_c} \quad (5)$$

According to [31], the ratio of compressive strength under biaxial loading to the uniaxial compressive strength (f_{b0}/f'_c) is obtained based on the available data in the literature as follows:

$$\frac{f_{b0}}{f'_c} = 1.5 \times (f'_c)^{-0.075} \quad (6)$$

The fracture energy (G_F) that is defined as the tensile softening response beyond the failure stress, depends on the maximum coarse aggregate size (d_{max}) and concrete strength (f'_c). The following equation can be deduced to determine G_F [32, 33]:

$$G_F = (0.0469 \times d_{max}^2 - 0.5 \times d_{max} + 26) \left(\frac{f'_c}{10}\right)^{0.7} \quad (7)$$

Authors in [29] stated that the compressive meridian (K_c) that is used to determine the yield surface of the material model, varied in the range from 0.5 to 1, and it is expressed as a function of equibiaxial concrete strength (f_{b0}) and concrete strength (f'_c) [34]:

$$K_c = \frac{5.5 \times f_{b0}}{3 \times f'_c + 5 \times f_{b0}} \quad (8)$$

The dilation angle (ψ) factor is used to define the plastic flow potential. This parameter is affected by the confining stress and plastic deformation of the material [27]. The factor ψ is determined as the function of the confinement factor (ξ_c) based on regression analysis [1]:

$$\psi = \begin{cases} 56.3 \times (1 - \xi_c) & \xi_c \leq 0.5 \\ 6.672e^{\left(\frac{7.4}{4.64 + \xi_c}\right)} & \xi_c > 0.5 \end{cases} \quad (9)$$

in which the confinement factor is calculated as: $\xi_c = A_s f'_y / A_c f'_c$, where A_s and A_c are the cross-sections of the steel tube and concrete respectively.

For the strain hardening-softening rule, authors in [12] proposed a $\sigma - \varepsilon$ curve for confined concrete through the tests. This model is revised in the present paper. The details of the proposed model are described below.

C. The Proposed Stress-Strain Relationship for Concrete

The $\sigma - \varepsilon$ model developed in [12] is modified to describe the effective $\sigma - \varepsilon$ relationship of confined concrete by the steel tube, as illustrated in Figure 2. The model includes three main stages, corresponding to ascending branch (OA), plateau branch (AB), and descending branch (BC). The initial stage OA is described as follows:

$$\frac{\sigma}{f'_c} = \frac{aX + bX^2}{1 + (a-2)X + (b+1)X^2} \quad \text{if } 0 < \varepsilon \leq \varepsilon_{c0} \quad (10)$$

where $X = \frac{\varepsilon}{\varepsilon_{c0}}$, a and b are factors that are defined in (11):

$$a = \frac{E_c \varepsilon_{c0}}{f'_c}, b = \frac{(a-1)^2}{0.55} - 1 \quad (11)$$

The parameter ϵ_{c0} of the unconfined concrete is expressed based on the regression analysis of over 100 test data from 17 references [35], in which the uniaxial compressive strength f'_c ranges from 10 to 100MPa:

$$\epsilon_{c0} = 0.00076 + \sqrt{(0.626f'_c - 4.33) \times 10^{-7}} \quad (12)$$

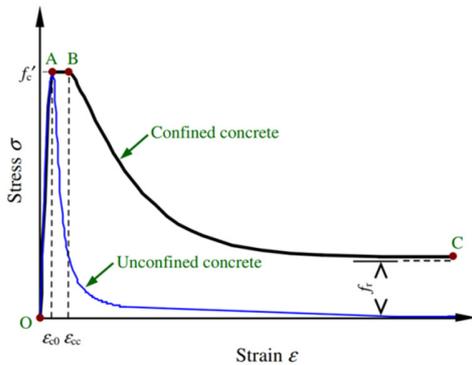


Fig. 2. The constitutive model of confined concrete.

The ultimate strain ϵ_{cc} of the confined concrete is expressed using the model of [12]. In this model, ϵ_{cc} depends on the compressive strength f'_c , and confining stress at the point B, f_B :

$$\frac{\epsilon_{cc}}{\epsilon_{c0}} = e^k$$

$$k = (2.9224 - 0.00367 \times f'_c) \left(\frac{f_B}{f'_c}\right)^{0.3124 + 0.002 \times f'_c} \quad (13)$$

It is worth mentioning that the confining stress only happens before and after the yielding location of the material. In this study, the confining stress of the circular column is determined based on the regression analysis as follows:

$$f_B = \frac{1 + 0.0324 \times \frac{f_y}{f'_c}}{e^{0.0212 \times \frac{D}{t}}} \quad (14)$$

The constant value of confining stress f_B remains from ϵ_{c0} to ϵ'_{cc} corresponding to the plateau branch (AB) that represents the increment of the peak strain of the concrete due to confinement. For the last stage of the concrete model (BC), the model of [36] is used, which is expressed as follows:

$$\sigma = f_{re} + (f'_c - f_{re}) \times \exp\left[-\left(\frac{\epsilon - \epsilon_{cc}}{\alpha}\right)^\beta\right] \quad \text{if } \epsilon > \epsilon_{cc} \quad (15)$$

where f_{re} is the residual stress [1] and α and β are factors determining the shape of the softening stage. The parameter β can be taken as 1.2 [1], and α is calculated as follows:

$$\alpha = 0.04 - \frac{0.036}{1 + e^{6.08\xi_c - 3.49}} \quad (16)$$

III. NUMERICAL MODELING

A. Finite Element Modeling

The current paper used the available model developed in [37]. The diameter and length of the CFST are 300mm and

900mm respectively. The thickness of the steel tube is 3.2mm. Concrete with 27.2MPa compressive strength and 0.351 of ultimate strain is assigned for the concrete core, while steel with yield strength of 232MPa is assigned for the steel tube.

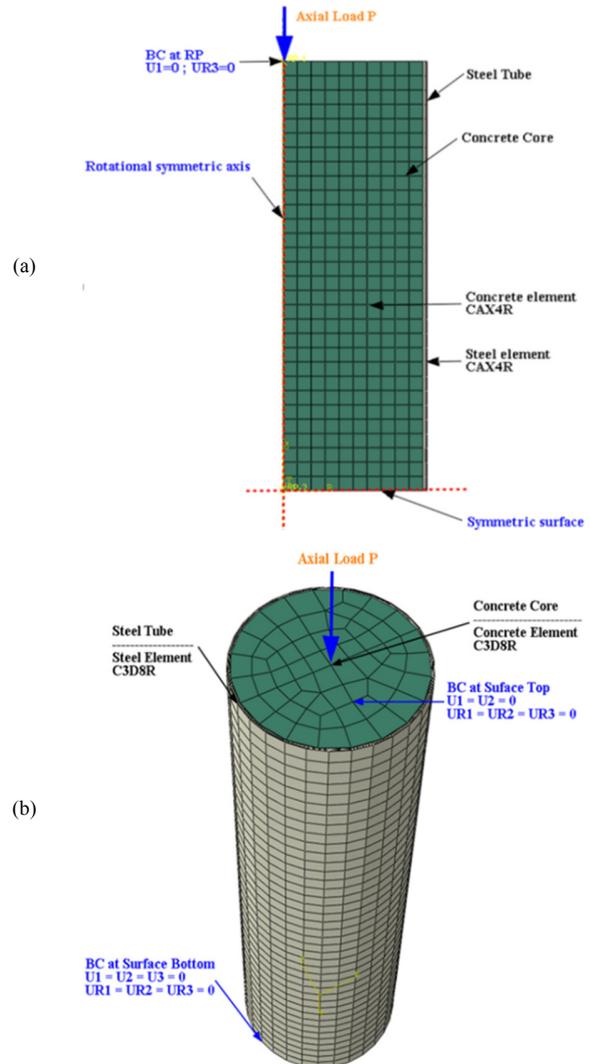


Fig. 3. Modeling of a CFST column. (a) Axisymmetric model, (b) full 3D model.

The finite element model of Figure 3 is developed in ABAQUS. The steel tube and concrete are simulated by a 4-node bilinear element with reduced integration with hourglass control (CAX4R). The mesh size of each element ranges from $D/20$ to $D/10$ which can provide accurate simulation. For the interaction between the steel tube and the concrete core, the surface-to-surface contact with *Hard contact* condition is used to model linear behavior, and the *Column friction model* is used to simulate the nonlinear behavior with a coefficient of 0.6 [3]. For boundary condition, the bottom surfaces are assumed to be fully fixed against all degrees of freedom. The incremental load is applied on the top of the column using the displacement control method to solve.

B. Effect of the Mesh Size on Finite Element Analysis

The accuracy and required computing time of the FE model are investigated in this section. Different meshing approaches, including the axisymmetric model and 3D model, are generated. Figure 4 illustrates the stress distribution from each model under the axial force. It can be seen that there is no obvious effect on the predicted Von Mises stress for the CFST column. For better consideration, the axial load-strain ($N - \epsilon$) relationship is obtained, as shown in Figure 5. The predicted responses from the two meshing approaches are almost the same as those from the experiment-tested CU-070 [20]. An important finding is that the computational time in the axisymmetric model is only 22sec which is less by a factor of 20 when compared to the 3D model. It can be concluded that the FE axisymmetric model can give accurate and efficient computation.

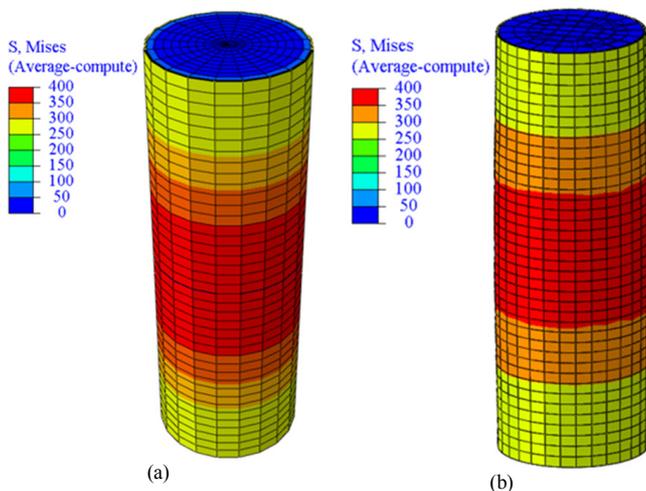


Fig. 4. Stress distribution of different models. (a) Axisymmetric model, (b) full 3D model.

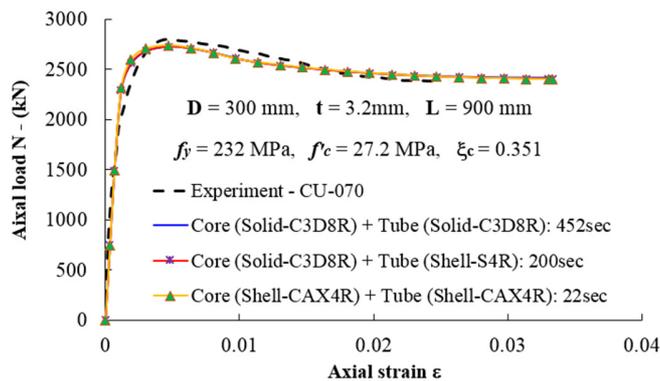


Fig. 5. Influence of the meshing and element types.

IV. VERIFICATION AND DISCUSSION

A. Ultimate Strength

The data of 663 circular CFST columns from the previous studies are selected for validation. The ultimate strength N_u is defined as the peak load of the $N - \epsilon$ curve. In this paper, the

ultimate strength from the proposed model N_u^p is compared with those obtained from the available models in [3, 20]. N_u^p is normalized with the measured ultimate strength N_u^m , which is denoted as N_u^p/N_u^m . The comparison is summarized in Table I. It can be seen that the result obtained from the FE model is more conservative than those produced from the available models. Additionally, the smallest value of standard deviation for the proposed model is found to be 0.055. The results indicate that the predicted outcomes from the current FE model are more reasonable than the previous models.

TABLE I. ULTIMATE STRENGTH OBTAINED FROM THE PROPOSED AND OTHER MODELS

	[3]	[20]	Proposed
Mean	0.975	0.980	1.011
Std.	0.090	0.129	0.055

B. $N - \epsilon$ Curves of the CFST Column

In terms of the prediction of the $N - \epsilon$ curves, the proposed model is compared to the available models from [3, 20]. The comparison is performed for different specimens obtained from [2, 16, 20]. Figure 6 illustrates the predicted $N - \epsilon$ relationship of normal CFST, with its modeling parameters taken from [20]. The observed result from the FE model agrees with the experiment very well.

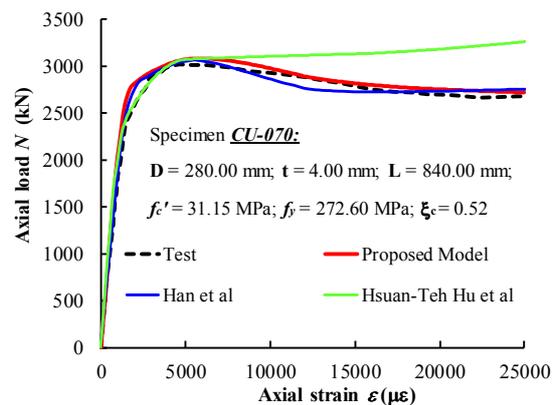


Fig. 6. $N - \epsilon$ curves of CFST column for normal specimens.

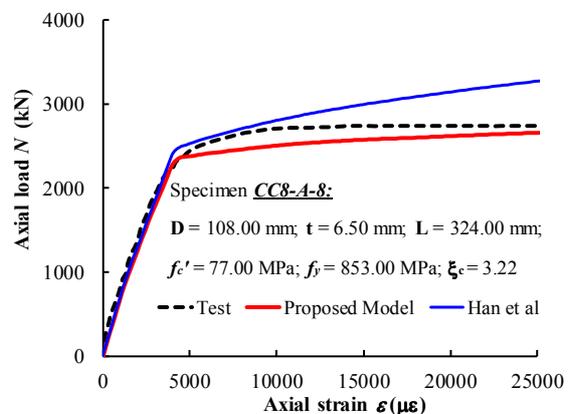


Fig. 7. $N - \epsilon$ curves of CFST column with $21.7 \leq D/t \leq 150$.

Figure 7 shows the prediction of the $N - \epsilon$ curve of the specimen reported in [2]. In this test, the ratio D/t varies from 21.7 to 150. As can be seen, the unsafety in the prediction is obtained for the model in [3] in the post-yielding strength branch. On the contract, the modified FE model gives a better prediction with the test. Furthermore, the predicted $N - \epsilon$ relationship was also performed for specimens with thin-walled steel tube and high-strength, as shown in Figure 8. The test curve is shown in [16]. The results indicate that the prediction from the FE modeling matches very well with those obtained from Han et al.'s model [16] and the test. In contrast, Hu et al.'s model [20] is an unsafe prediction in the post-yielding strength stage. The above discussion leads to the conclusion that the developed FE model can be used in simulating the CFST column under axial compression for accuracy, effectiveness, and safety predictions.

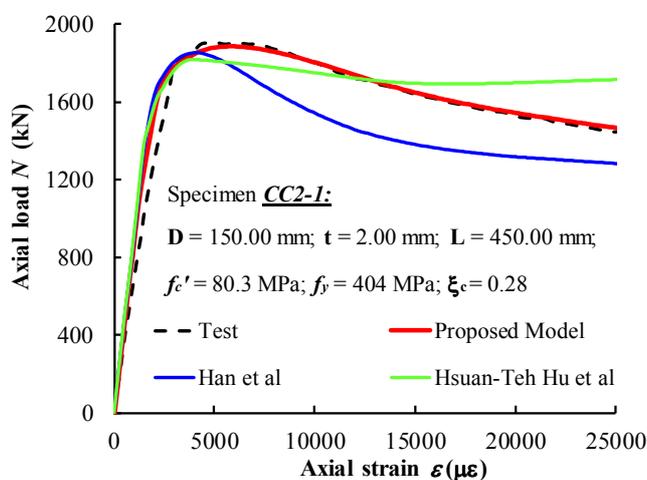


Fig. 8. $N - \epsilon$ curves of the CFST column with thin-walled and high-strength steel.

V. CONCLUSION

The current paper developed a FE modeling for axially loaded circular concrete-filled steel tube columns. The proposed model is compared with the available models in terms of accurate prediction. A new FE model is used to analyze the CFST subjected to axial compression. The model revised the $\sigma - \epsilon$ curve in [12]. The proposed model is more accurate than the models in [3, 20]. in predicting the ultimate strength. The predicted $N - \epsilon$ curve from the new model matches well with the test data, for both normal materials and thin-walled steel tubes, which indicates the versatility of the developed modeling in calculating the load-deformation curves. Regarding future work, the behavior of CFST columns with steel fiber reinforcement [21], concrete using fly ash [22], concrete with cement paste with including silica oxide nanoparticles [20], etc. can be studied by performing FEM simulation.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by the Scientific Research Fund of Ho Chi Minh City Open University, Vietnam. (No. E2019.02.2).

REFERENCES

- [1] Z. Tao, Z.-B. Wang, and Q. Yu, "Finite element modelling of concrete-filled steel stub columns under axial compression," *Journal of Constructional Steel Research*, vol. 89, pp. 121–131, Oct. 2013, <https://doi.org/10.1016/j.jcsr.2013.07.001>.
- [2] K. Sakino, H. Nakahara, S. Morino, and I. Nishiyama, "Behavior of Centrally Loaded Concrete-Filled Steel-Tube Short Columns," *Journal of Structural Engineering*, vol. 130, no. 2, pp. 180–188, Feb. 2004, [https://doi.org/10.1061/\(ASCE\)0733-9445\(2004\)130:2\(180\)](https://doi.org/10.1061/(ASCE)0733-9445(2004)130:2(180)).
- [3] L.-H. Han, G.-H. Yao, and Z. Tao, "Performance of concrete-filled thin-walled steel tubes under pure torsion," *Thin-Walled Structures*, vol. 45, no. 1, pp. 24–36, Jan. 2007, <https://doi.org/10.1016/j.tws.2007.01.008>.
- [4] P. K. Gupta, S. M. Sarda, and M. S. Kumar, "Experimental and computational study of concrete filled steel tubular columns under axial loads," *Journal of Constructional Steel Research*, vol. 63, no. 2, pp. 182–193, Feb. 2007, <https://doi.org/10.1016/j.jcsr.2006.04.004>.
- [5] Aci Committee 318, *Building Code Requirements for Structural Concrete (ACI 318-08) and Commentary*. Michigan, USA: American Concrete Institute, 2008.
- [6] R. P. Johnson and D. Anderson, *Designers' Guide to EN 1994-1-1: Eurocode 4: Design of Composite Steel and Concrete Structures. General rules and rules for buildings*. London, UK: Thomas Telford, 2004.
- [7] P. C. Nguyen, "Nonlinear Inelastic Earthquake Analysis of 2D Steel Frames," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6393–6398, Dec. 2020, <https://doi.org/10.48084/etasr.3855>.
- [8] P. C. Nguyen, B. Le-Van, and S. D. T. V. Thanh, "Nonlinear Inelastic Analysis of 2D Steel Frames: An Improvement of the Plastic Hinge Method," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5974–5978, Aug. 2020, <https://doi.org/10.48084/etasr.3600>.
- [9] V. Farhangi and M. Karakouzian, "Effect of Fiber Reinforced Polymer Tubes Filled with Recycled Materials and Concrete on Structural Capacity of Pile Foundations," *Applied Sciences*, vol. 10, no. 5, p. 1554, Jan. 2020, <https://doi.org/10.3390/app10051554>.
- [10] S.-E. Kim *et al.*, "Finite element simulation of normal - Strength CFST members with shear connectors under bending loading," *Engineering Structures*, vol. 238, Jul. 2021, Art. no. 112011, <https://doi.org/10.1016/j.engstruct.2021.112011>.
- [11] U. Katwal, Z. Tao, M. K. Hassan, and W.-D. Wang, "Simplified Numerical Modeling of Axially Loaded Circular Concrete-Filled Steel Stub Columns," *Journal of Structural Engineering*, vol. 143, no. 12, Dec. 2017, Art. no. 04017169, [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0001897](https://doi.org/10.1061/(ASCE)ST.1943-541X.0001897).
- [12] A. K. Samani and M. M. Attard, "A stress-strain model for uniaxial and confined concrete under compression," *Engineering Structures*, vol. 41, pp. 335–349, Aug. 2012, <https://doi.org/10.1016/j.engstruct.2012.03.027>.
- [13] M. D. Denavit and J. F. Hajjar, "Nonlinear Seismic Analysis of Circular Concrete-Filled Steel Tube Members and Frames," *Journal of Structural Engineering*, vol. 138, no. 9, pp. 1089–1098, Sep. 2012, [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0000544](https://doi.org/10.1061/(ASCE)ST.1943-541X.0000544).
- [14] G. D. Hatzigeorgiou, "Numerical model for the behavior and capacity of circular CFT columns, Part I: Theory," *Engineering Structures*, vol. 30, no. 6, pp. 1573–1578, Jun. 2008, <https://doi.org/10.1016/j.engstruct.2007.11.001>.
- [15] G. D. Hatzigeorgiou, "Numerical model for the behavior and capacity of circular CFT columns, Part II: Verification and extension," *Engineering Structures*, vol. 30, no. 6, pp. 1579–1589, Jun. 2008, <https://doi.org/10.1016/j.engstruct.2007.11.002>.
- [16] L.-H. Han, G.-H. Yao, and X.-L. Zhao, "Tests and calculations for hollow structural steel (HSS) stub columns filled with self-consolidating concrete (SCC)," *Journal of Constructional Steel Research*, vol. 61, no. 9, pp. 1241–1269, Sep. 2005, <https://doi.org/10.1016/j.jcsr.2005.01.004>.
- [17] D.-D. Pham and P.-C. Nguyen, "Finite Element Modelling for Axially Loaded Concrete-Filled Steel Circular Tubes," in *CIGOS 2019, Innovation for Sustainable Infrastructure*, C. Ha-Minh, D. V. Dao, F.

- Benboudjema, S. Derrible, D. V. K. Huynh, and A. M. Tang, Eds. Singapore: Springer, 2020, pp. 75–80.
- [18] D.-D. Pham, P.-C. Nguyen, D.-L. Nguyen, and H.-A. Le, "Simulation of Concrete-Filled Steel Box Columns," in *ICSCEA 2019*, Singapore, 2020, pp. 359–366, https://doi.org/10.1007/978-981-15-5144-4_30.
- [19] Z. Tao, X.-Q. Wang, and B. Uy, "Stress-Strain Curves of Structural and Reinforcing Steels after Exposure to Elevated Temperatures," *Journal of Materials in Civil Engineering*, vol. 25, no. 9, pp. 1306–1316, Sep. 2013, [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0000676](https://doi.org/10.1061/(ASCE)MT.1943-5533.0000676).
- [20] H.-T. Hu, C.-S. Huang, M.-H. Wu, and Y.-M. Wu, "Nonlinear Analysis of Axially Loaded Concrete-Filled Tube Columns with Confinement Effect," *Journal of Structural Engineering*, vol. 129, no. 10, pp. 1322–1329, Oct. 2003, [https://doi.org/10.1061/\(ASCE\)0733-9445\(2003\)129:10\(1322\)](https://doi.org/10.1061/(ASCE)0733-9445(2003)129:10(1322)).
- [21] B. Gebretsadik, K. Jdidirendi, V. Farhangi, and M. Karakouzian, "Application of Ultrasonic Measurements for the Evaluation of Steel Fiber Reinforced Concrete," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6662–6667, Jan. 2021, <https://doi.org/10.48084/etasr.3915>.
- [22] M. M. Roshani, S. H. Kargar, V. Farhangi, and M. Karakouzian, "Predicting the Effect of Fly Ash on Concrete's Mechanical Properties by ANN," *Sustainability*, vol. 13, no. 3, Jan. 2021, Art. no. 1469, <https://doi.org/10.3390/su13031469>.
- [23] M. Karakouzian, V. Farhangi, M. R. Farani, A. Joshaghani, M. Zadehmohamad, and M. Ahmadzadeh, "Mechanical Characteristics of Cement Paste in the Presence of Carbon Nanotubes and Silica Oxide Nanoparticles: An Experimental Study," *Materials*, vol. 14, no. 6, Jan. 2021, Art. no. 1347, <https://doi.org/10.3390/ma14061347>.
- [24] L.-H. Han, C.-C. Hou, and Q.-L. Wang, "Behavior of circular CFST stub columns under sustained load and chloride corrosion," *Journal of Constructional Steel Research*, vol. 103, pp. 23–36, Dec. 2014, <https://doi.org/10.1016/j.jcsr.2014.07.021>.
- [25] A. H. Varma, R. Sause, J. M. Ricles, and Q. Li, "Development and Validation of Fiber Model for High-Strength Square Concrete-Filled Steel Tube Beam-Columns," *Structural Journal*, vol. 102, no. 1, pp. 73–85, Jan. 2005, <https://doi.org/10.14359/13532>.
- [26] K. A. S. Susantha, H. Ge, and T. Usami, "Uniaxial stress-strain relationship of concrete confined by various shaped steel tubes," *Engineering Structures*, vol. 23, no. 10, pp. 1331–1347, Oct. 2001, [https://doi.org/10.1016/S0141-0296\(01\)00020-7](https://doi.org/10.1016/S0141-0296(01)00020-7).
- [27] T. Yu, J. G. Teng, Y. L. Wong, and S. L. Dong, "Finite element modeling of confined concrete-II: Plastic-damage model," *Engineering Structures*, vol. 32, no. 3, pp. 680–691, Mar. 2010, <https://doi.org/10.1016/j.engstruct.2009.11.013>.
- [28] J. G. Teng, Y. L. Huang, L. Lam, and L. P. Ye, "Theoretical Model for Fiber-Reinforced Polymer-Confined Concrete," *Journal of Composites for Construction*, vol. 11, no. 2, pp. 201–210, Apr. 2007, [https://doi.org/10.1061/\(ASCE\)1090-0268\(2007\)11:2\(201\)](https://doi.org/10.1061/(ASCE)1090-0268(2007)11:2(201)).
- [29] P. E. C. Seow and S. Swaddiwudhipong, "Failure Surface for Concrete under Multiaxial Load—a Unified Approach," *Journal of Materials in Civil Engineering*, vol. 17, no. 2, pp. 219–228, Apr. 2005, [https://doi.org/10.1061/\(ASCE\)0899-1561\(2005\)17:2\(219\)](https://doi.org/10.1061/(ASCE)0899-1561(2005)17:2(219)).
- [30] Aci Committee 318, *Building Code Requirements for Structural Concrete and Commentary (ACI 318-11)*. Michigan, USA: American Concrete Institute, 2012.
- [31] V. K. Papanikolaou and A. J. Kappos, "Confinement-sensitive plasticity constitutive model for concrete in triaxial compression," *International Journal of Solids and Structures*, vol. 44, no. 21, pp. 7021–7048, Oct. 2007, <https://doi.org/10.1016/j.ijsolstr.2007.03.022>.
- [32] Z. P. Bazant and E. Becq-Giraudon, "Statistical prediction of fracture parameters of concrete and implications for choice of testing standard," *Cement and Concrete Research*, vol. 32, no. 4, pp. 529–556, Apr. 2002, [https://doi.org/10.1016/S0008-8846\(01\)00723-2](https://doi.org/10.1016/S0008-8846(01)00723-2).
- [33] C. E.-I. D. B. Eton, *CEB-FIP Model Code 1990*. London, UK: Thomas Telford Publishing, 1993.
- [34] T. Yu, J. G. Teng, Y. L. Wong, and S. L. Dong, "Finite element modeling of confined concrete-I: Drucker-Prager type plasticity model," *Engineering Structures*, vol. 32, no. 3, pp. 665–679, Mar. 2010, <https://doi.org/10.1016/j.engstruct.2009.11.014>.
- [35] B. De Nicolò, L. Pani, and E. Pozzo, "Strain of concrete at peak compressive stress for a wide range of compressive strengths," *Materials and Structures*, vol. 27, no. 4, pp. 206–210, May 1994, <https://doi.org/10.1007/BF02473034>.
- [36] B. Binici, "An analytical model for stress-strain behavior of confined concrete," *Engineering Structures*, vol. 27, no. 7, pp. 1040–1051, Jun. 2005, <https://doi.org/10.1016/j.engstruct.2005.03.002>.
- [37] M.-X. Xiong, D.-X. Xiong, and J. Y. R. Liew, "Axial performance of short concrete filled steel tubes with high- and ultra-high- strength materials," *Engineering Structures*, vol. 136, pp. 494–510, Apr. 2017, <https://doi.org/10.1016/j.engstruct.2017.01.037>.

The Proposition of an EI Equation of Square and L-Shaped Slender Reinforced Concrete Columns under Combined Loading

Leila Hamzaoui

LGC-ROI, Civil Engineering Laboratory, Risks and Structures Interactions
Faculty of Science
University of Batna 2
Batna, Algeria
l.hamzaoui@univ-batna2.dz

Tayeb Bouzid

LGC-ROI, Civil Engineering Laboratory, Risks and Structures Interactions
Faculty of Science
University of Batna 2
Batna, Algeria
tayeb.bouzid@univ-batna2.dz

Abstract—The stability and strength of slender Reinforced Concrete (RC) columns depend directly on the flexural stiffness EI , which is a major parameter in strain calculations including those with bending and axial load. Due to the non-linearity of the stress-strain curve of concrete, the effective bending stiffness EI always remains variable. Numerical simulations were performed for square and L-shaped reinforced concrete sections of slender columns subjected to an eccentric axial force to estimate the variation of EI resulting from the actual behavior of the column, based on the moment-curvature relationship. Seventy thousand (70000) hypothetical slender columns, each with a different combination of variables, were used to investigate the main variables that affect the EI of RC slender columns. Using linear regression analysis, a new simple and linear expression of EI was developed. Slenderness, axial load level, and concrete strength have been identified as the most important factors affecting effective stiffness. Finally, the comparison between the results of the new equation and the methods proposed by ACI-318 and Euro Code-2 was carried out in connection with the experimental results of the literature. A good agreement of the results was found.

Keywords—flexural stiffness; reduction factor; reinforced concrete; rigidity; slender columns

I. INTRODUCTION

For a bi-articulated column supporting a load P applied with an eccentricity e , the corresponding moment which is equal to $P.e$ is called the "first order moment". Under the action of this moment the column deforms in simple curvature and the bending moment is at its maximum at half height and equal to $M=P.e+P.A$. The increase of the latter caused by the displacement A is called the PA effect or second order moment. The longer the column is compared to its cross section, the greater the lateral displacements will be and the greater the second order moment will be. Buckling failure will occur under a lower resistance than that which would have caused the short column to fail. In this case, the strength of slender Reinforced Concrete (RC) columns is remarkably reduced. Several factors contribute to this reduction. In addition to the PA effect,

shrinkage, creep, and cracking occur along the length of the column and are directly related to the supported part of the compressive stress block in the concrete. Leonhard Euler who based his work on Hook's law and the moment-curvature relationship studied the buckling problem. He developed (1) to predict the minimum buckling load of a perfect column:

$$p_{cr} = \frac{\pi^2 EI}{l^2} \quad (1)$$

This equation did not depend on the strength of the cross-section but rather on the bending stiffness EI . It was valid only if the moment of buckling obeys Hook's law, where the retained stiffness is constant. This is the case of a homogeneous isotropic and elastic material but this is not the case for reinforced concrete columns whose behavior in practice becomes clearly non-linear at the level of the descending branch of the σ - ϵ curve, therefore the buckling takes place in the non-linear domain. Euler's expression remains valid if the bending stiffness is replaced by a value representative of the effective stiffness at the moment of failure which is either with the crushing material (case of short columns) or with buckling (case of thin columns). During the loading and due to the inelastic behavior depicted by the stress/strain curve of concrete the value of the EI varies at each point along the section. The difficulty to choose a single value of EI appears because some methods consider that the column stiffness is constant along the section and the length of the column, when in addition it is necessary to consider shrinkage, creep, and lateral deflection. If we introduce the effects of concrete cracking, this can cause a large variation in bending stiffness. The objective of this work is to develop an analytical model for the study and analysis of the non-linear material and geometrical behavior of slender reinforced concrete columns of square and "L" cross-section under axial load and biaxial bending in order to find a simple relationship explicitly describing the real behavior of the effective bending stiffness EI as a function of parameters such as concrete strength f_c , steel yield strength f_y , cross-sectional dimensions, percentage of reinforced steel ρ , and column length L . In addition, this

Corresponding author: Leila Hamzaoui

stiffness is compared with experimental test results from the literature and design codes ACI-318 [1] and EC-2 [2]. The studied columns were bent into a single symmetrical curvature, in non-oscillating frames subjected to short-term loads.

The behavior of RC slender columns under axial loads and bending moments in the two main planes of inertia is of great interest for structural analysis and design. The difficulty in this case is in the fact that it is not possible to give a general mathematical formulation as in the case of uniaxial bending because the position of the neutral axis is not known in advance. The difficulty increases if the cross-section of the column is not symmetrical, and increases even more if the behavior law (σ - ϵ) is not linear. Some researchers have proposed an iterative method initiated in [3] which consists of establishing by successive trials the neutral axis for which the internal forces in the section are in equilibrium with the external loads. Other methods, based on the determination of interaction curves, interaction surfaces or interaction contours for loading, were used for calculation and verification [4]. In most computer programs for reinforced concrete structures, there are indications of the limit state justification of square and rectangular columns while the L-shaped cross-section is not treated. It is therefore interesting to fill in the gaps in the equations presented in the literature because they are valid only for ordinary sections, and in practice, there are sections of different shapes such as tee (T), simple cross (+), hexagonal or thin-walled boxes, and finally L-shaped. In fact, the latter presents several difficulties like the absence of vertical and horizontal axes of symmetry.

II. LITERATURE REVIEW

Flexural rigidity is a necessary component in the estimation of the deflection of RC elements subjected to bending moment and axial loading, especially slender columns. It should be noted that the deflection is important because it has a direct and strong influence on the second order moments induced by the axial loads. Recently, major progress has been made in the development of nonlinear methods for the analysis of slender RC columns. Numerous methods are used to evaluate the bending stiffness EI taking into account all the factors that contribute to its decrease at the level of the concrete stress block in compression (non-linear material behavior) and at the level of the lateral displacement length of the column (non-linear geometric behavior) to determine the critical buckling load nearest to the effective values. Accordingly, the effective bending stiffness of columns under short-term loading can be expressed as: $EI = \alpha \cdot E_c \cdot I_c + E_s \cdot I_s$ where α is the stiffness reduction factor, E_c and E_s are the elasticity modulus of concrete and steel respectively, I_c and I_s are the moments of inertia of the cross-section of concrete and steel about the central axis of the cross-section respectively. After determining the factor α , the EI can be easily calculated. The factor α is a function of the material properties, the steel ratio, the axial load, and other parameters. This reduction factor of bending stiffness has been evaluated and studied by a number of researchers and some current codes and according to [1], is always taken equal to 0.2 for columns. The Eurocode EC-2 [2] is provided with a coefficient of flexibility K_e which is a function of a set of parameters such as: the relative normal

force ν , the mechanical slenderness λ_m , the geometric reinforcement ratio ρ , the creep rate ϕ_{eff} , and the characteristic strength of the concrete compression cylinder at 28 days f_{ck} . Many formulations have been proposed in the literature. Authors in [5-7] have presented statistical evaluations of the flexural rigidity of RC columns and of rectangular cross-sections of composite RC columns. Authors in [8] proposed a reduction coefficient as a function of various parameters such as slenderness ratio (L/h), eccentricity ratio (e/h), and axial load ratio (P/P_0). Authors in [6] proposed a general expression to obtain an effective flexural stiffness applicable to any cross-sectional shape under short or long-term biaxial eccentric loading for normal and high concrete strength. Authors in [9] proposed a study for the reduction factor of the RC column with equi-axial T section and in [10] dealt with the cross-shaped cross-section. Authors in [11] presented new expressions for a beam with rectangular and T-shaped cross-sections, square, rectangular and circular cross-sections for columns in function of concrete strength f_c , steel percentage ρ , and level of charge P/P_0 . Authors in [12] presented their approach to find effective flexural stiffness for circular columns. In this research, the stiffness reduction factor was suggested to treat slender square and L-section columns with equal flanges. Some researchers have conducted studies on eccentrically loaded L-sections. Authors in [13] presented an experimental and analytical research on the load carrying capacity of L-shaped RC columns with equal and unequal flanges. Authors in [14] carried out a theoretical study on the calculation of L and square sections by the interaction approach. Authors in [15] based their study on the finite element method to analyze and draw the interaction surfaces of L and square sections. An interesting research concerning the behavior of L-section and square composite columns under biaxial eccentric loading according to different laws of concrete behavior is presented in [16]. Authors in [9] studied the L section with equal flanges and suggested a formula for the bending stiffness reduction factor of slender RC columns under different axial load levels and different seismic action levels. The current work is a contribution in this field of study and, in addition, a proposition of a simple solution for the calculation and verification of thin columns is presented.

III. ANALYSIS METHOD

In order to study the effective bending stiffness values of the slender RC column under combined biaxial bending and axial loading, nonlinear computer analysis was developed. Material and geometric nonlinearity were included in the research study for L and square sections. It should be noted that articulated end columns with equal load eccentricities at the ends acting in the same plane have been taken in account. The analysis can be divided into three parts: (a) the cross-sectional strength (analysis of cross section), (b) the effect of slenderness on the column strength (analysis of second order), (c) the third and main objective is the development of a new formula for estimating the EI bending stiffness of slender biaxially loaded RC columns in which these factors are included.

A. Material Behavior and Assumptions

The studied columns are all made of RC. This composite material consists of two elements with different characteristics.

In order to simulate the non-linear analysis of the complete response of RC columns of square and L-shaped cross-sections with equal flanges under biaxial bending and compression, a computer program compiled by an integral numerical method was used, based on the following assumptions:

- Plane sections before bending remain plane after bending.
- The tensile strength of concrete is neglected, and the stress-strain curve of concrete in compression refers to the unconfined diagram in [17], as shown in Figure 1(a), whose curve characteristics are calculated using (2) and (3):

$$\sigma = f_c \left[\left(2 \frac{\epsilon_c}{\epsilon_{c0}} \right) - \left(\frac{\epsilon_c}{\epsilon_{c0}} \right)^2 \right] \quad (2)$$

$$\text{for } 0 \leq \epsilon_c \leq \epsilon_{c0} = 0.002$$

$$\sigma = f_c [1 - Zu(\epsilon_c - \epsilon_{c0})] \quad (3)$$

$$\text{for } \epsilon_{c0} \leq \epsilon_c \leq \epsilon_{cu} = 0.004$$

with:

$$Zu = \frac{0.5}{\epsilon_{50u} - \epsilon_{c0}}, \quad \epsilon_{50u} = \frac{3 + 0.285 f_c}{142 f_c - 1000}, \quad \epsilon_{50u} \geq \epsilon_{c0} \quad (f_c \text{ mpa})$$

- The stress-strain curve of steel is taken as the usual bilinear relationship shown in Figure 1(b).

The section is divided into a number of small fiber elements. The concrete element stress distribution is also divided into a number of small fiber elements. Concrete shrinkage and creep are ignored.

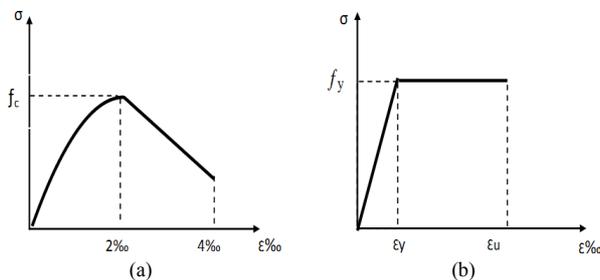


Fig. 1. Concrete and steel σ - ϵ diagram

1) Cross Section Analysis

This work presents an indirect method for the analysis of the ultimate limit states of RC columns of L and square cross-sections subjected to biaxial loading. We consider a square section of dimensions $B \times B$, and another square section of dimensions $B_1 \times B_1$ with $B_1 < B$ being away from the corner of the original square ($B \times B$). Removing the section $B_1 \times B_1$ results in an L-shaped section, as shown in Figure 2(a). For different values of B_1 , we can obtain different dimensions of the L-section. If ($B_1=0$) the L-section becomes a square section. The ratio (B_1/B) varies from 0 to 0.6. It is assumed that the steel reinforcement is evenly distributed around the perimeter of the cross-section (see Figure 2(b))[14].

2) Internal Forces

The cross section is divided into infinitely small slices (see Figure 3). For each one of them, the internal forces (concrete,

compression reinforcement and tensile reinforcement) are evaluated to give finally an axial load and an ultimate moment. The resulting axial load and the resulting moment must be equal respectively to the sum of the internal loads and the sum of elementary moments.

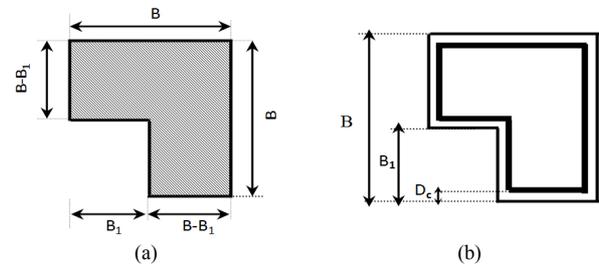


Fig. 2. L cross section and steel reinforcement distribution.

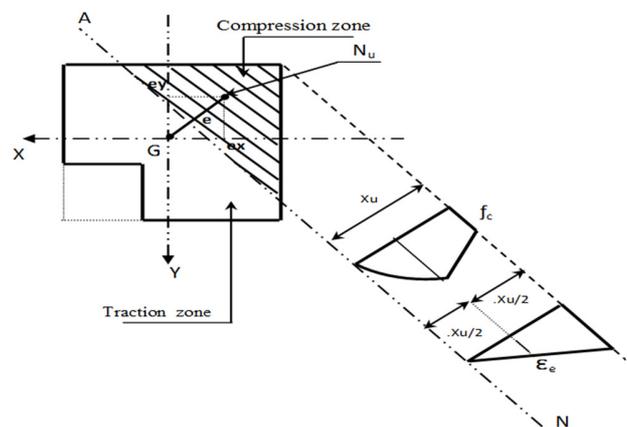


Fig. 3. L cross section subjected to biaxial load.

The resulting axial load and the resulting moment are equal respectively to the sum of the loads and elementary moment shown in (4), and (5):

$$P = \sum_{i=1}^n f_c \cdot A_{ci} \cdot \sum_{i=1}^n (f_y - f_c) \cdot \rho \cdot \frac{A_{ci}}{100} \quad (4)$$

$$M = \sum_{i=1}^n f_c \cdot Y_{ci} \cdot A_{ci} \cdot \sum_{i=1}^n (f_y - f_c) \cdot \rho \cdot Y_i \cdot \frac{A_{ci}}{100} \quad (5)$$

where n is the number of slices, A_{ci} is the concrete area of the slice i , $\rho = (As/A_c) \cdot 100$ presents the percentage of longitudinal reinforcement, As is the area of longitudinal reinforcement, and Y_i is the distance from the edge of steel to the center of the section.

Analysis is done by sweeping the section parallel to the neutral axis. For the L-shaped section, the computer program studies different positions of the neutral axis according to the angle of inclination θ , where $\theta = \tan^{-1}(ex/ey)$ with ex and ey being the eccentricities along the x and y axis respectively. The computer analysis has been prepared by considering the axis of bending. Figure 4 explains these cases:

- Case 1: Uniaxial bending, angle of inclination $\theta=0^\circ$.
- Case 2: Uniaxial bending, angle inclination $\theta=90^\circ$.

- Case3: Biaxial bending, angle of inclination $0 < \theta < 45^\circ$.
- Case4: Biaxial bending, angle of inclination $45^\circ < \theta < 90^\circ$.

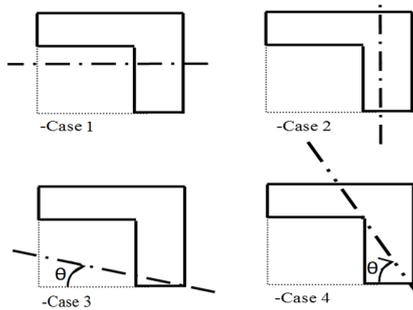


Fig. 4. The four cases of uniaxial and biaxial load.

The codes of the above mentioned cases were developed in FORTRAN and were used to obtain the ultimate resistance P_u and moments M_u as results of the different positions of the neutral axis. The input data include the size of the square B , the depth B of the section, the depth of embedding D_c , the characteristic strengths of concrete f_c and steel f_y , the moduli of elasticity of steel E_s and concrete E_c , the angle of inclination θ , and the percentage of reinforcement ρ .

IV. SECOND ORDER ANALYSIS

Columns with a moderate slenderness may have their secondary moments determined with an approximate analysis such as the moment magnifier method which is used in this study. This procedure uses a moment magnification factor δ to amplify the applied moment M_a sufficiently for the account of secondary moments and to ensure safe design. For columns subjected to axial loading and equal and opposite end moments, the maximum bending moment can be taken as:

$$M_c = \delta \cdot M_a \quad (6)$$

where:

$$\delta = \frac{1}{1 - \frac{P}{P_{cr}}}$$

where M_c is the design bending moment that includes second-order effects, M_a is the applied column bending moment, P is the factored axial load acting on the column, and P_{cr} is Euler's buckling strength. The secondary moment due to the column slenderness is calculated as the difference between the internal resisting moment and the applied moment. For the purpose of analysis, M_c is replaced by the cross section bending moment M_{CS} , and M_a is replaced by the overall column bending moment (M_{col}) as shown in (7):

$$M_{CS} = \left(\frac{1}{1 - \frac{P}{P_{cr}}} \right) \cdot M_{col} \quad (7)$$

V. CALCULATION OF THE FLEXURAL STIFFNES OF THE COLUMN

The design of the flexural stiffness as a function of the moment-to-curvature ratio M/Φ is the most reasonable relation of the EI representation. The analysis studies show various

cross-sectional parameters as well as axial loading influencing the bending stiffness. The computer program is structured according to the well-known (P, M, Φ) relation. A linear strain is assumed and the Kent & Park relation [17] relates stress to strain. The curvature is calculated directly from the geometry (Φ) and it is known as the exact value in this computer program and the extreme fiber is assumed to be at crush strain. The depth of the neutral axis X_u can be incremented throughout the program to provide the necessary strain variation. Therefore, this depth starts at a small fraction of the total section depth and is increased in defined increments until the maximum of the column capacity in pure axial compression is reached (P_0). The curvature is obtained by dividing the stress of the extreme fiber (usually 0.003) by the depth of the compression zone ($\Phi = 0.003/X_u$). Using this curvature and assuming a linear strain distribution, it is possible to find the stress in each fiber. For each increment, the axial load P , the bending moment of the cross section M_{CS} , and the curvature Φ are calculated. Equation (1) is used to find the critical load P_{cr} which allows calculating the bending moment of the column M_{col} through (7) whose slenderness effects were taken into account. The program uses the formulation M_{col}/Φ to calculate the effective bending stiffness of a slender column whose nonlinearities due to $P\Delta$ effects and nonlinear behavior of the material are taken into account at all times.

VI. PROPOSED FORMULA OF THE REDUCTION FACTOR OF FLEXURAL STIFFNESS

In practice, the engineer often has difficulties to give a reasonable interpretation of the bending stiffness that represents the real behavior of slender RC columns. This stiffness is easy to calculate with the M/Φ method, which gives satisfactory results, but it is not easy to solve manually making the use of a computer recommended which in this case makes it mandatory but not preferable. In order to simplify the estimation of EI bending stiffness under short-term loads, the well-known equation $EI = a \cdot E_c I_c + E_s I_s$ is obtained. Research and expressions have proposed to evaluate the reduction factor or the factor of effective stiffness which is of origin can be expressed by:

$$\alpha = \frac{EI - E_s I_s}{E_c I_c} \quad (8)$$

Thus, to evaluate the EI flexural stiffness, a simple expression is proposed to avoid sophisticated and complicated calculations. Multiple linear regression approaches, taking into account the simulated theoretical stiffness data, are conducted to evaluate the EI expression. Linear regression is chosen as the method of analysis because the objective is to develop a simple and accurate equation for the reduction factor α . This method is very useful in order to determine the effective behavior of certain parameters. For example the work in [18] was based on the same principle to give the equation of effective ductility as a function of shear reinforcement ratio and compressive strength of concrete. In this work the used analysis method consists in finding an expression describing the reduction factor α for different values of concrete strength f_c , steel yield f_y , percentage of reinforced steel ρ , steel ultimate strength f_u , L section dimensions, square or equal flanges presented by the ratio B_1/B , slenderness ratio L/B , the neutral axis angle θ which

presents the loading angle, and the axial loading ratio P/P_0 , where P_0 is the strength of the axial compression column under pure axial loading. This regression concerns a column of a square and equal L section presents the proprieties illustrated in Table I, based on the statistical analysis of the theoretical values of α for all simulated columns in which P/P_0 ranged from 0.1 to 1. Over 70000 isolated reinforced concrete columns were simulated to generate the effective stiffness factor. Each simulated column had a different combination of cross section, geometric, and material proprieties.

TABLE I. SPECIFIED PROPERTIES OF STUDIED RC COLUMNS

Proprieties	Specified values	Number of specified values
F_c	15 ;60 incr 5mPa	10
F_y	200;600 incr 100mPa	5
B_l/B	0 ; 0.6 incr 0.1	7
L/B	10 ; 30 incr 5	5
P	0.5 ;4% incr 0.5%	8
θ	0°; 90° incr 10°	8
Total number =10×5×7×5×8×8=70000 columns		

Finally the following equation of reduction factor α is obtained:

$$\alpha = -0.074 + 0.0003 \cdot f_c + 0.488 \cdot \frac{P}{P_0} - 0.002 \cdot \frac{L}{B} \quad (9)$$

The effect of f_c has an increasing influence on the flexural stiffness of the slender column. This was already predicted and proven in [19] which presented the effect and impact of this strength on the general behavior of corner columns. In addition to the axial load level, it has an increasing effect on the EI of the column. The expression also indicates that a decrease in the EI value occurs as slenderness ratio (L/B) ratio is increased.

VII. VALIDATION

For the experimentation of the proposed model a test series has been realized to determine the variation of the reduction factor according to the influence of different parameters on the behavior of the RC slender columns. The proposed method has been realized according to the expression of the reduction factor for an ultimate compressive concrete strain equal to 0.003 because it gives the best results among all other values of strain levels which varied between 0.002 and 0.004. The square cross section has dimensions $B_l=0$ and $B=300\text{mm}$ and the L section $B_l=100\text{mm}$ and $B=300\text{mm}$. Five methods are considered in this work to predict the nominal capacity of the RC columns. The curves in Figure 5 show the comparison between the proposed method, the codes [1, 2], and the methods proposed in [3, 6, 8]. All respective equations are presented in Table II. A comparison between the proposed method, and that suggested by Bonet et al. in [6] for the L-shaped cross-section is presented in Figure 6(a) showing the variation of the reduction factor as a function of the relative eccentricity η . This coefficient presented by Bonet has been calculated as $\eta=e/4.i_c$, where i_c is the radius of gyration of the concrete section with respect to the bending axis and e is the first order eccentricity. The effective stiffness factor shows a non-linear behavior as a function of η . A second comparison with Bonet equation presented in Figure 6(b) shows the effect of ρ on the α of the slender column with a different angle loading θ .

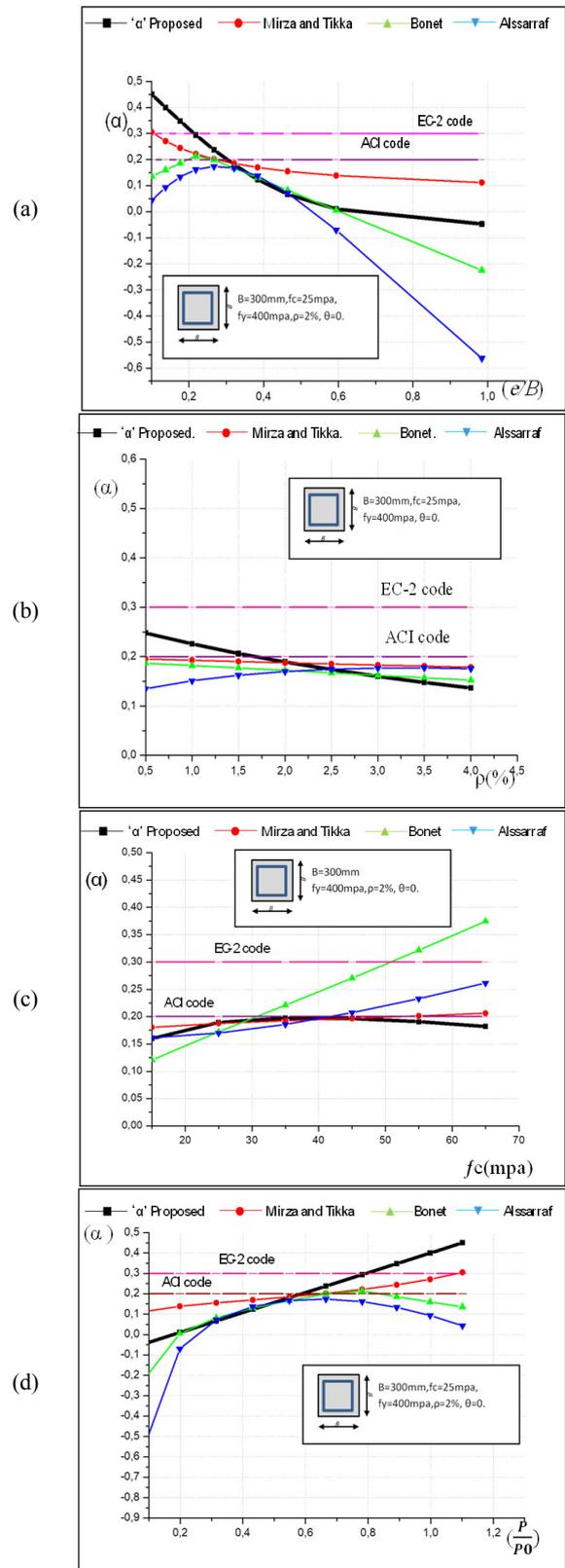


Fig. 5. Comparison of α with different authors and design codes for a square sections: Effect of (a) e/B , (b) ρ , (c) f_c , and (d) of the axial loading ratio on the α of a slender column.

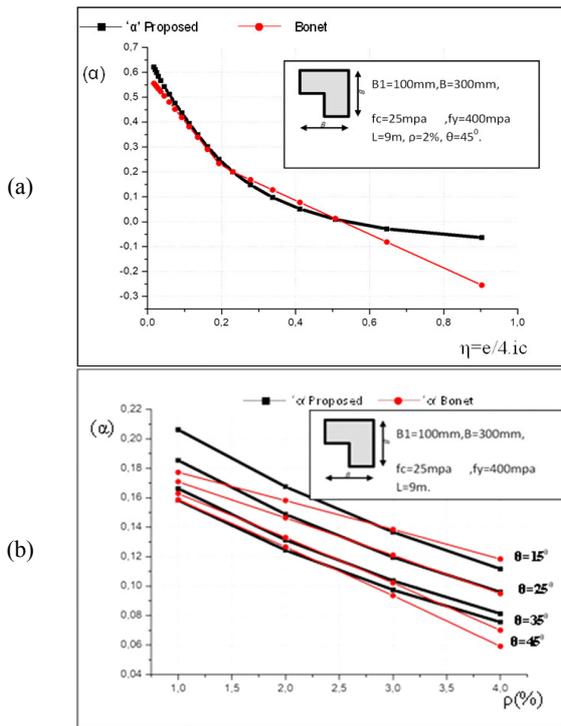


Fig. 6. Comparison between the Bonet equation and the proposed α for the L shaped section. (a) Effect of α of the slender column, (b) effect of ρ on α with a variation of angle loading θ .

TABLE II. EI EQUATION COMPARISON

Author	Effective flexural stiffness EI
Proposed	$EI = \alpha E_c I_c + E_s I_s$ $\alpha = -0.074 + 0.0003 f_c + 0.488 (P/P_0) - 0.002 (l/h)$
[1]	$EI = 0.2 E_c I_c + E_s I_s / 1 + \beta d$ $EI = 0.4 E_c I_c / 1 + \beta d$
[2]	$EI = K_c \cdot E_c \cdot I_c + K_s \cdot E_s \cdot I_s$ If $\rho \geq 0.2\%$, $K_s = 1$, $K_c = (K1 K2) / (1 + \phi_{eff})$, $K1 = \sqrt{f_c k} / 20$, $K2 = \nu \lambda_m / 170$ If $\rho \geq 1\%$, $K_s = 0$, $K_c = 0.3 / 1 + 0.5 \phi_{eff}$
[17]	$EI = (\alpha E_c I_c / 1 + \phi_{eff}) + (E_s I_s / 1 + \eta)$ $\alpha = (1.95 - 0.035 \lambda_m - 0.25 \phi_{eff}) (\eta - 0.2) + (f_c / 225) + 0.11 < 0.1$; if $\eta < 0.2$ $\alpha = (f_c / 225 + 0.45) (\eta - 0.2) + (f_c / 225 + 0.11) < 0.1$; if $\eta \geq 0.2$
[6]	$EI = \alpha E_c I_c + 0.8 E_s I_s$ $\alpha = 0.47 - 0.35 (e/h) (1 / 1 + \beta (e/h)) + 0.003 (l/h)$ where $\beta = 7.0$ for $p_{rs} \leq 2\%$; and $\beta = 8.0$ for $p_{rs} > 2\%$.
[8]	$EI = \alpha E_c I_c + E_s I_s$ $\alpha = 0.38 - 0.011 (l/h) - 1.3 (e/h) + 0.45 (1 - (P/P_0)^2)$ $0.1 \leq \alpha \leq 0.85$

EI = flexural stiffness of compression member, E_c = modulus of elasticity of concrete, E_s = modulus of elasticity of reinforcement, I_c = moment of inertia of gross concrete section, I_s = moment of inertia of reinforcement, P = axial load, P_0 = maximum load capacity, e/h = eccentricity ratio, l/h = geometrical slenderness ratio, βd = ratio of the maximum factored axial sustained load to the maximum factored axial load associated with the same load combination, ν = relative normal force, λ_m = mechanical slenderness ratio, ϕ_{eff} = effective creep ratio, f_c = concrete strength, ρ = geometrical reinforcement ratio, f_{ck} = characteristic compressive cylinder strength of concrete at 28 days.

VIII. RESULTS AND DISCUSSION

This paper presents a statistical evaluation of the parameters that affect the flexural stiffness of slender columns subjected to short term loads with biaxial bending. Five methods [1-3, 6, 8] for this comparison were considered to predict effective rigidity. In order to verify (9), variations of the e/B , P/P_0 , ρ , and

f_c were used. These parameters are the most appeared in the expressions proposed in the literature. It can be noticed that both design codes propose an effective stiffness factor α independent of all the parameters (Figure 5), however other authors include the dependence of α . In general an equation of α is proposed that decreases with e/B and ρ and increases with f_c and P/P_0 . The proposed equation confirms this behavior of α . The majority of tests considered the square cross section, with the exception of Bonet's expression which considered a cross section shape. A very large correspondence can be noticed between the Bonet's equation and the proposed equation (Figure 6) despite that Bonet's is nonlinear and the proposed expression is perfectly linear. Liu et al. [9] proposed a reduction stiffness factor α_e taking all the rigidity EI of the material, not only the concrete material of RC slender columns. The calculation model of columns has an L section of $600 \times 600 \times 200 \text{ mm}$. The cross section details and dimensions of the specimen are shown in the Table III. The stiffness reduction factors calculated with Liu formulation and the flexural stiffness EI corresponding for each axial load level are shown in Table IV. A comparison of the results of Liu with the flexural stiffness calculated with the proposed equation shows a good degree of accuracy between the results of Liu and the proposed equation (Figure 7). An average ratio of 0.983 was achieved.

TABLE III. GEOMETRY AND DIMENSIO OF THE SPECIMENS IN [9]

Column characteristic	Values
B_l (mm)	400
B (mm)	600
D_c (mm)	30
ρ (%)	2.512
f_c (mpa)	14.3
f_y (mpa)	360
L (mm)	2900
θ (°)	45
E_c (mPa)	30000
E_s (mPa)	200000
$EI = E_c I_c + E_s I_s$ (KN.m ²)	77686

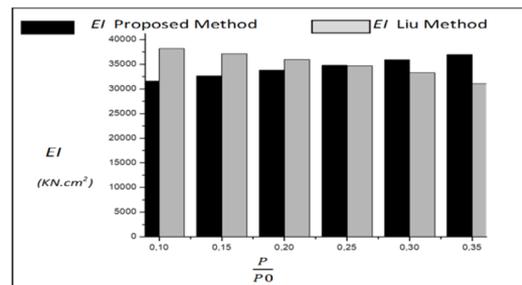


Fig. 7. Comparison of the proposed EI and the one of [9].

TABLE IV. VERIFICATION OF THE PROPOSED EQUATION WITH [9]

Axial load level	α_e	$EI_{Liu} = \alpha_e EI$ (KN.m ²)	EI_{prop} (KN.m ²)	EI_{prop}/EI_{Liu}
0.10	0.4915	38182	31587	0.83
0.15	0.4782	37149	32662	0.87
0.20	0.4630	35968	33787	0.94
0.25	0.4464	34679	34790	1.00
0.30	0.4281	33257	35914	1.07
0.35	0.3996	31043	36981	1.19
Average				0.983

IX. CONCLUSION

An iterative numerical procedure for the stiffness analysis and design of slender RC columns with square and L-shaped cross-sections under biaxial bending and axial loading using the unconfined Kent and Park stress is presented in this paper. The computational procedure takes into account the nonlinear behavior of the materials and includes the second order effects due to the additional eccentricity of the axial loading as applied by the moment-amplification method. The new equation suggested in this work is among the simplest proposed equations and its implementation is very easy to solve by manual calculation in a relatively short time. The concrete strength, the level of axial load, and the column slender ratio have been identified as the most important factors affecting the reduction coefficients and the effective bending stiffness EI . The formulation of the proposed method has been tested by comparison with other reported formulations. The results show that a good degree of accuracy and agreement has been achieved.

REFERENCES

- [1] *ACI 318-08: Building Code Requirements for Structural Concrete and Commentary*. ACI, 2008.
- [2] *EN 1992-1-1:2004 - Eurocode 2: Design of concrete structures - Part 1-1: General rules and rules for*. Brussels, Belgium: CEN, 2004.
- [3] C. S. Whitney and E. Cohen, "Guide for Ultimate Strength Design of Reinforced Concrete," *Journal Proceedings*, vol. 53, no. 11, pp. 455–490, Nov. 1956, <https://doi.org/10.14359/11524>.
- [4] B. Bresler, "Design Criteria for Reinforced Concrete Columns under Axial Load and Biaxial Bending," *ACI Journal, Proceedings*, vol. 10, no. 6, pp. 481–490, Jan. 2005.
- [5] M. Khuntia and S. K. Ghosh, "Flexural Stiffness of Reinforced Concrete Columns and Beams: Experimental Verification," *ACI Structural Journal*, vol. 10, no. 3, May 2004.
- [6] J. L. Bonet, M. L. Romero, and P. F. Miguel, "Effective flexural stiffness of slender reinforced concrete columns under axial forces and biaxial bending," *Engineering Structures*, vol. 33, no. 3, pp. 881–893, Mar. 2011, <https://doi.org/10.1016/j.engstruct.2010.12.009>.
- [7] M. Resheidat, M. Ghanma, C. Sutton, and W.-F. Chen, "Flexural rigidity of biaxially loaded reinforced concrete rectangular column sections," *Computers & Structures*, vol. 55, no. 4, pp. 601–614, May 1995, [https://doi.org/10.1016/0045-7949\(94\)00493-M](https://doi.org/10.1016/0045-7949(94)00493-M).
- [8] S. Z. Al-Sarraf, I. A. S. Al-Shaarbaf, B. R. Al-Bakri, and K. F. Sarsam, "Flexural Rigidity of Slender RC Columns," *Engineering and Technology Journal*, vol. 27, no. 1, pp. 96–116, 2009.
- [9] J. Liu, Y. Zhang, and P. Luo, "Flexural Stiffness Reduction Factor of Reinforced Concrete Column with Equal L-Shaped Section," in *The Twelfth COTA International Conference of Transportation Professionals*, Dec. 2012, pp. 3187–3193, <https://doi.org/10.1061/9780784412442.325>.
- [10] X. Feng, M. Shen, C. Sun, J. Chen, and P. Luo, "Research on Flexural Stiffness Reduction Factor of Reinforced Concrete Column with Equiaxial + Shaped Section," *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 168–174, Nov. 2013, <https://doi.org/10.1016/j.sbspro.2013.08.022>.
- [11] Ö. Avşar, B. Bayhan, and A. Yakut, "Effective flexural rigidities for ordinary reinforced concrete columns and beams," *The Structural Design of Tall and Special Buildings*, vol. 23, no. 6, pp. 463–482, 2014, <https://doi.org/10.1002/tal.1056>.
- [12] N. Caglar, A. Demir, H. Ozturk, and A. Akkaya, "A simple formulation for effective flexural stiffness of circular reinforced concrete columns," *Engineering Applications of Artificial Intelligence*, vol. 38, pp. 79–87, Feb. 2015, <https://doi.org/10.1016/j.engappai.2014.10.011>.
- [13] L. N. Ramamurthy and T. A. H. Khan, "L-Shaped Column Design for Biaxial Eccentricity," *Journal of Structural Engineering*, vol. 109, no. 8, pp. 1903–1917, Aug. 1983, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1983\)109:8\(1903\)](https://doi.org/10.1061/(ASCE)0733-9445(1983)109:8(1903)).
- [14] Mallikarjuna and P. Mahadevappa, "Computer aided analysis of reinforced concrete columns subjected to axial compression and bending—I L-shaped sections," *Computers & Structures*, vol. 44, no. 5, pp. 1121–1138, Aug. 1992, [https://doi.org/10.1016/0045-7949\(92\)90333-U](https://doi.org/10.1016/0045-7949(92)90333-U).
- [15] W. H. Taso and C.-T. T. Hsu, "Behaviour of biaxially loaded square and L-shaped slender reinforced concrete columns," *Magazine of Concrete Research*, vol. 46, no. 169, pp. 257–267, Dec. 1994, <https://doi.org/10.1680/mac.1994.46.169.257>.
- [16] C. Dundar, S. Tokgoz, A. K. Tanrikulu, and T. Baran, "Behaviour of reinforced and concrete-encased composite columns subjected to biaxial bending and axial load," *Building and Environment*, vol. 43, no. 6, pp. 1109–1120, Jun. 2008, <https://doi.org/10.1016/j.buildenv.2007.02.010>.
- [17] D. C. Kent and R. Park, "Flexural Members with Confined Concrete," *Journal of the Structural Division*, vol. 97, no. 7, pp. 1969–1990, 1971.
- [18] H. Wang, Y. Zhang, and S. Qin, "A Study on Ductility of Prestressed Concrete Pier Based on Response Surface Methodology," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1253–1257, Dec. 2016, <https://doi.org/10.48084/etasr.855>.
- [19] A. Ali, Z. Soomro, S. Iqbal, N. Bhatti, and A. F. Abro, "Prediction of Corner Columns' Load Capacity Using Composite Material Analogy," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2745–2749, Apr. 2018, <https://doi.org/10.48084/etasr.1879>.

The Influence of Fly Ash on the Compressive Strength of Recycled Concrete Utilizing Coarse Aggregates from Demolition Works

Van Tien Phan

Department of Construction

Vinh University

Vinh City, Vietnam

tienkxd@vinhuni.edu.vn, vantien.phan1010@gmail.com

Tien Hong Nguyen

Department of Construction

Vinh University

Vinh City, Vietnam

tienhongkxd@vinhuni.edu.vn

Abstract- The compressive strength of recycled concrete with coarse aggregates with and without fly ash was studied in this paper. The recycled concrete was designed to have a compressive strength of $R_n=15\text{Mpa}$. The ratio of conventional aggregates and recycled coarse aggregates was 50-50. Fifteen samples for each recycled concrete aggregate had been cast and were tested 7, 14, and 28 days after curing. The experimental results indicate that the strength development of the two recycled concrete mixes is relatively uniform. Fly ash increased the compressive strength of recycled concrete by 11% for 7-day aged samples. The recycled concrete with fly ash had 7% less compressive strength than the concrete consisting of natural aggregates.

Keywords- fly ash; recycled concrete; demolishing work

I. INTRODUCTION

Nowadays, reuse of solid waste from demolition works is a common issue, especially at developing cities. Construction solid waste is hardly reused due to various reasons such as the lack of solid waste treatment plants and concentration areas, strict environmental regulations, etc. In fact, many investors dump solid waste from demolition works to save time and money, which has a negative impact on the environment and the locals. Reusing construction solid waste will contribute to satisfy the growing demand for stone materials in construction industry.

Thermal power plants are generating a huge amount of fly ash and slag waste, which, if not treated and reused, cause serious environmental pollution. The burial of waste, such as fly ash, from thermal power plants causes long-term harm to the environment (polluting groundwater sources, damaging cultivated lands, etc.). The reuse of this type of waste is being widely supported and, in fact, fly ash has been used in many applications of construction material production such as the production of unburnt bricks. The problem of recycling concrete from demolition works has been studied extensively. Often, the structure of recycled concrete is similar to conventional concrete and even better in some cases.

There have been many scientific publications about recycled concrete and its uses [1-8]. The most recent research

aspects can be seen in [9], in which the authors investigate the effect of the strength of recycled concrete waste on the mechanical properties of recycled concrete. The results show that the reused concrete waste has almost no significant effect on compressive, flexural, and tensile strength of the recycled concrete. Authors in [10] assess the mechanical properties and environmental impacts of self-compacting concrete using recycled aggregates. Authors in [11] evaluate the use of recycled materials from waste concrete to self-compacted concrete. Although the use of recycled concrete is a much studied issue, according to our understanding, there are limited studies evaluating the use of fly ash to increase the efficiency of recycled concrete using construction solid waste [12-14]. Therefore, in this paper the effect of fly ash on the compressive strength of concrete with recycled aggregates from demolition works is studied.



Fig. 1. Construction waste dumped on the side of the road and into a drainage canal.

II. EXPERIMENTAL SETUP

A. Fly Ash

The used fly ash came from the Duyen Hai Thermal Power Plant (Tra Vinh province – Viet Nam). The fly ash belongs to type F because its total content of oxides $\text{SiO}_2 + \text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3$ is bigger than 70%. The technical parameters of the fly ash are shown in Table I.

B. Sand

The sand used in the experiments has the following mechanical properties: density 2.65g/cm^3 , modulus of magnitude 2.50, and volumetric mass 1660 kg/m^3 . Density is

the mass per unit volume of a material in a completely solid state and volumetric mass is the mass per unit volume of material in its natural state.

TABLE I. TECHNICAL PARAMETERS OF FLY ASH

Nam of indicator	Results (%)	Standard 10302:2014 (%)
Total content of $\text{SiO}_2+\text{Al}_2\text{O}_3+\text{Fe}_2\text{O}_3$	82.9	≥ 45
The content of sulfur and sulfur compounds converted to SO_3	0.03	≤ 6
Free calcium oxide content CaO	0.02	≤ 4
The loss content when heated	8.3	≤ 15
Harmful alkali content (soluble alkali)	< 0.01	≤ 0.1

C. Cement

Nghi Son Portland cement PCB40 was used with the following mechanical properties: actual strength 40Mpa and density 3.1g/cm^3 .

D. Coarse Aggregates

1) Natural Coarse Aggregates

Natural coarse aggregates were mixed with the recycled materials. In this study, aggregates of the most common size used in construction (20mm) have been used. The aggregates have the following mechanical properties: 2.61g/cm^3 density and 1430kg/m^3 volumetric mass.

2) Recycled Aggregates

The obtained concrete waste after the demolition of a project was crushed. The resulting coarse aggregates were washed, dried, and pre-screened to remove dust particles. Raw aggregates were screened to classify the particle sizes, then mixed again to be graded as standard for concrete in accordance with the provisions of Vietnamese standard No. 11969: 2018. The recycled coarse aggregate composition according to sieve analysis can be seen in Table II. To avoid strongly absorbing recycled aggregates, affecting the setting of recycled concrete, the coarse aggregates before being mixed were soaked in water in order to absorb it. These aggregates were then let to dry under normal conditions.



Fig. 2. Recycled coarse aggregates.

E. The Recycled Concrete

It is difficult to design high-grade recycled concrete. Therefore, in this study, the designed recycled concrete has design compressive strength of $R_n=15\text{Mpa}$, equivalent to grade

150 concrete. The design slump is 7cm. The composition of the mix of 1 batch of 1m^3 of concrete is shown in Table III.

TABLE II. GRANULAR COMPOSITION OF THE RECYCLED COARSE AGGREGATES

Sieve size	Quantity of sample per sieve (g)	Separate amount of residue on the sieve (%)	Amount accumulated on the sieve (%)	The amount accumulated on the sieve according to the standard
40mm	0	0	0	0
20mm	780	7.8	7.8	0-10
10mm	4930	49.3	57.1	40-70
5mm	3740	37.4	94.5	90-100
>5mm	550	-	-	-

TABLE III. CONCRETE COMPOSITION

Mixing compositions (1m^3)			
Cement (kg)	Sand (kg)	Coarse aggregates (kg)	Water (kg)
243.8	674.9	1231.1	195.0
Cement (kg)	Sand (m^3)	Coarse aggregates (m^3)	Water (lt)
243.8	0.407	0.861	195.0

In this study, fly ash was added to the concrete composition. To find out the effect of fly ash on the compressive strength of recycled concrete, the aggregate composition of concrete was fixed. The coarse aggregates consisted of natural and recycled aggregates in 50%-50% ratio. The amount of fly ash for each sample was:

- Sample 1 (CP1): Fly ash was not used (0%).
- - Sample 2 (CP2): Fly ash with a 5% content was used.

The Laryee UH5130 (6~300kN) universal hydraulic compressor was used to measure the compressive strength of the concrete samples (Figure 3).



Fig. 3. The Laryee UH5130 universal hydraulic compressor.

The compressive strength of concrete was tested according to the Vietnamese standard No. 3118:1993. Cylindrical samples with size of $D150 \times H300$ were cast and cured in water before being compressed to determine their compressive strength. The loading speed was set to 0.5kN/s . The compressive strength after curing of the concrete samples was investigated after periods of 7, 14, and 28 days. For obtaining accurate results, the minimum number of test samples was 3 for each type (Table III).

TABLE IV. SAMPLES USED IN THE COMPRESSION TEST

No.	Description	7 days	14 days	28 days
1	CP1: 50% recycled aggregates, 0% fly ash	3 samples	3 samples	9 samples
	CP2: 50% recycled aggregates, 5% fly ash	3 samples	3 samples	9 samples

III. RESULTS AND DISCUSSION

Table V shows the critical load values recorded in the compressive test. The compressive strengths were then calculated and the results are shown in Table VI. In order to evaluate the effect of fly ash on the compressive strength of recycled concrete, the average value, shown in Table VII, will be used.

TABLE V. CRITICAL LOAD RECORDED IN THE COMPRESSION TESTS

No.	Critical load (kN)				
	7 days	14 days	28 dasy (9 samples)		
CP1	174.77	195.80	212.23	227.96	231.67
	177.07	204.46	222.13	220.54	250.05
	173.18	211.70	230.61	213.47	259.77
CP2	191.74	229.55	248.28	251.11	259.24
	195.27	230.61	249.70	271.08	270.02
	195.80	227.25	245.99	239.45	244.22

TABLE VI. CALCULATED COMPRESSIVE STRENGTH

No.	Compressive strength (Mpa)				
	7 days	14 days	28 days (9 samples)		
CP1	9.89	11.08	12.01	12.9	13.11
	10.02	11.57	12.57	12.48	14.15
	9.80	11.98	13.05	12.08	14.7
CP2	10.85	12.99	14.05	14.21	14.67
	11.05	13.05	14.13	15.34	15.28
	11.08	12.86	13.92	13.55	13.82

TABLE VII. AVERAGE COMPRESSIVE STRENGTH

No.	Description	Compressive strength (Mpa)		
		7 days	14 days	28 days (9 samples)
CP1	50% recycled aggregates, 0% fly ash	9.90	11.54	13.00
CP2	50% recycled aggregates, 5% fly ash	10.99	12.97	14.33

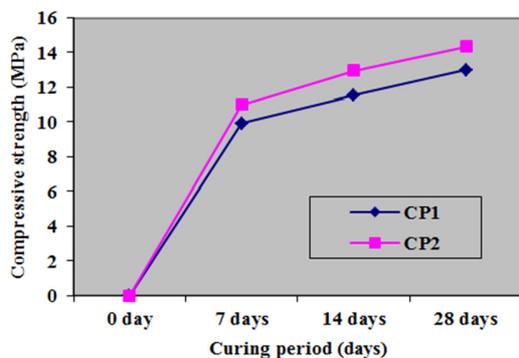


Fig. 4. Strength development of the recycled concrete.

Figure 4 shows the strength development of the experimental concrete samples of CP1 and CP2. The results show that in the first 7 days, the speed of development was very fast and gradually decreased for 7-28 days. At 7 days of age, concrete reaches 76% (CP1) and 77% R28 (CP2) respectively. Thus, it can be seen that the intensity development of the two experimental types of aggregates is relatively uniform.

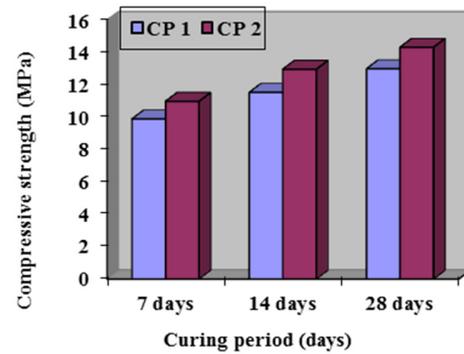


Fig. 5. Comparison of the strength of the two experimental recycled concrete mixes.

Figure 5 shows the result of the comparison of the compressive strength of the two types of the tested graded concrete samples. It can be easily seen that the strength of recycled concrete using 5% fly ash (CP2) is always higher than the compressive strength of recycled concrete without fly ash (CP1). This increase was 11% at 7 days old, 12% at 14 days old and 10.2% at 28 days old. Thus, it can be seen that the use of fly ash increases the strength of recycled concrete by at least 11%. To evaluate the compressive strength of recycled concrete, the compressive strength of recycled concrete using fly ash 5% (CP2) is compared with the compressive strength of conventional concrete by design ($R_n=15\text{Mpa}$). The result is shown in Figure 6.

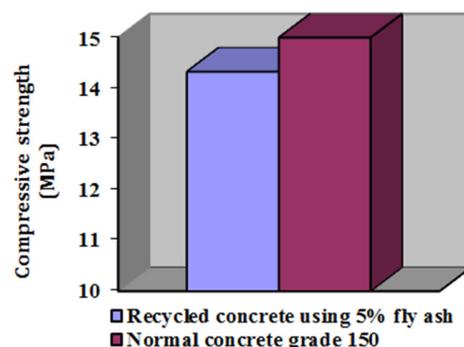


Fig. 6. Comparison of the strength of recycled concrete using fly ash with the conventional concrete.

It can be seen that when using recycled aggregates in a 50% ratio and 5% fly ash, the difference between the compressive strength between the two is not much. According to the experimental results, the compressive strength of recycled

concrete using fly ash is decreased by about 7% when compared to conventional concrete, according to the M150 design. This result is quite consistent with the results of [15], in which the authors conclude with:

- Fly ash content of 5% is the most optimal, giving the best compressive strength results among the experimental recycled concrete mixes.
- The difference in compressive strength between conventional and recycled concrete using 5% fly ash is 11%.

Authors in [12] investigated the use of 30% fly ash as a partial substitute of Portland cement for fly ash concrete production. They found that while embedding high amounts of recycled coarse aggregates could lower the resistance to chloride penetration, the carbonation of the concrete was still comparable and design strength similar to that of the control mix could still be achieved. Combined with the above analysis results, it can be concluded that it is possible to use fly ash at a higher concentration to ensure the compressive strength of recycled concrete.

IV. CONCLUSIONS

In this study, the compressive strength of two recycled concrete aggregates, with and without the utilization of fly ash, was studied. Recycled concrete was designed with compressive strength of $R_n=15\text{Mpa}$, equivalent to M150. The ratio between conventional and recycled coarse aggregates from waste concrete was 50-50. Fifteen samples for each recycled concrete aggregate had been casted, 3 were tested at the age of 7 days, 3 at 14 days, and 9 at 28 days. The drawn conclusions from the experimental results are:

- The strength development of these two recycled concrete mixes is relatively uniform. During the first 7 days, the speed of strength development was very fast. In the period of 7-28 days, the speed gradually decreases. At 7 days of age, concrete reaches 76% R28 (with recycled concrete using no fly ash) and 77% R28 (with recycled concrete using 5% fly ash).
- The strength of recycled concrete using 5% fly ash (CP2) is always higher than that of recycled concrete without the use of fly ash (CP1). This recorded increase was 11% at 7 days, 12% at 14 days, and 10.2% at 28 days.
- The compressive strength of recycled concrete using 5% fly ash is about 7% less than the compressive strength of conventional M150 concrete. Thus, it is possible to use fly ash at higher concentrations to ensure the compressive strength of the recycled concrete.

REFERENCES

- [1] M. F. Koondhar, B. A. Memon, M. Oad, F. A. Chandio, and S. A. Chandio, "Water Penetration of Concrete made with Coarse Aggregates from Demolishing Waste," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6445–6449, Dec. 2020, <https://doi.org/10.48084/etasr.3891>.
- [2] M. Oad, A. H. Buller, B. A. Memon, N. A. Memon, and S. Sohu, "Flexural Stress-Strain Behavior of RC Beams made with Partial Replacement of Coarse Aggregates with Coarse Aggregates from Old Concrete: Part-2: Rich Mix," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3338–3343, Oct. 2018, <https://doi.org/10.48084/etasr.2129>.
- [3] M. Oad, A. H. Buller, B. A. Memon, and N. A. Memon, "Impact of Long-Term Loading on Reinforced Concrete Beams Made with Partial Replacement of Coarse Aggregates with Recycled Aggregates from Old Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3818–3821, Feb. 2019, <https://doi.org/10.48084/etasr.2498>.
- [4] C. S. Poon, S. C. Kou, and L. Lam, "Use of recycled aggregates in molded concrete bricks and blocks," *Construction and Building Materials*, vol. 16, no. 5, pp. 281–289, Jul. 2002, [https://doi.org/10.1016/S0950-0618\(02\)00019-3](https://doi.org/10.1016/S0950-0618(02)00019-3).
- [5] *Quality protocol: Aggregates from inert waste*. Northern Ireland Environment Agency, 2013.
- [6] M. Soomro, "Use of Selective Recycled Materials in Civil Engineering Construction," 2013, pp. 635–684, https://doi.org/10.1142/9789814449175_0010.
- [7] K. Obla, H. Kim, and C. Lobo, "Crushed Returned Concrete as Aggregates for New Concrete," RMC Research & Education Foundation, 01649378, Sep. 2007.
- [8] *Erläuterungen zu den Normen DIN EN 206-1, DIN 1045-2, DIN 1045-3, DIN 1045-4 und DIN EN 12620: DAfStb-Heft 526*, 2nd ed. Berlin, Germany: Beuth Verlag, 2011.
- [9] J. Ahmad Bhat, "Effect of strength of parent concrete on the mechanical properties of recycled aggregate concrete," *Materials Today: Proceedings*, vol. 42, pp. 1462–1469, Jan. 2021, <https://doi.org/10.1016/j.matpr.2021.01.310>.
- [10] J. Li, J. Zhang, S. Ni, L. Liu, and L. F. Walubita, "Mechanical performance and environmental impacts of self-compacting concrete with recycled demolished concrete blocks," *Journal of Cleaner Production*, vol. 293, Apr. 2021, Art. no. 126129, <https://doi.org/10.1016/j.jclepro.2021.126129>.
- [11] C. Sun, Q. Chen, J. Xiao, and W. Liu, "Utilization of waste concrete recycling materials in self-compacting concrete," *Resources, Conservation and Recycling*, vol. 161, Oct. 2020, Art. no. 104930, <https://doi.org/10.1016/j.resconrec.2020.104930>.
- [12] M. Limbachiya, M. S. Meddah, and Y. Ouchagour, "Use of recycled concrete aggregate in fly-ash concrete," *Construction and Building Materials*, vol. 27, no. 1, pp. 439–449, Feb. 2012, <https://doi.org/10.1016/j.conbuildmat.2011.07.023>.
- [13] T. C. Hansen, "Recycled concrete aggregate and fly ash produce concrete without portland cement," *Cement and Concrete Research*, vol. 20, no. 3, pp. 355–356, May 1990, [https://doi.org/10.1016/0008-8846\(90\)90024-R](https://doi.org/10.1016/0008-8846(90)90024-R).
- [14] S. C. Kou, C. S. Poon, and D. Chan, "Influence of Fly Ash as Cement Replacement on the Properties of Recycled Aggregate Concrete," *Journal of Materials in Civil Engineering*, vol. 19, no. 9, pp. 709–717, Sep. 2007, [https://doi.org/10.1061/\(ASCE\)0899-1561\(2007\)19:9\(709\)](https://doi.org/10.1061/(ASCE)0899-1561(2007)19:9(709)).
- [15] 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>.

Investigating the Quality of Milk using Spectrometry Technique and Scattering Theory

Nur Ain Insyirah Muhamad Kamil
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Zatunnur Syakirah Nor'aini
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Wan Zakiah Wan Ismail
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Sharma Rao Balakrishnan
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Juliza Jamaludin
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Irneza Ismail
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Musab Sahrim
Advanced Devices and System (ADS)
Faculty of Engineering and Built Environment
Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia

Abstract-Milk is a dairy product that contains dissolved proteins, carbohydrates, fat, and many minerals. Milk enhances body growth and provides vital energy and fatty acids. Milk can turn bad after being kept at room temperature for several days. The endurance of milk could depend on its fat and protein composition. Our work aims to compare the quality of milk after being kept at room temperature for several days using spectroscopy methods. Modeling based on scattering theory is also provided to compare the light propagation in milk, water, and air. A VIS-NIR spectrometer was used to observe the light absorption, transmission, and reflectance whereas a modeling approach was applied to study the scattering, absorption, and extinction efficiencies. The milk samples consist of full cream milk kept at room temperature for 8 days, 11 days, 14 days, and 17 days. The results show that milk without fermentation has higher light absorbance and lower transmission compared to milk with fermentation, due to changes in milk composition after the fermentation process. Milk scatters more light compared to water and air due to its fat globule and protein ingredients. The output of this study can be used as a reference for studies involving bacteria or microorganisms in milk. It also can be used to compare the quality of milk with and without air exposure.

Keywords-light propagation; absorbance; transmittance; reflectance; scattering; milk; spectroscopy

I. INTRODUCTION

It is crucial to monitor the quality of milk in order to ensure we gain sufficient nutrients and minerals and prevent the occurrence of diseases. Cow milk consists of water (87%), fat (4%), proteins (3.4%), lactose (4.8%), and minerals (0.8%) [1]. Fat composition is not similar in full cream milk and skimmed milk. A layer of cream forms on the milk's surface if it is exposed for several days. This cream consists of spheres of various sizes floating in milk surrounded by a fat globule membrane. The membrane is responsible for fat protection against enzymes and prevents any globule coalescing into butter grains [1]. The spectroscopy technique can be used to observe the optical properties of milk based on light absorbance, transmission and scattering. Mie scattering theory is used to compute the absorption coefficient (μ_a), the scattering coefficient (μ_s), and the phase function $p(\theta)$, where θ is the scattering angle [2]. Mie theory is used to calculate the spectral dependence for the extinction cross section of nanoparticle suspensions [3]. The pump source energy passes

Corresponding author: Wan Zakiah Wan Ismail (drwanzakiah@usim.edu.my)

through the turbid media depending on optical properties such as the refractive index, scattering, anisotropic factor, and laser light absorption [4]. The optical properties of milk based on backscattering intensity can be used to study fat and protein concentrations [5]. The complex fluid of milk is made up of many components such as water, lipids, lactose and protein [6, 8]. Spectroscopy is widely used to measure the optical properties of samples based on light propagation and fluorescence. Color spectroscopy is used to obtain information about the atoms and molecules [8-9]. The absorbance spectroscopy is a technique used to measure the amount of absorbed light [10, 11] with the determination of solution concentration based on Beer's Law [12]. NIR spectrometer and VIS-NIR spectrometer with different wavelength ranges are used to determine the accuracy of the intensity spectrum in the spectroscopy analysis [13].

Many recent studies on light propagation in milk involve backscattering [5], external cavity-quantum cascade laser spectroscopy [6], and laser diffraction and centrifugation [7]. Authors in [8] introduced the simplified NIR spectroscopy in measuring the end of milk fermentation by transforming sugar to lactic acid. The key characteristic of the fermentation process is the pH end point value, in the range of 4.4-4.5 [8]. This technique is quite complicated and costly. To the best of our knowledge, no comparison has been conducted using milk after several days' exposure and water. The previous studies also do not provide a modeling approach on light scattering in milk. Our previous work [14] compared the optical properties of full cream and skimmed milk using different spectrometer types. We found that full cream milk has higher absorption due to its higher fat content. This research is continued in the current paper, which aims to study the light propagation in various milk samples for different exposure at room temperature durations based on spectroscopy techniques using Visible (VIS) and Near Infra-Red (NIR) spectrometers. The technique is simpler and cheaper than the ones used in previous studies as indicated in the experimental section. The output shows that the newly opened milk sample absorbs more light than the other samples. A modeling approach based on Mie theory was also provided to compare light scattering in milk, water, and air.

II. THEORETICAL FRAMEWORK

For the computation of Mie efficiencies, there are two input parameters which are the complex refractive index m and the parameter size x as shown in (1) and (2) [12].

$$m = m' + im'' \quad (1)$$

$$x = ka = \frac{2\pi}{\lambda} a \quad (2)$$

where m' is the real refractive index, im'' is the imaginary refractive index, k is wave number in the ambient medium, and a is the sphere radius.

The key parameters of Mie theory are the computed amplitudes of the scattered field. The coefficients a_n and b_n are required to obtain the Mie efficiency using Spherical Bessel function n ($n=1, 2, \dots$) of higher order and work well in the wider range of size parameters [15].

The efficiency of extinction Q_{ext} and scattering Q_{sca} can be identified in forward-scattering theorem and in the integration of the power scatters in all directions. The absorption efficiency Q_{abs} can be identified with the equation of energy conservation [16]. Meanwhile, the backscattering efficiency Q_b is applicable to monostatic radar [15]. Equations for absorption, scattering and backscattering efficiency are:

$$Q_{ext} = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \text{Re}(a_n^2 + b_n^2) \quad (3)$$

$$Q_{sca} = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) (|a_n|^2 + |b_n|^2) \quad (4)$$

$$Q_{ext} = Q_{sca} + Q_{abs} \quad (5)$$

$$Q_b = \frac{1}{x^2} \left| \sum_{n=1}^{\infty} (2n+1) (-1)^n (a_n - b_n) \right|^2 \quad (6)$$

where x is the parameter size and n is the spherical Bessel function order n .

The efficiency of radiation pressure can be proven by the Two-Stream Model and correlates with the asymmetry parameter [17].

$$Q_{pr} = Q_{ext} + Q_{sca}(\cos\theta) \quad (7)$$

where θ is the scattering angle.

Amplitude functions S_1 and S_2 indicate the scattering properties or the scattering of an electromagnetic wave from a spherical particle. The scattering function is required for the far field scatterer [16]:

$$S_1(\cos\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} (a_n \pi_n + b_n \tau_n) \quad (8)$$

$$S_2(\cos\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} (a_n \tau_n + b_n \pi_n) \quad (9)$$

where

$$\pi_n = \frac{2n-1}{n-1} \cos\theta \cdot \pi_{n-1} - \frac{n}{n-1} \pi_{n-2}$$

$$\tau_n = n \cos\theta \cdot \pi_n - (n+1) \pi_{n-1}$$

III. METHODOLOGY

The research is conducted using experimental and theoretical methods. The light absorption and scattering analysis in milk are based on Mie scattering theory. The scattering, absorption, extinction, and backscattering efficiencies are analyzed in a homogeneous dielectric sphere and its angular scattering using MATLAB. The analysis is also repeated for water and air.

A. Modeling Approach

The modelling part is used to determine the characteristics of light in a disordered medium using MATLAB. The light propagation efficiency with the justification of Mie coefficient matrix is computed. The angular functions are also computed to produce the Mie angular efficiency. Figure 1 shows the flowchart of the constructed modeling approach.

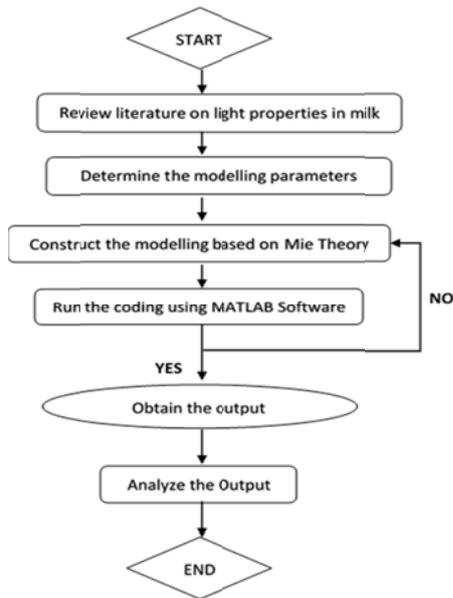


Fig. 1. The flowchart of the modeling approach.

B. Experimental Approach

The Ocean Optic Flame NIR spectrometer and VIS-NIR spectrometer were used to observe the characteristics of light propagation in milk. Every experiment was repeated 10 times to ensure the accuracy of the output. Five samples of milk with different days of exposure were used (the sample turns to yogurt after 14 days of air exposure). Figure 2 shows the flowchart of the experimental method.

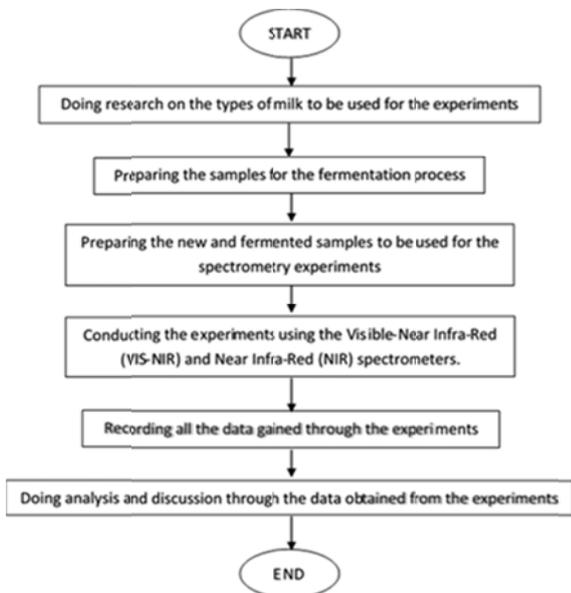


Fig. 2. The flowchart of the experiment.

The milk samples were kept at room temperature for 8, 11, 14, and 17 days. The samples were diluted with a ratio of 1:100 of milk and water respectively. The prepared samples were placed in a cuvette of 10mm and shaken lightly. The cuvette

was placed into a cuvette holder and was illuminated by a halogen lamp (the light source). The light from the cuvette was collected by the VIS-NIR spectrometer (Figure 3). The spectrum graph was displayed on a computer screen using the OceanView software. Figure 4 shows the milk samples used in the experiments. Milk samples of 8 and 11 days were still cloudy but milk samples after 14 days became transparent with large milk particles observed.

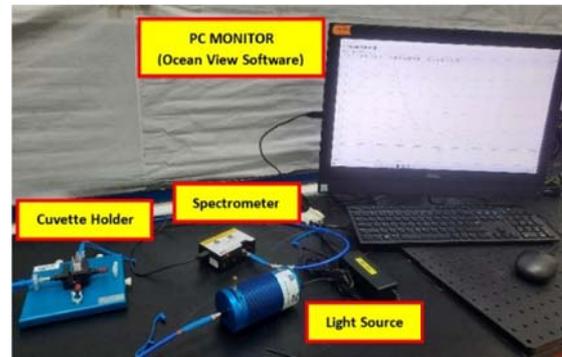


Fig. 3. The experimental set-up for the spectrometry technique.

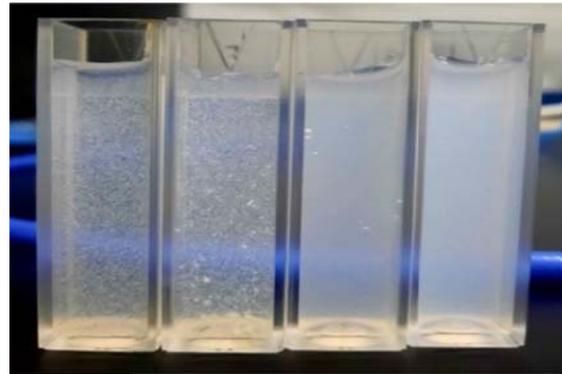


Fig. 4. Fermented full cream milk samples after dilution with water after 8, 11, 14, and 17 days (from right to left) of exposure.

The set-up for reflectance spectroscopy is different from the absorbance and transmission spectroscopy setup. It can be used in liquid samples and solids. NIR spectrometer was used to observe the reflectance intensity spectrum of the samples. The samples were placed on the stage RTL-T. The reflectance spectrum of a sample in the cuvette was observed by a fiber probe. The probe transmits the light from the halogen light and then it was placed on the stage RTL-T. The reflectance spectrum can be observed by using a NIR spectrometer when the light propagates inside the medium.

IV. RESULTS AND DISCUSSION

Previous works [5, 6] have analyzed fat and protein concentrations of milk through backscattering and spectroscopy. Authors in [7] studied the homogenization efficiency to control the development of the cream layer on fresh milk and authors in [8] used NIR spectroscopy to study yogurt fermentation automation. In this paper, we study the quality of milk which is kept at room temperature using VIS and NIR spectroscopy. We do a simple modeling to compare

the light propagation in milk, water and air. The modeling analysis uses Mie theory to compute the efficiency of scattering, absorption, extinction, backscattering, asymmetry parameter, and radiation pressure whereas the experimental section shows the output in terms of absorbance, transmission, and reflectance. The output from the theoretical and experimental study are analyzed and discussed thoroughly in this section.

A. Modeling based on Mie Scattering Theory

The measurements of scattering, extinction and absorption efficiency based on Mie theory were conducted in MATLAB. The input parameters were the complex refractive index and the parameter size x [18]. Modeling was done for milk, water, and air. Figures 5 and 6 summarize the modeling results.

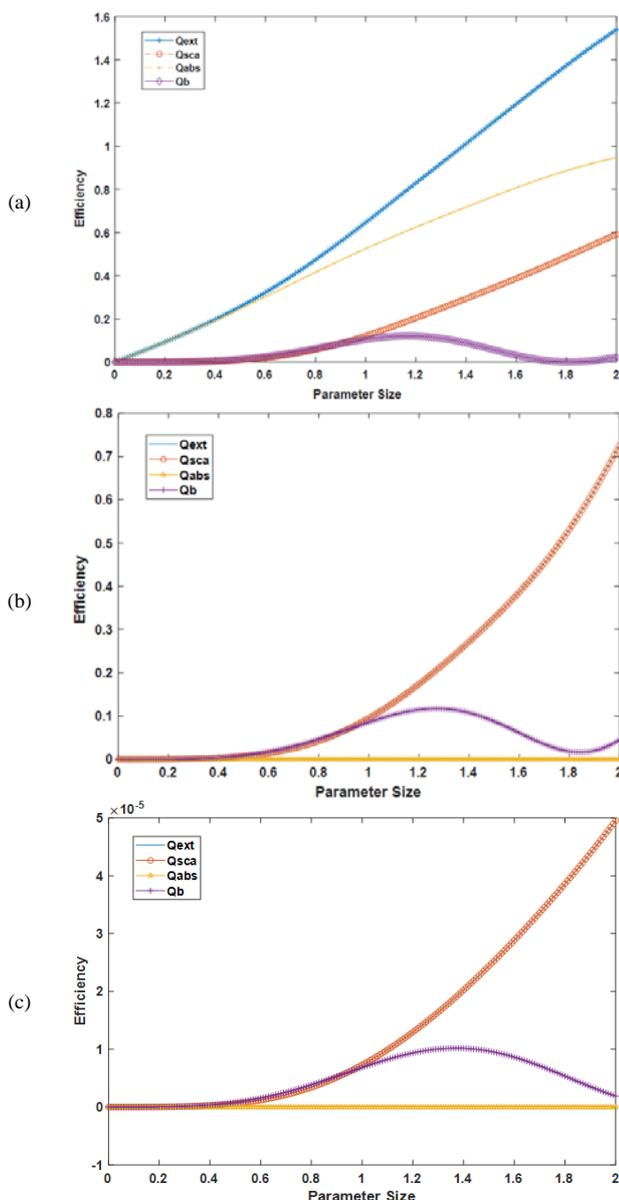


Fig. 5. Mie theory based efficiencies for (a) milk, (b)water, and (c) air.

Equations (3) to (7) were used in Figure 5. Figure 6 is plotted based on (8)-(9). The extinction, forward scattering, absorption, and backscattering efficiencies are represented by Q_{ext} , Q_{sca} , Q_{abs} and Q_b respectively. Figure 5 shows clearly that milk has better scattering efficiency than water and air. At parameter size 2, the scattering efficiency in milk (Figure 5(a)) reaches 0.6, while the scattering efficiency in water (Figure 5(b)) and air (Figure 5(c)) are 0.7 and 5×10^{-7} respectively, proving that the least light scattering occurs in the air, whereas milk and water consist of particles which can scatter the light. We presume that the light scattering and absorption are affected by the size and concentration of the particles, the incident light wavelength, and sample size [19]. Milk depicts the highest efficiency of light absorption due to its composition of fat globules and proteins. Figure 5 also shows that the forward scattering is more efficient compared to the backscattering for all samples due to the larger particles size of the samples. Figure 6 shows the scattering angle of milk, water and air respectively.

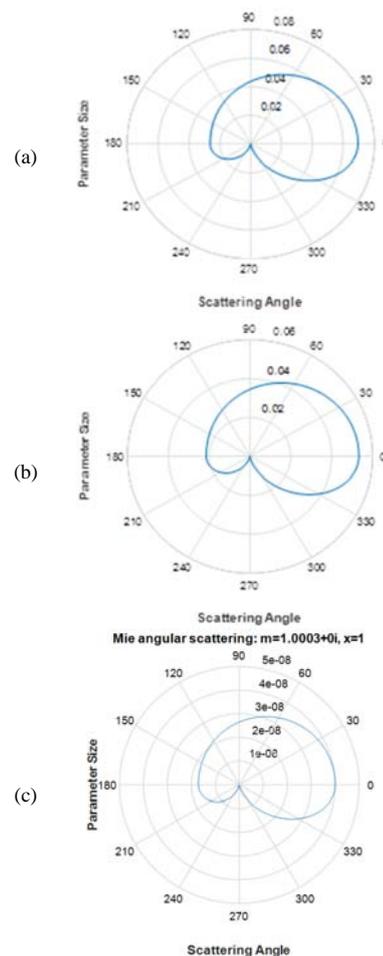


Fig. 6. Angular scattering for (a) milk, (b) water, and (c) air.

We observe that milk has larger value of angular scattering than water and air. It is clearly shown that milk has higher scattering effect. We attribute that to the milk contents which mostly consist of fat and proteins which can scatter light [20].

B. Experimental Results

The spectra of reflectance, transmission, and absorbance of the samples were monitored with the spectrometer. The samples of fresh full-cream milk were kept at room temperature for 8, 11, 14, and 17 days. The experiment utilized an NIR Spectrometer and a VIS-NIR Spectrometer with wavelength ranges of 950nm-1650nm and 350nm-1000nm respectively. Figure 7 depicts the absorbance, transmission and reflectance spectra of the milk samples.

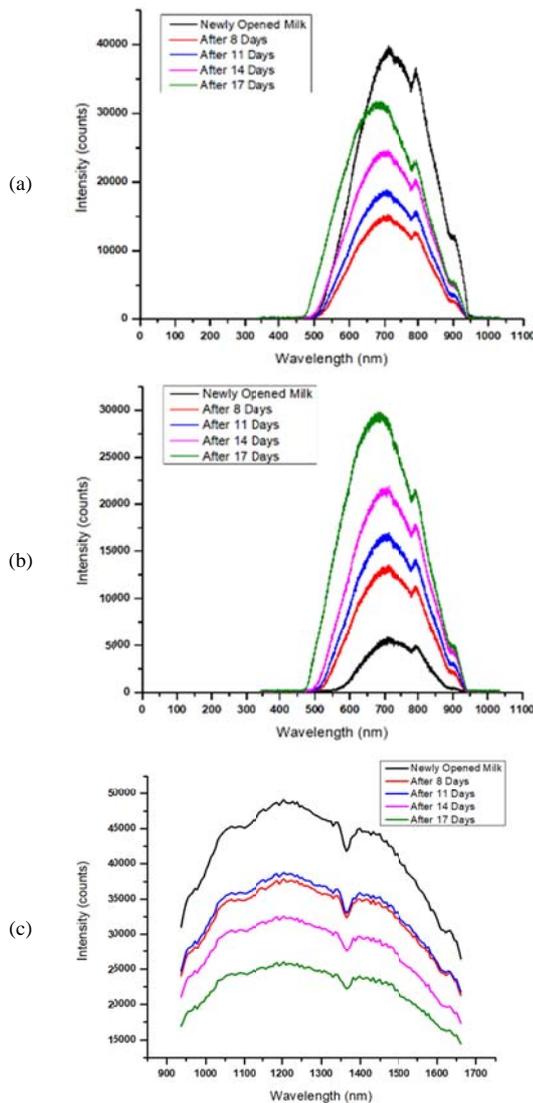


Fig. 7. (a) Absorbance, (b) transmission, and (c) reflectance of the milk samples for various days of fermentation.

Figure 7(a) shows the absorbance spectrum of the milk samples for different days of exposure. It is clearly shown that the absorbance spectra in freshly opened full-cream milk is higher than the milk that has been kept at room temperature for 17 days. The absorbance of milk reaches its peak at 700nm due to high attenuation coefficient which quickly absorbs the light

around 700nm. When milk samples are kept at room temperature, they undergo physical changes where the particles aggregate and create lumpy which affect the light absorption [21] (see Figure 4). Hence, the quality from fresh to fermented milk is gradually decreasing as the protein concentration decreases due to milk coagulation [22].

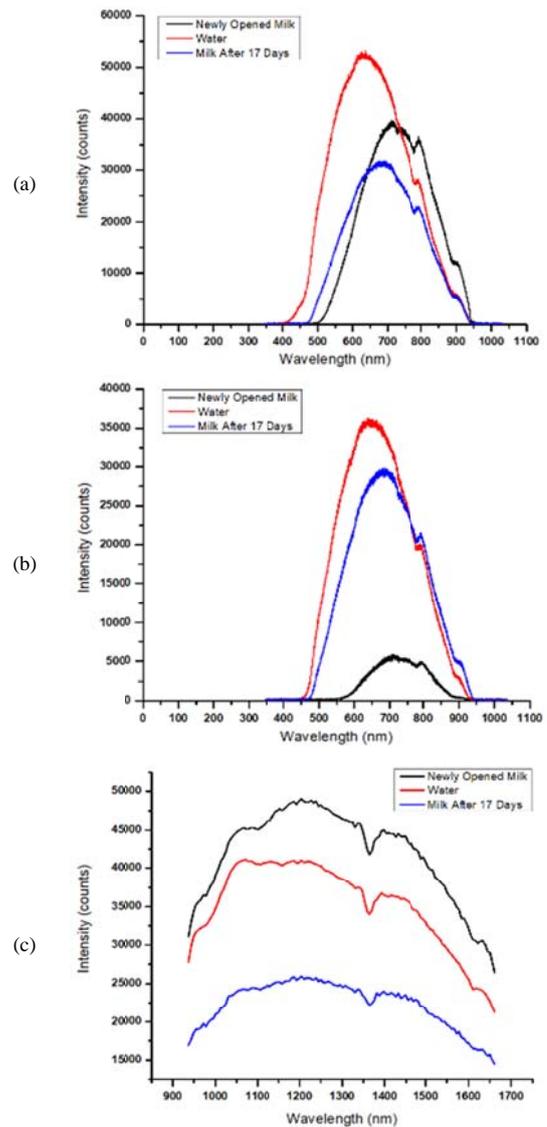


Fig. 8. (a) Absorbance, (b) transmission, and (c) reflectance of newly opened milk sample, water, and milk sample after 17 days.

Figure 7(b) shows the transmission spectra of freshly opened and fermented milks. The newly opened full-cream milk has lower light transmission than the fermented milk. The light transmission in fermented milk after 17 days is higher than in the rest of the samples as the particles in milk aggregate during the coagulation [21, 22]. The protein thickens due to the fermentation process which produces lactic acid [23]. The fermented milk experiences decay and the pH level decreases as the lactic acid bacteria grow to produce the acid [24].

Figure 7(c) shows the reflectance spectra of the milk samples. The reflectance of newly opened milk is higher than the fermented milk's. The low reflectance values for fermented milk are recorded due to its high water absorption [25]. Hence, we suppose that the presence of fat globules and protein micelles in milk affect the light reflectance. The reflectance intensity decreases over the fermentation process due to the changes of protein and fat globules [26].

Figure 8 shows the spectra comparison of milk samples and water. Figure 8(a) shows that water absorbs most of the light at 600nm whereas the absorption peak of milk is at 700nm. Milk and water depict similar transmission peaks at ~ 650nm (Figure 8(b)). Water sample shows higher transmission spectrum as it is more transparent than milk. Newly opened milk samples have higher reflectance than water due to their fat and protein composition. The size and shape of particles, the composition, and the concentration of the tested samples can affect the absorption, transmission, and reflectance of the samples respectively [26]. Newly opened milk samples consist of various particle compositions whereas the fermented milk samples have experienced physical state changes.

V. CONCLUSION

In conclusion, this research investigates the quality of milk for samples freshly opened and after being kept at room temperature for several days using spectroscopy and scattering theory. The optical properties of milk samples were investigated using VIS and NIR spectrometers. Newly opened milk samples have higher light absorbance and lower light transmission compared to the fermented milk, due to the aggregation of the fat and protein particles in milk during the fermentation process. Besides that, modeling based on scattering theory was done to compare the light propagation in milk, water, and air. The modeling shows that milk scatters more light compared to water and air due to the presence of fat globule, protein, and minerals. The outcome of the study shows that the quality of milk is reduced when it is kept at room temperature for several days. This is proved by both naked eye observation and spectroscopy. The outcome of this study can be useful in supporting future analysis studies on dairy products.

ACKNOWLEDGEMENT

We acknowledge the support of the Ministry of Education of Malaysia under the FRGS grant (FRGS/1/2018/STG02/USIM/02/2) and Universiti Sains Islam Malaysia (USIM) for the funding and support.

REFERENCES

- [1] G. H. Schmidt, L. D. V. Vleck, and M. F. Hutjens, *Principles of Dairy Science*, Subsequent edition. Englewood Cliffs, NJ, USA: Prentice Hall, 1988.
- [2] A. Kienle, M. S. Patterson, L. Ott, and R. Steiner, "Determination of the scattering coefficient and the anisotropy factor from laser Doppler spectra of liquids including blood," *Applied Optics*, vol. 35, no. 19, pp. 3404–3412, Jul. 1996, <https://doi.org/10.1364/AO.35.003404>.
- [3] A. A. Ramanenka, A. A. Lizunova, A. K. Mazharenko, M. F. Kerechanina, V. V. Ivanov, and S. V. Gaponenko, "Preparation and Optical Properties of Isopropanol Suspensions of Aluminum Nanoparticles," *Journal of Applied Spectroscopy*, vol. 87, no. 4, pp. 662–667, Sep. 2020, <https://doi.org/10.1007/s10812-020-01051-w>.
- [4] M. A. Ansari, M. Erfanzadeh, and E. Mohajerani, "Mechanisms of Laser-Tissue Interaction: II. Tissue Thermal Properties," *Journal of Lasers in Medical Sciences*, vol. 4, no. 3, pp. 99–106, 2013.
- [5] T. Katsumata, H. Aizawa, S. Komuro, S. Ito, and T. Matsumoto, "Quantitative analysis of fat and protein concentrations of milk based on fibre-optic evaluation of back scattering intensity," *International Dairy Journal*, vol. 109, Oct. 2020, Art. no. 104743, <https://doi.org/10.1016/j.idairyj.2020.104743>.
- [6] M. Montemurro, A. Schwaighofer, A. Schmidt, M. J. Culzoni, H. K. Mayer, and B. Lendl, "High-throughput quantitation of bovine milk proteins and discrimination of commercial milk types by external cavity-quantum cascade laser spectroscopy and chemometrics," *Analyst*, vol. 144, no. 18, pp. 5571–5579, Sep. 2019, <https://doi.org/10.1039/C9AN00746F>.
- [7] E. Ransmark, B. Svensson, I. Svedberg, A. Göransson, and T. Skoglund, "Measurement of homogenisation efficiency of milk by laser diffraction and centrifugation," *International Dairy Journal*, vol. 96, pp. 93–97, Sep. 2019, <https://doi.org/10.1016/j.idairyj.2019.04.011>.
- [8] A. Aljaafreh and H. Steiner, "Evaluation of Using NIR Simplified Spectroscopy in Yogurt Fermentation Automation," presented at the 8th International Conference on Researches in Engineering, Technology and Sciences, Istanbul, Turkey, Aug. 2015.
- [9] M. B. Alshammari, E. H. Anouar, and G. A. El-Hiti, "Spectroscopic Characterization, Hirshfeld Surface, DFT, and TD-DFT of tert-Butyl Phenethylcarbamate and 1,1-Dimethyl-3-Phenethylurea," *Journal of Applied Spectroscopy*, vol. 87, no. 4, pp. 736–744, Sep. 2020, <https://doi.org/10.1007/s10812-020-01063-6>.
- [10] F. Alorifi, S. M. A. Ghaly, M. Y. Shalaby, M. A. Ali, and M. O. Khan, "Analysis and Detection of a Target Gas System Based on TDLAS & LabVIEW," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4196–4199, Jun. 2019, <https://doi.org/10.48084/etasr.2736>.
- [11] M. M. Nadareishvili, G. Mamniashvili, D. Jishiashvili, G. Abramishvili, C. Ramana, and J. Ramsden, "Investigation of the Visible Light-Sensitive ZnO Photocatalytic Thin Films," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5524–5527, Apr. 2020, <https://doi.org/10.48084/etasr.3392>.
- [12] N. A. Bakar, Huize Cui, A. Abu-Siada, and Shengtao Li, "A review of spectroscopy technology applications in transformer condition monitoring," in *2016 International Conference on Condition Monitoring and Diagnosis (CMD)*, Xi'an, China, Sep. 2016, pp. 372–375, <https://doi.org/10.1109/CMD.2016.7757895>.
- [13] R. A. Viscarra Rossel, D. J. J. Walvoort, A. B. McBratney, L. J. Janik, and J. O. Skjemstad, "Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties," *Geoderma*, vol. 131, no. 1, pp. 59–75, Mar. 2006, <https://doi.org/10.1016/j.geoderma.2005.03.007>.
- [14] N. A. I. Muhamad Kamil *et al.*, "Investigating Light Propagation in Full and Skimmed Milk based on Spectroscopy and Monte Carlo Analysis," *Journal of Engineering and Applied Science*, vol. 67, no. 8, pp. 2265–2278, Dec. 2020.
- [15] C. Matzler, "MATLAB functions for Mie scattering and absorption," Institut für Angewandte Physik, Bern, Switzerland, Research Report 2002–08, Jun. 2002.
- [16] K. L. van der Molen, "Experiments on scattering lasers: from Mie to random," Ph.D. dissertation, University of Twente, Enschede, Netherlands, 2007.
- [17] W. E. Meador and W. R. Weaver, "Two-Stream Approximations to Radiative Transfer in Planetary Atmospheres: A Unified Description of Existing Methods and a New Improvement," *Journal of the Atmospheric Sciences*, vol. 37, pp. 630–643, Mar. 1980.
- [18] Y. Ren, H. Qi, X. Yu, and L. Ruan, "A forward-angle-scattering method for the determination of optical constants and particle size distribution by collimated laser irradiation," *Optics Communications*, vol. 389, pp. 258–264, Apr. 2017, <https://doi.org/10.1016/j.optcom.2016.12.060>.
- [19] D. J. Dahm, "Explaining Some Light Scattering Properties of Milk Using Representative Layer Theory," *Journal of Near Infrared Spectroscopy*, vol. 21, no. 5, pp. 323–339, Oct. 2013.

-
- [20] L. C. Andrews, *Field Guide to Atmospheric Optics, Second Edition*, 2nd ed. Washington, DC, USA: SPIE--The International Society for Optical Engineering, 2019.
- [21] C. Hahn, M. Sramek, S. Nöbel, and J. Hinrichs, "Post-processing of concentrated fermented milk: influence of temperature and holding time on the formation of particle clusters," *Dairy Science & Technology*, vol. 92, no. 1, pp. 91–107, Jan. 2012, <https://doi.org/10.1007/s13594-011-0046-1>.
- [22] S. L. Thomsen, S. L. Jacques, and S. T. Flock, "Microscopic correlates of macroscopic optical property changes during thermal coagulation of myocardium," in *Laser-Tissue Interaction*, Jun. 1990, vol. 1202, pp. 2–11, <https://doi.org/10.1117/12.17605>.
- [23] M. Lu *et al.*, "Milk Spoilage: Methods and Practices of Detecting Milk Quality," *Food and Nutrition Sciences*, vol. 4, no. 7, pp. 113–123, Jul. 2013, <https://doi.org/10.4236/fns.2013.47A014>.
- [24] R. Fernandes, Ed., *Microbiology Handbook: Dairy Products*, 3rd ed. Cambridge, UK: Royal Society of Chemistry, 2009.
- [25] B. Aernouts, E. Polshin, J. Lammertyn, and W. Saeys, "Visible and near-infrared spectroscopic analysis of raw milk for cow health monitoring: Reflectance or transmittance?," *Journal of Dairy Science*, vol. 94, no. 11, pp. 5315–5329, Nov. 2011, <https://doi.org/10.3168/jds.2011-4354>.
- [26] J. A. Raty and K.-E. Peiponen, "Reflectance Study of Milk in the UV-Visible Range," *Applied Spectroscopy*, vol. 53, no. 9, pp. 1123–1127, Sep. 1999.

The Effect of Moisture and Fine Grain Content on the Resilient Modulus of Sandy Clay Embankment Roadbed

Nguyen Anh Tuan
Faculty of Transportation Engineering
Ho Chi Minh City University of Transport
Ho Chi Minh City, Vietnam
tuanna@ut.edu.vn

Phan Quang Chieu
Faculty of Civil Engineering
Tien Giang University
Tien Giang, Vietnam
phanquangchieu5@yahoo.com

Abstract—This paper studies the effect of moisture and fine grain content on the resilient modulus of sandy clay embankment roadbed in the Mekong Delta, Vietnam. The study analyzed the grain content of 30 soil samples on the annually flooded routes of the Mekong Delta according to the AASHTO T88-97 standard. The triaxial compression test at room temperature was used to estimate the resilient modulus of the samples belonging to 6 moisture levels. The experiments were conducted using 3 levels of lateral pressure, 0, 21, and 41kPa. Five deflection stress tests of 14, 28, 41, 55, and 69kPa, were conducted for each lateral pressure. The results showed that as the percentage of grains finer than 0.075mm increased, the variation ratio of the resilient modulus also increased. The content of grains finer than 0.075mm was between 54.1%-93%, while the variation ratio of the resilient modulus ranged between 53.7% and 89.1%. Moreover, as the percentage of grains finer than 0.075mm increased, water absorption capacity increased and resilient modulus decreased. As moisture and fine grain content influence the resilient modulus of the roadbed, this study's results will help to limit and prevent the erosion of sandy clay embanked roadbeds, especially on frequently flooded areas such as the Mekong Delta.

Keywords—resilient modulus; moisture content; fine grain content; roadbed; in situ plate loading test

I. INTRODUCTION

The resilient modulus of the roadbed is one of the most important parameters for designing a new or restoring a soft road surface in case of deformation. The thickness of the road surface layer is determined based on the resilient modulus of the roadbed. Moisture content affects the resilient modulus of the roadbed significantly, as it increases the road surface's deformation causing cracks and subsidence, especially in sandy clay embanked roadbeds which are flooded for long periods. The resilient behavior of cohesive soils (fine-grained soils) related to moisture has been studied for over 40 years. The influence of density and water content on the resilient behavior of Florida subgrade soils was studied in [1], while the subgrade resilient modulus was estimated by using standard tests in [2]. The degree of saturation affecting the resilient modulus of Tennessee soils was studied in [3]. An improved evaluation

procedure of roadbed soil's resilient modulus was introduced in [4]. The effect of moisture on the resilient modulus of the Ohio roadbed was investigated in [5]. In [6], a correlation between the relative moisture and the resilient modulus of the roadbed was proposed. The elastic and deformation characteristics of bottom ash in road construction, in particular Young's modulus and Poisson's ratio, were studied in [7]. According to [8], the settlement response of the embankment dam was similar for the Mohr-Coulomb and the Hardening Soil Models for three material zones (clay core, sandy gravel, and random fill), having a modulus of elasticity in the range 25000-50000kPa. These studies showed that the resilient modulus of soil embanked roadbed depends heavily on soil type, moisture content, and soil condition [9].

The current study carried out laboratory experiments to observe the effects of moisture and fine grain content on the resilient modulus of a sandy clay embankment roadbed in the Mekong Delta of Vietnam.

II. MATERIALS AND METHODS

A. Determination of the Resilient Modulus

The resilient modulus is determined based on elastic strain. In road constructions, it is used to calculate the roadbed and road surface subsidence. Due to the transient workload, the loading and unloading time is extremely fast and repeated, while after several loads the accumulated plastic strain is reduced or eliminated. The subsidence of road works depends heavily on the elastic strain of the roadbed and road surface structure. Resilient modulus is defined, according to [10], as:

$$M_r = (\sigma_1 - \sigma_3) / \varepsilon_r = \sigma_d / \varepsilon_r \quad (1)$$

where σ_1 and σ_3 are the major and minor principal stresses, σ_d is the deviator stress, and ε_r is the accumulated plastic strain.

B. Factors Affecting Resilient Modulus

It has been shown that the resilient modulus of cohesive soil depends on soil type, moisture content, saturation, the content of the grains that can go through the No.200 sieve, deflection stress, suction force, plasticity index, pore water

pressure, lateral pressure, and lateral compression strength [11-30].

C. Effect of Moisture Content on the Resilient Modulus

Moisture content is the main parameter affecting the resilient modulus of sandy clay, as in low moisture the water binds soil grains and increases the effective stress between them through its suction and surface tension, leading to high resilient modulus values. Moreover, at low moisture content, the sandy clay produces strong suction in water enough to reproduce a significant temporary colloidal effect between soil grains. Increased moisture damages this phenomenon. The state of sandy clay depends heavily on water capacity, namely the physical bonding layers of water surrounding the coarse soil grains. When the soil is completely dry, the corresponding state is extremely hard or solid, the deformation is negligible, and the resilient modulus of the sandy clay increases due to the attraction between the opposite ions. Moisture content increases gradually with a strong water suction layer, increasing effective stress between soil grains through the suction and the surface tension of water. However, the soil sample volume remains unchanged until the grains are fully deposited in the water suction layer, the soil volume begins to increase due to the thickness of the water shells, the water takes up the voids that push the soil grains apart, the surface suction force decreases, the soil's resilient modulus value decreases, the soil becomes semi-hard and flexible, and when free water appears, the resilient modulus gradually decreases as the soil turns into the liquid state. Too much water causes the soil to enter a suspension state [31]. The resilient modulus value of the sample with the optimum moisture content is much greater than the resilient modulus value of the saturated sample.

D. Effect of Fine-Grained Content on Resilient Modulus

The resilient modulus of a cohesive embankment roadbed depends on the type of cohesive soil. The grain content is also an important factor affecting the resilient modulus of cohesive soils, especially the percentage of the grains that can go through the No. 200 sieve. When moisture is low, water binds soil grains (especially of fine-grained soil) and increases the effective stress between soil grains through the suction and the surface tension of water. In this case, the deformation of the soil is negligible. The clay may crack and become extremely hard when it is dry. Moisture gradually increases until the grains are fully deposited into the water suction layer, the soil volume begins to increase due to the thickness of water shells, the water takes up the hollow holes pushing the soil grains apart, and the soil becomes semi-hard and flexible [31, 32].

E. Laboratory Experiments

1) Purpose

The main purpose of the experiments was to identify the physical characteristics of moisture content, saturation, unconfined compressive strength, plasticity index, liquid limit, optimum moisture, and the content of grains finer than 0.075mm. The rapid compression test utilized a 3-axis compression chamber to determine the resilience of the soil samples and calculate the resilient modulus. A total of 124 rapid compression tests were conducted to determine the resilient modulus, while 30 experiments were conducted to

determine the physical characteristics and the grain content of the soil.

2) Standards and Methods

The soil samples were sandy clay samples with different moisture values. Soil samples were collected from the trunks of roads in annually flooded areas and classified according to AASHTO M 145-91 [33] based on grain composition and Atterberg limit. The resilient modulus was determined according to AASHTO T294-03 [34]. The simulation of vehicle load was performed on soil samples at 6 moisture values (2% and 3% dryer than the optimum, the optimum, 2% and 3% over the optimum, completely saturated). Liquid limit and plasticity index were determined according to AASHTO T89-07 [35] and T90-10 [36], while the grain content was analyzed according to AASHTO T88-04 [37]. The maximum dry density and the optimum moisture were determined according to AASHTO T180-01 [38], while the moisture was determined according to AASHTO T265-04 [39].

3) Sample Collection and Experimentation

Thirty soil samples were collected at 30cm depth in the road body of the annually flooded roads in Dong Thap, Long An, and Tien Giang provinces. The natural moisture content of the samples was 9.9-32.7%. The basic physical characteristics of the samples were determined, including liquid limit, plasticity limit, grain content, standard compaction, and unconfined compressive strength. Table I shows the results of liquid limit, plastic limit, and grain content analysis.

TABLE I. GRAIN CONTENT AND ATTERBERG LIMIT OF SAMPLES

No.	Sample	Liquid limit	Plasticity index	Sand (%)	Dust (%)	Clay (%)
1	LA842.1	33.0	14.8	15.2	48.9	33.9
2	LA842.2	34.3	12.1	12.4	48.9	35.9
3	LA842.3	39.7	14.6	14.0	49.8	35.4
4	DT942.1	39.6	12.0	16.6	45.4	35.1
5	DT942.2	36.0	11.3	17.7	43.0	32.1
6	DT942.3	38.3	11.6	15.3	43.5	39.8
7	DT942.4	39.0	11.9	21.5	38.4	34.8
8	DT942.5	38.8	11.6	10.8	53.4	34.6
9	DT942.6	35.3	11.4	10.7	50.9	31.6
10	DT942.7	39.7	11.5	5.2	54.4	38.6
11	DT942.8	38.2	11.8	15.7	51.9	31.9
12	DT942.9	39.6	11.6	11.3	44.2	39.2
13	DT942.10	39.9	11.7	7.3	55.1	36.2
14	DT847.1	38.3	12.4	14.0	49.4	36.2
15	DT847.2	38.9	11.5	10.6	53.8	35.2
16	DT847.3	38.8	11.3	12.6	43.4	41.4
17	DT847.4	39.9	11.5	16.9	42.4	38.1
18	DT847.5	39.5	16.4	24.8	38.7	33.6
19	DT847.6	39.4	15.3	34.5	30.1	31.6
20	DT847.7	39.4	15.3	16.6	45.2	33.7
21	DT847.8	38.0	13.8	27.1	36.5	33.7
22	DT847.9	38.4	13.2	13.8	50.1	35.1
23	DT847.10	37.3	12.3	19.7	38.2	37.3
24	DT847.11	38.3	12.0	20.5	46.6	32.6
25	DT847.12	39.1	11.6	20.0	47.6	31.6
26	DT847.13	38.0	14.8	33.9	31.2	31.2
27	DT867.1	38.2	13.3	33.9	29.1	33.1
28	DT867.2	39.0	11.6	23.5	33.7	37.3
29	DT867.3	38.8	13.3	42.2	22.9	31.2
30	DT867.4	39.0	12.1	36.8	29.5	32.5

An unconfined compressive strength test, according to AASHTO T208-05 [40], was carried out with a speed of 1mm/min right after compressing the sample to determine the value of the resilient modulus. The unconfined compressive strength was determined from the stress-strain curve. It is the maximum compressive stress value that the sample has to withstand or the value corresponding to 20% strain if this case happens first. Axial compressive stress σ_1 (kPa) was determined by:

$$\sigma_1 = \left[\frac{P(1-\epsilon)}{A_0} \right] \times 100 \quad (2)$$

F. Determining the Resilient Modulus of the Soil

The purpose of this test was to measure the resilient strain of the soil samples under the effect of fast compressive loading and determine the influence level of moisture and grains with a finer than 0.075mm size on the resilient modulus. For each soil sample collected, 15kg were selected through a 5mm sieve, they were divided into 6 equal weight parts, placed into 6 trays, and 6 different water quantities were sprayed on them to obtain the required moisture levels: 2% and 3% drier than the optimum, the optimum, 2% and 3% wetter than the optimum, and one saturated. Wet soil was mixed, covered with a damp cloth, and incubated for 12h. The preparation of a test sample m (g) with the desired moisture value W (%) requires the determination of the moisture content of the sample W_1 (%) and the calculation of the amount of water q (g) to be sprayed according to:

$$q = [0.01m / (1 + 0.01W_1)] \times (W - W_1) \quad (3)$$

A compacting mortar with 125mm diameter and 127mm height was placed on a hard and leveled ground. The prepared soil was taken into the mortar with 3 layers, each one occupying about 1/3 of the mortar’s volume. A 2.5kg hammer was used on free fall from a height of 300mm, evenly distributed over the surface of the soil layer. Each layer was compacted with 40 hammer drops. The compacted soil was removed from the mortar by pressing a cutting ring, having 36mm diameter and 76mm height, vertically into the core of the soil. The cutting ring was removed from the soil by whittling. The removed sample was weighted to determine its natural volume and placed in a rubber wrap. A 3-axis compression device model 28-T0401, shown in Figure 1, was used to perform the tests.



Fig. 1. The triaxial compression device system.

The prepared samples were placed in the laboratory chamber to eliminate the residual strain by applying lateral and vertical pressure. By loading and unloading many times with a strain rate of 1mm/min, the sample accumulated the residual

strain until there was only resilient strain left. The water inlet valve was opened until the chamber was full. Three levels of lateral pressure were applied (41, 21, and 0kPa) with 5 deviator stress levels (14, 28, 41, 55, and 69kPa). During the experiment, the bottom drain valve was unlocked. The results of loading for an unsaturated sample (DT.942.7-2) are shown in Figure 2.

TABLE II. LOADS ON UNSATURATED SAMPLE

No.	Lateral pressure σ_3 (kPa)	σ_d (kPa)	Times of loading (times)	Note
0	0	69	Many times	Residual strain elimination
1	41	14	3	Get the average value
2	41	28	3	
3	41	41	3	
4	41	55	3	
5	41	69	3	
6	21	14	3	
7	21	28	3	
8	21	41	3	
9	21	55	3	
10	21	69	3	
11	0	14	3	
12	0	28	3	
13	0	41	3	
14	0	55	3	
15	0	69	3	
16	0	Destroy the sample		Determine q_u

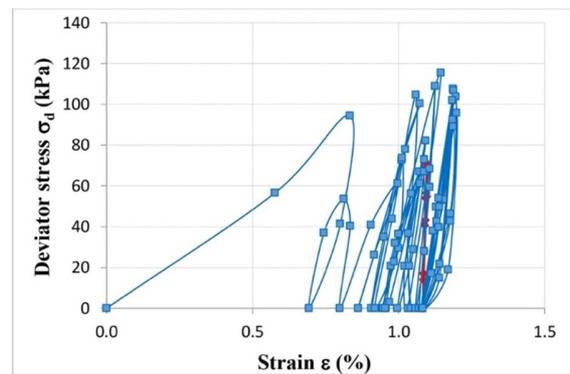


Fig. 2. Results of loading with unsaturated soil samples (DT.942.7-2).

In each deviator stress level test, it was necessary to load and unload 3 times to get the average of the deviator stresses corresponding to the average of the resilient strain. Figure 3 shows the deviator stress-resilient strain relationship of the sample. The average values of the deviator stress and resilient strain can be used to calculate the resilient modulus value according to (4). The resilient strain generated after each load level consists of two resilient strain parts: an instant resilient strain part generated when the load is applied to the soil and another due to the rheological behavior in load time. Slower speeds of loading (on the road when the car is running slower) cause greater resilient strains, as shown in Figure 4. Similar resilient values cause different P loads, resulting in different soil resilient modulus.

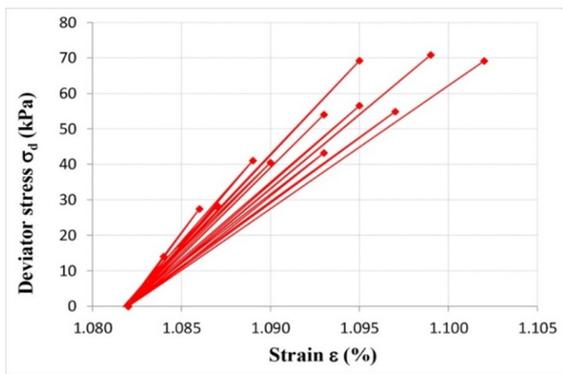


Fig. 3. The relationship of σ_d with the resilient strain σ_3 is unchanged (DT.942.7-2).

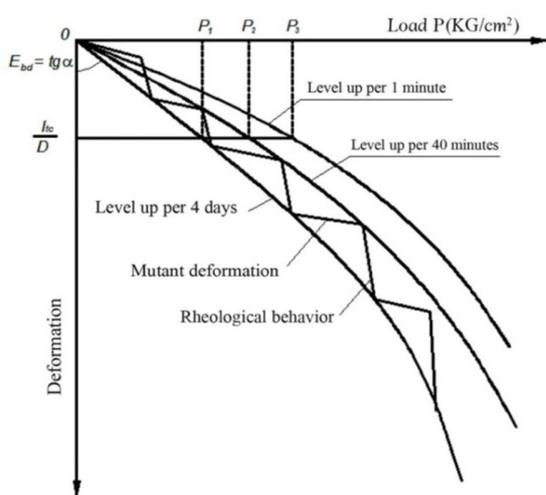


Fig. 4. The resilient strain curve of the compressing test when the load increases with each level at a uniform speed.

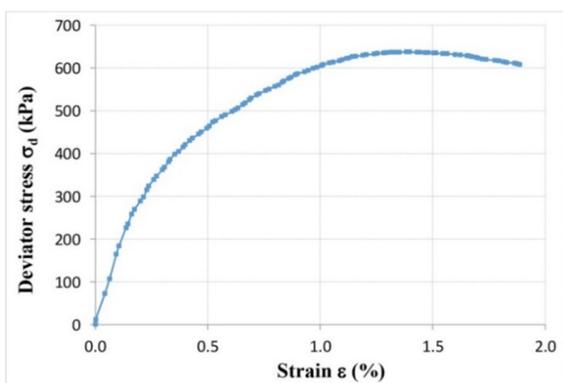


Fig. 5. Determining the unconfined compressive strength (DT.942.7-2).

The unconfined compressive strength test was carried out immediately after the completion of five deviator stress levels, with zero lateral pressure and vertical load at 1mm/min until the sample was damaged to determine the compressive strength q_u . Figure 5 shows a strain graph for determining q_u of 637.2kPa for sample DT942.7-2 having 15.7% moisture content. The unconfined compressive strength of the test samples was between 60.6 and 774.3kPa. The correlation

coefficient between the resilient modulus and the unconfined compressive strength (q_u) was 0.6788 (R^2). For saturated samples, the test was only performed with zero lateral pressure level and five levels of deviator stress: 14, 28, 41, 55, 69kPa. During the experiment, the bottom drain valve was locked. The loading method for saturated samples is shown in Table III, and the results of loading for the saturated samples (DT.942.7-5) are shown in Figure 5.

TABLE III. LOAD ON SATURATED SAMPLES

No.	Lateral pressure σ_3 (kPa)	σ_d (kPa)	Times of loading (times)	Note
1	0	14	3	
2	0	28	3	
3	0	41	3	
4	0	55	3	
5	0	69	3	
6	0	Destroy the sample		Determine q_u

The resilient modulus determined from a 3-axis compression test is the ratio of the deviator stress and the relative resilient strain of the sample:

$$M_r = \sigma_d / \epsilon \quad (4)$$

$$\sigma_d = \sigma_1 - \sigma_3 \quad (5)$$

$$\epsilon = \Delta h / h \quad (6)$$

The resilient strain of soil samples was recorded by the data collection system according to each load level. It was necessary to choose the average resilient strain value of the three load times for each load level to calculate the resilient modulus.

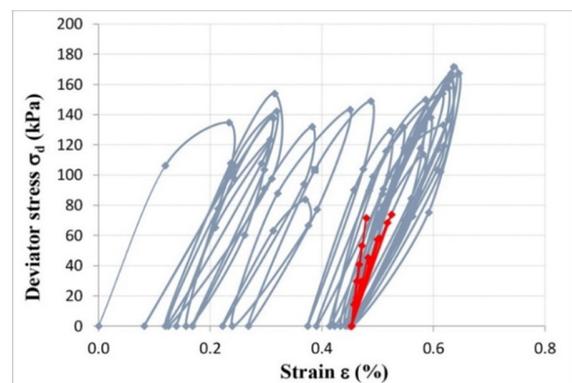


Fig. 6. Load result of unsaturated soil samples (DT.942.7-5).

III. RESULTS AND DISCUSSION

Figures 7-9 show the experimental results for the resilient modulus of the 30 soil samples. In Figure 7, a lateral pressure of 41kPa was applied during the experiment, with 5 samples having moisture content of 14.5%, 15.7%, 17.4%, 19.6%, and 20.3%. Five deviator stress levels were performed to determine the resilient modulus, at 14, 28, 41, 55, and 69 kPa. Figures 8 and 9 show similar experiments by applying lateral pressures of 21 and 0kPa respectively. The results illustrate the effect of moisture content, deviator stress, and lateral pressure on the resilient modulus. The resilient modulus is significantly affected by moisture content, as it decreased as the moisture

content increased at constant lateral pressure level. The resilient modulus decreased as the deviator stress increased, tending to change nonlinearly with the deviator stress at the same lateral pressure level. The resilient modulus increased as the lateral pressure increased at the optimum moisture value and the same level of deviator stress. The resilient modulus of the saturated samples decreased by 24.8-56.6% from the optimum moisture content. The correlation coefficient (R^2) between the resilient modulus and moisture was 0.8849. The R^2 between the resilient modulus and saturation was 0.6741.

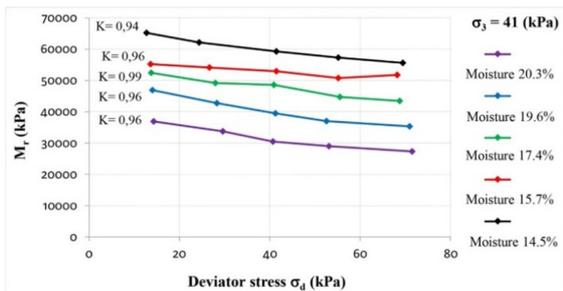


Fig. 7. M_r according to moisture and deviator stress, lateral pressure 41kPa (model DT942.7).

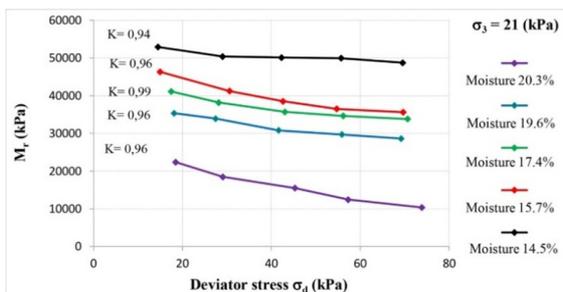


Fig. 8. M_r according to moisture and deviator stress, lateral pressure 21kPa (model DT942.7).

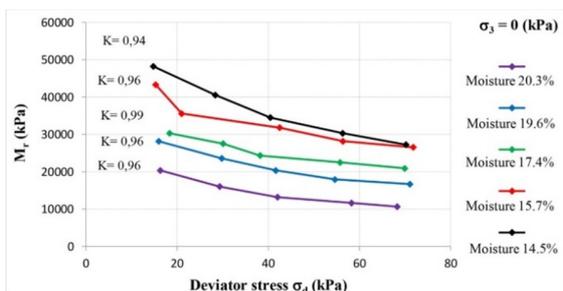


Fig. 9. M_r according to moisture and deviator stress, lateral pressure 0kPa (model DT942.7).

Figure 10 shows the effect of the grain content with a size finer than 0.075mm on the variation rate of resilient modulus due to the change of moisture content. The content of grains with size finer than 0.075mm varied from 54.1% (sample DT.867.4) to 93.0% (sample DT.942.7), while the value of the resilient modulus varied from 53.7% (60,548 - 28,043kPa) to 89.1% (95,533 - 10,426kPa). The more the content of grains having size finer than 0.075mm was, the more it affected the variation ratio of the resilient modulus. The content of grains

having size finer than 0.075mm accounts for 54.1% (sample DT.867.4) to 70.1% (sample DT.847.10) corresponding to the rate of value variation of the resilient modulus from 53.7% (60,548 - 28,043 kPa) to 65.6% (45,531 - 15,681 kPa). The content of grains having size finer than 0.075mm accounts from 72.3% (sample DT.867.6) to 93.0% (sample DT.942.7) corresponding to the variation rate value of the resilient modulus from 65.8% (48,903 - 16,709kPa) to 89.1% (95,533 - 10,426kPa). The R^2 between the resilient modulus and the content of grains having a size finer than 0.075mm was 0.7358. The correlation coefficient between the resilient modulus and plasticity index was 0.5412.

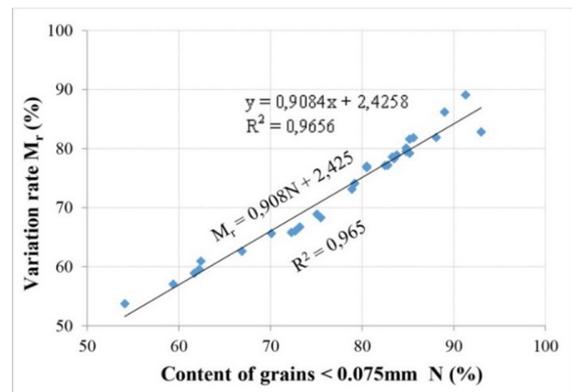


Fig. 10. Effect the content of grains having size finer than 0.075mm on M_r .

The obtained results demonstrate that the resilient modulus is significantly influenced by moisture and the content of grains having size finer than 0.075mm. The increased moisture content of the sample caused a drop in the resilient modulus. In many cases, the resilient modulus value of the saturated samples was reduced by more than 56.6% (35,698 - 15,489kPa) compared to the samples at optimum moisture content. The resilient modulus decreased nonlinearly when the deflection stress increased. Increasing lateral pressure led to an increase in the resilient modulus of the soil sample. The higher the percentage of clay grains, the larger the area they cover, and the greater the resilient modulus of the soil is due to the attraction between opposite ions. The increased moisture corresponds to the thickness of the water shells. The greater the volume of water that occupies the voids, the farther apart the soil grains are, and the lower the surface suction of soil grains decreases. Thus, the resilient modulus of the soil decreases significantly.

IV. CONCLUSIONS

The resilient modulus was significantly affected by the moisture content, as it decreased rapidly as the moisture content increased by 1-2%. The resilient modulus for saturated samples was reduced by 56.6% in comparison to the optimum moisture content (sample DT942.7). The resilient modulus reached the smallest value of 8.153kPa when the sample had a moisture content of 23.3% and the content of grains having size finer than 0.075mm accounted for 88.1% (sample DT942.5). The percentage of grains having size finer than 0.075mm of the

soil sample had an important influence on the value of the resilient modulus. Among the tested samples, those with a higher proportion of grains with size finer than 0.075mm had a greater drop of resilient modulus value with increased moisture, while their resilient modulus reached the smallest values on the maximum moisture. The percentage of grains having size finer than 0.075mm in the samples was between 54.1 and 79.2%. The resilient modulus varied with moisture content from 53.7% to 89.1%. The greater the percentage of grains that are finer than 0.075mm is, the higher the ratio, the larger the area ratio, the greater the water absorption capacity, and the more the resilient modulus decreases.

REFERENCES

- [1] M. K. Elfino and J. L. Davidson, "Modeling Field Moisture in Resilient Moduli Testing," in *Resilient Moduli of Soils: Laboratory Conditions*, David. J. Elton, Ed. New Orleans, Louisiana, USA: ASCE, 1989, pp. 31–51.
- [2] E. C. Drumm, Y. Boateng Poku, and T. Johnson Pierce, "Estimation of Subgrade Resilient Modulus from Standard Tests," *Journal of Geotechnical Engineering*, vol. 116, no. 5, pp. 774–789, May 1990, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1990\)116:5\(774\)](https://doi.org/10.1061/(ASCE)0733-9410(1990)116:5(774)).
- [3] E. C. Drumm, J. S. Reeves, M. R. Madgett, and W. D. Trolinger, "Subgrade Resilient Modulus Correction for Saturation Effects," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 123, no. 7, pp. 663–670, Jul. 1997, [https://doi.org/10.1061/\(ASCE\)1090-0241\(1997\)123:7\(663\)](https://doi.org/10.1061/(ASCE)1090-0241(1997)123:7(663)).
- [4] L. N. Mohammad, H. H. Titi, and A. Herath, "Evaluation of Resilient Modulus of Subgrade Soil by Cone Penetration Test," *Transportation Research Record*, vol. 1652, no. 1, pp. 236–245, Jan. 1999, <https://doi.org/10.3141/1652-30>.
- [5] D. G. Kim, "Development of a constitutive model for resilient modulus of cohesive soils," Ph.D. dissertation, Civil Engineering, The Ohio State University, Columbus, OH, USA, 2004.
- [6] D. C. Hai and X. T. Nguyen, *Road Design*. Education Publisher, 2012.
- [7] V. T. Phan and T. T. H. Nguyen, "Elastic and Deformation Characteristics of MSWI Bottom Ash for Road Construction," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6389–6392, Dec. 2020, <https://doi.org/10.48084/etasr.3817>.
- [8] A. H. Bhutto *et al.*, "Mohr-Coulomb and Hardening Soil Model Comparison of the Settlement of an Embankment Dam," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4654–4658, Oct. 2019, <https://doi.org/10.48084/etasr.3034>.
- [9] H. L. Wang *et al.*, "Effects of inclusion contents on resilient modulus and damping ratio of unsaturated track-bed materials," *Canadian Geotechnical Journal*, May 2017, <https://doi.org/10.1139/cgj-2016-0673>.
- [10] Y. H. Huang, *Pavement analysis and design*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
- [11] D.-S. Kim and S. Drabkin, "Accuracy improvement of external resilient modulus measurements using specimen grouting to end platens," *Transportation Research Record*, no. 1462, pp. 65–71, 1994.
- [12] T. Y. Chu, S. N. Chen, S. S. Guram, R. L. Stewart, and W. K. Humphries, "Soil Moisture as a Factor in Subgrade Evaluation," *Transportation Engineering Journal of ASCE*, vol. 103, no. 1, pp. 87–102, Jan. 1977, <https://doi.org/10.1061/TPEJAN.0000625>.
- [13] G. B. Thadkamalla and K. P. George, "Characterization of subgrade soils at simulated field moisture," *Characterization of subgrade soils at simulated field moisture*, no. 1481, pp. 21–27, 1995.
- [14] R. P. Elliott and S. I. Thornton, "Simplification of subgrade resilient modulus testing," *Transportation Research Record*, no. 1192, pp. 1–7, 1988.
- [15] R. Pezo and W. R. Hudson, "Prediction Models of Resilient Modulus for Nongranular Materials," *Geotechnical Testing Journal*, vol. 17, no. 3, pp. 349–355, Sep. 1994, <https://doi.org/10.1520/GTJ10109J>.
- [16] J. M. Burczyk, K. Ksaibati, R. Anderson-Sprecher, and M. J. Farrar, "Factors influencing determination of a subgrade resilient modulus value," *Transportation Research Record*, no. 1462, pp. 72–78, 1994.
- [17] M. R. Thompson and Q. L. Robnett, "Resilient Properties of Subgrade Soils," *Transportation Engineering Journal of ASCE*, vol. 105, no. 1, pp. 71–89, Jan. 1979, <https://doi.org/10.1061/TPEJAN.0000772>.
- [18] G. Rada and M. W. Witzczak, "Comprehensive evaluation of laboratory resilient moduli results for granular material," *Transportation Research Record*, no. 810, pp. 23–33, 1981.
- [19] D. Li and E. T. Selig, "Resilient Modulus for Fine-Grained Subgrade Soils," *Journal of Geotechnical Engineering*, vol. 120, no. 6, pp. 939–957, Jun. 1994, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1994\)120:6\(939\)](https://doi.org/10.1061/(ASCE)0733-9410(1994)120:6(939)).
- [20] D. G. Fredlund, A. T. Bergan, and P. K. Wong, "Relation between resilient modulus and stress conditions for cohesive subgrade soils," *Transportation Research Record*, no. 642, pp. 73–81, 1977.
- [21] R. Zhang, T. Ren, M. A. Khan, Y. Teng, and J. Zheng, "Back-Calculation of Soil Modulus from PFWD Based on a Viscoelastic Model," *Advances in Civil Engineering*, vol. 2019, p. e1316341, Nov. 2019, <https://doi.org/10.1155/2019/1316341>.
- [22] W. T. Oh and S. Vanapalli, "Modelling the stress versus displacement behavior of shallow foundations in unsaturated coarse-grained soils," *Deformation Characteristics of Geomaterials*, pp. 821–828, 2011, <https://doi.org/10.3233/978-1-60750-822-9-821>.
- [23] S. K. Vanapalli and F. M. O. Mohamed, "Bearing Capacity of Model Footings in Unsaturated Soils," in *Experimental Unsaturated Soil Mechanics*, Berlin, Heidelberg, Germany: Springer, 2007, pp. 483–493, https://doi.org/10.1007/3-540-69873-6_48.
- [24] W. T. O. T. Oh, S. K. V. K. Vanapalli, and A. J. P. J. Puppala, "Semi-empirical model for the prediction of modulus of elasticity for unsaturated soils," *Canadian Geotechnical Journal*, Jul. 2009, <https://doi.org/10.1139/T09-030>.
- [25] W. V. Ping and L. Ge, "Field Verification of Laboratory Resilient Modulus Measurements on Subgrade Soils," *Transportation Research Record*, vol. 1577, no. 1, pp. 53–61, Jan. 1997, <https://doi.org/10.3141/1577-07>.
- [26] Y. Yao, J. Qian, J. Li, A. Zhang, and J. Peng, "Calculation and Control Methods for Equivalent Resilient Modulus of Subgrade Based on Nonuniform Distribution of Stress," *Advances in Civil Engineering*, vol. 2019, Sep. 2019, Art. no. e6809510, <https://doi.org/10.1155/2019/6809510>.
- [27] L. Raad and B. A. Zeid, "Repeated load model for subgrade soils: model development," *Transportation Research Record*, no. 1278, pp. 72–82, 1990.
- [28] M. R. Thompson and Q. L. Robnett, "Resilient Properties of Subgrade Soils," *Transportation Engineering Journal of ASCE*, vol. 105, no. 1, pp. 71–89, Jan. 1979, <https://doi.org/10.1061/TPEJAN.0000772>.
- [29] L. N. Mohammad, H. H. Titi, and A. Herath, "Intrusion technology: An innovative approach to evaluate resilient modulus of subgrade soils," presented at the Application of Geotechnical Principles in Pavement Engineering, Proceedings of Sessions of Geo-Congress 98 American Society of Civil Engineers, Boston, MA, USA, 1998.
- [30] A. M. Rahim and K. P. George, "Automated Dynamic Cone Penetrometer for Subgrade Resilient Modulus Characterization," *Transportation Research Record*, vol. 1806, no. 1, pp. 70–77, Jan. 2002, <https://doi.org/10.3141/1806-08>.
- [31] K. P. George, "Prediction of resilient modulus from soil index properties," Department of Civil Engineering, University of Mississippi, Final report FHWA/MS-DOT-RD-04-172, Aug. 2004.
- [32] R. F. Carmichael III and E. Stuart, "Predicting resilient modulus: a study to determine the mechanic properties of subgrade soils (abridgment)," *Transportation Research Record*, no. 1043, pp. 145–148, 1985.
- [33] *Standard Specification for Classification of Soils and Soil-Aggregate Mixtures for Highway Construction Purposes, AASHTO M 145- 91*, American Association of State Highway and Transportation Officials, 1991.
- [34] *Standard Method of Test for Resilient Modulus of Subgrade Soils and Untreated Base/Subbase Materials – SHRP Protocol P46, AASHTO*

- T294-03, American Association of State Highway and Transportation Officials, 2003.
- [35] *Determining the Liquid Limit of Soils, AASHTO T89-07*, American Association of State Highway and Transportation Officials, 2007.
- [36] *Determining the Plastic Limit and Plasticity Index of Soils, AASHTO T90-10*, American Association of State Highway and Transportation Officials, 2010.
- [37] *Standard Method of Test for Particle Size Analysis of soils, AASHTO T88-04*, American Association of State Highway and Transportation Officials, 2004.
- [38] *Standard Method of Test for Moisture-Density Relations of Soils Using a 4.54-kg (10-lb) Rammer and a 457-mm (18-in.) Drop, AASHTO T180-01*, American Association of State Highway and Transportation Officials, 2001.
- [39] *Standard Method of Test for Laboratory Determination of Moisture Content of Soils, AASHTO T265-04*, American Association of State Highway and Transportation Officials, 2004.
- [40] *Standard Method of Test for Unconfined Compressive Strength of Cohesive Soil, AASHTO T208-05*, American Association of State Highway and Transportation Officials, 2005.

Development of an Ethernet-based Portable Embedded Isolated ECG System

Ali S. AlMejrad

Department of Clinical Laboratory Sciences

Faculty of Applied Medical Sciences

University of Hail

Hail, Saudi Arabia

a.almejrad@uoh.edu.sa

Abstract-Patient monitoring is an important issue in the medical field, requiring planning, effort, and time. One of the challenges that physicians, nurses, paramedics, and emergency medical technicians face is the shortage of specialized medical equipment and staff with the ability to measure, display, store, and transmit the vital signs of patients for remote diagnosis under critical conditions. Therefore, and based on our evaluation for such needs, our main research objective was to develop a new system to enable these professionals to perform such monitoring functions with a higher level of safety, performance, and time effectiveness with reduced cost. An advanced, integrated, effortless, and portable system was developed, which is capable to measure, monitor, display, store, and transmit data related to the Electrocardiography (ECG) signal along with its derivatives, such as the heart rate using a hybrid system that combines hardware and software.

Keywords-ECG; biomedical instrumentation; embedded system; medical isolation

I. INTRODUCTION

Electrocardiography (ECG) [1] is an interpretation of the electrical activity of the heart, detected by electrodes attached to the surface of the skin and recorded by a device, external to the body. Detection is conducted in the form of tiny rises and falls in the voltage between two electrodes placed at the sides of the heart. The display looks like a wavy line either on a screen or on paper. The electrodes are standardly applied on the four extremities and the chest surface. The display indicates the overall rhythm of the heart and weaknesses in different parts of the heart. The ECG device detects and amplifies the small electrical changes on the skin that are caused when the heart muscle depolarizes during each heartbeat. It provides information on the condition and performance of the heart. ECG is one of the vital signs monitored in many medical and intensive care procedures. Any deviation from the normal in a particular electrocardiogram is indicative of a possible heart disorder. Any ECG instrument is provided with an alert mechanism that warns the medical staff of any abnormal changes detected in the cardiac function. ECG equipment has been developed over the years [2]. Digital hardware programmable logic was used to develop an ECG system in late '60s [3] and computer based ECG systems came into use in

many hospitals [4, 5]. Wireless and digital technologies have been used to develop ECG home monitoring and telehealth systems [6-8]. To reduce size and cost, embedded technology has been used in portable medical systems [9-11].

Cardiovascular Diseases (CVDs) are one of the leading mortality causes accounting for 31% of deaths worldwide [12]. In Saudi Arabia and in the Gulf Countries, CVDs are becoming a major health concern as being contributing to more than 45% of total deaths [13]. This percentage increases in rural areas where the healthcare is provided in primary health centers with shortage of specialized medical equipment and staff. The most affected people with CVD are elderly and adults with low income. Patients of CVDs need emergency care as soon as they occur.

Our case study is the Hail region in the north of Saudi Arabia, having a high number of rural villages (59). The goal of our research, based on our evaluation of the needs of the Primary Health Centers (PHCs), is to develop a new system with the required features to enable healthcare professionals to perform monitoring functions with higher safety, performance, and time effectiveness with reduced cost using embedded technology and advanced signal processing that can be used remotely via the Virtual Instrumentation (VI) technology. This system should also overcome the lack of expert physicians in rural regions or some urban clinics or health centers, who can be consulted remotely in real-time if needed. The proposed dedicated new system consists of a low cost computer, with a developed portable device connected via Ethernet, and dedicated developed software (National Instruments Inc.). The Graphical User Interface (GUI) - VI system has been designed for easy measurement and analysis. In addition, it provides a user-friendly interface and the ability for advanced analysis of the acquired ECG signals using advanced signal processing techniques in time and frequency domains to help accurately diagnose such events at early stages without the need for consultation except from very critical situations.

The proposed integrated, effortless, and portable system measures, monitors, displays, stores, and transmits data related to the ECG signal along with its derivatives. The system is planned to have the capability to easily process, control, and send the resulted signal through the Ethernet technology.

Corresponding author: Ali S. AlMejrad

II. SYSTEM DEVELOPMENT

A. Hardware

The block diagram shown in Figure 1 consists of four stages: pre-amplification (surge arrester and low pass filter), amplification (ECG amplifier), controlling (μ c, EEPROM, and JTAG), and Interfacing (USB to RS232 converter, Serial to Ethernet convertor, power adaptor and power modules).

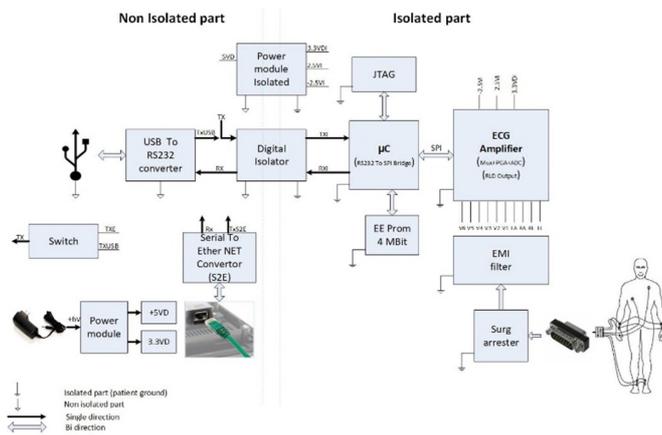


Fig. 1. System block diagram.

The elements of the system are:

- Surge arrester: It consists of a resistor and a gas discharge tube. It is used to prevent the circuit from high level defibrillation voltage (6000V) by reducing it into nearly 120V.
- Electro Magnetic Interference (EMI) Filter: It is built inside the ECG amplifier.
- Low Pass Filter: It is used to filter the weakly resulted ECG signal to become suited to processing, conditioning, and displaying. The ECG amplifier has 10 input leads: LL, RL, RA, LA, V1, V2, V3, V4, V5, and V6 with 3 power pens of -2.5V, 2.5V analog voltage, and 3.3V digital voltage.
- Serial Peripheral Interface (SPI): It is a protocol that allows bidirectional communication between the ECG amplifier and the microcontroller.
- Microcontroller (μ c): It is used to represent the controlling stage after the end of the previous pre-amplification and amplification stages. The microcontroller is programmed in C language. The microcontroller will convert the RS232 serial interface into the SPI.
- JTAG: It is linked with a bidirectional linker with the μ c and it is represented by a connector for the purpose of μ c programming.
- Electrical Erased Programmable Read Only Memory (EEPROM): It is used for the temporary storage of data before receiving a server request of transmitting these data by the μ c. It has a bidirectional linking with the μ c because it is available for both writing and reading, even though it is called Read Only Memory (ROM).

- Digital Isolator: It is represented by a small power transformer and it is used for the isolation of the transmitting and receiving buses between the computer interface region and the patient interface region.
- Power Module Isolated: It consists of power regulators distributed through two stages, the first stage converts the 5VD power module output into 5V and -5V output while the second stage converts the 5V primary output into 2.5V and 3.3VD secondary output and the -5V primary output into -2.5V secondary output.
- Transmit (TX) Switch: It is used to switch between the input from Ethernet and the input from the USB controller (Figure 2).



Fig. 2. The TX Switch.

- USB to RS232 Converter: It is used to convert the USB bus into the RS232 serial interface bus that is connected to the digital isolator.
- Power Module: It is one of the two input voltage sources. It converts the 6V input from the adapter into 5V and 3.3V (digital voltage). It is actually a low power transformer.
- Serial to Ethernet converter: It is used to convert the signal bus (RS232) which comes from the power module into Ethernet signal.

Our system has the ability to choose between two alternative input sources, the electrical adaptor and the USB adaptor. We can classify the components of our system into two classes: Isolated part components (patient-side components) and Non-Isolated part components (computer-side components). It is a high-safe system because it offers a safe protection to the patient by using a main digital isolator that separates the Non-Isolated from the Isolated parts. It also offers circuit protection from the high defibrillation voltage (6000V) by reducing it into a much lower voltage (120V) by using the surge arrester.

B. Software

To control the developed hardware, a graphical program has been developed in LabVIEW [14]. The Graphical User Interface (GUI) is termed as Virtual Instrument (VI) because its appearance and operation imitates physical instruments. LabVIEW contains a comprehensive set of tools for acquiring, analyzing, displaying, and storing data. LabVIEW is used to communicate with hardware for data acquisition via ports such as GPIB, RS232 and RS485. The front panel in Figure 3 shows 9 leads of normal ECG signals while the front panel in Figure 4 shows normal ECG with a red indicator detecting that the V1 lead is off. Heart rate derivation from the ECG is conducted as shown in the block diagram in Figure 5.



Fig. 3. Normal ECG signal.

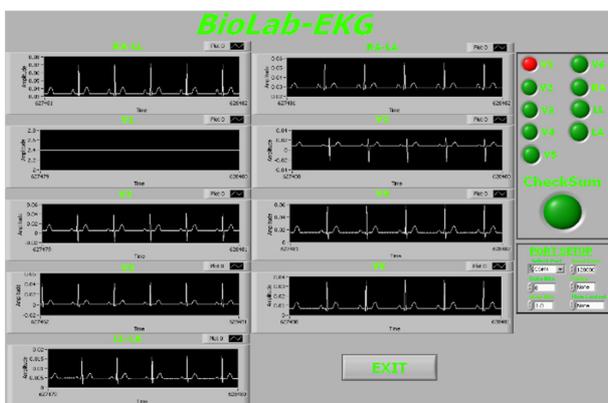


Fig. 4. Normal ECG with a red indicator alerting that the V1 lead is off.

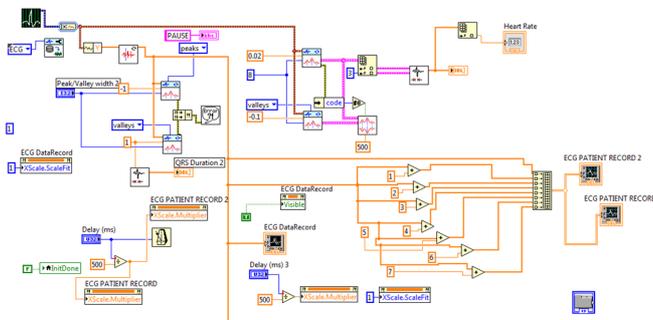


Fig. 5. VI block diagram.

Wavelet-based peak detection is applied. The peaks in an ECG signal indicate the rhythm of the heart. Typically, one lead of an ECG signal (such as lead II) consists of 5 peaks and valleys. The first positive deflection is referred to as the P wave. The duration between the Q and S waves is called the QRS duration and indicates the ventricular condition. The peaks in ECG signals usually are corrupted by environment noise and cannot be detected properly by traditional curve-fitting methods. The wavelet-based peak detection method is naturally immune to noise and can detect both sharp and mild peaks/valleys accurately. The wavelet-based method also is less affected by trends in the signal. This item uses the WA Multi-scale Peak Detection VI to detect the peaks in an ECG signal

and computes the QRS duration as shown in the block diagram in Figure 6.

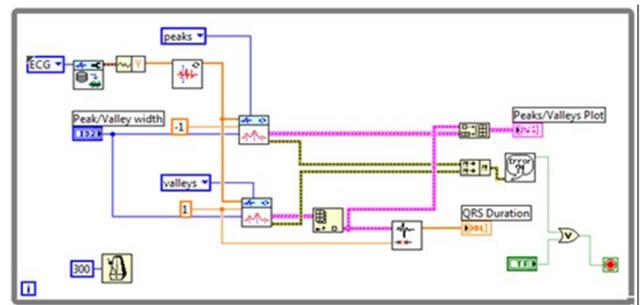


Fig. 6. ECG QRST complex detection block diagram.

III. RESULTS AND DISCUSSION

To test the developed system we did various tests using a patient simulator for normal (Figure 7) and abnormal cases (Figures 8-10).

A. Normal Case

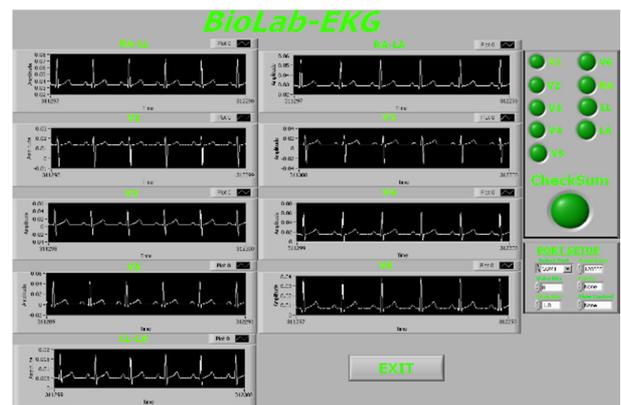


Fig. 7. Normal ECG signal.

B. Abnormal Cases

1) Tachycardia

The heart rate is controlled by electrical signals sent across heart tissues. Tachycardia [15] occurs when an abnormality in the heart produces rapid electrical signals. In some cases, tachycardia may cause no symptoms or complications. However, tachycardias can seriously disrupt normal heart function, increase the risk of stroke, or cause sudden cardiac arrest or death. Treatments may help control a rapid heartbeat or manage diseases contributing to tachycardia. A healthy adult heart normally beats 60 to 100 times a minute when a person is at rest. If you have tachycardia the rate in the upper chambers or lower chambers of the heart, or both, are increased significantly as shown in Figure 8.

2) Atrial Flutter

Atrial Flutter (AFL) [15] is an abnormal heart rhythm that occurs in the atria of the heart. When it first occurs, it is usually associated with a fast heart rate or tachycardia (beats over 100

per minute), and falls into the category of supra-ventricular tachycardia's as shown in Figure 9.

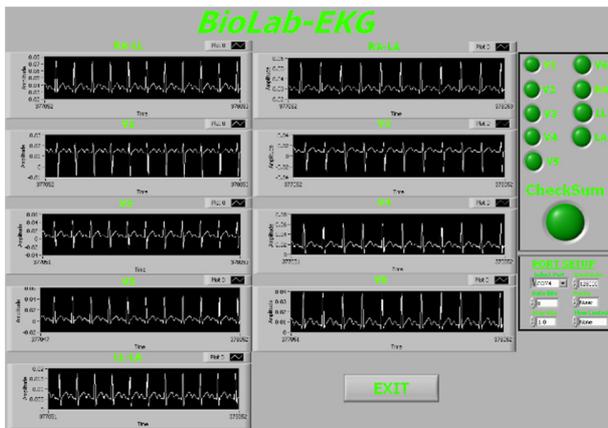


Fig. 8. Tachycardia.

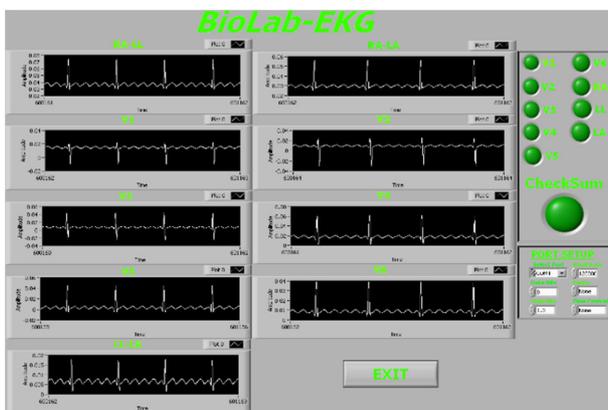


Fig. 9. Atrial flutter.

While this rhythm occurs most often in individuals with CVDs such as hypertension, coronary artery disease, and cardiomyopathy, it may occur spontaneously in people with otherwise normal hearts. It is typically not a stable rhythm, and frequently degenerates into Atrial fibrillation. Rarely, it persists for months to years. Atrial flutter is usually well tolerated initially (a high heart rate is for most people just a normal response to exercise), however, people with other underlying heart diseases or poor exercise tolerance may rapidly develop symptoms, which can include shortness of breath, chest pains, lightheadedness or dizziness, nausea and, in some patients, nervousness and feelings of impending doom. Prolonged fast flutter may lead to decompensation with loss of normal heart function (heart failure). This may manifest as effort intolerance (exertional breathlessness), nocturnal breathlessness, or swelling of the legs or abdomen.

3) Ventricular Fibrillation

Ventricular Fibrillation (V-fib or VF) [15] is a condition in which there is uncoordinated contraction of the cardiac muscle of the ventricles in the heart, making them quiver rather than contract properly as shown in Figure 10. Ventricular fibrillation is a medical emergency and the most commonly identified

arrhythmia in cardiac arrest patients. While there is activity, it is undetectable by palpation (feeling) at major pulse points of the carotid and femoral arteries especially by the layperson. Such an arrhythmia is only confirmed by ECG. VF is a medical emergency that requires prompt basic life support interventions because should the arrhythmia continue for more than a few seconds, it will likely degenerate further into a systole ("flat line"). The condition results in cardiogenic shock, cessation of effective blood circulation, and Sudden Cardiac Death (SCD) will result in a matter of minutes.

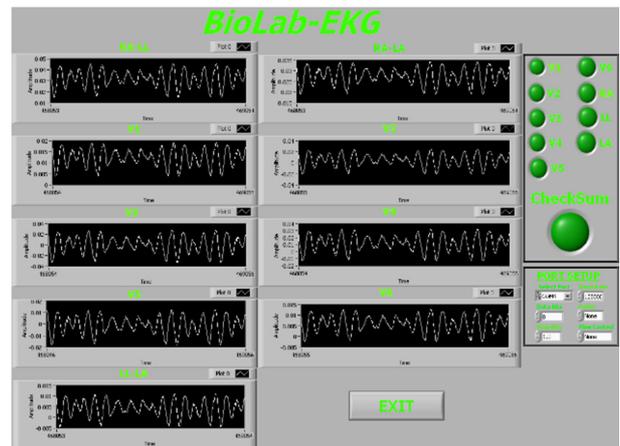


Fig. 10. Ventricular fibrillation.

C. The Realized System

The main designed and developed Embedded Isolated ECG system board with size of 75mm×75mm is shown in Figure 11.



Fig. 11. The developed main embedded isolated ECG system board.

IV. CONCLUSION

The developed ECG system development stages consist of designing the board, creating the layout, fabricating the printed circuit board, and assembling the components, in order to get the final system. The developed system has the capability to acquire different ECG signals from different sites on the body, protect the patient at this side, amplify the acquired signal, control it, transfer it from an isolated safely system's part

(patient-end part) to a non-isolated one (computer-end part) with perfect isolation. A GUI was built for displaying the acquired signals using the LabVIEW program with the properties of displaying 9 ECG leads simultaneously or any selected specific lead and storing and transferring them through the Ethernet. The presence of physicians, nurses, and paramedics in the PHCs and private clinics ensures the developed system to be of low cost, small, portable, safe, biocompatible, effective, precise, simple, and durable. The system will be applied and validated in future work after Saudi Food and Drug Authority approval.

ACKNOWLEDGEMENT

The author would like to thank the University of Hail, Kingdom of Saudi Arabia.

REFERENCES

- [1] P. Kligfield *et al.*, "Recommendations for the standardization and interpretation of the electrocardiogram: part I: The electrocardiogram and its technology: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology," *Circulation*, vol. 115, no. 10, pp. 1306–1324, Mar. 2007, <https://doi.org/10.1161/CIRCULATIONAHA.106.180200>.
- [2] W. Einthoven, "The string galvanometer and the human electrocardiogram," in *KNAW, Proceedings*, Amsterdam, Netherlands, 1904, vol. 6, pp. 107–115.
- [3] A. Rosado, M. Bataller, J. F. Guerrero, J. Calpe, J. V. Francés, and J. R. Magdalena, "High performance hardware correlation coefficient assessment using programmable logic for ECG signals," *Microprocessors and Microsystems*, vol. 27, no. 1, pp. 33–39, Feb. 2003, [https://doi.org/10.1016/S0141-9331\(02\)00083-2](https://doi.org/10.1016/S0141-9331(02)00083-2).
- [4] A. C. Zoric and S. S. Ilic, "PC-Based System For Electrocardiography And Data Acquisition," in *The SIKS 2005 - 2005 uth International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services*, Nis, Serbia, Sep. 2005, vol. 2, pp. 619–622, <https://doi.org/10.1109/TEL.SKS.2005.1572189>.
- [5] A. Josko and R. J. Rak, "Effective simulation of signals for testing ECG analyzer," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 3, pp. 1019–1024, Jun. 2005, <https://doi.org/10.1109/TIM.2005.847350>.
- [6] S. Korsakas *et al.*, "The mobile ECG and motion activity monitoring system for home care patients," in *2006 Computers in Cardiology*, The Mobile ECG and Motion Activity Monitoring System for Home Care Patient, Sep. 2006, pp. 833–836.
- [7] F. Bagheri, N. Ghafarnia, and F. Bahrami, "Electrocardiogram (ECG) Signal Modeling and Noise Reduction Using Hopfield Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 3, no. 1, pp. 345–348, Feb. 2013, <https://doi.org/10.48084/etasr.243>.
- [8] K. H. Hii, V. Narayanamurthy, and F. Samsuri, "ECG Noise Reduction with the Use of the Ant Lion Optimizer Algorithm," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4525–4529, Aug. 2019, <https://doi.org/10.48084/etasr.2766>.
- [9] S. Kara, S. Kemaloğlu, and Ş. Kirbaş, "Low-Cost Compact ECG With Graphic LCD and Phonocardiogram System Design," *Journal of Medical Systems*, vol. 30, no. 3, pp. 205–209, Jun. 2006, <https://doi.org/10.1007/s10916-005-7989-6>.
- [10] A. S. AlMejrad, "A Single Supply Standard 8051 Microcontroller Based Medical K-grade Isolation ECG Module with Graphics LCD," in *2012 Second International Conference on Intelligent System Design and Engineering Application*, Sanya, China, Jan. 2012, pp. 1184–1187, <https://doi.org/10.1109/ISdea.2012.584>.
- [11] A. S. AlMejrad, "Design of an intelligent system for speech monitoring and treatment of low and excessive vocal intensity," *Artificial Life and Robotics*, vol. 15, no. 3, pp. 320–324, Sep. 2010, <https://doi.org/10.1007/s10015-010-0816-5>.
- [12] "Cardiovascular diseases (CVDs)," *WHO*, May 2017. [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) (accessed Apr. 20, 2021).
- [13] N. Aljefree and F. Ahmed, "Prevalence of Cardiovascular Disease and Associated Risk Factors among Adult Population in the Gulf Region: A Systematic Review," *Advances in Public Health*, vol. 2015, Jan. 2015, Art. no. e235101, <https://doi.org/10.1155/2015/235101>.
- [14] G. W. Johnson, *Labview Graphical Programming: Practical Applications in Instrumentation and Control/Book and Disk*. New York, NY, USA: McGraw-Hill, 1994.
- [15] B. S. M. MACC and T. K. MD, *Chou's Electrocardiography in Clinical Practice: Adult and Pediatric*, 6th edition. Philadelphia, PA, USA: Saunders, 2008.

AUTHORS PROFILE

Ali S. AlMejrad received the B.Sc. with 1st honor in Electrical Engineering from King Saud University, Saudi Arabia, M.Sc. in Bioengineering from Strathclyde University, UK in 1987 and 1990 respectively and the Ph.D. in Medical Engineering from the University of Edinburgh, UK, in 1996. He is currently an Associate Professor of Medical Engineering and Instrumentation in Clinical Laboratory Sciences at the College of Applied Medical Sciences, University of Hail, Saudi Arabia. His research interests include the Development of Biomedical Instrumentation and Intelligent Systems for Health Care using advanced Technologies. He is a member of IEEE and IPEM

Evolutionary Algorithm-based Feature Selection for an Intrusion Detection System

Devendra Kumar Singh

Department of Computer Science & Engineering
Guru Ghasidas University
Bilaspur, India
devendra.singh1700@gmail.com

Manish Shrivastava

Department of Computer Science & Engineering
Guru Ghasidas University
Bilaspur, India
mannbsp@gmail.com

Abstract—Keeping computer reliability to confirm reliable, secure, and truthful correspondence of data between different enterprises is a major security issue. Ensuring information correspondence over the web or computer grids is always under threat of hackers or intruders. Many techniques have been utilized in intrusion detections, but all have flaws. In this paper, a new hybrid technique is proposed, which combines the Ensemble of Feature Selection (EFS) algorithm and Teaching Learning-Based Optimization (TLBO) techniques. In the proposed, EFS-TLBO method, the EFS strategy is applied to rank the features for choosing the ideal best subset of applicable information, and the TLBO is utilized to identify the most important features from the produced datasets. The TLBO algorithm uses the Extreme Learning Machine (ELM) to choose the most effective attributes and to enhance classification accuracy. The performance of the recommended technique is evaluated in a benchmark dataset. The experimental outcomes depict that the proposed model has high predictive accuracy, detection rate, false-positive rate, and requires less significant attributes than other techniques known from the literature.

Keywords—classification; feature selection; teaching learning-based optimization; intrusion detection

I. INTRODUCTION

Despite the increasing alertness in cyber-security problems, the present ongoing solutions are not suitable for shielding computer applications or undertaking security frameworks against the risk from consistently ever-propelling organized assaults [1]. Adaptable security methods were developed to solve this issue, but they turned out to have further issues. Typical cyber-security methods are insufficient to entirely defend web computer security, since they face problems from intruder attempts at the initiation of the security procedure, e.g. user authentication and firewalling [2, 3]. Thus, a different line of threat protection is acutely mentioned in the Intrusion Detection Systems (IDSs). An IDS is a program that observes the internet for venomous actions and policy contraventions [4]. At present, IDS along with safe-guard applications have resolved into an indispensable element of computer security of most companies. The union of the above-mentioned security machines provides increased opposition in network-attacks and improves network security.

II. RELATED WORK

Gradually, inexhaustible applications, e.g. choice and order models have been put in to intrusion datasets (i.e. KDD CUP 1999) for detecting network problems and attacks. Attribute selection with learning algorithms couldn't control or scale to very large volumes of datasets [5]. To beat this impediment, authors in [6, 7] proposed another hybrid feature selection technique that diminishes the non-applicable features and selects the best ideal component subsets. The recurring pattern study in [8] indicated that individual hunt calculation locates the most suitable subsets that amplify information over-fitting, while a probing interest is less prone to information over-fitting in the part assurance, developing a modest number of tests [9]. Authors in [10] proposed the use of ELM and alpha profiling to diminish the required time while superfluous highlights were disposed of utilizing a group of separated, relationship and consistency-based feature selection procedures.

A. Filter-based Methods

Optimum and appropriate feature subset selection is a task accomplished by choosing the qualities dependent on the high connections of concerning classes and uncorrelated features. From the Conditional Mutual Information Maximization (CMIM) method, Feature/attributes Subset Selection (FSS) is conducted depending on maximizing conditional mutual information [11] regarding the class. In addition, it is extremely close with class attributes and uncorrelated to attributes. It makes a compromise between the predictive power of the nominated competitor (significance for the class carrier) and its freedom from all recently chosen attributes. Mutual Information (MI) estimation between the class label y and attributes X is calculated in:

$$I(y; X) = H(y) - H\left(\frac{y}{X}\right) \quad (1)$$

where $H(y)$ and $H\left(\frac{y}{X}\right)$ show the entropy and conditional entropy of the class change respectively. Some writers have mentioned issues using the Mutual Information-based Feature Selection (MIFS) technique [7, 12]. Therefore, we used this strategy to decrease the readability between class y and data attributes as shown in (2). The primary objective of CMIM is to choose the final feature subset that conveys as much information as possible from the sample S :

$$M_{CMIM}(X) = \min_{x_j \in S} I(y; x_k/x_j) \quad (2)$$

where M_{CMIM} estimates the mutual information between the full features set x_k and certain features x_j regarding class label y , whereas S shows the subsets of the selected features. $I(y; x_k/x_j)$ measures the quantity of the classification information that x_k affords when x_j has been selected [13]. Selected feature subset S cannot provide this information. As comparison to $I(y; x_k)$, $I(y; x_k/x_j)$ does not contain the superfluous data of pair wise attributes for categorization.

The importance of the input attributes defined by the JMI is shown in (3):

$$M_{JMI}(X) = \sum_{x_j \in S} I(x_k; x_j; y) \propto \sum_{x_j \in S} I(y; x_k/x_j) \quad (3)$$

where $I(x_k; x_j; y)$ signifies the mutual information between the novel attribute subset x_k and the selected attributes x_j regarding class y . In linguistics of mutual information, the determination of attribute choice is to reduce attribute subsets S with N attributes with a maximum holding on the target class c . This structure, called Max-Dependency, has the form of:

$$\max w(X, y) = I(y; x_1, x_2, \dots, x_N) = H(y)H\left(\frac{y}{x_1, x_2, \dots, x_N}\right) \quad (4)$$

In (3), the holding among attribute X is resultant and can have a high value [14]. The correspondence between readability between attributes is expressed in (5) and (6):

$$\min Z(X, c) = 1/s^2 \mid \sum_{x_j \in S} I(x_j; x_k) \quad (5)$$

$$\text{Max } \emptyset(w, Z) = w - Z \quad (6)$$

The incorporation (i.e. integration) of (5) and (6) is known as the Minimal-Redundancy-Maximal-Relevance (mRMR) [15]:

$$j_{mRMR}(\emptyset) = I(c; X) - 1/s^2 \mid \sum_{x_j \in S} I(x_j; x_k) \quad (7)$$

where x_j is a selected subset of attributes S and x_k is a native feature set.

B. The Proposed Ensemble Feature Selection

The pre-owned Feature Selection strategies are mRMR, JMI, and CMIM which can relegate the position of the IDS datasets and the output is aggregated utilizing a combination strategy [7].

C. Frequency Vote

Frequency Vote (FV) is a cooperative decision making framework that has been proposed as more useful than other increasingly complex plans [16]. Thus, we can follow the most voted prediction as to the last prediction or expectation as per (8):

$$\sum_{n=1}^{\mathcal{L}} d_{n,j} = \operatorname{argmax}_{i \in \{1,2,\dots,L\}} \sum_{n=1}^{\mathcal{L}} d_{n,i} \quad (8)$$

where \mathcal{L} shows the number of attribute choice methods, and L is a selection of some attributes. For attribute j , the sum $\sum_{n=1}^{\mathcal{L}} d_{n,j}$ tabulates the number of votes for j .

D. Using Teaching Learning-Based Optimization (TLBO)

TLBO [17-19] is the best and most powerful metaheuristic method to apply high convergence rate with less adjusting parameters. It is an easy and simple computation of tuning the control parameters with less memory requirements. The working methodology of the TLBO algorithm can create better evaluation outcome [20]. The position of the i th learner is :

$$X_{i,k} = \{X_{i,1}, X_{i,2}, \dots, X_{i,D}\} \quad (9)$$

where L_b shows the lower limit and U_b shows the upper limit of the D dimension in the search area $X_{i,D} \in [L_b, U_b]$ [21]. The learner X is unplanned, initialized in the search area [22]. The development (i.e. evolution) of $X_{i,k}$ is generated by:

$$X_{i,k} = L_{b,k} + r_1 * (U_{b,k} - L_{b,k}) \quad (10)$$

where $i=1, 2, 3, \dots, nPop$, $k=1,2,3,\dots,D$, r_1 signifies the unplanned variable, $L_{b,k}$ shows the lower limit and $U_{b,k}$ shows the upper limit value, and $nPop$ denotes the population count [23]. The simulation of an old-style initiation procedure is arranged into two critical stages of the TLBO calculation: the teacher stage and the learner stage.

In TLBO algorithm, the teacher is a quantification of obtaining an ideal output gained from optimization problems. Therefore, the teacher can grow the mean result of a classroom to a specific result which relies on the ability of the complete classroom. Let $M_{i,k} = (1/nPop) (\sum X_{i,k})$ be the mean value of the particular subject where $k=1, 2, \dots, D$. Equation (11) shows the updating equation process:

$$X_{i,k}^{new} = X_{i,k}^{old} + r_2 * (X_{teacher} - T_f * M_{i,k}) \\ \& T_f = \text{round}[1 + \text{rand}(0,1)] \quad (11)$$

where $X_{teacher}$ is the greatest beginner of the embrance group (i.e. population) at the current duplication of the algorithm, r_2 represents random numbers, T_f behaves as a teaching element that chooses the merit of the mean to be changed. In each iteration, $X_{i,k}^{new}$ is updated from the old merit $X_{i,k}^{old} \cdot X_{i,k}^{new}$ and $X_{i,k}^{old}$ denote the k -th beginner choice after or before it is modernized by the teacher.

E. The Fitness Function

The fitness function must maximize the categorization Accuracy of the calculations accomplished by the best attributes during the progressive (i.e. evolutionary) process, which is defined as:

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

TP , TN , FP , and FN stand for True Positive, True Negative, False Positive, and False Negative respectively.

III. EXPERIMENTAL RESULTS AND DISCUSSION

During the experiments, every record had features such as feature name, records, and feature portrayal. Most IDS numerical studies have been performed on NSL-KDD [24]. This data set have varying data importance and feature integrity. Authors in [25] analyzed the deliberate intrusion dataset called KDD Cup 1999 [36]. Every record is tagged as

normal or as an attack type in the dataset. The flowchart of the proposed method can be seen in Figure 1.

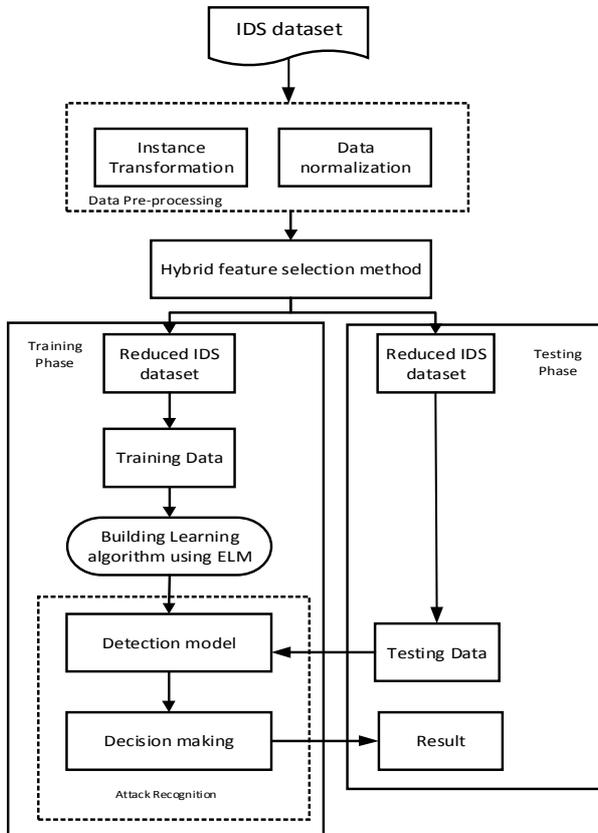


Fig. 1. General structure of the proposed model for intrusion detection.

A. Results and Discussion

Different exploratory tools and techniques were used on the NSL-KDD dataset (see Table I) [14, 25, 26]. The classification performance is estimated with the assistance of the support vector machine categorization with four execution variables. These exhibition measures, along with Accuracy, are [27-29]:

$$\text{Detection Rate: DR} = \frac{TP}{TP + FN}$$

$$\text{Precision: Pr} = \frac{TP}{TP + FP}$$

$$\text{F-measure} = \frac{2 * Pr * Re}{Pr + Re}$$

$$\text{False Alarm Rate: FAR} = \frac{FP}{TN + FP}$$

B. Result Comparison

Tenfold cross-validation was applied to ELM [7] and other classifiers, namely SVM [30, 31] and NB in the IDS dataset. Table I shows the comparison of the performance of the proposed algorithm with existing known algorithms. The result shows that the proposed algorithm performs better on the basis

of parameters like feature, DR, FPR and Accuracy in the same data set. Only five of the attributes have been selected by the proposed method which can identify intrusion attacks in the network with maximum Accuracy.

TABLE I. PERFORMANCE COMPARISON OF THE PROPOSED AND EXISTING ALGORITHMS IN THE SAME DATABASE

Algorithm	Feature	DR	FPR	Accuracy
LSSVM-IDS + FMIFS [32]	18	98.93	0.28	99.94
TUIDS [33]	All	98.88	1.12	96.55
HTTP based IDS [34]	13	99.03	1.0	99.38
Hybrid IDS [35]	All	99.10	1.2	*
Proposed	5	99.31	0.19	99.95

IV. CONCLUSION

In this study, a novel hybrid model called EFS-TLBO with ELM is proposed, to easily identify threats by using the attribute choice algorithm [7] which increases the perceptive power for better class distinction. For exhibiting the superiority of the proposed technique, the NSL-KDD database of intruders was employed. The results show that the proposed technique provides an important depletion to the required features and outperforms the advanced attribute selection techniques from the literature. The practical results show that the suggested technique achieved an accuracy of 99.95% in the NSL-KDD data set of intruders [36, 37], surpassing the other techniques.

Future work is going to be focused on multi-objective algorithms that combine ensemble filter and classification methods for pattern analysis and intrusion attack detection. Also, some different optimization algorithms for ELM parameter optimization are going to be researched.

REFERENCES

- [1] A. Praseed and P. S. Thilagam, "DDoS Attacks at the Application Layer: Challenges and Research Perspectives for Safeguarding Web Applications," *IEEE Communications Surveys Tutorials*, vol. 21, no. 1, pp. 661–685, Firstquarter 2019, <https://doi.org/10.1109/COMST.2018.2870658>.
- [2] S. Pontarelli, G. Bianchi, and S. Teofili, "Traffic-Aware Design of a High-Speed FPGA Network Intrusion Detection System," *IEEE Transactions on Computers*, vol. 62, no. 11, pp. 2322–2334, Nov. 2013, <https://doi.org/10.1109/TC.2012.105>.
- [3] N. Fallahi, A. Sami, and M. Tajbakhsh, "Automated flow-based rule generation for network intrusion detection systems," in *24th Iranian Conference on Electrical Engineering*, Shiraz, Iran, May 2016, pp. 1948–1953, <https://doi.org/10.1109/IranianCEE.2016.7585840>.
- [4] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, "A System for Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 447–456, Feb. 2014, <https://doi.org/10.1109/TPDS.2013.146>.
- [5] Y. Wang, W. Meng, W. Li, J. Li, W.-X. Liu, and Y. Xiang, "A fog-based privacy-preserving approach for distributed signature-based intrusion detection," *Journal of Parallel and Distributed Computing*, vol. 122, pp. 26–35, Dec. 2018, <https://doi.org/10.1016/j.jpdc.2018.07.013>.
- [6] S. A. Medjehed, T. A. Saadi, A. Benyettou, and M. Ouali, "Kernel-based learning and feature selection analysis for cancer diagnosis," *Applied Soft Computing*, vol. 51, pp. 39–48, Feb. 2017, <https://doi.org/10.1016/j.asoc.2016.12.010>.
- [7] A. K. Shukla and P. Singh, "Building an Effective Approach toward Intrusion Detection Using Ensemble Feature Selection," *International*

- Journal of Information Security and Privacy*, vol. 13, no. 3, pp. 31–47, 2019.
- [8] T. F. Ghanem, W. S. Elkilani, and H. M. Abdul-kader, "A hybrid approach for efficient anomaly detection using metaheuristic methods," *Journal of Advanced Research*, vol. 6, no. 4, pp. 609–619, Jul. 2015, <https://doi.org/10.1016/j.jare.2014.02.009>.
- [9] S. Dwivedi, M. Vardhan, and S. Tripathi, "Incorporating evolutionary computation for securing wireless network against cyberthreats," *The Journal of Supercomputing*, vol. 76, no. 11, pp. 8691–8728, Nov. 2020, <https://doi.org/10.1007/s11227-020-03161-w>.
- [10] R. Singh, H. Kumar, and R. K. Singla, "An intrusion detection system using network traffic profiling and online sequential extreme learning machine," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8609–8624, Dec. 2015, <https://doi.org/10.1016/j.eswa.2015.07.015>.
- [11] A. K. Shukla, P. Singh, and M. Vardhan, "A two-stage gene selection method for biomarker discovery from microarray data for cancer classification," *Chemometrics and Intelligent Laboratory Systems*, vol. 183, pp. 47–58, Dec. 2018, <https://doi.org/10.1016/j.chemolab.2018.10.009>.
- [12] C. Liu, W. Wang, Q. Zhao, X. Shen, and M. Konan, "A new feature selection method based on a validity index of feature subset," *Pattern Recognition Letters*, vol. 92, pp. 1–8, Jun. 2017, <https://doi.org/10.1016/j.patrec.2017.03.018>.
- [13] P. E. Meyer, C. Schretter, and G. Bontempi, "Information-Theoretic Feature Selection in Microarray Data Using Variable Complementarity," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 261–274, Jun. 2008, <https://doi.org/10.1109/JSTSP.2008.923858>.
- [14] A. K. Shukla, P. Singh, and M. Vardhan, "A hybrid gene selection method for microarray recognition," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 4, pp. 975–991, Jan. 2018, <https://doi.org/10.1016/j.bbe.2018.08.004>.
- [15] A. K. Shukla, P. Singh, and M. Vardhan, "Gene selection for cancer types classification using novel hybrid metaheuristics approach," *Swarm and Evolutionary Computation*, vol. 54, May 2020, Art. no. 100661, <https://doi.org/10.1016/j.swevo.2020.100661>.
- [16] L. Barolli and O. Terzo, Eds., *Complex, Intelligent, and Software Intensive Systems*, 1st edition. New York, NY, USA: Springer, 2017.
- [17] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems," *Computer-Aided Design*, vol. 43, no. 3, pp. 303–315, Mar. 2011, <https://doi.org/10.1016/j.cad.2010.12.015>.
- [18] A. Rajasekhar, R. Rani, K. Ramya, and A. Abraham, "Elitist Teaching Learning Opposition based algorithm for global optimization," in *IEEE International Conference on Systems, Man, and Cybernetics*, Seoul, Korea (South), Oct. 2012, pp. 1124–1129, <https://doi.org/10.1109/ICSMC.2012.6377882>.
- [19] Y. Oubbati and S. Arif, "Transient stability constrained optimal power flow using teaching learning based optimization," in *8th International Conference on Modelling, Identification and Control*, Algiers, Algeria, Nov. 2016, pp. 284–289, <https://doi.org/10.1109/ICMIC.2016.7804124>.
- [20] A. K. Shukla, S. K. Pippal, and S. S. Chauhan, "An empirical evaluation of teaching–learning-based optimization, genetic algorithm and particle swarm optimization," *International Journal of Computers and Applications*, pp. 1–15, Nov. 2019, <https://doi.org/10.1080/1206212X.2019.1686562>.
- [21] M. M. Polycarpou, A. de Carvalho, J.-S. Pan, M. Wozniak, H. Quintian, and E. Corchado, Eds., *Hybrid Artificial Intelligence Systems*. New York, NY, USA: Springer International Publishing, 2014.
- [22] A. K. Shukla, P. Singh, and M. Vardhan, "An adaptive inertia weight teaching–learning-based optimization algorithm and its applications," *Applied Mathematical Modelling*, vol. 77, pp. 309–326, Jan. 2020, <https://doi.org/10.1016/j.apm.2019.07.046>.
- [23] P. K. Nayak, S. Mishra, P. K. Dash, and R. Bisoi, "Comparison of modified teaching–learning-based optimization and extreme learning machine for classification of multiple power signal disturbances," *Neural Computing and Applications*, vol. 27, no. 7, pp. 2107–2122, Oct. 2016, <https://doi.org/10.1007/s00521-015-2010-0>.
- [24] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," *IEEE Communications Surveys Tutorials*, vol. 18, no. 2, pp. 1153–1176, Secondquarter 2016, <https://doi.org/10.1109/COMST.2015.2494502>.
- [25] M. Tavallae, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *IEEE Symposium on Computational Intelligence for Security and Defense Applications*, Ottawa, Canada, Jul. 2009, pp. 1–6, <https://doi.org/10.1109/CISDA.2009.5356528>.
- [26] G. V. Nadiammai and M. Hemalatha, "Effective approach toward Intrusion Detection System using data mining techniques," *Egyptian Informatics Journal*, vol. 15, no. 1, pp. 37–50, Mar. 2014, <https://doi.org/10.1016/j.eij.2013.10.003>.
- [27] K. Ray, S. N. Sharan, S. Rawat, S. K. Jain, S. Srivastava, and A. Bandyopadhyay, Eds., *Engineering Vibration, Communication and Information Processing*, 1st edition. New York, NY, USA: Springer, 2018.
- [28] N. Shakhovska, Ed., *Advances in Intelligent Systems and Computing: Selected Papers from the International Conference on Computer Science and Information Technologies, ... in Intelligent Systems and Computing, 512*, 1st ed. New York, NY: Springer, 2016.
- [29] S. R. Basha and J. K. Rani, "A Comparative Approach of Dimensionality Reduction Techniques in Text Classification," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4974–4979, Dec. 2019, <https://doi.org/10.48084/etasr.3146>.
- [30] F. Kuang, S. Zhang, Z. Jin, and W. Xu, "A novel SVM by combining kernel principal component analysis and improved chaotic particle swarm optimization for intrusion detection," *Soft Computing*, vol. 19, no. 5, pp. 1187–1199, May 2015, <https://doi.org/10.1007/s00500-014-1332-7>.
- [31] J. Gu, L. Wang, H. Wang, and S. Wang, "A novel approach to intrusion detection using SVM ensemble with feature augmentation," *Computers & Security*, vol. 86, pp. 53–62, Sep. 2019, <https://doi.org/10.1016/j.cose.2019.05.022>.
- [32] S. Senthamarai Kannan and N. Ramaraj, "A novel hybrid feature selection via Symmetrical Uncertainty ranking based local memetic search algorithm," *Knowledge-Based Systems*, vol. 23, no. 6, pp. 580–585, Aug. 2010, <https://doi.org/10.1016/j.knsys.2010.03.016>.
- [33] P. Gogoi, M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Packet and Flow Based Network Intrusion Dataset," in *International Conference on Contemporary Computing*, Noida, India, Aug. 2012, pp. 322–334, https://doi.org/10.1007/978-3-642-32129-0_34.
- [34] M. M. Abd-Eldayem, "A proposed HTTP service based IDS," *Egyptian Informatics Journal*, vol. 15, no. 1, pp. 13–24, Mar. 2014, <https://doi.org/10.1016/j.eij.2014.01.001>.
- [35] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications*, vol. 41, no. 4, Part 2, pp. 1690–1700, Mar. 2014, <https://doi.org/10.1016/j.eswa.2013.08.066>.
- [36] "NSL-KDD Datasets," *Canadian Institute for Cybersecurity | UNB*. <https://www.unb.ca/cic/datasets/nsl.html> (accessed Apr. 23, 2021).
- [37] A. S. A. Aziz, S. E.-O. Hanafi, and A. E. Hassanien, "Comparison of classification techniques applied for network intrusion detection and classification," *Journal of Applied Logic*, vol. 24, pp. 109–118, Nov. 2017, <https://doi.org/10.1016/j.jal.2016.11.018>.

AUTHORS PROFILE



Devendra Kumar Singh got his B.E. from the B.U. University of Bhopal (MP) in 2000 and his M.E from AAI-DU Allahabad (UP) in 2006. He is currently pursuing his Ph.D. and works as an Assistant Professor in the Department of Computer Science & Engineering, SOS, E&T, GGV Bilaspur (CG) since 2005. He is a member of AIENG. He has published more than 10 research papers in reputed international journals. His main research work focuses on Intrusion Detection Systems. He has 15 years of teaching experience and 6 years of research experience.



Manish Shrivastava, is an Assistant Professor at the Department of Computer Science & Engineering, Institute of Technology, Guru Ghasidas University, Bilaspur. he obtained his M.Tech. Degree from DAVV, Indore, and Ph.D. from Guru Ghasidas University, Bilaspur. He has about 15 years of teaching and research experience, and a number of papers in various national and international journals to his credit. His field of interest is network security. He has served as a chairman of the board of studies, computer science & engineering, Guru Ghasidas, Bilaspur. He is a life member of the Indian Society for Technical Education and a senior member of IACSIT .

Classification of Dementia in EEG with a Two-Layered Feed Forward Artificial Neural Network

G. Anuradha

Department of ECE
BSA Crescent Institute of Science and Technology
Chennai, India
bhavanandhi@yahoo.com

D. Najumnessa Jamal

Department of EIE
BSA Crescent Institute of Science and Technology
Chennai, India
najumnessa.d@crescent.education

Abstract—Dementia has become a global public health issue. The current study is focused on diagnosing dementia with Electro Encephalography (EEG). The detection of the advancement of the disease is carried out by detecting the abnormal behavior in EEG measurements. Assessment and evaluation of EEG abnormalities is conducted for all the subjects in order to detect dementia. EEG feature analysis, namely dominant frequency, dominant frequency variability, and frequency prevalence, is done for abnormal and normal subjects and the results are compared. For dementia with Lewy bodies, in 85% of the epochs, the dominant frequency is present in the delta range whereas for normal subjects it lies in the alpha range. The dominant frequency variability in 75% of the epochs is above 4Hz for dementia with Lewy bodies, and in normal subjects at 72% of the epochs, the dominant frequency variability is less than 2Hz. It is observed that these features are sufficient to diagnose dementia with Lewy bodies. The classification of Lewy body dementia is done by using a feed-forward artificial neural network which proved to have a 94.4% classification accuracy. The classification with the proposed feed-forward neural network has better accuracy, sensitivity, and specificity than the already known methods.

Keywords—Lewy body dementia; EEG; dementia; neural network; dominant frequency

I. INTRODUCTION

The term dementia refers to a syndrome that causes a decline in cognitive functions, mainly the person's memory and intelligence, due to the death of brain cells. Different types of dementia have their own characteristic features. In Alzheimer's disease, the pathology involves the reduction of neurons which help in stimulating cortex and hippocampus regions, deposits in blood vessels, and formation of amyloid plaques [1]. An intermediate stage called mild cognitive impairment was found between the normal cognitive and dementia. Subjects in this stage are not demented but exhibit a decline in memory beyond that expected at their age and education [2, 3]. Mild cognitive impairment is an earlier stage, which may progress to dementia [4]. Psychometric predictors and clinical evaluation in elderly subjects can predict mild cognitive impairment [5]. Mild cognitive impairment and dementia can be detected using EEG biomarkers [6-8]. Mini Mental State Examination (MMSE), Clinical Dementia Rating (CDR), and Global Deterioration

Scale (GDS) are the subjective measures used to evaluate the early stages of mild dementia [3, 4, 11].

Alzheimer's Disease (AD) is the most common form of dementia [9, 10]. In AD patients, the decrease in consistency of fast rhythms and shift of power spectrum to lower frequencies, are the abnormalities in EEG signals [13]. The second most common dementia is Vascular Dementia (VaD) [14]. VaD is caused by the ischemic brain lesions due to cardiovascular pathology resulting in decline of the cognitive functions [15-17]. Another type of dementia, Lewy Body Dementia (LBD) is characterized by Lewy bodies with large old plaques. Fronto Temporal Dementia (FTD) might occur in middle aged people and is characterized by the degeneration of prefrontal and anterior temporal cortex associated with personality changes and changes in social conduct [18]. Parkinsonism consists of problems in motor nerves namely, resting tremor, rigidity and freezing [19].

Finding a biomarker to detect dementia at an earlier stage is a major challenge. The biomarker should detect the neuropathological processes caused by dementia to help the effective treatment of the disease [10]. EEG can function as such a biomarker to identify those physiological and pathological conditions. EEG in addition to functioning as a diagnosis tool also differentiates the stages of dementia [20]. The reduced cost and non-invasiveness of the EEG electrodes makes it suitable for detecting dementia and its progression [10, 11]. The challenges in an EEG biomarker are the accurate and early detection of the disease [6]. The identification of different stages of dementia and the precise detection is made possible by EEG [21]. Slowing of EEG dominant posterior rhythm frequency and reduced alpha and beta activities are observed in visual characteristics of AD. Reductions in mean frequency and alpha and beta power are observed in the analysis of EEG signals of AD patients, in earlier stages of dementia [4]. Mild dementia is characterized by an increase in theta activity and severe dementia by a decrease in beta activity [22, 23]. Resting state eyes-closed EEG recording is used to assess the progression from mild to severe AD. The stage of mild cognitive impairment is characterized by increase in delta power and decrease in alpha and beta power [24]. Early stages of LBD show reduction in cognitive functions like attention, problem solving, and visuospatial skills. Changes in cognitive

function and Parkinsonism are the characteristic features that differentiate DLB from AD and other dementias [25]. LBD and AD are distinguished in early stages by quantitative analysis on the EEG signal. The dominant frequency of AD patients were in alpha range, in greater than 60% of the epochs and dominant frequency variability in AD was less than 1.6Hz whereas LBD patients had dominant frequency variability greater than 4Hz [26-28].

In our present study, classification of LBD and control (normal) subjects is achieved with the help of a feed forward neural network with features extracted in the frequency domain [33]. This paper aims in finding whether these abnormalities are present in the EEG and whether these can help in the detection and classification of dementia to improve diagnostic accuracy.

II. MATERIALS AND METHODS

The proposed method consists of EEG acquisition, preprocessing, feature extraction, and classification as shown in Figure 1. EEG recording was done with subjects awake, relaxed, and with eyes closed. The sampling rate was 173.6Hz. Two-second long epochs were chosen for analysis [28, 33]. A total of 12 abnormal and 18 normal EEG records were obtained. Fast Fourier Transform (FFT), $M(K)$ was performed for the EEG time series $m(n)$. Spectral amplitude Y_j was evaluated for each epoch, with j being the number of epochs. The frequency at which Y_j is maximum, among all epochs is the dominant frequency value F_j . The Mean Dominant Frequency (MDF) of all epochs was found. Dominant Frequency Range (DFR), the range of maximum and minimum value of F_j , Dominant Frequency Variability (DFV), the variation of the dominant frequency, i.e. the difference between the maximum and minimum value of F_j , Frequency Prevalence (FP) which indicates the existence of a frequency F_j in alpha, theta, or delta range were evaluated. The above features are extracted from the preprocessed signal which is applied for classification to the feed forward neural network as normal and demented signals. In the proposed method, Levenberg algorithm is used to update the weights of the feed forward neural network.

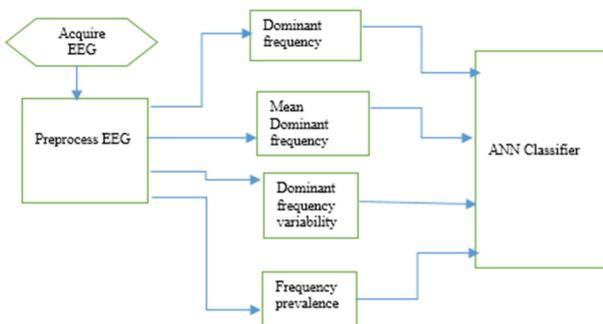


Fig. 1. Flow chart of the proposed method.

Machine learning algorithms learn from the past experience and provide optimized performance criteria. The most important aspect of a learning algorithm is its accuracy.

Statistical learning techniques are developed and compared using pattern recognition which is one of the key aspects of artificial intelligence. Linear Discriminant Analysis (LDA) and neural networks, Support Vector Machine (SVM) [29, 30], Linear Support Vector Machine (LSVM), Balanced Communication-Avoiding Support Vector Machine (BCA-SVM) [32], and Levenberga algorithm [33] are some of the available classifiers.

In the proposed method, accuracy, i.e. the differentiation of subjects with LBD from normal subjects, sensitivity which determines LBD correctly, and specificity which determines the healthy subjects correctly, were obtained with a two-layered feed forward neural network. Simulation of the above study is done using the EEG data obtained from the Bonn data base [32].

III. RESULTS AND DISCUSSION

Analysis of Dominant Frequency (DF), MDF, DFR, DFV, and FP in both normal and abnormal EEGs was conducted. Table I shows the features extracted from 12 abnormal EEGs. For every subject, the maximum DF among all epochs was chosen. It is observed that, the DF in almost all subjects, except subject 9, falls in the delta range. MDF is present in the delta range in all the analyzed subjects. The DFRs of subjects 2, 4, and 8 fall completely in the delta range. The DFV of 50% of the subjects is present in the delta and theta range. Above 80% of the subjects have FP in the delta range. From the feature analysis of abnormal subjects, it is found that, the MDF of all the epochs is found to be greater than 0.9Hz and the DFV is greater than 4Hz in 75% of epochs which confirms LBD similar to pattern 4 [27].

TABLE I. FEATURES OF LDB

S.NO	DF (Hz)	MDF (Hz)	DFR (Hz)	DFV (Hz)	FP (%)
1	1.3	1.0	0-4.4	4.0	D=90 T=10
2	3.3	0.9	0-3.3	3.0	D=100
3	0.6	1.6	0-11.5	12.0	D=90 A=10
4	0.6	1.0	0.3-2.0	2.0	D=100
5	1.3	1.6	0-5-7.0	6.0	D=80 T=20
6	0.3	1.4	0.0-11.8	12.0	D=80 A=20
7	0.3	3.5	0.3-11.1	11.0	D=73 A=27
8	1.0	1.29	0.3-3.0	3.0	D=100
9	10.8	2.9	0.0-11.8	12.0	D=80 A=20
10	0.6	2.2	0.3-11.5	12.0	D=90 A=10
11	0.339	4.0	0.0-11.5	11.5	D=58 A=42
12	0.339	4.2	0.33-12.2	11.86	D=67 A=33

D: delta 1-3.9Hz, T: theta 4-5.5Hz, A: alpha 8-12Hz

Table II shows the features extracted from normal EEGs. It is observed that, the DF in more than the 80% subjects falls in the alpha range. MDF is present in alpha range 80% of the analyzed subjects. The DF range of 83% of the subjects falls in the delta range. The DFV of 80% of the subjects is present in the delta range. Above 70% of the subjects have FP in the alpha range. Figure 2 shows the comparison of the DF for LBD and normal subjects. The DF for LBD is present in the 85% of the epochs in delta range, between 0.3Hz and 3.3 Hz. The DF of subject 9 alone falls in the alpha (10.8Hz) in LBD, whereas subjects, except 3, 15, and 16, exhibit DF in alpha, between 10Hz and 11Hz. The 85% of the epochs in normal subjects have DF in the alpha range.

TABLE II. FEATURES OF NORMAL SUBJECTS

S.NO	DF (Hz)	MDF (Hz)	DFR(Hz)	DFV(Hz)	FP (%)
1	10.5	3.2	0.0-10.5	10.5	D=75 A=25
2	11.0	11.0	11.0-12.0	1.0	A=100
3	14.0	14.0	12.0-14.0	2.0	A=100
4	11.0	11.5	11.0-12.0	1.0	D=25 A=75
5	11.0	11.5	11.0-12.0	1.0	D=100
6	10.8	11.0	10.5-11.0	1.02	A=100
7	11.0	11.29	11.0-12.0	1.0	A=100
8	11.0	11.25	11.0-12.0	1.0	D=75 A=25
9	11.0	11.13	11.0-12.0	1.0	A=100
10	11.0	11.55	11.0-12.0	1.0	A=100
11	11.0	11.61	11.0-12.0	1.0	A=100
12	11.0	11.5	11.0-12.0	1.0	D=8 A=92
13	11	10	9-11	2	D=8 A=92
14	11	11	11-12	1	D=25 A=75
15	15	13	10-16	6	A=100
16	14	14	8-15	7	A=100
17	11	11.5	11-12	1	A=100
18	11	11.5	11-12	1	A=100

D: delta 1-3.9Hz, T: theta 4-5.5Hz, A: alpha 8-12Hz

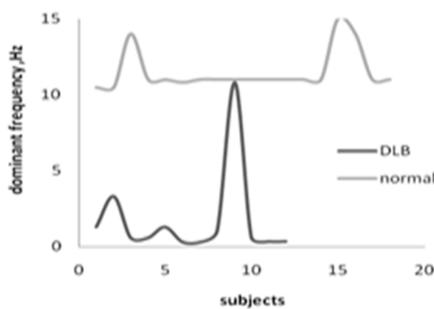


Fig. 2. DF of LBD and normal subjects.

The calculated MDF for LBD and normal subjects is shown in Figure 3. All the normal subjects, except 1, 3, 15, and 16, have MDF in alpha, between 11Hz and 12Hz, whereas in LBD, the MDF falls in the delta range of 0.9Hz and 3.5 Hz in subjects 1 to 10 and theta range of 4Hz and 4.2Hz in subjects 11 and 12 respectively.

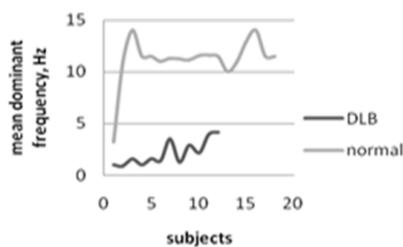


Fig. 3. MDF of LBD and normal subjects.

Figure 4 shows the comparison of DFV for subjects with LBD and normal subjects. In 75% of the epochs, DFV above 4Hz is observed in dementia and in normal subjects 83% of the epochs have DFV less than 2Hz. All LBD subjects, except 2, 4, and 8, have DFV above 4Hz. All normal subjects, except 1, 15, and 16, have DFV less than 2Hz. The calculated FP for LBD and normal subjects is shown in Figure 5. More than 80% of the epochs in LBD have FP in the delta range [24], whereas for

normal subjects 80% of the epochs are in the alpha range. It is therefore found that EEG abnormalities can function as a marker for the detection of LBD.

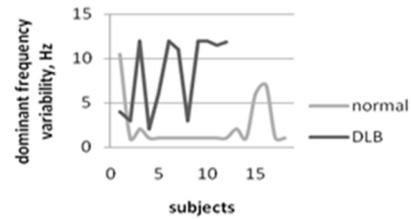


Fig. 4. DFV for LBD and normal subjects.

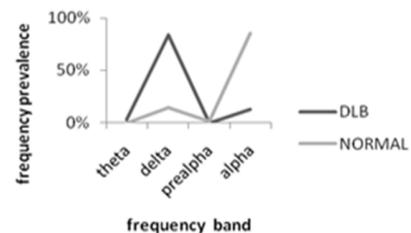


Fig. 5. FP for LBD and normal subjects.

IV. CLASSIFICATION

The extracted features DF, MDF, and DFV were utilized in a feed forward neural network implemented in MatLab. Levenberg algorithm was used to update the weights of the neural network. Accuracy, sensitivity, and specificity were used to evaluate the classification performance. These parameters are calculated as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (3)$$

where TP, TN, FP, FN stand for the true positive, true negative, false positive and false negative classifications respectively.

The confusion matrix evaluates the overall performance of a classification method. Accuracy, sensitivity, and specificity of each feature are obtained from the confusion matrix. Figure 6 shows the confusion matrix obtained for classification using the feature DFV. Accuracy of 88.9%, sensitivity of 83.3%, and specificity of 100% were achieved. Figure 7 shows the confusion matrix obtained for classification using DF, with accuracy of 94.4%, sensitivity of 90.9%, and specificity of 100% achieved. Figure 8 shows the confusion matrix obtained for MDF. The acquired accuracy, sensitivity, and specificity were 94.4%, 88.9% and 100% respectively.

Table III shows the accuracy, sensitivity, and specificity using the classifiers spectral SVM, decision tree with Fourier coefficients, and SVM with spectral peak [30, 35, 36]. The classifier performance is poor in Fourier coefficients with decision tree, where the accuracy, sensitivity and specificity are less than 75%, whereas spectral svm shows good classifier performance, with accuracy of 86%, sensitivity of 88% and

specificity of 84%. The accuracy, sensitivity, and specificity of the proposed feed forward neural network classifier with features like MDF, DFV and DF obtained from the confusion matrix is shown in Table IV.

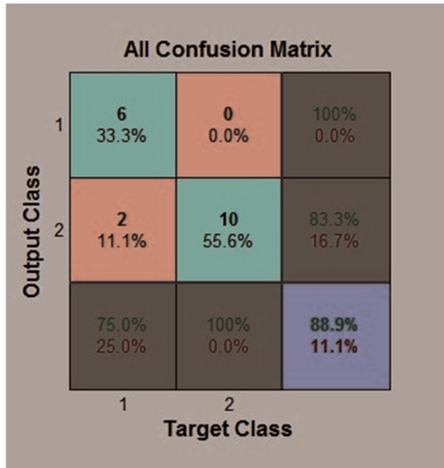


Fig. 6. Confusion matrix for DFV.

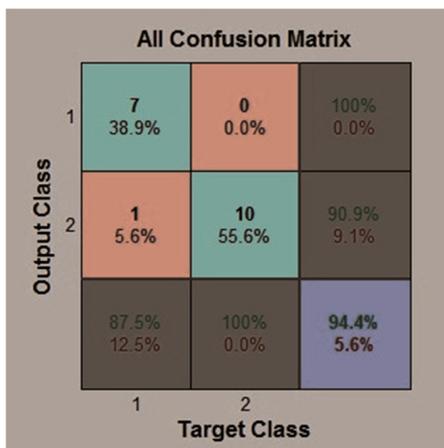


Fig. 7. Confusion matrix for DF.

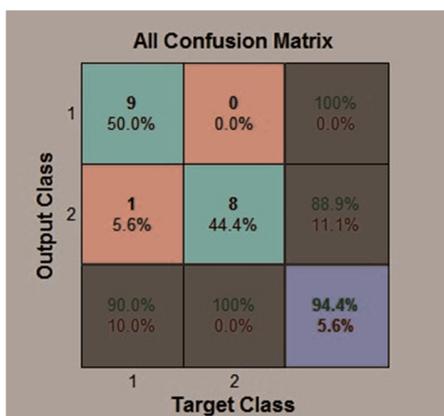


Fig. 8. Confusion matrix for MDF.

TABLE III. ACCURACY, SENSITIVITY, AND SPECIFICITY OF OTHER CLASSIFIERS

Feature	Accuracy (%)	Sensitivity (%)	Specificity (%)
Spectral SVM	86	88	84
Fourier coefficients with decision tree	72.2	72.2	59.4
Spectral peak SVM	79.9	83.2	76.4

TABLE IV. ACCURACY, SENSITIVITY, AND SPECIFICITY OF THE NN CLASSIFIER

Feature	Accuracy (%)	Sensitivity (%)	Specificity (%)
NN with MDF	94.4	88.9	100
NN with DFV	88.9	83.3	100
NN with DF	94.4	90.9	100

The highest percentage of sensitivity of the proposed classifier is 90.9% with the DF feature. The specificity obtained for the proposed classifier is 100% for MDF, DFV, and DF, while the accuracy is 94.4% in both MDF and DF. Hence, DF produces better results when compared to the other features, MDF and DFV. This type of classification method with a two-layered feed forward NN using dominant frequency feature gives better sensitivity, specificity, and accuracy when compared with other existing methods.

V. CONCLUSION

In this study, EEG analysis was performed on both abnormal and healthy subjects. Dominant frequency analysis was proposed to extract spectral features from EEG signals. The spectral features obtained are sufficient for the diagnosis of LBD [30]. An artificial neural network was used for classification. The experimental results show that the proposed two-layered feed forward artificial neural network-based classifier can achieve better classification performance than the existing LBD diagnosis methods.

REFERENCES

- [1] J. Wu, Y. Yu, and J. Yang, "Neuropsychological Parameters as Potential Biomarkers for Alzheimer's Disease," *Current Translational Geriatrics and Experimental Gerontology Reports*, vol. 1, no. 2, pp. 68–75, Jun. 2012, <https://doi.org/10.1007/s13670-012-0007-4>.
- [2] S. Fahn, "Description of Parkinson's disease as a clinical syndrome," *Annals of the New York Academy of Sciences*, vol. 991, pp. 1–14, Jun. 2003, <https://doi.org/10.1111/j.1749-6632.2003.tb07458.x>.
- [3] P. M. Rossini, S. Rossi, C. Babiloni, and J. Polich, "Clinical neurophysiology of aging brain: From normal aging to neurodegeneration," *Progress in Neurobiology*, vol. 83, no. 6, pp. 375–400, Dec. 2007, <https://doi.org/10.1016/j.pneurobio.2007.07.010>.
- [4] R. C. Petersen *et al.*, "Apolipoprotein E Status as a Predictor of the Development of Alzheimer's Disease in Memory-Impaired Individuals," *Journal of the American Medical Association*, vol. 273, no. 16, pp. 1274–1278, Apr. 1995, <https://doi.org/10.1001/jama.1995.03520400044042>.
- [5] C. Flicker, S. H. Ferris, and B. Reisberg, "Mild cognitive impairment in the elderly: Predictors of dementia," *Neurology*, vol. 41, no. 7, pp. 1006–1006, Jul. 1991, <https://doi.org/10.1212/WNL.41.7.1006>.
- [6] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 115, no. 7, pp. 1490–1505, Jul. 2004, <https://doi.org/10.1016/j.clinph.2004.01.001>.
- [7] G. Henderson *et al.*, "Development and assessment of methods for detecting dementia using the human electroencephalogram," *IEEE*

- Transactions on Biomedical Engineering*, vol. 53, no. 8, pp. 1557–1568, Aug. 2006, <https://doi.org/10.1109/TBME.2006.878067>.
- [8] Y. E. Geda, "Mild Cognitive Impairment in Older Adults," *Current Psychiatry Reports*, vol. 14, no. 4, pp. 320–327, Aug. 2012, <https://doi.org/10.1007/s11920-012-0291-x>.
- [9] M. P. Mattson, "Pathways towards and away from Alzheimer's disease," *Nature*, vol. 430, no. 7000, pp. 631–639, Aug. 2004, <https://doi.org/10.1038/nature02621>.
- [10] A. Cedazo-Minguez and B. Winblad, "Biomarkers for Alzheimer's disease and other forms of dementia: Clinical needs, limitations and future aspects," *Experimental Gerontology*, vol. 45, no. 1, pp. 5–14, Jan. 2010, <https://doi.org/10.1016/j.exger.2009.09.008>.
- [11] B. Reisberg, S. H. Ferris, M. J. de Leon, and T. Crook, "Global Deterioration Scale (GDS)," *Psychopharmacology Bulletin*, vol. 24, no. 4, pp. 661–663, 1988.
- [12] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen, "Mild Cognitive Impairment: Clinical Characterization and Outcome," *Archives of Neurology*, vol. 56, no. 3, pp. 303–308, Mar. 1999, <https://doi.org/10.1001/archneur.56.3.303>.
- [13] P. Luu, D. M. Tucker, R. Englander, A. Lockfeld, H. Lutsep, and B. Oken, "Localizing Acute Stroke-related EEG Changes: Assessing the Effects of Spatial Undersampling," *Journal of Clinical Neurophysiology*, vol. 18, no. 4, pp. 302–317, Jul. 2001.
- [14] P. D. Meek, E. K. McKeithan, and G. T. Schumock, "Economic Considerations in Alzheimer's Disease," *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, vol. 18, no. 2P2, pp. 68–73, 1998, <https://doi.org/10.1002/j.1875-9114.1998.tb03880.x>.
- [15] J. T. Coyle, D. L. Price, and M. R. DeLong, "Alzheimer's disease: a disorder of cortical cholinergic innervation," *Science*, vol. 219, no. 4589, pp. 1184–1190, Mar. 1983, <https://doi.org/10.1126/science.6338589>.
- [16] A. V. Terry and J. J. Buccafusco, "The Cholinergic Hypothesis of Age and Alzheimer's Disease-Related Cognitive Deficits: Recent Challenges and Their Implications for Novel Drug Development," *Journal of Pharmacology and Experimental Therapeutics*, vol. 306, no. 3, pp. 821–827, Sep. 2003, <https://doi.org/10.1124/jpet.102.041616>.
- [17] C. McVeigh and P. Passmore, "Vascular dementia: prevention and treatment," *Clinical Interventions in Aging*, vol. 1, no. 3, pp. 229–235, Sep. 2006.
- [18] G. C. Roman, "Vascular Dementia: Distinguishing Characteristics, Treatment, and Prevention," *Journal of the American Geriatrics Society*, vol. 51, no. 5s2, pp. S296–S304, 2003, <https://doi.org/10.1046/j.1532-5415.5155.x>.
- [19] D. R. Thal, L. T. Grinberg, and J. Attems, "Vascular dementia: Different forms of vessel disorders contribute to the development of dementia in the elderly brain," *Experimental Gerontology*, vol. 47, no. 11, pp. 816–824, Nov. 2012, <https://doi.org/10.1016/j.exger.2012.05.023>.
- [20] H. Hampel *et al.*, "Perspective on future role of biological markers in clinical therapy trials of Alzheimer's disease: A long-range point of view beyond 2020," *Biochemical Pharmacology*, vol. 88, no. 4, pp. 426–449, Apr. 2014, <https://doi.org/10.1016/j.bcp.2013.11.009>.
- [21] U. P. Mosimann and I. G. McKeith, "Dementia with lewy bodies—diagnosis and treatment," *Swiss medical weekly*, vol. 133, no. 9–10, pp. 131–142, Mar. 2003, <https://doi.org/10.7892/boris.43041>.
- [22] L. A. Coben, W. L. Danziger, and L. Berg, "Frequency analysis of the resting awake EEG in mild senile dementia of Alzheimer type," *Electroencephalography and Clinical Neurophysiology*, vol. 55, no. 4, pp. 372–380, Apr. 1983, [https://doi.org/10.1016/0013-4694\(83\)90124-4](https://doi.org/10.1016/0013-4694(83)90124-4).
- [23] L. A. Coben, W. Danziger, and M. Storandt, "A longitudinal EEG study of mild senile dementia of Alzheimer type: changes at 1 year and at 2.5 years," *Electroencephalography and Clinical Neurophysiology*, vol. 61, no. 2, pp. 101–112, Aug. 1985, [https://doi.org/10.1016/0013-4694\(85\)91048-X](https://doi.org/10.1016/0013-4694(85)91048-X).
- [24] C. Babiloni *et al.*, "Cortical sources of resting state electroencephalographic alpha rhythms deteriorate across time in subjects with amnesic mild cognitive impairment," *Neurobiology of Aging*, vol. 35, no. 1, pp. 130–142, Jan. 2014, <https://doi.org/10.1016/j.neurobiolaging.2013.06.019>.
- [25] I. G. McKeith *et al.*, "Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium for DLB international workshop," *Neurology*, vol. 47, no. 5, pp. 1113–1124, Nov. 1996, <https://doi.org/10.1212/WNL.47.5.1113>.
- [26] L. Bonanni, A. Thomas, P. Tiraboschi, B. Perfetti, S. Varanese, and M. Onofri, "EEG comparisons in early Alzheimer's disease, dementia with Lewy bodies and Parkinson's disease with dementia patients with a 2-year follow-up," *Brain*, vol. 131, no. 3, pp. 690–705, Mar. 2008, <https://doi.org/10.1093/brain/awn322>.
- [27] S.-C. Du, D.-L. Huang, and H. Wang, "An Adaptive Support Vector Machine-Based Workpiece Surface Classification System Using High-Definition Metrology," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 10, pp. 2590–2604, Oct. 2015, <https://doi.org/10.1109/TIM.2015.2418684>.
- [28] N. N. Kulkarni and V. K. Bairagi, "Extracting Salient Features for EEG-based Diagnosis of Alzheimer's Disease Using Support Vector Machine Classifier," *IETE Journal of Research*, vol. 63, no. 1, pp. 11–22, Jan. 2017, <https://doi.org/10.1080/03772063.2016.1241164>.
- [29] S. Du, C. Liu, and L. Xi, "A Selective Multiclass Support Vector Machine Ensemble Classifier for Engineering Surface Classification Using High Definition Metrology," *Journal of Manufacturing Science and Engineering*, vol. 137, Feb. 2015, Art. no. 011003, <https://doi.org/10.1115/1.4028165>.
- [30] L. Bonanni *et al.*, "Quantitative electroencephalogram utility in predicting conversion of mild cognitive impairment to dementia with Lewy bodies," *Neurobiology of Aging*, vol. 36, no. 1, pp. 434–445, Jan. 2015, <https://doi.org/10.1016/j.neurobiolaging.2014.07.009>.
- [31] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, Nov. 2001, Art. no. 061907, <https://doi.org/10.1103/PhysRevE.64.061907>.
- [32] "EEG time series download page," *Department of Epileptology, University of Bonn*. <http://epileptologie-bonn.de/cms/upload/workgroup/lehnertz/eegdata.html> (accessed Mar. 24, 2021).
- [33] G. Anuradha, N. Jamal, and S. Rafiammal, "Detection of dementia in EEG signal using dominant frequency analysis," in *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering*, Chennai, India, Sep. 2017, pp. 710–714, <https://doi.org/10.1109/ICPSCI.2017.8391806>.
- [34] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, Nov. 1994, <https://doi.org/10.1109/72.329697>.
- [35] G. Fisco *et al.*, "Combining EEG signal processing with supervised methods for Alzheimer's patients classification," *BMC Medical Informatics and Decision Making*, vol. 18, no. 1, May 2018, Art. no. 35, <https://doi.org/10.1186/s12911-018-0613-y>.
- [36] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. M. Kanda, R. Anghinah, and R. Nitri, "Improving Alzheimer's Disease Diagnosis with Machine Learning Techniques," *Clinical EEG and Neuroscience*, vol. 42, no. 3, pp. 160–165, Jul. 2011, <https://doi.org/10.1177/155005941104200304>.
- [37] S. S. Rafiammal, D. N. Jamal, and S. K. Mohideen, "Reconfigurable Hardware Design for Automatic Epilepsy Seizure Detection using EEG Signals," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5803–5807, Jun. 2020, <https://doi.org/10.48084/etasr.3419>.
- [38] 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>.

A Holistic Approach to Urdu Language Word Recognition using Deep Neural Networks

Hashim Raza Khan

Department of Electronic Engineering
NED University of Engineering and Technology
Karachi, Pakistan
hashim@neduet.edu.pk

Majida Kazmi

Dpt. of Computer and Information Systems Engineering
NED University of Engineering and Technology
Karachi, Pakistan
majidakazmi@neduet.edu.pk

Hamza Khalid

Dpt. of Computer and Information Systems Engineering
NED University of Engineering and Technology
Karachi, Pakistan
hamza@neduet.edu.pk

Muhammad Abul Hasan

Department of Bio-Medical Engineering
NED University of Engineering and Technology
Karachi, Pakistan
abulhasan@neduet.edu.pk

Nabeel Fayyaz

Department of Electrical Engineering
NED University of Engineering and Technology
Karachi, Pakistan
nabeelfayyaz@neduet.edu.pk

Saad Ahmed Qazi

Dpt. of Computer and Information Systems Engineering
NED University of Engineering and Technology
Karachi, Pakistan
saadqazi@neduet.edu.pk

Abstract-Urdu is one of the most popular languages in the world. It is a Persianized standard register of the Hindi language with considerable and valuable literature. While digital libraries are constantly replacing conventional libraries, a vast amount of Urdu literature is still handwritten. Digitizing this handwritten literature is essential to preserve it and make it more accessible. Nevertheless, the scarcity of Urdu Optical Character Recognition (OCR) research limits a digital library's scope to a manual document search. The limited research work in this area is mainly due to the complexity of Urdu Script. Unlike the English language, the Urdu writing style is cursive, bidirectional, and character shapes and sizes highly vary depending on their position. Holistic word recognition is found to be a better solution among many other text segmentation techniques as it takes the complete word into account instead of segmenting it explicitly or implicitly. For this project, the data of five different Urdu words were collected for training and testing a convolutional neural network and 96% recognition accuracy was achieved.

Keywords-word recognition; Urdu; deep learning; CNN; cursive writing

I. INTRODUCTION

The Urdu language is an Indo-European language derived from the Farsi alphabet, which stems from the Arabic alphabet [1]. Having more than a hundred (100) million native speakers in the South Asian region, Urdu is one of the two official languages of Pakistan. Similar to Arabic, the Urdu language is bidirectional. However, Urdu has more isolated letters (38) than Persian (32) and Arabic (28). A great deal of history,

literature, and lore of the South Asian region is present in the form of handwritten manuscripts in Urdu. To make them accessible to larger audiences through electronic media, the only possible solution is the digitization of the literature in searchable form. An obvious solution is manually sorting/annotation, which is a tedious and time-consuming task [2]. Usually, if we want to bring any hard copy document in e-format, scanning is performed, bringing the data in image format. Searching keywords from that image is not directly possible in Urdu, Persian, and Arabic. This problem can be addressed with the help of Optical Character Recognition (OCR) systems. OCR is defined as a process of classification of optical patterns in a digital image or as the conversion of images of typed, handwritten, or printed text into machine-encoded text [3]. Unlike English OCRs that first hit the market in the early 90s [4, 5], the earliest research conducted in Urdu OCRs was reported in 2003 [6]. The complications associated with Urdu script are:

- Urdu is a bidirectional language. It is written from right to left, and in some cases, such as numbers, it follows the left to the right direction.
- Each Urdu letter can take four distinct shapes due to four different places, i.e. initial, middle, final, and standalone, as shown in Figure 1 [7].
- Urdu is always written cursively. Inter- and intra- word spaces are present in the Urdu script.

Corresponding author: Hashim Raza Khan

- Words sometimes overlap each other and make the recognition process difficult.
- Some Urdu characters have dots associated with them which can vary in numbers and locations.
- Sometimes in handwritten documents, open loops may get closed [8].

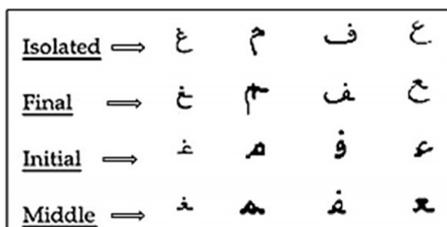


Fig. 1. Character variation depending on the position.

Research has been done on similar problem types, as mentioned above, in numerous languages, and text recognition has been experimented on text images using OCR [9]. OCR technique is used for converting images into editable text. However, for cursive scripts like Urdu, results are not satisfactory as isolated scripts. Hence, cursive scripts' pattern recognition is challenging [10]. Moreover, since Arabic and Persian languages' characters are subsets of Urdu language, it is practically impossible to use Arabic/Persian OCRs for Urdu language. On the other hand, the general problem with OCRs is that the text that occurs in natural scene images, historical documents, or images hugely varies in appearance and layout or images with background objects cause false-positive detections [11]. Due to these complications, converting an image with cursive text into editable text is not flawless, and accuracy is not satisfactory. Word (keyword) spotting is another technique that can be defined as a process of finding all instances of a query word that exist in a document image without wholly recognizing the document. In the past years, research in handwriting recognition with OCRs has developed to a level that generates commercial applications. However, the success rate is much higher in online than offline recognition [9]. Hence, in cases where the OCR system cannot assure human reader satisfaction, word spotting is a viable solution for information retrieval [2].

Over the last few decades, there has been rapid development in text spotting methods. The majority of methods split the complete process into two major stages [12]: Segmentation/text detection and word recognition. Segmentation/text detection generates candidate character or word bounding boxes/region proposals [3]. It can be defined as decomposing an image into many sub-images, each containing a character/word. Text detection tackles the standard text spotting pipeline and contributes to the whole process's error rate [12]. Several segmentation strategies can be broadly classified into three main categories: implicit segmentation, explicit segmentation, and holistic approach. Implicit segmentation (recognition-based segmentation) is a technique in which the system searches the images for components that match classes in its alphabet. However, this technique is

adopted to integrate the processes of segmentation and recognition. Accordingly, Hidden Markov Models (HMMs) based approaches have emerged, and authors in [13] proposed a method based on HMMs using an implicit segmentation approach and claimed word recognition accuracy of 88.2% on a lexicon of size 3,771. In explicit segmentation (classical approach), an input image is first partitioned into sub-images containing characters and then it is classified. Explicit segmentation needs to find all the interconnection between words, also known as ligatures, and dissect the word images through all detected ligatures. Authors in [14] proposed a direct segmentation approach for offline handwriting recognition. The accuracy of the segmentation approach was reported at 83.6%. However, over-segmentation and wrong segmentation are considerably high [15].

A holistic method tries to identify words instead of characters and considers them as units. It is therefore termed as Intelligible Word Recognition (IWR) instead of a character recognition technique. There are other significant reasons for its popularity as well: the holistic approach is parallel to human reading and can be applied to historical documents whose poor quality makes it challenging to analyze with the classical approach [4]. Due to the cursive handwriting style, changing the shape of characters depending on the Urdu script's position, the classification using a holistic approach is more efficient and optimal [1]. Besides these advantages, its major drawback is that the holistic approach is always restricted to a predefined lexicon since it does not deal directly with letters, but with words, so a specific lexicon of the word is a necessary constraint. Whenever we want to modify the lexicon, the training stage is mandatory. This property makes this approach suitable where the lexicon is statistically defined [15]. After applying machine learning techniques in the image recognition domain, the lexicon size can be made quite large, making this technique suitable for many applications. A lexicon of size of 90,000 words has been claimed in recent works of word recognition [7]. A typical optical character recognition pipeline starts with looking at the picture, locating text areas, identifying the lines of the text, and at the end, it tries to recognize each character or word present in that image. The text recognition step receives a single word or character's cropped image and identifies the depicted object. Character-based classifiers are used to classify individual characters, which are then integrated to generate an entire word. Different techniques are used for this classification task. Authors in [16, 17] use Convolutional Neural Networks (CNN) as character classifiers. Authors in [17] used a complicated combination of CNNs and HMMs for a fixed lexicon to generate the final recognition results. An alternate approach for this text detection task is whole word-based recognition or holistic word recognition, extracting features from the entire word before performing word classification [7].

Since we cannot use Arabic and Persian language recognition systems, our primary focus is to develop a word recognition system using a deep neural network for the Urdu language. To that end, we have used a holistic or whole word-based recognition approach. We have collected around 5298 samples of five Urdu words: tawhid, namaz, hajj, rehmat, and jannat for our project.

II. DATA PREPARATION

A. Training and Testing Data

Data were collected from different people, with a diversity of writing styles, containing 5298 samples of 5 different words (approximately every word is written around 1000 times). Seventy five percent of the samples were selected for training, and 25% were set for testing purposes.

B. Data Pre-processing

Recognition results rely on the preprocessing phase. Unwanted noise can degrade the recognition accuracy. Our work in the preprocessing stages involves binarization and converting of the cropped images into a uniform size while creating the dataset. For data collection, we have used A4-sized papers with 26 rows and 10 columns. These papers were scanned at 300 dpi, and then individual characters were cropped using an open-source OCR helper tool. These cropped images were saved in binarized format in a text file, as shown in Figure 2. This text file has three channels (RGB), which carries redundant information, so information from only one channel was extracted, and then these arrays were concatenated to form a complete dataset. The OCR helper tool facilitates the dataset preparation process by providing an adjustable binarization threshold, height and width sensitivity, and the export size of the blob.



Fig. 2. OCR helper tool for cropping and binarization.

III. THE DEEP LEARNING MODEL

A. Model Architecture

Neural networks are typically organized in layers, consisting of many interconnected nodes known as neurons. Generally, there are 3 types of layers: input, hidden, and output layers. The input layer receives the input pattern and communicates it to the next (hidden) layer. A hidden layer is an intermediate layer between the input and the output layers. The processing is performed here via weighted connections. The last output layer of the neural network usually consists of n nodes, where n is the total number of classes. A Convolutional Neural Network (CNN) is one of the most common types of artificial neural networks and it is considered a state-of-the-art technique in pattern recognition, outperforming several classifiers [18]. The CNN's architecture contains convolutional layers, pooling layers, and a fully connected layer. A CNN finds its application in a wide range of areas, e.g. natural language processing, video analysis, speech recognition, and pattern recognition [19].

The convolutional layer is the first layer in a CNN. This layer uses a number of n different filters or kernels (edge or

curve detectors). Each of these can be intuitively modeled as a feature identifier. These filters then convolve with the original image's sub-images according to their dimension and extract features from that image. To understand the process of convolution, consider a 6×6 image and a 3×3 filter. To extract features, we must convolve them. This operation is shown in Figure 3.

$$\begin{array}{|c|c|c|c|c|c|} \hline 3 & 0 & 1 & 2 & 7 & 4 \\ \hline 4 & 5 & 6 & 5 & 3 & 2 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \end{array} * \begin{array}{ccc} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{array} = \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \\ X_5 & X_6 & X_7 & X_8 \\ X_9 & X_{10} & X_{11} & X_{12} \\ X_{13} & X_{14} & X_{15} & X_{16} \end{array}$$

Fig. 3. A simple convolution operation.

The values of the output can be easily calculated as follows:

$$X_1 = 3*1 + 4*1 + 1*1 + 0*0 + 0*8 + 0*2 + 1*-1 + 6*-1 + 3*-1 = -2$$

Output dimensions can be found as: $n-f+1 * n-f+1$

where n is the input image dimension (6 in this case) and f is filter's dimension (3 in this case). The output dimensions will be 4×4 .

Pooling or subsampling layer is used to decrease the features' resolution, making the feature more robust against noise and distortion and making the system computationally effective. A fully connected layer is often used as the last layer in a CNN. The working principle of a fully connected layer is the same as traditional multi-layer perceptrons. Every neuron present in this layer receives input from every neuron of the previous layer and finally classifies the output.

B. Model Training and Testing

To develop an Urdu handwritten word recognition system, we created a deep feed-forward neural network architecture using the TensorFlow library. We started our experiment with 5 different word classes, each having approximately 1000 samples. We trained our network on a non-GPU machine with varying image sizes, i.e. (60×60 , 50×50 , and 30×30 pixels) and different numbers of neurons and got meaningful results comparable with [20].

C. Results

We trained our model for 5 different words and tested it for accuracy and loss by changing the number of neurons and input image sizes. Increasing the number of neurons resulted in better accuracy but increased training time. The highest accuracy achieved was 95.81%, with 60×60 image size and 2000 neurons in each layer. The words used to train the model were tawhid, namaz, hajj, rehmat, and jannat which were labeled as 0,1,2,3, and 4 respectively. Figure 4 shows the random inferences from the test set, where 'Tr' shows the original word and 'Pred' shows the predicted output.

1) 30×30 Pixel Images

The accuracies of individual words and of the complete development set on 30×30 images with different number of neurons is shown in Table I. Figure 5 shows the loss decrease and the accuracy increase with the number of epochs. Figure 6 shows the individual accuracy of each word with a CNN containing 500 neurons.

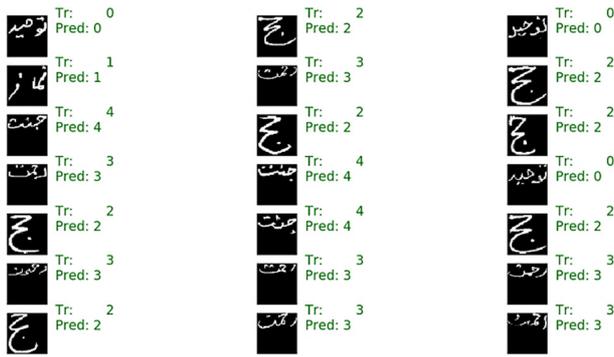


Fig. 4. Prediction on the test set with randomly selected words.

TABLE I. ACCURACY ON 30×30 IMAGES

Neurons	Accuracy (%)					
	Test set	Tawhid	Namaz	Haji	Rehmat	Jannat
450	88.36	91.63	79.15	94.59	99.97	76.53
500	90.22	88.44	84.55	93.82	99.90	84.23
600	89.68	88.84	89.57	91.50	99.68	78.46
300	86.03	80.87	87.64	94.59	99.61	67.30
350	89.21	86.05	86.48	94.20	98.46	80.76

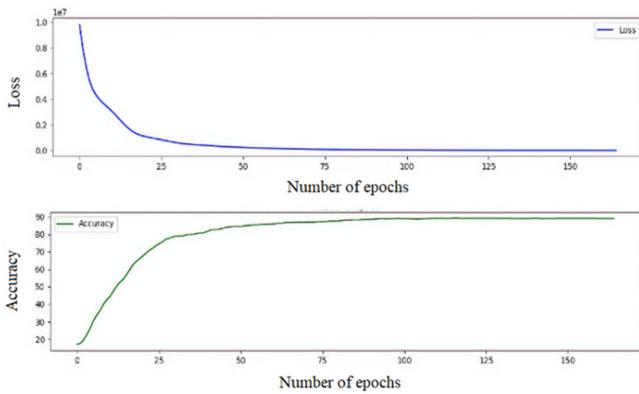


Fig. 5. Accuracy and loss with 500 neurons.

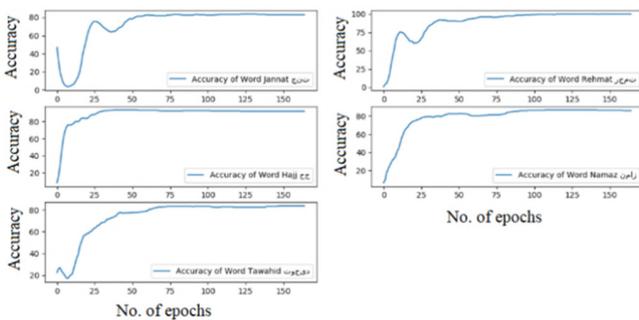


Fig. 6. Individual accuracies with 500 neurons.

2) 50×50 Pixel Images

The accuracies of individual words and of the complete development set on images of size 50×50 with different number of neurons is shown in Table II. Figure 7 shows the loss decrease and the accuracy increase with the number of epochs. Figure 8 shows the individual accuracy of each word with a CNN containing 1250 neurons.

TABLE II. ACCURACY ON 50×50 IMAGES

Neurons	Accuracy (%)					
	Test set	Tawhid	Namaz	Haji	Rehmat	Jannat
1250	94.33	94.82	93.05	97.29	97.66	86.53
1400	94.25	97.61	94.59	98.45	97.50	80.76
1000	94.24	94.02	93.05	92.66	97.46	81.53
800	91.93	92.82	92.27	96.91	97.21	77.69

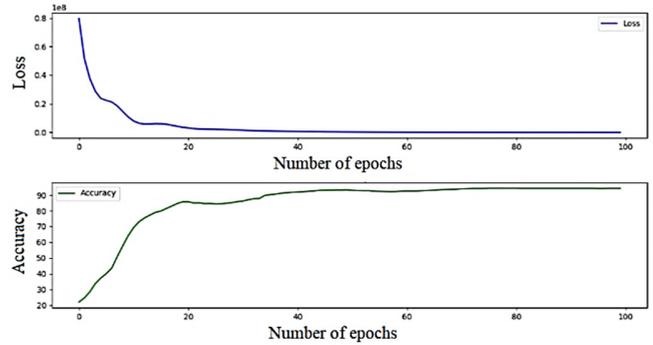


Fig. 7. Accuracy and loss of a CNN with 1250 neurons.

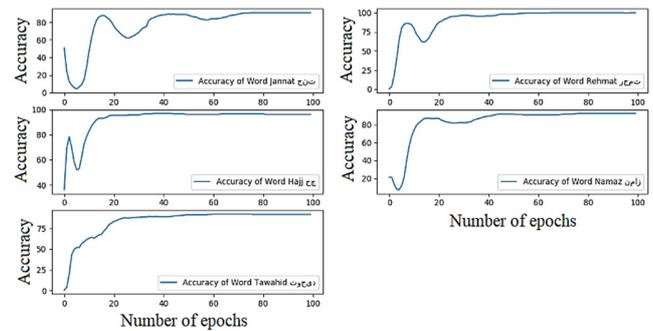


Fig. 8. Individual accuracies with 1250 neurons.

3) 60×60 Pixel Images

The accuracies of individual words and of the complete development set on images of size 60×60 with different number of neurons is shown in Table III. Figure 9 shows the loss decrease and the accuracy increase with the number of epochs. Figure 10 shows the individual accuracy of each word with a CNN containing 2000 neurons.

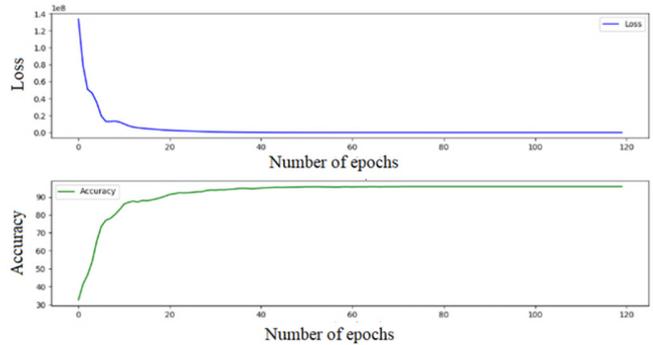


Fig. 9. Accuracy and loss with 2000 neurons.

TABLE III. ACCURACY ON 60×60 IMAGES

Neurons	Accuracy (%)					
	Test set	Tawhid	Namaz	Haji	Rehmat	Jannat
400	87.82	80.87	87.25	94.98	98.60	75.76
800	92.63	93.22	88.03	97.68	98.94	84.23
1600	94.33	90.04	95.75	96.13	99.32	89.61
2000	95.81	95.21	96.52	98.45	99.61	88.84
2400	94.49	95.61	93.82	93.82	99.59	89.23

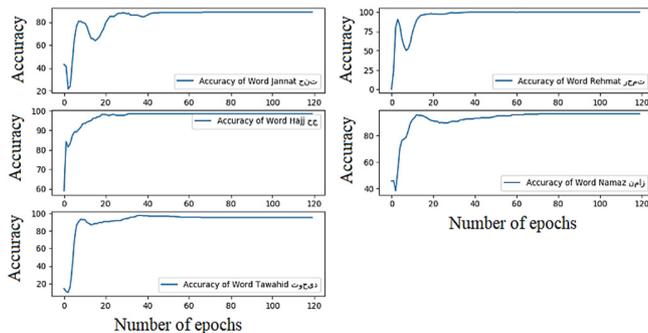


Fig. 10. Individual accuracies with 2000 neurons.

IV. CONCLUSION

This report describes a system for handwritten Urdu holistic word recognition using a deep neural network. This work's motivation was the existence of extensive valuable literature in the Urdu language in many South Asian libraries and the ongoing effort to digitalize these scripts. Research work has already been conducted in this area on different languages, but the digitization of Urdu script is a scarcely developed area. Some work has been done for Urdu word/character recognition using SVMs and HMMs.

In this work, a holistic Urdu handwritten word recognition has been developed using a deep neural network. We have collected more than 5000 handwritten samples to include a diversity of styles in our study. Approximately every word has 1000 samples. The images were saved in binarized format and the dataset was divided to training (75%) and testing (25%) subsets. A feed-forward deep neural network was trained for recognition/classification, and showed promising results. We have achieved accuracy up to 96% on the complete test set.

V. FUTURE RECOMMENDATIONS

Many different experiments, tests, and architectures can be tried in the future, including:

- Training neural networks for a larger lexicon of Urdu words.
- Using different architectures and configurations of neural networks for improved accuracy.
- Segmentation of Urdu words in handwritten script.
- Developing a complete end-to-end system of keyword search in handwritten Urdu script.

REFERENCES

- [1] M. W. Sagheer, C. L. He, N. Nobile, and C. Y. Suen, "Holistic Urdu Handwritten Word Recognition Using Support Vector Machine," in *2010 20th International Conference on Pattern Recognition*, Aug. 2010, pp. 1900–1903, <https://doi.org/10.1109/ICPR.2010.468>.
- [2] A. Abidi, A. Jamil, I. Siddiqi, and K. Khurshid, "Word Spotting Based Retrieval of Urdu Handwritten Documents," in *2012 International Conference on Frontiers in Handwriting Recognition*, Bari, Italy, Sep. 2012, pp. 331–336, <https://doi.org/10.1109/ICFHR.2012.289>.
- [3] H. Bunke and P. S. P. Wang, *Handbook of Character Recognition and Document Image Analysis*. World Scientific, 1997, <https://doi.org/10.1142/2757>.
- [4] A. Chaudhuri, K. Mandaviya, P. Badelia, and S. K. Ghosh, *Optical Character Recognition Systems for Different Languages with Soft Computing*. Springer International Publishing, 2017.
- [5] N. H. Khan and A. Adnan, "Urdu Optical Character Recognition Systems: Present Contributions and Future Directions," *IEEE Access*, vol. 6, pp. 46019–46046, 2018, <https://doi.org/10.1109/ACCESS.2018.2865532>.
- [6] U. Pal and A. Sarkar, "Recognition of Printed Urdu Script," in *Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 2*, USA, Aug. 2003, pp. 1183–1187.
- [7] Q. U. A. Akram and S. Hussain, "Improving Urdu Recognition Using Character-Based Artistic Features of Nastaliq Calligraphy," *IEEE Access*, vol. 7, pp. 8495–8507, 2019, <https://doi.org/10.1109/ACCESS.2018.2887103>.
- [8] Z. Ahmad, J. K. Orakzai, and I. Shamsheer, "Urdu compound Character Recognition using feed forward neural networks," in *2009 2nd IEEE International Conference on Computer Science and Information Technology*, Beijing, China, Aug. 2009, pp. 457–462, <https://doi.org/10.1109/ICCSIT.2009.5234683>.
- [9] V. Lavrenko, T. M. Rath, and R. Manmatha, "Holistic word recognition for handwritten historical documents," in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, Jan. 2004, pp. 278–287, <https://doi.org/10.1109/DIAL.2004.1263256>.
- [10] N. P. T. Kishna and S. Francis, "Intelligent tool for Malayalam cursive handwritten character recognition using artificial neural network and Hidden Markov Model," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*, Coimbatore, India, Nov. 2017, pp. 595–598, <https://doi.org/10.1109/ICICI.2017.8365201>.
- [11] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Reading Text in the Wild with Convolutional Neural Networks," *International Journal of Computer Vision*, vol. 116, no. 1, pp. 1–20, Jan. 2016, <https://doi.org/10.1007/s11263-015-0823-z>.
- [12] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, Washington, DC, USA, Jun. 2004, vol. 2, <https://doi.org/10.1109/CVPR.2004.1315187>.
- [13] P. R. Cavalin, A. de Souza Britto, F. Bortolozzi, R. Sabourin, and L. E. S. Oliveira, "An implicit segmentation-based method for recognition of handwritten strings of characters," in *Proceedings of the 2006 ACM symposium on Applied computing*, New York, NY, USA, Apr. 2006, pp. 836–840, <https://doi.org/10.1145/1141277.1141468>.
- [14] M. Ghosh, R. Ghosh, and B. Verma, "A fully automated offline handwriting recognition system incorporating rule based neural network validated segmentation and hybrid neural network classifier," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, no. 7, pp. 1267–1283, Nov. 2004, <https://doi.org/10.1142/S0218001404003654>.
- [15] N. Sivashanmugam, "A Study of Various Segmentation Techniques for Cursive Handwritten Words Recognition," *International Journal of Modern Trends in Engineering and Research*, 2015.
- [16] O. Alsharif and J. Pineau, "End-to-End Text Recognition with Hybrid HMM Maxout Models," *arXiv:1310.1811 [cs]*, Oct. 2013, Accessed: Apr. 23, 2021. [Online]. Available: <http://arxiv.org/abs/1310.1811>.

- [17] M. Jaderberg, A. Vedaldi, and A. Zisserman, "Speeding up Convolutional Neural Networks with Low Rank Expansions," *arXiv:1405.3866 [cs]*, May 2014, Accessed: Apr. 23, 2021. [Online]. Available: <http://arxiv.org/abs/1405.3866>.
- [18] S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition." Cadence Design Systems Inc, 2015.
- [19] S. Sahel, M. Alsahafi, M. Alghamdi, and T. Alsubait, "Logo Detection Using Deep Learning with Pretrained CNN Models," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6724–6729, Feb. 2021, <https://doi.org/10.48084/etasr.3919>.
- [20] 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>.

LVRT Enhancement of a Grid-tied PMSG-based Wind Farm using Static VAR Compensator

Shoaib Ahmed Dayo

Department of Electrical Engineering
Mehran University of Engineering and Technology
SZAB Campus, Khairpur Mirs, Pakistan
shoaibahmed-dayo@muetkhp.edu.pk

Mohammad Aslam Uqaili

Department of Electrical Engineering
Mehran University of Engineering and Technology
Jamshoro, Pakistan
aslam.uqaili@faculty.muet.edu.pk

Shafqat Hussain Memon

Department of Electrical Engineering
Mehran University of Engineering and Technology
SZAB Campus, Khairpur Mirs, Pakistan
shafqatmemon@muetkhp.edu.pk

Zubair Ahmed Memon

Department of Electrical Engineering
Mehran University of Engineering and Technology
Jamshoro, Pakistan
zubair.memon@faculty.muet.edu.pk

Abstract-This paper presents an efficient Low Voltage Ride Through (LVRT) control scheme for a 10.0MW grid-tied Permanent Magnet Synchronous Generator (PMSG)-based wind farm. The proposed control strategy plans to enhance the power quality and amount of injected power to satisfy the grid code requirements. The proposed approach utilizes a static Shunt Var Compensator (SVC) to enhance the LVRT capability and to improve power quality. It has been observed from the outcomes of the study that the proposed SVC controller ensures safe and reliable operation of the considered PMSG-based power system. The proposed system not only improves power quality but also it provides voltage stability of the Wind Energy Conversion System (WECS) under abnormal/fault conditions. The results show the superiority of the proposed control strategy.

Keywords-LVRT; PMSG; SVC; WECS; power quality; voltage stability

I. INTRODUCTION

The demand for renewable energy sources that produce clean energy with low carbon footprint is increasing. Fossil fuel consumption emits CO₂ resulting in enhanced environmental pollution. Research and development in the energy sector have boosted the extension of renewable energy resources like wind, hydro, and solar [1, 2]. Wind power generation is one of the most promising contributors to the usage of renewable energy [3, 4]. The wind generation potential integrated with the grid is increasing at an optimal level in electricity production fulfilling the energy gap [5] to an extent. Past studies have considered the disconnections of wind farms from the grid in the case of fault occurrence [6]. However, the use of modern grid codes is a prominent solution that ensures the grid system's operational capability even during the occurrence of faults [6-10]. Among multiple types of wind turbines, Permanent Magnet Synchronous Generators (PMSGs) [11] and Double Fed Induction Generators (DFIGs) are widely used in wind farms [1, 5, 12-14]. Development in semiconductor switching devices

and high efficiency demand the extension of the use of PMSGs [5, 8, 14]. Authors in [8] explored the LVRT capability enhancement along with the controlled supply of reactive power. Authors in [2] compared FACTS Static Var Compensator (SVC) controller and STATCOM regarding the provision of faster dynamic response to the wind system to control and enhance the LVRT. The SVC was installed at the end of the transmission line to enhance the LVRT control capability during a fault condition in order to maintain supply continuity. SVC was mostly used when voltage swings were the greatest.

The conventional back-to-back converter control strategy of Rotor Side Converter (RSC) was introduced for Maximum Power Point Tracking (MPPT) [15] and Grid Side Converter (GSC) was implemented to regulate the DC link voltage resulting from the grid fault occurrence [7, 8, 16]. When an LVRT occurs, the DC link voltage can be increased because MSC does not sense the grid fault and GSC fails to control the DC link voltage [5, 6, 17, 18] which results in the tripping of the wind farm. The use of a non-linear, SVC controller is rapidly growing because of its robustness with respect to external disturbances. The prime issue of controllers using traditional back-to-back converters is blathering. To reduce blathering, various LVRT methods have been introduced [1-5, 17-19]. The most attractive LVRT method is the SVC because of shunt connection and longer life cycle [2]. Moreover, many authors have proposed different control techniques to enhance the LVRT capability of wind farms [18]. The SVC controller is robust and cost-effective and can be implemented in wind farms. During fault conditions, SVC is capable of enhancing the LVRT performance by supplying the reactive power and thus improve voltage stability [1, 5, 6, 17]. According to the National Electric Power Regulatory Authority (NEPRA) standard, the wind system must be able to withstand voltage sags of around 30% of rated voltage for durations of 80ms [9].

Corresponding author: Shoaib Ahmed Dayo

In this paper, the conventional SVC controller is employed; the LVRT features have been enhanced and maintain voltage supply continuously. The improvement in the wind farm performance has been analyzed. The proposed model design achieves excellent power quality, optimal voltage stability, and reasonable reactive power compensation. The main contributions of this work are:

- The study provides detailed modeling and analysis of a 10.0MW, 120kV grid-connected PMSG wind farm.
- The SVC controller is employed to compensate for reactive power and to enhance the LVRT capability of the developed PMSG-based wind farm during normal and abnormal conditions.
- The proposed system provides uninterrupted power supply to the grid once the fault is removed, thus validating the effectiveness of the proposed control strategy.

II. SYSTEM MODEL AND CONTROL STRATEGY

Figure 1 shows the model layout and the control strategy of the proposed system. The wind farm is designed with 5 wind turbines, each with a rating of 2.0MW. The model design comprises of a PMSG based grid integrated wind farm along with an SVC controller. The 10.0MW wind farm has been integrated to a 120kV grid. The proposed model is designed at 50Hz to validate the performance of the system in LVRT conditions. Further, a variable speed PMSG machine is used because of its cost-effectiveness and high efficiency. During the LVRT condition, an SVC controller is installed at the Point of Common Coupling (PCC) to maintain the supply continuity. The bus voltage at PCC is 575V. As per the NEPRA standard, if a voltage dip occurs, the system must not drop below 30% of the rated voltage. The designed control strategy of the SVC controller comprises of a voltage measurement unit, a regulator, a distribution unit, and a synchronizing unit. The voltage measurement module measures the output voltage of the system. It can also be used for filtering purposes. To achieve a smooth voltage response, the voltage regulator is used to control and regulate the voltage. The distribution unit is used to control the thyristor susceptance module and is interfaced with the power system. It also determines the number of Thyristor Switched Capacitor (TSC) units and the absorption level of Thyristor-Controlled Reactor (TCR) power.

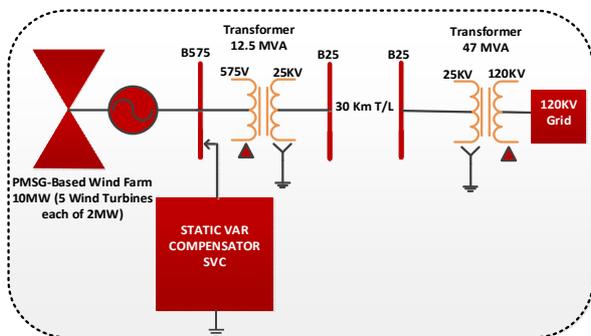


Fig. 1. The PMSG based grid integrated wind farm.

A synchronizing unit contains a Phase-Locked Loop (PLL) circuit which is synchronized against secondary voltages. A pulse generator produces the pulses to energize the thyristors. The SCR starts conducting at a particular point of the voltage known as the firing angle. Lowering the firing angle increases output power. It can be observed that the installation of an SVC controller results in LVRT operation enhancement and as a result, the system performance is improved.

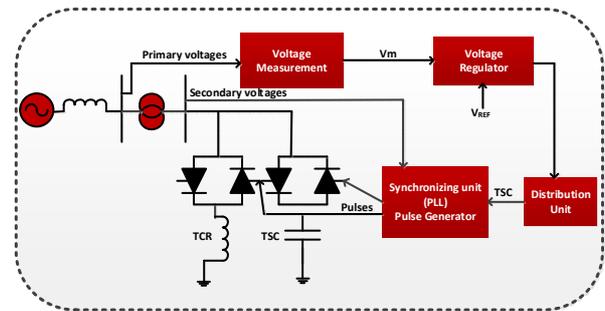


Fig. 2. The designed strategy of the SVC controller.

The result shows the voltage profile is maintained at 0.75p.u. during the LVRT condition. Hence, the continuity of supply is ensured. Power quality and voltage stability during disturbances are also improved. The proposed model was designed and simulated in MATLAB/ Simulink version 2016b. The parameters of the proposed model are listed in Table I.

TABLE I. PROPOSED MODEL SYSTEM PARAMETERS

Parameter name	Value	Parameter name	Value
C_p (max)	11.0	P_n (VA)	2000000/0.9
ρ (K_g/m^3)	1.225	f_n (Hz)	50
ω_m (rad/s)	0.3	R_s (p.u.)	0.006
H (s)	4.32	V_{dc}	1100
K_p	1.1	K_i	27.5
C (nF)	250	L (mH)	1.13
R_s	4.26	Quality factor	50

III. SIMULATED RESULTS ANALYSIS

In this section, the performance analysis of the proposed system at different stages is evaluated. The steady-state and transient responses during fault conditions, such as SLG and L-L fault, are analyzed in the simulation results. Furthermore, the outcomes of multiple fault conditions on voltage stability and reactive power compensation are also taken into consideration. The Grid-tied PMSG-based wind farm model in MATLAB/Simulink is given in Figure 3. The model consists of a 10MW Wind farm with 1 TCR and 3 TSCs. The rating of TCR is 109Mvar while that of each TSC unit is 94Mvar.

A. System Response at Steady State Condition

This section investigates the operation and control of the proposed model where the voltage is set at 1.0p.u. with generator rating of 5×2 MW and $V=575$ V at B575. The active generated power increases steadily and reaches the nominal value of 10MW as shown in Figure 4. The simulation results confirm that the proposed system is running smoothly and shows a good performance during the steady-state condition.

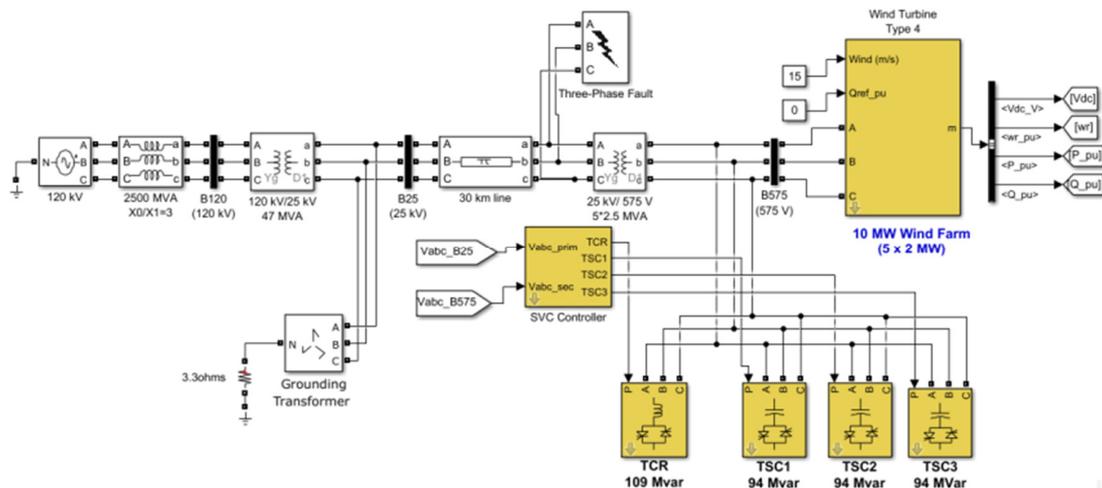


Fig. 3. MATLAB/Simulink model of the proposed system.

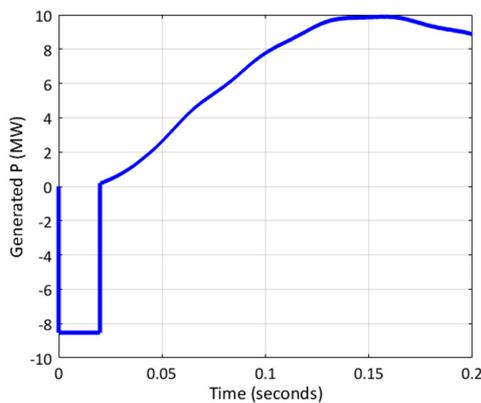


Fig. 4. Active power generation against time.

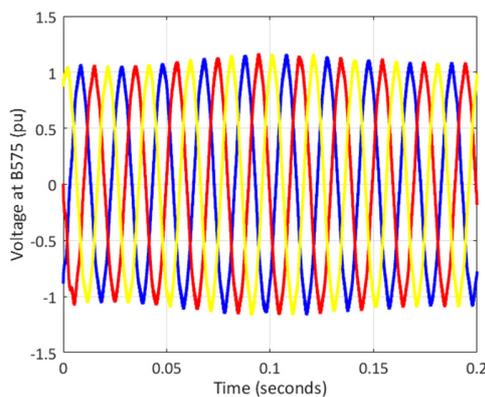


Fig. 5. Voltage response at PCC B575.

The simulation results confirm that the voltage is maintained at 1.0p.u., when the stop time is set at 0.2s. This result is obtained by running iterative rigorous simulations. Figure 5 shows that the optimized value is obtained at 1.0p.u.. During the steady-state condition, the current is not increased beyond the limit. However, from simulation results, it can be seen that the current increases and reaches 0.81p.u. as shown in

Figure 5. This shows an excellent performance and proves that the proposed system runs smoothly without any disturbance. The speed turbine starts from 1.0p.u. and reaches up to 1.027p.u. as demonstrated in Figure 6.

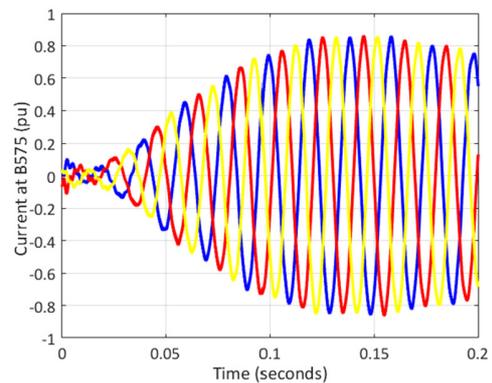


Fig. 6. Generated current at PCC B575.

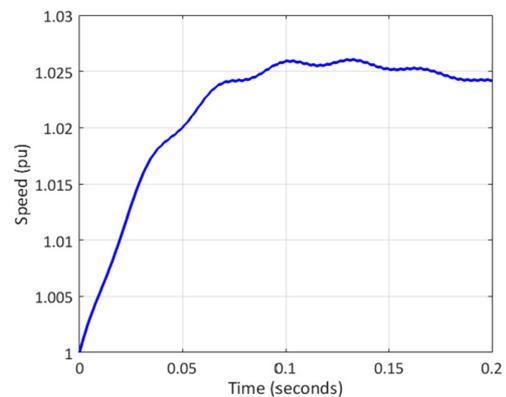


Fig. 7. Turbine speed over the specified time.

It can be observed that the turbine speed is in the safe range and shows the good performance of the proposed system. Figure 8 elucidates the voltage at PCC B575 which is

maintained at 1.0p.u. by controlling the amount of reactive power and the wind farm has equivalent dc-link voltage of 1100V. It is increased due to the control process of back-to-back converters in the conventional design.

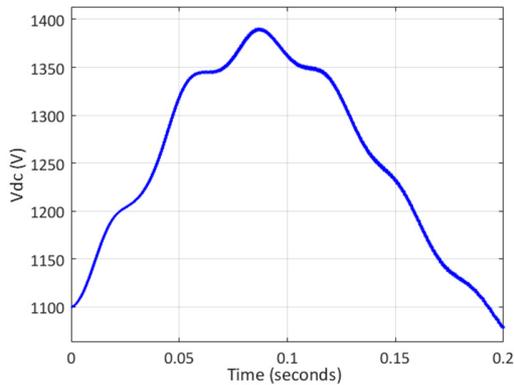


Fig. 8. DC link equivalent voltage across time.

B. System Response at Unsymmetrical SLG Fault

1) Without SVC

The unsymmetrical Single Line to Ground (SLG) fault injected at PCC is demonstrated in Figure 9.

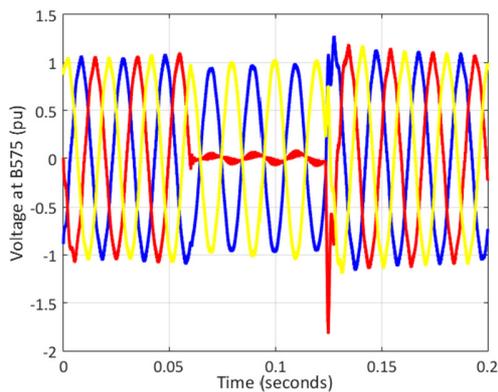


Fig. 9. Generated voltage at PCC B575 during SLG fault.

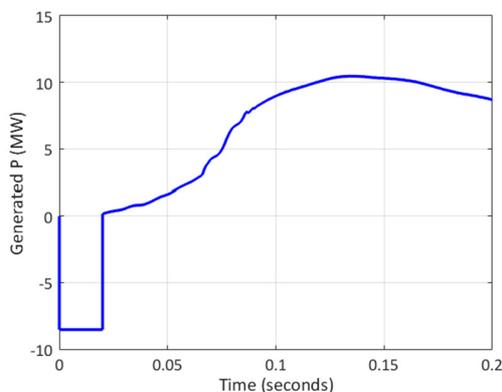


Fig. 10. Generated active power during SLG fault.

One phase is dropped (red line) below 0.5p.u. and the proposed wind farm system is tripped. Hence, the output power of the system becomes zero. Figure 10 shows the variation in the generated power from 0.04 to 0.08s. It can be seen that the optimized 10MW power has certain variations during fault switching time. Further, the variation in the generated power has been observed mainly due to the SLG fault condition. The result also shows the overall performance of the proposed model during temporary injected fault time. During the fault injection, very small reactive power is generated. The reactive power provides some power to run the proposed system model smoothly. Figure 11 shows the poor performance of the system due to the SLG fault.

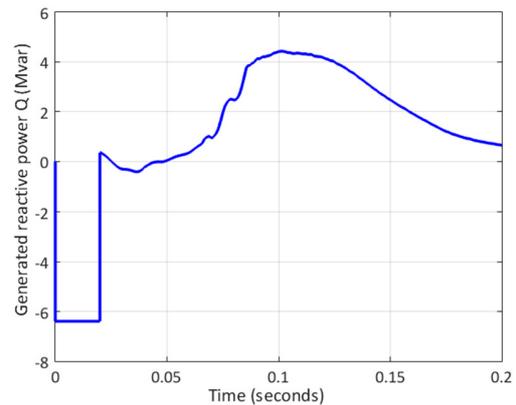


Fig. 11. Generated reactive power during SLG fault.

2) With SVC

Figure 12 shows the bus voltages with an SVC implemented at bus-3 during SLG fault.

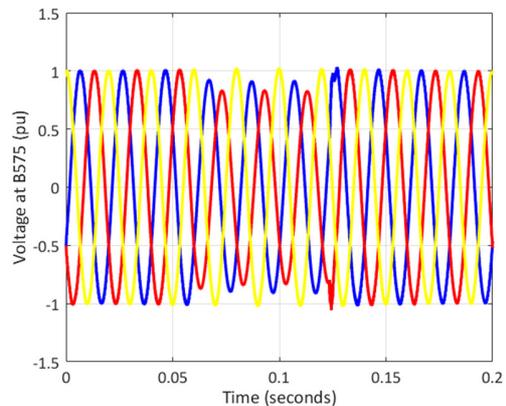


Fig. 12. Voltage at PCC during SLG fault with SVC.

The SVC ensures that the grid integrated wind farm maintains its voltage profile above the value of 0.75p.u.. Therefore, voltage stability and power quality are improved. The wind system continues its connection during fault switching time from 0.06 to 0.12s with the grid by supplying reactive power. The wind farm is not disconnected from the grid and the continuity of supply is maintained. The performance evaluation of the SVC during the temporarily injected SLG fault is depicted in Figure 13. During the

switching fault time from 0.02 to 0.06s, the reactive power (Q) is increased, making the system stable. Power quality is improved, and the supply continuity is maintained. The reactive power is generated by the SVC controller and reaches its optimized value. It can be observed that the proposed model generates a large amount of reactive power when the system voltage is low and achieves better performance in terms of voltage stability.

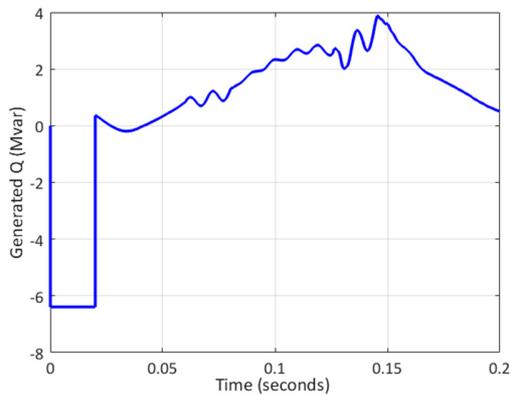


Fig. 13. Generated reactive power (Q) during SLG fault.

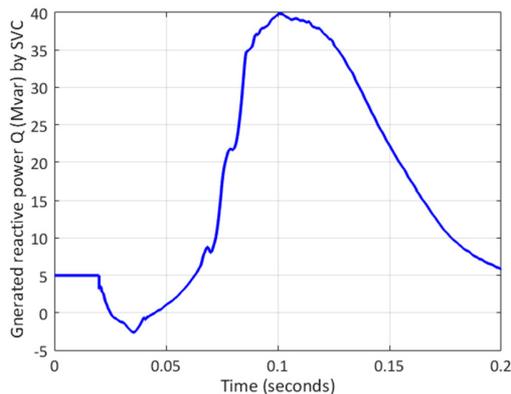


Fig. 14. Generated reactive power (Q) during SLG fault by the SVC.

C. System Response at Unsymmetrical L-L Fault

1) Without SVC

It can be seen in Figure 15 that two phases (red and blue phase) drop from 1.0 to 0.5p.u.. In Line to Line (L-L) fault, it is confirmed that the system is in severe condition which can make it stop from running. Figure 15 demonstrates the phase condition in which the system is disturbed, and the wind farm is tripped. Fault switching occurs between 0.03 and 0.13s. As per grid code requirements, the voltage dip should not fall below 0.75p.. Voltage dip is increased during the fault switching time. It can be noticed that the voltage sag produced by the L-L fault is more severe than the SLG fault. Furthermore, during the fault switching time, from 0.03 to 0.06s, the reactive power is decreased as shown in Figure 16.

2) With SVC

The static SVC controller is the key part of the proposed

model. After the installation of SVC at PCC and during the fault period, voltage profile, voltage stability, and power quality are improved. It is also noticed that the voltage profile is improved from 0.45 to 0.75s as shown in Figure 17.

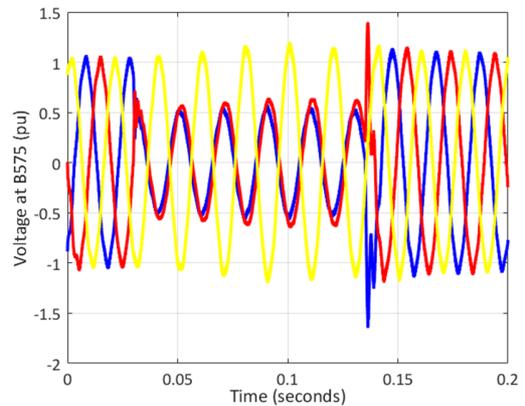


Fig. 15. Generated voltage at PCC during L-L fault.

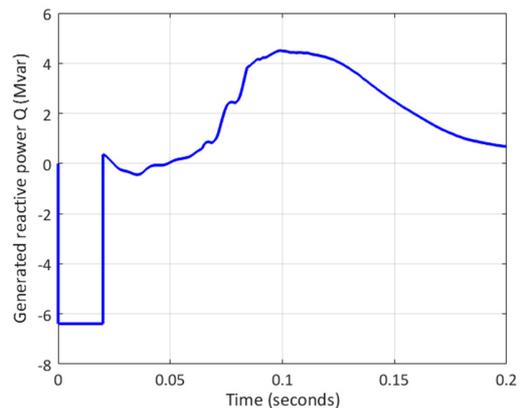


Fig. 16. Generated reactive power (Q) during L-L fault.

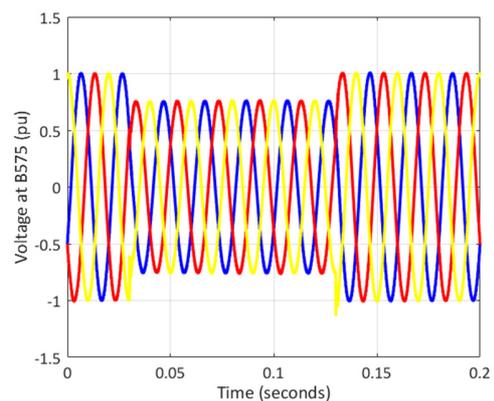


Fig. 17. Voltage at PCC during L-L fault with SVC.

Finally, it is confirmed that the simulation results show better performance in terms of power quality. The wind farm runs in a smooth way and continuity of supply is not disturbed. Figure 18 shows the optimized result of the proposed model. It

is noted that SVC has generated a large amount of reactive power during the L-L fault. The proposed model shows enhanced performance and meets the required conditions. Further, it is observed that the model system is in operational condition and the supply remains uninterrupted.

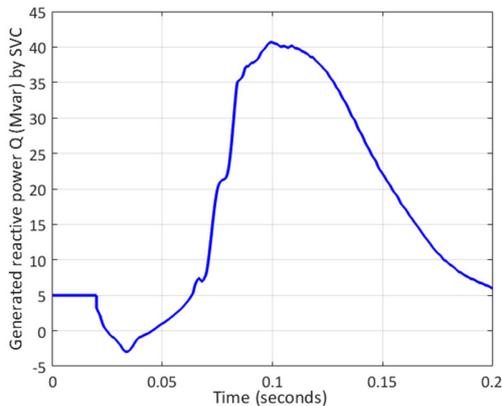


Fig. 18. Generated reactive power (Q) during L-L fault by the SVC.

IV. CONCLUSION

An LVRT enhancement of a grid-tied PMSG-based wind farm system has been proposed in this paper. It consists of 5 wind turbine generators integrated with the grid system. An LVRT control technique is implemented which has achieved balanced grid voltage. Reactive power is supplied during the fault condition to secure the protection of WECS of the wind turbine and as a result, it determines the enhanced LVRT capability. The proposed system has been designed with the use of the SVC controller. When a grid fault occurs, the SVC is automatically activated to regulate the grid voltage and compensate the power loss. The system model has been designed and simulated in MATLAB/Simulink. The simulation results reveal that the proposed system has superior performance in terms of power quality, voltage stability, and compensation of the reactive power in severe disturbances.

REFERENCES

- [1] A. M. Howlader and T. Senjyu, "A comprehensive review of low voltage ride through capability strategies for the wind energy conversion systems," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 643–658, Apr. 2016, <https://doi.org/10.1016/j.rser.2015.11.073>.
- [2] B. Khan and M. Kassas, "FSIG-Based Wind Power Plant Transient Stability Margin Improvement, a STATCOM/SVC Comparison," in *2019 IEEE Texas Power and Energy Conference (TPEC)*, College Station, TX, USA, Feb. 2019, <https://doi.org/10.1109/TPEC.2019.8662129>.
- [3] M. Cheah-Mane, J. Liang, and N. Jenkins, "Permanent magnet synchronous generator for wind turbines: Modelling, control and Inertial Frequency Response," in *2014 49th International Universities Power Engineering Conference (UPEC)*, Cluj-Napoca, Romania, Sep. 2014, <https://doi.org/10.1109/UPEC.2014.6934799>.
- [4] M. Hussain, M. H. Baloch, A. H. Memon, and N. K. Pathan, "Maximum Power Tracking System Based on Power Electronic Topology for Wind Energy Conversion System Applications," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3392–3397, Oct. 2018, <https://doi.org/10.48084/etasr.2251>.
- [5] M. Nasiri, J. Milimonfared, and S. H. Fathi, "A review of low-voltage ride-through enhancement methods for permanent magnet synchronous generator based wind turbines," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 399–415, Jul. 2015, <https://doi.org/10.1016/j.rser.2015.03.079>.
- [6] Z. Memon, M. Uqaili, and M. Unar, "Design of Three-Phase Hybrid Active Power Filter for Compensating the Harmonic Currents of Three-Phase System," *Mehran University Research Journal of Engineering and Technology*, Jamshoro, Pakistan, vol. 31, no. 2, pp. 347–354, Apr. 2012.
- [7] H. S. Dhiman and A. S. Deshpande, "Fault Ride-Through study of PMSG based offshore wind farms during grid faults," in *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, Delhi, India, Jul. 2016, <https://doi.org/10.1109/ICPEICES.2016.7853258>.
- [8] X. Yang, X. Duan, F. Feng, and L. Tian, "Low Voltage Ride-Through of Directly Driven Wind Turbine with Permanent Magnet Synchronous Generator," in *2009 Asia-Pacific Power and Energy Engineering Conference*, Wuhan, China, Mar. 2009, <https://doi.org/10.1109/APPEEC.2009.4918470>.
- [9] M. Nasiri, B. Faridpak, and M. Farrokhifar, "Low Voltage Ride Through Enhancement in PMSG-based Wind Turbines using De-loading Droop," in *2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC)*, Tehran, Iran, Feb. 2020, <https://doi.org/10.1109/PEDSTC49159.2020.9088374>.
- [10] I. A. Ethman, A. Yahfhdou, A. K. Mahmoud, and M. Maaroufi, "Statcom integration in a power grid to enhance voltage stability," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 7, no. 4, pp. 620–627, Dec. 2019, <https://doi.org/10.11591/ijeeci.v7i4.990>.
- [11] S. H. Memon, M. Kumar, A. H. Memon, Z. A. Memon, and S. A. Soomro, "Total Harmonic Distortion (THD) Analysis of Grid Integrated Permanent Magnet Synchronous Generator (PMSG) With Full Scale Converter (FSC) Based Wind Farm," *International Journal of Computer Science and Network Security*, vol. 18, no. 12, pp. 232–238, Dec. 2018.
- [12] S. L. S. Louarem, D. E. C. Belkhat, T. Bouktir, and S. Belkhat, "An Efficient Active and Reactive Power Control of DFIG for a Wind Power Generator," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4775–4782, Oct. 2019, <https://doi.org/10.48084/etasr.3007>.
- [13] T. Ahmed, M. H. Baloch, N. Khan, G. Mehr, B. A. Mirjat, and Y. A. Memon, "Experimental Analysis and Control of a Wind-Generator System through a DC-DC Boost Converter for Extremum Seeking," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6714–6718, Feb. 2021, <https://doi.org/10.48084/etasr.3948>.
- [14] S. H. Memon, Z. A. Memon, M. A. Uqaili, and Shan-E-Farooq, "Voltage Stability and Reactive Power Compensation of 132kv Grid Integrated 50MW Wind Farm Using Statcom," *Journal of Applied Environmental and Biological Sciences*, vol. 8, no. 1, pp. 232–240, 2018.
- [15] L. Sartika, M. Rosyadi, A. Umamura, R. Takahashi, and J. Tamura, "Cooperated stabilizing control of PMSG based grid connected wind farm," in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, Montreal, Canada, Sep. 2015, pp. 3482–3488, <https://doi.org/10.1109/ECCE.2015.7310152>.
- [16] M. Nasiri, S. Mobayen, and Q. M. Zhu, "Super-Twisting Sliding Mode Control for Gearless PMSG-Based Wind Turbine," *Complexity*, vol. 2019, Apr. 2019, Art. no. e6141607, <https://doi.org/10.1155/2019/6141607>.
- [17] A. Mahrouch, M. Ouassaid, and K. Elyalaoui, "LVRT Control for Wind Farm Based on Permanent Magnet Synchronous Generator Connected into the Grid," in *2017 International Renewable and Sustainable Energy Conference (IRSEC)*, Tangier, Morocco, Dec. 2017, <https://doi.org/10.1109/IRSEC.2017.8477281>.
- [18] M. I. Hossain and M. A. Abido, "Positive-Negative Sequence Current Controller for LVRT Improvement of Wind Farms Integrated MMC-HVDC Network," *IEEE Access*, vol. 8, pp. 193314–193339, 2020, <https://doi.org/10.1109/ACCESS.2020.3032400>.
- [19] V. F. Mendes, F. F. Matos, S. Y. Liu, A. F. Cupertino, H. A. Pereira, and C. V. De Sousa, "Low Voltage Ride-Through Capability Solutions for Permanent Magnet Synchronous Wind Generators," *Energies*, vol. 9, no. 1, Jan. 2016, Art. no. 59, <https://doi.org/10.3390/en9010059>.

Assessing the Risk Level of the Challenges Faced In Construction Projects

Abdulla Alameri

Faculty of Business and Technology
Management, Universiti Tun Hussein
Onn Malaysia (UTHM)
Johor, Malaysia

Ahmed S. A. Marey Alhammadi

Faculty of Business and Technology
Management, Universiti Tun Hussein
Onn Malaysia (UTHM)
Johor, Malaysia

Aftab Hameed Memon

Department of Civil Engineering
Quaid-e-Awam University of
Engineering, Science and Technology
Nawabshah, Pakistan

Ismail Abdul Rahman

Faculty of Civil and Environmental Engineering
Universiti Tun Hussein Onn Malaysia (UTHM)
Johor, Malaysia

Nur Ain Ngah Nasaruddin

Faculty of Civil and Environmental Engineering,
Universiti Tun Hussein Onn Malaysia (UTHM)
Johor, Malaysia

Abstract-The failure of construction projects could be attributed to several challenges that emerged during the construction process. Nine major challenges were uncovered from the literature: resource allocation, time, cost, quality, safety, project complexity, changes, uncertainties, and communication. A survey was carried out among professionals having a minimum of 10 years of experience in the construction industry to identify the occurrence and severity of these challenges. A total of 117 responses were analyzed and plotted on a risk matrix. The mean values indicated that time, cost, and quality are the top three challenges faced in construction projects, while the risk matrix revealed that all challenges are high-risk challenges. Thus, the professionals should consider suitable measures to address these challenges for improving a project's performance.

Keywords-construction industry; construction challenges; risk; risk matrix; Malaysia

I. INTRODUCTION

The construction industry is considered a key industry in a country's development and contributes 18% of the worldwide GDP [1]. The construction industry is referred to as vibrant due to uncertainties in technology, budgets, and development processes [2]. The operations of the construction industry require sufficient labor force, skills and commitment, and effective manpower organization [3]. Conventionally, the performance of a construction project is assessed based on three parameters i.e. time, incurred cost, and quality [4-6]. Other studies have stated that the performance also includes safety as an additional criterion [7-9]. Once the project size becomes larger it is becoming more challenging, especially during the implementation stage and thus it will likely disrupt the targeted performance [10]. To deal with this challenge, specific skills, expertise, experience, and knowledge are needed to avoid any potential project overruns [11]. Several approaches have been introduced to address these challenges, and leadership is one of them [12]. Failure or success of construction projects highly depends on the ability to handle

challenges that emerge along with the construction processes. For this, the knowledge of the challenges occurring in construction projects is essential. This study aims to investigate the various challenges faced in construction projects and to assess the level of risk for each of them.

II. IDENTIFICATION OF CONSTRUCTION CHALLENGES

Challenges are inevitable in a construction project and impact its performance. These challenges may emanate from improper communication among the staff, low quality workmanship, and insufficient amount of materials available [13]. Moreover, the organizational drawbacks and the incapability to deal with critical situations also fuel these challenges [14]. This study investigated the constraints faced by construction leaders to keep the project well organized and running smoothly, ensuring its successful completion.

In [15] the challenges were classified into construction issues and peripheral pressures. In [16] the challenges were classified into 4 categories: engineering, human development, managerial and political, and sustainability challenges. Seven construction management challenges were pointed out in [17]. Financial constraints and awareness & knowledge were identified as two major challenges of the construction field in [18]. In [19], three project challenges were pointed out: late approvals, sudden changes of rules, and non-uniform compliance. In [20], the challenges in construction practices were abridged to resource allocation, scheduling and budget management, meeting the quality and safety standards, maintaining complexity, organization management, change management, risk and uncertainties management, and communication system. In [21], the challenges were classified into resource management, risks and uncertainties, support from top management, and restraint of budget. In [22], uncertainty, bad worker attitude, financing matters, contractor selection, and adoption of the latest technology and advanced techniques were pointed out as major construction challenges.

Corresponding author: Aftab Hameed Memon (aftabm78@hotmail.com)

Moreover, claim challenges are faced by every contractor in the construction project [23]. A literature review highlighted the nine common challenges faced in construction works shown in Table I.

TABLE I. CHALLENGES FACED IN CONSTRUCTION PROJECTS

Challenges in construction projects	References
Resource allocation	[24-29]
Time	[30-33]
Cost	[33-38]
Quality	[39, 40]
Safety	[41, 42]
Project complexity	[43, 44]
Changes	[45, 46]
Uncertainties	[24, 47-49]
Communication	[50-52]

A. Resource Allocation

Construction activities depend on the availability of resources [53]. Resource management is essential to achieve the success of any project [54]. Resources in the construction project usually consist of labor, materials, machines, and money [55]. The success of infrastructure and megaprojects needs a high amount of financial and human resources [56]. Resources should be properly allocated due to limited availability and high demands in the construction industry, while all resources are equally important. Construction materials typically contribute about 40-45% to the total cost of a project [57, 58]. Hence, appropriate planning is crucial to ensure the timely availability of adequate and suitable materials [59]. Planning the construction project manpower is important for determining the size of the project workforce, organizing into functional groups, and hiring suitable manpower [58, 59]. Construction machinery contributes about 20-30% to a project's cost, with additional costs for maintenance, repair, and operation. Resource planning also ensures the availability of adequate and suitable equipment [59].

B. Time

Construction time is a major concern, as each project has a predetermined period with a definable beginning and an identifiable end. Inefficient time management increases a project's cost, spoils a company's reputation, and reduces construction productivity [55].

C. Cost

The construction process is capital intensive due to the large amounts incurred at every project stage [29]. The cost of labor, materials, equipment, professional fees, and total profit are major components of a project's cost. Cost management can be affected by capital shortages, impractical profit margin, high debits, and improper management of resources [37]. Besides, cost escalation is also a ubiquitous issue faced in construction projects [60].

D. Quality

Quality is defined as meeting the legal, aesthetic, and functional requirements of a project [61]. Quality problems cause delays and cost overruns. Effective quality management

increases product quality, improves workmanship and efficiency, decreases wastage, and increases profit [62]. Quality is an essential consideration in selecting the proper contractor for any project [63].

E. Safety

Safety is a major concern in the success of the construction industry, especially when considering work-related accidents [64]. Safety measures account for 1-2% of the total contract cost in a construction project [58]. Safety management issues in a construction site depend on workers' attitudes [65].

F. Project Complexity

Project complexity involves the employment of individuals from different organizations and combinations of technology [66]. Project complexity has a direct relation with uncertainty in the term of its element [67].

G. Changes

A change in a construction project is an alteration or a modification in the pre-existing conditions, assumptions, or requirements [68]. It occurs when the scope of work differs from the scope of the work outlined in the contract documents [69]. Changes in construction affect delivery time and cost significantly [69-70].

H. Uncertainties

Uncertainty is the chance of any event with a genuinely unknown probability distribution [71]. The ambiguity of a project's uncertainty is considered a threat to construction performance [72]. Managing uncertainties involves risk and opportunities, where the activities and processes could become worse or better than planned [73].

I. Communication

Previous studies indicated that over 50% of construction projects were unsuccessful due to poor or insufficient communication [74]. The complex communication environment in large or mega construction projects requires frequently changing sets of relationships [50].

III. DATA COLLECTION AND ANALYSIS

This study performed a quantitative survey on the perceptions of various professionals engaged in handling construction projects. It aimed to uncover the possible challenges in the construction industry and identify their respective level of risk. Professionals' perceptions were collected utilizing a structured questionnaire. Data analysis involved statistical methods using mean score calculation with the formula adopted by [75]:

$$M = \frac{\sum(X)}{N} \quad (1)$$

where M is the mean, X are the individual data points, and N is the sample size (number of the data points).

The risk level of the challenges was assessed based on the risk matrix. A risk matrix is a tool to assess the risk level using a graphical representation based on two axes: the level of severity and occurrence. A 5-point scale was used for each axis, as shown in Table II: 1-very low, 2-low, 3-moderate, 4-high, and 5-very high. The factors could be classified based on

risk level or using a risk matrix where the factors are placed based on occurrence and severity levels and assessed accordingly. This study utilized the risk matrix to assess the risk category of construction's challenges. Based on the two parameters and the scale, the drawn risk matrix can be seen in Figure 1 [76].

TABLE II. MEASUREMENT SCALE

Scale	Level of Severity	Frequency of Occurrence/likelihood
1	Very Low (VL)	Very Low (VL)
2	Low (L)	Low (L)
3	Moderate (M)	Moderate (M)
4	High (H)	High (H)
5	Very High (VH)	Very High (VH)

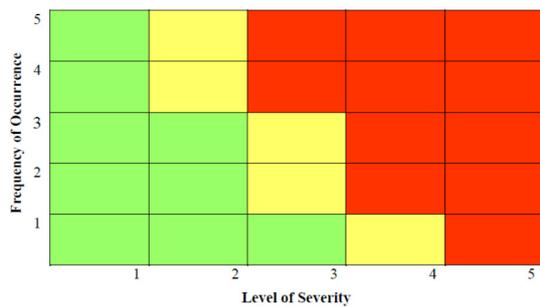


Fig. 1. Risk matrix.

The risk matrix is divided into three zones:

- Green zone: Risks are of low level and can be ignored.
- Yellow zone: Risks are of moderate importance and should be controlled.
- Red zone: Risks are of critical importance. These are the top priorities and close attention should be paid.

IV. RESULTS AND DISCUSSION

A total of 117 completed questionnaires were collected throughout Peninsular Malaysia. Data collection was performed through an in-person/site visit (face-to-face) approach.

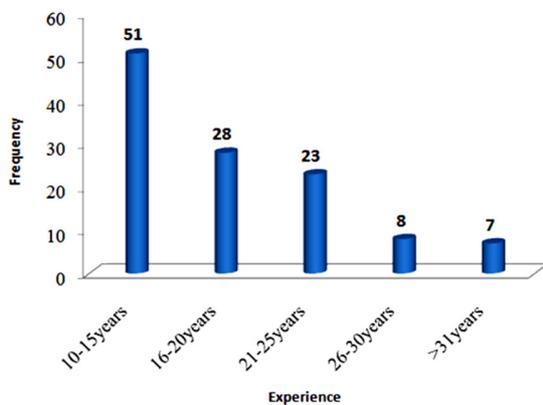


Fig. 2. Respondents' experience.

Initially, an appointment was set with the respondents to participate in a telephone interview and an e-mail survey. On the agreed date, the respondents were approached and a short briefing was given about the study's purpose and the contents of the questionnaire. The survey was completed in 3 months. The respondents were professionals with experience in several project types and sizes. A summary of the respondents' project experience and size is shown in Figures 2 and 3 respectively.

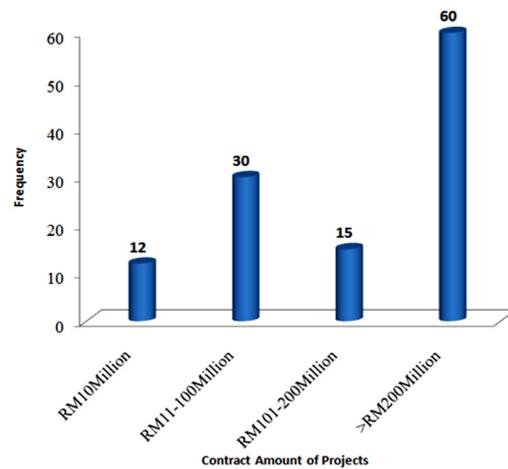


Fig. 3. Size of the projects participated.

Respondents' experience is a very crucial parameter as they could provide better judgment on the challenges and the required leadership characteristics due to their long-time involvement. From Figure 2, it can be noted that the minimum experience of the respondents was 10 years while 38 respondents were working for more than 20 years in the construction industry. This indicates that the participants were capable enough to provide insight views of the problems faced in construction. The respondents were handling projects of several sizes. The minimum size of the projects handled by respondents was RM 10 million, while the majority was engaged in projects with more than RM 200 million contracts. Collected questionnaires were checked for reliability to assess their consistency. Cronbach's alpha value was computed to check the internal consistency, which ranges from 0 and 1, where 1 means that the data are highly reliable and valid while 0 means the opposite [77]. In this study, Cronbach's alpha value of data for severity level was found at 0.851, and the same for occurrence level was found as 0.844. As these values are higher than 0.6, they are deemed suitable for further analysis as suggested in [78]. The mean values of the responses on the occurrence and severity levels of construction challenges were calculated, and the results are shown in Table III. From Table III, it can be observed that time challenge is the most commonly occurring and severe challenge in construction projects of Malaysia, as it was also shown in [79]. The second major challenge was the cost of the construction project. Cost is a very serious issue and is given the highest priority by the owners and subsidizers. Many researchers have emphasized that cost must be given a high priority. Quality was reported as the third major challenge from the occurrence point of view, while from the severity perspective it was reported as fourth.

On the contrary, the practitioners reported safety as the fourth and third major challenge from the occurrence and severity aspects respectively. The mean values of all challenges were plotted on a risk matrix, as shown in Figure 4.

TABLE III. LEVEL OF OCCURRENCE AND SEVERITY OF THE CHALLENGES

S.No	Challenges in handling the construction projects	Frequency of occurrence	Level of severity
1	Time	3.67	3.48
2	Cost	3.51	3.29
3	Quality	3.44	3.19
4	Safety	3.33	3.24
5	Resources allocation	3.29	3.09
6	Changes	3.19	3.08
7	Uncertainties	3.15	3.06
8	Project complexity	3.03	2.84
9	Communication	3.13	2.83

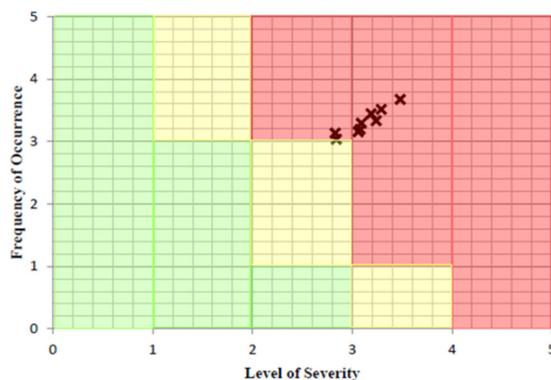


Fig. 4. Risk level of the challenges.

Figure 4 illustrates the 9 challenges marked with "x" on the risk matrix. Since all marks are located in the red zone/region, it indicates that all challenges belong to the high-risk category. The challenges need to be addressed carefully by the construction leaders during the implementation of a project. Hence, appropriate measures should be considered to handle these challenges successfully. One of the appropriate ways for controlling these challenges is effective leadership. Leadership can be used effectively by adopting several characteristics such as imparting information effectively, knowledgeable, work according to the plan, setting priorities, and willingness to take responsibility [80].

V. CONCLUSION

Every construction project faces unavoidable challenges. The level and effects of a challenge might vary from project to project as well as from region to region. This study focused on examining the various challenges faced in construction projects of Malaysia through a questionnaire survey among experienced practitioners. These challenges included resource allocation, time, cost, quality, safety, project complexity, changes, uncertainties, and communication. Statistical analysis revealed that all investigated challenges were high-risk challenges. The practitioners reported that time, cost, and quality were the most critical challenges. The practitioners need to take the necessary

measures for avoiding time, cost, and quality problems that could fail a construction project.

REFERENCES

- [1] R. Udhayakumar and D. P. Karthikeyan, "Expected Leadership Qualities for a Project Manager to Manage Construction Projects," *International Journal of Innovative Research and Development*, vol. 3, no. 10, Oct. 2014.
- [2] M. Jain and K. K. Pathakx, "Critical Factors Affecting the Success of a Construction Projects: A Review," *Asian Journal of Educational Research and Technol*, vol. 4, no. 2, pp. 420–426, 2014.
- [3] R. Morton and A. Ross, *Construction UK: Introduction to the Industry*, 2nd edition. Oxford ; Malden, MA: Wiley-Blackwell, 2007.
- [4] A. P. C. Chan, D. C. K. Ho, and C. M. Tam, "Design and Build Project Success Factors: Multivariate Analysis," *Journal of Construction Engineering and Management*, vol. 127, no. 2, pp. 93–100, Apr. 2001, [https://doi.org/10.1061/\(ASCE\)0733-9364\(2001\)127:2\(93\)](https://doi.org/10.1061/(ASCE)0733-9364(2001)127:2(93)).
- [5] A. S. Ali and I. Rahmat, "The performance measurement of construction projects managed by ISO-certified contractors in Malaysia," *Journal of Retail & Leisure Property*, vol. 9, no. 1, pp. 25–35, Feb. 2010, <https://doi.org/10.1057/rlp.2009.20>.
- [6] Y. Singh, T. Abdelhamid, T. Mrozowski, and M. A. El-Gafy, "Investigation of Contemporary Performance Measurement Systems for Production Management of Renovation Projects," *Journal of Construction Engineering*, vol. 2014, Jul. 2014, Art. no. e417853, <https://doi.org/10.1155/2014/417853>.
- [7] A. Enshassi, "Factors Affecting Safety on Construction Projects," *Islamic University of Gaza, Gaza Strip, Palestine*, 2003, Accessed: Apr. 23, 2021. [Online]. Available: <https://iugspace.iugaza.edu.ps/handle/20.500.12358/26611>.
- [8] G. John and A. Ganah, "Integrating BIM and Planning Software for Health and Safety Site Induction," presented at the COBRA 2011, The Royal Institution of Chartered Surveyors International Research Conference, University of Salford, UK, Sep. 2011.
- [9] J. Wanberg, C. Harper, M. R. Hollowell, and S. Rajendran, "Relationship between Construction Safety and Quality Performance," *Journal of Construction Engineering and Management*, vol. 139, no. 10, Oct. 2013, Art. no. 04013003, [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000732](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000732).
- [10] A. S. Ali and S. N. Kamaruzzaman, "Cost performance for building construction projects in Klang valley," *Journal of Building Performance*, vol. 1, no. 1, pp. 110–118, 2010.
- [11] W. H. N. W. Muda, W. H. Libunao, K. M. Salleh, and N. Sulaiman, "Developing a Leadership Capability for Team Leaders in the Construction Industry: A Concept for Organizational Success," *Journal of Technical Education and Training*, vol. 8, no. 2, Dec. 2016.
- [12] R. B. Isa, R. A. Jimoh, and E. Achuen, "An overview of the contribution of construction sector to sustainable development in Nigeria," *Net Journal of Business Management*, vol. 1, no. 1, pp. 1–6, Nov. 2013.
- [13] S. T. Ng, R. M. Skitmore, K. C. Lam, and A. W. C. Poon, "Demotivating factors influencing the productivity of civil engineering projects," *International Journal of Project Management*, vol. 22, no. 2, pp. 139–146, Feb. 2004, [https://doi.org/10.1016/S0263-7863\(03\)00061-9](https://doi.org/10.1016/S0263-7863(03)00061-9).
- [14] M. Yadollahi, M. Mirghasemi, R. Mohamad Zin, and B. Singh, "Architect Critical Challenges as a Project Manager in Construction Projects: A Case Study," *Advances in Civil Engineering*, vol. 2014, Aug. 2014, Art. No. e205310, <https://doi.org/10.1155/2014/205310>.
- [15] B. Muir, "Challenges Facing Today's Construction Manager," University of Delaware, Supplemental Reading for CIEG 486-010 Construction Methods & Management, 2005.
- [16] A. A. E. Othman, "Challenges of mega construction projects in developing countries," *Organization, Technology and Management in Construction - An International Journal*, vol. 5, no. 1, pp. 730–746, 2013.

- [17] D. Q. Tran, M. R. Hallowell and K. R. Molenaar, "Construction management challenges and best practices for rural transit projects," *Journal of Management in Engineering*, vol. 31, no. 5, 2014.
- [18] Y. H. Chan, B. C. T. Lee, and J. C. Lee, "Sustainability in the Construction Industry in Malaysia: The Challenges and Breakthroughs," *International Journal of Economics and Management Engineering*, vol. 8, no. 4, pp. 1218–1222, Aug. 2014.
- [19] A. R. Nuruddin, S. P. S. A. Bakar, and M. Jaafar, "Unveiling the Challenges Faced By Malaysian Housing Developers through Government Policy Changes," *Journal of Construction in Developing Countries*, vol. 20, no. 2, pp. 37–52, 2015.
- [20] P. U. Okoye, C. Ngwu, and S. C. Ugochukwu, "Evaluation of management challenges facing constructino practice in Nigeria," *International Journal of Application or Innovation in Engineering & Management*, vol. 4, no. 1, pp. 19–28, Jan. 2015.
- [21] K. Subramanyam and M. K. Haridharan, "Examining the Challenging Hindrances facing in the Construction Projects: South India's Perspective," *IOP Conference Series: Earth and Environmental Science*, vol. 80, Jul. 2017, Art. no. 012045, <https://doi.org/10.1088/1755-1315/80/1/012045>.
- [22] M. S. Kiu, S. R. Mehdi Riazi, M. N. Mohd Nawi, and M. Esa, "The roles, challenges & improvement of project manager in project financial management in building development in Penang, Malaysia," *International Journal of Supply Chain Management (IJSCM)*, vol. 6, no. 1, pp. 238–245, Mar. 2017.
- [23] B. N. Azmi, A. Hamimah, and I. Azmi, "Construction claim problems in Malaysia: from the contractors perspective," *MATEC Web of Conferences*, vol. 192, 2018, Art. no. 02004, <https://doi.org/10.1051/mateconf/201819202004>.
- [24] C. Hendrickson and T. Au, *Project Management for Construction: Fundamental Concepts for Owners, Engineers, Architects, and Builders*, Facsimile edition. Englewood Cliffs, N.J: Prentice-Hall, 1989.
- [25] A. Laskar and V. R. Murty, "Challenges before Construction Industry in India," Department of Civil Engineering, Indian Institute of Technology Kanpur, 2004.
- [26] T. Ala-Risku and M. Kärkkäinen, "Material delivery problems in construction projects: A possible solution," *International Journal of Production Economics*, vol. 104, no. 1, pp. 19–29, Nov. 2006, <https://doi.org/10.1016/j.ijpe.2004.12.027>.
- [27] A. C. S. Dudgikar, E. M. B. Kumthekar, and E. S. R. Khot, "Development of ERP Module for Quality Management in Construction Industry," *International Journal of Electronics and Communications*, vol. 1, no. 1, Aug. 2012, Art. no. 12.
- [28] A. O. Sanni and M. Hashim, "Assesing the chalenges of cost control practices in Nigerian construction industry," *Interdisciplinary Journal of Contemporary Research in Business*, vol. 4, no. 9, Jan. 2013, Art. no. 9.
- [29] S. Chukwudi and T. Onyekwena, "Participation of indigenous contractors in Nigerian public sector projects and their challenges in managing working capital," *International Journal of Civil Engineering, Construction and Estate Management*, vol. 1, no. 1, Mar. 2014, Art. no. 23.
- [30] A. Al-Kharashi and M. Skitmore, "Causes of delays in Saudi Arabian public sector construction projects," *Construction Management and Economics*, vol. 27, no. 1, pp. 3–23, Jan. 2009, <https://doi.org/10.1080/01446190802541457>.
- [31] A. H. Memon, I. A. Rahman, M. R. Abdullah, and A. A. A. Azis, "Factors Affecting Construction Cost in Mara Large Construction Project: Perspective of Project Management Consultant," *International Journal of Sustainable Construction Engineering and Technology*, vol. 1, no. 2, pp. 41–54, 2010.
- [32] K. A. Mohammed and A. D. Isah, "Causes of delay in Nigeria construction industry," *Interdisciplinary Journal of Contemporary Research in Business*, vol. 4, no. 2, pp. :785-94, 2012
- [33] A. Q. Memon, A. H. Memon, and M. A. Soomro, "Contractor's Perception on Factors Causing Cost Overrun in Construction Works of Pakistan," *International Journal of Sustainable Construction Engineering and Technology*, vol. 11, no. 3, pp. 84–92, Dec. 2020.
- [34] S. Ogunlana, Z. Siddiqui, S. Yisa, and P. Olomolaiye, "Factors and procedures used in matching project managers to construction projects in Bangkok," *International Journal of Project Management*, vol. 20, no. 5, pp. 385–400, Jul. 2002, [https://doi.org/10.1016/S0263-7863\(01\)00017-5](https://doi.org/10.1016/S0263-7863(01)00017-5).
- [35] A. Omoregie and D. Radford, "Infrastructure delays and cost escalation: Causes and effects in Nigeria," in *Proceedings of the 6th International Postgraduate Research Conference in the Built and Human Environment*, Delft, Netherlands, Apr. 2006.
- [36] B. Chigara, T. Moyo, and F. H. Mudzengerere, "An Analysis of Cost Management Strategies Employed By Building Contractors on Projects in Zimbabwe," *International Journal of Sustainable Construction Engineering and Technology*, vol. 4, no. 2, pp. 1–13, Nov. 2013.
- [37] M. S. A. Halim, M. S. M. Haniff, M. Z. M. Junoh, and A. Osman, "Financial Performance and the Management Issues of Bumiputera Construction Firms in the Malaysian Construction Industry," *Journal of Scientific Research and Reports*, pp. 1190–1202, Mar. 2014, <https://doi.org/10.9734/JSRR/2014/8418>.
- [38] A. S. A. M. Alhammadi and A. H. Memon, "Inhibiting Factors of Cost Performance in UAE Construction Projects," *International Journal of Sustainable Construction Engineering and Technology*, vol. 11, no. 2, pp. 126–132, Oct. 2020.
- [39] P. Hoonakker, P. Carayon, and T. Loushine, "Barriers and benefits of quality management in the construction industry: An empirical study," *Total Quality Management & Business Excellence*, vol. 21, no. 9, pp. 953–969, Sep. 2010, <https://doi.org/10.1080/14783363.2010.487673>.
- [40] A. H. Memon *et al.*, "Risk Level of the Factors Affecting the Quality of Construction Projects in Pakistan," *Indian Journal of Science and Technology*, vol. 11, no. 48, pp. 1–8, Dec. 2018, <https://doi.org/10.17485/ijst/2018/v11i48/131226>.
- [41] K. Dorji and B. H. W. Hadikusumo, "Safety Management Practices in the Bhutanese Construction Industry," *Journal of Construction in Developing Countries*, vol. 11, no. 2, pp. 53–75, 2006.
- [42] G. Muiruri, "Health and safety management on construction project sites in Kenya: a case of construction projects in Nairobi County," M.S. thesis, University of Nairobi, Kenya, 2012.
- [43] H. Wood and K. Gidado, "Project Complexity in Construction," in COBRA 2008 - Construction and Building Research Conference of the Royal Institution of Chartered Surveyors, Jan. 2008.
- [44] Z. Shehu and A. Akintoye, "Major challenges to the successful implementation and practice of programme management in the construction environment: A critical analysis," *International Journal of Project Management*, vol. 28, no. 1, pp. 26–39, Jan. 2010, <https://doi.org/10.1016/j.ijproman.2009.02.004>.
- [45] I. Motawa, "A Systematic Approach to Modelling Change Processes in Construction Projects," *Construction Economics and Building*, vol. 5, no. 1, pp. 23–31, 2005, <https://doi.org/10.5130/AJCEB.v5i1.2940>.
- [46] Q. Hao, W. Shen, J. Neelamkavil, and R. Thomas, "Change management in construction projects," presented at the International Conference on Information Technology in Construction, Santiago, Chile, Jul. 2008.
- [47] J. E. Okema, "Risk and Uncertainty Management of Projects: Challenges of Construction Industry," presented at the 2nd International Conference on Construction in Developing Countries: Challenges Facing the Construction Industry in Developing Countries, Gaborone, Botswana, Nov. 2000.
- [48] M. Rahman and M. Kumaraswamy, "Revamping risk management in Hong Kong construction industry," in *Proceedings of the RICS Foundation - Construction & Building research Conference (COBRA 2001)*, Glasgow, UK, Sep. 2001.
- [49] N. Banaitiene and A. Banaitis, "Risk Management in Construction Projects," in *Risk Management - Current Issues and Challenges*, Intech: Rijeka, Croatia, pp. 429–448.
- [50] M. Hoezen, I. Reymen, and G. Dewulf, "The problem of communication in construction," presented at the CIB W96 Adaptable Conference, Enschede, Netherlands, Jul. 2006.
- [51] B. Sidawi, "Potential Use of Communications and Project Management Systems in Remote Construction Projects: The Case of Saudi Electric Company," *Journal of Engineering, Project, and Production*

- Management, vol. 2, no. 1, pp. 14–22, Jan. 2012, <https://doi.org/10.32738/JEPPM.201201.0003>.
- [52] F. Emuze and M. James, "Exploring communication challenges due to language and cultural diversity on South African construction sites," *Acta Structilia*, vol. 20, no. 1, pp. 44–65, Sep. 2013.
- [53] A. Zin, A. Memon, and R. Mohamad zin, "Comparative Study of Resource Scheduling with Primavera Project Planner and Microsoft Project," in *SEPKA 2006 - National Seminar on Civil Engineering Research*, Johor, Malaysia, Dec. 2006.
- [54] A. Hameed, "Resource-driven scheduling: barriers to implementation," M.S. thesis, University Teknologi Malaysia, 2005.
- [55] S. M. Patil, D. B. Desai, and A. K. Gupta, "Earned value analysis in construction industry," *International Journal of Informative & Futuristic Research*, no. Special, pp. 19–26, Mar. 2015.
- [56] M. A. Akhund, A. R. Khoso, A. A. Pathan, H. U. Imad, and F. Siddiqui, "Risk Attributes, Influencing the Time and Cost Overrun in Joint Venture Construction Projects of Pakistan," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3260–3264, Aug. 2018, <https://doi.org/10.48084/etasr.2203>.
- [57] M. M. A. El Al Kass, "A Construction Resources Management System for Gaza Strip Building Contractors," M.S. thesis, Islamic University of Gaza, Gaza, Palestine, 2012.
- [58] S. Rowlinson, *Cost escalation in the Hong Kong construction industry report*. Hong Kong, China: The University of Hong Kong, 2014.
- [59] K. S. Kumari, "A Study On Resource Planning In Highway Construction Projects," *International Journal of Engineering Research and*, vol. 2, pp. 1960–1967, 2012.
- [60] S. Ahmed, A. H. Memon, N. A. Memon, A. N. Laghari, M. A. Akhund, and H. U. Imad, "Common Factors of Cost Escalation in Construction Industry of Pakistan," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3508–3511, Dec. 2018, <https://doi.org/10.48084/etasr.2278>.
- [61] D. Arditi and H. M. Gunaydin, "Total quality management in the construction process," *International Journal of Project Management*, vol. 15, no. 4, pp. 235–243, Aug. 1997, [https://doi.org/10.1016/S0263-7863\(96\)00076-2](https://doi.org/10.1016/S0263-7863(96)00076-2).
- [62] A. A. Bubshait and T. H. Al-Atiq, "ISO 9000 Quality Standards in Construction," *Journal of Management in Engineering*, vol. 15, no. 6, pp. 41–46, Nov. 1999, [https://doi.org/10.1061/\(ASCE\)0742-597X\(1999\)15:6\(41\)](https://doi.org/10.1061/(ASCE)0742-597X(1999)15:6(41)).
- [63] N. U. I. Soomro, A. H. Memon, N. A. Memon, and K. R. Memon, "Contractor's Selection Criteria in Construction Works in Pakistan," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5520–5523, Apr. 2020, <https://doi.org/10.48084/etasr.3334>.
- [64] M. Saïdani, "Health and Safety in the construction industry: Challenges and solutions in UAE," presented at the Third International World of Construction Project Management Conference, Coventry University, UK, Jan. 2010, <https://doi.org/10.13140/RG.2.1.4352.8168>.
- [65] E. M. K. Ghani, E. D. Z. B. A. Hamid, E. M. Z. M. Zain, E. A. H. A. R. Rahim, E. K. A. M. Kamar, and M. A. A. Rahman, "Safety in Malaysian Construction: The Challenges and Initiatives," *JURUTERA*, pp. 16–19, May 2008.
- [66] R. F. Fellows and A. M. M. Liu, *Research Methods for Construction*, 4th ed. Hoboken, NJ, USA: John Wiley & Sons, 2015.
- [67] C. Brockmann, "Evaluating Construction Project Complexity," in *Management of Construction: Research to Practice (MCRP) Conference Proceedings, Joint CIB International Symposium of W055, W065, W089, W118, TG76, TG78, TG81 and TG84*, Montreal, Canada, 2012, pp. 716–727.
- [68] D. Lazarus and R. Clifton, *Managing Project Change - A Best Practice Guide: A Best Practice Guide C556*. London: Construction Industry Research & Information Association, 2001.
- [69] J. J. Egan, J. E. Seder, and D. L. Anderson, "Practices in Construction Change Order Management," *Cost Engineering*, Apr. 2012.
- [70] B.-G. Hwang and L. K. Low, "Construction project change management in Singapore: Status, importance and impact," *International Journal of Project Management*, vol. 30, no. 7, pp. 817–826, Oct. 2012, <https://doi.org/10.1016/j.ijproman.2011.11.001>.
- [71] H. Wood and P. Ashton, "The factors of project complexity," in *18th CIB World Building Congress*, Jan. 2010, pp. 69–80.
- [72] M. Mentis, "Managing project risks and uncertainties," *Forest Ecosystems*, vol. 2, no. 1, Jan. 2015, Art. no. 2, <https://doi.org/10.1186/s40663-014-0026-z>.
- [73] S. Lichtenberg, "The Successive Principle—a scientific crystal ball for management," *The IPSI BgD Transactions on Internet Research*, vol. 2, no. 1, pp. 33–39, Jul. 2006.
- [74] L. G. Tipili, P. O. Ojeba, and M. S. Ilyasu, "Evaluating the effects of communication in construction project delivery in Nigeria," *Global Journal of Environmental Science and Technology*, vol. 2, no. 5, pp. 48–54, Jun. 2014.
- [75] N. A. Weiss, *Elementary Statistics*, 8th edition. Boston, MA, USA: Pearson, 2011.
- [76] I. Mahamid, "Common risks affecting time overrun in road construction projects in Palestine: Contractors' perspective," *Australasian Journal of Construction Economics and Building, The*, Jan. 2013.
- [77] M. Tavakol and R. Dennick, "Making sense of Cronbach's alpha," *International Journal of Medical Education*, vol. 2, pp. 53–55, Jun. 2011, <https://doi.org/10.5116/ijme.4dfb.8dfd>.
- [78] A. Yahaya, s Haslim, J. Ramli, Y. Boon, and A. R. Hamdan, *Menguasai Penyelidikan Dalam Pendidikan: Teori, Analisis & Interpretasi Data*. Kuala Lumpur, Malaysia: PTS Professional, 2007.
- [79] I. A. Rahman, A. H. Memon, A. Q. Memon, M. A. Shaikh, and F. Siddiqui, "Reasons of poor labour productivity in construction works," *Malaysian Construction Research Journal*, vol. 8, no. 3, 2019.
- [80] N. A. Nasaruddin, I. A. Rahman, and M. M. Jaber, "PLS-SEM Model of Leadership Characteristics facing Challenges in Malaysia Construction Industry," *International Journal of Engineering & Technology*, vol. 7, pp. 620–624, 2018.

A New Model for Enhancing Student Portal Usage in Saudi Arabia Universities

Salem Suliman T. Alatawi
School of Computing
Universiti Teknologi Malaysia
Johor, Malaysia
salem.net1@hotmail.com

Suraya Miskon
Azman Hashim International Business
School, Universiti Teknologi Malaysia
Johor, Malaysia
suraya@utm.my

Norris Syed Abdullah
Azman Hashim International Business
School, Universiti Teknologi Malaysia
Johor, Malaysia
norris@utm.my

Fahad Ghabban
College of Computer Science and Engineering
Taibah University
Medina, Saudi Arabia
fghaban@taibahu.edu.sa

Faisal Saeed
College of Computer Science and Engineering
Taibah University
Medina, Saudi Arabia
fsaeed@taibahu.edu.sa

Ibrahim Alfadli
College of Computer Science and Engineering
Taibah University
Medina, Saudi Arabia
ialfadli@taibahu.edu.sa

Omair Ameerbakhsh
College of Computer Science and Engineering
Taibah University
Medina, Saudi Arabia
oameerbakhsh@taibahu.edu.sa

Abstract—Portals are gateways that provide users with the information they need from different sources and display it on a single page. It is important to see that universities utilize the resources and services provided by their student portals. With the rapid development of Information and Communication Technology (ICT), the Ministry of Education in Saudi Arabia aims to develop and improve student portals by providing high-quality teaching services through the university portal systems. This paper discusses the importance of student portal usage in Saudi Arabian universities and investigates the factors that influence the utilization of student portals as perceived by the students of the Saudi universities. Based on these factors, a model is proposed which identifies students' expectations about the Saudi university portals. A quantitative methodology was employed to develop the model. The results revealed that 8 out of 10 factors of the model are significant and positively affect student portal usage. The enhancement of student portals based on the identified significant factors will assist the universities to increase their utilization and their provided services.

Keywords—university portal; higher education; e-services; e-learning; quality of service; utilization of student portals; ICT

I. INTRODUCTION

High-quality web portals are critical to commercial and non-commercial organizations [1, 2]. A portal is defined as a general knowledge management system that helps companies to share, reuse, create or exchange information [3]. A university portal provides its users a specialized view that matches their needs and meets their software and hardware

requirements. To build a university portal that is updated regularly with the latest information and can accommodate additional services is a significant challenge [4]. The main concern is to build a portal that controls, coordinates, interacts and takes feedbacks from customers (visitors, staff and students). The Ministry of Higher Education in Saudi Arabia works close with Saudi universities to improve the quality of higher education in order to ensure that university graduates have skills and learning outcomes in line with the international standards and job market requirements. Despite the good Information and Communication Technology (ICT) infrastructure in Saudi universities, the usage of university student portals remains low [5]. According to [6], a big challenge that Saudi universities face is the lack of awareness regarding the use of ICT among students, faculty members, and administrative staff. Many universities in Saudi Arabia spend generously on ICT, but this expenditure does not lead to significant improvements in the usage of ICT services. Many universities have adequate ICT infrastructure, but they do not use it effectively [7]. Higher education institutions, particularly universities in developed countries, are heavily involved in portals which have become an essential tool for service delivery and communication between faculty members, administrative staff, and, most important, students [8].

Several studies have been conducted in Saudi Arabia on the importance and usage of university student portals. Authors in [9] showed that the official university portal for King Abdul-Aziz University is well-below the expected standard. The study

Corresponding author: Salem Suliman T. Alatawi

claims that King Abdul-Aziz University's portal does not match the standards of design, content, user support, and navigation and the undergraduate students are dissatisfied with it [10]. The aim of [10] was to explore usability problems within Saudi university websites in comparison with the UK university websites by applying the SUS methods and Thinking aloud method to measure the usability of the websites. The results illustrated that Saudi students faced various usability problems related with satisfaction, integration, and confidence. The author in [2] evaluated the EduGate, an online academic portal of King Saud University and problems were found concerning its design and content. These problems reduced the usage of the portal and dissatisfied academic staff. The author in [11] revealed that Saudi university portals need to focus on their content to meet the needs of staff and students and to increase their level of usage. Furthermore, authors in [12] revealed that the score for using King Khalid University's website is 53.5 while the average industry score of using a website is 65. This indicates that there are problems impeding the effectiveness and efficiency of the university portal resulting in minimal usage.

Given the above background, this study investigates the factors that affect students' usage of their university portal by proposing a model that meets student expectations and increases their satisfaction. This model will help the universities to communicate more efficiently with their students.

II. RELATED WORK

Web portals are general knowledge management systems that help institutions to share, reuse, create, or exchange information or knowledge [3]. Portals have been employed increasingly to manage the communication between different stakeholders in an organization or initiative. Educational web portals have become more responsive and turned out to be very dynamically based on the demands and the requirements of the varying academic community. The communication between teachers and students is often conducted through these websites and even more these days due to the Covid 19 pandemic. These portals actually integrate the application, content, and information together in order to help the end users. University student portals provide customers with a specialized view that matches their needs and meets their software and hardware requirements. The understanding of the users' needs is a key element in the student portal usage. It mainly includes four important things, namely the accessibility standards, page layout, graphic design, and, most important, content design [13]. Therefore, the main concern of the university is to build a portal that is able to control, coordinate, and receive feedback from the users (visitors, staff, and students). Such input contributes to developing a portal that will meet the expectations of end-users and the industry standards [2, 13].

King Saud University and King Abdul-Aziz University are among the top five universities of Saudi Arabia [14, 15]. Authors in [14] indicated that information technology infrastructure is a key success factor that increases students' confidence in using Saudi university portals. As such, the author in [15] recommended updating the information of the portals, e-services, and redesigning in order to increase their

use. He examined the attitudes of faculty members towards the use of virtual learning portals in teaching language programs at major Saudi universities. Authors in [16] evaluated and compared the awareness and the level of the use of portals among the students of the Al-Jouf University. The study concluded that the undergraduate students of the University lack the necessary information skills needed to meet their academic and research requirements in order to continue to use the portal. Authors in [17] examined different factors that affect the use of the university portals and determined the level of the use of e-learning portals in Saudi Arabia while they compared the results with the results from studies conducted in other countries. It was revealed that the technical infrastructure of the portal, instructor knowledge of the use of portal technologies, and student knowledge of using computer systems are important to the success and increase of the usage of the portal. The author in [18] conducted a comprehensive literature study to identify the current state of using the health information systems' portals in Saudi Arabia. The findings revealed that factors such as: design, privacy, and efficiency affect portal usage. Author in [19] examined the current situation of King Abdulaziz University portal content and proposed a content strategy, which can help the university to solve portal content issues and encourage users to use it. Also, this study recommended examining the students' perceptions regarding the use of the portal. The author in [20] examined the level of the use of academic portals by the staff of Alqassim University in Saudi Arabia. The results showed that the academic staff members lack awareness about the use of the available services in the portal. This study revealed that there is a need to increase the level of usage by training the academic members on how to use the services of the portal. Authors in [21] examined the level of use of university portals by identifying the factors that influence lecturers to use it. The result revealed that there are some issues such as: navigation, response time, minimum download time, and portal accessibility. This study recommended university decision makers and web designers to take into consideration these issues in order to increase the level of portal's usage. The author in [22] examined students' attitudes towards using electronic information resources of the library portal at Princess Nora University in Saudi Arabia. The results revealed that the management should attempt to decrease the gap between student attitudes and their use in order to increase the portal usage. In addition, authors in [10] examined the level of usage of the portal of the King Abdul-Aziz University among undergraduate students. It was found that students were dissatisfied with the portal. The study recommended the improvement of the level of the portal usage by solving the problems the students face.

In spite of the advancements being made in the domain of mobile portals in Saudi Arabia, there are aspects that present obstacles in the progress. For instance, authors in [23] ascertained several limitations related to mobile learning in Saudi Arabia, limitations related to processing performance, memory size, battery life, user interface, and the existence of various platforms. Other concerns encompass the availability of dependable and inexpensive web access from home, because the IT security in Saudi universities makes it difficult for students who are away from campus to access the university

systems. The author in [24] highlighted the substandard wireless network and technical concerns. Utilizing mobile phones for education would render likely issues regarding the hardware specifications of the devices involved: small screen sizes, inadequate input options, small battery life, inadequate memory capacity, and inadequate video quality [26, 27].

Based on these limitations, this study focuses on student web portals in Saudi Arabia universities. Social Cognitive Theory (SCT), which is regarded as a very powerful theory related to human behavior, was utilized. The SCT was applied for investigating computer utilization in [28, 29]. The nature of the proposed model along with its theory allowed its extension to the acceptance and usage of web portals. The SCT was regarded as the base model because: first, the constructs that were used in the model represented the variables that were analyzed as the factors responsible for improving the web portals. Second, since the SCT was a relatively less-represented model in the adoption of the web portals, it was important to investigate its performance and application in the existing context. Finally, this model included all the appropriate and vital constructs that helped in understanding the behavior of an individual. According to [25], user behavior is an important key to enhance and continue using a website. In addition, user behavior and perceived enjoyment on websites are important factors regarding the increase of its usage [26]. Also, the consideration of user behavior provides useful guidelines for designers to deliver the optimum website design. According to [32], student behavior is the most important key factor to improve the use of an educational portal. Therefore, understanding students' expectations has significant influence on their performance and increases user satisfaction [33, 34].

Based on the previous studies, it was found that there are issues regarding the usage of portals of Saudi Arabia universities. These issues decrease the student portals usage and make the students dissatisfied with them. Therefore, this study will investigate the factors that prevent the undergraduate students from using the university student portals.

III. MATERIALS AND METHODS

A. Population and Study Context

This study uses the quantitative approach to achieve its objectives by investigating the importance of student portals and the most important factors that affect students to use the portals of Saudi universities. The collected data were from 399 participants and were gathered from students of four different universities, namely the University of Tabuk, Taibah University, Taif University, and King Faisal University, which are located in different states of Saudi Arabia. Also, the study used the PLS algorithm in SmartPLS3 to analyze the collected data.

B. Factor Selection Method

Five databases were used, namely ScienceDirect, Taylor & Francis, Springer, Emerald, and Google Scholar to determine the importance of student portals and the most important factors which affect the students' usage. Moreover, this study attempts to investigate the importance of student portals among students based on their attitude in order to increase their

performance. There are works indicating the importance of studying personal and organizational factors to enhance users' performance. The factors that enhance academic staff in Nigeria to use ITC in the library were examined in [59]. The findings of the study revealed that organization factors are the most important. Organization factors include the lack of ICT strategy and the lack of commitment by the institutional management. According to [60], personal and behavioral factors are key success factors in improving the performance of employees. It is essential to examine the impact of personal, behavioral, and organizational factors on the academic staff in universities in order to improve the performance of research productivity [61]. Personal factors, such as the level of skills for using ICT, are important. Therefore, this study identified the factors that affect students' use of the university portals by using Microsoft Excel to extract and filter the factors from previous studies. This approach consisted of four steps namely, search for suitable factors, filtering factors, naming three classification categories, and classifying the factors into these categories.

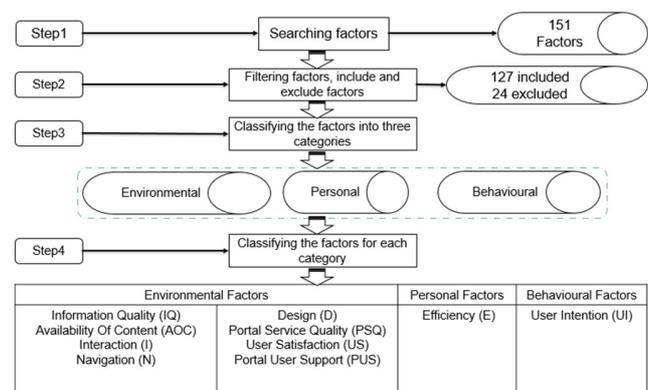


Fig. 1. Factor derivation process.

Firstly, 54 papers were reviewed. Twenty eight papers were found that were related to this study domain. From these included papers, 151 factors were identified. In the second step, the researchers filtered all identified factors to determine the included and excluded factors. In this step, the researchers identified 127 included sub-factors and excluded 24 factors. At the third step, the researchers extracted and classified the identified 127 sub-factors and factors into 3 categories, namely environmental, personal, and behavioral. Then, the researchers divided each category's sub-factors into the main factor. In the fourth step, the researchers identified the main factors related to each category. Regarding the environmental category, the researchers found that there are 8 main factors related with it. These factors are: Information Quality (IQ), Availability Of Content (AOC), Interaction (I), Navigation (N), Design (D), Portal Service Quality (PSQ), User Satisfaction (US), and Portal User Support (PUS). The personal category consists of 1 main factor namely Efficiency (E), and the behavioral category also consists of 1 main factor namely User Intention (UI), as shown in Figure 1.

1) Environmental Factors

Environmental factors refer to the degree an individual trusts the technical and organizational infrastructure that should exist and the level of support for using the system [27] or the availability of technological and external resources of the organization such as money, time, and effort needed to ease the performance of a specific behavior.

a) User Satisfaction

In information technology, user satisfaction of web portals has been studied extensively. Authors in [28, 36] studied the direct influence of competing resources, user satisfaction, and the voluntariness on the portal usage. It has been used as a surrogate measure for the effectiveness of information systems [29]. User satisfaction is defined as the degree of successful interaction between users and the information system [30]. Authors in [31] described satisfaction as the degree of the comparison of the expectations with the perceived performance. The first hypothesis regarding user satisfaction is:

- Ha1: US positively affects students' use of the student portal.

b) Information Quality

Information quality is the quality of the information system output that reflects user needs such as accuracy, content, format and timing [33], which is critical to user acceptance [33]. IQ can predict consumer satisfaction [34]. Accordingly, providing a high-quality student portal will help provide innovative services to students. The assumption made regarding information quality is:

- Ha2: IQ positively affects students' use of the student portal.

c) Interaction

According to [35], interaction refers to the ease by which users can interact with the input and output of the system and their ability to learn through it. Human-computer interaction is vital to improving user performance [36]. Better human-computer interaction provides good learning systems to create portals that are easy to use and increases user satisfaction [35]. As such, interaction helps improving student portal usage.

- Ha3: Interaction positively affects students' use of the student portal.

d) Availability Of Content

AOC can affect students' usage of the university portal [37]. Educational and non-educational services provided in the portal, such as online library services, email services, lecturer evaluation, and registration services are the key areas. The student portal should facilitate the accessibility to all services and be flexible, quick, and allow users to perform their tasks efficiently [38]. The students will be more satisfied when they get more available content.

- Ha4: AOC positively affects students' use of the student portal.

e) Portal Service Quality

Authors in [39] defined PSQ as the support from the information technology department concerning the user system. PSQ is also defined as the level of customers' expectations with the degree of the services delivered via the user information system [40]. According to [41], PSQ can be determined within the context of e-services as a portal that is responsive, available, accessible, and flexible. Therefore, PSQ is a key determinant of use and user satisfaction about the e-services provided in a portal [42].

- Ha5: PSQ positively affects students' use of the student portal.

f) Design

Portal design is considered an important factor that determines the usage of any online portal [43]. Portals through simple user interface design provide services and information that users need [44]. So, a student portal should include design features that help students enjoy their visit, which will lead to increased activity [45, 46].

- Ha6: Design positively affects students' use of the student portal.

g) Navigation

Navigation is a significant factor in using web portals [47]. Users should be able to move through portal links without getting lost. The ease of navigation is a critical component of portal usage [48]. Therefore, the definition of navigation refers to the ease of finding what the user wants in the portal [49]. The navigation hyperlinks should be available and must work properly [50].

- Ha7: Navigation positively affects students' use of the student portal.

h) User Support

US is the ability of the IT department to provide users with assistance when needed [51]. Moreover, because an academic portal is usually complex and needs qualified staff, it is necessary to have online US tools such as online chatting services provided via an instant messenger to help users resolve problems should they arise [28].

- Ha8: US positively affects students' use of the student portal.

Figure 2 presents the environmental factors which affect the students' use of a portal.

2) Personal Factors

This study suggests that the personal factors have a significant impact on the students' use of the portal. Moreover, personal factor efficiency refers to the judgements of people on their capabilities to organize and implement the required courses of action to achieve the required action [52].

a) Efficiency

A student portal is a good learning environment that helps students increase their abilities to learn [38]. Efficiency is a kind of self-evaluation that can assist the understanding of human behavior and its performance in specified tasks [53]. Even though student portals offer good information, their

efficiency will decrease if the students' ability does not meet the expectations [54]. The students' ability to use the portal is vital for them to achieve their needs, as shown in Figure 3.

- Hb1: Efficiency positively affects students with using the students' portal.

3) Behavioral Factors

Authors in [55] defined UI as the degree by which a person can formulate clear plans to perform a determined future behaviour (Figure 4).

a) User Intention

UI is defined by the individual's behavioral intention and, in turn, the behavioral intentions are a function of an individual's attitude to the behavior. This is suggested by the Theory of Reasoned Action (TRA). The attitude towards a type of behavior is defined by a person's feelings toward that action and can be negative or positive [56]. Authors in [57] stated that behavioral intention is an important independent variable that predicts the actual use. Similarly, authors in [58] stated that behavioral intention is likely to be linked with the actual usage.

- Hc1: UI positively affects user with using the students' portal.

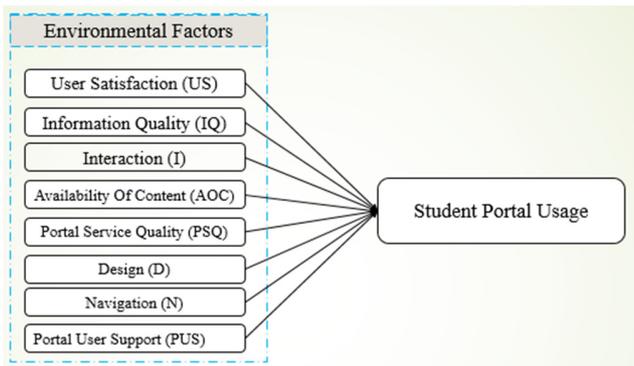


Fig. 2. Environmental factors.

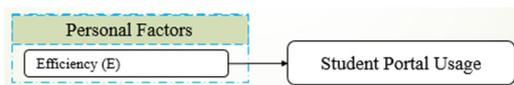


Fig. 3. Personal factors.

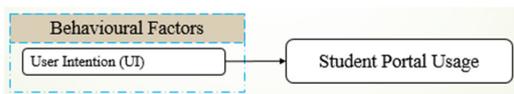


Fig. 4. Behavioral factors.

C. Survey Instruments

In this study, a questionnaire-based survey method was adopted and used to collect study data and test the hypotheses. The method includes data collection, data analysis, and report writing. Authors in [62] identified the questionnaire as recording the respondents' opinions on specific topic as a pre-formulated set of questions. Using a questionnaire method is efficient, less expensive, and less time-consuming compared to

other methods such as interviews and observation. The goal of the developed data collection survey instruments was to measure the validity and reliability of each factor in the model. According to [63], the developing of data collection survey instruments includes three steps: defining the constructs and content domain, generating and judging measurement items, and constructing studies to develop and refine the scale. The first step in data collection is to define each construct definition in the model. The importance in a theory to identify the construct definition and its domain was also highlighted, and conducting a comprehensive literature review to determine them was recommended. Therefore, the authors of this study did a thorough literature review and defined each construct in the model based on the content domain of the study. The second step of generating and judging measurement items follows. In addition, using the existing literature is considered a good scientific method to determine the items of the study [63]. Therefore, the authors of the current paper identified the items from previous studies that are related to this domain.

D. Face and Content Validity Assessment

In this step, the researchers measured the face and content validity of the measurement items [64]. According to [65], via referring to the literature or an expert, content and face validity can be achieved. A content and face validity form to test the validity of the selected measurement items was developed based on the validity form in [66]. The validity form consists of the research title, a short description of the research, the goal of this form, construct definitions, and items choice type. The items include three types of choice, Perfect Match (maintain items as it is), Moderate Match (maintain item but some refining is needed), and Poor Match (remove item) as shown in Appendix C. The researchers distributed the validation form to 9 experts as shown in Table I. The researchers invited experts from different disciplines to examine whether the items of each construct represent the construct's definition. Some items were changed and deleted based on the experts' comments and feedback.

TABLE I. EXPERTS' PROFILES

University	Department/School	Area of Expertise
Universiti Teknologi Malaysia	School of Computing	1 expert in Information systems, e-commerce, m-commerce.
Universiti Teknologi Malaysia	Azman Hashim International Business School	1 expert in Information systems, IT project management, information system planning.
Taibah University	Community College	2 experts in information systems and computer science.
Taibah University	Faculty of Computer Science and Information Systems	5 experts in information system and computer science, educational technology, e-learning.

E. Data Analysis Tools

To analyze the data, Partial Least Square (PLS) technique of Structural Equation Modeling (SEM) was used. By applying PLS-SEM, we were able to examine the measurement and the structural process of the proposed model. PLS-SEM is considered as a data analysis tool that validates predictive models [73]. SmartPLS software tool is appropriate for small sizes of data [73]. In this study, a two-type assessment to

analyze the data has been applied by using SmartPLS V.3. The first assessment is the measurement model, which is accomplished by the assessment of indicator reliability, internal consistency, convergent validity, and discriminant validity. The second assessment is the structural model, which is used to assess path analysis hypotheses testing by examining the size and significance of the path coefficient and coefficient determination [73].

IV. RESULTS AND DISCUSSION

This study examines the importance of student portal usage in Saudi universities by examining the most important factors affecting students' use of these portals. Ten factors were identified, 8 of which with a positive impact on students' usage of the student portal while the other 2 factors had no effect. These factors were divided into 3 main categories, namely environmental, personal, and behavioral. These 3 categories have three primary hypotheses. Furthermore, each primary hypothesis includes a sub-hypothesis depending on how many factors belong to each category in order to investigate the relationship between the constructs.

A. Demographic Background

As mentioned above, the study was conducted in 4 universities, in different states of Saudi Arabia. The total number of participants was 399. The distribution of the participants' gender indicates that the highest percentage of participants were female ($n=295$, 73%), while males were $n=104$, 26.1%. Regarding the participants' age distribution, 299 participants (75%) were aged between 18 and 24 years, 50 participants (12.5%) were in the range of 25 and 29 years, and the same number and percentage were more than 30 years old (50 participants, 12.5%). The distribution of participants in the selected universities is: University of Tabuk $n=77$, 19.3%, Taibah University $n=131$, 32.8%, Taif University $n=91$, 22.8%, and King Faisal University $n=100$, 25.1%. As for the participants' experience in using computer and internet, 51 participants (13%) had less than 1 year of experience, 57 participants 14.2% had 1 to 3 years, 67 participants 16.7% had 3 to 5 years, 103 participants 25.8% had 5 to 10 years, and 121 participants 30.3% had more than 10 years of experience. Two hundred and thirty (27.6%) participants had good computer skills, 154 participants had excellent computer skills (38.6%), and 15 had low computer skills (3.8%). Most (218) participants had excellent internet skills (54.6%), 171 had good internet skills (42.9%), and 10 had low internet skills (2.5%).

B. Data Analysis

This study used the PLS algorithm in SmartPLS3 to identify the reliability and validity of the study items. Measurement and structural models were used to measure the study model and hypotheses. The measurement model examines the constructs and their items. It also determines the degree to which the indicators measure the construct. The structural model evaluated the reliability and validity of the measurement model, determined the measurement of each item and concluded that all are reflective items. Authors in [67] summarize the standard criteria tests used to evaluate the measurement model. Cronbach's Alpha and Composite Reliability (CR) of 0.60 or greater are accepted. The Average

Variance Extracted (AVE) is another criterion for the assessment of the measurement model, and its accepted value is 0.5 or greater. As shown in Table II, the criteria of the measurement model of all construct items meet the standard criteria. Figure 5 presents the study measurement model.

TABLE II. THE CRITERIA OF THE MEASUREMENT MODEL

Types of Analysis	Test name	Standard criteria
Reliability of internal consistency	Cronbach's alpha	≥ 0.6 acceptable
	Composite reliability	≥ 0.6 acceptable
Convergent validity	Factor loading	≥ 0.7 acceptable
	AVE	≥ 0.5 acceptable
Discriminant validity	Cross loading	The outer loading of each construct should be more than the cross loading of the next constructs

1) Reliability of Internal Consistency

The reliability of internal consistency of the constructs was the first indicator that was evaluated. Moreover, composite reliability test and Cronbach's alpha test were used. Cronbach's alpha is a convenient test that uses inter-correlations to estimate the reliability indicator variables. The second test composites reliability which is considered another measure for the reliability of internal consistency that takes into consideration the different outer loading of indicator variables. Cronbach's alpha and composite reliability should be more than or equal to 0.60 for each construct [73]. As shown in Table III, the result of both tests indicated the reliability of internal consistency.

TABLE III. RELIABILITY OF INTERNAL CONSISTENCY

Construct name	Construct code	Cronbach's alpha	Composite reliability
Availability of Content	AOC	0.952	0.959
Design	D	0.954	0.961
Efficiency	E	0.929	0.945
Interaction	I	0.922	0.939
Information Quality	IQ	0.892	0.921
Navigation	N	0.916	0.947
Portal Service Quality	PSQ	0.953	0.962
Portal User Support	PUS	0.967	0.972
Student Portal Usage	SPU	0.942	0.954
User Intention	UI	0.933	0.946
User Satisfaction	US	0.964	0.97

2) Convergent Validity

According to [67], the convergent validity indicates the degree of which alternative procedures for the same construct are correlated positively. In this regard, there are two tests: the AVE and outer loadings of indicators that should be taken into account in this stage to evaluate the convergent validity of the constructs. Outer loading value should be 0.7 and above [67]. The AVE value for each construct should be 0.5 and above [67-69] as shown in Table IV.

3) Discriminant Validity

Discriminant validity is defined as the degree of which a specific construct is different from others [73]. Cross loadings is the first measure to evaluate the discriminant validity indicators. The outer loadings indicators should be greater than

any of its correlations to other constructs (see Appendix A). Fornell-Larcker is the second measure to evaluate discriminant validity. It compares the square root of the AVE values with the construct correlations. Fornell-Larcker criterion determines whether the AVE is bigger than the squared construct correlation with the other constructs [67]. The Fornell-Larcker results satisfy the criterion (Appendix B).

TABLE IV. CONVERGENT VALIDITY TESTS

Independent variable	Items	Loading	AVE
IQ	IQ1	0.804	0.699
	IQ2	0.847	
	IQ3	0.9	
	IQ4	0.842	
	IQ5	0.783	
AOC	AOC2	0.811	0.724
	AOC3	0.775	
	AOC4	0.879	
	AOC5	0.864	
	AOC6	0.836	
	AOC7	0.852	
	AOC8	0.883	
	AOC9	0.862	
I	I2	0.785	0.721
	I3	0.842	
	I5	0.897	
	I6	0.86	
	I7	0.842	
	I8	0.864	
N	N1	0.916	0.856
	N3	0.94	
	N4	0.92	
D	D2	0.805	0.733
	D3	0.871	
	D4	0.862	
	D8	0.871	
	D9	0.842	
	D10	0.859	
	D11	0.882	
	D12	0.846	
E	E1	0.764	0.741
	E2	0.867	
	E3	0.888	
	E4	0.868	
	E5	0.909	
	E6	0.861	
UI	UI1	0.788	0.715
	UI2	0.846	
	UI3	0.837	
	UI4	0.889	
	UI5	0.869	
	UI6	0.852	
	UI7	0.833	
PSQ	PSQ2	0.878	0.783
	PSQ3	0.758	
	PSQ4	0.906	
	PSQ5	0.883	
	PSQ6	0.915	
	PSQ7	0.928	
PUS	PUS1	0.843	0.815
	PUS2	0.889	
	PUS3	0.919	

US	PUS4	0.915	
	PUS5	0.922	
	PUS6	0.905	
	PUS7	0.91	
	PUS8	0.918	
	US1	0.872	0.822
	US2	0.923	
	US3	0.935	
US4	0.904		
US5	0.926		
US6	0.895		
US7	0.891		

4) Goodness of Fit (GoF)

GoF is a global fit measure [70]. It is the geometric mean of both the AVE and the average of R^2 of the endogenous variables [71]. The calculation formula of GoF is:

$$GoF = \sqrt{R^2 \times AVE} \quad (1)$$

The criteria to determine whether GoF values are no fit, small, medium, or large to be considered as global valid PLS model [72] are shown in Table V.

TABLE V. GOODNESS OF FIT

≤ 0.1	No fit
0.1 to 0.25	Small
0.25 to 0.36	Medium
≥ 0.36	Large

The value of the Gof of this study is 0.771, which means the GoF model of this study is large enough to be considered of sufficient global PLS model validity.

The next stage examines the hypotheses of this study using the structural model. This stage tests the predictive capability of the model and the relationship between the constructs. The path coefficients of the structural model are determined after the collinearity assessment, which represents the relationship of the hypotheses among the structural model constructs. The standardized values of the path coefficients are between -1 and +1. If the path coefficients are close to 1, it means that there is a strong and statistically significant positive relationship among the constructs. If the path coefficients are close to 0, then the relationship between the constructs is considered weak [68]. In this study, the one-tailed test was used because the relationships of the hypotheses are positive and direct. The critical values and p-values for one-tailed tests are 1.28 (with significance level=10%), 1.65 (with significance level=5%), and 2.33 (with significance level=1%) [67] as shown in Table VI. Figure 6 presents the path coefficients of the structural model using smartPLS3.

TABLE VI. PATH RELATIONSHIPS

Analysis	Test	Standard
Path relationship	p value	* < 0.10 ** < 0.05 *** < 0.01
	t value	>1.28 level of significance =10% >1.96 level of significance = 5% >2.33 level of significance = 1%

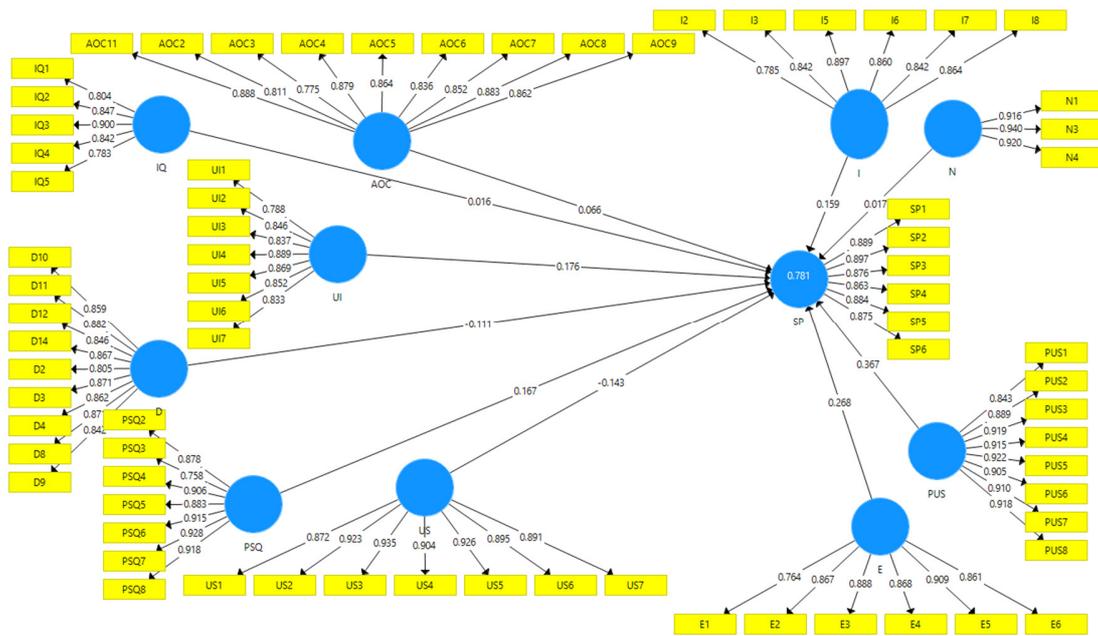


Fig. 5. Study measurement model.

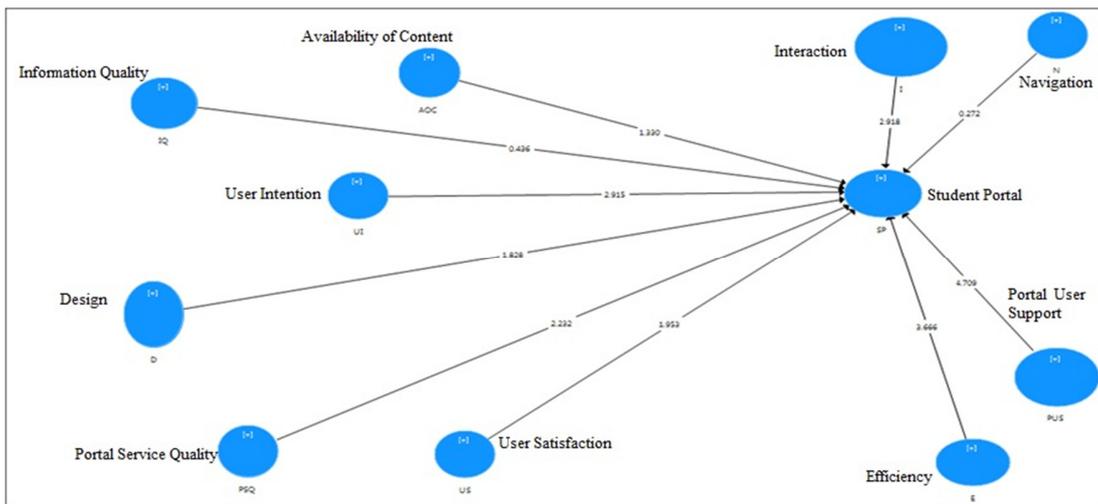


Fig. 6. Path coefficients of the structural model.

TABLE VII. RESULTS OF HYPOTHESIS TESTING

Factor path	t value	p value	Significant level	Decision
AOC → SP	1.33	0.092	*	Supported
D → SP	1.828	0.034	**	Supported
E → SP	3.666	0	***	Supported
I → SP	2.918	0.002	***	Supported
IQ → SP	0.436	0.331	NS	Not supported
N → SP	0.272	0.393	NS	Not supported
PSQ → SP	2.232	0.013	***	Supported
PUS → SP	4.709	0	***	Supported
UI → SP	2.915	0.002	***	Supported
US → SP	1.953	0.025	**	Supported

The result of hypotheses testing shows that most of the hypotheses are significant. Moreover, by using a bootstrapping test in smart-PLS, the result shows that 8 out of 10 hypotheses

are significant. Table VII presents the results of hypotheses testing.

As presented above, this study has 10 hypotheses. Ha1 stated that US has a significant and positive affect on students' use of the student portal. Previous studies indicated that the satisfaction of users will provide a surrogate which is significant to the critical product of the existing or new information system [30]. US is in terms of the use of the system and the acceptance of the portal, which is important in measuring the success of the information system [73]. Therefore, in this study, US about the portal is considered significant based on preliminary studies and the final result of data collection which reveals that there is a positive relationship between US and the use of the student portal (t-value=1.953, p-value=0.018). Saudi universities should take

into account the US when they improve their portals based on their students' needs and demands.

Ha2 states that IQ has a positive effect on students' use of the student portal. According to [38], IQ is the output quality of an information system, which reflects dimensions such as content, accuracy, timing, and format. It is an important factor because providing high-quality information to users in a university student portal can help in delivering innovative services to students. The results show a negative relationship between IQ and the use of the students' portal (t-value=0.436, p-value=0.331) because the student portal does not create new information nor meet students' needs based on students' point of view. However, the quality of the information provided on the portal should meet the students' needs. Thus, universities in Saudi Arabia have to improve their student portals based on their students' needs and demands.

Ha3 states that Interaction has a positive and significant effect on students' use of the student portal. Authors in [35] defined interaction as the ease by which users interact with the input and output of the system and their ability to learn through it. Interaction is one of the main factors that contribute to positive learning outcomes in a learning environment [74]. The findings of this study reveal that there is a positive relationship between Interaction and using the student portal (t-value=2.918, p-value=0.002). Thus, Interaction is considered an important indicator that can help improve student portal usage.

Ha4 states that the AOC has a positive effect on students' use of the student portal [37]. This is shaped by the educational and non-educational services provided in the portal, such as online library services, email services, lecturer evaluation, and registration services. A student portal should facilitate easy access to all services and be flexible, quick, and allow users to perform their tasks [38]. The findings of this study showed that there is a positive relationship between the availability of content and the usage of the student portals (t-value=1.33, p-value=0.092). Thus, the AOC will help Saudi universities to improve their student portals based on their students' demands and expectations.

Ha5 states that PSQ has a significant and positive effect on students' use of the student portal. The author in [40] defined PSQ as a measure of the level of customers' expectations with how well the information system offers services to users. Service quality within the context of the portal can be determined by prompt, responsive, available, accessible, and flexible e-services [41]. Therefore, PSQ is a key determinant for the use and US about e-services provided via a portal [42]. The result of this study reveals a positive relationship between PSQ and the use of a students' portal (t-value=2.232, p-value=0.013). Thus, PSQ is a significant indicator that can help Saudi universities improve their student portals.

Ha6 states that portal design has a significant and positive effect on the use of the student portal. According to [75], university student portals are designed to make the member's experience with the portal as personalized as possible. The design features that attract users and increase their satisfaction include enjoyment, excitement, participation, and charm [45].

Design can motivate them to participate and promote excitement to enjoy while visiting the website. Therefore, student portal design features that help students to enjoy their visit lead to increased activity [46]. This study reveals that there is a positive relationship between design and use of the students' portal (t-value=1.828, p-value=0.034). Thus, the portal design is paramount to improving student portal usage.

HA7 states that navigation has a positive effect on students' use of the student portal. Usually, a navigation bar offers the main tasks of the website [76]. According to [47], navigation is a significant factor in learning the web portal. Users should be able to move through portal links without getting lost. The ease of navigation is a critical component of portal usage [48]. The result of this study shows that there is a negative relationship between Navigation and use of portals (t-value=0.272, p-value=0.393) which means that Navigation is not regarded as an essential predictor for the use of a student portal from the viewpoint of students. Anyhow, the student portal navigation does not attract students or meet their needs, so it will be useful for Saudi universities to improve it, provide more links that students need in their study, and make the portal easier to use and easier to find the necessary information.

Ha8 states that PUS has a significant and positive effect on students' use of the student portal. Authors in [51] define user support as the ability of the IT department to assist users when needed. Moreover, because an academic portal is usually complex and needs qualified staff, it is necessary to have online user support tools to help the users [28]. Students (users) should be given specified training and user support, to increase the ease of use and learnability to effectively and efficiently fulfil a specified range of task scenarios on the web portal while management achieves their economic and global attainments [77]. This study reveals that there is a positive relationship between PUS and the use of the student portals (t-value=4.709, p-value=0).

Hypothesis Hb1 states that Efficiency has a significant and positive effect on students' use of the student portal. As a personal factor, Efficiency is defined as the judgements of people on their capabilities to organize and implement the required courses of action to achieve a required goal [52]. Efficiency was shown to be an important indicator that determines the actual behavior of a person [78]. Furthermore, the student portal is a good learning environment [38]. Even though student portals offer good information, their efficiency will decrease if the student ability does not meet the suitable level [54]. This study reveals that there is a positive relationship between the efficiency factor and the use of students' portals (t-value=3.666, p-value=0).

The last hypothesis is that UI has a significant impact on students' use of the student portal. Moreover, as a behavioral factor, UI is defined as a person's ability to formulate clear plans to perform a determined future behaviour [55]. Many factors might affect students' behavioral intention to use the student portal. Behavioral intention predicts or at least is linked with the actual usage [57, 58]. This study revealed that there is a positive relationship between UI and using the students' portal (t-value=2.915, p-value=0.002).

This study developed and validated a model based on data collected via a survey of students in 4 Saudi universities. SmartPLS3 was used to analyze the data. The findings indicate that 8 out of 10 considered factors are significant and have a positive effect on student portal usage. These factors are User Satisfaction, Design, Availability Of Content, User Intention, Efficiency, Portal User Support, Portal Service Quality, and Interaction. On the other hand, Information Quality and Navigation had no significant effect on portal usage. The model was able to discover students' expectations about the portals of Saudi universities, so it can help them to improve their students' portal usage. Figure 7, presents the model of this study along with the significance results.

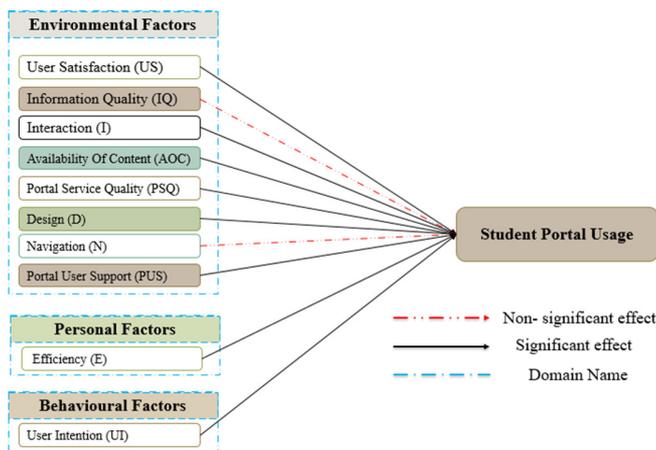


Fig. 7. The model of this study.

The model of this study has established the context of what students need to use the university portals. This model recommends university management and IT departments to provide an appropriate electronic e-service and suitable ICT environment that support the portal users. The model of this study also identified the most important factors that affect the students to use the portals. By taking into account these factors in the portal development process, universities can improve their portals. In addition, this study is one of the few studies that provide the factors that help universities build suitable electronic systems to develop and improve their university portal usage.

V. CONCLUSION AND FUTURE WORK

This study studied the importance of student portals by investigating the most important factors that attract undergraduate students in Saudi universities to use them. The authors divided the identified factors into environmental, personal, and behavioral. Each category included sub-factors. The environmental category included 8, the personal category included 1, and the behavioral category also included 1. This study found that the 3 categories positively affect students' portal usage. Moreover, the 10 hypotheses regarding the factors and their results presented the relationship between independent variables and the dependent variable, which indicates that UI, D, AOC, UI, E, PUS, PSQ, and I have a positive effect on students' use of the student portal. On the

other hand, IQ and N do not have a significant effect on students' portal usage. Future work can investigate these factors from other perspectives, such as the academic and administrative staff point of views. This study was conducted on undergraduate students in Saudi Arabia universities. Future work can extend and consider postgraduate students. This study can be considered a good start for researchers to investigate more on the usage of electronic resources and ICT in university portals to improve the needs of other university users.

APPENDIX A

Cross Loadings of Discriminant Validity

	AOC	D	E	I	IQ	N	PSQ	PUS	SPU	UI	US
AOC2	0.811	0.611	0.593	0.605	0.592	0.662	0.658	0.511	0.58	0.581	0.63
AOC3	0.775	0.512	0.526	0.667	0.528	0.543	0.543	0.426	0.553	0.552	0.503
AOC4	0.879	0.589	0.591	0.666	0.598	0.701	0.621	0.531	0.603	0.58	0.605
AOC5	0.864	0.626	0.625	0.638	0.533	0.669	0.645	0.585	0.602	0.638	0.64
AOC6	0.836	0.623	0.578	0.616	0.503	0.611	0.596	0.535	0.568	0.585	0.617
AOC7	0.852	0.655	0.604	0.672	0.553	0.657	0.61	0.535	0.573	0.607	0.607
AOC8	0.883	0.662	0.625	0.681	0.571	0.723	0.654	0.586	0.623	0.608	0.657
AOC9	0.862	0.648	0.648	0.689	0.597	0.663	0.629	0.619	0.667	0.666	0.65
AOC11	0.888	0.645	0.633	0.655	0.556	0.712	0.656	0.619	0.616	0.676	0.669
D2	0.598	0.805	0.592	0.611	0.464	0.591	0.601	0.555	0.544	0.614	0.609
D3	0.645	0.871	0.671	0.641	0.538	0.658	0.652	0.616	0.603	0.672	0.68
D4	0.614	0.862	0.627	0.633	0.503	0.63	0.662	0.573	0.584	0.653	0.687
D8	0.692	0.871	0.664	0.679	0.514	0.681	0.727	0.593	0.591	0.688	0.735
D9	0.55	0.842	0.611	0.565	0.455	0.566	0.62	0.571	0.562	0.659	0.64
D10	0.615	0.859	0.65	0.631	0.56	0.635	0.636	0.616	0.631	0.688	0.637
D11	0.633	0.882	0.661	0.632	0.49	0.671	0.669	0.59	0.602	0.698	0.681
D12	0.666	0.846	0.686	0.642	0.472	0.679	0.663	0.638	0.587	0.629	0.677
D14	0.6	0.867	0.679	0.612	0.431	0.666	0.692	0.623	0.594	0.677	0.72
E1	0.541	0.605	0.764	0.561	0.498	0.573	0.599	0.635	0.645	0.649	0.611
E2	0.628	0.677	0.867	0.602	0.443	0.647	0.755	0.723	0.69	0.658	0.787
E3	0.635	0.697	0.888	0.63	0.509	0.659	0.7	0.73	0.7	0.675	0.75
E4	0.619	0.618	0.868	0.582	0.436	0.561	0.677	0.726	0.718	0.666	0.739
E5	0.647	0.683	0.909	0.629	0.49	0.657	0.754	0.746	0.731	0.708	0.805
E6	0.592	0.637	0.861	0.582	0.501	0.594	0.645	0.736	0.742	0.704	0.64
I2	0.62	0.585	0.521	0.785	0.514	0.565	0.589	0.51	0.537	0.607	0.544
I3	0.606	0.57	0.519	0.842	0.558	0.591	0.536	0.474	0.612	0.613	0.535
I5	0.69	0.619	0.58	0.897	0.53	0.679	0.601	0.527	0.615	0.652	0.579
I6	0.62	0.608	0.587	0.86	0.543	0.69	0.578	0.549	0.589	0.59	0.58
I7	0.667	0.659	0.636	0.842	0.495	0.732	0.637	0.597	0.66	0.629	0.623
I8	0.711	0.683	0.68	0.864	0.518	0.728	0.691	0.59	0.653	0.645	0.663
IQ1	0.575	0.501	0.508	0.535	0.804	0.489	0.493	0.478	0.523	0.523	0.49
IQ2	0.546	0.442	0.436	0.499	0.847	0.434	0.408	0.376	0.428	0.47	0.406
IQ3	0.594	0.522	0.505	0.541	0.9	0.498	0.483	0.444	0.495	0.534	0.47
IQ4	0.518	0.459	0.46	0.502	0.842	0.449	0.42	0.399	0.465	0.479	0.418
IQ5	0.505	0.472	0.399	0.505	0.783	0.367	0.4	0.381	0.396	0.489	0.372
N1	0.714	0.69	0.653	0.722	0.527	0.916	0.667	0.595	0.642	0.653	0.657
N3	0.716	0.677	0.658	0.74	0.498	0.94	0.683	0.612	0.663	0.643	0.664
N4	0.729	0.717	0.674	0.719	0.473	0.92	0.696	0.624	0.633	0.671	0.669
PSQ2	0.676	0.7	0.731	0.678	0.478	0.711	0.878	0.621	0.675	0.738	0.792
PSQ3	0.574	0.642	0.635	0.632	0.498	0.572	0.758	0.62	0.7	0.702	0.661
PSQ4	0.656	0.681	0.703	0.651	0.46	0.661	0.906	0.641	0.663	0.676	0.798
PSQ5	0.66	0.685	0.684	0.617	0.477	0.649	0.883	0.605	0.619	0.647	0.772
PSQ6	0.644	0.656	0.701	0.599	0.451	0.622	0.915	0.615	0.638	0.663	0.803
PSQ7	0.658	0.686	0.742	0.625	0.446	0.676	0.928	0.642	0.67	0.68	0.827
PSQ8	0.668	0.698	0.751	0.609	0.468	0.661	0.918	0.648	0.664	0.669	0.845
PUS1	0.615	0.633	0.746	0.551	0.47	0.608	0.622	0.843	0.714	0.634	0.679
PUS2	0.638	0.634	0.768	0.622	0.481	0.624	0.644	0.889	0.771	0.679	0.692
PUS3	0.578	0.654	0.759	0.585	0.431	0.595	0.64	0.919	0.728	0.661	0.683
PUS4	0.576	0.618	0.752	0.563	0.445	0.568	0.634	0.915	0.716	0.647	0.656
PUS5	0.606	0.654	0.791	0.609	0.472	0.623	0.675	0.922	0.769	0.694	0.714
PUS6	0.543	0.654	0.745	0.549	0.413	0.562	0.648	0.905	0.7	0.628	0.7
PUS7	0.551	0.584	0.712	0.556	0.428	0.586	0.621	0.91	0.743	0.651	0.649
PUS8	0.575	0.608	0.743	0.575	0.474	0.593	0.652	0.918	0.747	0.637	0.675
SPU1	0.647	0.625	0.737	0.65	0.504	0.653	0.656	0.717	0.889	0.685	0.657
SPU2	0.647	0.599	0.726	0.648	0.487	0.643	0.649	0.709	0.897	0.678	0.633
SPU3	0.55	0.542	0.662	0.584	0.468	0.535	0.578	0.674	0.876	0.629	0.562

SPU4	0.553	0.556	0.649	0.611	0.487	0.547	0.598	0.662	0.863	0.653	0.575
SPU5	0.657	0.642	0.768	0.645	0.493	0.634	0.724	0.768	0.884	0.697	0.728
SPU6	0.654	0.656	0.771	0.668	0.498	0.659	0.738	0.768	0.875	0.722	0.728
UI1	0.641	0.724	0.692	0.607	0.455	0.663	0.755	0.612	0.615	0.788	0.768
UI2	0.665	0.708	0.684	0.638	0.486	0.686	0.709	0.63	0.638	0.846	0.727
UI3	0.515	0.565	0.598	0.566	0.478	0.511	0.547	0.555	0.636	0.837	0.55
UI4	0.643	0.72	0.701	0.687	0.528	0.632	0.687	0.639	0.697	0.889	0.687
UI5	0.594	0.629	0.662	0.608	0.538	0.553	0.624	0.601	0.652	0.869	0.614
UI6	0.588	0.621	0.668	0.602	0.517	0.579	0.655	0.643	0.689	0.852	0.646
UI7	0.612	0.629	0.65	0.632	0.537	0.573	0.607	0.61	0.631	0.833	0.622
US1	0.688	0.753	0.748	0.637	0.481	0.692	0.795	0.651	0.661	0.711	0.872
US2	0.67	0.703	0.764	0.615	0.465	0.638	0.833	0.68	0.681	0.697	0.923
US3	0.709	0.742	0.783	0.669	0.499	0.701	0.822	0.665	0.671	0.715	0.935
US4	0.643	0.689	0.768	0.639	0.465	0.657	0.801	0.715	0.683	0.694	0.904
US5	0.658	0.735	0.766	0.634	0.463	0.65	0.802	0.682	0.653	0.724	0.926
US6	0.627	0.692	0.779	0.592	0.459	0.609	0.794	0.68	0.631	0.694	0.895
US7	0.638	0.685	0.785	0.613	0.471	0.602	0.795	0.709	0.708	0.706	0.891

APPENDIX B

Fornell-Larcker of Discriminant Validity

	AOC	D	E	I	IQ	N	PSQ	PUS	SPU	UI	US
AOC	0.851										
D	0.729	0.856									
E	0.71	0.759	0.861								
I	0.77	0.733	0.694	0.849							
IQ	0.658	0.575	0.557	0.619	0.836						
N	0.778	0.75	0.715	0.785	0.54	0.925					
PSQ	0.734	0.769	0.801	0.714	0.531	0.737	0.885				
PUS	0.649	0.698	0.833	0.639	0.501	0.66	0.711	0.903			
SPU	0.705	0.688	0.819	0.722	0.556	0.698	0.75	0.816	0.881		
UI	0.719	0.776	0.786	0.733	0.599	0.708	0.774	0.725	0.771	0.846	

APPENDIX C

The questionnaire

Please answer the following questions by choosing the correct information representing your demographic.
1. Your gender: (Male/Female)
2. Your age: (18 - 24 years/25 - 29 years/ More than 30)
3. Your University Name: (Taibah University/ Taif University/ King Faisal University/ University of Tabuk)
4. Years of experience using computer and internet: (less than 1 year/1 - 3 years/3 - 5 years/5 - 10 years/ More than 10 years)
5. Computer skills: (Low/Good/ Excellent)
6. Internet skills: (Low/Good/ Excellent)

B. For the following questions, please rate your perception with regard to your experience in using student portals. (Possible choices: Strongly disagree, Disagree, Neutral, Agree, Strongly Agree)
1. Information Quality (IQ): Refers to the system's output value, as perceived by the users
I believe that the easy to understand information on the student portal encourages me to use the portal.
I believe the student portal provides accurate information to me as a student.
I believe the information displayed on the student portal is useful to me as a student.
I believe the student portal provides information related to my study needs.
I believe the rich information available on the student portal attracts me to use the portal.
2. Availability of Content (AOC): Refers to the availability of portal content and services which provide to student in their portal.
I believe that if the availability of information content in the student portal meets my needs, it will encourage me to use the portal.
I believe that if online discussions were available in the student portal, that would help me increase my interaction with the portal.
The availability of online course catalogues in the student portal will help me in my study.
I believe that if the student portal provides complete information about the university and the various events, it will attract me to use it.
I believe having contact information such as phone number, fax number, e-mail, and postal address in the portal is helpful to me as a student.
I believe that giving feedback and facilitated question forms in the student portal will increase the interaction with the portal.
I believe that if a help facility was provided in the student portal, that will make the portal more usable.
I believe that if the student portal provides e-mail services to students will help me interact with the portal administration.
I believe that if the student portal contains an e-learning material service, will make my materials easy to access on the portal.
3. Interaction (I): Refers to the interaction between the students, instructors, and the portal services.
I believe that the interaction in the student portal enhances the effectiveness of my academic decisions as a student.
I believe that the interaction in the student portal enhances my productivity in my academic life as a student.
I believe that if the student portal makes online discussion tasks easy to use, will increase my interaction with it.
I believe the interaction allows me to share experiences with other students in the portal.
I believe the instructor can encourage students to increase their interaction with the student portal.
I believe interactive feedback between students and teachers in the student portal will encourage both of them to use it.
4. Navigation (N): Assesses whether a portal includes tools (i.e. navigation menu, internal search facility) and links which facilitate users' navigation.
I believe that if all links are working in my student portal, that will make it more interesting to use
I believe that if the internal search is effective in the student portal, that will help me in my study.
I believe that if the internal search is working in the student portal, that will help me to use it.
5. Design (D): Refers to the visual attractiveness of the portal design, the appropriate design of a portal pages and the appropriate use of images, fonts, colors in the portal design.
I believe that if the images in the student portal are interesting, that will increase my satisfaction of the portal
I believe that if there are consistent colors throughout the student portal, that will make the portal more attractive.
I believe that if the student portal menu design is consistent, that will make it more interesting to use.
I believe that if the content text is simple in the student portal, that will make the portal easier to use.
I believe that if the layout of the student portal is very simple, that will make the portal more attractive.
I believe that if the alignment of the header on each page is consistent in the student portal, that will make it more attractive.
I believe that if all pages have proper headings, that will make the portal more interesting to use.

I believe that if consistent language is used throughout the student portal, that will make the portal easier to use.
I believe that the design makes pleasurable for students to use the student portal services.
6. User Intention (UI): Refers to the behavioral of user intention to use the student portal services.
I believe that if the portal services are worthwhile for students, that will make me use the portal more.
If I get good services, I will use the student portal many times per week.
I plan to use the student portal very often in the future.
If the student portal is enjoyable, i intend to use it in the future.
I intend to use the student portal to conduct university related study tasks.
I predict i will use the student portal in the future.
If I get to use the student portal, i intend to use all of its services.
7. Portal Service Quality (PSQ): Refers to the users evaluation that the service they are receiving from the portal is the service they expect.
I believe the services offered in the student portal should be reliable.
I believe the services offered in the student portal save my time.
I believe the pages should respond fast when I use the portal services.
I believe the student portal should provide sufficient information to achieve my tasks.
I believe the student portal should be flexible to use.
I believe the services offered in the student portal should be comprehensive for all student needs.
I believe the student portal should provide up-to-date information.
8. User Satisfaction (US): Refers to student satisfaction about the provided services.
If the quality of information provided through the student portal is good, i will feel satisfied.
If the quality of the student portal system is good, i will feel satisfied.
If the services offered in the student portal meet my needs, that will increase my satisfaction with the portal.
I believe that if the student portal is enjoyable, it will make me satisfied.
If the services offered by the student portal are suited to my needs, that will make me satisfied.
I believe the services offered by the student portal should be easy to use.
I believe that if the process of using the student portal is pleasant, that will make me satisfied.
9. Efficiency (E): Mean users have the ability to organize and perform the actions on the portal in order to achieve their tasks.
I have the ability to download programs and use student portal services.
I believe the student portal services should be easy to use.
I believe that if the information in the student portal is interesting to me, it should be obtained with minimal effort.
I believe that if i find that anyone can use the student portal easily, that will encourage me to use the portal.
I believe that if I can easily navigate the student portal, that will help me to find what i need easily.
I believe the student portal serves its purpose by producing good services.
10 Portal User Support (PUS): The degree to which the usersupport departments are willing to serve the users and provide them with the required services.
I believe the portal loads its pages quickly if user support provides high internet speed.
I believe the user support will help users by providing training in the use of the student portal.
I believe the user support will help users by providing quick responses to their service requests.
I believe the user support will help users by providing correct information.
I believe the user support will help users by providing different browsers where portal pages work.
I believe the user support will help users by providing a help desk.
I believe the user support will help users by providing feedback forms.
I believe the user support will help users by providing orientation in the use of students.
11. Student Portal (SP): Student portal provides services which matche the study needs from different resources in a single page.
I believe using the student portal will provide me with information that would lead me to produce better research.
I believe using the student portal will make it easier for me to do my assignments and prepare for examinations.
I believe using the student portal will give me access to information that I cannot find elsewhere.
I believe using the student portal will provide me with information that would lead me to produce better research.

REFERENCES

- [1] N. Nasirun, S. M. Noor, Z. M. Nor, H. Ahmat, and Z. Ahmad, "Perceived Web Service Quality for Students' Portal in Higher Learning Institution," in *International Proceedings of Economics Development and Research*, 2012, vol. 56, pp. 52–56.
- [2] H. Al-Dossari, "A Heuristic Based Approach for Usability Evaluation of Academic Portals," *International Journal of Computer Science & Information Technology*, vol. 9, no. 3, pp. 15–30, 2017.
- [3] M. Mansourvar and N. M. Yasin, "Web portal As A Knowledge Management System In The Universities," *World Academy of Science, Engineering and Technology*, vol. 70, pp. 968–974, 2010.
- [4] N. Thomas-Alvarez and L. Mahdjoubi, "Testing the effectiveness of a web-based portal system for the building control sector," *Automation in Construction*, vol. 29, pp. 196–204, Jan. 2013, <https://doi.org/10.1016/j.autcon.2012.02.018>.
- [5] A. Altameem, "What drives successful e-learning? An empirical investigation of the key technical issues in Saudi Arabian universities," *Journal of Theoretical and Applied Information Technology*, vol. 53, no. 1, pp. 63–70, 2013.
- [6] S. M. Al-Turki, "Assessment of information technology awareness and usage in higher education in Saudi Arabia: sample study in Jazan and King Faisal University KSA," *International Journal of Advanced Research in IT and Engineering*, vol. 3, no. 9, pp. 1–17, 2014.
- [7] M. Althonayan and A. Papazafeiropoulou, "Evaluating the Performance of ERP Systems in Saudi Arabian Higher Education: A Stakeholders' Perspective," in *2nd International Conference on Information Management and Evaluation*, London, UK, Apr. 2011, pp. 473–482.
- [8] K. A. Al-Busaidi, "The payoff of corporate portal usage in an academic institution," *Campus-Wide Information Systems*, vol. 29, no. 5, pp. 368–379, Jan. 2012, <https://doi.org/10.1108/10650741211275125>.
- [9] H. Gul, S. Z. Iqbal, and M. Saqib, "Usability Evaluation of an Educational Website in Saudi Arabia," *VAWKUM Transactions on Computer Sciences*, vol. 3, no. 1, pp. 84–92, Dec. 2015, <https://doi.org/10.21015/vtcs.v8i2.382>.
- [10] M. Benaïda, A. Namoun, and A. Taleb, "Evaluation of the Impact of Usability in Arabic University Websites: Comparison between Saudi

- Arabia and the UK," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 8, pp. 365–375, 2018.
- [11] K. Al-Omar, "Evaluating the Internal and External Usability Attributes of E-Learning Websites in Saudi Arabia," *Advanced Computing: An International Journal*, vol. 8, no. 3/4, pp. 1–12, Jul. 2017.
- [12] Naiflabi, H. Alghamdi, and F. Demir, "The Usability Of King Khalid University Website: Assessing Effectiveness, Efficiency, And Satisfaction," *International Journal of Arts Humanities and Social Sciences*, vol. 3, no. 7, pp. 10–16, 2018.
- [13] Y. F. Hassan, "Cellular Automata For Adaptive Web Portal Structure In Egyptian Universities," *International Journal of Engineering Sciences & Emerging Technologies*, vol. 6, no. 2, pp. 133–141, 2013.
- [14] S. Bahaj, K. Aljaaidi, and T. Ahmed, "Using TAM model to empirically examine students' attitudes towards e-services in college of business administration," *Management Science Letters*, vol. 9, no. 5, pp. 651–660, 2019.
- [15] M. Al Harbi, "Faculty Attitudes Toward and Motivation for Virtual Learning Environments (VLE) for Language Studies: A Cross-National Study in Saudi Arabian Universities," *Journal of Psychology and Behavioral Science*, vol. 4, no. 2, pp. 100–114, 2016.
- [16] A. Ahmed and S. Al-Ryaeae, "Knowledge and Use of Electronic Information Resources by Medical Students at Al-Jouf University in Saudi Arabia," *Library Philosophy and Practice*, vol. 1524, Mar. 2017.
- [17] A. Alhabeeb and J. Rowley, "Critical success factors for eLearning in Saudi Arabian universities," *International Journal of Educational Management*, vol. 31, no. 2, pp. 131–147, Jan. 2017.
- [18] M. Alshamari, "Usability Factors Assessment in Health Information System," *Intelligent Information Management*, vol. 8, no. 6, pp. 170–180, Nov. 2016, <https://doi.org/10.4236/iim.2016.86012>.
- [19] A. M. Elsayed, "Web content strategy in higher education institutions: The case of King Abdulaziz University," *Information Development*, vol. 33, no. 5, pp. 479–494, Nov. 2017, <https://doi.org/10.1177/0266666916671387>.
- [20] N. J. H. Al-Mashaqba, "Use of Blackboard Collaborate at Qassim University: Evaluating Teachers' Perceptions," *Sino-US English Teaching*, vol. 14, no. 2, pp. 76–85, 2017, <https://doi.org/10.17265/1539-8072/2017.02.002>.
- [21] S. M. Almahamid, A. F. Tweiqat, and M. S. Almanaseer, "University website quality characteristics and success: lecturers' perspective," *International Journal of Business Information Systems*, vol. 22, no. 1, pp. 41–61, Jan. 2016, <https://doi.org/10.1504/IJBIS.2016.075717>.
- [22] L. Alkahtani, "The Attitudes of Princess Nora University Students Towards Using Electronic Information Resources of the Library," *Proceedings from the Document Academy*, vol. 3, no. 2, Dec. 2016, Art. no. 8, <https://doi.org/10.35492/docam/3/2/8>.
- [23] A. Badwelan, S. Drew, and A. A. Bahaddad, "Towards Acceptance M-Learning Approach in Higher Education in Saudi Arabia," *International Journal of Business and Management*, vol. 11, no. 8, pp. 12–30, 2016.
- [24] S. S. Alkhalaf, "Evaluating m-learning in Saudi Arabian higher education: a case study," *arXiv:1510.03189 [cs]*, Oct. 2015, [Online]. Available: <http://arxiv.org/abs/1510.03189>.
- [25] M. Y. Wu, H. P. Chou, Y. C. Weng, and Y. H. Huang, "TAM-2 based study of website user behavior-using web 2.0 websites as an example," *WSEAS Transactions on Business and Economics*, vol. 8, no. 4, pp. 133–151, 2011.
- [26] A. Benlian, "Web Personalization Cues and Their Differential Effects on User Assessments of Website Value," *Journal of Management Information Systems*, vol. 32, no. 1, pp. 225–260, Jan. 2015, <https://doi.org/10.1080/07421222.2015.1029394>.
- [27] 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>.
- [28] Y.-H. Chen and I. Chengalur-Smith, "Factors influencing students' use of a library Web portal: Applying course-integrated information literacy instruction as an intervention," *The Internet and Higher Education*, vol. 26, pp. 42–55, Jul. 2015, <https://doi.org/10.1016/j.iheduc.2015.04.005>.
- [29] W. H. DeLone and E. R. McLean, "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research*, vol. 3, no. 1, pp. 60–95, Mar. 1992, <https://doi.org/10.1287/isre.3.1.60>.
- [30] P. Longinidis and K. Gotzamani, "ERP user satisfaction issues: insights from a Greek industrial giant," *Industrial Management & Data Systems*, vol. 109, no. 5, pp. 628–645, Jan. 2009, <https://doi.org/10.1108/02635570910957623>.
- [31] P. T. Kotler and K. L. Keller, *Marketing Management*, 14th edition. Upper Saddle River, NJ, USA: Pearson, 2011.
- [32] A. Tella and M. T. Bashorun, "Undergraduate Students' Satisfaction with the Use of Web Portals," *International Journal of Web Portals*, vol. 4, no. 2, pp. 56–73, Apr. 2012, <https://doi.org/10.4018/jwp.2012040104>.
- [33] V. McKinney, K. Yoon, and F. "Mariam" Zahedi, "The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach," *Information Systems Research*, vol. 13, no. 3, pp. 296–315, Sep. 2002, <https://doi.org/10.1287/isre.13.3.296.76>.
- [34] L. C. Schupp, "Web Site Success: Antecedents of Web Site Satisfaction and Re-use," *Journal of Internet Commerce*, vol. 9, no. 1, pp. 42–64, Jun. 2010, <https://doi.org/10.1080/15332861.2010.487414>.
- [35] S. Thuseethan, S. Achchuthan, and S. Kuhanesan, "Usability Evaluation of Learning Management Systems in Sri Lankan Universities," *arXiv:1412.0197 [cs]*, Feb. 2015, [Online]. Available: <http://arxiv.org/abs/1412.0197>.
- [36] E. Sung and R. E. Mayer, "Affective impact of navigational and signaling aids to e-learning," *Computers in Human Behavior*, vol. 28, no. 2, pp. 473–483, Mar. 2012, <https://doi.org/10.1016/j.chb.2011.10.019>.
- [37] M. Bajec, "Educational Portals: A Way to Get an Integrated, User-Centric University Information System," in *Web Portals: The New Gateways to Internet Information and Services*, Melbourne, Australia: Idea Group, 2005, pp. 252–269.
- [38] A. M. Shaltoni, H. Khraim, A. Abuhamad, and M. Amer, "Exploring students' satisfaction with universities' portals in developing countries: A cultural perspective," *The International Journal of Information and Learning Technology*, vol. 32, no. 2, pp. 82–93, Jan. 2015, <https://doi.org/10.1108/IJILT-12-2012-0042>.
- [39] S. Petter, W. DeLone, and E. McLean, "Measuring information systems success: models, dimensions, measures, and interrelationships," *European Journal of Information Systems*, vol. 17, no. 3, pp. 236–263, Jun. 2008, <https://doi.org/10.1057/ejis.2008.15>.
- [40] A. H. R. Zaied, "An Integrated Success Model for Evaluating Information System in Public Sectors," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 3, no. 6, pp. 814–825, 2012.
- [41] S. Lee, "Academic Library E-Service User Motivation: A Korean Case Study," *Libri*, vol. 62, no. 3, pp. 233–247, Sep. 2012, <https://doi.org/10.1515/libri-2012-0019>.
- [42] X. Zhang and V. R. Prybutok, "A consumer perspective of E-service quality," *IEEE Transactions on Engineering Management*, vol. 52, no. 4, pp. 461–477, Nov. 2005, <https://doi.org/10.1109/TEM.2005.856568>.
- [43] S. Vairamuthu and S. M. Anuncia, "Reconnoitering Students' Satisfaction of an Online Based Assessment System to Improve Usability using PSO: An Examination into a Problem Solving and Programming Course," *Engineering, Technology & Applied Science Research*, vol. 6, no. 5, pp. 1207–1211, Oct. 2016, <https://doi.org/10.48084/etasr.800>.
- [44] K. P. Asiimwe, D. Machuve, and M. A. Dida, "Usability Testing of a Web Portal for Ornamental Plants and Flowers in Arusha, Tanzania," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5627–5631, Jun. 2020, <https://doi.org/10.48084/etasr.3412>.
- [45] C. Liu and K. P. Arnett, "Exploring the factors associated with Web site success in the context of electronic commerce," *Information & Management*, vol. 38, no. 1, pp. 23–33, Oct. 2000, [https://doi.org/10.1016/S0378-7206\(00\)00049-5](https://doi.org/10.1016/S0378-7206(00)00049-5).
- [46] L. Cobus, V. F. Dent, and A. Ondrusek, "How Twenty-Eight Users Helped Redesign an Academic Library Web Site: A Usability Study," *Reference & User Services Quarterly*, vol. 44, no. 3, pp. 232–246, 2005.

- [47] A. Aggarwal, *Web-Based Learning and Teaching Technologies: Opportunities and Challenges: Opportunities and Challenges*. London, UK: Idea Group, 1999.
- [48] J. M. Pearson and A. M. Pearson, "An Exploratory Study into Determining the Relative Importance of Key Criteria in Web Usability: A Multi-Criteria Approach," *Journal of Computer Information Systems*, vol. 48, no. 4, pp. 115–127, Jun. 2008, <https://doi.org/10.1080/08874417.2008.11646041>.
- [49] P. N. Sindhuja and S. G. Dastidar, "Impact of the Factors Influencing Website Usability on User Satisfaction," *IUP Journal of Management Research*, vol. 8, no. 12, pp. 54–66, Dec. 2009.
- [50] D. U. Bolliger, "Key Factors for Determining Student Satisfaction in Online Courses," *International Journal on E-Learning*, vol. 3, no. 1, pp. 61–67, 2004.
- [51] R. Aziz and B. Shahzad, "Factors for Measurement of ITES Quality for Higher Education Institutions in Saudi Arabia," *Global Journal of Computer Science and Technology*, vol. 15, no. 3, May 2015.
- [52] A. Bandura, *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1986.
- [53] A. Bandura, *Self-efficacy: the exercise of control*. New York, NY, USA: W.H. Freeman and Company, 1997.
- [54] W. Liu *et al.*, "Internet-Enabled User Interfaces for Distance Learning," *International Journal of Technology and Human Interaction*, vol. 5, no. 1, pp. 51–77, Jan. 2009, <https://doi.org/10.4018/jthi.2009010105>.
- [55] P. R. Warshaw and F. D. Davis, "Disentangling behavioral intention and behavioral expectation," *Journal of Experimental Social Psychology*, vol. 21, no. 3, pp. 213–228, May 1985, [https://doi.org/10.1016/0022-1031\(85\)90017-4](https://doi.org/10.1016/0022-1031(85)90017-4).
- [56] M. A. Mahmood, L. Hall, and D. L. Swanberg, "Factors Affecting Information Technology Usage: A Meta-Analysis of the Empirical Literature," *Journal of Organizational Computing and Electronic Commerce*, vol. 11, no. 2, pp. 107–130, Jun. 2001, https://doi.org/10.1207/S15327744JOCE1102_02.
- [57] 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>.
- [58] M. Turner, B. Kitchenham, P. Brereton, S. Charters, and D. Budgen, "Does the technology acceptance model predict actual use? A systematic literature review," *Information and Software Technology*, vol. 52, no. 5, pp. 463–479, May 2010, <https://doi.org/10.1016/j.infsof.2009.11.005>.
- [59] Y. Haliso, "Factors Affecting Information and Communication Technologies (ICTs) Use by Academic Librarians in Southwestern Nigeria," *Library Philosophy and Practice (e-journal)*, Apr. 2011, [Online]. Available: <https://digitalcommons.unl.edu/libphilprac/571>.
- [60] F. Alturise and B. Alojaiman, "Benefits and challenges of using ICT in Saudi Arabia universities: A literature review," in *International Conference on Advanced in Computing, Engineering and Learning Technologies*, 2013, vol. 2.
- [61] M. Verbree, E. Horlings, P. Groenewegen, I. Van der Weijden, and P. van den Besselaar, "Organizational factors influencing scholarly performance: a multivariate study of biomedical research groups," *Scientometrics*, vol. 102, no. 1, pp. 25–49, Jan. 2015, <https://doi.org/10.1007/s11192-014-1437-x>.
- [62] U. Sekaran and R. Bougie, *Research Methods For Business: A Skill Building Approach*, 7th edition. Chichester, WS, UK: Wiley, 2016.
- [63] R. G. Netemeyer, W. O. Bearden, and S. Sharma, *Scaling Procedures: Issues and Applications*, 1st edition. Thousand Oaks, CA, USA: SAGE, 2003.
- [64] D. F. Polit and C. T. Beck, "The content validity index: Are you sure you know what's being reported? critique and recommendations," *Research in Nursing & Health*, vol. 29, no. 5, pp. 489–497, 2006, <https://doi.org/10.1002/nur.20147>.
- [65] M.-C. Boudreau, D. Gefen, and D. W. Straub, "Validation in Information Systems Research: A State-of-the-Art Assessment," *MIS Quarterly*, vol. 25, no. 1, pp. 1–16, 2001, <https://doi.org/10.2307/3250956>.
- [66] V. Zamanzadeh, A. Ghahramanian, M. Rassouli, A. Abbaszadeh, H. Alavi-Majd, and A.-R. Nikanfar, "Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication," *Journal of Caring Sciences*, vol. 4, no. 2, pp. 165–178, Jun. 2015, <https://doi.org/10.15171/jcs.2015.017>.
- [67] J. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling*, 2nd edition. Los Angeles, CA, USA: SAGE, 2016.
- [68] J. F. Hair Jr, M. Sarstedt, L. Hopkins, and V. G. Kuppelwieser, "Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research," *European Business Review*, vol. 26, no. 2, pp. 106–121, Jan. 2014, <https://doi.org/10.1108/EBR-10-2013-0128>.
- [69] S. B. MacKenzie, P. M. Podsakoff, and N. P. Podsakoff, "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," *MIS Quarterly*, vol. 35, no. 2, pp. 293–334, 2011, <https://doi.org/10.2307/23044045>.
- [70] M. Tenenhaus, V. E. Vinzi, Y.-M. Chatelin, and C. Lauro, "PLS path modeling," *Computational Statistics & Data Analysis*, vol. 48, no. 1, pp. 159–205, Jan. 2005, <https://doi.org/10.1016/j.csda.2004.03.005>.
- [71] J. Henseler and M. Sarstedt, "Goodness-of-fit indices for partial least squares path modeling," *Computational Statistics*, vol. 28, no. 2, pp. 565–580, Apr. 2013, <https://doi.org/10.1007/s00180-012-0317-1>.
- [72] M. Wetzels, G. Odekerken-Schroder, and C. van Oppen, "Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration," *MIS Quarterly*, vol. 33, no. 1, pp. 177–195, 2009, <https://doi.org/10.2307/20650284>.
- [73] M. Zviran, C. Glezer, and I. Avni, "User satisfaction from commercial web sites: The effect of design and use," *Information & Management*, vol. 43, no. 2, pp. 157–178, Mar. 2006, <https://doi.org/10.1016/j.im.2005.04.002>.
- [74] "Learner-Centered Psychological Principles: Guidelines for School Redesign and Reform," American Psychological Association, Washington, DC, USA, Jan. 1993.
- [75] S. M. Abdulhamid and I. Idris, "Design Evaluation of Some Nigerian University Portals: A Programmer's Point of View," *arXiv:1402.1241 [cs]*, Feb. 2014, [Online]. Available: <http://arxiv.org/abs/1402.1241>.
- [76] K. M. Alsharif, "Towards quality teacher education: productive pedagogies as a framework for Saudi pre-service teachers' training in Mathematics education," Ph.D. dissertation, Curtin University, Perth, Australia, 2011.
- [77] F. O. Oliha, "Web Portal Usability among Nigerian University Students: A Case Study of University Of Benin, Nigeria," *Nigerian Journal of Technology*, vol. 33, no. 2, pp. 199–206, Apr. 2014, <https://doi.org/10.4314/njt.v33i2.9>.
- [78] J. C. Roca, C.-M. Chiu, and F. J. Martinez, "Understanding e-learning continuance intention: An extension of the Technology Acceptance Model," *International Journal of Human-Computer Studies*, vol. 64, no. 8, pp. 683–696, Aug. 2006, <https://doi.org/10.1016/j.ijhcs.2006.01.003>.

An Effective Combination of Textures and Wavelet Features for Facial Expression Recognition

Syed Muhammad Hassan

Department of AI and Mathematical Sciences
SMI University
Karachi, Pakistan
m.hassan@smiu.edu.pk

Abdullah Alghamdi

College of Computer Science and Information Systems
Najran University
Najran, Saudi Arabia
abdulresearch@hotmail.com

Abdul Hafeez

Department of Software Engineering
SMI University
Karachi, Pakistan
ahkhan@smiu.edu.pk

Mohammad Hamdi

College of Computer Science and Information Systems
Najran University
Najran, Saudi Arabia
mahamdi@nu.edu.sa

Imtiaz Hussain

Department of AI and Mathematical Sciences
SMI University
Karachi, Pakistan
imtiaz@smiu.edu.pk

Mesfer Alrizq

College of Computer Science and Information Systems
Najran University
Najran, Saudi Arabia
msalrizq@nu.edu.sa

Abstract—In order to explore the accompanying examination goals for facial expression recognition, a proper combination of classification and adequate feature extraction is necessary. If inadequate features are used, even the best classifier could fail to achieve accurate recognition. In this paper, a new fusion technique for human facial expression recognition is used to accurately recognize human facial expressions. A combination of Discrete Wavelet Features (DWT), Local Binary Pattern (LBP), and Histogram of Gradients (HoG) feature extraction techniques was used to investigate six human emotions. K-Nearest Neighbors (KNN), Decision Tree (DT), Multi-Layer Perceptron (MLP), and Random Forest (RF) were chosen for classification. These algorithms were implemented and tested on the Static Facial Expression in Wild (SFEW) dataset which consists of facial expressions of high accuracy. The proposed algorithm exhibited 87% accuracy which is higher than the accuracy of the individual algorithms.

Keywords—ANN; FER; DWT; LBP; HOG; K-Nearest Neighbors

I. INTRODUCTION

Facial expressions are a way of sentiment expression and non-verbal correspondence. There are various systems that deal with human attitude and point of view recognition. Facial Expression Recognition (FER) transforms is one of the most discussed scientific areas nowadays. This issue is furthermore incredibly noteworthy in Human-Computer Collaboration (HCI) [1, 2]. FER is being utilized to provide a description for the mental state of human beings [3]. Meanwhile, modifications in the look of photos can occur by disturbances

in the pixels. Illumination troubles might also occur in indoor or outdoor photos. The exploration indicates those issues and proposes a combination strategy for different accessible highlights that surpasses these issues [4].

II. LITERATURE SURVEY AND THEORETICAL FRAMEWORK

It is very difficult to identify human facial regions. In order to handle this efficiently, a technique should be implemented to recognize facial indicators. One clause that is vital to know is the dynamic angle on transferring video [5]. Lower and top face method extends the spatial pyramid histogram of edges which give 3-dimensional facial acknowledgment. Fundamentally in this method, elements are researched for cheerful and pity indicators [6]. LBP and Improved Local Binary Pattern have been applied alongside Coordinate Bunching Representation [7]. Face recognition using an optimized algorithm chain for both 2D and 3D images gives an accuracy about 96% with SVM classifier using LBP and PCA. Further testing on 2D and 3D images using LBP and PCA with FFBNP (Feed Forward Back Propagation Neural Network) is less effective and efficient as compared to the SVM classifier [8]. Locality Preserving Projections (LPPs) have been used for manifold systems originated from Local Binary Pattern (LBP) subjects [9]. At first, a pyramid change is utilized to divide the test photographs into different areas. So, the goal pictures are isolated. After this, the ELBP is applied upon the little pictures to compute the ELBP pyramid and the community photo decided qualities are utilized to the little pictures from AWM which can ascertain the importance of the facts they got.

Corresponding author: Syed Muhammad Hassan

Finally, the AWELBPP highlight is assembled from the blend of the little ELBP pyramid and the AWM [10]. Support Vector Machine (SVM) has been applied in dispensing boisterous pictures for highlight extraction [11, 12].

Background Subtraction [13] showed good results by applying background subtraction on real-time feeds. In this work, a model based on Gaussian Mixture was used for unfolding the pixels of images and the variables of the pattern were calculated with the Expectation-Maximization (EM) algorithm. The shades were also spotted effectively. Background subtraction was also very effective and met the requirements of drowning detection. Authors in [14] reviewed earlier approaches and tried to cover up the issue of recognizing actions and behaviors and the problem of dealing with a moderate crowded situation with a good modeling technique. The conventional techniques where mixed and a Gaussian distribution was used to design the temporal changing of the background pixels in [15]. This has been proven to be insufficient for extremely non-stationary environment. However, the thresholding method with hysteresis dealt with the issue of choosing thresholds in the background subtraction context. Stationary cameras have also been used to find drowning persons in swimming pools [6, 17, 18]. In contrast to previous works based on geometrical and 3D Mahalanobis distance features, the presented method in [18] captured the temporal and spatial correlation of the swimmers along with color information using the Markov Random Field (MRF) context to give better performance. Promising outcomes for drowning detection were achieved using an exclusive functional link net which fused the descriptors of extracted swimmers optimally. An improved descriptor fusion technique associated with the hierarchical technique was proposed in [19]. The current drowning detection techniques can be broadly classified into the vision-based schemes and the systems based on wearable sensors [20-22]. On the other hand, the combination of aerial and underwater cameras to monitor the postures of FER was utilized in [23], whereas the CNN model achieved 99.78% accuracy [24]. An even more successful accuracy level was achieved in learning similarities and dissimilarities among the faces of dataset using FDREnet in [25].

III. SYSTEM METHODOLOGY

Usually, cameras are present in most areas for security purposes. The already installed cameras can be utilized for the purpose of monitoring and expression detection. So, few critical frames are extracted from the video or can be utilized. A facial expression video is divided into frames to be processed. The image frames extracted from the video are utilized for feature extraction. Then, classification is carried out.

A. Feature Extraction

The input dataset is very large to be handled and processed. It is supposed to be redundant (enough data, but not abundant information), so, the input dataset will be converted into a reduced depiction set containing features. This set is named as Features vector (Fv). This process is known as feature extraction. Therefore, taking out the prejudiced features from

the images enhances the decline of the dimension of the Fv by removing the redundancy in images and squeezing the relevant data into the Fv to a much smaller size.

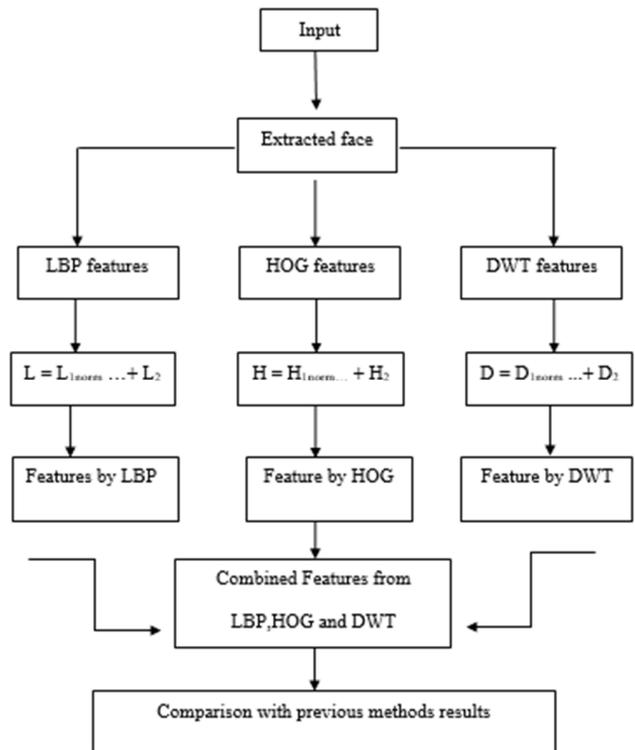


Fig. 1. The flowchart of the proposed method.

B. Feature Extraction via Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) is utilized to extract features from an image on various levels of low pass (g), and high pass filtering (h). A signal x is calculated by passing through these series of filters, at first through the low pass filter (g) and then through the high pass (h):

$$h[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k] \quad (1)$$

The low pass filter (g) and the high pass filter (h) with h cut-off frequency are described by:

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \quad (2)$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \quad (3)$$

The wavelet coefficients are the consecutive persistence of the estimate and the detail coefficients. The basic process of features extraction consists of:

- Mouldering the image using DWT in N-levels using decimation and filtering to get the detailed coefficients and approximation.
- Feature extraction using the DWT coefficients output.
- The features that were taken out from the DWT coefficients of the images are considered as the input to classifiers because of their operative representation.

The algorithmic steps for feature extraction from the dataset are:

- Step 1: The image data are decomposed into 4 detailed sub-bands by DWT.
- Step 2: The coefficients of approximation are further been decomposed by DWT to obtain localized data from the sub-band of the detailed coefficients of approximation (horizontal, vertical, and diagonal).
- Step 3: Aimed at processing and analyzing, all of the 4 levels detailed coefficients are calculated.
- Step 4: Finally, the features are analyzed and tabulated to be used as the input of the classifier.

C. Feature Extraction via Histogram of Gradients

The Histogram of Oriented Gradients (HOG) is the shape of the "function descriptor". The motive behind a feature descriptor is to generalize the item on a way that the same item (in this case a person) produces the same feature descriptor at the same time as considered under specific situations. This makes the class assignment simpler. Static Facial Expressions in the Wild (SFEW) has been utilized for selecting frames from AFEW. Regarding the block normalization for HOG, we consider \mathbf{v} as the non-normalized vector containing all histograms in a given block, $\|\mathbf{v}\|_k$ be its k -norm for $k=1, 2$, and ϵ is some small constant. The normalized factor is defined as:

$$f = \frac{\mathbf{v}}{\sqrt{\|\mathbf{v}\|_2^2 + \epsilon^2}} \quad (4)$$

The dataset covers unconstrained facial expressions, numerous head poses, massive age variety, occlusions, numerous poses, and near actual global illumination. Frames had been extracted from AFEW sequences and were labeled based on the label of the series. Typically, SFEW includes seven-hundred snapshots which have been classified for six fundamental expressions: anger, disgust, fear, happiness, sadness, and surprise, and were categorized by unbiased labelers.

D. Feature Extraction via Local Binary Pattern (LBP)

The LBP method is applied on facial images in order to extract features that may be used to get a degree of similarity. Firstly, the pictures have been divided into several blocks. After that, the LBP histogram was calculated for each block. The value of the LBP code of a pixel (x_c, y_c) is considered as:

$$LBP_{(p,r)} = \sum_{p=0}^{p-1} s(g_p, g_c) 2^p \quad (5)$$

where $s(x) = \{1, \forall x \geq 0\}$ and $s(x) = \{0, \text{otherwise}\}$. The notation $LBP_{(p,r)u2}$ is used for the LBP operator, where (p, r) represents the neighborhood, and $u2$ stands for uniform patterns and labeling all remaining patterns with a single label. The histogram for the image $f_1(x, y)$ is defined as:

$$H_i = \sum_{(x,y)} I\{f_1(x, y) = i\}, i = 0, \dots, n-1, \quad (6)$$

The number of different labels produced by the LBP operator, and $I\{A\}$ is 1, if A is true and 0 if it is false. Further, the image patches whose histograms are to be compared must be normalized in order to get a coherent description:

$$N_i = \frac{H_i}{\sum_{j=0}^{n-1} H_j} \quad (7)$$

Then, the block LBP histograms were concatenated into an unmarried vector. The histograms have then been evaluated by using space similarity [16]. Moreover, each bin in histograms consists of the variety of its look within the region. Lastly, the feature vector is constructed with the useful data by concatenating the community histograms to one massive histogram.

IV. RESULTS AND DISCUSSION

In this study, the SFEW dataset was used for testing, which is close to real world environment, having 300 color images with 6 emotion categories, consisting of 50 pictures each with dimensions of 143×181 pixels. The classes are Surprise, Fear, Anger, Sadness, Disgust, and Happiness represented by SU, F, A, S, D, and H respectively. The results were evaluated with assessment metrics, including confusion matrix, precision, recall, and F1 score. To compute the overall precision, we used micro-averages to combine the consequences across the 6 categories. We divided our dataset into 80% training and 20% testing subsets. The sets were fed to the distinctive learning system which utilized algorithms such as K-Nearest Neighbor (KNN), Decision Tree (DT), Multilayer Perceptron (MLP), and Random Forest (RF). Our experimental model was divided into four parts. The mentioned machine learning algorithms were applied directly to the first part of the dataset. Table I shows the original dataset accuracies.

TABLE I. ORIGINAL IMAGES WITHOUT FEATURE EXTRACTION

Algorithm	KNN	DT	MLP	RF
Accuracy	27%	14%	22%	32%

TABLE II. LBP FEATURE EXTRACTION

Algorithm	KNN	DT	MLP	RF
Accuracy	50%	96%	22%	95%

Maximum accuracy was achieved by the RF and was only 32%. Then, all algorithm accuracies were computed using DWT, LPB, and HOG for the other parts of the dataset (Tables II-IV), and finally combined them and achieved 87% maximum accuracy with MLP and 29% minimum accuracy with KNN (Table V), which are respectively shown in the confusion matrices of Figures 2 and 3.

TABLE III. DWT FEATURE EXTRACTION

Algorithm	KNN	DT	MLP	RF
Accuracy	12%	17.5%	36%	21%

TABLE IV. HOG FEATURE EXTRACTION

Algorithm	KNN	DT	MLP	RF
Accuracy	12%	12%	37%	14%

TABLE V. COMBINATION OF LBP, DWT, AND HOG

Algorithm	KNN	DT	MLP	RF
Accuracy	29%	80%	87%	79%

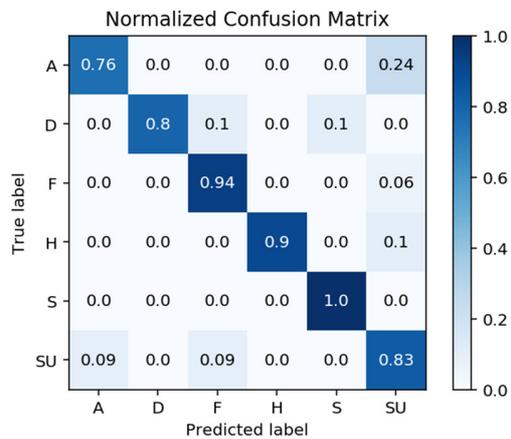


Fig. 2. MLP confusion matrix.

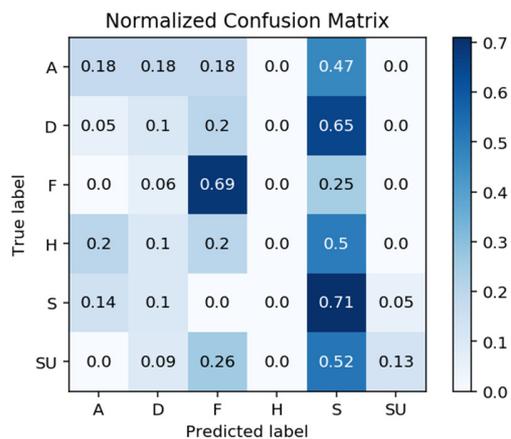


Fig. 3. KNN confusion matrix.

Further, we also calculated some edges of the face generated by DWT, LBP, and HOG. The original image is reconstructed using Harr DWT techniques.

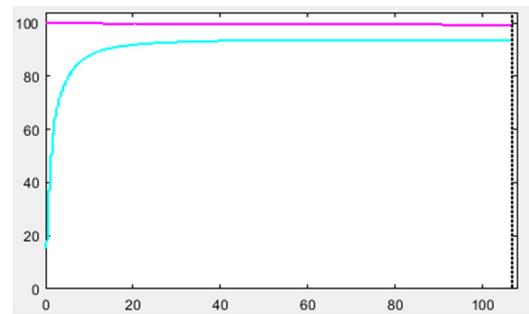


Fig. 4. Retained energy is 99.40%.

Histogram of Oriented Gradients

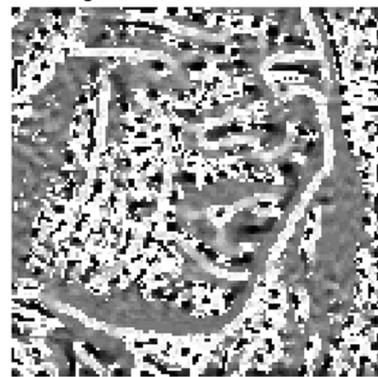


Fig. 5. HOG edges.



Fig. 6. LBP visualized surprised face.

V. CONCLUSION AND FUTURE WORK

The proposed model combines DWT, HOG, and LBP capabilities in a feature extraction technique with system learning algorithms in an excellent way of enhancing the accuracy of facial feature recognition. Six facial expressions from the SFEW database had been used for training and validation. The results indicated that the accuracy of the use of blended methods is 87%, which is higher from the individual accuracies of the combined algorithms. However, the proposed combination has the issue of generalization which may be addressed in our future work.

FER is one of the most well-known regions in image processing. Generally, FER has been given more attention nowadays. The proposed technique gives an exquisite overview of facial recognition methods. The extraction of functions is vital as it decreases the very massive amount of data to only a required set. Thus, it reduces the processing time of the machine and the results are more correct. In future work, the accuracy may be augmented by using more learning algorithms. A similar approach to the usage of the Convolution Natural Community can be combined with the prevailing support vector classifier.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Abhinav Dhall, Australian National University for the provision of the SFEW dataset.

REFERENCES

- [1] W.-L. Chao, J.-J. Ding, and J.-Z. Liu, "Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection," *Signal Processing*, vol. 117, pp. 1–10, Dec. 2015, <https://doi.org/10.1016/j.sigpro.2015.04.007>.
- [2] F. Long and M. S. Bartlett, "Video-based facial expression recognition using learned spatiotemporal pyramid sparse coding features," *Neurocomputing*, vol. 173, pp. 2049–2054, Jan. 2016, <https://doi.org/10.1016/j.neucom.2015.09.049>.
- [3] H. Fang *et al.*, "Facial expression recognition in dynamic sequences: An integrated approach," *Pattern Recognition*, vol. 47, no. 3, pp. 1271–1281, Mar. 2014, <https://doi.org/10.1016/j.patcog.2013.09.023>.
- [4] J. Hussain Shah, M. Sharif, M. Raza, M. Murtaza, and S. Ur-Rehman, "Robust Face Recognition Technique under Varying Illumination," *Journal of applied research and technology*, vol. 13, no. 1, pp. 97–105, 2015.
- [5] S. Arya, N. Pratap, and K. Bhatia, "Future of Face Recognition: A Review," *Procedia Computer Science*, vol. 58, pp. 578–585, Jan. 2015, <https://doi.org/10.1016/j.procs.2015.08.076>.
- [6] M.-Y. Chen and C.-C. Chen, "The contribution of the upper and lower face in happy and sad facial expression classification," *Vision Research*, vol. 50, no. 18, pp. 1814–1823, Aug. 2010, <https://doi.org/10.1016/j.visres.2010.06.002>.
- [7] A. Fernandez, O. Ghita, E. Gonzalez, F. Bianconi, and P. F. Whelan, "Evaluation of robustness against rotation of LBP, CCR and ILBP features in granite texture classification," *Machine Vision and Applications*, vol. 22, no. 6, pp. 913–926, Nov. 2011, <https://doi.org/10.1007/s00138-010-0253-4>.
- [8] S. Shankar and V. R. Udipi, "Recognition of Faces – An Optimized Algorithmic Chain," *Procedia Computer Science*, vol. 89, pp. 597–606, Jan. 2016, <https://doi.org/10.1016/j.procs.2016.06.020>.
- [9] R. K. Nagar, R. Manazhy, and P. Sankaran, "Sparse Manifold Discriminant Embedding for Face Recognition," *Procedia Computer Science*, vol. 89, pp. 743–748, Jan. 2016, <https://doi.org/10.1016/j.procs.2016.06.050>.
- [10] T. Gao, X. L. Feng, H. Lu, and J. H. Zhai, "A novel face feature descriptor using adaptively weighted extended LBP pyramid," *Optik*, vol. 124, no. 23, pp. 6286–6291, Dec. 2013, <https://doi.org/10.1016/j.ijleo.2013.05.007>.
- [11] K. Yu, Z. Wang, L. Zhuo, J. Wang, Z. Chi, and D. Feng, "Learning realistic facial expressions from web images," *Pattern Recognition*, vol. 46, no. 8, pp. 2144–2155, Aug. 2013, <https://doi.org/10.1016/j.patcog.2013.01.032>.
- [12] R. A. Khan, A. Meyer, H. Konik, and S. Bouakaz, "Framework for reliable, real-time facial expression recognition for low resolution images," *Pattern Recognition Letters*, vol. 34, no. 10, pp. 1159–1168, Jul. 2013, <https://doi.org/10.1016/j.patrec.2013.03.022>.
- [13] S. Ali Khan, A. Hussain, and M. Usman, "Facial expression recognition on real world face images using intelligent techniques: A survey," *Optik*, vol. 127, no. 15, pp. 6195–6203, Aug. 2016, <https://doi.org/10.1016/j.ijleo.2016.04.015>.
- [14] S. Ali Khan, A. Hussain, A. Basit, and S. Akram, "Kruskal-Wallis-Based Computationally Efficient Feature Selection for Face Recognition," *The Scientific World Journal*, vol. 2014, May 2014, Art. no. e672630, <https://doi.org/10.1155/2014/672630>.
- [15] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, May 2009, <https://doi.org/10.1016/j.imavis.2008.08.005>.
- [16] W.-H. Chen, P.-C. Cho, P.-L. Fan, and Y.-W. Yang, "A framework for vision-based swimmer tracking," in *International Conference on Uncertainty Reasoning and Knowledge Engineering*, Bali, Indonesia, Aug. 2011, vol. 1, pp. 44–47, <https://doi.org/10.1109/URKE.2011.6007835>.
- [17] D. Zecha, T. Greif, and R. Lienhart, "Swimmer detection and pose estimation for continuous stroke-rate determination," in *Multimedia on Mobile Devices 2012; and Multimedia Content Access: Algorithms and Systems VI*, California, United States, Jan. 2012, vol. 8304, Art. no. 830410, <https://doi.org/10.1117/12.908309>.
- [18] S. S. Intille, J. W. Davis, and A. F. Bobick, "Real-time closed-world tracking," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, USA, Jun. 1997, pp. 697–703, <https://doi.org/10.1109/CVPR.1997.609402>.
- [19] K.-A. Toh, W.-Y. Yau, and X. Jiang, "A reduced multivariate polynomial model for multimodal biometrics and classifiers fusion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 224–233, Feb. 2004, <https://doi.org/10.1109/TCSVT.2003.821974>.
- [20] M. Kharat, Y. Wakuda, N. Koshizuka, and K. Sakamura, "Automatic waist airbag drowning prevention system based on underwater time-lapse and motion information measured by smartphone's pressure sensor and accelerometer," in *IEEE International Conference on Consumer Electronics*, Las Vegas, NE, USA, Jan. 2013, pp. 270–273, <https://doi.org/10.1109/ICCE.2013.6486891>.
- [21] M. Kharat, Y. Wakuda, S. Kobayashi, N. Koshizuka, and K. Sakamura, "Near drowning detection system based on swimmer's physiological information analysis," presented at the World Conference on Drowning Prevention (WCDP), May 2011.
- [22] E. McAdams *et al.*, "Wearable sensor systems: The challenges," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, MA, USA, Sep. 2011, pp. 3648–3651, <https://doi.org/10.1109/IEMBS.2011.6090614>.
- [23] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, "Support Vector Clustering," *Journal of Machine Learning Research*, vol. 2, pp. 125–137, 2001.
- [24] Y. Said, M. Barr, and H. E. Ahmed, "Design of a Face Recognition System based on Convolutional Neural Network (CNN)," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5608–5612, Jun. 2020, <https://doi.org/10.48084/etasr.3490>.
- [25] D. Virmani, P. Girdhar, P. Jain, and P. Bamdev, "FDREnet: Face Detection and Recognition Pipeline," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3933–3938, Apr. 2019, <https://doi.org/10.48084/etasr.2492>.

Recommendations for the Implementation of ISO 9001:2015 in the Manufacturing Industry of Pakistan

Zohaib Khan

Faculty of Mechanical and
Manufacturing Engineering
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Johor, Malaysia

Yusri Bin Yusof

Faculty of Mechanical and
Manufacturing Engineering
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Johor, Malaysia

Nor Haslinda Binti Abass

Faculty of Civil Engineering and Built
Environment
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Johor, Malaysia

Maznah Binti Illyas Ahmed

Faculty of Mechanical and Manufacturing Engineering
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Johor, Malaysia

Qadir Bakhsh Jamali

Department of Mechanical Engineering
Quaid-e-Awam University of Engineering Science & Technology
Nawabshah Sindh, Pakistan

Abstract-ISO 9001:2015 is known as the highest quality management standard, enhancing customer and company satisfaction. While the manufacturing industry is playing a vital role in Pakistan's GDP, the implementation of ISO 9001:2015 is not widespread. This study aims to examine the factors that hinder the implementation of ISO 9001:2015 in Pakistan's manufacturing industry and recommend suggestions to overcome them. An in-depth literature review identified 32 general factors that hinder the implementation of ISO 9001:2015. A structured questionnaire was designed and distributed to 135 respondents to examine the factors hindering the implementation of ISO 9001:2015 in Pakistan's manufacturing industry. The gathered data were analyzed using the average index formula, and 8 significant factors were identified. These factors were used in the design of a semi-structured questionnaire distributed among high qualified and experienced respondents working in top managerial posts of manufacturing industries in Pakistan. The collected data were analyzed using content analysis. The findings can be helpful in the determination of recommendations regarding the implementation of ISO 9001:2015 in the manufacturing industries of Pakistan.

Keywords-recommendations; manufacturing industry; ISO 9001:2015; significant factors

I. INTRODUCTION

ISO 9001 is based on seven quality administration standards. These standards and codes are elementary convictions, imperatives, norms, and values that form a foundation that ultimately leads firms, companies, and organizations to performance enhancement. Before, on ISO 9001:2008, the principles were based on 8 quality administration standards, and later they were decreased to the seven quality executive code standards [1, 2]. The previous ISO 9001 edition referred to "products" which encompassed every kind of product, while the definition "purchasing" was altered to "control of externally provided processes, products, and services". So, the organizations that subcontract these functions

or processes can ensure that their external suppliers comply with the definitions of the administration system [3]. The fundamental model could be interpreted as an organization or a company incorporating the new processes or implementing the new ISO standard. The PDCA cycle proposes a test preparation process that guides companies or organizations during changes or problem-solving procedures. The PDCA cycle includes test data and feedback [4]. The major parts of this sequence involve numerical procedure control strategies in the production process, simplifying quality enhancement. The PDCA cycle was applied by Japanese industrial engineers with enormous success. For instance, the Toyota Production System (TPS) is partially associated with PDCA scientific schemes and quality control [5]. In underdeveloped countries, fewer organizations are certified by ISO 9001 in comparison to the developed countries. As North American and European countries demand trustworthy infrastructure for the ISO 9001 list and registration, uncompliant organizations cannot access these markets [6]. In many countries, the production sector has an important and vibrant position [7], while unsteadiness in the production division may cause turbulence on the whole economy. Several studies investigated the business sector in numerous, developing and developed, countries. Previous studies reported that although there is significant literature on the ISO 9001, only a few publications highlighted the problems, the issues, and the challenges of the ISO 9001, especially in the developing countries [8, 9]. Besides, progress in the implementation of ISO 9001 has been noticed in Middle East countries among numerous groups of companies in the United Arab Emirates, Kuwait, Lebanon, Qatar, and Oman, while Pakistani companies are trying to incorporate the codes and the standards to obtain the official certification.

Experimental research was carried out in Egypt, by selecting a sample of industrial organizations to explore the critical aspects of the application process of ISO 9001 [10]. The study highlighted the problems arising during the application of ISO 9001 principles, the administration body's actions to apply

them successfully, and reported the necessity to alter the organizational system to conform. Hence, the employees' confrontation during the implementation of the standards was the most momentous issue faced. Furthermore, 11 probable elements that influence the application of ISO 9001 were recommended. The responders were requested to rate the intensity to which every dimension influenced them, and the scale ranged from "very helpful" to "not helpful at all". Thus, by investigating the outcomes, it was shown that the main advantageous elements in the successful application of ISO 9001:2015 were the commitment from apex supervision, the devotion of the company's in-house auditor, the well-designed system of the procedures, and the support from the mother organization or the partner.

A study conducted on ISO 9001 certified goods and manufacturing companies in Pakistan reported 8 prominent difficulties during the procedure [11]. Another study reported the following significant aspects that assist the ISO 9001 implementation in Pakistan: a well-organized process system, the commitment of higher management to the goal, and the useful views from domestic auditors [12]. Moreover, the major problems connected to the standards' application were the requirements to change the prevailing systems, the uncertainty reported by the employees concerning the alteration, and the lack of understanding and respect to the ISO principles illustrated by all departments. Furthermore, several problems and issues concerning the function of the quality structure factors in the Pakistani manufacturing field were reported, such as the inspiration for the registration, the cost of the official ISO certification and the training plans, the quality guidelines, the registration companies, the ISO advantages, and the formal auditing.

II. RESEARCH METHODOLOGY

The main objective of this study was to identify the significant factors which hinder the implementation of ISO 9001:2015 in the manufacturing industry of Pakistan and determine recommendations to overcome them. An in-depth literature review was conducted which identified 32 common factors that hinder the implementation of ISO 9001:2015 in the manufacturing industry worldwide. A structured questionnaire was designed and developed based on the identified common factors shown in Table I. The questionnaire was distributed among 130 experienced respondents who were working in the manufacturing industry of Pakistan. The respondent's experience profile is shown in Figure 1.

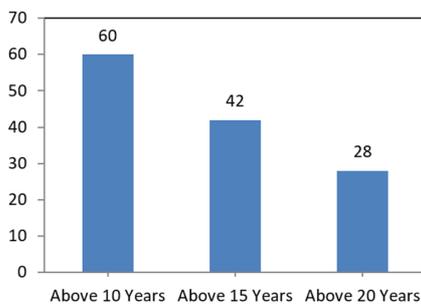


Fig. 1. Respondents' experience.

TABLE I. FACTORS HINDERING ISO 9001 IMPLEMENTATION

	Factors hindering ISO implementation	1	2	3	4	5
1	Absence of ISO 9001 certification					
2	Unwillingness to change organizational culture					
3	Inflating the size of documents					
4	Unwillingness to change work system					
5	Weak Interdepartmental relationship					
6	Management involvement					
7	Insufficient knowledge in quality					
8	Insufficient employee training					
9	Training to top management by industries					
10	Insufficient knowledge in quality					
11	Insufficient Human Resources					
12	Shortage of financial resources					
13	ISO requirements as unrealistic					
14	Absence of consulting boards					
15	Difficult to implement ISO					
16	Employee resistance					
17	Top management involvement					
18	Leadership from top management of company					
19	Accountability in industry					
20	Action orientation by industry					
21	Credibility of the industry					
22	Collaboration among Industries					
23	Technology resources in the company					
24	Big quality standards					
25	Complexity in industries					
26	Rules & Regulations					
27	Training of staff					
28	Previous record of Quality					
29	Inspection and Testing					
30	Unavailability of Quality Manuals					
31	Employees behavior					
32	Corrective and preventive action					

The gathered data were statistically analyzed using the average index formula. Factors having an average score of more than 3.6 were marked as significant factors that hinder the implementation of ISO 9001:2015 in the manufacturing industry of Pakistan [13]. These factors are shown in Table II.

TABLE II. LIST OF SIGNIFICANT FACTORS

No	Significant Factors	AI Score	Ranking
01	Absence of ISO 9001 Certification	4.874	01
02	Unavailability of Quality Manuals	4.764	02
03	Training of Staff	4.697	03
04	Top Management involvement	4.518	04
05	Employees behavior	4.473	05
06	Management involvement	4.197	06
07	Rules and Regulations	3.954	07
08	Difficult for implementation of ISO	3.879	08

After identifying the significant factors, a semi-structured questionnaire was developed and distributed to 50 top managerial respondents to investigate recommendations for each significant factor hindering the implementation of ISO

9001:2015 in Pakistan's manufacturing industry. The academic qualification of the selected respondents is shown in Figure 2. Gathered data from the semi-structured questionnaire were analyzed using content analysis. Content analysis is a research method involving a thorough inspection of human conversation and a systematic, objective, quantitative analysis of human characteristics. After revising the collected data of every aspect that may be relevant, the coding frame was developed [14].

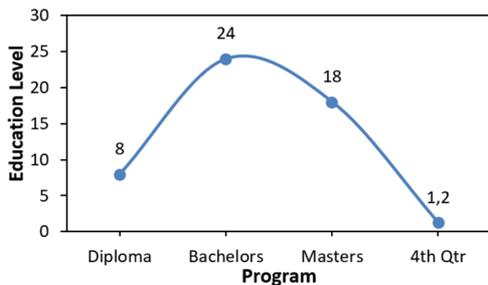


Fig. 2. Respondents' academic qualification.

III. DATA ANALYSIS AND RESULTS

Cronbach's alpha measures the internal consistency or reliability of survey data, and it was utilized to evaluate the reliability of the data set using SPSS v27. As Cronbach's alpha was measured at 0.914, it was reliable to proceed further [15]. Eight significant factors were found to hinder the implementation of ISO 9001:2015 in the manufacturing industry of Pakistan. Content analysis revealed some recommendations and suggestions to counteract those hindering factors, as shown in Table III.

TABLE III. RECOMMENDATIONS FOR SIGNIFICANT FACTORS

Factor	Recommendations	Frequency
1	1. ISO certification should be clear	47
	2. Process of ISO certification should be explained	45
	3. Owner should be informed of ISO certification	42
2	1. Quality manuals should be available easily	48
	2. Quality manuals should be explained	45
	3. Quality manuals should be revised	41
3	1. Training of staff	48
	2. Staff should be informed of ISO 9001 benefits	45
	3. Training should be done regularly	41
4	1. Top management should be involved in ISO 9001 certification	48
	2. Brief should be given for advantages of ISO 9001	46
	3. Top management should be given examples of ISO 9001 certified industries	43
5	1. Employees should be informed of ISO 9001 advantages	49
	2. Employees should follow the standards of ISO 9001	46
	3. Employees should be trained on ISO-9001	44
6	1. ISO certification training should be given to management	48
	2. Management should be informed of ISO 9001 advantages	45
	3. Quality manuals should be revised	41
7	1. Rules and regulations should be informed to all staff	48
	2. Policies of ISO 9001 should be easily understood	45
	3. Training should be done regularly	41
8	1. Clear concept of ISO 9001 should be explained	48
	2. Implementation procedure of ISO 9001 should be made easy	46
	3. Top management should be informed on the advantages of ISO-9001	43

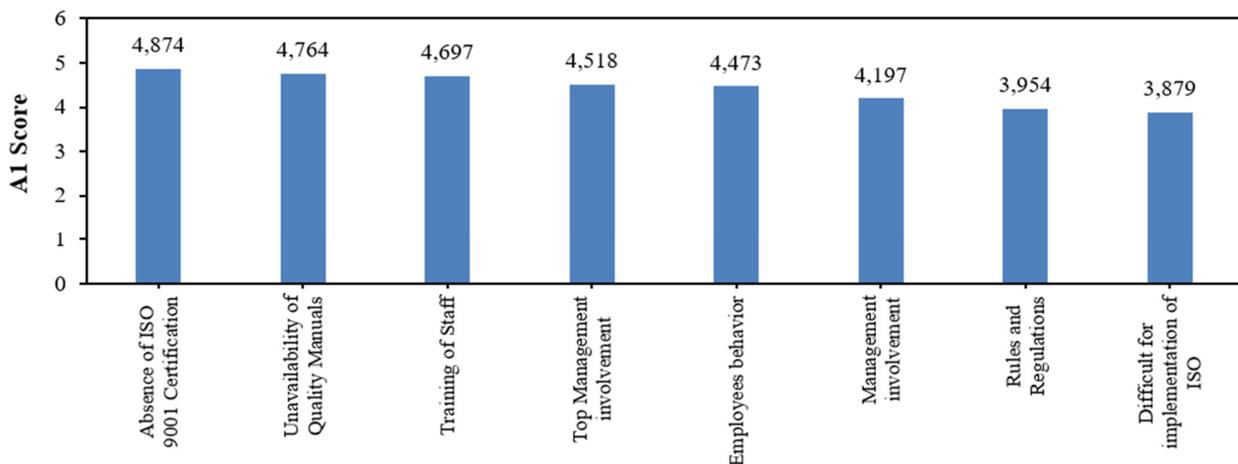


Fig. 3. Significant factors hindering the implementation of ISO (with AI score).

IV. CONCLUSION

ISO certification plays an important role in any country's GDP development. The implementation of ISO 9001:2015 in Pakistan's manufacturing industry is a serious concern. Eight

significant factors were found to hinder the implementation of ISO 9001:2015 in Pakistan's manufacturing industry. Recommendations and suggestions were pointed out by the qualified and experienced professionals of Pakistan's manufacturing industry. The proposed suggestions for the

implementation of ISO 9001-2015 include training to employees and administrative staff, and informing the top management of the advantages of ISO implementation. This study will be helpful for the implementation of ISO 9001-2015 in the manufacturing industry of Pakistan.

REFERENCES

- [1] M. M. Chhikara, D. M. S. Narwal, and M. P. Dahiya, "Implementation of Six Sigma in Indian Manufacturing Industries," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 3, no. 1, pp. 22–30, Feb. 2017.
- [2] D. Hoyle, *ISO 9000 Quality Systems Handbook - Updated for the ISO 9001:2008 Standard*, 6th ed. New York, NY, USA: Routledge, 2009.
- [3] S. I. S. A.- Hawary and F. M. F. Aldaihani, "Customer Relationship Management and Innovation Capabilities of Kuwait Airways," *International Journal of Academic Research in Economics and Management Sciences*, vol. 5, no. 4, pp. 201–226, Dec. 2016, <https://doi.org/10.6007/IJAREMS/v5-i4/2463>.
- [4] J. R. Evans and W. M. Lindsay, *An Introduction to Six Sigma and Process Improvement*, 2nd edition. Stamford, CT, USA: Cengage Learning, 2014.
- [5] A. Jönsson and M. Berglund, "ISO 9001:2015 implementation at a manufacturing company," M.S. thesis, Halmstad University, Halmstad, Sweden, 2016.
- [6] H. Latan, C. J. Chiappetta Jabbour, A. B. Lopes de Sousa Jabbour, P. de Camargo Fiorini, and C. Foropon, "Innovative efforts of ISO 9001-certified manufacturing firms: Evidence of links between determinants of innovation, continuous innovation and firm performance," *International Journal of Production Economics*, vol. 223, May 2020, Art. no. 107526, <https://doi.org/10.1016/j.ijpe.2019.107526>.
- [7] N. S. Siddharthan and K. Narayanan, Eds., *Innovation and Global Competitiveness: Case of India's Manufacturing Sector*, 1st ed. New York, NY, USA: Routledge, 2018.
- [8] A. S. A. Al-mijrab, M. E. Elgharib, and M. A. Al-Griw, "Critical Success Factors of ISO/IEC 17025 Implementation within Arabic Countries: A Case Study of Libyan Research Centres and Laboratories (LRCL)," presented at the 23rd International Conference on ISO & TQM 23-ICIT, Zhuhai, China, May 2019.
- [9] A. Kakouris and E. Sfakianaki, "Motives for implementing ISO 9000 – does enterprise size matter?," *International Journal of Productivity and Performance Management*, vol. 68, no. 2, pp. 447–463, Jan. 2019, <https://doi.org/10.1108/IJPPM-03-2018-0096>.
- [10] H. Magd, "Quality Management Standards (QMS) Implementation in Egypt: ISO 9000 Perspectives," *Global Business and Management Research; An International Journal*, vol. 2, no. 1, pp. 57–68, 2010.
- [11] M. Shafiq, "Implementation of Quality Management Systems and Business Excellence Frameworks in Pakistani Textile Companies," *Journal of Quality and Technology Management*, vol. 8, no. 2, pp. 11–23, Dec. 2012.
- [12] W. Ahmed, "ISO 9001 Transition and its Impact on the Organizational Performance: Evidence from Service Industries of Pakistan," *International Journal of Research in Business Studies and Management*, vol. 4, no. 3, 2017, <https://doi.org/10.22259/ijrbsm.0403004>.
- [13] S. Sohu, A. H. Abdullah, S. Nagapan, T. A. Rind, and A. A. Jhatial, "Controlling Measures for Cost Overrun Causes in Highway Projects of Sindh Province," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4276–4280, Jun. 2019, <https://doi.org/10.48084/etasr.2749>.
- [14] G. Durak *et al.*, "Content Analysis of Master Theses and Dissertations Based on Action Research," *Journal of Education and Training Studies*, vol. 4, no. 12, pp. 71–80, Nov. 2016, <https://doi.org/10.11114/jets.v4i12.1906>.
- [15] S. Sohu, A. H. Abdullah, S. Nagapan, A. A. Jhatial, K. Ullah, and I. A. Bhatti, "Significant Mitigation Measures for Critical Factors of Cost Overrun in Highway Projects of Pakistan," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2770–2774, Apr. 2018, <https://doi.org/10.48084/etasr.1916>.

Convective Heat Transfer During Melting in a Solar LHTES

Fatima Zohra Mecieb

Laboratory of Reactive Systems and Materials
University Djillali liabes of Sidi Bel Abbas
Sidi Bel Abbas, Algeria
fatima.mecieb@univ-sba.dz

Francisco García Bermejo

Thermal and Fluid Engineering Department
Polytechnic University of Cartagena
Cartagena, Spain
pacoaitpaco@gmail.com

Juan Pedro Solano Fernández

Thermal and Fluid Engineering Department
Polytechnic University of Cartagena
Cartagena, Spain
JuanP.Solano@upct.es

Samir Laouedj

Laboratory of Reactive Systems and Materials
University Djillali liabes of Sidi Bel Abbas
Sidi Bel Abbas, Algeria
s.laouedj@gmail.com

Abstract-Melting combined with natural convection in a shell and Latent Heat Thermal Energy Storage (LHTES) tube driven by a solar collector was analyzed numerically in the present work. This work's particularity lies in the fact that the HTF temperature varies at each moment following the solar irradiance curve. A program (UDF) has been developed and integrated into Ansys to meet this requirement. The use of this coupling strategy allows obtaining realistic unsteady LHTES results. Several numerical investigations were carried out to analyze the effect of the heat sources' power on the accumulator's performance. The obtained results show that natural convection considerably influences the heat transfer as well as the melting kinetics of the Phase Change Material (PCM). Besides, the results show that increasing the heat transfer fluid's thermal load can increase the melting rate of the PCM and the stored energy and reduce the entire melting time.

Keywords-natural convection; solar coupling; PCM melting

I. INTRODUCTION

Among the wide variety of renewable energy sources, solar energy is the largest one and simultaneously an example of a time-variable energy source that is mismatched with the needs of use. Energy storage is the most appropriate way to overcome these inadequacies. Although there are several storage systems, latent heat storage using Phase Change Material (PCM) has clear advantages over others: a greater density of energy storage and a relatively constant storage temperature. However, these materials present many problems because of their low thermal conductivity. Improving heat transfer by developing a useful heat exchanger device in the Latent Heat Thermal Energy Storage (LHTES) unit is a solution to the PCM's low thermal conductivity problem. This calls for a powerful knowledge of the heat transfer in the LHTES process. The absence of comprehension of heat transfer modes during the phase change is an issue frequently mentioned in the literature. Potential heat transfer mechanisms are convection, conduction,

or a mixture of both. Therefore, the question which mode dominates at different melting levels has been discussed for decades [1, 2]. Some authors thought that conduction took an essential part in phase transition operations [3, 4]. However, several researchers have reported that natural convection, especially during melting, is the essential heat transfer mode in phase transformation.

Among the first works carried out in this context, we can cite the experimental work in [5, 6] as well as the numerical simulations in [7, 8] that proved that natural convection could have a powerful impact on the rate of heat transfer and the form of the melting front. Later, authors in [9] developed a scale theory and identified the boundaries of the different regimes of the heat transfer mechanisms present in the melting of PCM in a laterally heated square enclosure. The research results demonstrated that the conduction heat transfer mechanism is dominant in the melting process's initial phases. The natural convection induced by the buoyancy of the liquid is intensified as melting progresses. They mentioned the existence of a transition region between these two. Several recent experimental and computational studies have shown that a mixture of conduction and convection is the heat transfer mechanism in an LHTES system [10–14].

As discussed above, the work concentrating on heat transfer with a fixed hot wall temperature in a basic LHTES device configuration was substantially different from the actual situation where the hot source's temperature is not constant. Thus, this paper focuses on PCM's heat transfer mechanics during the charging processes in the horizontal cylindrical shell-and-tubes LHTES system driven by a solar collector in which the temperature of the HTF changes every moment during the day according to a journal solar irradiation curve.

Corresponding author: Fatima Zohra Mecieb

II. NUMERICAL SIMULATION

A. Physical Model

Figure 1 shows a schematic illustration of a horizontal shell and tubes LHTES unit. The designed LHTES unit consists mainly of 12 stainless steel tubes with a 17.2mm diameter and a thickness of 2.31mm. The heat transfer fluid flows in series through the inner tubes, and the PCM is uniformly distributed on the shell side.

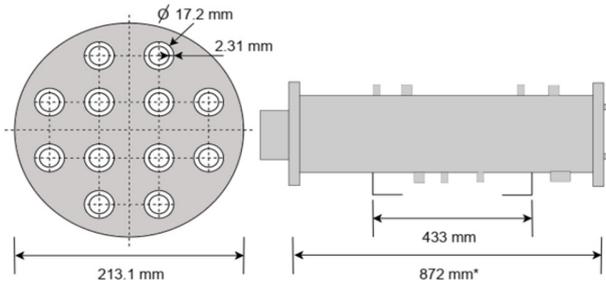


Fig. 1. Characteristic dimensions of the geometry.

The enthalpy-porosity method developed in [15] is employed in modeling and simulating the phase change process. In this method, the enthalpy H is written as the sum of the sensible enthalpy $h_{sensible}$ and the latent enthalpy h_{latent} .

$$H = h_{sensible} + h_{latent} \quad (1)$$

with:

$$h_{sensible} = h_{ref} + \int_{T_{ref}}^T Cp \cdot dT \quad (2)$$

The latent heat enthalpy can vary between zero (for a solid) and ΔH_L (for a liquid).

$$h_{latent} = L_f \cdot \Delta H_L \quad (3)$$

The liquid fraction L_f can be defined as:

$$\begin{cases} L_f = 0 & T < T_{solidus} \\ L_f = 1 & T < T_{liquidus} \\ L_f = \frac{T - T_{solidus}}{T_{liquidus} - T_{solidus}} & T_{solidus} < T < T_{liquidus} \end{cases} \quad (4)$$

The correct characterization of material properties is one of the essential parts of performing simulations. The construction material of the accumulator is steel, whose characteristics are presented in Table I.

TABLE I. STEEL PROPERTIES

Properties	Values
Density [kg/m ³]	8030
Specific heat [J/kg. K]	502.48
Thermal conductivity [W/m.K]	16

The other material to be defined in the system is the paraffin RT70HC, with a melting point around 70°C. This paraffin is used as phase change material, storing the installation's solar collector's absorbed thermal energy. The

principal reason for choosing paraffin wax is that its melting temperature conforms with the HTF temperature coming from the solar collectors at medium temperatures. This material has been analyzed and detailed correlations of the material's descriptive parameters have been provided in [16]. The melting and solidification temperatures of the paraffin and the latent thermal energy must be added to these correlations.

$$T_{solidus} = 66.76 \text{ }^\circ\text{C}$$

$$T_{liquidus} = 78.85 \text{ }^\circ\text{C}$$

$$\Delta H_L = 244970 \text{ J/kg}$$

Note that paraffin should be treated as a fluid that is initially in a solid-state.

B. Initial and Boundary Conditions

A two-dimensional domain is considered saving computational cost. Since there is no depth, there will be no HTF flow, and the boundary condition "convective heat transfer" is applied to the tubes' wall. Therefore, only the "free stream temperature" is considered. A UDF is employed to ensure the three-dimensional aspect of the solution by calculating the outlet temperature of each tube using the following equation:

$$T_{out} = \frac{\dot{Q}}{\dot{m}_{HTF} \cdot Cp_{HTF}} + T_{in} \quad (5)$$

where T_{in} is the output temperature of the previous tube and \dot{Q} is the heat that the PCM absorbed from the HTF at the current time which is calculated by:

$$\dot{Q} = h_{int} \cdot A(T_{fs} - T_w) \quad (6)$$

The heat transfer coefficient h_{int} between water and the internal tubes' walls is calculated in the UDF according to [17] by the following equations:

$$h_{int} = Nu_D \frac{k_{HTF}}{D} \quad (7)$$

$$Nu_D = \frac{(F/8)(Re-1000)Pr}{1+12.7(F/8)^{1/2}(Pr^{2/3}-1)} \quad (8)$$

The correlation is valid for $0.5 < Pr < 2000$ and $3000 < Re < 5 \cdot 10^6$.

The friction factor F is calculated by:

$$F = (1.58 \ln Re - 3.28)^{-2} \quad (9)$$

Initially, the solid PCM is at a temperature of 295°K. The HTF flow rate in tubes is 1.84L/min with Reynold number $Re=3,089,37$. The heat in (6) is exploited to update the HTF temperature at the outlet of one tube (9) that corresponds to the boundary condition (inlet) temperature of the next tube.

Since a series flow is adopted, the fluid will flow from one tube to the other by gradually decreasing the temperature. At each time step, a different fluid temperature would be imposed for each tube, which implies calculating the solution in each of them. The accumulator will be directly connected to a solar collector (Figure 2), with an efficiency given by (10), (11) and an absorption surface of $A_{col}=3.03\text{m}^2$.

$$\eta = 0.514 - 1.158.T_{corr} - 0.005.G.T_{corr}^2 \quad (10)$$

with:

$$T_{corr} = (T_{COL_{in}} - T_a) / G \quad (11)$$

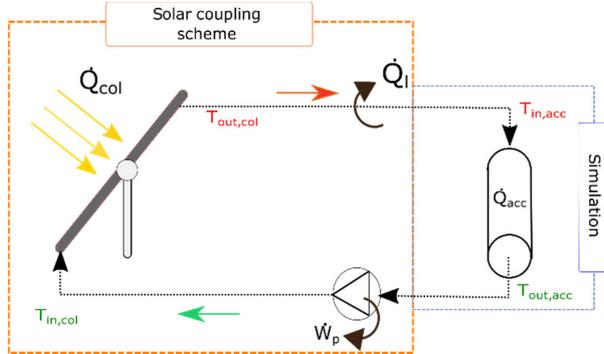


Fig. 2. Schematic diagram of the primary circuit of the installation.

As a general rule, under ideal conditions (without cloud interference), solar radiation can be expressed by a positive sinusoidal function whose period corresponds to the duration of the day. According to the Murcia region's climatic data (in the south of Spain), 3 solar curves can be formed depending on the period (winter, mixed, and summer). The heat is ideally absorbed by the solar collector per unit area, G .

$$G = 270. \left(1 - \cos\left(\frac{\pi.t}{18000}\right)\right) \quad (12)$$

$$G = 392. \left(1 - \cos\left(\frac{\pi.t}{21000}\right)\right) \quad (13)$$

$$G = 480. \left(1 - \cos\left(\frac{\pi.t}{27000}\right)\right) \quad (14)$$

The progressive rise of the accumulator fluid inlet temperature, as a function of the irradiation dynamics involved in (12)-(14), drives the storage of thermal energy in the PCM accumulator through a solar coupling algorithm programmed in the UDF. Equation (15) gets the temperature at the solar collector's output at each moment:

$$T_{COL_{out}} = T_{COL_{in}} + Q_{COL} \cdot \dot{m}_{HTF} \cdot Cp_{HTF} \quad (15)$$

with:

$$Q_{COL} = A_{COL} \cdot G \cdot \eta \quad (16)$$

The resulting temperature of the HTF at the solar collector outlet will be specified at the storage system inlet. The solar coupling scheme will update the heat-transfer fluid's temperature at each moment, considering the solar radiation using the algorithm described above.

C. Numerical Procedure

Considering the shell-and-tube LHTES design, hexahedron quad and map elements are adopted to discretize the PCM domain. The cells near the inner tube are refined to get accurately the velocity and temperature gradients. The second-order discretization scheme is used for the momentum and energy equations' spatial discretization to have higher accuracy in the solution. A pressure-velocity coupling by the algorithm

SIMPLEC is applied in this calculation, with a modification of the under-relaxation factors. The following simplifications are applied to the 2D model:

- The liquid paraffin and the heat transfer fluid are incompressible Newtonian fluids.
- PCM is assumed to be isotropic and homogeneous and the volumetric expansion of PCM during melting is neglected.

In the present simulation, the boundary condition by convective heat transfer and the strategies followed for the coupling of the solar curve by UDF represent the fundamental basis for this accumulator's modeling. The system's performances are analyzed by modifying the thermal load of the HTF through varying the irradiance curve according to the climatic conditions.

III. VALIDATION

The simulation results are compared with the experimental data from [16] to verify the actual numerical simulations. The comparison of measured irradiation and liquid fraction and the simulation results for the cases of a sunny day of May 31, 2019 are shown in Figure 3. A reasonable agreement can be seen between the numerical and the experimental results (with a maximum liquid fraction absolute error of 0.1812), confirming the combined enthalpy porosity and natural convection model's feasibility in the present investigation.

IV. RESULTS AND DISCUSSION

A. Comparison of Conduction and Conduction-Convection Models

Figures 4 and 5 show the temperature and liquid fraction contours in the storage tank's paraffin, whose heat is supplied by two solar collectors for a winter curve. The melting process is simulated using a convection and conduction model and a pure conduction model. As illustrated in Figure 4(a), the temperature contours are symmetrical along the horizontal and vertical axes. The conduction heat transfer mechanism could interpret that. The phase change material temperature is monitored by the energy absorbed from the HTF by the PCM. Thermal energy is transmitted from the hot paraffin to the adjacent layer of cold paraffin. The temperature at one location in the accumulator could remain stable or increase depending on the two energy transfer processes' rates. The total heat transfer rate in the conduction model seems to be the same in all directions. The melting front in the radial direction passes from the tube wall to the shell's internal surface at the same velocity (Figure 5(a)). Nevertheless, the results found from the convection-conduction model are distinct. Figure 4(b) displays the temperature contours through the melting phase. The PCM temperature varies considerably in the vertical axis at distinct moments during the melting process. In the convection-conduction model, heat transfer from the tube surface to the PCM proceeds the conduction just before PCM begins to melt. As the paraffin melts, a thin liquid film forms between the tubes' outer wall and the solid PCM and natural convection begins to occur. As the molten area expands, the convection gets more significant, and the interface front moves quicker (Figure 5(b)).

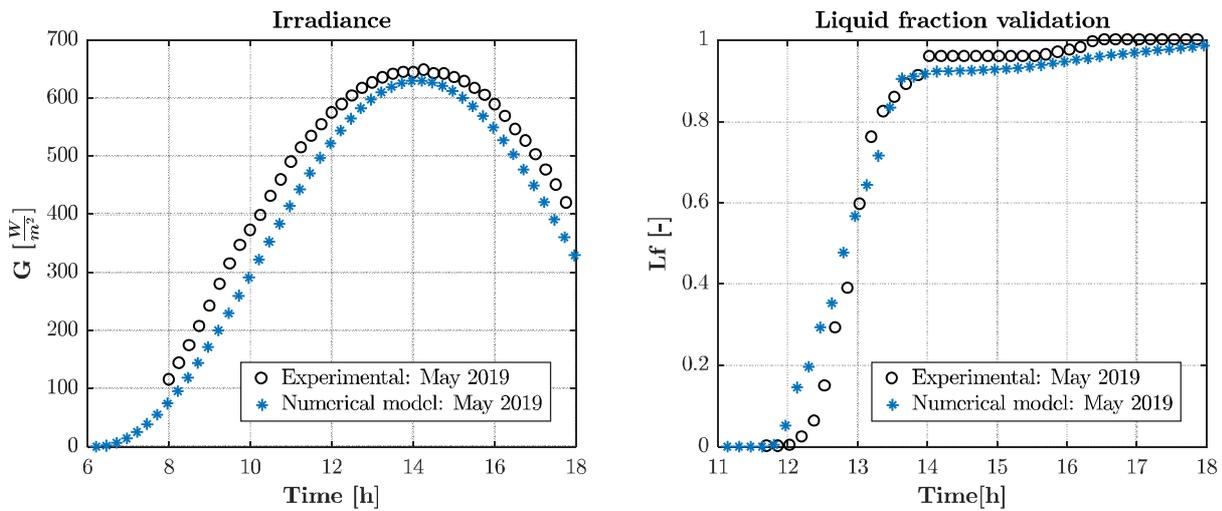


Fig. 3. Irradiance and PCM liquid fraction by current numerical model in comparison with the experimental results of [16].

As shown in **Error! Reference source not found.**(b), the liquid PCM increases and progressively occupies the storage tank's upper section due to buoyancy. In the top zone of the accumulator, PCM takes less time to melt than the bottom. This result is compatible with that reported in [16].

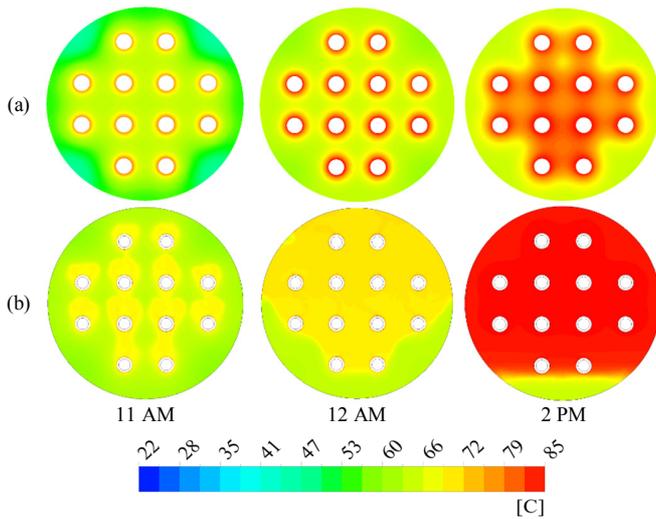


Fig. 4. Temperature contours: (a) conduction model, (b) convection-conduction model.

B. Nusselt Number.

Nusselt identifies the relationship between convective and conductive heat transmission. It is expressed according to (17):

$$Nu = \frac{h_{conv} \cdot D}{k} \quad (17)$$

The convection coefficient on this LHTES side h_{conv} is such that:

$$h_{conv}^i = \frac{\dot{q}^i}{T_w^i - T_{f,PCM}^i} \quad (18)$$

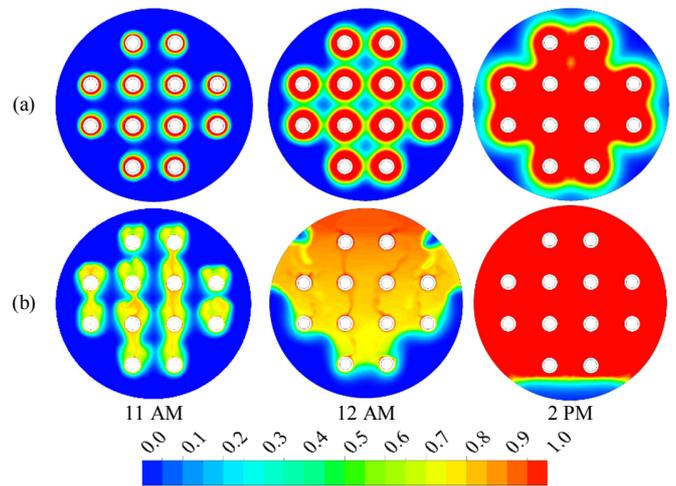


Fig. 5. Liquid fraction contours: (a) conduction model, (b) convection-conduction model.

At each moment, only the fluid portion (liquid) of the PCM is considered getting an average PCM liquid temperature value, $T_{f,PCM}^i$.

Figure 6 shows the solution obtained for the simulations using different solar curves. The Nusselt number is represented as a function of dimensionless time θ , which is given by:

$$\theta = Fo \cdot Ste \quad (19)$$

The temporal evolution of the Nusselt number is represented as a function of different heat transfer regimes (Figure 6):

- A pure conduction regime for $\theta > 0$ (corresponding to zone I). As the temperature gradient initially has significant values due to the increased tube wall temperature, the Nusselt number decreases quickly within the initial melting step. The weak heat transfer, distinguished by a monotonous drop of the Nusselt number, is the particularity of this conduction regime.

- A mixed convection-conduction regime for $0.017 \leq \theta \leq 0.043$ (corresponding to zone II) where the conduction process is gradually replaced by convection. One crucial finding associated with convection melting is that the impact of natural convection ascends with θ [9].
- A convection-dominated regime for $\theta > 0.043$ (corresponding to Zone III). Within the asymptotic limit of θ , we can get the simplified Nu law $\sim Ra^{1/4}$ [9]. The plateau at $Ra^{1/4}$ refers to pure convective heat transfer and is shown in Figure 6 for $0.043 < \theta < 0.11$. Besides, the average Nusselt

number between the heated tubes and the PCM liquid zone is relatively independent of time in the convection-dominated regime. Therefore, the convection-dominated regime is also known as the "quasi-steady" regime.

- The numerical results indicate a decrease of Nu in the ultimate phase ($\theta > 0.11$) when the liquid PCM fills the majority of the accumulator ($L_f > 0.8$) (corresponding to zone IV). Stratification dominates in this advanced evolution of the melting process.

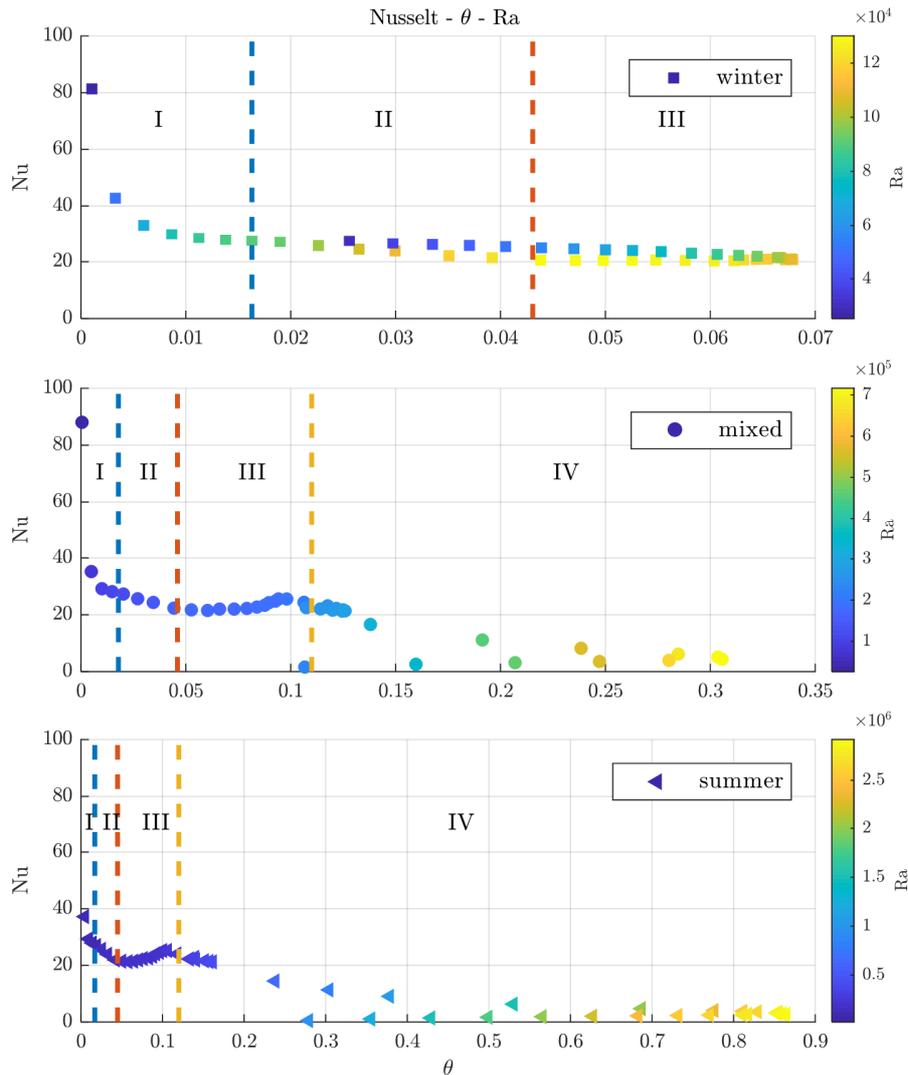


Fig. 6. Nusselt number according to θ .

V. CONCLUSION

A paraffin melting process has been carried out in a tube-and-shell LHTES integrated with a solar collector. The results can be summarized as follows:

- In the convection-conduction model's temperature field, the rise in temperature in the shell's upper zone was faster than

in the low zone, whereas in the conduction model, there was a simultaneous rise in temperature at the shell's top and bottom. The convection-conduction model better interprets the heat transfer mechanism during PCM melting.

- The performance of the LHTES in summer conditions is higher than in mixed and winter conditions, but the PCM

temperature exceeds 100°C, which must be avoided to prevent PCM degradation.

- The melting of the complete paraffin passes via 4 regimes, and each regime has its own Nusselt number. The relation between the Nusselt number and dimensionless time initially indicates a significant drop (conduction), a transition (conduction and convection), a plateau (convection), and, eventually, a gradual decrease towards the disappearance of the solid (stratification).

NOMENCLATURE

D	Diameter (m)
C_p	Specific heat capacity ($J \cdot kg^{-1} \cdot K^{-1}$)
Fo	Fourier number, $Fo = \frac{\alpha t}{D^2}$
ΔH_L	Latent heat ($J \cdot kg^{-1}$)
k	Thermal conductivity ($W \cdot m^{-1} \cdot K^{-1}$)
\dot{m}	Mass flow ($kg \cdot h^{-1}$)
Pr	Prandtl number
T	Temperature (K)
\bar{T}	Average temperature (K)
T_w	Wall temperature (K)
Ste	Stefan number, $Ste = \frac{C_p \Delta T}{\Delta H_L}$
t	Time (s)
α	Thermal diffusivity ($m^2 \cdot s^{-1}$)
μ	Viscosity, $kg \cdot m^{-1} \cdot s^{-1}$
ρ	Density ($kg \cdot m^{-3}$)

REFERENCES

- [1] H. Hu and S. A. Argyropoulos, "Mathematical modelling of solidification and melting: a review," *Modelling and Simulation in Materials Science and Engineering*, vol. 4, no. 4, pp. 371–396, Jul. 1996, <https://doi.org/10.1088/0965-0393/4/4/004>.
- [2] Y. Dutil, D. R. Rousse, N. B. Salah, S. Lassue, and L. Zalewski, "A review on phase-change materials: Mathematical modeling and simulations," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 1, pp. 112–130, Jan. 2011, <https://doi.org/10.1016/j.rser.2010.06.011>.
- [3] N. B. Khedher, "Numerical Study of the Thermal Behavior of a Composite Phase Change Material (PCM) Room," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2663–2667, Apr. 2018, <https://doi.org/10.48084/etasr.1824>.
- [4] Z. Tan, K. M. Lim, and B. C. Khoo, "An adaptive mesh redistribution method for the incompressible mixture flows using phase-field model," *Journal of Computational Physics*, vol. 225, no. 1, pp. 1137–1158, Jul. 2007, <https://doi.org/10.1016/j.jcp.2007.01.019>.
- [5] E. M. Sparrow, S. V. Patankar, and S. Ramadhyani, "Analysis of Melting in the Presence of Natural Convection in the Melt Region," *Journal of Heat Transfer*, vol. 99, no. 4, pp. 520–526, Nov. 1977, <https://doi.org/10.1115/1.3450736>.
- [6] C. Benard, D. Gobin, and F. Martinez, "Melting in Rectangular Enclosures: Experiments and Numerical Simulations," *Journal of Heat Transfer*, vol. 107, no. 4, pp. 794–803, Nov. 1985, <https://doi.org/10.1115/1.3247506>.
- [7] H. Rieger, U. Projahn, M. Bareiss, and H. Beer, "Heat Transfer During Melting Inside a Horizontal Tube," *Journal of Heat Transfer*, vol. 105, no. 2, pp. 226–234, May 1983, <https://doi.org/10.1115/1.3245567>.
- [8] M. Okada, "Analysis of heat transfer during melting from a vertical wall," *International Journal of Heat and Mass Transfer*, vol. 27, no. 11, pp. 2057–2066, Nov. 1984, [https://doi.org/10.1016/0017-9310\(84\)90192-3](https://doi.org/10.1016/0017-9310(84)90192-3).
- [9] P. Jany and A. Bejan, "Scaling theory of melting with natural convection in an enclosure," *International Journal of Heat and Mass Transfer*, vol. 31, no. 6, pp. 1221–1235, Jun. 1988, [https://doi.org/10.1016/0017-9310\(88\)90065-8](https://doi.org/10.1016/0017-9310(88)90065-8).
- [10] X. Sun, Y. Chu, Y. Mo, S. Fan, and S. Liao, "Experimental investigations on the heat transfer of melting phase change material (PCM)," *Energy Procedia*, vol. 152, pp. 186–191, Oct. 2018, <https://doi.org/10.1016/j.egypro.2018.09.079>.
- [11] A. Memon, G. Mishra, and A. K. Gupta, "Buoyancy-driven melting and heat transfer around a horizontal cylinder in square enclosure filled with phase change material," *Applied Thermal Engineering*, vol. 181, Nov. 2020, Art. no. 115990, <https://doi.org/10.1016/j.applthermaleng.2020.115990>.
- [12] N. Mallya and S. Haussener, "Buoyancy-driven melting and solidification heat transfer analysis in encapsulated phase change materials," *International Journal of Heat and Mass Transfer*, vol. 164, Jan. 2021, Art. no. 120525, <https://doi.org/10.1016/j.ijheatmasstransfer.2020.120525>.
- [13] A. S. Soliman, S. Zhu, L. Xu, J. Dong, and P. Cheng, "Numerical simulation and experimental verification of constrained melting of phase change material in cylindrical enclosure subjected to a constant heat flux," *Journal of Energy Storage*, vol. 35, Mar. 2021, Art. no. 102312, <https://doi.org/10.1016/j.est.2021.102312>.
- [14] M. A. Aichouni, N. F. Alshammari, N. B. Khedher, and M. Aichouni, "Experimental Evaluation of Nano-Enhanced Phase Change Materials in a Finned Storage Unit," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5814–5818, Jun. 2020, <https://doi.org/10.48084/etasr.3616>.
- [15] V. R. Voller and C. Prakash, "A fixed grid numerical modelling methodology for convection-diffusion mushy region phase-change problems," *International Journal of Heat and Mass Transfer*, vol. 30, no. 8, pp. 1709–1719, Aug. 1987, [https://doi.org/10.1016/0017-9310\(87\)90317-6](https://doi.org/10.1016/0017-9310(87)90317-6).
- [16] A. Egea, J. P. Solano, J. Pérez-García, and A. García, "Solar-driven melting dynamics in a shell and tube thermal energy store: An experimental analysis," *Renewable Energy*, vol. 154, pp. 1044–1052, Jul. 2020, <https://doi.org/10.1016/j.renene.2020.03.078>.
- [17] F. P. Incropera, D. P. DeWitt, T. L. Bergman, and A. S. Lavine, *Fundamentals of Heat and Mass Transfer*, 6th ed. Hoboken, NJ, USA: John Wiley & Sons, 2006.

Electromagnetic Torque Analysis of SRM 12/8 by Rotor/Stator Pole Angle

Dinh Bui Minh

School of Electrical Engineering
Hanoi University of Science and Technology
Hanoi, Vietnam

Tuan Le Anh

Hanoi University of Industry
Hanoi, Vietnam

Linh Dinh Hai

School of Electrical Engineering
Hanoi University of Science and Technology and
College of Electromechanical and Civil Engineering
Vietnam National University of Forestry
Hanoi, Vietnam

Vuong Dang Quoc

School of Electrical Engineering
Hanoi University of Science and Technology
Hanoi, Vietnam

Abstract—This paper presents the harmonic torque reduction by the different rotor pole angles of a three-phase 12/8 switched reluctance motor via an analytical model and simulation method. Improving torque performance by stator and rotor angles was applied for three-phase switched reluctance motor at stator pole/rotor pole ratios of 6/4, 8/12, 18/12, and 24/18. The average torque and the torque ripple effect by stator and rotor pole embrace have been recently studied in many projects. Due to the fact that leakage flux, flux density, and inductance are affected by the stator and rotor pole angles non-linear and linear leakage flux curves occur. Many stator and rotor pole angle combinations for the three-phase switched reluctance motor have already been done via a finite element method. In this paper, turn-on and turn-off angles will be figured based on stator and rotor pole embraces.

Keywords—Switched Reluctance Motor (SRM); pole arcs; torque; Finite Element Method (FEM)

I. INTRODUCTION

Nowadays, the Switched Reluctance Motor (SRM) is considered as the most appropriate candidate to drive small scale electric vehicles due to advantages such as its simple construction, wide constant power regions, and effective torque speed characteristics. The SRMs have been applied in various fields [1-5]. A convenient design of the SRM has a significant importance for an efficient performance of the motor. Some design parameters must be considered such as the number of phases, stator pole arc, air gap thickness, and other SRM geometry elements. The effect of the SRM geometry on the motor performance has been examined thoroughly. In [6], the influence of pole embrace has been investigated on the performance of the SRM by the Finite Element Method (FEM) with stator and rotor pole embraces changing from 0.2 to 0.5. According to the highest average torque and lowest torque

ripple, the proper pole embraces have been selected for the optimum design of the SRM [7-9].

II. ELECTROMAGNETIC TORQUE ANALYSIS

The rotor and stator pole arc angles are a critical part in the design of the SRMs. The rotor pole arc (β_r) must be equal or greater than the stator pole arc (β_s), because the number of rotor poles is less than the number of stator poles ($\beta_s \leq \beta_r$). When the β_s is smaller than the step angle, then none of the phases can have rising inductance slope. The rotor pole angle should be greater than the sum of stator and rotor pole arcs [10]:

$$\beta_r \geq \beta_s; \quad \beta_r + \beta_s < \frac{360^\circ}{N_r}; \quad \min(\beta_r, \beta_s) \geq \frac{360^\circ}{m \cdot N_r} \quad (1)$$

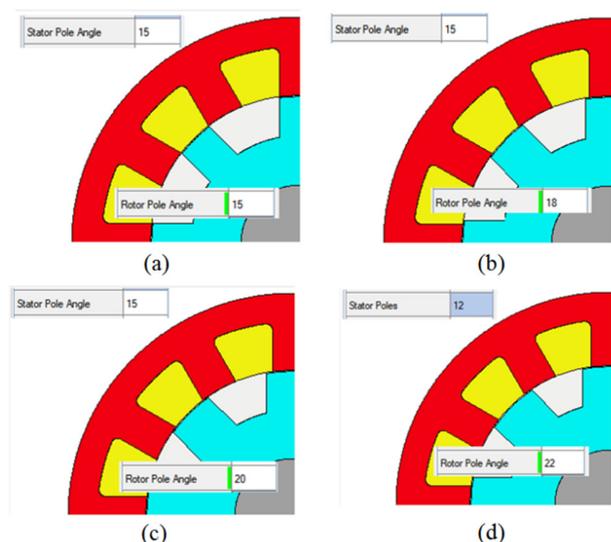


Fig. 1. Cross-section of a three-phase SRM 12/8 β_r equal to (a) 15°, (b) 18°, (c) 20°, (d) 22° for $\beta_s=15^\circ, 15^\circ, 15^\circ$, and 12° respectively.

The initial size L is estimated from the output power and torque with speed of 1500rpm. With an initial assumption for the L and with respect to the average torque proportionality with LD^2 , the values of L and D have been calculated via the analytical method. The three phase SRMs 12/8 with rotor angle ($15^\circ \div 22^\circ$) are listed in Table I.

TABLE I. TECHNICAL PARAMETERS OF SRM 12/8

Technical Parameters	Value	Units
Stator poles	12	
Stator pole angle	15	degrees
Stator lam dia	140	mm
Stator bore	90	mm
Stator pole depth	15	mm
Stator duct layers	2	
Rotor pole angle	15, 18, 20, 22	degrees
Rotor slot depth	11	mm
Airgap	0.3	mm
Number strands in hand	9	
Phases	3	
Turns per phase	12	
Power	1200	W
Speed	1500	rpm
Torque	12	N.m
Rotor pole τ_r ($360/Nr=360/8$)	45	degrees
Stator pole τ_s ($360/Ns=360/12$)	30	degrees

The static torque curves (T_c) have been implemented with FEM with current ranging from 0 to 50A depending on the rotor angle, i.e.:

$$T_c = \frac{1}{2} i^2 \frac{dL(\theta, i)}{d\theta} \quad (2)$$

The rotor embrace ratio a_s and stator embrace ratio a_r are determined by:

$$a_s = \frac{\beta_s}{\tau_s}, \quad a_r = \frac{\beta_r}{\tau_r} \quad (3)$$

In order to maximize the static electromagnetic torque, the stator embrace angle values of 0.45, 0.5, 0.55, 0.6, or 13.5° , 15° , 16.5° , 18° have been applied to the finite element simulation to get the static torque, and the rotor angle is verified from 0 to 45° as shown in Figure 2.

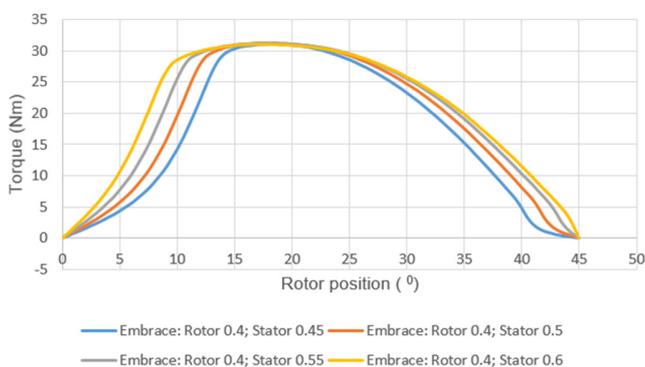


Fig. 2. The static torque curves and rotor position stator pole embrace for rotor embrace of 0.4.

The static torque curves have been also evaluated with rotor embrace ratio of 0.5 and stator pole embrace values of 0.45, 0.5, 0.55, 0.6 or 13.5° , 15° , 16.5° , 18° . The stator pole embrace is changed from 0.45 to 0.6 with 0.05 steps for constant rotor pole embrace of 0.5 as shown in Figure 3. It is obtained that the combination of 0.40 rotor pole and 0.5 stator pole embraces produces the highest peak output torque which is 31.046Nm. Besides, the combination of stator and rotor pole embraces at 0.40 value has the minimum output torque which is 30.920Nm.

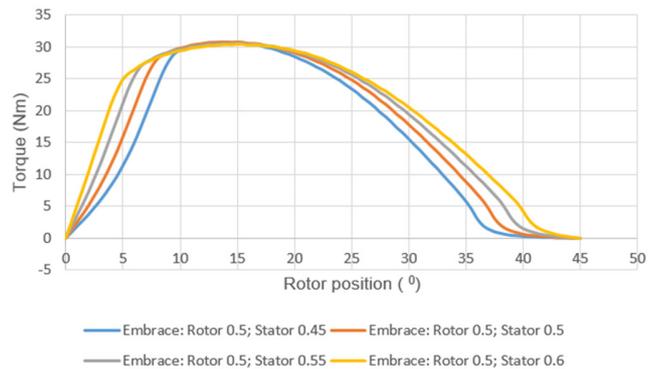


Fig. 3. The static torque curves and rotor position stator pole embrace for rotor embrace of 0.5.

The stator pole embrace is changed from 0.45 to 0.6 with steps of 0.05 for a constant of 0.40 rotor pole embrace depicted in Figure 3. It is seen that the combination of 0.5 rotor pole and 0.50 stator pole embraces produces the highest peak output torque which is 31.191Nm. In the same way, the combination of stator pole embrace and 0.50 rotor pole embrace has the minimum output torque which is 31.094Nm. As can be seen from the results, while the highest output torque is obtained by the combination of 0.40 rotor pole and 0.5 stator pole embraces, the lowest output torque is obtained by the combination of 0.6 rotor pole and 0.6 stator pole embraces. According to these combinations, the optimum values of pole arcs stator/rotor can be selected for better motor design. The average torque is calculated and the best combination of stator and rotor pole embrace can be selected for optimum analysis. The results on the effect of pole embrace are given in Figure 4 for different combinations of the stator and rotor pole embraces.

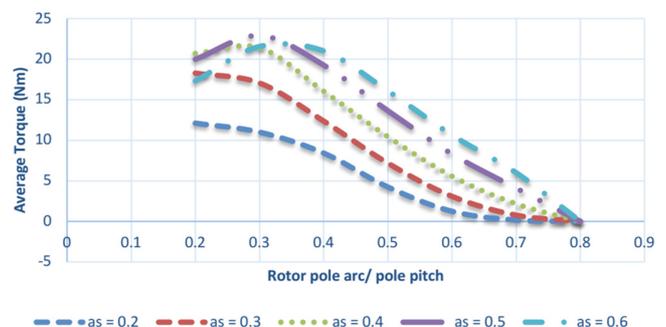


Fig. 4. Average Torque vs rotor embrace.

Figure 5 shows the average torque and torque ripple, and the embrace stator for various values of the embrace rotor. The average torque is maximum or on top of the curve ($a_s=0.5$) at rotor pole arc $a_r=0.3$. With the 4 rotor poles and 6 SRM stator poles, the rotor/stator pitches are 45/30 degrees. Thus, the rotor embrace $a_r=0.3$ is 30 mechanical degrees and the stator embrace is $a_s=0.5$. This means that the rotor pole angle and stator angle have the same value of 15 degrees. The electromagnetic torques with different rotor and stator embraces are presented in Table II. From those results, the best combination of $a_s=0.3$ and $a_r=0.5$ is selected to verify the dynamic torque performance.

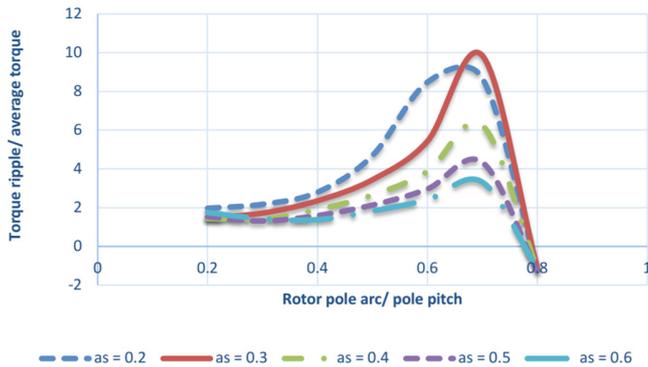


Fig. 5. Torque ripple/average torque rate vs rotor embrace.

TABLE II. AVERAGE TORQUE (Nm).

a_r	$a_s=0.2$	$a_s=0.3$	$a_s=0.4$	$a_s=0.5$	$a_s=0.6$
0.2	12.09608	18.2667	20.69018	19.97587	17.31098
0.3	10.9917	17.04873	21.39854	22.89821	21.54776
0.4	8.411142	12.35922	16.03515	19.25603	21.00559
0.5	4.228949	7.167675	10.4344	13.63371	16.0756
0.6	1.233799	3.079451	5.558719	8.299413	10.63121
0.7	0.150363	0.761016	2.154871	4.085025	5.991545

III. DYNAMIC SIMULATION AND EXPERIMENTAL RESULTS

The finite element model has been presented to validate the dynamic torque of 1500rpm and phase current $I=50A$. The turn on and turn off angles are determined in (4). The torque curve plot is depicted in Figure 6. The torque ripple ranges from 5 to 9N.m.

$$\begin{cases} T_{on} = \theta_{on} = -(\beta_r - \beta_s) \\ T_{off} = \theta_{off} = -(\beta_r - \beta_s) + \frac{\pi}{N_r} \end{cases} \quad (4)$$

The dynamic torque and efficiency with different rotor and stator angles are compared in Table III. The SRM 12/8 with $\beta_s=15^\circ$, $\beta_r=18^\circ$ is the maximum average torque and minimum torque ripple. The harmonic torque of SRM 12/8 is classified with multiples of 3 orders such as 3th, 6th, 9th, 12th and 15th because the SRM 12/8 has 3 phases in stator and 4 coils/phase (Figure 7). The total torque of SRM 12/8 can be calculated by 3th, 6th, 9th, 12th and 15th. The 3th harmonic component is still significant on the torque ripple. Therefore, this study will investigate different stator and rotor embrace angles on the harmonic torque and total torque.

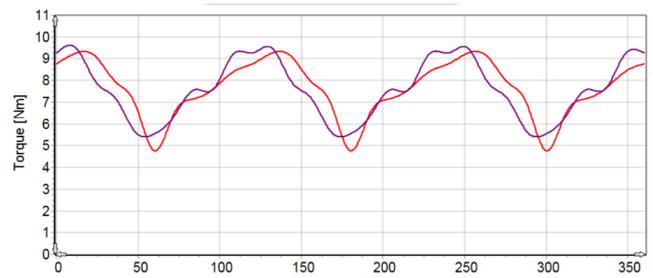


Fig. 6. Dynamic torque curve at 1500rpm.

TABLE III. SRM 12/8 ELECTROMAGNETIC PERFORMANCE COMPARISON

Parameters	Unit	15 ^o /15 ^o	15 ^o /16.5 ^o	15 ^o /18 ^o	15 ^o /20 ^o
Maximum possible torque	Nm	8.466	8.6722	8.6466	8.5649
Average torque	Nm	7.642	8.0111	8.1424	7.9893
Average torque	Nm	7.496	7.9467	8.159	8.1269
Torque ripple	Nm	4.223	1.9376	1.9682	2.4121
Torque ripple [%]	%	57.93	23.812	23.958	29.622
Electromagnetic power	W	916.1	1022.5	1032.4	1023.3
Input power	W	1092	1198.8	1207.6	1195.1
Output power	W	872.7	976.22	985.65	976.53
Total losses (on load)	W	219.4	222.55	221.99	218.59
System efficiency	%	79.91	81.435	81.618	81.71
Shaft torque	Nm	6.945	7.7685	7.8435	7.771

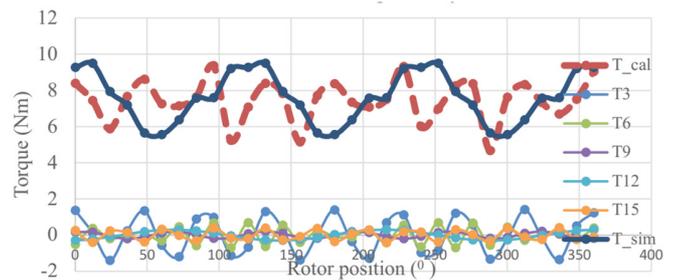


Fig. 7. Harmonic torque.

The dynamic torque test bench installation for SRM 12/8 and permanent magnet generator in the torque transduce is shown in Figure 8. The torque sensor TM300 is connected to PC to get the data during testing.

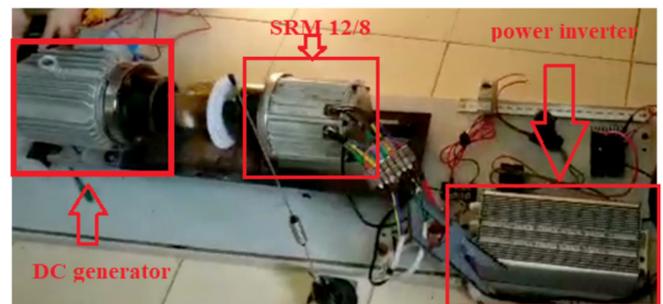


Fig. 8. Dynamic torque test bench.

The speed and torque values with different time steps are presented in Figure 9. In the time interval from 1s to 5s, the

speed and torque are very stabilized and vary trivially. However, torque ripple appears in the beginning time (0.02s). These values are also recorded in starting stage and constant speed. The dynamic torque with different rotor positions is shown in Figure 10. The values on the torque at stator 15 and rotor 22 reach approximately 10Nm.

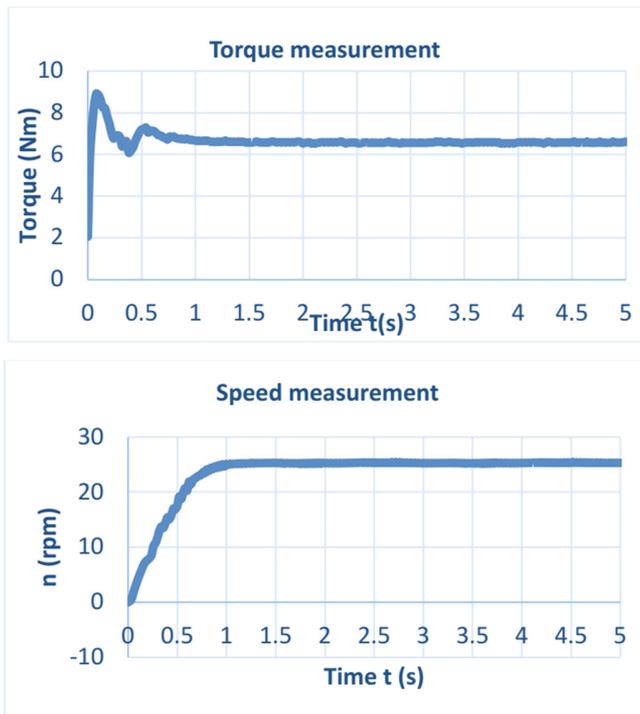


Fig. 9. Dynamic speed (top) and torque (bottom).

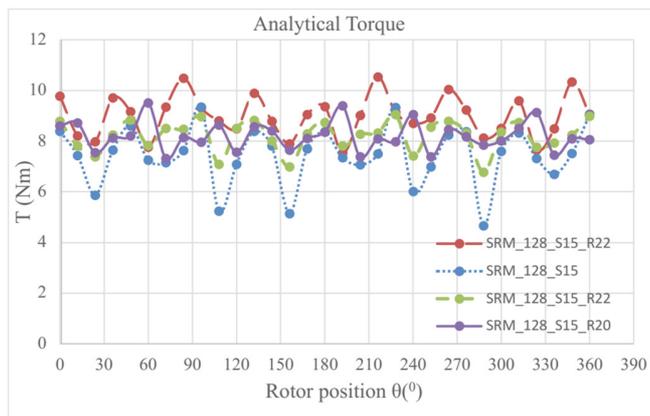


Fig. 10. Dynamic torque with different rotor positions.

IV. CONCLUSION

The current paper has developed stator and rotor embrace/pitch influenced on the average torque and torque ripples via FEM [10, 11]. For high speed of the SRM, the control method is single voltage or current pulse because the magnetic circuit is saturated. Thus, it is difficult for chopping current to apply and the phase current cannot reach reference

values. The electromagnetic design will help improving torque and current controller. The static torque curves have been also evaluated by the FEM. This paper also demonstrates a method to improve and design the SRMs in practice.

REFERENCES

- [1] R. M. Azhagar and A. Kavitha, "Effect of rotor geometry on Peak and Average Torque of direct drive External-Rotor Synchronous Reluctance Motor (Ex-R SynRM) in comparison with Switched Reluctance Motor for low speed domestic application," in *2017 IEEE International Magnetics Conference (INTERMAG)*, Dublin, UK, Apr. 2017, pp. 1–2, <https://doi.org/10.1109/INTMAG.2017.8007641>.
- [2] J. Han, B. Ge, K. Zhang, Y. Wang, and C. Wang, "Influence of Control and Structure Parameters on the Starting Performance of a 12/8 Pole Switched Reluctance Motor," *Energies*, vol. 13, no. 14, Jan. 2020, Art. no. 3744, <https://doi.org/10.3390/en13143744>.
- [3] N. K. Sheth and K. R. Rajagopal, "Optimum pole arcs for a switched reluctance motor for higher torque with reduced ripple," *IEEE Transactions on Magnetics*, vol. 39, no. 5, pp. 3214–3216, Sep. 2003, <https://doi.org/10.1109/TMAG.2003.816151>.
- [4] M. Yildirim and H. Kurum, "Influence of Poles Embrace on In-Wheel Switched Reluctance Motor Design," in *2018 IEEE 18th International Power Electronics and Motion Control Conference (PEMC)*, Budapest, Hungary, Aug. 2018, pp. 562–567, <https://doi.org/10.1109/EPEPEMC.2018.8521859>.
- [5] A. Tap, L. Xheladini, T. Asan, M. Imeryuz, M. Yilmaz, and L. T. Ergene, "Effects of the rotor design parameters on the torque production of a PMSynRM for washing machine applications," in *2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP)*, Brasov, Romania, May 2017, pp. 370–375, <https://doi.org/10.1109/OPTIM.2017.7974998>.
- [6] R. Krishnan, *Switched Reluctance Motor Drives: Modeling, Simulation, Analysis, Design, and Applications*, 1st ed. Boca Raton, FL, USA: CRC Press, 2001.
- [7] Z. Yueying, Y. Chuantian, Y. Yuan, W. Weiyan, and Z. Chengwen, "Design and optimisation of an In-wheel switched reluctance motor for electric vehicles," *IET Intelligent Transport Systems*, vol. 13, no. 1, Jan. 2019.
- [8] A. Siadatan, M. Roohisankestani, and S. Farhangian, "Design and Simulation of a new Switched Reluctance Motor with changes in the shape of stator and rotor in order to reduce torque ripple and comparison with the conventional motor," in *2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, Amalfi, Italy, Jun. 2018, pp. 353–358, <https://doi.org/10.1109/SPEEDAM.2018.8445245>.
- [9] O. M. Al-Barbarawi, "Improving Performance of the Braking Process, and Analysis Torque-Speed Characteristics of the Induction Motor," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3585–3591, Dec. 2018, <https://doi.org/10.48084/etasr.2325>.
- [10] V. D. Quoc, "Robust Correction Procedure for Accurate Thin Shell Models via a Perturbation Technique," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5832–5836, Jun. 2020, <https://doi.org/10.48084/etasr.3615>.
- [11] V. D. Quoc, "Accurate Magnetic Shell Approximations with Magnetostatic Finite Element Formulations by a Subdomain Approach," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5953–5957, Aug. 2020, <https://doi.org/10.48084/etasr.3678>.

Study of the Compressive Strength of Concrete with Partial Replacement of Recycled Coarse Aggregates

Xuan Hung Vu
Department of Construction
Vinh University
Vinh City, Viet Nam
hungvukxd@vinhuni.edu.vn

Trong Cuong Vo
Department of Construction
Vinh University
Vinh City, Viet Nam
cuongvqc@gmail.com

Van Tien Phan
Department of Construction
Vinh University
Vinh City, Viet Nam
vantienkxd@vinhuni.edu.vn

Abstract-This paper presents a study on the compressive strength of concrete using recycled aggregates. The concrete was designed to have a 25MPa compressive strength and an 8cm slump. The rates of replacing natural aggregates with recycled coarse were 0%, 10%, and 20%. The test samples were compressed to determine their compressive strength value after 7, 14, and 28 days of curing. The results showed that the concrete slump did not change effectively at a 10% replacement rate. When using 20% recycled aggregates, the concrete was too hard and the homogeneity of the concrete mixture could not be guaranteed. The compressive strength slightly decreased using 10% of recycled aggregates and decreased significantly using 20%. Therefore, 20% of recycled aggregate replacement is not suitable. The results showed that using recycled aggregates at a rate of 10% is optimal.

Keywords-compressive strength; recycled concrete; demolishing work

I. INTRODUCTION

In recent years, developed countries deal with the recycling and treatment of construction solid waste. Construction's solid waste treatment, solid waste minimization, pressure to change, and the approach to resolving these issues in the UK were studied in [1]. China faces a similar problem as the amount of solid construction waste was more than 1.5 billion tons per year in 2018, and it was predicted to reach 2.5 billion per year in 2020 [2]. The solid waste treatment in current construction sites, the life cycle of sustainable development, and the evaluation of the environmental impact of the construction materials consumed in China during 2000-2015 were studied in [3]. Many studies have been conducted on recycled aggregates, such as their use in green concrete [4, 5], coarse aggregates for concrete [6], and fine concrete [7, 8]. While the reuse of construction solid waste as aggregates for concrete has been widely used and achieved remarkable efficiency, many studies have been conducted recently on the mechanical properties of concrete with recycled aggregates. The influence of the recycled aggregates on the compressive strength of concrete was studied in [9], concluding that they did not have a significant influence on compressive, flexural, and tensile strength. The use of recycled aggregates from demolished works after crushing and grading, helps to save natural aggregate sources and protect the environment, was studied in

[10]. The compressive strength of concrete made from recycled coarse aggregates was studied in [11], with consideration of the source of the recycled aggregates and the strength of the target concrete. The toughness and soundness test results on the recycled coarse aggregates showed a higher percentage loss than the natural, but it remained within the acceptable limits. The compressive and splitting tensile strengths of concrete with recycled coarse aggregate depend on the mix proportions. In general, the strength of the recycled concrete can be 10-25% lower than the conventional made with natural coarse aggregates.

This paper presents a study on the compressive strength of concrete having a partial replacement of natural with recycled coarse aggregates.

II. COMPOSITION OF RECYCLED CONCRETE

The concrete obtained after a project's demolition was crushed and the resulting coarse aggregates were washed, dried, and pre-screened to remove any dust particles. These raw aggregates were screened to classify the particles and then mixed again to be graded as standard for concrete according to [12]. To avoid strongly absorbing recycled aggregates affecting the setting of the concrete, the coarse aggregates were soaked in water and then dried under normal conditions. In this study, a natural concrete aggregate was selected as a reference. The coarse aggregate was replaced partially by solid waste from the crushed demolition site. After casting according to the standard, the samples were stored under normal conditions and tested for compressive strength after 7, 14, and 28 days.



Fig. 1. Recycled coarse aggregates

A. Sand

The mechanical properties of the sand used in the experiment were:

- Density: 2.65g/cm³
- Modulus of magnitude: 2.50
- Volumetric mass: 1660kg/m³

B. Cement

A commercial Portland cement PCB40 was used, having the following mechanical properties:

- Actual strength: 40MPa
- Density: 3.1g/cm³

C. Coarse Aggregates

Natural coarse aggregates were mixed with recycled materials. In this study, 20mm aggregates, the most common size of aggregates used in construction, was used having the following mechanical properties:

- Density: 2.61g/cm³
- Volumetric mass: 1430kg/m³

D. Concrete

The designed recycled concrete had the grade of B20, and the design slump was 8cm. The compressive strength of the concrete was tested according to the Vietnamese standard 3118:1993. Cylindrical samples with size D150×H300 were cast and cured in water before being compressed to determine their compressive strength. The loading speed was set to 0.5kN/s. The replacement rates of coarse aggregates were 0%, 10%, and 20% of the standard reference sample, respectively. The mix components for 1m³ of concrete are presented in Table I.

TABLE I. MIX COMPONENTS FOR 1M³ OF CONCRETE

Mix components for 1m ³ of concrete				
Aggregates by mass	Cement (kg)	Sand (kg)	Aggregate (kg)	Water (kg)
	292.5	648.3	1216.3	195.0
Aggregates by volume	Cement (kg)	Sand (m ³)	Aggregate (m ³)	Water (l)
	292.5	0.391	0.851	195.0

TABLE II. EXPERIMENTAL DETAILS OF CONCRETE SAMPLES

No.	Notation of samples	Description	Quantity
1	CP0	Reference sample: Mix according to Table I	3 samples for 7 days, 3 samples for 14 days, and 3 samples for 28 days.
2	CP10	Recycled aggregate replacement rate: 10%	
3	CP20	Recycled aggregate replacement rate: 20%	

III. RESULTS AND DISCUSSION

The results of measuring the slump of the concrete with recycled coarse materials are shown in Table III and Figure 2.

TABLE III. THE SLUMP OF THE RECYCLED CONCRETE

No.	Replacement rate	Slump (mm)
1	0%	73
2	10%	70
3	20%	12

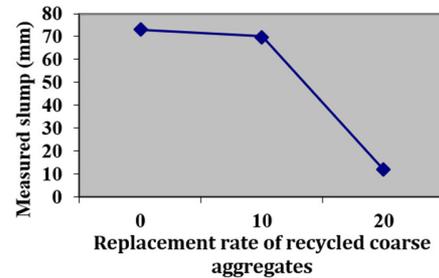


Fig. 2. The slump of recycled concrete at various recycled coarse aggregates replacement rates.

From Figure 2, it can be noted that the slump of concrete did not change significantly at 10% replacement rate. When using 20% recycled coarse aggregates, the drop was markedly obvious as it dropped from 73mm at 0% to 70mm at 10%, and finally to 12mm at 20%. Thus, when using 20% recycled aggregates, the concrete was too hard and its homogeneity could not be guaranteed.



Fig. 3. Samples after the compressive experiments.

Table IV shows the destructive force (kN) results obtained from the compression test. Compressive strength (MPa) was calculated from the destructive force P (kN) for each sample having a 15cm diameter, using:

$$S = \pi \times R^2 = \pi \times 7.5^2 = 176.71cm^2 \quad (1)$$

Compressive strength was calculated, using (1), as:

$$R = a \times P/S \quad (2)$$

where P is the destructive load of the sample, S is the compressive area, and a is the coefficient of converting the experimental results when compressing the samples with

different sizes from the standard samples (150×150×150mm). For a cylinder sample having 150mm diameter and 300mm height, α was calculated to 1.2. Table V shows the compressive strengths of the experimental samples after 7, 14, and 28 days.

TABLE IV. DESTRUCTIVE FORCE RESULTS

No.	Replacement rate of recycled aggregates (%)	Sample destructive force (kN)		
		7 days	14 days	28 days
1	0	185.55	231.94	309.25
2	0	196.30	199.47	316.61
3	0	200.32	239.73	328.39
4	10	195.36	224.52	291.58
5	10	186.26	228.32	300.41
6	10	176.80	240.82	304.83
7	20	130.95	178.79	251.82
8	20	135.02	173.24	254.76
9	20	126.93	161.77	248.87

TABLE V. COMPRESSIVE STRENGTH RESULTS

No.	Replacement rate of recycled aggregates (%)	Compressive strength (MPa)		
		7 days	14 days	28 days
1	0	12.6	15.75	21
2	0	13.33	13.545	21.5
3	0	13.603	16.279	22.3
4	10	13.266	15.246	19.8
5	10	12.648	15.504	20.4
6	10	12.006	16.353	20.7
7	20	8.892	12.141	17.1
8	20	9.169	11.764	17.3
9	20	8.619	10.985	16.9

As it can be noted, the experimental concrete samples did not reach the design strength. The reference sample reached only 86.4% of the design strength. This rate was 81.2% when using 10% and 68.4% when using 20% recycled aggregates. Thus, using 10% recycled aggregates did not affect significantly the compressive strength of the concrete. When using 20% recycled aggregates, the compressive strength dropped significantly. For a clearer observation, the compressive strength results of concrete at various recycle aggregate replacement rates are shown in Figure 4 and Table VI.

TABLE VI. COMPRESSIVE STRENGTH MEAN VALUES FOR EACH EXPERIMENT

No.	Replacement rate of recycled aggregates (%)	Compressive strength (MPa)	Mean value	Achieved compared with design (25 MPa)
1	0%	21	21.6	86.4 %
		21.5		
		22.3		
2	10%	19.8	20.3	81.2 %
		20.4		
		20.7		
3	20%	17.1	17.1	68.4%
		17.3		
		16.9		

The decrease of compressive strength when using recycled aggregates was predicted, as noticed in previous studies [11,

13, 14]. The higher compressive strength of the concrete having a lower recycled aggregate replacement rate may be attributed to the greater bonding force and strength when using the same type of aggregate. Figure 5 shows the strength development of recycled concrete, which indicates that concrete strength reached 51-67% after 7 days and about 63-75% after 14 days of curing. Details are shown in Table VII.

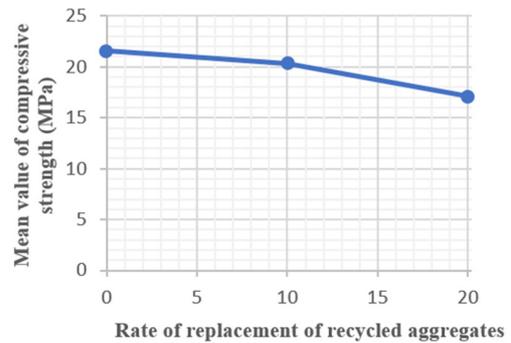


Fig. 4. Compressive strength of concrete at various recycled coarse aggregate replacement rates.

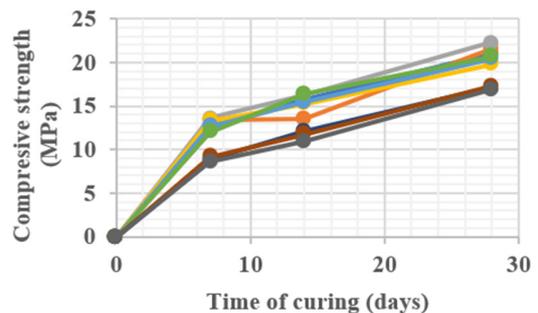


Fig. 5. The development of concrete's compressive strength.

TABLE VII. COMPARISON OF THE COMPRESSIVE STRENGTH OF CONCRETE WITH THE STRENGTH AFTER 28 DAYS OF CURING (R₂₈)

No.	Compressive strength (MPa)			
	7 days	Ratio compared to R ₂₈	14 days	Ratio compared to R ₂₈
1	12.6	60%	15.75	75%
2	13.33	62%	13.545	63%
3	13.603	61%	16.279	73%
4	13.266	67%	15.246	77%
5	12.648	62%	15.504	76%
6	12.006	58%	16.353	79%
7	8.892	52%	12.141	71%
8	9.169	53%	11.764	68%
9	8.619	51%	10.985	65%

As it can be noted, the strength development of the concrete using recycled aggregates at a low rate ($\leq 20\%$) is similar to ordinary concrete. However, a few notes can be pinpointed:

- Normal concrete grows up to 65% of its maximum design strength at the age of 7 days. The experimental samples did not reach this level. This growth rate was slowed down when using 20% recycled aggregates, with a ratio of 51-53% to R₂₈.

- At 14 days of curing, the experimental concrete achieved about 70-75% of the maximum design strength. However, most of the concrete samples using 20% recycled aggregate had slightly lower strength than normal.

IV. CONCLUSION

This paper presents the results of a study on the compressive strength of concrete using recycled aggregate from demolition works. The concrete was designed to have 25MPa compressive strength and 8cm slump. The rates of replacing natural aggregates with recycled coarse were 0%, 10%, and 20%. The test samples were compressed to determine their compressive strength values at 7, 14, and 28 days of curing. The results showed that the concrete slump did not change significantly at samples having 10% recycled aggregates. When using 20% recycled coarse aggregates, the drop was markedly obvious. Thus, using 20% recycled aggregate resulted in too hard concrete, while the homogeneity of the concrete mixture could not be guaranteed. The compressive strength decreased slightly when using 10% recycled aggregates and significantly when using 20%. Therefore, it can be concluded that the 20% replacement rate is not appropriate. The results showed that using recycled aggregates at a rate of 10% is an optimal solution.

REFERENCES

- [1] M. Osmani, "Construction Waste Minimization in the UK: Current Pressures for Change and Approaches," *Procedia - Social and Behavioral Sciences*, vol. 40, pp. 37–40, Jan. 2012, <https://doi.org/10.1016/j.sbspro.2012.03.158>.
- [2] "People's Republic of China: Construction and Demolition Waste Management and Recycling," AECOM Asia Company Limited, Shatin, HK, TA-8906 PRC, Apr. 2018.
- [3] J. Lei, B. Huang, and Y. Huang, "Chapter 6 - Life cycle thinking for sustainable development in the building industry," in *Life Cycle Sustainability Assessment for Decision-Making*, J. Ren and S. Toniolo, Eds. Elsevier, 2020, pp. 125–138.
- [4] V. Radonjanin, M. Malešev, S. Marinković, and A. E. S. Al Maly, "Green recycled aggregate concrete," *Construction and Building Materials*, vol. 47, pp. 1503–1511, Oct. 2013, <https://doi.org/10.1016/j.conbuildmat.2013.06.076>.
- [5] N. Jain, M. Garg, and A. K. Minocha, "Green Concrete from Sustainable Recycled Coarse Aggregates: Mechanical and Durability Properties," *Journal of Waste Management*, vol. 2015, Feb. 2015, Art. no. e281043, <https://doi.org/10.1155/2015/281043>.
- [6] L. Butler, J. S. West, and S. L. Tighe, "Effect of recycled concrete coarse aggregate from multiple sources on the hardened properties of concrete with equivalent compressive strength," *Construction and Building Materials*, vol. 47, pp. 1292–1301, Oct. 2013, <https://doi.org/10.1016/j.conbuildmat.2013.05.074>.
- [7] R. Sri Ravindrarajah and C. T. Tam, "Recycling concrete as fine aggregate in concrete," *International Journal of Cement Composites and Lightweight Concrete*, vol. 9, no. 4, pp. 235–241, Nov. 1987, [https://doi.org/10.1016/0262-5075\(87\)90007-8](https://doi.org/10.1016/0262-5075(87)90007-8).
- [8] L. Evangelista and J. de Brito, "Concrete with fine recycled aggregates: a review," *European Journal of Environmental and Civil Engineering*, vol. 18, no. 2, pp. 129–172, Feb. 2014, <https://doi.org/10.1080/19648189.2013.851038>.
- [9] J. Ahmad Bhat, "Effect of strength of parent concrete on the mechanical properties of recycled aggregate concrete," *Materials Today: Proceedings*, vol. 42, pp. 1462–1469, Jan. 2021, <https://doi.org/10.1016/j.matpr.2021.01.310>.
- [10] J. Junak and A. Sicakova, "Concrete Containing Recycled Concrete Aggregate with Modified Surface," *Procedia Engineering*, vol. 180, pp. 1284–1291, Jan. 2017, <https://doi.org/10.1016/j.proeng.2017.04.290>.
- [11] S. W. Tabsh and A. S. Abdelfatah, "Influence of recycled concrete aggregates on strength properties of concrete," *Construction and Building Materials*, vol. 23, no. 2, pp. 1163–1167, Feb. 2009, <https://doi.org/10.1016/j.conbuildmat.2008.06.007>.
- [12] "About the coarse recycled aggregates for concrete," Vietnam, 11969:2018, 2018.
- [13] V. T. Phan, T. H. Nguyen, "The influence of fly ash on the compressive strength of recycled concrete utilizing coarse aggregate from demolition works," *Engineering, Technology and Applied Science Research*, vol 11, no. 3, pp. 7107-7110, Jun. 2021.
- [14] A. A. Mohammed Ali, R. S. Zidan, and T. W. Ahmed, "Evaluation of high-strength concrete made with recycled aggregate under effect of well water," *Case Studies in Construction Materials*, vol. 12, Jun. 2020, Art. no. e00338, <https://doi.org/10.1016/j.cscm.2020.e00338>.

Cetane Number Improvement of Distilled Diesel from Tawke Wells

Shwan M. Noori Khalifa Majeed
Zakho Technical Institute
Duhok Polytechnic University
Zakho, Kurdistan Region, Iraq
shwan.khalifa@dpu.edu.krd

Shinwar A. Idrees
Chemistry Department, Faculty of Science
University of Zakho
Zakho, Kurdistan Region, Iraq
shinwar.idrees@uoz.edu.krd

Veyan A. Musa
Mechanical Department, College of Engineering
University of Zakho
Zakho, Kurdistan Region, Iraq
veyan.musa@uoz.edu.krd

Sherwan M. Simo
Petroleum Engineering Department, College of Engineering
University of Zakho
Zakho, Kurdistan Region, Iraq
sherwan.simo@uoz.edu.krd

Abraheem A. Mohammed
Chemistry Department, Faculty of Science
University of Zakho
Zakho, Kurdistan Region, Iraq
ibraheem.mohammed@uoz.edu.krd

Lokman A. Abdulkareem
Petroleum Engineering Department, College of Engineering
University of Zakho
Zakho, Kurdistan Region, Iraq
lokman.abdulkareem@uoz.edu.krd

Abstract—The current research aims to improve the cetane number of diesel extracted from the crude oil of Tawke region in Iraqi Kurdistan. A specific mixture of chemical compounds was prepared which included *m*-nitrophenol, 4-nitro toluene, and nitrobenzene. The components' effects were investigated with regard to the cetane number, flash point, viscosity, and refractive index of diesel. The quantity of each compound mixed with diesel was prepared based on the statistical analysis of the experiment device (Box–Behnken Designs-BBDs). The tested mixture showed a good agreement and improvement of cetane and flash point and a very low effect on viscosity and refractive index. According to the statistical analysis, the main influence on cetane number and the flashpoint was from *m*-nitrophenol. The investigation showed that the best results were acquired from the samples of 25PPM 4-nitro toluene and 50PPM *m*-nitrophenol with a cetane number of 65.3. The correlation and the interaction of the regression equation were linear with all cases. It is worth mentioning that all additives positively influenced the cetane number in the regression equation. The sulfur content was measured as well, and the obtained weight percentage of sulfur was 0.8404%.

Keywords—diesel improvers; cetane number; flash point; viscosity

I. INTRODUCTION

Additives are mixed with fuel to improve its quality and to enhance its efficiency. This process is termed as mixing with fuel of the trace elements. Fuel in various forms like diesel and benzene does not perform well or reach the required international standards without improvers [1]. Therefore, the

search for new improvers is in great demand. The utilizing of influential improvers is very essential in order to reach the requested mechanical and environmental standards [2]. In the past, attempts have been made to ignite fuel readily and thus upgrade ignition quality. Nitro compounds are commonly used as cetane number additives and such chemicals make the delay time of fuel ignition shorter and as a result the rate of knocking decreases [2]. Such additives can be utilized with bio-diesel, which commonly has low cetane number [3]. The cetane number is affected by physical and chemical properties of fuel such as density, viscosity, surface tension, and vaporization. The cetane number is also influenced by the molecular structure of fuel and additives [4].

The octane number is a parameter that determines the quality of a fuel. Fuels with high octane numbers, close to 100, knock less. Oxygenated compounds, aromatic hydrocarbons, aromatic amines, and organometallic compounds have been exploited for this purpose as octane number modifiers [4-6]. Flash point temperature is another important factor to improve the valid ability of fuel in car engines [8], where lower flash point temperature means lower ignition temperature (piloted temperature) for the fuel oil ranging from 60 to 93°C. However, this research article mainly considers cetane number improvement. The influence on the cetane number of adding dicyclopentadiene to diesel has been examined in [9]. The authors proposed a novel additive of hydrocarbon-based, dicyclopentadiene (DCPD) for diesel fuel which leads to minimized particulate emissions and also enhances cetane

Corresponding author: Veyan A. Musa.

number. The improvement of diesel properties like viscosity, flash point, density, cloud point, water content, and sulfur content were investigated in [10]. The Design of Experiment (DOE) to study any oil sample represents a significant part of chemo measurements (effects of chemicals on the yield). The universal benefit of DOE is to ensure that the conditions between any test parameters and their yield of the reactions can be evaluated dependably with low cost and exertion with an insignificant number of trials and less amount of chemicals. DOE can be isolated into a few subtopics, for example, confirming factors from an enormous arrangement of factors which is called screening structures, finding the impact of a blend organization on the reaction factors tended to blend plans, and discovering wellsprings of error in estimation frameworks. This leads to concocting ideal conditions inconsistent procedures (evolutionary operation), batch process (response surface methodology), or planning tests for ideal parameter estimation in numerical models (optimal design or optimization) [11]. Different statistical DOE models have been utilized to smooth out factors in the exploratory plan. Accordingly, this helps to select the impact of exploratory factors by conventional methodologies. Tests have been conducted with conscious changes of the specific boundaries. These assessments should be repeated to each boundary effect achieving a reliable number of runs [12].

Numerous studies have been carried out to study the extracted crude oil from the wells of Tawke region in Iraqi Kurdistan [13], but the investigation in crude oil and its components from the specific wells requires more investigations to reach the desired characteristics and to improve its properties. The tests described in the current paper were performed using a sample consisting of diesel mixed with a measured amount of m-nitrophenol, 4-nitro toluene, and nitrobenzene to aggregate the effect of those components on the diesel.

II. EXPERIMENTAL SETUP AND METHODS

The portable analyzer instrument SHATOX SX-100M was used as the main device to analyze the octane/cetane numbers and to determine cetane and octane number, while a device NCL 120 - CLEVELAND OPEN CUP FLASH POINT TESTER MANUAL120 ASTM D 92, IP 36, ISO 2592 was utilized for flash point temperature determination as shown in Figure. 1.

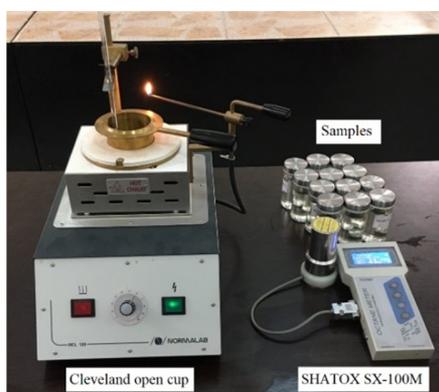


Fig. 1. The experimental devices.

TABLE I. LEVELS OF THE PARAMETERS STUDIED IN THE CCD STATISTICAL EXPERIMENT.

Chemical compound	Unit	Levels		
		-1 (low)	0 (middle)	+1 (high)
m-nitrophenol	PPM	0	25	50
4-nitro toluene	PPM	0	25	50
nitro benzene	PPM	0	25	50

TABLE II. PHYSICAL AND CHEMICAL PROPERTIES

Chemical compound	m-nitrophenol	4-nitro toluene	nitro benzene
IUPAC name	3-Hydroxynitrobenzene	1-methyl-4-nitrobenzene	Nitro benzene
Chemical structure			
Formula	C ₆ H ₅ NO ₃	C ₇ H ₇ NO ₂	C ₆ H ₅ NO ₂
Appearance	Colorless to yellow crystals	Yellowish crystals	Yellowish, oily, aromatic nitro-compound
Molecular weight	139.11g/mol	137.14g/mol	123.11g/mol
Solubility	Soluble in hot and dilute acids and in caustic solutions. Insoluble in petroleum ether. Very soluble in ethanol, ether, and acetone. In water, 13.550mg/L at 25°C	Soluble in alcohol, benzene, ether, chloroform, acetone. In water, 442mg/L at 30°C	Slightly soluble in carbon tetrachloride. Very soluble in ethanol, diethyl ether, acetone, benzene. In water, 2.09×10 ⁻³ mg/L at 25°C
Boiling point	194°C/70mmHg	238.3°C	210.8°C
Melting point	96-98°C	51.6°C	5.7°C
Density	1.485g/cm ³ at 20°C	1.29g/cm ³	1.2037g/cm ³
Vapour pressure	1.5×10 ⁻⁴ mm Hg at 25°C	0.0157mm Hg at 25°C	0.245mm Hg at 25°C
Toxicity	LD50=328mg/kg (Rat)	LD50=2250mg/kg body weight	LD50=640mg/kg (Rat)

M-nitro phenol, 4-nitro toluene, and nitrobenzene supplied from Germany were used as received. Solutions of m-nitrophenol, 4-nitro toluene, and nitrobenzene 50mL/L (PPM) were prepared in diesel and then were diluted to solutions according to Table I. The prepared solutions were covered by aluminum foil and kept in dark to avoid the photodegradation of the 13 samples. Diesel was brought from Tawke oil well distillery and was used as received. Different concentrations of improvers were prepared according to the needs of the experiments as shown in Table I. Table II shows the physico-chemical properties of the additives.

It is essential to fit a logical model in order to depict the response directly in the test field by choosing the DOE. Generally, this model is suitable for illustrating a plane surface, as indicated by:

$$R = \beta_0 + \sum \beta_i X_i + \varepsilon \quad (1)$$

where R is the response, β_0 is the constant term, β_i represents the coefficients of the linear parameters, X_i represents the parameters, and ε is the irregular error or commotion to the response. On specific events, it is called the essential impacts model since it incorporates just the principle impacts of the

factors as explained intensively in [14]. If the interaction between the parameters is contained, then the first-order model becomes:

$$R = \beta_0 + \sum \beta_i X_i + \sum \beta_{ij} X_i X_j + \sum \beta_{ii} X_i^2 + \varepsilon \quad (2)$$

where β_{ii} indicates the quadratic coefficients of the variables and $i < j$.

Response Surface Methodology (RSM) was applied to identify the positive and negative effects of each factor. The experimental Box Behnken Design (BBD) was also utilized to select the number of required experiments for the present research as represented in [14, 15]. The effect of cetane number was studied. All tests were carried out based on the experimental design that is derived from the Minitab16 program (version 2018) as in Table III, where the estimated exact cetane number and the quantities of the mixed additives are stated. Flash point, viscosity, and the refractive index were measured. The solutions were taken for each run according to the run order and levels as illustrated in Table III and Table I. The mixtures were then put aside for 24 hours in order to assure solubility. In each run, before testing, the mixtures were shaken well for total miscibility. The readings were taken from the octane meter and flash point temperature meter. All runs were performed at ambient temperature and on the same day to avoid any environmental effect on the outcome.

TABLE III. CENTRAL COMPOSITE DESIGN MATRIX -BBD.

No. of run order	Nitro benzene (PPM)	4-nitro toluene (PPM)	M-nitro phenol (PPM)
1	25	0	0
2	25	0	50
3	25	50	0
4	25	50	50
5	0	25	0
6	0	25	50
7	50	25	0
8	50	25	50
9	0	0	25
10	0	50	25
11	50	0	25
12	50	50	25
13	0	0	0

III. RESULTS AND DISCUSSION

Statistical analysis was employed to determine the effect of each additive and their interactions in order to indicate the best model. All models were studied. During the present study, the linear model showed the best agreement among other models and regression equations as in (1) and (2) were obtained. The regression equation shows that all parameters have a positive effect on the cetane number. However, the effect is rational on octane number and flash point which means there is a negative effect from some additives. The effect of each parameter is determined by a factor in regression equations.

A. Effect of the Added Chemicals on the Cetane Number

Figure 2 demonstrated the relationship between the contour plot of the determined additives and cetane number, while Figure 3 showed the surface plot of the obtained additives and cetane number. The effect of the additives m-nitrophenol, 4-

nitro toluene, and nitrobenzene on cetane number was positive. The greatest estimated impact was recorded from nitro phenol followed by 4-nitro toluene, and nitrobenzene by 0.3545, 0.0054, and 0.0035 respectively as shown in regression equation (1). The correlation between the parameters is very strong and equal to 98.92%. Table IV contains the coefficients and ANOVA parameters obtained from the current experiment, and Table V demonstrates the cetane numbers results from the tested samples.

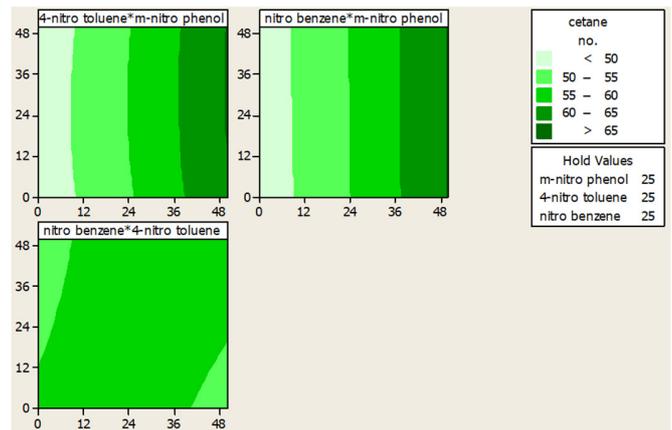


Fig. 2. Contour plot for the determination of additives effect on cetane number.

TABLE IV. COEFFICIENTS AND ANOVA PARAMETERS

Term	Coefficient	SE coefficient	T	F	P
Constant	46.6417	0.534102	87.3273		0.000
Regression				336.62	0.000
m-nitro phenol	0.3545	0.011157	31.7737	1009.57	0.000
4-nitro toluene	0.0050	0.011157	0.4481	0.20	0.666
Nitro benzene	0.0035	0.011157	0.3137	0.10	0.762
Model summary	S = 0.788921, R-Sq = 99.21%, R-Sq(adj) = 98.92%, PRESS = 11.2031, R-Sq(pred) = 98.23%				

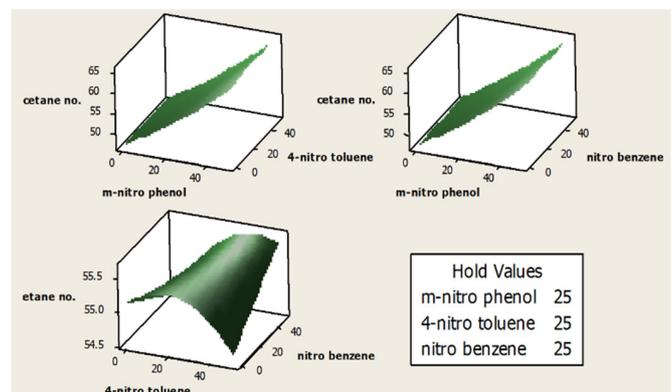


Fig. 3. Surface plot for the determination of additives effect on cetane number.

According to the test, the correlation equation between the cetane number and the additive becomes:

$$\text{Cetan number} = 46.6417 + 0.3545 m - \text{nitro phenol} + 0.005 4 \text{ nitro toluene} + 0.0035 \text{ nitro benzene} \quad (3)$$

TABLE V. CETANE NUMBER RESULTS

No. of run order	Nitro benzene (PPM)	4-nitro toluene (PPM)	M-nitro phenol (PPM)	Cetane number
1	25	0	0	46.9
2	25	0	50	64.1
3	25	50	0	47.3
4	25	50	50	65.2
5	0	25	0	47.3
6	0	25	50	65.3
7	50	25	0	47.4
8	50	25	50	65.2
9	0	0	25	55.3
10	0	50	25	54.3
11	50	0	25	54.9
12	50	50	25	55.4
13	0	0	0	46.9

B. Effect of the Added Chemicals on Flash Point

Figures 4 and 5 exhibit the contour and surface plots on the flash point obtained from the present work. The effect of the additives on flash point temperature has been investigated. The results show that the additives present a negative effect on flash point temperature. This means that when the concentration of additives is increasing, the flash point temperature decreases by factors of 0.135, 0.005, and 0.06 respectively as shown in regression equation (3). The resulted correlation was about 45.41%. Table VI illustrates the coefficients and ANOVA parameters estimated from the runs, whereas Table VII consists of the results of all tests with their run order. It is worth mentioning that the mixture flash point reaches the requirement of the international standards which is about 51-60 [17].

The equation between the flashpoint and the additive becomes:

$$\text{Flash point} = 90.25 - 0.135 m - \text{nitro phenol} - 0.005 4 \text{ nitro toluene} - 0.06 \text{ nitro benzene} \quad (4)$$

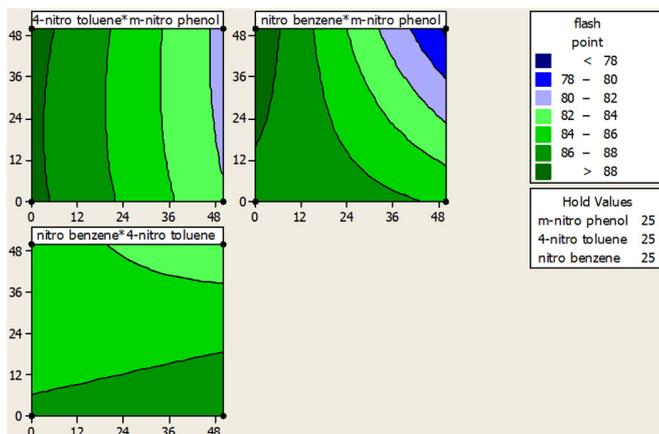


Fig. 4. Contour plot for the determination of additives effect on cetane number.

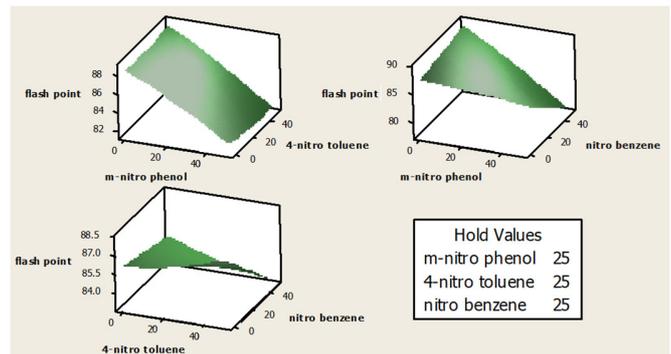


Fig. 5. Surface plot for the determination of additives effect on flash point temperature.

TABLE VI. COEFFICIENTS AND ANOVA PARAMETERS

Term	Coefficient	SE coefficient	T	F	P
Constant	90.250	1.41605	63.7335		0.000
Regression				8.3238	0.007655
m-nitro phenol	-0.135	0.02958	-4.5638	20.8286	0.002
4-nitro toluene	-0.005	0.02958	-0.1690	0.0286	0.870
Nitro benzene	-0.060	0.02958	-2.0284	4.1143	0.077
Model summary	S = 2.09165 R-Sq = 75.74% R-Sq(adj) = 66.64% PRESS = 78.75 R-Sq(pred) = 45.41%				

TABLE VII. FLASH POINT TEMPERATURE RESULTS

No. of run order	Flash point	Nitro benzene (PPM)	4-nitro toluene (PPM)	M-nitro phenol (PPM)
1	89	25	0	0
2	83	25	0	50
3	88	25	50	0
4	81	25	50	50
5	88	0	25	0
6	86	0	25	50
7	89	50	25	0
8	77	50	25	50
9	85	0	0	25
10	88	0	50	25
11	85	50	0	25
12	84	50	50	25
13	91	0	0	0

C. Effect of the Added Chemicals on Viscosity

The effect of the selected additives are exhibited in Figure 6. All additives had a negative effect on viscosity. This means that when the concentration of additives increased, viscosity decreased by the factors of -0.000080, -0.00044, and -0.00008 as shown in (5). Nevertheless, the effect is very weak. Table VIII shows the coefficients and ANOVA parameters obtained from the tests, while Table IX states the results of the viscosity values. The regression equation is:

$$\text{Viscosity} = 3.15 - 0.00008 m - \text{nitro phenol} - 0.00044 4 \text{ nitro toluene} - 0.00008 \text{ nitro benzene} \quad (5)$$

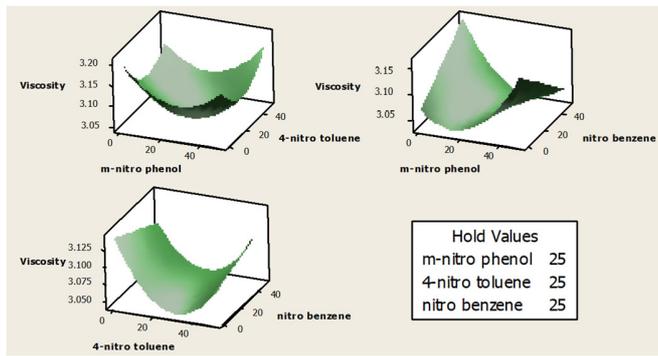


Fig. 6. Surface plot for the determination of additives effect on kinematic viscosity.

TABLE VIII. COEFFICIENTS AND ANOVA PARAMETERS

Term	Coefficient	SE coefficient	T	F	P
Constant	3.14867	0.04942	63.71		0.000
Regression				0.06	0.977
m-nitro phenol	-0.000080	0.001032	-0.08	20.8286	0.940
4-nitro toluene	-0.000440	0.001032	-0.43	0.0286	0.681
Nitro benzene	-0.000080	0.001032	-0.08	4.1143	0.940
Model summary	S = 0.0729989 R-Sq = 2.4% R-Sq(adj) = 0.0%				

TABLE IX. KINEMATIC VISCOSITY RESULTS

No. of run order	Kinematic Viscosity Cts/s	Nitro benzene (PPM)	4-nitro toluene (PPM)	M-nitro phenol (PPM)
1	3.256	25	0	0
2	3.172	25	0	50
3	3.148	25	50	0
4	3.132	25	50	50
5	3.08	0	25	0
6	3.22	0	25	50
7	3.104	50	25	0
8	3.048	50	25	50
9	3.08	0	0	25
10	3.076	0	50	25
11	3.112	50	0	25
12	3.176	50	50	25
13	3.182	0	0	0

D. Effect of the Added Chemicals on Refractive Index

The influence of the chosen additives is shown in Figure 7. M-nitrophenol, 4-nitro toluene, and nitrobenzene were studied with regard on their effect on the refractive index. The additives presented a very low impact on the refractive index which means their effect is negligible as shown in regression equation (6). Table X demonstrates the coefficients and ANOVA values obtained from the runs, and Table XI shows the summed values of the tests with their run order. The regression equation is:

$$\text{Refractive index} = 1.46 - 0.000013 m - \text{nitro phenol} - 0.0000224 \text{ nitro toluene} + 0.000007 \text{ nitro benzene} \quad (6)$$

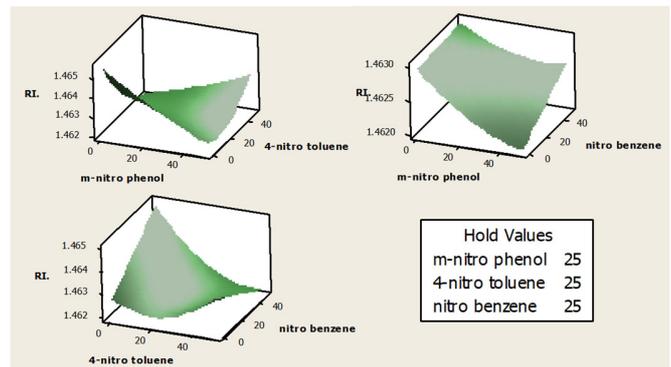


Fig. 7. Surface plot for the determination of additives effect on refractive index.

TABLE X. COEFFICIENTS AND ANOVA PARAMETERS

Term	Coefficient	SE coefficient	T	F	P
Constant	3.14867	0.04942	63.71	----	0.000
Regression				0.06	0.977
m-nitro phenol	-0.000080	0.001032	-0.08	20.8286	0.940
4-nitro toluene	-0.000440	0.001032	-0.43	0.0286	0.681
Nitro benzene	-0.000080	0.001032	-0.08	4.1143	0.940
Model summary	S = 0.0729989 R-Sq = 2.4% R-Sq(adj) = 0.0%				

TABLE XI. REFRACTIVE INDEX RESULTS

No. of run order	Refractive index	Nitro benzene (PPM)	4-nitro toluene (PPM)	M-nitro phenol (PPM)
1	1.4623	25	0	0
2	1.4624	25	0	50
3	1.4621	25	50	0
4	1.4631	25	50	50
5	1.4632	0	25	0
6	1.463	0	25	50
7	1.462	50	25	0
8	1.4624	50	25	50
9	1.4618	0	0	25
10	1.4632	0	50	25
11	1.4653	50	0	25
12	1.4629	50	50	25
13	1.4653	0	0	0

IV. CONCLUSION

The current statistical and experimental analyses have been performed to study the effect of three aromatic compounds as diesel cetane number improvers and their impact on flash point, refractive index, and viscosity of Tawke oil well diesel. Sample 6 gave the best result. To the best of our knowledge, no similar study has been conducted with the certain components to contrast the present outcome, however, many investigations to improve the cetane number of fuel have been performed including the transesterification and higher alcohol-diesel blends [17], the oxygenated compound for cetane number improvement di-n-pentyl-ether (DNPE) [18], and the ethanol with 2-ethyl hexyl nitrate (EHN) component [3].

Mixtures of additives were added to the diesel samples and were measured regarding the cetane number improvement and

flash point. Three different organic chemicals were selected for the current research, namely m-nitrophenol, 4-nitro toluene, and nitrobenzene. Statistical analysis was performed to determine the effect of each component and to calculate the regression equation. Besides, the experimental design was utilized to calculate the most active mixture composition that presents the best performance and attains the international standards. The best mixture composition turn out to be the additive consisting of run number 6 that consisted of no nitrobenzene, 25PPM 4-nitro toluene, and 50PPM m-nitrophenol. Viscosity and refractive index were measured as well and their effect was significant.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Higher Education and Scientific Research in Kurdistan, the Chemistry Department of the University of Zakho, and the Petroleum Department, Polytechnic University of Duhok/Zakho Institute for their support. Also, the authors declare that no competing interests exist regarding the current work.

REFERENCES

- [1] P. E. Oguntunde, O. O. Ojo, O. A. Oguntunde, and H. I. Okagbue, "Crude Oil Importation and Exportation in Nigeria: An Exploratory and Comparative Study," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3329–3331, Oct. 2018, <https://doi.org/10.48084/etasr.2172>.
- [2] M. M. Khudhair, S. A. Husain, S. M. Salih, and Z. M. Jassim, "Preparation of Cetane Improver for Diesel Fuel and Study It's Performance," *Journal of Al-Nahrain University-Science*, vol. 20, no. 3, pp. 42–50, 2017, <https://doi.org/10.22401/jnus.20.1.06>.
- [3] M. Semakula and F. Inambao, "Cetane Improvers and Ethanol Performance and Emissions Characteristics Using Pyrorated Biodiesel," Jun. 2018, <https://doi.org/10.20944/preprints201806.0106.v1>.
- [4] E. G. Giakoumis and C. K. Sarakatsanis, "A comparative assessment of biodiesel cetane number predictive correlations based on fatty acid composition," *Energies*, vol. 12, no. 3, pp. 1–30, 2019, <https://doi.org/10.3390/en12030422>.
- [5] A. S. Hamadi, "Selective Additives for Improvement of Gasoline Octane Number," *Tikrit Journal of Engineering Sciences*, vol. 17, no. 2, pp. 22–35, 2010.
- [6] A. Oseev, M. Zubtsov, and R. Lucklum, "Octane number determination of gasoline with a phononic crystal sensor," *Procedia Engineering*, vol. 47, pp. 1382–1385, 2012, <https://doi.org/10.1016/j.proeng.2012.09.414>.
- [7] R. Stradling, J. Williams, H. Hamje, and D. Rickeard, "Effect of Octane on Performance, Energy Consumption and Emissions of Two Euro 4 Passenger Cars," *Transportation Research Procedia*, vol. 14, pp. 3159–3168, 2016, <https://doi.org/10.1016/j.trpro.2016.05.256>.
- [8] A. Álvarez, M. Lapuerta, and J. R. Agudelo, "Prediction of Flash-Point Temperature of Alcohol/Biodiesel/Diesel Fuel Blends," *Industrial and Engineering Chemistry Research*, vol. 58, no. 16, pp. 6860–6869, 2019, <https://doi.org/10.1021/acs.iecr.9b00843>.
- [9] M. M. Alrefaai, G. D. J. Guerrero Peña, A. Raj, S. Stephen, T. Anjana, and A. Dindi, "Impact of dicyclopentadiene addition to diesel on cetane number, sooting propensity, and soot characteristics," *Fuel*, vol. 216, pp. 110–120, Mar. 2018, <https://doi.org/10.1016/j.fuel.2017.11.145>.
- [10] S. Naman, M. Ali, and S. Simo, "Evaluation and Improvement of Diesel Cut from Tawke Crude Oil Wells, Zakho," *Science Journal of University of Zakho*, vol. 5, no. 1, pp. 93–100, 2017, <https://doi.org/10.25271/2017.5.1.308>.
- [11] S. A. Idrees and M. Khalil Ibrahim, "Optimization of Congo-Red Photocatalytic Degradation by Central Composite Design," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, Duhok, Iraq, Oct. 2018, pp. 389–393, <https://doi.org/10.1109/ICOASE.2018.8548919>.
- [12] S. L. C. Ferreira *et al.*, "Statistical designs and response surface techniques for the optimization of chromatographic systems," *Journal of Chromatography A*, vol. 1158, no. 1, pp. 2–14, Jul. 2007, <https://doi.org/10.1016/j.chroma.2007.03.051>.
- [13] I. M. Kareem, L. A. Abdulkareem, and H. I. Al-Barudi, "Surface and Deep Soil 222Rn Gas Exhalation Comparison: A Case Study in Tawke, Duhok, Northern Iraq," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4741–4744, Oct. 2019, <https://doi.org/10.48084/etasr.3086>.
- [14] M. Vaez, A. Zarringhalam Moghaddam, and S. Alijani, "Optimization and Modeling of Photocatalytic Degradation of Azo Dye Using a Response Surface Methodology (RSM) Based on the Central Composite Design with Immobilized Titania Nanoparticles," *Industrial & Engineering Chemistry Research*, vol. 51, no. 11, pp. 4199–4207, Mar. 2012, <https://doi.org/10.1021/ie202809w>.
- [15] V.-M. T. Taavitsainen, "Experimental Optimization and Response Surfaces," in *Chemometrics in Practical Applications*, London, UK: IntechOpen, 2012.
- [16] P. Calza *et al.*, "Multivariate experimental design for the photocatalytic degradation of imipramine. Determination of the reaction pathway and identification of intermediate products," *Applied Catalysis B: Environmental*, vol. 84, no. 3–4, pp. 379–388, 2008, <https://doi.org/10.1016/j.apcatb.2008.04.015>.
- [17] F. Lujaji, A. Bereczky, L. Janosi, C. Novak, and M. Mbarawa, "Cetane number and thermal properties of vegetable oil, biodiesel, 1-butanol and diesel blends," *Journal of Thermal Analysis and Calorimetry*, vol. 102, no. 3, pp. 1175–1181, 2010, <https://doi.org/10.1007/s10973-010-0733-9>.
- [18] M. Marchionna, R. Patrini, D. Sanfilippo, A. Paggini, F. Giavazzi, and L. Pellegrini, "From natural gas to oxygenates for cleaner diesel fuels," in *Studies in Surface Science and Catalysis*, vol. 136, Paris, France: Elsevier, 2001, pp. 489–494.

Day-ahead Coordinated Operation of a Wind-Storage System Considering Wind Forecast Uncertainty

Nhi Thi Ai Nguyen

Faculty of Electrical Engineering
The University of Danang - University of Science and
Technology
Danang, Vietnam
ntanhi@dut.udn.vn

Dinh Duong Le

Faculty of Electrical Engineering
The University of Danang - University of Science and
Technology
Danang, Vietnam
ldduong@dut.udn.vn

Abstract-This paper proposes an optimal operation for coordinated Battery Energy Storage (BES) and wind generation in a day-ahead market under wind uncertainty. A comprehensive AC Optimal Power Flow (AC OPF) model was established to incorporate wind and storage into a power system. To take into account wind forecast uncertainty, preprocessing technique, time series model, and fast forward selection method were applied for scenario generation and reduction processes. Tests were performed on a modified IEEE 14-bus system and the results show that the use of BESs is an alternative to guarantee a more efficient and flexible operation of wind power plants.

Keywords-AC OPF; Battery Energy Storage (BES); operation; wind; uncertainty

I. INTRODUCTION

Wind energy has gained considerable interest globally during the last decades due to the increased environment concern. The overall capacity of wind turbines installed worldwide by the end of 2019 reached 650.8GW, according to statistics by the World Wind Energy Association (WWEA). In 2019, 59,667MW were added, substantially more than the 50,252MW of 2018. The sum of the wind turbines installed by the end of 2019 can cover more than 6% of global electricity demand [1]. Unfortunately, wind power output is well documented for its variable and non-dispatchable nature. The output of a wind turbine depends on wind speed, which varies daily and seasonally. Thus, the high penetration level of wind generation brings new challenges to the power system. From the operational aspect, it is essential to find a solution regarding the provision of controllability to wind generation, improving power quality of grid-connected wind farms and hence, facilitating higher wind installation capacity. Energy storage systems (ESSs) can be an economically favorable option to mitigate the variability of renewable generation. ESSs can be used by system operators for reliability improvement, ancillary services, and transmission congestion relief [2-4]. For wind power producers, ESSs are utilized in capacity firming and energy time shifting [5]. In electricity markets, storage systems can take advantage of price difference and gain profits for wind power plant using energy arbitrage. Wind power output is usually high during night-time and low at daytime, whereas electricity demand and price are usually low at night and get

higher during daytime. If wind energy is stored during low-price periods and is discharged back to the supply load at high-price periods, higher profit can be obtained. Energy storage can benefit by providing other services to the grid, but energy arbitrage is by far the largest business opportunity. The co-operation strategy of a wind-storage system is substantially important in achieving optimal tradeoff between operation cost and profit. This operation problem is challenging due to the stochastic behavior of wind power output.

The issue of coupling operation of wind and storage has been extensively studied with regard to operation and planning aspects [6-13]. Authors in [9] presented a dynamic programming algorithm employed to determine optimal energy exchange with the market for a specified scheduling period, taking into account transmission constraints. In [10], research was conducted to analyze the possibility of coupling wind and storage systems for time-shifting application, reducing grid congestions and making renewables more controllable on grid operator side. Authors in [11] investigate the problem of planning and operation of a combined wind-storage system. Specifically, a procedure is proposed, aiming to determine an hourly operation schedule of the combined system. In [12], a joint planning problem is established for transmission congestion and wind curtailment, including wind power installed capacity and location, transmission network expansion, and ESS locating and sizing. The problem was formulated into a linearized MILP model. An optimized output strategy of wind storage system is formulated in [13]. The storage is shown to effectively improve the efficiency of the wind farm while reducing wind power output fluctuations. There are many publications discussing the incorporation of an energy storage system in day-ahead market. In [14], a deterministic centralized unit commitment problem is proposed to optimally schedule storage operation in power systems with high wind penetration. Authors in [15] formulate an optimal joint bidding problem of wind hydro system under deterministic scenarios of wind generation, and illustrate its solution with a simple three-reservoir cascade. In addition to those deterministic formulations, wind uncertainty has also been dealt with, applying scenario-based stochastic programming [16-20]. Authors in [16] present a stochastic unit commitment model with energy storage and obtained a reduced

Corresponding author: Dinh Duong Le

system constraint violation in the N-1 analysis. Authors in [17] adopt a scenario-based stochastic unit commitment model to address uncertainties of renewables and demands in evaluating the reserves provided by ESSs and generators. The problem is solved using the progressive hedging algorithm with heuristic approaches. In [18], a stochastic unit commitment model with ESS and wind for system scheduling at day-ahead and real-time stages is proposed. A stochastic unit commitment model is presented in [19] for optimal energy and reserve bids in systems with high wind penetration. In [20], a stochastic battery arbitrage model for day-ahead and real-time prices is proposed. Authors in [21] formulated a robust unit commitment model with pumped hydro storage units. This approach models the randomness using an uncertainty set to protect the system against worst- case scenarios.

Most previous studies explored the value of storage systems with wind generation in day-ahead market, either using deterministic models [11, 14, 15, 22-23] or focusing on price uncertainties [24-26]. However, with the increasing level of wind energy integration, the variable nature of wind should be examined in any optimization model. Therefore, in this paper, we establish a day-ahead optimal dispatch model of wind-storage system while taking into account wind power uncertainties. This is implemented by formulating an AC OPF model with BES and wind integration and applying wind uncertainty modeling techniques, including preprocessing technique, time series model, and fast forward selection method. The model focuses on the use of BES to time-shift wind energy to higher price periods, i.e. arbitrage, which is a particularly important service that energy storage systems can provide in day-ahead energy markets. This model is an effective tool for wind farm operators to determine the optimal day-ahead operating strategies of wind-storage systems under wind uncertainties. Extensive tests were performed on a modified IEEE 14-bus system.

II. WIND UNCERTAINTY MODELING

In this section an approach to capture wind uncertainty is presented. Combined preprocessing technique and time series analysis are used to build a model representing the uncertainty associated with wind speed. Taken as an example, three-year measured hourly wind speed values of a real wind farm (rated at 85MW) in Italy, are used. Wind speed is usually not stationary with distinct diurnal and seasonal patterns [27] while a time series model such as Auto-Regressive Moving Average (ARMA) requires stationary data. To handle this issue, we adopt preprocessing techniques [27]. After that we make use of the process described in detail in [28] to build the ARMA model for the obtained stationary data. By sampling from the resulting time series model, the set of 10,000 hourly wind speed scenarios for day-ahead operation is generated as in Figure 1. Equal probability is assigned to each scenario of the set. To capture the uncertainty, a large number of scenarios should be generated, leading to increased complexity and computational burden. In this research, the fast forward selection approach [29] was used to select a limited number of representative scenarios for the set. Figure 2 depicts 10 selected wind speed scenarios. For determining an optimal operation for coordinated BES and wind generation in a day-ahead market,

wind power data are necessary. For this purpose, an aggregate power curve for the entire wind farm is needed to map wind speed scenarios into wind power scenarios. In this paper, we make use of the method of bins [27, 30, 31] using wind power-wind speed pairs at the wind farm to obtain the aggregate power curve as in Figure 3. Using the power curve, 10 representative wind power scenarios are obtained from 10 representative wind speed scenarios as shown in Figure 4.

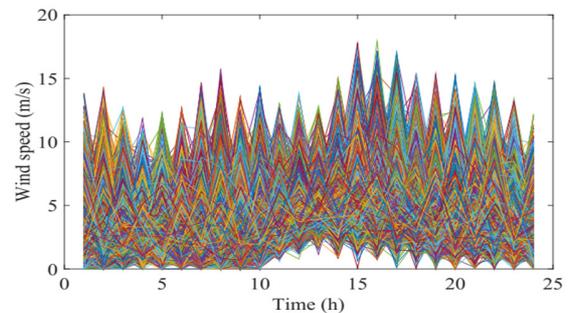


Fig. 1. Wind speed scenarios.

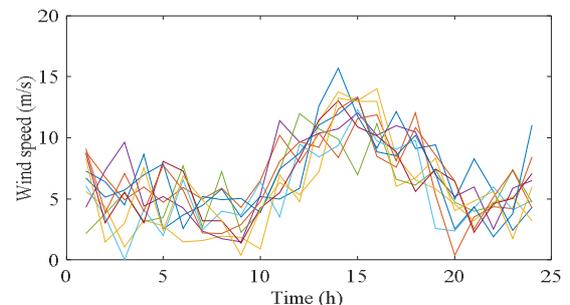


Fig. 2. Representative wind speed scenarios.

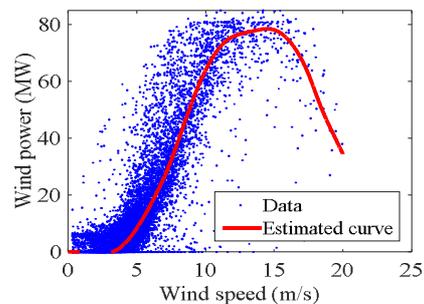


Fig. 3. Estimated power curve.

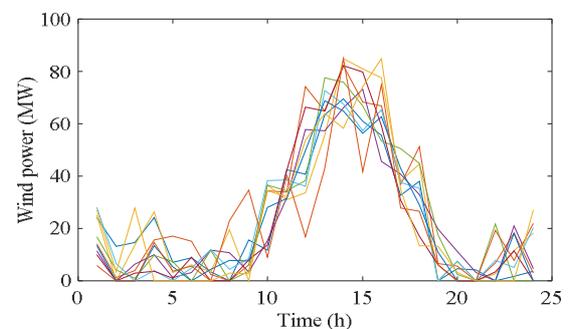


Fig. 4. Representative wind power scenarios.

III. THE AC OPF MODEL

A. Objective Function

The problem is formulated into an AC OPF model with an optimization goal of minimizing total system operating cost, which includes generating cost of all generating units and BES operation cost.

$$\text{Min } \sum_{t=1}^T \sum_{i=1}^{NG} (c_{0,i} + c_{1,i}P_{G,i}(t) + c_{2,i}[P_{G,i}(t)]^2) + \sum_{t=1}^T \sum_{j=1}^{NS} (c_{d,j}P_{d,j}(t) - c_{ch,j}P_{ch,j}(t)) \quad (1)$$

where $P_{G,i}(t)$ is real power generating at bus i and period t , $P_{ch,j}(t)$ is the charging power of BES at bus i and period t , $P_{d,j}(t)$ is the discharging power of BES at bus i and period t , $c_{0,i}$; $c_{1,i}$; $c_{2,i}$ are the cost coefficients of generating unit at bus i , $c_{ch,j}$; $c_{d,j}$ are the cost coefficients for charging and discharging power of BES at bus j , NG is the number of generating units, NS the number of BESSs, and T is the optimization time horizon.

B. System Constraints

The objective function must fulfill the following network constraints:

- Power balance:

$$P_{G,i}(t) - P_{L,i}(t) + P_{d,i}(t) - P_{ch,i}(t) = V_i(t) \sum_{k=1}^{NB} V_k(t) \{G_{ik} \cos[\theta_i(t) - \theta_k(t)] + B_{ik} \sin[\theta_i(t) - \theta_k(t)]\} \quad (2)$$

$$Q_{G,i}(t) - Q_{L,i}(t) + Q_{d,i}(t) - Q_{ch,i}(t) = V_i(t) \sum_{k=1}^{NB} V_k(t) \{G_{ik} \sin[\theta_i(t) - \theta_k(t)] + B_{ik} \cos[\theta_i(t) - \theta_k(t)]\} \quad (3)$$

where $Q_{G,i}(t)$ is the reactive power generating at bus i and period t , $P_{L,i}(t)$ is the real load power at bus i and period t , $Q_{L,i}(t)$ is the reactive load power at bus i and period t , $Q_{d,i}(t)$ is the reactive discharging power of BES bus i and period t , $Q_{ch,i}(t)$ is the reactive charging power of BES at bus i and period t , $V_i(t)$ is the magnitude of voltage at bus i and period t , $V_k(t)$ the magnitude of voltage at bus k and period t , $\theta_i(t)$ the angle of voltage at bus i and period t , $\theta_k(t)$ the angle of voltage at bus k , period t , G_{ik} the line conductance of branch ik , B_{ik} the line susceptance of branch ik , and NB the total system bus number.

- Voltage magnitude limits:

$$V_{i,min} \leq V_i(t) \leq V_{i,max} \quad (4)$$

where $V_{i,min}$ is the lower limit and $V_{i,max}$ is the upper limit for voltage bus i .

- Generator power limits:

$$P_{G,i-min} \leq P_{G,i}(t) \leq P_{G,i-max} \quad (5)$$

$$Q_{G,i-min} \leq Q_{G,i}(t) \leq Q_{G,i-max} \quad (6)$$

where $P_{G,i-min}$, $P_{G,i-max}$ are the active and $Q_{G,i-min}$, $Q_{G,i-max}$ the reactive generating power limits.

- Branch current limits:

$$I_{ij}(t) \leq I_{ij,max} \quad (7)$$

$$I_{ji}(t) \leq I_{ji,max} \quad (8)$$

where $I_{ij}(t)$ is the magnitude of current flowing from i to j , in period t , $I_{ji}(t)$ the magnitude of current flowing from bus j to bus i , in period t , and $I_{ij,max}$, $I_{ji,max}$ the current limits.

C. BES Constraints

- Energy balance:

$$B_i(t) = B_{i-1}(t) + (\eta_{ch}P_{ch,i}(t) - P_{d,i}(t)/\eta_d)\Delta t \quad (9)$$

where $B_i(t)$ is the energy level of BES at bus i in period t , $B_{i-1}(t)$ the energy level of BES at bus i in period $t - 1$, η_{ch} the charging efficiency, η_d the discharging efficiency, and Δt the time interval between two consecutive periods.

- Charging/discharging power limits:

$$P_{d,i-min} \leq P_{d,i}(t) \leq R_{i,max} \quad (10)$$

$$P_{ch,i-min} \leq P_{ch,i}(t) \leq R_{i,max} \quad (11)$$

where $R_{i,max}$ is the power rating of BES at bus i and $P_{ch,i-min}$, $P_{d,i-min}$ the minimum charging/discharging power of BES at bus i .

- ESS energy limits:

$$B_{i,min} \leq B_i(t) \leq B_{i,max} \quad (12)$$

where $B_{i,min}$ is the minimum energy limit of BES at bus i and $B_{i,max}$ the energy rating of BES at bus i .

IV. THE PROPOSED APPROACH

The overall procedure for the proposed approach is described in Figure 5, where N_s is the number of representative wind power scenarios. According to this procedure, wind data are first preprocessed to create wind power scenarios. These wind scenarios and load data will be the input of the AC OPF model which optimizes the operation of wind-storage systems. The model is run for all wind scenarios and the obtained results will be the optimal dispatching schedules for the systems.

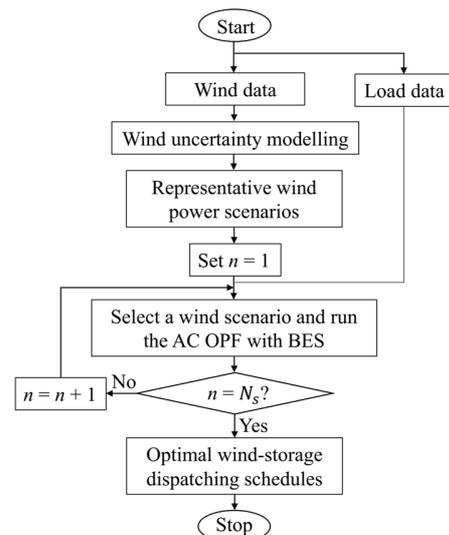


Fig. 5. Overall procedure of the proposed approach.

V. TESTS AND RESULTS

In this section, a case study is carried out on a modified IEEE 14-bus network (Figure 6), with a wind-storage system located at bus 6. The BES is used for energy arbitrage. It is charged from both wind and conventional sources during low price periods and discharged at peak price hours. The wind farm has an installed capacity of 85MW, accounting for 40% wind penetration level. There are 4 conventional generators, at buses 1, 2, 3, and 8, with total capacity of 292.4MW. The typical load per day, with peak value of 212MW, is calculated through statistical average data. Wind scenarios are generated as described in Section II. The optimization horizon is 24h. Simulations are performed for both deterministic and scenario-based cases, using Matlab 2016a, on a PC with Intel Core i7 – 3.4GHz CPU and 8.0GB of memory. First, a base case is run with BES parameters as shown in Table I and wind penetration level of 40% to obtain optimal day-ahead operating schedules of the wind-storage system. Then, in order to show the effect of wind uncertainties on optimization results, sensitivity analysis is performed on BES capacities with different wind penetration levels in both deterministic and scenario-based cases. The energy arbitrage profits of the system are also calculated for all wind penetration levels.

TABLE I. BES PARAMETERS

R_{max} [MW]	B_{max} [MWh]	η_{ch}	η_d
80	200	0.85	0.85

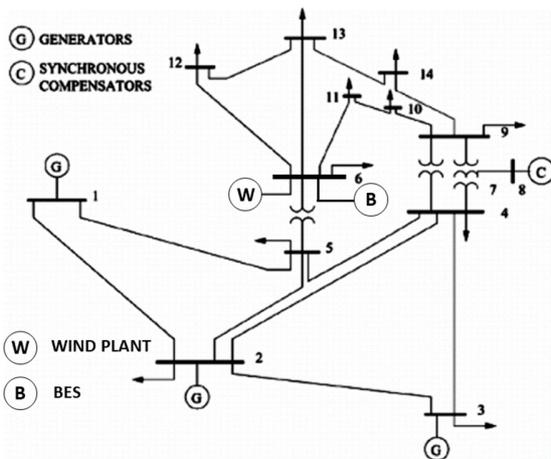


Fig. 6. The modified IEEE 14-bus system.

As a result, the optimal output of wind - storage system in a typical day can be observed in Figure 7. As can be seen in this Figure, the available wind generation in a single day is negatively correlated to the market price. However, with the use of BES, the actual wind-storage system output is made to positively follow market prices. Accordingly, less power is generated during low-price hours (from 12 to 16), instead, more power is generated at peak-price periods (hours 7, 8 and 18 to 21). Figure 8 presents the storage system operation. As can be seen, the BES has a large possibility of charging during valley hours to store more power and transfer that to peak hours for energy arbitrage. This operation effectively follows system market prices.

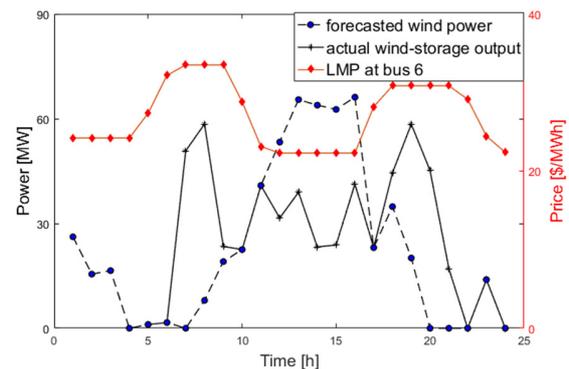


Fig. 7. The output of wind-storage system in a typical day.

The optimal day-ahead power dispatches of BES in deterministic and scenario-based cases are shown in Figure 9. This figure shows that BES operation follows the market price trend in all wind scenarios, thus effectively supports wind generation.

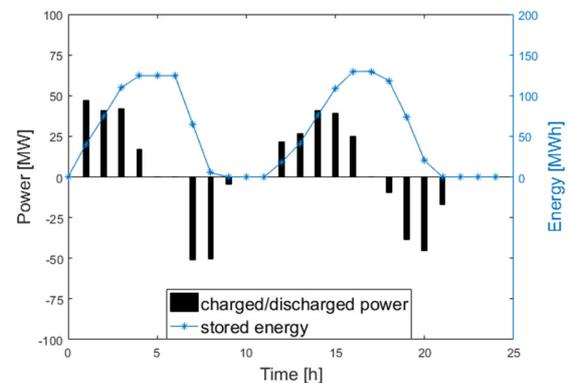


Fig. 8. BES operation.

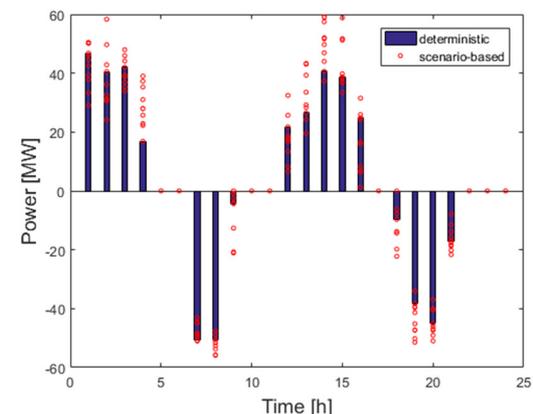


Fig. 9. BES optimal day-ahead power dispatch.

In order to see the impact of wind uncertainty on storage capacities, sensitivity analysis was performed. Wind penetration level varied from 20% to 50% and the rest was kept to the same level as in the base case. Deterministic and scenario-based models were run. The simulation results are shown in Figure 10. It is seen that wind uncertainty has a high impact on BES capacities. For example, at 40% wind

penetration level, a BES capacity of approximately 45MW and 130MWh is required for the deterministic case while a BES capacity of about 59MW and 150MWh is found in the scenario-based case. Besides, wind penetration level leads to increasing BES capacities. Table II shows the energy arbitrage profit of wind-storage system at different wind penetration levels, with and without BES. This Table shows the obvious benefit of using energy storage systems along with a wind farm for energy arbitrage. In addition, considerably higher profits are gained for the wind-storage system with higher wind penetration levels.

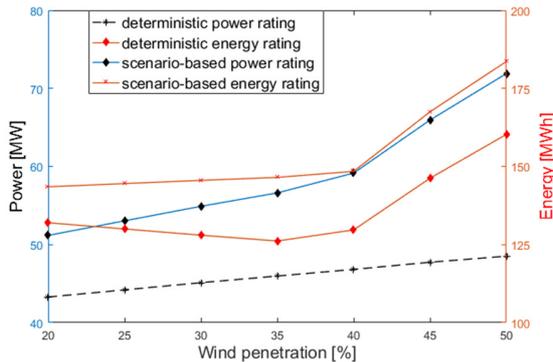


Fig. 10. BES capacities.

TABLE II. WIND-STORAGE SYSTEM ENERGY ARBITRAGE PROFIT [\$]

Wind penetration [%]	20	25	30	35	40	45	50
no BES	7.14 x 10 ³	8.75 x 10 ³	1.028 x 10 ⁴	1.056 x 10 ⁴	1.058 x 10 ⁴	1.028 x 10 ⁴	7.68 x 10 ³
with BES	7.24 x 10 ³	8.95 x 10 ³	1.064 x 10 ⁴	1.215 x 10 ⁴	1.362 x 10 ⁴	1.517 x 10 ⁴	1.644 x 10 ⁴

VI. CONCLUSIONS

In this paper, a model was developed to simulate the BES operation for energy arbitrage in power systems with high wind integration. The problem was formulated into an AC OPF model. Wind uncertainty was taken into account by applying generation and reduction techniques. Tests were carried out on a modified IEEE 14-bus system and day-ahead optimal operation schedule for wind-storage system was obtained. The simulation results show that the use of BES with wind generation can significantly improve profit for wind farm operators, especially at high wind penetration level. Furthermore, when considering wind uncertainty, higher capacities of BES are noted for the system as compared with deterministic cases, which means wind uncertainty should be examined in any optimization model with wind penetration.

ACKNOWLEDGMENT

This research was funded by the Ministry of Education and Training, Vietnam under project number B2020-DNA-02.

REFERENCES

- [1] "World wind capacity at 650,8 GW, Corona crisis will slow down markets in 2020, renewables to be core of economic stimulus programmes," *World Wind Energy Association*, Apr. 16, 2020. <https://wwindea.org/world-wind-capacity-at-650-gw/> (accessed May 04, 2021).
- [2] Y. Xu and C. Singh, "Adequacy and Economy Analysis of Distribution Systems Integrated With Electric Energy Storage and Renewable Energy Resources," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2332–2341, Nov. 2012, <https://doi.org/10.1109/TPWRS.2012.2186830>.
- [3] A. Oudalov, D. Chartouni, and C. Ohler, "Optimizing a Battery Energy Storage System for Primary Frequency Control," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1259–1266, Aug. 2007, <https://doi.org/10.1109/TPWRS.2007.901459>.
- [4] Y. Zhang, S. Zhu, and A. A. Chowdhury, "Reliability Modeling and Control Schemes of Composite Energy Storage and Wind Generation System With Adequate Transmission Upgrades," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 4, pp. 520–526, Oct. 2011, <https://doi.org/10.1109/TSTE.2011.2160663>.
- [5] J. M. Eyer, G. P. Corey, and J. Iannucci, "Energy storage benefits and market analysis handbook: a study for the DOE Energy Storage Systems Program," Sandia National Laboratories, SAND2004-6177, Dec. 2004, <https://doi.org/10.2172/920774>.
- [6] K. Alqunun, "Strength Pareto Evolutionary Algorithm for the Dynamic Economic Emission Dispatch Problem incorporating Wind Farms and Energy Storage Systems," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5668–5673, Jun. 2020, <https://doi.org/10.48084/etasr.3508>.
- [7] S. Gope, A. K. Goswami, and P. K. Tiwari, "Transmission Congestion Management using a Wind Integrated Compressed Air Energy Storage System," *Engineering, Technology & Applied Science Research*, vol. 7, no. 4, pp. 1746–1752, Aug. 2017, <https://doi.org/10.48084/etasr.1316>.
- [8] M. Korpaas, A. T. Holen, and R. Hildrum, "Operation and sizing of energy storage for wind power plants in a market system," *International Journal of Electrical Power & Energy Systems*, vol. 25, no. 8, pp. 599–606, Oct. 2003, [https://doi.org/10.1016/S0142-0615\(03\)00016-4](https://doi.org/10.1016/S0142-0615(03)00016-4).
- [9] P. Denholm and R. Sioshansi, "The value of compressed air energy storage with wind in transmission-constrained electric power systems," *Energy Policy*, vol. 37, no. 8, pp. 3149–3158, Aug. 2009, <https://doi.org/10.1016/j.enpol.2009.04.002>.
- [10] E. Casalini and S. Leva, "Feasibility analysis of storage systems in wind plants — an Italian application," in *IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe*, Milan, Italy, Jun. 2017, pp. 1–6, <https://doi.org/10.1109/EEEIC.2017.7977669>.
- [11] M. Dicorato, G. Forte, M. Pisani, and M. Trovato, "Planning and Operating Combined Wind-Storage System in Electricity Market," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 2, pp. 209–217, Apr. 2012, <https://doi.org/10.1109/TSTE.2011.2179953>.
- [12] X. Wu and Y. Jiang, "Source-Network-Storage Joint Planning Considering Energy Storage Systems and Wind Power Integration," *IEEE Access*, vol. 7, pp. 137330–137343, 2019, <https://doi.org/10.1109/ACCESS.2019.2942134>.
- [13] W. Ai et al., "Study on Transmission Planning of Combined Wind and Storage System," in *2nd IEEE Conference on Energy Internet and Energy System Integration*, Beijing, China, Oct. 2018, pp. 1–4, <https://doi.org/10.1109/EI2.2018.8582476>.
- [14] C. A. Silva Monroy and R. D. Christie, "Energy storage effects on day-ahead operation of power systems with high wind penetration," in *North American Power Symposium*, Boston, MA, USA, Aug. 2011, pp. 1–7, <https://doi.org/10.1109/NAPS.2011.6025193>.
- [15] A. Cerejo, S. J. P. S. Mariano, P. M. S. Carvalho, and M. R. A. Calado, "Hydro-wind Optimal Operation for Joint Bidding in Day-

- ahead Market: Storage Efficiency and Impact of Wind Forecasting Uncertainty," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 1, pp. 142–149, Jan. 2020, <https://doi.org/10.35833/MPCE.2018.000689>.
- [16] N. Li and K. W. Hedman, "Economic Assessment of Energy Storage in Systems With High Levels of Renewable Resources," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 1103–1111, Jul. 2015, <https://doi.org/10.1109/TSTE.2014.2329881>.
- [17] Z. Tang, Y. Liu, L. Wu, J. Liu, and H. Gao, "Reserve Model of Energy Storage in Day-Ahead Joint Energy and Reserve Markets: A Stochastic UC Solution," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 372–382, Jan. 2021, <https://doi.org/10.1109/TSG.2020.3009114>.
- [18] N. Li, C. Uckun, E. M. Constantinescu, J. R. Birge, K. W. Hedman, and A. Botterud, "Flexible Operation of Batteries in Power System Scheduling With Renewable Energy," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 685–696, Apr. 2016, <https://doi.org/10.1109/TSTE.2015.2497470>.
- [19] H. Akhavan-Hejazi and H. Mohsenian-Rad, "Optimal Operation of Independent Storage Systems in Energy and Reserve Markets With High Wind Penetration," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 1088–1097, Mar. 2014, <https://doi.org/10.1109/TSG.2013.2273800>.
- [20] H. Mohsenian-Rad, "Optimal Bidding, Scheduling, and Deployment of Battery Systems in California Day-Ahead Energy Market," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 442–453, Jan. 2016, <https://doi.org/10.1109/TPWRS.2015.2394355>.
- [21] R. Jiang, J. Wang, and Y. Guan, "Robust Unit Commitment With Wind Power and Pumped Storage Hydro," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 800–810, May 2012, <https://doi.org/10.1109/TPWRS.2011.2169817>.
- [22] A. Daggett, M. Qadrdan, and N. Jenkins, "Feasibility of a battery storage system for a renewable energy park operating with price arbitrage," in *IEEE PES Innovative Smart Grid Technologies Conference Europe*, Turin, Italy, Sep. 2017, pp. 1–6, <https://doi.org/10.1109/ISGTEurope.2017.8260249>.
- [23] Z. Shu and P. Jirutitijaroen, "Optimal Operation Strategy of Energy Storage System for Grid-Connected Wind Power Plants," *IEEE Transactions on Sustainable Energy*, vol. 1, no. 5, pp. 190–199, 2014, <https://doi.org/10.1109/TSTE.2013.2278406>.
- [24] D. Krishnamurthy, C. Uckun, Z. Zhou, P. R. Thimmapuram, and A. Botterud, "Energy Storage Arbitrage Under Day-Ahead and Real-Time Price Uncertainty," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 84–93, Jan. 2018, <https://doi.org/10.1109/TPWRS.2017.2685347>.
- [25] S. Vejdani and S. Grijalva, "The Value of Real-Time Energy Arbitrage with Energy Storage Systems," in *IEEE Power Energy Society General Meeting*, Portland, OR, USA, Aug. 2018, pp. 1–5, <https://doi.org/10.1109/PESGM.2018.8585767>.
- [26] X. Fang, B.-M. Hodge, L. Bai, H. Cui, and F. Li, "Mean-Variance Optimization-Based Energy Storage Scheduling Considering Day-Ahead and Real-Time LMP Uncertainties," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7292–7295, Nov. 2018, <https://doi.org/10.1109/TPWRS.2018.2852951>.
- [27] D. D. Le, G. Gross, and A. Berizzi, "Probabilistic Modeling of Multisite Wind Farm Production for Scenario-Based Applications," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 748–758, Jul. 2015, <https://doi.org/10.1109/TSTE.2015.2411252>.
- [28] G. E. P. Box and G. M. Jenkins, *Time series analysis, forecasting and control*. San Francisco, CA, USA: Holden-Day, 1976.
- [29] J. M. Morales, S. Pineda, A. J. Conejo, and M. Carrion, "Scenario Reduction for Futures Market Trading in Electricity Markets," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 878–888, May 2009, <https://doi.org/10.1109/TPWRS.2009.2016072>.
- [30] IEC–International Electrotechnical Commission, *Wind Turbines – Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines*, Geneva, Switzerland: IEC-International Electrotechnical Commission, IEC-61400-12, 2005.
- [31] Y. H. Wan, E. Ela, and K. Orwig, "Development of an Equivalent Wind Plant Power-Curve: Preprint," National Renewable Energy Lab, NREL/CP-550-48146, Jun. 2010. Accessed: May 04, 2021. [Online]. Available: <https://www.osti.gov/biblio/983731>.

Preparation of Polypropylene/Bentonite Composites of Enhanced Thermal and Mechanical Properties using L-leucine and Stearic Acid as Coupling Agents

Nadia Mohammedi

Department of Process Engineering
Ferhat Abbas Setif University 1
Sétif, Algeria
chimiebatna2016@gmail.com

Fouzia Zoukrami

Emerging Materials Research Unit
Ferhat Abbas Setif University 1
Sétif, Algeria
fzoukrami@univ-setif.dz

Nacereddine Haddaoui

Department of Process Engineering
Ferhat Abbas Setif University 1
Sétif, Algeria
n_haddaoui@univ-setif.dz

Abstract—The compatibilization of raw bentonite (bent) with a polymer matrix of polypropylene (PP) can improve the performance of the material in terms of thermal and mechanical properties. In this study, two kinds of untreated bentonite, bentonite-Maghnia (bent-m) and bentonite-Mostaganem (bent-M), that differ in the proportion of Al_2O_3 and in the particle size distribution were coupled to typical maleic anhydride grafted polypropylene PP-MA. Stearic Acid (SA) and L-leucine Amino Acid (AA) were selected as new coupling modifiers at a 5/5 ratio of bentonite/coupling agent. All PP/bent composites were prepared by melt mixing at 190°C. Morphological observation revealed a good dispersion of bentonite into the PP matrix in the presence of AA, SA, and PP-MA. Mechanical properties showed an increase in stiffness as bent-m or bent-M were associated with AA. For instance, PP/bent-m/AA composite underwent an improvement of about 13% in Young's modulus as compared to neat PP. On the other hand, the addition of SA into bent-m maintained stiffness and tensile strength at an acceptable level. An increase of around 40°C and 37% in the decomposition temperature and elongation at break was respectively observed for the PP/bent-m/SA composite. All coupled composites showed high degradation temperatures.

Keywords—bentonite; surfactant; coupling agent; mechanical properties; layered composite

I. INTRODUCTION

Recently, the development and characterization of polymer/clay nanocomposites has been the subject of increasing interest because these systems habitually show significant improvement in thermal, mechanical, and barrier properties when compared with virgin polymers or conventional micro or macro composites for the same amount of filler. Usually, the improvement is due to the high surface area of layered silicates which is more than $700\text{m}^2/\text{g}$, its aspect

ratio of about 50–200, in addition to the possible exfoliation or intercalation and good dispersion of clay into polymer matrix [1, 2]. At the melt state, the dispersion of clay particles into polyolefin matrix depends on several factors such as processing time, cationic surfactant modified clay, processing method, and interface modifiers. Recent studies have reported the effects of these factors on the properties and clay basal spacing of PP/EVA/nanoclay composites [3], PVC/bentonite organoclay system [4] and PP/clay composites [5, 6].

Polypropylene (PP) is selected in polymer composite or nanocomposite systems because it is widely used in engineering materials, electronic cases and interior decoration for its excellent insulation properties, low cost, ease of processing [7], and recyclability [8]. To enhance the thermal stability and other properties of PP, researchers have tried to use nanotechnology [9]. On the other hand, bentonite, which is a geological material, is often used as filler in polymer matrix. Bentonite has a structure that generally consists of two tetrahedral sheets (silicon/oxygen) separated by an octahedral sheet (aluminum/oxygen/hydroxyl) [10, 11]. The ions with positive charges on the clay surface can be adsorbed onto the bentonite structure due to the interaction between the negative and positive charges. Bentonite is used in diverse branches of industry for the fabrication of several ceramic products mainly due to its rheological properties. Also, it is used as a component in dyes, pharmaceuticals, paper [12–15], and in polymer composites [16]. The industrial and environmental applications of bentonite have showed an increase of interest in recent years [13]. For example, in [17], two different bentonite/PP composites were synthesized, with and without pore-enlarging treatment. These granular composite adsorbents were effective in eliminating Pb^{2+} in aqueous solutions.

Corresponding author: Nadia Mohammedi

Moreover, the preparation of PP/bentonite micro or nanocomposites is not straightforward due to the poor interaction between the non-polar polymers like PP and polar bentonite. Therefore, by adding PP-MA to the system as a coupling agent, the improvement in the interfacial adhesion is possible [8]. PP-MA is often used with a large number of polyolefin materials [12, 15]. Authors in [18] reported that for PP/modified montmorillonite (MMT) and PP/treated sodium bentonite, diethyl maleic grafted PP (PP-DEM) compatibilizer has lower polarity compared to maleic anhydride grafted polypropylene (PP-MA). They concluded that clay and matrix modification are the key factors and must be carefully and properly used to get the appropriate properties in the non-polar polymer based nanocomposites [19]. The ion exchange treatment of layered silicate is often influenced by the length of the alkyl chain, and the quality of multiple interactions [20, 21]. Authors in [22] indicated that the exfoliated nanocomposite based cloisite 15A compatibilized with acrylic acid grafted high density polyethylene (HDPE-AA) and maleic anhydride grafted high density polyethylene (HDPE-MA) displayed a higher level of clay dispersion. However, HDPE compatibilized with ethylene-acrylic acid copolymer (EAA) leads to poor dispersion due to its immiscibility with the HDPE matrix. Recently, authors in [23] confirmed that the composites containing modified bentonite in presence of stearic acid used as interface modifier induce a better dispersion of the filler, a decrease in viscosity and an increase in elongation at break of the composites based on PP matrix. Also, authors in [24] established that stearic acid favored the dispersion of PP/MMT nanocomposites when used as interface modifier and clay surface treatment. The samples exhibited crystallization temperatures comparable to pure PP, with a good dispersion. In [25], the authors synthesized L-leucine micro/nanocrystals and showed that these samples have a large potential in pharmaceutical applications as dispersibility improvers. This coating design concept can also be used for diverse active pharmaceutical ingredients.

In our study, two bentonites, bentonite-Maghnia (bent-m) and bentonite-Mostaganem (bent-M) from two different regions of Algeria were incorporated into a PP matrix. The traditional reactive maleic anhydride-grafted polypropylene PP-MA was used as a coupling agent to produce, via direct melt intercalation, PP/bent-m/PP-MA and PP/bent-M/PP-MA composites. Also, PP/bent-m/SA, PP/bent-M/SA, PP/bent-m/AA, and PP/bent-M/AA composites were obtained using stearic acid, SA, and L-leucine amino acid AA, respectively as coupling agents. Large utilization of PP-based layered composites is possible if a more effective or a less expensive compatibilizer is available. It is important to point out that stearic acid has a low cost and only small amounts are required. On the other hand, its hydrocarbon end is compatible with PP [24]. The purpose of this study is the use of raw bentonites, which have not previously been subjected to chemical treatments to improve their compatibility, as these treatments usually involve the use of chemicals difficult to eliminate [26]. Stearic acid and L-leucine have been incorporated directly in PP/bent composites at the melt state. The preparation, characterization, and mechanical properties of the prepared materials are reported here. Some formulations exhibited

significant improvements in the mechanical properties and degradation temperature when compared to pure PP. To the best of our knowledge, there are no previous reports of the use of this amino acid in improving the mechanical and thermal properties of PP/natural microbentonite composites obtained in molten state and the use of stearic acid as interface modifier for this system when it is introduced directly with polymer matrix and filler in internal mixer has not been reported.

II. EXPERIMENTAL PART

A. Materials

The PP used in this study (homopolymer Adstif HA740N) was supplied by Lyon dellBasell industries with Melt Flow Index (MFI) of 12g/10min and density of 0.9g/cm^3 . The clays used as reinforcement filler are natural bentonites from Algeria, supplied by BENTAL Society: bentonite of Maghnia (bent-m) with average diameter of $10.7\mu\text{m}$, $\%Al_2O_3$ of 17.79, and cation-exchange capacity (CEC) of 65 meq/100g [27], and bentonite of Mostaganem (bent-M) with average diameter of $24.8\mu\text{m}$, $\%Al_2O_3$ of 13.78, and CEC of 48meq/100g [28]. On the other hand, the ratio of SiO_2/Al_2O_3 in bent-m is 3.62 and in bent-M 4.49, which classifies it as Si-bentonite. Three different coupling agents were used, maleic anhydride-grafted polypropylene PP-MA, which was provided by Crompton-Uniroyal Chemical (Polybond3200), with MFI of 90-120g/10min and a density of 0.91g/cm^3 , stearic acid (SA), supplied by Henkel (Germany), and L-Leucine $C_6H_{13}NO_2$ interface amino acid (AA), purchased from MERCK. All the materials were used as received.

B. Elemental Analysis of Raw Bentonite

The elemental analysis of raw bentonite shows that the bent-m has higher alumina content ($Al_2O_3=17.79$) than bent-M ($Al_2O_3=13.78$), however the iron content of bent-M is more. The stoichiometric analysis results are:

bent-m (%mass): SiO_2 -64.44, Al_2O_3 -17.79, CaO-3.16, Fe_2O_3 -2.84, MgO-5.48, Na_2O -1.12, K_2O -1.27, SO_3 -0.02.

bent-M (%mass): SiO_2 -61.89, Al_2O_3 -13.78, CaO-7.06, Fe_2O_3 -3.66, MgO-3.12, Na_2O -0.92, K_2O -1.69, SO_3 -0.14.

C. Preparation of PP/Bentonite Composites

PP-MA pellets were pre-mixed in a bag with bentonite (bent-m/PP-MA) at 5:5 weight ratio as well as bent-m/AA, bent-M/AA, bent-m/SA, and bent-M/SA and then, with PP at 5% wt. All the compounds were dried in a vacuum oven for 24h at 80°C , except SA and were melted in a Brabender mixing chamber (Plasti-corder EC), which was pre-heated at 190°C . The rotor speed was set at 40rpm, and the time of mixing was fixed at 20min. Mixing times between 8 and 50min and rotating speeds between 40 and 150rpm have been used in different laboratory polymer mixers [29–33]. The material was withdrawn from the mixer chamber with a spatula, compressed between hot plates at 190°C in a Collin press, under 100bar, followed by cold pressing at 20°C .

D. Characterization

1) XRD

X-Ray Diffraction (XRD) spectra were recorded using a Bruker D8 Advance A25 diffractometer with a copper laser and a SSD160 TM detector. Ni-filtered $\text{CuK}\alpha$ radiation (wavelength of 0.1542nm) was produced at 40kV and 25mA. Scattered radiation was detected at the angular range of (2θ) 0–45° with a step of 0.02°.

2) FTIR

The different composites studied were analyzed by Fourier Transform Infrared Spectroscopy (FTIR) in the transmission mode using KBr pellets for the powdered sample and films prepared by compression-molded for PP/bentonite composites. The spectra were recorded on a Perkin Elmer Spectrum 1000 spectrometer at a resolution of 4cm^{-1} in the range of 4000 to 400cm^{-1} .

3) SEM

For Scanning Electron Microscopy (SEM) micrographs, several acquisitions of images were performed on the surface of the sample in several zones and with different magnifications from 500x to 5000x in SEM-EDXQuanta 250 tungsten filament equipment.

4) DSC

The melting points (T_m) of the PP composites and the virgin PP were analyzed with a Differential Scanning Calorimeter (DSC) Q 2000 thermal analyser. The samples were heated from -50°C to 200°C at the rate of $20^\circ\text{C}/\text{min}$ under an inert atmosphere of nitrogen (50ml/min), after decreasing the temperature from room temperature to -50°C . Then, the samples were heated up to 200°C and were maintained at this temperature for 5min. Consecutive to the heating process, samples were cooled down to 25°C , and after that, new successive heating and cooling runs were performed. The crystallinity percentage (X_c) was calculated with:

$$X_c = \frac{(\Delta H_m)}{\phi(\Delta H_0)} \times 100 \quad (1)$$

where ΔH_m is the measured heat of melting per gram of polymer during the second heating scan, ΔH_0 (207.1J/g) is the theoretical heat of crystallization of 100% crystalline isotactic polypropylene [34], and ϕ is the weight fraction of PP in the composite. Thermal analysis was performed in a Q2000 TA instrument. The samples ($\pm 10\text{mg}$) were weighed to 0.002mg with an electronic balance (Perkin-Elmer AD4). The samples were heated from 25°C to 900°C at a heating rate of $20^\circ\text{C}/\text{min}$ in air. All the samples were previously dried at 80°C for 24h.

5) Mechanical Properties

Tensile tests were carried out according to UNE-ENN ISO 527-1 and 527-2 with an Instron Model 5500R60025 apparatus. One mm thick specimens were cut from the sheets with a Wallace die cutter. Across-head speed of 10mm/min was used.

III. RESULTS AND DISCUSSIONS

A. Processability (Torque Value)

Figures 1 and 2 show the processing torque during the mixing of PP/bent composites. It was observed that the addition of 5% of AA and PP-MA to the polymer matrix loaded with 5% bent-m increased the torque values of the

systems (31N.m, 25N.m), suggesting that the composites required more severe processing conditions in its presence. In contrast, 5% of SA conducted PP/bent-m system to processing torque reduction (15N.m) and improved flow properties.

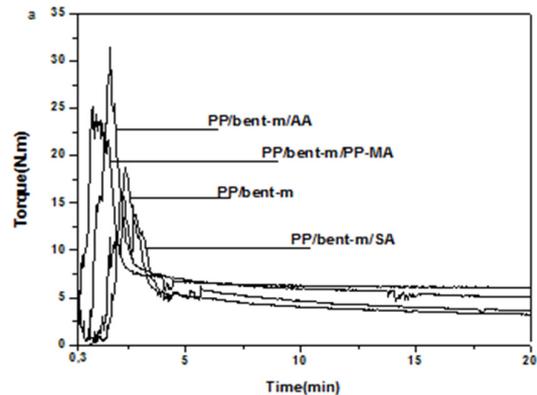


Fig. 1. Torque behavior of PP/bent-m composites

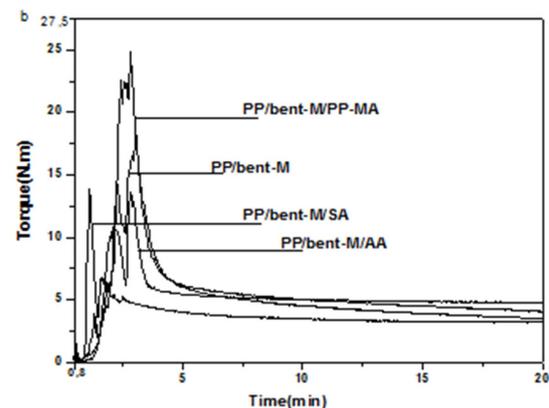


Fig. 2. Torque behavior of PP/bent-M composites

These results give possible confirmation that the stearic acid is acting as a coupling modifier and lubricant. The enhancement of lubricity through the processing provided fine dispersion of the filler, and it was also in accord with the decrease in dimension and quantity of agglomerates showed by the SEM. This was in agreement with the findings of the authors of [23], who reported that during the processing of the PP/modified bentonite composites, the stearic acid is acting as an interface modifier and lubricant and thus decreases the viscosity of the melt compound due to the free movement of the polymer chains. These effects enhanced the mechanical behavior by improving the elongation at break of the composite. On the other hand, bent-M showed a different behavior than the bent-m with the addition of 5% of AA coupling agent. The torque decreased to 14N.m in PP/bent-M/SA and to 11N.m in PP/bent-M/AA composites. The reduction in the torque values observed in this case is important and can be attributed to a decrease in particle/particle interaction which showed that different SA and AA coupling

agents act as external lubricants that are able to facilitate the processing of the PP/bent-M composites.

Authors in [35] studied the effect of palm oil fatty acid and polypropylene grafted-maleic anhydride compatibilizers on the peak torque of bentonite filled PP. They showed that the incorporation of these coupling agents decreases the torque generated during mixing and the agents serve as external lubricants. The torque values of PP/bent-m composites in the presence of AA and PP-MA are higher. This behavior can be attributed to the good adhesion between the PP matrix and bent-m in the presence of these coupling agents, which may inhibit the chain flexibility and increase the composite torque. Authors in [36] showed that the viscosity of hexamethylenediamine modified PP/montmorillonite nanocomposites is higher than that of the pure PP in the presence of maleic anhydride grafted polypropylene. Clearly, in PP/bent-M composites, the integration of AA reduces the quantity of the generated torque during mixing as compared to PP/bent-m composites. The possible reason for the differences between the behaviors of the two bentonites (bent-m and bent-M) in the presence of AA is the dissimilarity in particle size distribution: bent-M presents large particles with average particle size of 24.8 μm which can form agglomerates. The introduction of AA as a coupling agent reduces particle-particle interaction resulting in the decrease of particle agglomeration and an enhancement of dispersion and flow properties.

B. XRD

Figure 3 presents the XRD patterns of coupled PP/bent-m samples, revealing reflections assigned to planes (110), (040), (130), and (041) as reported in [37, 38].

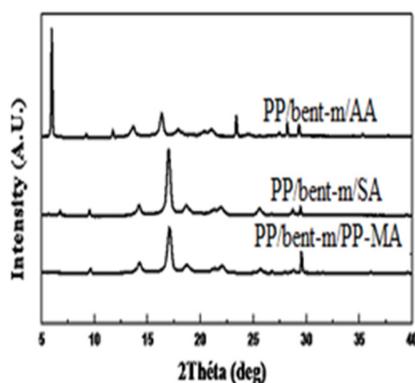


Fig. 3. XRD spectra of PP/bent-m composites.

The addition of AA in PP/bent-m/AA samples exhibited reflections with diffractions angles around 11.75° and 23.4°. This result suggests the occurrence of partial decomposition of the organic modifier of AA during the mixing process with bentonite [39] or implies that partially intercalated structure took place in composite containing the AA coupling modifier. PP/bent-m/PP-MA composite was not influenced after blending and forming. The peak that appears for the diffraction angle around 6° in PP/bent-m/SA composite is due to the amount of stearic acid introduced as a coupling modifier. These results are in good agreement with the findings of the authors

in [24] who reported that the addition of stearic acid as interface modifier showed the same peak at around 6° and provided better intercalation of the silicate layers in its nanocomposites when compared to the rest of the PP/bentonite nanocomposites. Similar diffraction patterns were also observed for the composites with bent-M and with different coupling agents (not shown for brevity).

C. FTIR Spectra

Figure 4 shows the FT-IR spectra of bentonite (bent-m and bent-M). The bands obtained at 3600-3200 cm^{-1} are attributed to the OH stretching mode of Si-OH groups [40]. The band between 1640-1620 cm^{-1} belongs to the deformation of the adsorbed H₂O molecules. The intense broad-band with the maximum at 1041 cm^{-1} ranged at 1040-1080 cm^{-1} corresponds to the Si-O-Si valence vibration of the SiO₄. In addition, the narrow peak at 797 cm^{-1} corresponds to the silanol group [41] and at around 600 cm^{-1} corresponds to Si-O-Al [23]. The peak recorded at 1439 cm^{-1} is associated with the presence of calcite in the sample [42]. Bent-M shows a much larger peak than the bent-m as the CaO content is greater in the case of bent-M (7.06% mass) than of the bent-m (3.16% mass). On the other hand, bent-m is more hygroscopic than bent-M, because the intensity of the absorption bands at 3432.8 cm^{-1} and 1635 cm^{-1} assigned to the stretching and bending vibrations of the OH groups for the water molecules adsorbed on bentonite surface [43] are superior in the case of bent-m than bent-M (1619 cm^{-1} , 3432.8 cm^{-1}).

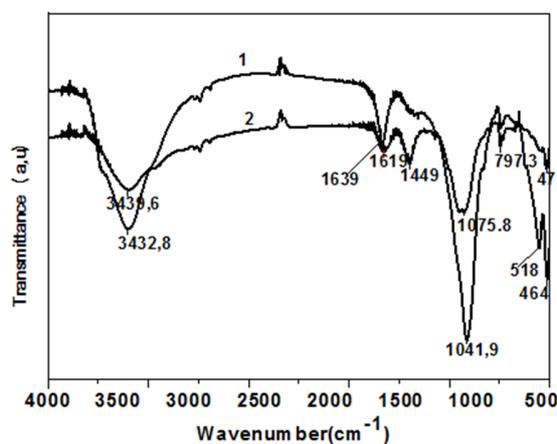


Fig. 4. FTIR spectra of (1) bent-m and (2) bent-M.

The FT-IR spectra of PP (Figure 5(a)) display absorption peaks which coincide with the literature. Methylene groups vibrations are registered in the range 1445-1485 cm^{-1} and methyl group vibrations are registered in the range 1430-1470 cm^{-1} or 1365-1395 cm^{-1} [44]. These peaks in our spectrum appear at 1455 and 1355 cm^{-1} respectively. At 2838-2982 cm^{-1} , the peaks are attributed to the asymmetric and symmetric CH bands [40]. The characteristic vibrations of terminal unsaturated CH₂ group absorbs at 840, 1000, and 1170 cm^{-1} [44]. These peaks are detected in our spectrum at 840, 980, and 1170 cm^{-1} (Figure 5(a)). The FT-IR spectra in Figure 5(c) and Figure 5(d) display the contribution of AA in PP/bent-m and PP/bent-M composites. Again the bands at 950-1080 cm^{-1} are

characteristic of the asymmetric and symmetric vibrations of Si-O-Si. The bands ranged generally at $1540\text{-}1640\text{cm}^{-1}$ attest of the stretching vibrations of CO_2 [45]. In our spectrum, there were registered at $1512\text{-}1608\text{cm}^{-1}$. The bending vibrations due to the symmetric CH_3 and due to HCH band absorb at 1385.5cm^{-1} were generally detected at 1408cm^{-1} [45]. This one disappeared in our composites. On the other hand, a differentiated absorption at 1295.5cm^{-1} associated with the bending vibration of CH_3 , stretching of CC and rocking of CH also disappeared in our composites. The peak in the spectra of pure AA at 1239cm^{-1} (Figure 5(b)) is attributed to the rocking of NH_3 and bending of COH, the intensity of the corresponding peak in our composites was decreased at this range. The peak generally located at 1133cm^{-1} was attributed to the CC stretch [45], and was seen at 1146cm^{-1} . There is a displacement toward the high frequency region 1150cm^{-1} in our composites (Figure 5(c)-(d)). While 769cm^{-1} is the peak of the bending of CO_2 , a weak signal in the corresponding band was observed in our composites and these remarks suggested the immobilization of AA on the surface of bentonite.

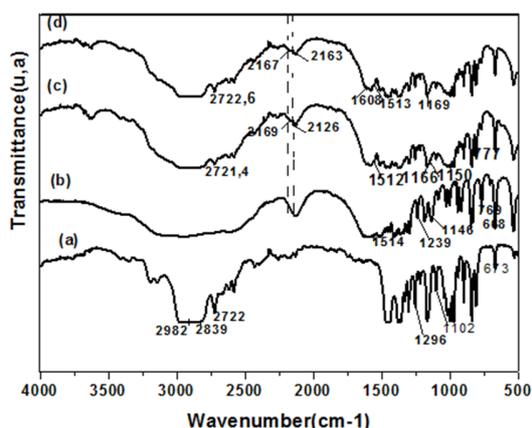


Fig. 5. FTIR spectra of (a) neat PP, (b) AA, (c) PP/bent-m/AA, and (d) PP/bent-M/AA.

The FTIR spectra of the pure PP-MA film (Figure 6(b)) are characterized by the presence of two bands typical of the pure PP-MA at 1783cm^{-1} and 1717cm^{-1} which are attributed to the symmetric and asymmetric stretching vibration of the C=O group. The FTIR spectra of the composite in Figure 6(c)-(d) shows very weak peaks at 1783cm^{-1} and 1717cm^{-1} which attest the small amount of the carbonyl in a polymer molecule. C-O group is registered in the range of $1190\text{-}960\text{cm}^{-1}$ in pure PP-MA [44]. In our case, C-O absorption was detected at 1165cm^{-1} . It was observed that this absorption intensity is less strong in PP/bent-m/PP-MA composite and diminished in PP/bent-M/PP-MA composite. Also, the peak characteristic of C-O group shifted to 1167cm^{-1} in PP/bent-m/PP-MA. This suggested the presence of new hydrogen bonds between PP-MA and bent-m. In the case of neat SA (Figure 7(b)), broad and intense bands observed at 2916cm^{-1} and 2858cm^{-1} are attributed to the stretching vibration of C-H bands located generally in range of $2800\text{-}2900\text{cm}^{-1}$. The absorption in this region showed very weak peaks in the spectrum of composite (Figure 7(c)-(d)). Also, the original C=O absorption at 1704cm^{-1}

cm^{-1} of SA shifted by 4cm^{-1} toward the high frequency region in the composites with SA. Apparently these changes are due to the adsorption of SA on bentonite as reported in [24].

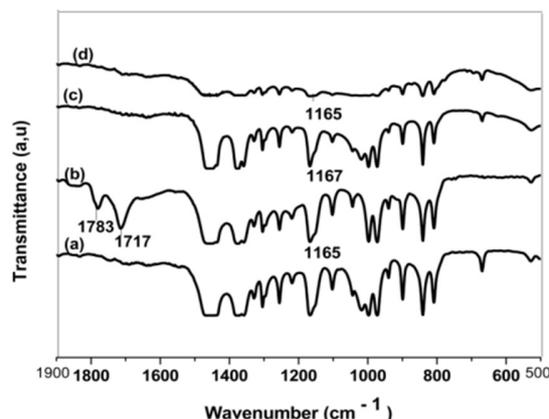


Fig. 6. FTIR spectra of (a) neat PP, (b) PP-MA, (c) PP/bent-m/PP-MA, and (d) PP/bent-M/PP-MA (d).

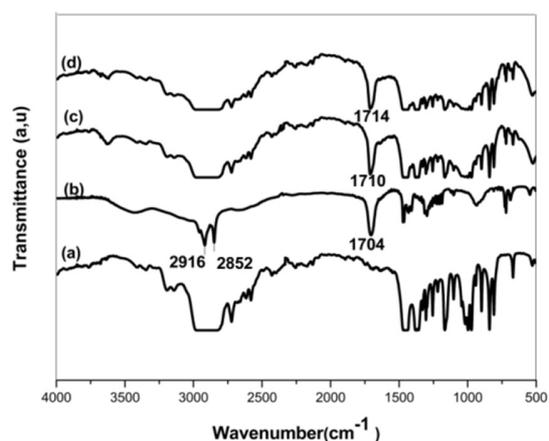


Fig. 7. FTIR spectra of (a) neat PP, (b) SA, (c) PP/bent-m/SA, and (d) PP/bent-M/SA.

D. Scanning Electron Microscopy of Film Surface

The morphologies of the film surface of uncoupled and coupled PP/bent-m and PP/bent-M composites are presented in Figures 8 and 9 respectively. SEM micrographs show micro-sized particles of 5wt% clay in PP/bent-m composite (Figure 10), which is larger as compared to particles in coupled composites. This indicates an inhomogeneous distribution of clay in PP in the absence of a coupling agent. With further inclusion of SA and AA coupling agents, the presence of agglomerates in bent-m is almost unnoticeable, which is attributed to the good dispersion of bentonite bent-m in the PP matrix in the presence of these coupling modifiers and clarified the enhancement in some mechanical properties. Stearic acid as coupling agent introduces an improved wetting of bentonite particles through the matrix of PP, owing to interactions between clay and PP. In this case, bentonite layers were aggregated with a size of 195nm in the polypropylene matrix.

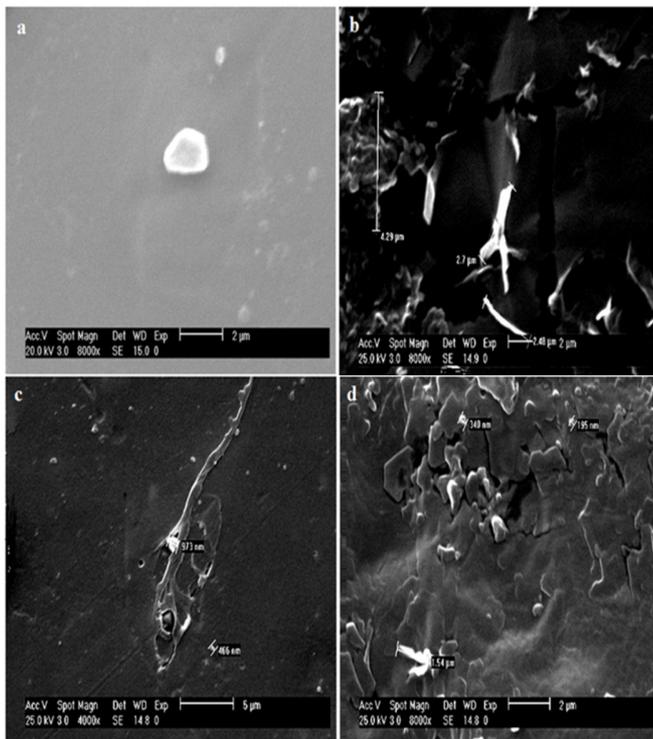


Fig. 8. SEM micrographs of (a) PP/bent-m, (b) PP/bent-m/PP-MA, (c) PP/bent-m/AA, and (d) PP/bent-m/SA.

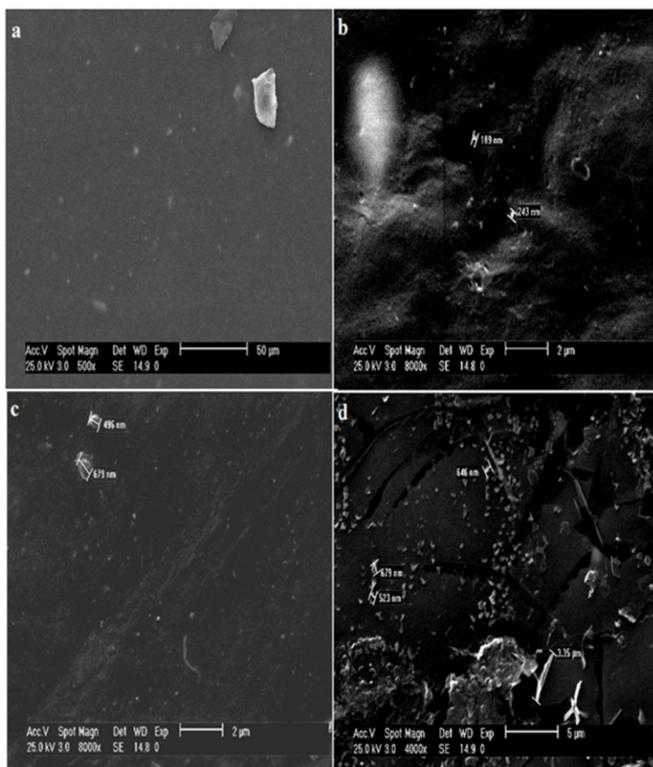


Fig. 9. SEM micrographs of (a) PP/bent-M, (b) PP/bent-M/PP-MA, (c) PP/bent-M/AA, and (d) PP/bent-M/SA.

The SEM micrographs of PP/bent-M composites (Figure 9) with a clay content of 5wt% showed that the uncoupled composite included a small quantity of agglomerates of clay of size less than 50µm. By adding AA and SA coupling agents in the composite, no detectable particle agglomerates were observed at 2µm scale. This behavior indicates a good dispersion of bent-M in PP matrix in the presence of amino acid and stearic acid coupling agents. However, the PP/bent-M/PP-MA micrograph shows large aggregates. The structure (Figure 9(b)) observed in PP/bent-m/SA can be attributed to the needles structure of stearic acid and to the good dispersion of bent-m in the PP matrix in the presence of this coupling agent.

E. Differential Scanning Calorimeter Properties

Table I illustrates the values of the crystallinity content (X_c) calculated from ΔH_m which is the enthalpy of melting of polymer in the composite, compared with the enthalpy of melting of 100% crystalline PP, the melting temperature (T_m) and the crystallization temperature (T_c) of all prepared samples. Each composite in the presence of coupling agent exhibits similar T_m regardless of the kind of bentonite. It is noticeable that T_m and T_c remained unaffected with the incorporation of the clay in the presence of different coupling agents in all composites, indicating that the crystal size of PP did not change [46]. The degree of crystallinity increased slightly in PP-MA and SA coupled composites as compared to that in pure PP (51%). The degree of crystallinity increased slightly in PP-MA and SA coupled composites as compared to that in pure PP (51%) and in uncoupled composite. PP/bent-m/PP-MA and PP/bent-M/SA present the higher degree of crystallinity, around 57%, while PP/bent-M/AA provides the lowest crystallinity due to the lower enthalpy of melting of the compatibilizer. This result confirms how the crystallization of PP matrix varied with the further inclusion of PP-MA, AA, and SA coupling agents [39]. The percentage of crystallinity did not change with the addition of bentonite. This signifies that the bentonite did not act as a nucleating agent in this compound. According to [36], the presence of coupling agents appears to decrease the crystallization process in bentonite filled PP composite and bentonite did not serve as nucleating agent in this system.

TABLE I. CHARACTERISTIC THERMAL PARAMETERS OF PP/BENT COMPOSITES WITH AND WITHOUT COUPLING AGENTS

Samples	T_m (°C)	ΔH_m (J/g)	T_c (°C)	ΔH_c (J/g)	X_c %
Neat PP	165	107	126	113	51
PP/bent-m	166	103	127	110	52
PP-bent-M	166	102	127	109	52
PP/bent-m/PP-MA	165	106	127	109	57
PP/bent-m/SA	164	101	127	104	54
PP/bent-m/AA	165	99	127	108	53
PP/bent-M/PP-MA	165	103	126	109	55
PP/bent-M/SA	164	107	127	107	57
PP/bent-M/AA	165	93	127	103	50

F. Mechanical Properties

The results of mechanical properties of PP and uncoupled and coupled PP/bent composites are exhibited in Table II. We observe a decrease in Young's modulus caused by the bent-M filler which appears to be related to the amount of clay and the size filler. The Young's modulus of PP/bent-M/AA composite

was still improved even when compared to the other composites of PP/bent-M coupled to PP-MA or SA. The changes in Young's modulus reflect the nature of the interface between polymer and filler. The introduction of bent-m in PP provides a slight decrease in Young's modulus and PP/bent-m/SA had the same stiffness as the uncoupled composite. No changes of Young's modulus in PP/bent-m coupled to PP-MA were observed, while this property was more improved for PP/bent-m/AA composite due to the effective adhesion between PP and clay in the presence of AA coupling agent, conducted by the homogeneous dispersion as shown by the SEM images, providing an easy transfer of load throughout the composite [47]. As expected at 5 wt.% of clay content, there is a significant decrease in elongation at break. The elongation at break for neat PP is above 140% and decreases to 103% especially in PP/bent-m. This decrease could be caused by the fact that the inorganic particles of filler are rigid and cannot be deformed by external stress in the samples but acted only as stress concentrators during the deformation process [36]. A significant decrease in elongation at break was reported in [48] for LDPE/TPS/SNC nanocomposites by adding 5wt.% of starch nanocrystals (SNC) without DCP compared to pure LDPE. By introducing PP-MA coupling agent, an enhancement in elongation at break of about 21% is observed compared to uncoupled PP/bent-m composite. However, the percentage of improvement in elongation at break by using SA in PP/bent-m compound is 37%. This increase in ductility (i.e. elongation at break) might be due to the existence of fatty acids in SA which act as interface modifiers for bentonite into the PP matrix. Consequently, the ductility of bentonite filled PP composites can be improved [35]. These results prove the results of torque values. According to [24, 35], fatty acids promote an improvement in elongation at break and the dispersion state of nanoclay and bentonite in a polypropylene matrix. It has been noted [35] that the elongation at break of composites under the presence of Palm Oil Fatty Acid (POFA) additive shows the highest value compared to composites with the presence of PP-MA. A similar trend is observed for the elongation at break and the strain at break change for all samples. Table II shows a reduction of the maximum stress by the incorporation of the unmodified bent-m or bent-M filler. This behavior can be explained by the low adhesion between bentonite and the PP matrix. The unmodified bentonite particles in uncoupled composite can serve as sites to initiate and activate the deformation mechanism and as a result, the yield stress decreases. On the other hand, there are some bentonite layers with good matrix adhesion based on SEM micrographs in PP/bent-m and PP/bent-M composites coupled with PP-MA coupling agent. These ones can withstand the movement of the polymer chains and therefore increase the yield stress. Improvement in yield values by the introduction of PP-MA compatibilizer in PP/modified clay has been reported in [49]. For coupled composites in the presence of SA and AA coupling agents and because they act as external lubricants in PP/bent-M composites (as seen on the torque results), free movement of the polymer chains is possible. The latter makes the physical sliding between the molecular segments easier which results in the decrease of the maximum stress. The same remark was noted for the coupled PP/bent-m/SA composite. From the results of this study, it can be supposed that the AA

and SA coupling modifiers show certain degrees of interactions in the PP/bentonite system as observed in some improvement on mechanical properties. These interactions are superior for bent-m because of its low average particle size compared to bent-M.

TABLE II. VALUES OF STIFFNESS, ELONGATION AT BREAK, STRAIN BREAK, AND MAXIMUM STRESS OF PP/BENTONITE

Samples	Young's Modulus (MPa)	Elongation at break %	Strain break %	Maximum stress (MPa)
Neat PP	1538±20	145±4	7.3±0.2	37.9±0.2
PP/bent-m	1516±6	103±3	5.3 0.4	34.0±0.6
PP-bent-M	1467±7	136±8	6.8±0.4	35.5±0.1
PP/bent-m/PP-MA	1539±5	125±3	6.3±0.2	35.8±0.3
PP/bent-m/SA	1516±24	141±1	7±1	32.5±0.2
PP/bent-m/AA	1719±52	83±4	4.2±0.2	30.5±0.4
PP/bent-M/PP-MA	1435±23	123±0.1	6.1±0.3	37.0±0.4
PP/bent-M/SA	1481±31	79±2	3.9±0.5	31.5±0.8
PP/bent-M/AA	1532±13	94 ±0.1	4.7±0.5	29.6±0.2

G. Thermogravimetric Analysis (TGA)

TGA was employed to evaluate the thermal stability of each sample. Figure 10 shows the TGA curves of neat PP, PP/bent-m, and PP/bent-M composites in the presence of different coupling agents. We do not observe an improvement in $T_{5\%}$ in all cases of coupled PP/bentonite systems (excepted PP/bent-m/AA). The temperature at 20% weight loss of all composites enhanced (excepted PP/bent-M/SA). The coupling agent can improve the stability at high temperatures [50]. The temperature at 50% weight loss of PP/bent-m/SA displays the highest value. This composite is more stable than neat PP, while a slight decrease is shown in PP/bent-M/SA. Adding AA to PP/bent-m and PP/bent-M results in a lower improvement of $T_{50\%}$ compared to SA (there is about 13°C of $T_{50\%}$ improvement as compared to neat PP). PP-MA displays a slight increase in $T_{50\%}$ in both PP/bent-m and PP/bent-M. However, there is an improvement in overall coupled composites in T_{max} , the highest being recorded for PP/bent-m/SA composite. It is noted that these materials become more thermally stable due to an improvement in the miscibility between all the added coupling agents [51]. Also, thermal stability enhancement is a consequence of the fine dispersion of bentonite within the PP matrix and the interface which resulted in slower escape of decomposed smaller molecules[39, 52-53].

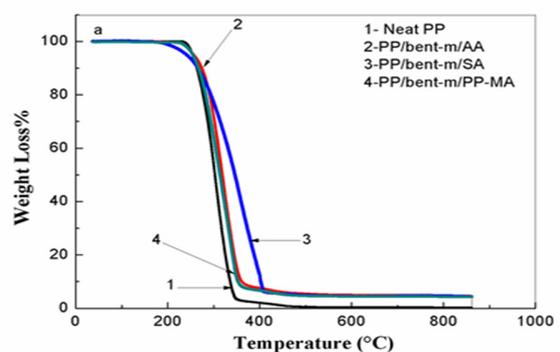


Fig. 10. TGA curves of selected samples of PP with PP/bent-m composites

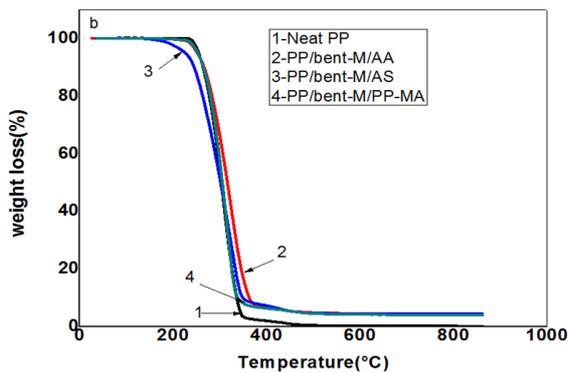


Fig. 11. TGA curves of selected samples of PP with PP/bent-M composites

TABLE III. THERMOGRAVIMETRIC PARAMETERS OF PP AND PP/BENTONITE COPOSITES IN THE AIR

Samples	$T_{5\%}$ (°C)	$T_{20\%}$ (°C)	$T_{50\%}$ (°C)	T_{max} (°C)
Neat PP	254	278	303	306
PP/bent-m/PP-MA	252	279	305	336
PP/bent-m/SA	241	294	347	353
PP/bent-m/AA	256	291	319	324
PP/bent-M/PP-MA	251	280	305	316

This increase in the decomposition temperature of about 40°C, as compared with the neat polymer, might be taken as a sign of the incidence of clay intercalation, since the intercalated or exfoliated clay will act as a barrier to the diffusion of atmospheric oxygen in the compound, inhibiting the polymer decomposition [54]. Recently, it was reported that the incorporation of fatty acids like stearic acid between the layers of mineral ceramic facilitates polymer-ceramic interaction [55]. The degradation temperatures of all samples are listed in Table III.

IV. CONCLUSIONS

This study has investigated the effect of the addition of different coupling agents such as PP-MA, SA, and AA on the mechanical properties, thermal stability, crystallinity, and rheological properties of PP/raw bentonite composites. Fillers such as bentonite without treatment can be used as good reinforcement. According to XRD results, polar molecules of AA exhibited reflections that shift to smaller angles. This is an indication that AA was introduced in bentonite. PP/bent-m/PP-MA and PP/bent-M/SA composites present the higher degree of crystallinity of around 57% as compared to neat PP and uncoupled composites. PP/bent-m/SA composite is more stable than neat matrix while the improvement in decomposition temperature is about 13°C in PP/bent-m/AA composite. The presence of the different coupling agents in natural bentonite increases T_{max} in all the composites. The PP/bent-m/SA system can be considered as an important material, which improves the thermal stability and kept the modulus and elongation at break at a value almost similar to that of PP. In addition, PP/bent-m/AA composite can also be selected as a material with high Young modulus and good thermal stability. These enhancements were important as this behavior has not been reported in the literature. In most cases the elongation at break properties are reduced in PP filled with nanoclay. Also, the

increase in thermal stability and modulus of elasticity were observed after the addition of nanobentonite in the presence of stearic acid as interface modifier and not for PP/raw microbentonite composite containing 5wt.% of filler and 5wt.% of stearic acid or amino acid as coupling agents. The behavior of processing composites showed a decrease in torque values above the uncoupled system by using SA and AA coupling agents in PP/bent-M composites. The decrease is more significant in presence of SA in PP/bent-m composites.

ACKNOWLEDGMENT

The authors appreciate the financial support from the Setif 1 University.

REFERENCES

- [1] T. G. Mofokeng, S. S. Ray, V. Ojijo, and R. Salehiyan, "Tuning the nano/micro-structure and properties of melt-processed ternary composites of polypropylene/ethylene vinyl acetate blend and nanoclay: The influence of kinetic and thermodynamic parameters," *Journal of Applied Polymer Science*, vol. 135, no. 1, 2018, Art. no. 45585, <https://doi.org/10.1002/app.45585>.
- [2] N. Zaman, S. Ahmed, M. Sanaullah, A. U. Rehman, A. R. Shar, and M. R. Luhur, "Fabrication and Characterization of Organoclay Reinforced Polyester Based Hybrid Nanocomposite Materials," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 3038–3040, Jun. 2018, <https://doi.org/10.48084/etasr.1977>.
- [3] B. A. A. Majeed and D. A. Sabar, "Preparations of Organoclay Using Cationic Surfactant and Characterization of PVC/ (Bentonite and Organoclay) Composite Prepared via Melt Blending Method," *Iraqi Journal of Chemical and Petroleum Engineering*, vol. 18, no. 1, pp. 17–36, Mar. 2017.
- [4] R. S. Cardoso, V. O. Aguiar, and M. de F. V. Marques, "Masterbatches of polypropylene/clay obtained by in situ polymerization and melt-blended with commercial polypropylene," *Journal of Composite Materials*, vol. 51, no. 25, pp. 3547–3556, Oct. 2017, <https://doi.org/10.1177/0021998317690444>.
- [5] Y. Hernandez *et al.*, "Improvement of toughness properties of polypropylene filled with nanobentonite using stearic acid as interface modifier," *Journal of Composite Materials*, vol. 51, no. 3, pp. 373–380, Feb. 2017, <https://doi.org/10.1177/0021998316644852>.
- [6] D. D. J. Rousseaux, M. Sclavons, P. Godard, and J. Marchand-Brynaert, "Tuning the functionalization chemistry of polypropylene for polypropylene/clay nanocomposites," *Reactive and Functional Polymers*, vol. 72, no. 1, pp. 17–24, Jan. 2012, <https://doi.org/10.1016/j.reactfunctpolym.2011.10.005>.
- [7] S. K. Sharma and S. K. Nayak, "Surface modified clay/polypropylene (PP) nanocomposites: Effect on physico-mechanical, thermal and morphological properties," *Polymer Degradation and Stability*, vol. 94, no. 1, pp. 132–138, Jan. 2009, <https://doi.org/10.1016/j.polydegradstab.2008.09.004>.
- [8] M. Alexandre and P. Dubois, "Polymer-layered silicate nanocomposites: preparation, properties and uses of a new class of materials," *Materials Science and Engineering: R: Reports*, vol. 28, no. 1, pp. 1–63, Jun. 2000, [https://doi.org/10.1016/S0927-796X\(00\)00012-7](https://doi.org/10.1016/S0927-796X(00)00012-7).
- [9] B. Uzbas and A. C. Aydin, "Microstructural Analysis of Silica Fume Concrete with Scanning Electron Microscopy and X-Ray Diffraction," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5845–5850, Jun. 2020, <https://doi.org/10.48084/etasr.3288>.
- [10] G. Bulut, M. Chimeddorj, F. Esenli, and M. S. Çelik, "Production of desiccants from Turkish bentonites," *Applied Clay Science*, vol. 46, no. 2, pp. 141–147, Oct. 2009, <https://doi.org/10.1016/j.clay.2009.07.013>.
- [11] S. Cara, G. Carcangiu, G. Padalino, M. Palomba, and M. Tamanini, "The bentonites in pelotherapy: chemical, mineralogical and technological properties of materials from Sardinia deposits (Italy)," *Applied Clay Science*, vol. 16, no. 1, pp. 117–124, Jan. 2000, [https://doi.org/10.1016/S0169-1317\(99\)00049-6](https://doi.org/10.1016/S0169-1317(99)00049-6).

- [12] E. M. Araújo *et al.*, "The influence of organo-bentonite clay on the processing and mechanical properties of nylon 6 and polystyrene composites," *Materials Science and Engineering: B*, vol. 112, no. 2, pp. 175–178, Sep. 2004, <https://doi.org/10.1016/j.mseb.2004.05.027>.
- [13] B. Abu-Jdayil, K. Al-Malah, and R. Sawalha, "Study on Bentonite-Unsaturated Polyester Composite Materials," *Journal of Reinforced Plastics and Composites*, vol. 21, no. 17, pp. 1597–1607, Nov. 2002, <https://doi.org/10.1177/0731684402021017493>.
- [14] B. Abu-Jdayil and K. Al-Malah, "Jordanian Clay-Based Heat Insulator Composites: Mechanical Properties," *Journal of Reinforced Plastics and Composites*, vol. 27, no. 14, pp. 1559–1568, Sep. 2008, <https://doi.org/10.1177/0731684407087371>.
- [15] N. Hasegawa, M. Kawasumi, M. Kato, A. Usuki, and A. Okada, "Preparation and mechanical properties of polypropylene-clay hybrids using a maleic anhydride-modified polypropylene oligomer," *Journal of Applied Polymer Science*, vol. 67, no. 1, pp. 87–92, 1998, [https://doi.org/10.1002/\(SICI\)1097-4628\(19980103\)67:1<87::AID-APP10>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-4628(19980103)67:1<87::AID-APP10>3.0.CO;2-2).
- [16] R. Doufnoune, F. Riahi, S. Bouchareb, and A. Ourari, "Silanization of natural Na⁺-magnite algerian clay pre-intercalated with a cationic alkylphosphonium salt," *Journal of Adhesion Science and Technology*, vol. 33, no. 6, pp. 616–634, Mar. 2019, <https://doi.org/10.1080/01694243.2018.1551752>.
- [17] W. Mo, Q. He, X. Su, S. Ma, J. Feng, and Z. He, "Preparation and characterization of a granular bentonite composite adsorbent and its application for Pb²⁺ adsorption," *Applied Clay Science*, vol. 159, pp. 68–73, Jun. 2018, <https://doi.org/10.1016/j.clay.2017.12.001>.
- [18] D. García-López, O. Picazo, J. C. Merino, and J. M. Pastor, "Polypropylene-clay nanocomposites: effect of compatibilizing agents on clay dispersion," *European Polymer Journal*, vol. 39, no. 5, pp. 945–950, May 2003, [https://doi.org/10.1016/S0014-3057\(02\)00333-6](https://doi.org/10.1016/S0014-3057(02)00333-6).
- [19] H. Baniasadi, A. Ramazani S.A., and S. Javan Nikkhab, "Investigation of in situ prepared polypropylene/clay nanocomposites properties and comparing to melt blending method," *Materials & Design*, vol. 31, no. 1, pp. 76–84, Jan. 2010, <https://doi.org/10.1016/j.matdes.2009.07.014>.
- [20] H. Md. Akil, M. F. A. Rasyid, and J. Sharif, "Effect of Compatibilizer on Properties of Polypropylene Layered Silicate Nanocomposite," *Procedia Chemistry*, vol. 4, pp. 65–72, Jan. 2012, <https://doi.org/10.1016/j.proche.2012.06.010>.
- [21] G. Lagaly and T. J. Pinnavia, "Clay mineral - Polymer nanocomposites," *Applied Clay Science*, vol. 15, no. 1–2, 1999.
- [22] L. Minkova, Y. Peneva, E. Tashev, S. Filippi, M. Pracella, and P. Magagnini, "Thermal properties and microhardness of HDPE/clay nanocomposites compatibilized by different functionalized polyethylenes," *Polymer Testing*, vol. 28, no. 5, pp. 528–533, Aug. 2009, <https://doi.org/10.1016/j.polymertesting.2009.04.001>.
- [23] C. D. Angel *et al.*, "Mechanical and rheological properties of polypropylene/bentonite composites with stearic acid as an interface modifier," *Journal of Applied Polymer Science*, vol. 132, no. 30, 2015, <https://doi.org/10.1002/app.42264>.
- [24] L. Gonzalez *et al.*, "Mechanical and Thermal Properties of Polypropylene/Montmorillonite Nanocomposites Using Stearic Acid as Both an Interface and a Clay Surface Modifier," *Polymer Composites*, vol. 35, no. 1, Jan. 2014, <https://doi.org/10.1002/pc.22627>.
- [25] Q. Han, R. Yang, J. Li, Y. Zhang, and C. Wang, "Synthesis of leucine micro/nanocrystals for pharmaceutical applications," *CrystEngComm*, vol. 13, no. 20, pp. 6157–6162, Sep. 2011, <https://doi.org/10.1039/C1CE05519D>.
- [26] E. G. Pradas, M. V. Sánchez, F. C. Cruz, M. S. Viciano, and M. F. Pérez, "Adsorption of cadmium and zinc from aqueous solution on natural and activated bentonite," *Journal of Chemical Technology & Biotechnology*, vol. 59, no. 3, pp. 289–295, 1994, <https://doi.org/10.1002/jctb.280590312>.
- [27] H. Khalaf, O. Bouras, and V. Perrichon, "Synthesis and characterization of Al-pillared and cationic surfactant modified Al-pillared Algerian bentonite," *Microporous Materials*, vol. 8, no. 3, pp. 141–150, Feb. 1997, [https://doi.org/10.1016/S0927-6513\(96\)00079-X](https://doi.org/10.1016/S0927-6513(96)00079-X).
- [28] O. Bouras, J.-C. Bollinger, M. Baudu, and H. Khalaf, "Adsorption of diuron and its degradation products from aqueous solution by surfactant-modified pillared clays," *Applied Clay Science*, vol. 37, no. 3, pp. 240–250, Sep. 2007, <https://doi.org/10.1016/j.clay.2007.01.009>.
- [29] Y. Tang *et al.*, "Preparation and thermal stability of polypropylene/montmorillonite nanocomposites," *Polymer Degradation and Stability*, vol. 82, no. 1, pp. 127–131, Jan. 2003, [https://doi.org/10.1016/S0141-3910\(03\)00173-3](https://doi.org/10.1016/S0141-3910(03)00173-3).
- [30] K.-N. Kim, H. Kimm, and J.-W. Lee, "Effect of interlayer structure, matrix viscosity and composition of a functionalized polymer on the phase structure of polypropylene-montmorillonite nanocomposites," *Polymer Engineering & Science*, vol. 41, no. 11, pp. 1963–1969, 2001, <https://doi.org/10.1002/pen.10892>.
- [31] D. Marchant and K. Jayaraman, "Strategies for Optimizing Polypropylene-Clay Nanocomposite Structure," *Industrial & Engineering Chemistry Research*, vol. 41, no. 25, pp. 6402–6408, Dec. 2002, <https://doi.org/10.1021/ie011022c>.
- [32] W. Lertwimolnun and B. Vergnes, "Influence of compatibilizer and processing conditions on the dispersion of nanoclay in a polypropylene matrix," *Polymer*, vol. 46, no. 10, pp. 3462–3471, Apr. 2005, <https://doi.org/10.1016/j.polymer.2005.02.018>.
- [33] M. J. Solomon, A. S. Almusallam, K. F. Seefeldt, A. Somwangthanaroj, and P. Varadan, "Rheology of Polypropylene/Clay Hybrid Materials," *Macromolecules*, vol. 34, no. 6, pp. 1864–1872, Mar. 2001, <https://doi.org/10.1021/ma001122e>.
- [34] B. Wunderlich, "Atlas table of thermal properties of linear macromolecules," in *Thermal Analysis*, London, UK: Academic Press Limited, 1990, pp. 417–431.
- [35] N. Othman, H. Ismail, and M. Mariatti, "Effect of compatibilisers on mechanical and thermal properties of bentonite filled polypropylene composites," *Polymer Degradation and Stability*, vol. 91, no. 8, pp. 1761–1774, Aug. 2006, <https://doi.org/10.1016/j.polymdegradstab.2005.11.022>.
- [36] S. Sánchez-Valdes *et al.*, "Effect of different amine modified clays on the compatibility and clay dispersion of polypropylene nanocomposites," *e-Polymers*, vol. 9, no. 1, Dec. 2009, <https://doi.org/10.1515/epoly.2009.9.1.1499>.
- [37] B. Kim, S.-H. Lee, D. Lee, B. Ha, J. Park, and K. Char, "Crystallization Kinetics of Maleated Polypropylene/Clay Hybrids," *Industrial & Engineering Chemistry Research*, vol. 43, no. 19, pp. 6082–6089, Sep. 2004, <https://doi.org/10.1021/ie049825y>.
- [38] P. Liborio, V. A. Oliveira, and M. de F. V. Marques, "New chemical treatment of bentonite for the preparation of polypropylene nanocomposites by melt intercalation," *Applied Clay Science*, vol. 111, pp. 44–49, Jul. 2015, <https://doi.org/10.1016/j.clay.2015.04.003>.
- [39] F.-C. Chiu, H.-Z. Yen, and C.-E. Lee, "Characterization of PP/HDPE blend-based nanocomposites using different maleated polyolefins as compatibilizers," *Polymer Testing*, vol. 29, no. 3, pp. 397–406, May 2010, <https://doi.org/10.1016/j.polymertesting.2010.01.004>.
- [40] E. M. Essassi, M. Yahya, S. El Kazzouli, and M. Bousmina, "Synthesis and characterization of new organophilic clay. Preparation of polystyrene/clay nanocomposite," *Scientific Study & Research: Chemistry & Chemical Engineering, Biotechnology, Food Industry*, vol. 19, no. 2, pp. 193–202, Jul. 2018.
- [41] S. Yu. Khashirova, Yu. I. Musaeu, A. K. Mikitaev, Yu. A. Malkanduev, and M. Kh. Ligidov, "Hybrid nanocomposites based on guanidine methacrylate monomer and polymer and layered aluminosilicates: Synthesis, structure, and properties," *Polymer Science Series B*, vol. 51, no. 9, Nov. 2009, Art. no. 377, <https://doi.org/10.1134/S1560090409090085>.
- [42] H. Olphen and J. J. Fripiat, *Data handbook for clay materials and other non-metallic minerals*. Oxford, UK: Pergamon Press, 1979.
- [43] M. El Bouraie and A. A. Masoud, "Adsorption of phosphate ions from aqueous solution by modified bentonite with magnesium hydroxide Mg(OH)₂," *Applied Clay Science*, vol. 140, pp. 157–164, May 2017, <https://doi.org/10.1016/j.clay.2017.01.021>.
- [44] S. Kumar, "Spectroscopic studies of valine and leucine molecules a comparative study," vol. 2011, no. 39, pp. 4996–4999, 2011.

- [45] B. H. Stuart, *Polymer Analysis*, 1st ed. Chichester, NY, USA: Wiley, 2002.
- [46] V. S. Kumar, "Thermal Properties of Polypropylene/Montmorillonite Nanocomposites," *Indian Journal of Science and Technology*, vol. 7, no. Suppl. 7, Jan. 2015, <https://doi.org/10.17485/ijst/2014/v7sp7.3>.
- [47] E. Kontou, M. Niaounakis, and P. Georgiopoulos, "Comparative study of PLA nanocomposites reinforced with clay and silica nanofillers and their mixtures," *Journal of Applied Polymer Science*, vol. 122, no. 3, pp. 1519–1529, 2011, <https://doi.org/10.1002/app.34234>.
- [48] S. Chaoui, D. Smail, A. Hellati, and D. Benachour, "Effect of Starch Nanocrystals on the Properties of Low Density Polyethylene/Thermoplastic Starch Blends," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5875–5881, Aug. 2020, <https://doi.org/10.48084/etasr.3608>.
- [49] B. Akbari and R. Bagheri, "Influence of PP-g-MA on morphology, mechanical properties and deformation mechanism of copolypropylene/clay nanocomposite," *Journal of Applied Polymer Science*, vol. 114, no. 6, pp. 3751–3759, 2009, <https://doi.org/10.1002/app.30847>.
- [50] J. Son, D. J. Gardner, S. O'Neill, and C. Metaxas, "Understanding the viscoelastic properties of extruded polypropylene wood plastic composites," *Journal of Applied Polymer Science*, vol. 89, no. 6, pp. 1638–1644, 2003, <https://doi.org/10.1002/app.12372>.
- [51] J. M. Quiroz-Castillo *et al.*, "Preparation and Characterization of Films Extruded of Polyethylene/Chitosan Modified with Poly(lactic acid)," *Materials*, vol. 8, no. 1, pp. 137–148, Jan. 2015, <https://doi.org/10.3390/ma8010137>.
- [52] H. Qin *et al.*, "Thermal stability and flammability of polypropylene/montmorillonite composites," *Polymer Degradation and Stability*, vol. 85, no. 2, pp. 807–813, Aug. 2004, <https://doi.org/10.1016/j.polydegradstab.2004.03.014>.
- [53] J. Morawiec, A. Pawlak, M. Slouf, A. Galeski, E. Piorkowska, and N. Krasnikowa, "Preparation and properties of compatibilized LDPE/organo-modified montmorillonite nanocomposites," *European Polymer Journal*, vol. 41, no. 5, pp. 1115–1122, May 2005, <https://doi.org/10.1016/j.eurpolymj.2004.11.011>.
- [54] C. Ding, D. Jia, H. He, B. Guo, and H. Hong, "How organo-montmorillonite truly affects the structure and properties of polypropylene," *Polymer Testing*, vol. 24, no. 1, pp. 94–100, Feb. 2005, <https://doi.org/10.1016/j.polymertesting.2004.06.005>.
- [55] N. el H. Belkham, D. Benachour, and A. Mehamha, "Elaboration and Physico-Chemical Characterization of the Gibbsite Li(OH)3 Hybrid Material," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6740–6744, Feb. 2021, <https://doi.org/10.48084/etasr.3988>.

An End-to-End Machine Learning based Unified Architecture for Non-Intrusive Load Monitoring

Syed Wali

Faculty of Electrical and Computer Engineering
NED University of Engineering and Technology
Karachi, Pakistan
Syedwali110@hotmail.com

Muhammad Hassan ul Haq

Faculty of Electrical and Computer Engineering
NED University of Engineering and Technology
Karachi, Pakistan
hassanulhaq@neduet.edu.pk

Majida Kazmi

Faculty of Electrical and Computer Engineering
NED University of Engineering & Technology
Karachi, Pakistan
majidakazmi@neduet.edu.pk

Saad Ahmed Qazi

Neurocomputation Lab, National Centre of Artificial
Intelligence, NED University of Engineering and
Technology, Karachi, Pakistan
saadqazi@neduet.edu.pk

Abstract-Non-Intrusive Load Monitoring (NILM) or load disaggregation aims to analyze power consumption by decomposing the energy measured at the aggregate level into constituent appliances level. The conventional load disaggregation framework consists of signal processing and machine learning-based pipelined architectures, respectively for explicit feature extraction and decision making. Manual feature selection in such load disaggregation frameworks leads to biased decisions that eventually reduce system performance. This paper presents an efficient End-to-End (E2E) approach-based unified architecture using Gated Recurrent Units (GRU) for NILM. The proposed approach eliminates explicit feature engineering and has a unified classification and prediction model for appliance power. This eventually reduces the computational cost and enhances response time. The performance of the proposed system is compared with conventional algorithms' with the use of recall, precision, accuracy, F1 score, the relative error in total energy and Mean Absolute Error (MAE). These evaluation metrics are calculated on the power consumption of top priority appliances of Reference Energy Disaggregation Dataset (REDD). The proposed architecture with an overall accuracy of 91.2 and MAE of 25.23 outperforms conventional methods for all electrical appliances. It has been showcased through a series of experiments that feature extraction and event-based approaches for NILM can readily be replaced with E2E deep learning techniques allowing simpler and cost-efficient implementation pathways.

Keywords-non-intrusive load monitoring; gated recurrent units; end-to-end machine learning; reference energy disaggregation dataset

I. INTRODUCTION

Energy demand is increasing drastically with the increase in industrial development. This raises the need of managing energy usage effectively at consumer end. Efficient demand management is possible by analyzing the appliance level power consumption in buildings [1]. Today, this financially feasible solution, provided in 1992 [2], is known as Non-Intrusive Load

Monitoring (NILM). The basic idea of NILM revolves around the decomposition of total demand into appliance level power consumption. Since this load disaggregation approach does not depend on several data recorder sensors, it is therefore a cost-effective solution adopted for demand reduction and load forecasting. The interest of researchers in this domain is significantly increasing with the development of smart meters, capable of delivering aggregate power information to the customer. With the advancements in Machine Learning (ML) domain, it is expected that NILM with high accurate power consumption analysis capabilities will serve as the backbone in the development of innovative smart grid services [3].

Generally, NILM can be categorized into two types: event-based approaches and event-less or state dependent approaches. In the event-based approaches, any significant change in the signal that is considered during load disaggregation is regarded as an event. All event-based approaches depend on previous training, thus supervised ML approaches are mostly adopted in this category [4]. The second category, i.e. the event-less approach, does not rely on event detection. It primarily uses statistical and probability-models to match consumption signal of single or group appliances to the aggregate power signal [3]. Thus, label transitions are not required in this category. Event-based NILM methods primarily depend on finding edges in order to observe change in power demand. Initially, Hart formulated different clusters based on similar power change and characteristics of appliances using Combinatorial Optimization (CO) for disaggregating power demand [2]. Conventional edge detection approaches were then replaced by probability-based methods. These methods were comparatively less complex than the conventional methods [5, 6]. Standard deviation was calculated in [5] instead of using a single fixed parameter. Also, authors in [6] utilized a statistical approach for detecting the edges and power change signature of different appliances. Later, classifiers such as Support Vector Machines (SVMs) [7], Decision Tree [8], and other hybrid approaches

Corresponding author: Majida Kazmi

were investigated to serve the purpose. Hidden Markov Models (HMMs) and their several variants were also utilized to model multi-state appliances and different possibilities of their combination [9-11], since the complexity of modeling multistate appliances increases with increasing number of appliances installed in customer premises [12]. Therefore, inherent complexity of this methodology was reduced via the introduction of Viterbi algorithm in [13].

The latest development in ML shifts the paradigm to deep learning based NILM. Multilayer Perceptron (MLP), Convolution Neural Networks (CNNs), Deep Neural Networks (DNNs), Recurrent Neural Networks (RNN) K-Nearest Neighbor (k-NN) and Naïve Bayes classifiers are a few of the most widely used supervised ML techniques for load disaggregation [3]. CNNs and RNN-based Long Short-Term Memory (LSTM) have been explored for NILM in [12]. Authors in [14] implemented 3 different DNN architectures for short term load forecasting. 1D CNN was implemented in [15] for examining the effect of variables dependent on power demand. Hybrid CNN was also proposed in [16] where impact of reactive power, current, and apparent power were assessed on the performance of NILM disaggregation. The window of aggregate power signal was utilized in [15] to predict the power of the targeted appliance. The input sequence utilized for generating the sequence of output power is termed as sequence-to-sequence approach. If the same sequence predicts power at specific time instant only then it is termed as sequence-to-point. Both sequence-to-point and sequence-to-sequence NILM approaches [15] were based on CNN architecture. The load disaggregation problem was treated as noise reduction in [17] using denoising autoencoders. This approach showed improved performance under different types of loads.

One of the major drawbacks in the above stated deep learning based NILM architectures is their dependency on explicit features extraction from signal. This manual feature(s) extraction leads to biased decisions that eventually reduce the overall performance of NILM. In order to improve performance, these methods deploy extremely dense neural architecture with large number of layers. These dense architectures are time consuming and computationally expensive. Moreover, dependency on separate classification and regression networks [15, 18] further increases computational power, thus making these solutions extremely expensive. The manhours required in feature engineering and in the collection of contextual information for deep neural architectures comprise a time taking procedure and there lies a strong probability of losing important load signatures in manual feature extraction. In order to address all of the above limitations in previously proposed NILM approaches, this paper presents an efficient end-to-end (E2E) ML based unified architecture using Gated Recurrent Units (GRU).

The main characteristics of the proposed architecture are:

- The E2E ML approach is adopted which does not depend on explicit feature extraction. The feed input of this E2E architecture is complete aggregate power signal, ensuring reliable and better prediction even under different load categories.

- A unified module is an inherent characteristic of the proposed E2E architecture. Since the proposed architecture considers both classification of appliance and prediction of consumption as a single problem, there is no need to use separate modules.
- Low computational cost due to the unified architecture as compared to the conventional pipelined architecture. Real time load disaggregation is also possible due to the fast response time of the proposed architecture. It allows easy integration in modern smart metering devices.
- Improved performance of the proposed E2E architecture as compared to previously proposed DNN architectures despite of using comparatively lesser number of layers and neurons. The performance edge of the proposed approach is showcased on REDD, which is a renowned load disaggregation dataset.

II. THE PROPOSED E2E MACHINE LEARNING BASED UNIFIED ARCHITECTURE

An E2E ML based unified architecture for NILM is proposed in this work to completely eliminate the reliance on feature extraction. The proposed framework is presented in Figure 1. It consists of dataset and preprocessing, E2E ML model, and evaluation metrics.



Fig. 1. The proposed E2E ML-based unified architecture for NILM

A. Load Disaggregation Dataset and Preprocessing

Training and development of E2E ML Model depends on datasets prepared for load disaggregation. The Residential Energy Disaggregation Dataset (REDD) [4] is utilized in this research work. REDD was made publicly available in 2011 [4] with the aim of fast paced research and development in load disaggregation domain. It was developed with two major objectives. Firstly, it helps the researchers to apply algorithms and techniques directly on the available data instead of investing extensive efforts on the data acquisition stage of the NILM. Secondly, this dataset provides globally accepted reference data for comparing different algorithms and techniques implemented by different researchers.

The REDD dataset contains information of the aggregate power signals and the power of each individual appliance installed in 6 different homes in Massachusetts, USA. These 6 different homes cover almost all types of appliances used in consumer premises like washing machine, microwave, fridge, lights, air conditioning, electric stove, smoke detectors, etc. This dataset includes two-state, finite state, and continuously varying type of electrical loads. Low frequency power data from the first building of REDD dataset are selected for evaluation of the proposed algorithm. Six top priority appliances with respect to power consumption from House1 of REDD were considered. The selected dataset is preprocessed for removing erroneous readings, detecting gaps and downtime

using NILM tool kit (NILMTK). The data are separated as training and testing subsets with a ratio of 50:50

B. The E2E ML Model

The Gated Recurrent Unit (GRU), a variant of RNN, is selected as the basic ML model for the proposed E2E architecture. RNN is selected due to its efficient information handling with smaller context [19]. GRU is computationally simpler as compared to other RNN variants. It controls the flow of contextual information using just two gates as illustrated in Figure 2.

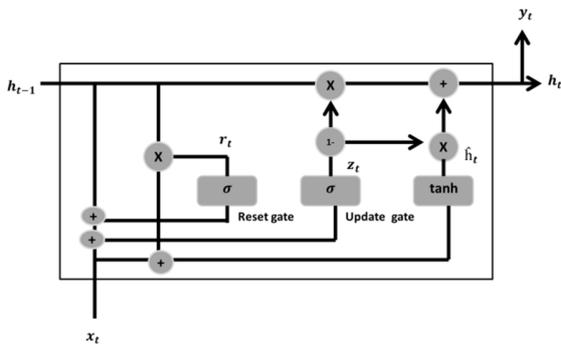


Fig. 2. The GRU architecture.

The first gate of GRU is the update gate. This gate is responsible to decide the extent to which information content from previous history should be passed for the determination of the future state. The output vector of the update gate (z_t) depends on previous cell output (h_{t-1}), present input (x_t), calculated weights (W_z , U_z), and biased vector (b_z). Mathematically, the output of the update gate depends on the sigmoid function and can be represented as:

$$z_t = \sigma [W_z * X_t + U_z * h_{t-1} + b_z] \quad (1)$$

The second GRU gate is the forget gate. This gate is responsible to filter and remove the flow of information from cells. It depends on current input (x_t), previous output (h_{t-1}), corresponding weights (W_r , U_r) and biased vector (b_r):

$$r_t = \sigma [W_r * X_t + U_r * h_{t-1} + b_r] \quad (2)$$

The final output produced by the GRU depends on intermediate memory state (\hat{h}_t). This intermediate memory depends on weights (W_h , U_h), current input (x_t), previous output (h_{t-1}), and biased vector (b_h). The mathematical model of this hidden memory state is shown in (3). It depends on tanh which is used as the activation function.

$$\hat{h}_t = \tanh [W_h * X_t + r_t * U_h * h_{t-1} + b_h] \quad (3)$$

The GRU output depends on the hidden memory state and update gate as shown in (4):

$$h_t = (1 - z_t) * h_{t-1} + z_t * \hat{h}_t \quad (4)$$

The proposed deep neural architecture consists GRU hidden layers, whereas the convolution layer and the dense layer with linear activation function are the input and output layers. The number of layers and neurons for GRU is selected for optimal performance using constructive approach in multiple passes. It

starts with a small or undersized network having a smaller number of neurons and layers, and gradually its size increases until it achieves optimal performance. During the first pass, the number of neurons is steadily increased from 64 to 2100 with a step size of 100. Better performance in terms of accuracy and MAE is observed in the range of 500 to 700 neurons as shown in Figure 3. In the second pass, neurons are gradually increased from 500 to 700 with a reduced step size of 10. It was found that the load disaggregation model suffered from overfitting when the number of neurons increased above 650. The best point in this region in terms of reduced MAE and improved accuracy was 630 neurons, so it was selected as the optimum number of neurons.

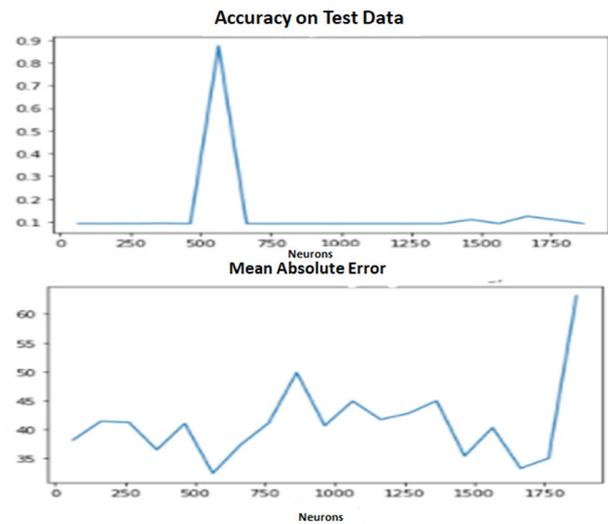


Fig. 3. GRU accuracy and MAE for different numbers of neurons.

Single hidden layered architecture is insufficient as a network with a smaller number of layers or neurons often fails to extract details from the training data [20]. Thus, 6 different architectures were tested comprising of 2, 4, 6, 8, 10, and 12 hidden GRU layers. Increasing architecture after 4 layers leads to overfitting and loss of generalization. Table I tabulates the accuracy on training data against the number of GRU layers. Optimal performance was achieved with the model with 4 GRU hidden layers of 1, 630, 1, and 1 neurons respectively.

TABLE I. ACCURACY ON TRAINING DATA AGAINST THE NUMBER OF GRU LAYERS

No. of layers	Accuracy						
	Overall test data	Fridge	Microwave	Washer dryer	Dish washer	Light	Socket
2	0.86	0.82	0.92	0.88	0.83	0.72	1
4	0.95	0.97	0.95	0.98	0.89	0.91	1
6	0.87	0.92	0.91	0.91	0.78	0.67	1
8	0.85	0.88	0.91	0.81	0.79	0.73	1
10	0.82	0.79	0.88	0.77	0.75	0.73	1
12	0.80	0.77	0.86	0.75	0.71	0.72	1

The learned model was then applied to the testing data. Table II tabulates the accuracy on the test data. The training and testing phases of the proposed E2E ML model are elaborated using the pseudocode shown in Figure 4.

TABLE II. ACCURACY OF THE LEARNED MODEL ON TEST DATA

Appliance	Accuracy
Fridge	0.88
Microwave	0.93
Washer dryer	0.98
Dish washer	0.81
Light	0.87
Socket	1
Overall	0.912

Algorithm: Pseudocode for training and testing of proposed

```

Import libraries
Import REDD dataset
Select House 1
Train Data ← Set train window in REDD
Test Data ← Set test window in REDD
Loads=[top appliances in House 1]
Algorithms = Proposed E2E Model
Initialize Arrays & Dataframe for storing results
For A in Algorithms:
    Initialize Disaggregator( ) w.r.t A
    For L in Loads:
        Train Disaggregator( Train data, L, sample period)
    end
    For L in Loads:
        R=Test Disaggregator( Test data, sample period )
        Results.append (R)
        Calculate True Positive, False Positive, True Negative
        Calculate Recall, Precision, Accuracy, F1 Score, MAE
        Store Accuracy metrics in Array
    end
    Pandas Dataframe ← Appliance Results
end
Display Results of Proposed E2E Model stored in Dataframe
    
```

Fig. 4. Training and testing pseudocode of the proposed E2E ML model.

C. Evaluation Metrics

The prediction of each load consumption at a certain time instant is not purely a regression problem. The load disaggregator first identifies the presence of certain appliances

in the aggregate power signal and then predicts their individual power consumption. Thus, performance evaluation should not be based on regression indices only, classification accuracy must also be evaluated. Recall, precision, accuracy, and F1 score are used to account classification performance, whereas, Mean Absolute Error (MAE) and relative error (RE) in total energy are used for the assessment of values predicted by the load disaggregator:

$$\text{Recall} = \frac{\text{True Positive}}{\text{False Negative} + \text{True Positive}} \quad (5)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{False Positive} + \text{True Positive}} \quad (6)$$

$$\text{F1 score} = 2 * \frac{\text{Precision} * \text{recall}}{\text{Precision} + \text{recall}} \quad (7)$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Ground Positive} + \text{Ground Negative}} \quad (8)$$

$$\text{RE} = \frac{|\text{Total predicted energy} - \text{Total actual energy}|}{\max(\text{Predicted Energy}, \text{Actual Energy})} \quad (9)$$

$$\text{MAE} = \frac{\sum_{t=1}^{T} |\text{estimated power at } t - \text{actual power at } t|}{\text{Time}(T)} \quad (10)$$

III. RESULTS AND COMPARISON

The proposed E2E ML model was evaluated and compared with conventional load disaggregation algorithms. NILMTK is equipped with benchmark algorithms. Two of the most widely used benchmark algorithms, i.e. CO and FHMM, are selected for evaluation on the REDD dataset, whereas one state of the art algorithm based on RNN architecture proposed in [21] is reproduced using the details provided by the authors. For fair comparison, all baseline algorithms were trained and tested with similar train tests split on 6 top priority appliances of REDD, i.e. fridge, microwave, dishwasher, washer dryer, sockets, and lights. These top 6 appliances, with respect to power consumption, were used to evaluate the performance of the conventional and the proposed algorithms as shown in Figure 5.

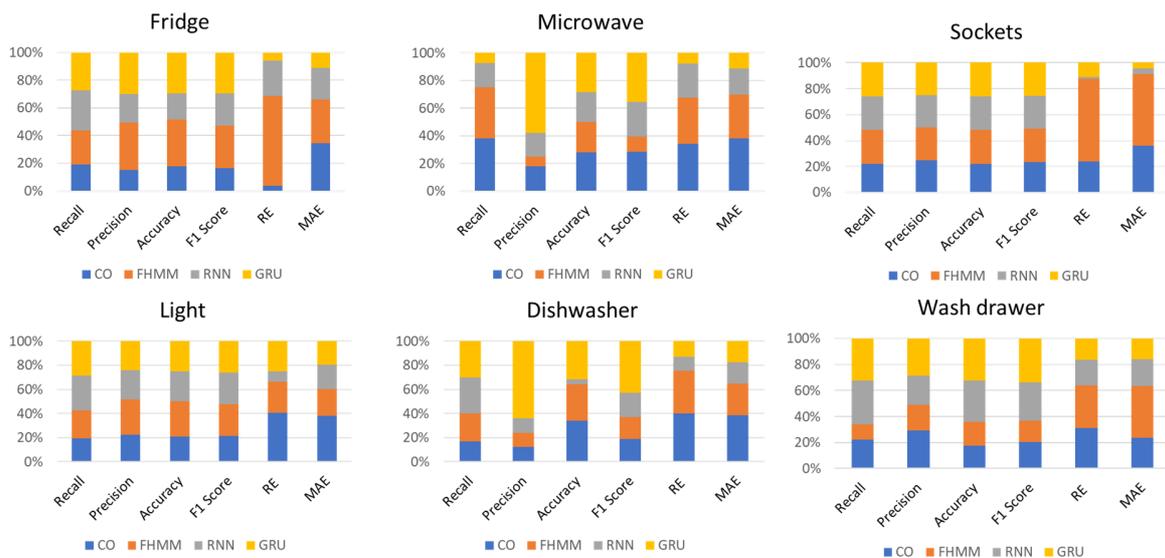


Fig. 5. Performance Evaluation of the NILM algorithms.

The proposed GRU shows a consistent performance in all the appliances, in terms of MAE. The overall MAE of the proposed model is 25.23, whereas, the MAE of CO, FHMM, and RNN is 61.88, 56.14, and 33.92 respectively. Thus, the proposed approach shows reduced MAE without being dependent on feature engineering approach. Similarly, the overall accuracy of CO, FHMM, RNN, and the proposed model is calculated as 66, 74, 64, and 91 respectively. FHMM shows better accuracy in fridge and light only, whereas, CO is better in the case of the dishwasher. This clearly indicates a better and consistent performance of the proposed model in terms of accuracy.

Beside baseline algorithms, the proposed model is also compared with few of the latest DNN approaches Convolution Sequence to point (Seq2point) [15], Convolution sequence to Sequence (Seq2Seq) [15], and Denoising Autoencoder (DAE) [17]. These DNN approaches are compared in terms of MAE and computational cost for two major appliances (fridge and microwave) of REDD's House1. The number of layers depicts the computational complexity and time required for power disaggregation. Table III indicates that the proposed E2E architecture shows better performance in comparison to the modern approaches in [15, 17], whereas this unified and less dense architecture is also computationally efficient and fast to be used in real time applications.

TABLE III. COMPARISON WITH LATEST DNN APPROACHES

Ref.	Algorithm	Layers	Comp. cost	Fridge MAE	Microwave MAE
[15]	Seq2Point	8	Medium	28.10	28.19
[15]	Seq2Seq	8	Medium	30.63	33.27
[17]	DAE	10	Highest	29.11	23.26
Proposed	E2E	6	lowest	22.12	23.19

IV. CONCLUSION

NILM is dominated by pipeline approaches with explicit feature extraction for load identification. The developments made in the domain of ML can eliminate the dependency on inherently low performing and computationally complex architectures. The proposed E2E ML based unified architecture using GRUs was evaluated on top 6 appliances of REDD. Comparative analysis of the proposed and previously reported algorithms showed that the proposed architecture has the ability to replace the feature-based load disaggregation approach. Since the proposed architecture performs better in all cases irrespective of the appliance type, it depicts that it can improve performance of the load disaggregator and due to its less computational demand and fast response time it can also be integrated in modern smart metering solutions.

REFERENCES

[1] M. Sun, F. M. Nakoty, Q. Liu, X. Liu, Y. Yang, and T. Shen, "Non-Intrusive Load Monitoring System Framework and Load Disaggregation Algorithms: A Survey," in *2019 International Conference on Advanced Mechatronic Systems (ICAMechS)*, Kusatsu, Japan, Aug. 2019, pp. 284–288, <https://doi.org/10.1109/ICAMechS.2019.8861646>.

[2] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992, <https://doi.org/10.1109/5.192069>.

[3] L. Pereira and N. Nunes, "Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review," *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 6, 2018, Art. no. e1265, <https://doi.org/10.1002/widm.1265>.

[4] J. Kolter and M. Johnson, "REDD: A Public Data Set for Energy Disaggregation Research," *Artificial Intelligence*, vol. 25, pp. 59–62, Jan. 2011.

[5] M. E. Berges, E. Goldman, H. S. Matthews, and L. Soibelman, "Enhancing Electricity Audits in Residential Buildings with Nonintrusive Load Monitoring," *Journal of Industrial Ecology*, vol. 14, no. 5, pp. 844–858, 2010, <https://doi.org/10.1111/j.1530-9290.2010.00280.x>.

[6] D. Luo, L. K. Norwood, S. R. Shaw, and S. Leeb, "Monitoring HVAC Equipment Electrical Loads from a Centralized Location: Methods and Field Test Results," *ASHRAE Transactions 2002*, vol. 108, no. 1, pp. 841–857, 2002.

[7] H. Altrabalsi, V. Stankovic, J. Liao, L. Stankovic, and Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK, "Low-complexity energy disaggregation using appliance load modelling," *AIMS Energy*, vol. 4, no. 1, pp. 1–21, 2016, <https://doi.org/10.3934/energy.2016.1.1>.

[8] J. Liao, G. Elafoudi, L. Stankovic, and V. Stankovic, "Non-intrusive appliance load monitoring using low-resolution smart meter data," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Venice, Italy, Nov. 2014, pp. 535–540, <https://doi.org/10.1109/SmartGridComm.2014.7007702>.

[9] W. Kong, Z. Y. Dong, D. J. Hill, J. Ma, J. H. Zhao, and F. J. Luo, "A Hierarchical Hidden Markov Model Framework for Home Appliance Modeling," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3079–3090, Jul. 2018, <https://doi.org/10.1109/TSG.2016.2626389>.

[10] D. Bajović, K. He, L. Stanković, D. Vukobratović, and V. Stanković, "Optimal Detection and Error Exponents for Hidden Semi-Markov Models," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 5, pp. 1077–1092, Oct. 2018, <https://doi.org/10.1109/JSTSP.2018.2851506>.

[11] R. Bonfigli, E. Principi, M. Fagiani, M. Severini, S. Squartini, and F. Piazza, "Non-intrusive load monitoring by using active and reactive power in additive Factorial Hidden Markov Models," *Applied Energy*, vol. 208, pp. 1590–1607, Dec. 2017, <https://doi.org/10.1016/j.apenergy.2017.08.203>.

[12] L. Mauch and B. Yang, "A new approach for supervised power disaggregation by using a deep recurrent LSTM network," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, FL, USA, Dec. 2015, pp. 63–67, <https://doi.org/10.1109/GlobalSIP.2015.7418157>.

[13] S. Makonin, F. Popowich, I. V. Bajić, B. Gill, and L. Bartram, "Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2575–2585, Nov. 2016, <https://doi.org/10.1109/TSG.2015.2494592>.

[14] A. F. Ebrahim and O. A. Mohammed, "Energy Disaggregation Based Deep Learning Techniques: A Pre-processing Stage to Enhance The Household Load Forecasting," in *2018 IEEE Industry Applications Society Annual Meeting (IAS)*, Portland, OR, USA, Sep. 2018, <https://doi.org/10.1109/IAS.2018.8544664>.

[15] C. Zhang, M. Zhong, Z. Wang, N. Goddard, and C. Sutton, "Sequence-to-point learning with neural networks for nonintrusive load monitoring," *arXiv:1612.09106 [cs, stat]*, Sep. 2017, Accessed: May 07, 2021. [Online]. Available: <http://arxiv.org/abs/1612.09106>.

[16] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis, "Multi-Channel Recurrent Convolutional Neural Networks for Energy Disaggregation," *IEEE Access*, vol. 7, pp. 81047–81056, 2019, <https://doi.org/10.1109/ACCESS.2019.2923742>.

[17] R. Bonfigli, A. Felicetti, E. Principi, M. Fagiani, S. Squartini, and F. Piazza, "Denoising autoencoders for Non-Intrusive Load Monitoring: Improvements and comparative evaluation," *Energy and Buildings*, vol. 158, pp. 1461–1474, Jan. 2018, <https://doi.org/10.1016/j.enbuild.2017.11.054>.

[18] V. Piccialli and A. Sudoso, "Improving Non-Intrusive Load Disaggregation through an Attention-Based Deep Neural Network,"

- Energies*, vol. 14, no. 4, Feb. 2021, Art. no. 847, <https://doi.org/10.3390/en14040847>.
- [19] L. B. Salah and F. Fourati, "Systems Modeling Using Deep Elman Neural Network," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3881–3886, Apr. 2019, <https://doi.org/10.48084/etasr.2455>.
- [20] S. Abid, M. Chtourou, and M. Djemel, "A Pruning Algorithm Based on Relevancy Index of Hidden Neurons Outputs," *Engineering, Technology & Applied Science Research*, vol. 6, no. 4, pp. 1067–1074, Aug. 2016, <https://doi.org/10.48084/etasr.704>.
- [21] J. Kelly and W. Knottenbelt, "Neural NILM: Deep Neural Networks Applied to Energy Disaggregation," in *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, New York, NY, USA, Nov. 2015, pp. 55–64, <https://doi.org/10.1145/2821650.2821672>.

Corruption Risk Analysis at the Project Planning Stage in the Iraqi Construction Sector using the Bowtie Methodology

Semaa Hazim Najim Alani
Civil Engineering Department
University of Baghdad
Baghdad, Iraq
S.Najim0901@coeng.uobaghdad.edu.iq

Ahmed Mohammed Raof Mahjoob
Civil Engineering Department
University of Baghdad
Baghdad, Iraq
ahmed.mahjoob@coeng.uobaghdad.edu.iq

Abstract-In this paper, the bowtie method was utilized by a multidisciplinary team in the Federal Board of Supreme Audit (FBSA) for the purpose of managing corruption risks threatening the Iraqi construction sector. Corruption in Iraq is a widespread phenomenon that threatens to degrade society and halt the wheel of economic development, so it must be reduced through appropriate strategies. A total of eleven corruption risks have been identified by the involved parties in corruption and were analyzed by using probability and impact matrix and their priority has been ranked. Bowtie analysis was conducted on four factors with high score risk in causing corruption in the planning stage. The number and effectiveness of the existing proactive measures to prevent threats from resulting in corruption and were examined for each threat.

Keywords-corruption risk; bowtie; prevention; risk analysis

I. INTRODUCTION

Corruption is an old phenomenon that threatens the administrative apparatus in Iraq. As it was stated in the reports of Transparency International [1], Iraq is third among 180 countries in corruption and the phenomenon is increasing. The phenomenon of corruption largely stands in the way of the process of construction and progress at all economic levels. It costs money, time, energy, and impedes the performance of responsibilities and the fulfillment of jobs and services. Corruption can be defined as a social habit which is deep-rooted in mankind's history. It is considered to be comparable to other crime types, which occur likely in the procurement of works through local authorities along with the complexity to monitor the expenditures of a project [2]. Authors in [3] provided a thorough overview of the corruption causes from certain studies in identified construction management journals for addressing the already-indicated issues. Forty-four corruption causes were indicated from a total of 37 publications and were examined with regard to current corruption factors, thematic categorizations of observed variables, and yearly trends of publications. The major causes were bad professionalism, ethical standards, close relationships, bad working and industrial conditions, insufficient sanctions, and bad role models. Authors in [4] identified factors which lead to fraudulent practices in every stage of the construction life

cycle. Those extracted fraudulent aspects have been mapped into a matrix form for checking the factor's similarities. A total of 42 contributing factors were verified and divided into 5 phases of the construction project life-cycle. With regard to the frequency, the main factors with regard to each stage were project approvals in the stage of planning, collusions between public officers and tenderers in the stage of design and tendering, insufficient compliances with the design for audit and report in the stage of construction, avoiding difficulties in the contract inspections in the stage of finishing, and utilizing sub-standard materials and services in the stage of maintenance.

II. LITERATURE REVIEW

Corruption is one of the biggest problems the governments face and it stands as an obstacle to sustainable development. Most projects in Iraq have suffered high deviations in cost, time, and quality which lead to wasting public money through exploiting public officers or authorities to achieve personal benefits that destroy the country's economy. Authors in [5] proposed a model for the performance and application control of rational decision-making with the best use of available resources and ongoing performance monitoring at all stages of the implementation of projects in order to eliminate aspects of financial and administrative corruption and reduce wastage in one of the most important formations carried out by the government in Iraq. Authors in [6] developed an Artificial Neural Network (ANN) model for assessing the collusion risks in the construction projects in China. The ANN has been utilized in a real life metro project, while the evaluated results established a complete framework regarding the collusive practices in the construction projects in China, that includes a total of 20 collusive practices associated with all the contracting entities. Authors in [7] presented the critical risk factors for design-build projects in the construction industry in Vietnam, factors affecting the on schedule implementation completion. The top 3 critical risk factors were: delays in project approval and licensing, interest rate fluctuations, and design or technical specifications deficiency and changes. Authors in [8] stated time overrun in construction projects is a very dangerous issue in developed countries. A quantitative

Corresponding author: Semaa Hazim Najim Alani

approach was adopted to present the main causes of time overrun in the construction projects in Pakistan and the RIW (Relative Importance Weight) approach was used for collecting data.

The precise beginnings of the bowtie method are unclear [9], but it probably was originated in the 1970s via the Imperial Chemical Industries (ICI). The first company integrating bowties totally in business practices is the Royal Dutch/Shell Group. Now, the use of bowties has been spread between nations, industries, and companies. Authors in [10] indicated that the bowtie analysis is a qualitative risk-assessment methodology providing an approach of efficient communication in complicated risk situations in simple graphical formats that shows the relations between damage escalation potential and the unwanted events' causes.

III. METHODOLOGY

The methodology that will be used in this paper is as follows:

- Identification of the corruption risks during the conception (planning) stage, which actually occurred in some completed construction projects by reviewing the Quarterly reports of the FBSA of projects from 2004-2013 and after some interviews with construction industry experts.
- Distribution of a prepared questionnaire on the FBSA staff in order to determine the probability and the effect of such risks.
- Implementing qualitative risk analysis on the corruption risks by using a probability - impact matrix to identify the priority of these risks.
- Application of bowtie methodology on these risks (threats).

IV. RISK RATING

A. Risk Probability

For each of the risk occurrences, the probability must be set at a single value which is representing the best team judgment from the provided available data. In addition, the team must not be considering the effect while assessing the probability. The probability values and terms utilized in a 5-point scale to suit the likelihood of getting the risk can be seen in Table I [11]. Other classifications can be used according to what suits a project.

TABLE I. RISK PROBABILITY

Term	Value	Risk probability (%)
(1) Very high	0.9	> 70
(2) High	0.7	50-70
(3) Moderate	0.5	30-50
(4) Low	0.3	10-30
(5) Very low	0.1	0-10

B. Risk Impact

The risk impact is measuring the consequences on the project when a risk occurs. A scale that is comparable to the one of risk probabilities should be set up with regard to risk impact (Table II [12]).

TABLE II. RISK IMPACT

Term	Value
(1) Very high	0.8
(2) High	0.4
(3) Moderate	0.2
(4) Low	0.1
(5) Very low	0.05

C. Probability and Impact Matrix

A probability and impact matrix has been utilized for looking up a combined risk rating, also referred to as the risk score, with regard to specific combinations related to probability and impact values. In addition, the score was utilized for prioritizing the risks for later activities of risk management. A graphical probability and impact matrix is efficiently communicating the relative severity regarding risk scores. The module is addressing the details regarding evaluating the risks and impacts, along with the way for applying such measures. Furthermore, a measure regarding the risk effect might be evaluated via multiplying risk probability (P) by risk impact (I) for producing risk score (RS) with regard to each one of the risks.

$$RS=P \times I \quad (1)$$

In the heat map of Figure 1, the resulting high risks are colored red, the moderate risks yellow, and the low risks green [12].

Probability	Threats				
	Risk Score =Probability × Impact				
0.9 Very high	0.05	0.09	0.18	0.38	0.72
0.7 High	0.04	0.07	0.14	0.28	0.56
0.5 Moderate	0.03	0.05	0.10	0.12	0.4
0.3 Low	0.02	0.03	0.06	0.12	0.24
0.1 Very low	0.01	0.01	0.02	0.04	0.08
	0.05 Very low	0.1 Low	0.2 Moderate	0.4 High	0.8 Very high

High (Red) / Med (YEL)/ low (GRN)

Fig. 1. Probability and impact matrix (heat map).

D. The Questionnaire

A questionnaire survey has been carried out for obtaining perception-based data related to corruption measurement items from 2 viewpoints, which are (1) probability (the occurrence possibility regarding each one of the measurement items), and (2) severity (the consequence's impact regarding each one of the measurement items), with the use of a 5-point Likert scale from 1 for very high to 5 for very low. About 60 questionnaire forms about corruption risk were distributed by hand to resident engineer's offices of FBSA projects in Iraq. Twenty two answered forms were not considered because they were partially answered and 8 were not received. A total of 30 answered questionnaires were considered. The responders belong to senior management staff, having 8 to 40 years of

experience. Table III shows some details about the questionnaire responders whose responses were considered in this study.

E. Responders Characteristics Summary

Table III shows the experience of the respondents, Table IV their level of education, and Table V the scientific areas of their education.

TABLE III. RESPONDENTS' EXPERIENCE

No.	Experience (years)	Frequency	%
1	8-14	8	27
2	15-18	15	50
3	20-22	3	10
4	36-40	4	13
	Total	30	100%

TABLE IV. RESPONDENTS' LEVEL OF EDUCATION

No.	Level of education	Frequency	%
1	B.Sc.	18	60
2	Diploma	1	3
3	M.Sc.	7	23
4	Ph.D.	2	7
5	CPA	2	7
	total	30	100%

TABLE V. RESPONDENTS' SCIENTIFIC AREAS

No.	Area	Frequency	%
1	Civil Engineering	15	50
2	Mechanical Engineering	4	13.33
3	Electrical Engineering	2	6.67
4	Architecture Engineering	1	3.33
5	Law	5	16.67
6	Control and Auditing	2	6.67
7	Accounting	1	3.33
	Total	30	100

V. RESULTS

The areas of risk concern at the stage of planning came from project documents and are provided in Table VI and Figure 2.

We can see that the CR3 (the estimated cost of the project compared to the building area was exaggerated) is ranked at the first place with a risk score of 0.4196, which makes it a high risk rating. This result is in contrast with the findings in [15]. The second ranked risk is CR2 (there is a lack of a careful study of the estimated cost of the project within the feasibility study report), scoring 0.3067, which also makes it a high risk which leads to inaccurate cost estimations of the project. This result is in line with the findings of [14]. The third risk CR6 (there is no ownership of the land on which the project is located, and some lands were transgressed) scored 0.1956, which is also a high risk rating, in line with the findings in [2]. The fourth risk is CR1 (there is no technical and economic feasibility study to the project) that scored 0.182, which is a high risk rating, and is one of the requirements for approving the financial allocations for the implementation of any project. This result is in contradiction to Article (A-3) of the instructions for implementation of governmental contracts No. 1 in 2008 [13] and is line with the findings of [15].

TABLE VI. CORRUPTION RISKS RATING IN THE PLANNING STAGE

ID	Involved parties	Corruption risk	Risk score	Risk rating	Rank
CR1	Client	No technical and economic feasibility study to the project.	0.182	High	4
CR2	Client	There is a lack of a careful study of the estimated cost of the project within the feasibility study report.	0.3067	High	2
CR3	Client	The estimated cost of the project compared to the building area is exaggerated.	0.4196	High	1
CR4	Client	The technical and economic feasibility study has not been sent out to the Ministry of Planning for the purpose of obtaining approval to include the project in the investment budget table.	0.0472	Low	9
CR5	Client	There is no holistic vision in the department about the project requirements to be translated and escalated according to a careful planning process into engineering plans.	0.1469	Moderate	7
CR6	Client	There is no ownership of the land on which the project is located, and some lands were transgressed	0.1956	High	3
CR7	Client	The work is divided into several contracts and a feasibility study is prepared with an estimated cost for each contract, even though the project was included in the ministry's plan according to one feasibility study and one holistic cost.	0.176	Moderate	5
CR8	Client	A technical study has not been prepared about the contracting party's ability to provide the necessary staff.	0.0261	Low	11
CR9	Client, consultant	Soil was investigated for the project long before the work was carried out. This leads to geological changes to the site's soil, and this causes these investigations mismatch the reality of the project.	0.1493	Moderate	6
CR10	Client	Project planning that requires a true contemplation and practical forecasting is poor.	0.0496	Low	8
CR11	Client	The project in the investment budget is included before specifying estimated and total costs of the project that was later completed.	0.0271	Low	10

The fifth risk is CR7 (the work is divided into several contracts and a feasibility study is prepared with an estimated cost for each contract, even though the project was included in

the ministry's plan according to one feasibility study and one holistic cost) that scored 0.176 which is a moderate risk rating. This result is in line with the findings in [2]. The sixth risk is CR9 (soil was investigated for the project long before the work was carried out), scoring 0.1493 which is a moderate risk rating. This result is in line with the findings in [2]. The seventh risk is CR5 (there is no holistic vision in the department about the project requirements to be translated and escalated according to a careful planning process into engineering plans) that scored 0.1469 which is a moderate risk rating. This result is in line with the findings in [15].

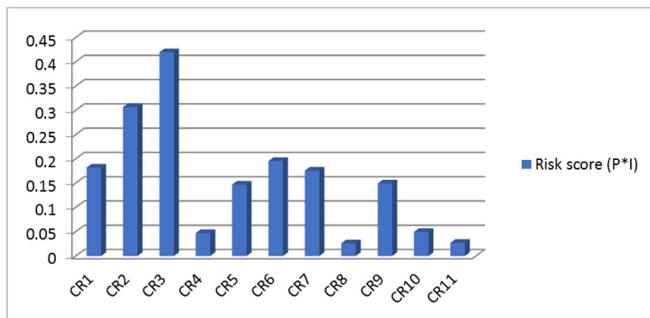


Fig. 2. The risk score for the corruption risk in planning stage.

The other risks have low risks ratings. CR10 (project planning that requires a true contemplation and practical forecasting is poor) scored 0.0496, CR4 (the technical and economic feasibility study has not been sent out to the Ministry of Planning for the purpose of obtaining approval to include the project in the investment budget table) scored 0.0472 in line with the findings in [2], CR11 (the project in the investment budget is included before specifying estimated and total costs of the project that was later completed-in contradiction to the instructions for implementation of governmental contracts (No. 1) in 2008 that stated that an updated study on the estimated cost or the required contract to be executed within the feasibility study for the purpose is to be used as a standard when analyzing bids and awarding contracts) scored 0.0271, and CR8 (a technical study has not been prepared about contracting party's ability to provide the necessary staff) scored 0.0261.

VI. APPLICATION OF BOWTIE XP ON THE CORRUPTION RISK

After identifying the probability and impact (consequence) for each risk, the corruption risks in the planning stage of construction project were ranked. It was adopted that the risks with low and moderate rating can usually be managed using routine procedures, and the BowTie XP software [16] was applied on the risks with high rating. In a bowtie, the representation is that a threat (on diagram's left hand side) causes the top event (hazard release), which results in a consequence (harm to someone of something). We then have the opportunity to put in preventive barriers on diagram's left hand side to prevent the hazards, and barriers on diagram's right hand side to mitigate the harm from occurring. Practical recommendations were led via applying bowtie methodology, which leads to excellent understanding regarding the hazard

protection and prevention needed for corruption risks in the sector construction. The causal factors that were used in BowTie XP to manage corruption risks threatening the construction sector in the planning stage were identified by a multidisciplinary team in the FBSA as the prevention barrier, effectiveness, and barrier type for each causal factor (threat), and corrective barrier, effectiveness, and barrier type for each consequence as shown in Figure 3.

To the best of our knowledge, the current research is the first one that uses the application of bowtie methodology with corruption risk in the construction sector.

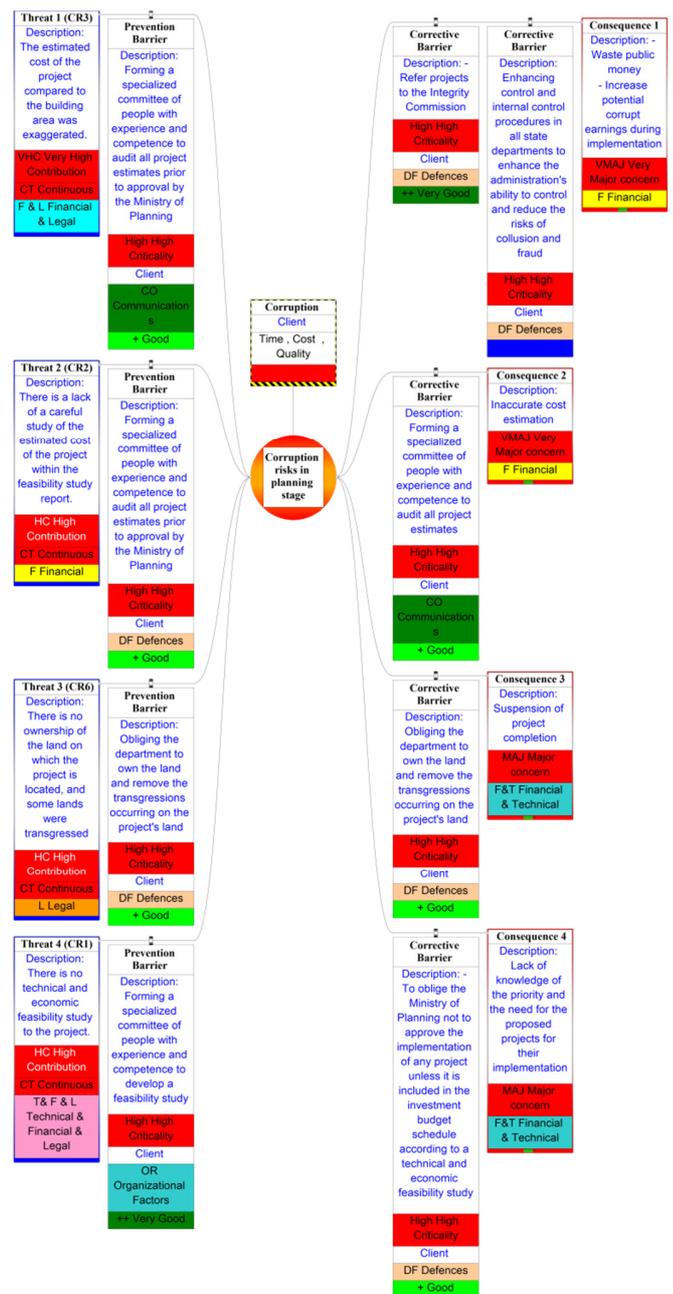


Fig. 3. Bowtie diagram of corruption risks in the planning stage .

VII. CONCLUSION

The bowtie methodology is considered to have high flexibility regarding corruption risk management. At the same time, its adaptability might be improved by the use of some adequate tools. On the basis of the system approach, bowtie methodology helps auditors and managers responsible for creating the project plan with the necessary requirements from the estimation cost of the project, offering a technical and economic feasibility study. The conception of threats, consequences, and barriers in the bowtie diagrams facilitate the comprehension regarding the protection and prevention actions needed for corruption risks. Typically, the bowtie has the ability to be vital in managing the corruption risks that we can process before they appear by taking proactive actions for these threats. The bowtie methodology was used to demonstrate the effectiveness of corruption risk management in construction projects in Iraq. The use of BowTie software assists in the analysis and management of hazard and risks in any sector. It provides reports and documents all information that will be used in risk management and supports the audits in the government institutions.

VIII. RISK MANAGEMENT RECOMMENDATIONS

A set of recommendations regarding risk management for corruption-related risks follows:

- A study combining bowtie methodology and simulation model should be conducted.
- There is a need to identify the corruption risks in other stages of the construction sector and appropriate quantitative risk analysis should be conducted.
- The application of the bowtie methodology should be expanded in other sectors such as the safety risk management in construction projects.

ACKNOWLEDGEMENT

This paper was produced in cooperation with the Department of Civil Engineering, University of Baghdad.

REFERENCES

- [1] "Corruption Perception Index (CPI) 2015," *Transparency International*. <https://www.transparency.org/en/cpi/2015> (accessed May 08, 2021).
- [2] P. Zou, "Strategies for Minimizing Corruption in the Construction Industry in China," *Journal of Construction in Developing Countries*, vol. 11, no. 2, pp. 15–29, Dec. 2006.
- [3] E. K. Owusu, A. P. C. Chan, and M. Shan, "Causal Factors of Corruption in Construction Project Management: An Overview," *Science and Engineering Ethics*, vol. 25, no. 1, pp. 1–31, Feb. 2019, <https://doi.org/10.1007/s11948-017-0002-4>.
- [4] N. A. I. M. Saim, I. A. Rahman, and M. F. Ismail, "Factors Contributing to Fraudulent Practices in Construction Project Life Cycle," *International Journal of Engineering & Technology*, vol. 7, pp. 647–651, Aug. 2018, <https://doi.org/10.14419/ijet.v7i3.20.22965>.
- [5] T. Y. Al-Tai, "The role of performance audit in rationalizing investment decisions for infrastructure projects," Ph.D. dissertation, Baghdad University, Baghdad, Iraq, 2014.
- [6] M. Shan, Y. Le, K. T. W. Yiu, A. P. C. Chan, Y. Hu, and Y. Zhou, "Assessing collusion risks in managing construction projects using artificial neural network," *Technological and Economic Development of Economy*, vol. 24, no. 5, pp. 2003–2025, Oct. 2018, <https://doi.org/10.3846/20294913.2017.1303648>.
- [7] P. T. Nguyen and P. C. Nguyen, "Risk Management in Engineering and Construction: A Case Study in Design-Build Projects in Vietnam," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5237–5241, Feb. 2020, <https://doi.org/10.48084/etasr.3286>.
- [8] F. A. Soomro, M. J. Memon, A. F. Chandio, S. Sohu, and R. Soomro, "Causes of Time Overrun in Construction of Building Projects in Pakistan," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3762–3764, Feb. 2019, <https://doi.org/10.48084/etasr.2449>.
- [9] A. Lidstone, "The Use of Bowties to Aid the Management of Risk," 2017.
- [10] I. Voicu, F. V. Panaitescu, M. Panaitescu, L. G. Dumitrescu, and M. Turof, "Risk management with Bowtie diagrams," *IOP Conference Series: Materials Science and Engineering*, vol. 400, p. 082021, Sep. 2018, <https://doi.org/10.1088/1757-899X/400/8/082021>.
- [11] D. Allen, *Risk Management in Business*. Cambridge, UK: Cambridge University Press, 1995.
- [12] *PMBOK® Guide - A Guide to the Project Management Body of Knowledge*, 6th ed. PMI (Project Management Institute), 2107.
- [13] "Ministry of Planning Government Implementation Instructions No. (1) for the year 2008." Ministry of Planning, 2008, [Online]. Available: <https://mop.gov.iq/static/uploads/7/pdf/151012302174946107c03fdff610a3ea1241cbc53a2--file1.pdf>.
- [14] M. Sohail and S. Cavill, "Accountability to Prevent Corruption in Construction Projects," *Journal of Construction Engineering and Management*, vol. 134, no. 9, pp. 729–738, Sep. 2008, [https://doi.org/10.1061/\(ASCE\)0733-9364\(2008\)134:9\(729\)](https://doi.org/10.1061/(ASCE)0733-9364(2008)134:9(729)).
- [15] J. Hawkins, "How to Note: Reducing corruption in infrastructure sectors," Department for International Development, Jan. 2013, http://dx.doi.org/10.12774/eod_cr.may2013.hawkins.
- [16] "BowTieXP Virtual Risk Assessment," *CGE Risk*. <https://www.cgerisk.com/products/bowtiexp/> (accessed May 08, 2021).

Influence of Starch Content on the Thermal and Viscoelastic Properties of Syndiotactic Polypropylene/Starch Composites

Naveed Ahmad

Department of Chemical and Material Engineering
College of Engineering
Northern Border University
Arar, Saudi Arabia
naveed.ahmad@nbu.edu.sa

Farooq Ahmad

Department of Chemical and Material Engineering
College of Engineering
Northern Border University
Arar, Saudi Arabia
farooq.ahmad@nbu.edu.sa

Ibrahim Alenezi

Department of Chemical and Material Engineering College of Engineering
Northern Border University
Arar, Saudi Arabia
i.alenezi@nbu.edu.sa

Abstract—In this study, syndiotactic Polypropylene/Starch (sPP/starch) composites were prepared using a solution mixing technique. The thermal characterization was performed using Differential Scanning Calorimetry (DSC), and the melting point was measured for all polymer composites. The thermal degradation temperature was measured using thermal gravimetric analysis. The viscoelastic measurements were performed using the Atomic Rheometric Expansion System (ARES). Both melting point and thermal degradation temperatures were found to decrease with increasing starch content. Moreover, the elastic modulus was found to decrease when the starch content increased.

Keywords—aeration; melting point; sPP/starch composites; thermal degradation temperature; elastic modulus; frequency sweep test

I. INTRODUCTION

Polymer composites are synthesized by embedding natural or synthetic materials into a polymer to increase its desired properties [1, 2]. Several studies have been conducted on enhancing the properties of polymer products [3-23]. Polymer composites with different combinations of filler and polymer are studied in order to obtain the desired and targeted properties of polymer products [3]. However, the final desired microstructure and macrostructure properties of polymers and composites depend on the nature, amount, geometry, and interfacial interactions of the components [4]. Rheology is used as a tool to study a polymer's or a polymer composite's microstructure, as the rheological response is highly sensitive to the molecular structure of polymers and their composites [5-9]. While several studies have been conducted on the synthesis and the rheology of polymer and polymer/clay composites [10-

18], only a few have targeted the viscoelasticity of polymer/clay composites.

The rheological properties of polyethylene oxide/montmorillonite clay gels and multilayered films were studied in [14], using an Atomic Rheometric Expansion system (ARE-2 rheometer), and noticing a slight increase in both the loss modulus (G'') and the storage modulus (G') when increasing the clay content. The thermal properties were examined using Differential Scanning Calorimetry (DSC) and Thermal Gravimetric Analysis (TGA), while the results were related to the microstructural properties of the composites and the films. The effect of clay content on the rheology of polypropylene/clay nanocomposites was investigated in [15], examining both linear and non-linear rheology, finding that loss modulus (G''), storage modulus (G'), and dynamic viscosities increased monotonically with organophilic montmorillonite nanocomposites. The viscoelastic properties of syndiotactic polypropylene (sPP) were investigated in [16], examining the effect of the degree of syndiotacticity on the rheological parameters, including plateau modulus and entanglement molecular weight. This study indicated that increasing the degree of syndiotacticity enhanced the plateau modulus causing a decrease in the molecular weight between entanglements. In [17], the influence of clay contents on plateau modulus and entanglement molecular weight was studied on polypropylene/clay composites, concluding that an increase in the clay content increased the plateau modulus and consequently decreased the entanglement molecular weight. The rheological properties of a drilling fluid polymer treating agent named Driscal-D using a Fann 50SL rheometer were investigated in [24]. The effects of adding polymer, electrolyte, clay type, and antioxidant on the rheological properties of a

Corresponding author: Naveed Ahmad

Driscald-D solution were examined, finding that solution's viscosity tends to increase when increasing polymer addition and decreased in the presence of an electrolyte. Furthermore, adding clay in a Driscald-D solution enhanced its rheological properties, which could improve cuttings transportation.

The current study examines the viscoelasticity of syndiotactic polypropylene/starch composites (sPP/starch), with starch content ranging from 0 to 8%, using an Atomic Rheometric Expansion System (ARES) to investigate the effect of starch content on the melt rheological response of composites.

II. METHODOLOGY

A. Polymer Composite Synthesis

1) Materials

Syndiotactic Polypropylene (sPP) having a 60% degree of syndiotacticity, supplied by Zhongwei Industrial Co. Ltd., was used as a matrix. Starch, originating from potato, was used as the filler. Xylene, provided by Shanxi Zhongwei Industrial Co. Ltd., was used as a solvent in mixing. The samples' details are listed in Table I.

TABLE I. SAMPLES' DETAILS

Number of samples	Name of samples	Percentage of starch content	Degree of syndiotacticity (%rrr)
1	sPP/S-8	8%	
2	sPP-6	6%	
3	sPP/S-4	4%	
4	sPP/S-2	2%	
5	sPP	0	60

2) Preparation of Composite Samples

Solution mixing was applied to create a set of syndiotactic polypropylene/starch composite samples with four different starch content proportions. The starch was accurately weighed and mixed with the polypropylene matrix using xylene as a solvent in a 1000ml beaker. The mixing temperature was kept at 120°C using a heating mantle until all syndiotactic polypropylene granules dissolved in a clear solution. The solvent containing syndiotactic polypropylene was initially shaken vigorously by hand and then stirred by a high-speed electrical stirrer keeping the heating conditions. The starch, dissolved in xylene, was added slowly to the syndiotactic polypropylene and xylene mixture. Vigorous and continuous stirring helped the uniformity of starch mixing and the intercalation of sPP to starch layers. The beaker operated for a few hours to evaporate the xylene. Then, the size of the prepared polymer-starch composite was mechanically reduced. After the completion of xylene evaporation, the composite samples were created by compression molding at 170°C temperature and 7Psi pressure for one and a half hour [10].

B. Polymer Composite Analysis

1) Differential Scanning Calorimetry (DSC)

The melting temperature (T_m) of all samples was measured using Differential Scanning Calorimetry (DSC). The samples, weighing 4mg, were heated up to 350°C at a constant heating

rate of 40°C/min. During the heating cycle, the melting point peak was obtained. Afterward, the sample was kept under this annealing process for 7 minutes to eliminate crystallinity and remove thermal history. The analysis was performed using the instrument software TA-60 to find the melting temperature (T_m) of all samples.

2) Rheometry

The sPP/starch composites were pressed using Hot Press to prepare a uniform film for each composite. An Atomic Rheometric Expansion System (ARES) (TA Instruments), having a geometry of 8mm diameter parallel plate, was used to perform the viscoelastic tests under nitrogen atmosphere to reduce sample degradation. All experiments were carried out at a different gap depending on the sample's thickness. At first, strain sweep and stability tests were carried out. After choosing the suitable strain value (0-10%), dynamic frequency sweep tests were performed at temperatures ranging from 100°C to 200°C to obtain both elastic and viscous moduli at different frequencies and the master curves using the time-temperature superposition principle. No tests were performed below 100°C due to the faster crystallization kinetic.

III. RESULTS AND DISCUSSION

Frequency sweep tests were performed on all samples. The frequency sweep tests for samples having 2% and 6% starch contents are shown in Figures 1 and 2 respectively. It can be noted that the elastic modulus increased along with the frequency range (0 to 100rad/s), while the viscous modulus tended to decrease. Figure 1 shows a cross-over frequency at almost 1rad/s indicating that relaxation time was a little greater than 1s. Figure 2 shows that viscous was greater than elastic modulus at lower frequencies, while the elastic modulus was higher than viscous at higher frequencies. The cross-over frequency of the 6% sample was found to be greater than the one in the 2% sample, indicating that relaxation time for the 6% sample was less than for the 2%. This result also indicates that increasing starch content decreased the elastic and increased the viscous response. In other words, a strength decrease occurs when the starch content increases.

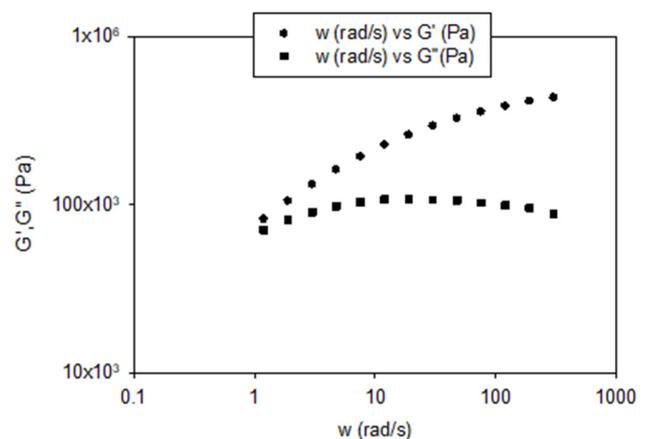


Fig. 1. Frequency sweep test for sPP/S-2 at 200°C.

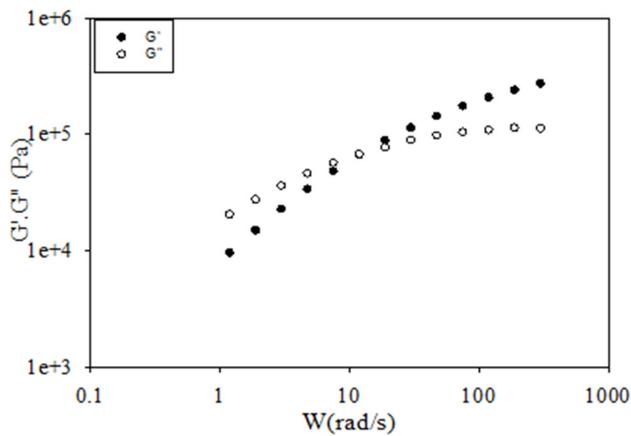


Fig. 2. Frequency sweep test for sPP/S-6 at 200°C.

A. Effect of Starch on Melting Point

The effect of starch on the melting point temperature of sPP/starch composites is shown in Figure 3. As it can be noted, increasing starch content decreased the melting point temperature of sPP/starch composites. The melting point of the composite was actually the melting point of starch. Starch's thermal stability is low compared to the syndiotactic polypropylene polymer. Therefore, increasing starch decreases the melting point temperature of sPP/starch composites.

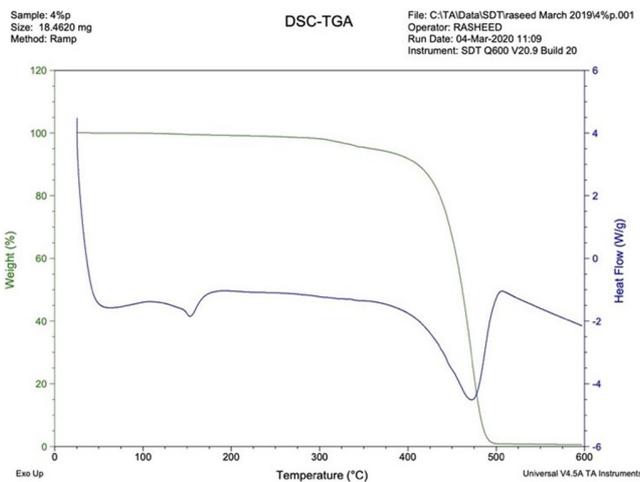


Fig. 3. Relationship between starch and melting point (T_m).

B. Effect of Starch on Thermal Degradation Temperature

The thermal degradation temperature of all composites was measured using Thermal Gravimetric Analysis (TGA). The TGA analysis for the sPP sample containing 4% starch is shown in Figure 4. This Figure depicts two drops in mass: The first drop indicates the removal of water from the composite sample, i.e. the dryness of the sample, while the second drop indicates the degradation of the sample. Thermal degradation temperature decreased when increasing starch content, as shown in Figure 5.

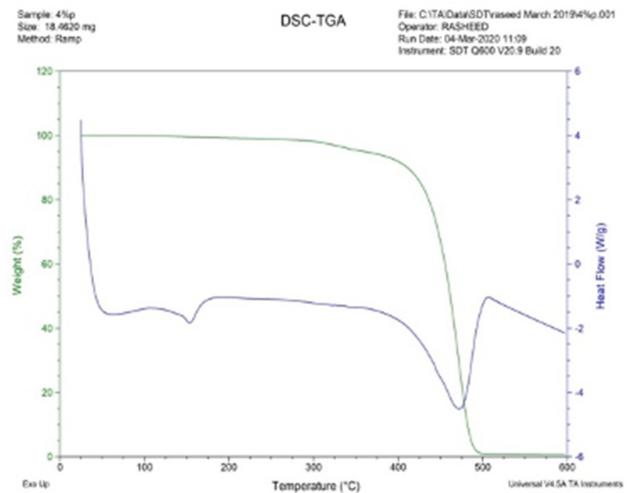


Fig. 4. FDSC and TGA analysis of sPP/S-4.

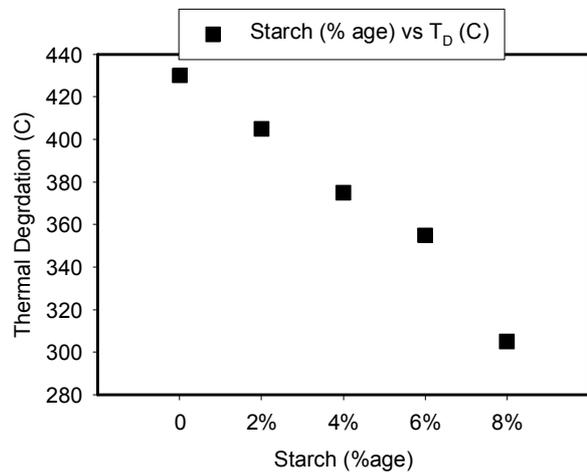


Fig. 5. Relationship of thermal degradation temperature and starch content.

C. Effect of Starch on Elastic Modulus

The elastic modulus was calculated using frequency sweep tests for all samples. The maximum value for each sample was noted. The elastic modulus indicates the mechanical response of a sample, and it was examined as a function of starch content. Increasing starch content decreased the elastic modulus, as shown in Figure 6. Therefore, increasing starch content decreased the strength of the composites. This result comes in agreement with the findings of [17], which investigated the effects of clay content on the thermal and rheological properties of sPP/clay composites. Clay content and starch affected the rheological and thermal properties of sPP/clay composites oppositely. Clay content increased the melting point temperature and the plateau modulus of sPP/clay composites. On the other hand, starch decreased both the thermal properties and the elastic modulus. An increase in clay content increased the mechanical response of the sPP/clay composites, while an increase in starch content decreased the mechanical and increased the viscous response of syndiotactic polypropylene/starch composites.

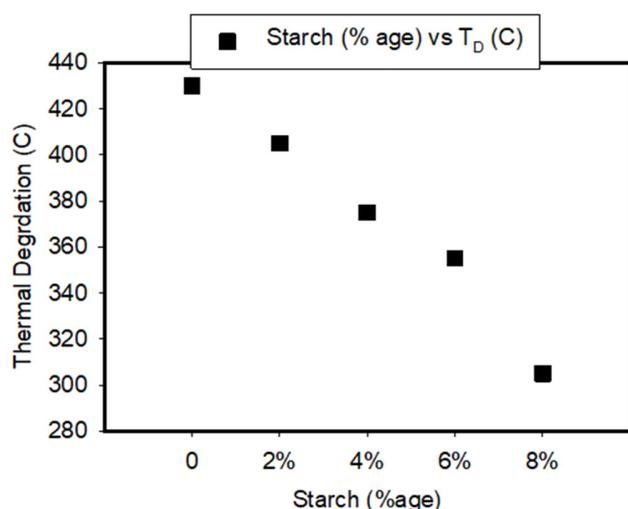


Fig. 6. Relationship between starch contents and elastic modulus.

IV. CONCLUSION

The viscoelastic and thermal analysis of sPP/starch on melting point, thermal degradation temperature, and elastic modulus depend on starch content. This study's findings showed that viscoelasticity is a powerful tool to investigate the microstructure and the chain parameters of polymer/starch composites. Thermal and viscoelastic properties were found sensitive to starch content proportions. Starch alters both the melting point and the thermal degradation temperatures, as increasing starch content decreased both of them. The same trend was found for the elastic modulus, as it decreased when the starch content increased. In conclusion, increasing starch content results in a decrease of the mechanical and elastic responses of the composite.

ACKNOWLEDGMENT

The authors wish to acknowledge the approval and the support of this research study by the grant no.7178-ENG-2017-1-7-F from the Deanship of Scientific Research in Northern Border University, Arar, K.S.A.

REFERENCES

- [1] D. J. Arriola, E. M. Carnahan, P. D. Hustad, R. L. Kuhlman, and T. T. Wenzel, "Catalytic Production of Olefin Block Copolymers via Chain Shuttling Polymerization," *Science*, vol. 312, no. 5774, pp. 714–719, May 2006, <https://doi.org/10.1126/science.1125268>.
- [2] V. C. Gibson, "Shuttling Polyolefins to a New Materials Dimension," *Science*, vol. 312, no. 5774, pp. 703–704, May 2006, <https://doi.org/10.1126/science.1127258>.
- [3] A. Ghanbari, M.-C. Heuzey, P. J. Carreau, and M.-T. Ton-That, "Morphological and rheological properties of PET/clay nanocomposites," *Rheologica Acta*, vol. 52, no. 1, pp. 59–74, Jan. 2013, <https://doi.org/10.1007/s00397-012-0667-1>.
- [4] G. K. Jatav, R. Mukhopadhyay, and N. De, "Characterization of Swelling Behaviour of Nanoclay Composite," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 2, no. 5, pp. 1560–1563, 2013.
- [5] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*. Oxford, UK: Clarendon Press, 1988.
- [6] L. J. Fetters, D. J. Lohse, D. Richter, T. A. Witten, and A. Zirkel, "Connection between Polymer Molecular Weight, Density, Chain Dimensions, and Melt Viscoelastic Properties | *Macromolecules*," *Macromolecules*, vol. 27, no. 17, pp. 1639–4647, Aug. 1994, <https://doi.org/10.1021/ma00095a001>.
- [7] A. Eckstein *et al.*, "Determination of Plateau Moduli and Entanglement Molecular Weights of Isotactic, Syndiotactic, and Atactic Polypropylenes Synthesized with Metallocene Catalysts," *Macromolecules*, vol. 31, no. 4, pp. 1335–1340, Feb. 1998, <https://doi.org/10.1021/ma971270d>.
- [8] J. M. Carella, W. W. Graessley, and L. J. Fetters, "Effects of chain microstructure on the viscoelastic properties of linear polymer melts: polybutadienes and hydrogenated polybutadienes," *Macromolecules*, vol. 17, no. 12, pp. 2775–2786, Dec. 1984, <https://doi.org/10.1021/ma00142a059>.
- [9] S. Bagheri-Kazemabad *et al.*, "Morphology, rheology and mechanical properties of polypropylene/ethylene-octene copolymer/clay nanocomposites: Effects of the compatibilizer," *Composites Science and Technology*, vol. 72, no. 14, pp. 1697–1704, Sep. 2012, <https://doi.org/10.1016/j.compscitech.2012.06.007>.
- [10] M. Sarkar, K. Dana, S. Ghatak, and A. Banerjee, "Polypropylene-clay composite prepared from Indian bentonite," *Bulletin of Materials Science*, vol. 31, no. 1, pp. 23–28, Feb. 2008, <https://doi.org/10.1007/s12034-008-0005-5>.
- [11] Y. Xiang, Z. Peng, and D. Chen, "A new polymer/clay nano-composite hydrogel with improved response rate and tensile mechanical properties," *European Polymer Journal*, vol. 42, no. 9, pp. 2125–2132, Sep. 2006, <https://doi.org/10.1016/j.eurpolymj.2006.04.003>.
- [12] C. Liu, J. He, E. van Ruymbeke, R. Keunings, and C. Bailly, "Evaluation of different methods for the determination of the plateau modulus and the entanglement molecular weight," *Polymer*, vol. 47, no. 13, pp. 4461–4479, Jun. 2006, <https://doi.org/10.1016/j.polymer.2006.04.054>.
- [13] J. D. Ferry, *Viscoelastic Properties of Polymers*, 3rd ed. New York, NY, USA: John Wiley & Sons, 1980.
- [14] E. Stefanescu, "A Study of Rheological and Thermodynamic Properties of Polymer-Clay Gels and Multilayered Films," Ph.D. dissertation, Louisiana State University, 2008.
- [15] S.-Y. Gu, J. Ren, and Q.-F. Wang, "Rheology of poly(propylene)/clay nanocomposites," *Journal of Applied Polymer Science*, vol. 91, no. 4, pp. 2427–2434, 2004, <https://doi.org/10.1002/app.13403>.
- [16] N. Ahmad, R. Di Girolamo, F. Auriemma, C. De Rosa, and N. Grizzuti, "Relations between Stereoregularity and Melt Viscoelasticity of Syndiotactic Polypropylene," *Macromolecules*, vol. 46, no. 19, pp. 7940–7946, Oct. 2013, <https://doi.org/10.1021/ma401469a>.
- [17] N. Ahmad and E. Fouad, "Influence of Clay Contents on Rheology of Syndiotactic Polypropylene/Clay Composites," *Arabian Journal for Science and Engineering*, vol. 42, no. 4, pp. 1537–1543, Apr. 2017, <https://doi.org/10.1007/s13369-016-2389-7>.
- [18] N. Ahmad, E. Fouad, and F. Ahmad, "Effect of Shear Flow on Crystallization of Syndiotactic Polypropylene/Clay Composites," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3108–3112, Aug. 2018, <https://doi.org/10.48084/etasr.2079>.
- [19] J. Ren, A. S. Silva, and R. Krishnamoorti, "Linear Viscoelasticity of Disordered Polystyrene-Polyisoprene Block Copolymer Based Layered-Silicate Nanocomposites," *Macromolecules*, vol. 33, no. 10, pp. 3739–3746, May 2000, <https://doi.org/10.1021/ma992091u>.
- [20] A. Leygue, C. Bailly, and R. Keunings, "A differential tube-based model for predicting the linear viscoelastic moduli of polydisperse entangled linear polymers," *Journal of Non-Newtonian Fluid Mechanics*, vol. 133, no. 1, pp. 28–34, Jan. 2006, <https://doi.org/10.1016/j.jnnfm.2005.10.003>.
- [21] D. K. Bangwar, M. A. Soomro, N. A. Laghari, M. A. Soomro, and A. A. Buriro, "Improving the Bond Strength of Rice Husk Ash Concrete by Incorporating Polymer: A New Approach," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2595–2597, Feb. 2018, <https://doi.org/10.48084/etasr.1791>.
- [22] Y. H. Lin, "Number of entanglement strands per cubed tube diameter, a fundamental aspect of topological universality in polymer viscoelasticity," *Macromolecules*, vol. 20, no. 12, pp. 3080–3083, May 1987, <https://doi.org/10.1021/ma00178a024>.

- [23] S. Ng and H. Justnes, "Influence of dispersing agents on the rheology and early heat of hydration of blended cements with high loading of calcined marl," *Cement and Concrete Composites*, vol. 60, pp. 123–134, Jul. 2015, <https://doi.org/10.1016/j.cemconcomp.2015.04.007>.
- [24] W. Zheng, X. Wu, and Y. Huang, "Impact of polymer addition, electrolyte, clay and antioxidant on rheological properties of polymer fluid at high temperature and high pressure," *Journal of Petroleum Exploration and Production Technology*, vol. 10, no. 2, pp. 663–671, Feb. 2020, <https://doi.org/10.1007/s13202-019-0732-8>.

Resource Assessment of Tidal Stream Power in Pakiputan Strait, Davao Gulf, Philippines

Eleonor V. Palconit
School of Engineering
Ateneo de Davao University and
Mindanao Renewable Energy R&D Center
Davao City, Philippines
evpalconit@addu.edu.ph

Josef Rene Villanueva
School of Engineering
Ateneo de Davao University and
Mindanao Renewable Energy R&D Center
Davao City, Philippines
jrlvillanueva@addu.edu.ph

Nelson Enano Jr.
School of Engineering
Ateneo de Davao University and
Mindanao Renewable Energy R&D Center
Davao City, Philippines
nhenanojr@addu.edu.ph

Mario Buhali Jr.
Mindanao Renewable Energy R&D Center
Davao City, Philippines
mariobuhali jr@gmail.com

Arby Christy Mascariñas
Mindanao Renewable Energy R&D
Center
Davao City, Philippines
acpmascarinas@addu.edu.ph

Gabriel Earl Galope
Mindanao Renewable Energy R&D
Center
Davao City, Philippines
gedgalope@addu.edu.ph

Johna Marie Tabanguil
School of Engineering
Ateneo de Davao University and
Mindanao Renewable Energy R&D
Center
Davao City, Philippines
johmrtabanguil@addu.edu.ph

Abstract-During the last years, there are ongoing efforts on the development of tidal energy conversion systems in the Philippines. This study conducts tidal energy resource assessment in the Pakiputan Strait following a methodology outlined as stage 2a tidal resource assessment published by the European Marine Energy Centre (EMEC). The study assessed the preliminary results of the tidal velocities at Pakiputan Strait with a mean spring peak velocity (V_{msp}) of 1.7m/s at 3m from the water surface from 15 days of continuous data collection using a seabed-mounted acoustic Doppler current profiler. This corresponded to an estimated Annual Energy Production (AEP) of 1350kWh/y for 1m² of capture area of the generic device. Sensitivity analysis showed that the spatial distribution of hydrodynamic model results does not vary significantly with variations in certain input parameters. It further showed that a 10% decrease in the nominal value of V_{msp} on-site led to a 15% decrease in the nominal value of AEP, while a 10% increase in the nominal value of V_{msp} led to a 30% increase in the nominal value of AEP, assuming that the considered V_{msp} still corresponded to the velocity distribution from observations. A static survey and the use of computational fluid dynamics modeling are recommended to further enhance the analysis of the study.

Keywords-tidal energy; ocean renewable energy; resource assessment; Davao Gulf

I. INTRODUCTION

The global interest in renewable energy has increased over the past decade, mainly due to the global mandate to mitigate anthropogenic climate change. One of the renewable energy technologies that are exploited nowadays is ocean energy. The ocean provides a vast source of potential energy that can be captured through tidal, wave, salinity gradient, and ocean thermal energy conversion systems [1]. Energy from tidal currents is one of the most promising renewable energies due to the several advantages compared to other renewable energy sources [2]. Among its advantages are the high predictability of tidal currents [3], the higher density of seawater (over 800 times that of air) that assures larger potential energy as compared to wind [2], lesser environmental impact compared to tidal barrage facilities [4], and d) the absence of extreme flow velocities that might damage the equipment or complicate its maintenance [4]. The last point mentioned depends on the strong flow profile brought by the local bathymetry near the TEC deployment site [5]. With highly predictable tidal current benefits in tidal stream energy conversion, it has a high potential to contribute to the decarbonization of the energy supply and, at the same time, provide energy security [1].

In exploring a potential site for ocean energy extraction, the first consideration is the energy resource assessment [6]. One

method of identifying potential sites is by analyzing tidal range distributions and resources. This method was particularly followed by the authors in [7] in order to select tidal range sites in East Malaysia suitable for power generation. Another method is by producing a hydrodynamic model and analyzing its outputs. In the northern part of the Gulf of California in Mexico, a detailed qualification of the theoretical tidal range energy, performed using level from a depth-averaged barotropic hydrodynamic model and 0-D operation modeling approach was applied to determine the power that can be extracted at four key sites [3]. In [1], Delft3D modeling suite was used in assessing the Strait of Larantuka, Indonesia, wherein the flow conditions of the tidal stream were considered for the energy conversion. The bathymetric data have been obtained with the use of a Single-beam Echo Sounder (SBES) and SRTM15 Plus to have a more accurate determination of the potential energy to be harnessed, and the resource assessment showed that the Strait of Larantuka in Indonesia has a potential for the production of ocean renewable energy with current velocities up to 3–4m/s and average kinetic power densities around 10kW/m² [1]. The same modeling suite was also used in [8] along with MIKE system, to observe the morphological behavior of the Nhat Le estuary which is located in the north-central coast of Vietnam.

Numerous numerical modeling methods have been used in many researches in order to understand the hydrodynamic characteristic of the tidal stream energy sites and the current turbine technology. An example of this is the investigation of the tidal induced Lagrangian circulation in the north-western coastal waters of Brittany, France, including islands, shoals, and straits with strong potential for tidal turbine farm implementation [9]. With the increase of interest in different renewable energy sources, the Philippines is now exploring ocean renewable technologies as alternative energy sources that can be implemented in potential areas.

The current study aims to assess the tidal stream power in Pakiputan Strait, Davao Gulf, Philippines and to locate a potential site for the implementation of the tidal energy converter system. The structure of the assessment follows the steps described by the EMEC, which start from the regional assessment and site selection through the estimation of current speed and spatial distribution using hydrodynamic modeling. The preliminary results of this study show that tidal current velocities above 0.5m/s exist in the identified area. These currents could drive a generic energy harvesting device and occur frequently enough to ensure significant availability of power.

II. BACKGROUND INFORMATION OF THE STUDY SITE

A. The Philippine Oceans and Coasts

The Philippines, is an archipelago with 7,641 islands, with a total coastline stretching at 36,289km [10]. It is bounded by the Philippines Sea to the east, the Celebes or Sulawesi Sea to the south, the Sulu Sea to the southwest, the West Philippine Sea to the west, and the South China Sea to the north [11]. It has three major island groups, namely Luzon, Visayas, and Mindanao. The Sibuyan Sea separates Luzon and the Visayas. The islands in the Visayas are surrounded by the Visayan and

Camotes Seas. The Bohol Sea extends down to the north of Mindanao, with several straits connecting the internal and marginal seas to the ocean [10].

The total land area amounts to 298,170km², with a sea area of 1,830km², 6% of which are considered marine protected areas [12]. The population has reached 109,763,425 in 2020, with around 62% living in the coastal areas [12]. It was ranked 83rd among 221 countries with exclusive economic zones (EEZ) with an overall Ocean Health Index score of 71 [13]. The national ocean economy contributed US\$ 11.9 billion in 2015, amounting to 7% of the Gross Domestic Product (GDP) [12]. Some of the key contributors to this belong to coastal and marine tourism (25%), fisheries and aquaculture (20%), manufacturing (19%), and port and shipping (12%) sectors [12]. The country's coastal and marine habitats include coral reefs (25,060km²), natural mangroves (2,472.68km²), mudflats, and tidal swamps (2,000km²), seagrass (978km²), and planted mangroves (765.14km²) [12]. The blue economy initiatives that the country focuses on are sustainable fisheries and aquaculture, sustainable tourism, sustainable ports, marine renewable energy, marine biotechnology, pollution reduction, and habitat restoration and management [12].

The Davao Gulf is a 308,000-hectare body of water found in the southeastern part of Mindanao [14]. It is one of the richest zones for fish production in the country and is considered as one of the priority conservation areas of the Sulu-Sulawesi Marine Ecoregion. There are currently 19 marine protected areas in the 5 cities and 5 municipalities surrounding the gulf [15]. It is home to a diverse species of cetaceans with at least 15 identified species identified in 2018 [16]. According to the Permanent Service for Mean Sea Level (PSMSL), there is a tide gauge station located at 7.08333°N, 125.63333°E, in the Davao Gulf. It was considered a Revised Local Reference (RLR) station in 2010. There are two neighboring RLRs in the Davao Gulf, one in Mati, Davao Oriental and the other in Makar, General Santos City [17]. Aside from being home to diverse marine life, marine traffic is a commonality in the Davao Gulf. There were usually around 5 to 11 arrivals per day and about the same number of departures from July 22, 2020, to August 19. The majority of the vessels arriving in the port are cargo vessels with several tankers. Other ports in the Davao Gulf include the Davao Therma South port (6.9621405°N, 125.48615°E), and the Davao Anchorage port (7.0515755°N, 125.65395°E) [18].

B. Oceanographic Conditions

In the Philippine Straits Dynamics Experiment (PhilEx), global ocean models were used to investigate the internal circulation in the Philippine archipelago, both spatially and temporally. However, realistic modeling was difficult for a global model due to the numerous straits and small interior seas [19]. In general, the bathymetry of the Philippine waters between islands is highly complicated, consisting of narrow shelves, steep slopes, and deep basins connected by shallow sills. Ocean circulation is highly influenced by the bathymetry, the seasonally reversing northeast monsoon, locally known as Amihan, and southwest monsoon, locally known as Habagat, and the tidal and nontidal circulation between the South China Sea and the Western Pacific [20]. In addition to the reversing

winds, the monsoon brings a change in the rainfall pattern, which affects the surface layer thermohaline seasonal stratification [21]. The Amihan brings less rainfall while the Habagat more. Comparing actual rainfall for a specific year to the average annual cycle, however, reveals precipitation anomalies, which may affect the surface layer thermohaline stratification [21]. This can be further complicated due to the interannual El Niño or La Niña phenomena [21]. Numerous factors make the analysis of the circulation within the Philippines challenging. However, with further studies, a better understanding of these dynamics can be achieved.

C. Ocean Energy Potential

The Philippines' Department of Energy (DOE) estimates the theoretical ocean energy potential of the country to about 170GW, which was based on an assessment done by the Mindanao State University – Iligan Institute of Technology (MSU-IIT) [22]. Since then, several studies have been conducted in line with ocean energy resource assessment in the country. One such study reviews the ocean energy status in Southeast Asia [23]. It focuses not only on the potential resource but also the enabling and hindering factors, particularly in government initiatives that affect its development. In the study, a map of the potential sites for tidal in-stream energy, wave energy, and Ocean Thermal Energy Conversion (OTEC) in the Philippines was presented. For tidal-instream energy, four sites were identified, namely the Luzon Strait, San Bernardino Strait, Surigao Strait, and Basilan Strait.

In its National Renewable Energy Program (NREP), which started implementing in 2011, the Philippines' DOE has targeted to increase the country's renewable energy capacity to an estimated 15,304MW by 2030. This included an ocean energy share of 70.5MW, with a target of achieving the country's first ocean energy facility, a 10MW pilot ocean thermal energy conversion plant in Cabangan, Zambales, in 2018 [24, 25]. However, in 2019, the DOE moved to update its target after falling short of its original goals. Instead, it has focused on increasing the share of renewables in the total generation, instead of capacity, with a targeted 20,000MW installed capacity by 2040 [25, 26]. There are 8 awarded ocean projects by the DOE as of June 30, 2020, as presented in Table I. All the projects, however, are still in the pre-development stage, with 3 projects that have to determine their target capacity. The total capacity of the projects amounts to 26.00MW [27]. It is worth noting that 7 of these awarded projects involve tidal in-stream energy technology. This is a testament to its commercial readiness and maturity in terms of research and development.

III. METHODOLOGY AND APPLICATION

A. Methodology

The process flow diagram of the methodology for the pre-feasibility assessment is shown in Figure 1. The inputs for the hydrodynamic model, such as bathymetry and grid resolution, are mainly based on the area of interest. A desktop was conducted to identify the potential sites in Davao Gulf. The area of interest was narrowed down into Pakiputan Strait with a two-day transect survey.

TABLE I. AWARDED OCEAN ENERGY PROJECTS BY THE DOE AS OF JUNE 30, 2020

Region	Location	Technology	Developer	Type of contract	Total capacity (MW)
III	Cabangan, Zambales	OTEC	Bell Pirie Power Corp.	Pre-development	5.00
V	San Bernardino Strait, Matnog, Sorsogon	TISE	H&WB Corp.	Pre-development	5.00
VIII	Northwest Capul, Northern Samar	TISE	Oceantera Energy Corp.	Pre-development	TBD
	San Bernardino Strait, between Bicol Peninsula and Samar Leyte Corridor, Area 2OP	TISE	H&WB Corp.	Pre-development	5.00
	Capul Pass, Dalupiri Island, San Antonio, Northern Samar, TISEC-Project Site (Area 6)	TISE	Poseidon Renewable Energy Corp.	Pre-development	TBD
	San Bernardino Strait and Northern Samar, TISEC-Project Site (Area 4&5)	TISE	Poseidon Renewable Energy Corp.	Pre-development	TBD
	San Bernardino Strait, between Bicol Peninsula and Samar Leyte Corridor, Area 3OP	TISE	H&WB Corp.	Pre-development	5.00
XIII	Gaboc Channel, Surigao del Norte	TISE	Adnama Power Resources Corp.	Pre-development	6.00

The surface and depth-averaged velocities, along with their GPS coordinates, obtained from the transect survey were post-processed to identify potential locations for a tidal turbine. Bathymetry data were collected. Continuous on-site data collection of tidal currents was also carried out to collect empirical data as inputs for a harmonic analysis. The bathymetry data and the results of the harmonic analysis were used as inputs to the hydrodynamic modeling. Finally, a forecast estimate of the mean electrical power and annual energy production was determined from the collected tidal current data.



Fig. 1. Methodology flow diagram.

B. Site Selection

This research closely follows the methodology of site definition and selection described in Assessment of Tidal Energy Resource, a document prepared in consultation with The European Marine Energy Centre Ltd (EMEC) and with other interested parties in the UK marine energy community [6]. The site selection methodology used in the study was the Stage 2a, which is the first stage of the pre-feasibility of the tidal resource assessment [6]. The determined region was Davao Gulf. The channel considered was Pakiputan Strait in the northern part of the gulf, with the aim of selecting a localized area for the tidal energy converter implementation.

C. Estimation of Current Speeds

The modeling suite used in the hydrodynamic modeling was Delft3D [28]. Other inputs, such as the grid, land boundary file, and enclosure, were also generated from the modules included in Delft3D. Rectangular structured grids were used to coincide with the given grid resolution for each stage of the resource assessment. For stages 1 and 2a, the grid resolutions should be less than 5km and 500m, respectively [6]. As shown in Figures 2, 3, the model has a grid resolution of 1km for the regional assessment (Stage 1) and a grid resolution of 100m for the pre-feasibility assessment (Stage 2a). In Stage 2a, the area of interest was narrowed down into a channel. The bathymetry data were acquired from the French Research Institute for Exploitation of the Sea (IFREMER) [29-31] and the General Bathymetric Chart of the Oceans (GEBCO) [32]. The raw bathymetric data were processed with DELFT 3D as shown in Figures 4-5. The bathymetric data from the 2020 gridded bathymetry of GEBCO have soundings with spacing of approximately 450m. However, the soundings for IFREMER bathymetric data were not specified in the obtained raw data. As for the harmonic analysis, 15 tidal constituents (O1, K1, P1, N2, NU2, M2, L2, S2, K2, MK3, M4, S4, M6, S6, and M8) were extracted from the empirical tidal velocity data that were continuously collected on site for 15 days. The amplitudes and phases of these tidal constituents were used in the tidal forcing needed by the model. Aside from the open boundary, physical parameters and flow conditions were also added. Parameters such as salinity, temperature, and other meteorological inputs were not yet considered in the simulation of the model.

Observation points were added in the model to check the modeled tidal currents in specific areas.

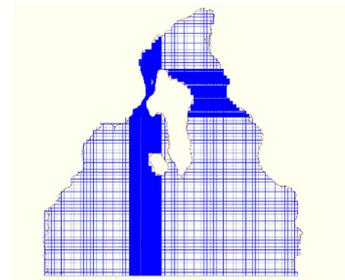


Fig. 2. Davao Gulf 1km grid resolution.

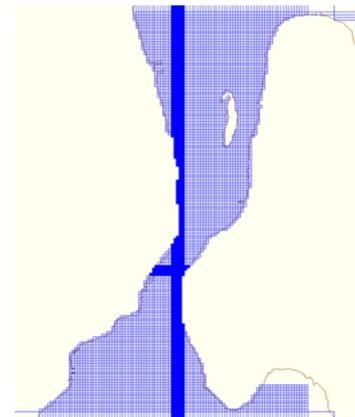


Fig. 3. Pakiputan Strait 100m grid resolution.

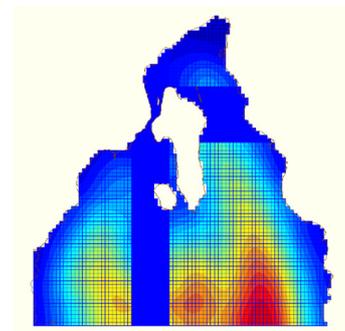


Fig. 4. Bathymetric data for Davao Gulf.

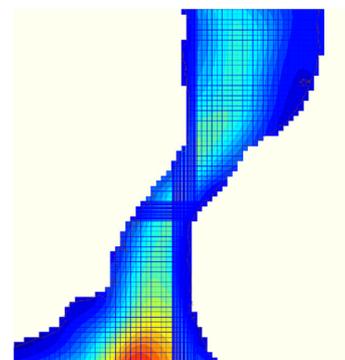


Fig. 5. Bathymetry for Pakiputan Strait.

Several simulations were made, as a part of the sensitivity analysis, with varied inputs. It was found out that non-astronomical attribute files have a significant influence on the output of the model. Three simulations showed more realistic results. From these three simulations, it was determined, using their respective East vs. North graphs (see Figure 6), that the current in Pakiputan Strait is reversing. This was also evident in the results of the harmonic analysis for the eccentricity of all the harmonic constituents that are almost equal to 1.

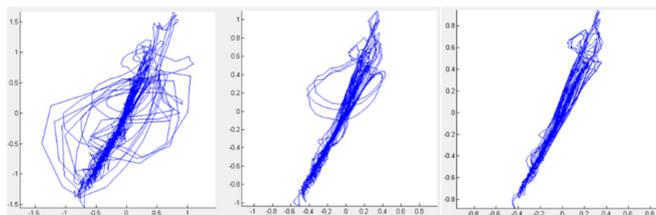


Fig. 6. East vs. North components (m/s).

The potential sites from the spatial variation of the depth-averaged velocities, as shown in Figure 7, were similar to those that can be identified from the desktop study. However, though the spatial variation gave similar sites to those from the desktop study, the current velocities from the model did not show a positive linear relationship with the actual current velocities obtained from the 15-day data collection. Hence, to improve the model performance and outputs, the grid, location of the open boundary, tidal forcing, and other parameters of the model must be reviewed. In the new model, the number of layers will be increased to generate predicted surface velocities and not just depth-averaged velocities. The new hydrodynamic model will then be validated using in-situ measurements.

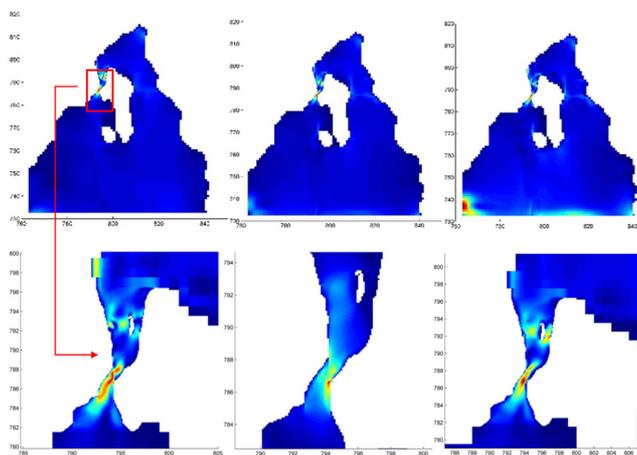


Fig. 7. Spatial variation for depth-averaged velocities.

D. Energy Yield

An estimate of the mean electrical power and annual energy production from tidal in-stream energy was obtained simulating a generic device, as described in [6]. A generic device was used for the study since no device had yet been tested. The available

power $P_{av(i)}$ for every vertical bin i across the depth of the water is given by:

$$P_{av(i)} = 0.5\rho AU_i^3 \quad (1)$$

where ρ is the density of seawater, assumed constant for simplicity at 1025kg/m^3 , A is the capture area of the device, assumed to be 1m^2 (bin size), and U_i^3 is the tidal stream velocity at the bin.

Combining this with the rotor efficiency η_R , the electrical power per bin $P(U_i)$ is given by:

$$P(U_i) = P_{av(i)} \cdot \eta_R \quad (2)$$

For a generic device, η_R was considered to rise from 38% at cut-in speed to reach 45% at the rated velocity. The rated velocity was taken as 71% of the mean spring peak velocity V_{msp} at the hub height. The V_{msp} is defined in [6] as the peak velocity that has been reached for 10min during the whole month from the results of a static field survey, model, or harmonic analysis. From the results of a static field survey conducted during February 2020 to collect empirical data for harmonic analysis, the V_{msp} was estimated to be at 1.7m/s at a hub height of 3m below the water surface (i.e. $V_{msp} = U_3$). The data set was comprised of velocity data collected at 1min intervals for each 1m vertical bin across the depth continuously over a 15-day period on the same site. This V_{msp} of 1.7m/s was significantly higher compared to that which had been determined from an earlier transect survey using a boat-mounted Acoustic Doppler Current Profiler (ADCP) on the same site conducted in August 2019, generating a value of 1.5m/s. Hence, the rated velocity for the generic device was determined as $0.71 \times 1.7 \approx 1.2\text{m/s}$.

The cut-in speed (the minimum velocity required for device operation) was assumed constant at 0.5m/s. This assumption greatly simplified the analysis without imposing significant limitations on the accuracy, since the available energy from tidal currents at speeds below 0.5m/s is usually less than 5% of the total available energy. A velocity distribution $f(U_i)$ is also defined to be derived from the tidal harmonic analysis [6]. The data set from the 15-day data collection was instead used to develop $f(U_i)$. As part of the preliminary results, this distribution was assumed to represent an annual distribution. The chart in Figure 8 displays the energy yield from a generic device, showing $P_{av(i)}$, $P(U_i)$, and $f(U_i)$.

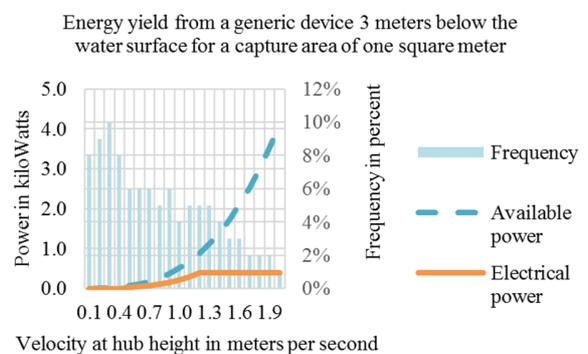


Fig. 8. Velocity distribution and estimated power from the site.

The mean annual electrical power P_{mean} was obtained by combining the velocity distribution $f(U_i)$ with the average absorbed power for each velocity bin $P(U_i)$ by using (3):

$$P_{mean} = \sum_{i=1}^{N_B} P(U_i) \cdot f(U_i) \quad (3)$$

The number of vertical bins N_B considered was only one (i.e. the 1m vertical bin at 3m below the water surface). After taking the summation, P_{mean} was determined to be at 0.154kW for 1m² of the capture area.

E. Sensitivity Analysis

Performing sensitivity analysis on the hydrodynamic model showed that varying the input parameters to the model resulted in a visually similar spatial distribution of the currents in Pakiputan Strait. Figure 9 illustrates several results from these variations. Certain inputs to the model, particularly the tidal constituents and the bathymetry data, were kept constant during the sensitivity analysis.

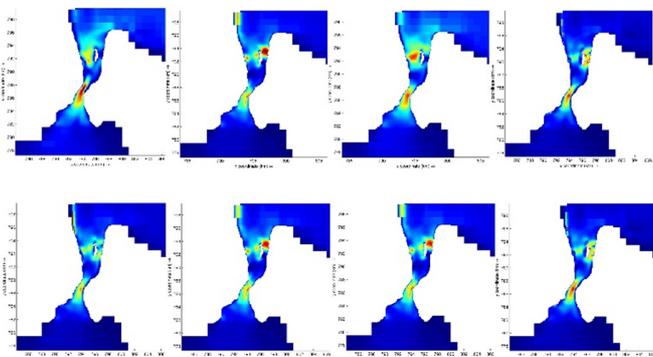


Fig. 9. The hydrodynamic model spatial results did not vary significantly with varying input parameters.

Due to the inherent dependence of the nominal value of energy and power to the cube of the nominal value of tidal velocities, as shown in (1), it was reasonable to expect significant variations of power corresponding to any changes in observed velocity values.

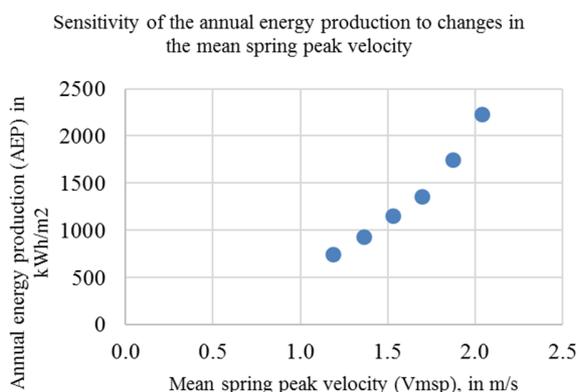


Fig. 10. AEP varied significantly with varying Vmsp.

Sensitivity analysis (Figure 10) showed that a 10% decrease in the nominal value of the Vmsp on site will lead to a 15%

decrease in the nominal value of the AEP. This observation was also true for every 10% decrease up to 30%. On the other hand, a 10% increase in the nominal value of the Vmsp on site led to a 30% increase in the nominal value of AEP. Checking for sensitivity beyond a 20% increase in the nominal value of the Vmsp was not realistic because the velocity distribution shows no instances of such velocities occurring. This analysis assumed that the considered Vmsp still corresponded to the velocity distribution derived from observations.

IV. CONCLUSION

Based on preliminary results, Pakiputan Strait is a candidate site for the deployment of a tidal energy conversion system. From the results of a static field survey conducted during February 2020 and the collected empirical data for harmonic analysis, the Vmsp was to be at 1.7m/s at a hub height of 3m below the water surface (i.e. $V_{msp} = U_3$). The data set was comprised of velocity data collected at 1 min intervals for each 1m vertical bin across the depth continuously over a 15-day period on the same site. The velocities from modeled and empirical data show that tidal power from a generic device will be potentially available for the 60% of the time. The results of the study are not yet conclusive and could be improved by the collection of more data and conducting a field survey as a validation step. Much motivation could be derived from the preliminary results, as it responds to technological questions by showing tidal velocities that can run a generic device. Similarly, it beckons the need to design devices suitable to harness the resource given in this site. Additional sections of the channel could be used as generation sites once the identification of constraints is completed at the end of the resource assessment and during its further stages.

ACKNOWLEDGMENT

The Ocean Renewable Energy (ORE) Project of MREC is made possible by the generous funding of the Philippine Department of Science and Technology Philippine Council for Industry, Energy, and Emerging Technology (DOST-PCIEERD) and also of the University Research Council of the Ateneo de Davao University. Much of this research has been the output of the first year of the project. The authors would like to thank Dr. Michael Abundo and Weslie Capute of the Ocean Pixel Ltd for the technical assistance during the conduct of the tidal current survey in the Pakiputan Strait, Davao Gulf, and its project staff, namely Janice Ivy Jumao, Vince Peter Cortez, and Hazel Dioso.

REFERENCES

- [1] K. Orhan and R. Mayerle, "Assessment of the tidal stream power potential and impacts of tidal current turbines in the Strait of Larantuka, Indonesia," *Energy Procedia*, vol. 125, pp. 230–239, Sep. 2017, <https://doi.org/10.1016/j.egypro.2017.08.199>.
- [2] M. Mestres, P. Cerralbo, M. Grifoll, J. P. Sierra, and M. Espino, "Modelling assessment of the tidal stream resource in the Ria of Ferrol (NW Spain) using a year-long simulation," *Renewable Energy*, vol. 131, pp. 811–817, Feb. 2019, <https://doi.org/10.1016/j.renene.2018.07.105>.
- [3] C. J. Mejia-Olivares, I. D. Haigh, A. Angeloudis, M. J. Lewis, and S. P. Neill, "Tidal range energy resource assessment of the Gulf of California, Mexico," *Renewable Energy*, vol. 155, pp. 469–483, Aug. 2020, <https://doi.org/10.1016/j.renene.2020.03.086>.

- [4] G. Iglesias, M. Sanchez, R. Carballo, and H. Fernandez, "The TSE index – A new tool for selecting tidal stream sites in depth-limited regions," *Renewable Energy*, vol. 48, pp. 350–357, Dec. 2012, <https://doi.org/10.1016/j.renene.2012.05.012>.
- [5] P. Evans, A. Mason-Jones, C. Wilson, C. Wooldridge, T. O'Doherty, and D. O'Doherty, "Constraints on extractable power from energetic tidal straits," *Renewable Energy*, vol. 81, pp. 707–722, Sep. 2015, <https://doi.org/10.1016/j.renene.2015.03.085>.
- [6] C. Legrand, *Assessment of Tidal Energy Resource*. London, UK: BSI, 2009.
- [7] 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>.
- [8] D. D. Cham, N. T. Son, N. Q. Minh, N. T. Hung, and N. T. Thanh, "Hydrodynamic Condition Modeling along the North-Central Coast of Vietnam," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5648–5654, Jun. 2020, <https://doi.org/10.48084/etasr.3506>.
- [9] N. Guillou and G. Chapalain, "Assessing the impact of tidal stream energy extraction on the Lagrangian circulation," *Applied Energy*, vol. 203, pp. 321–332, Oct. 2017, <https://doi.org/10.1016/j.apenergy.2017.06.022>.
- [10] W. Y. Licuanan, R. W. Cabreira, and P. M. Aliño, "Chapter 23 - The Philippines," in *World Seas: an Environmental Evaluation*, 2nd ed., C. Sheppard, Ed. Cambridge, MA, USA: Academic Press, 2019, pp. 515–537.
- [11] G. C. Borlaza, M. Cullinane, and C. G. Hernandez, "Philippines | History, Map, Flag, Population, Capital, & Facts," *Encyclopedia Britannica*. <https://www.britannica.com/place/Philippines>
- [12] "State of Oceans and Coasts of the Philippines," *PEMSEA*. <http://pemsea.org/publications/reports/state-oceans-and-coasts-philippines>
- [13] "Philippines," *Ocean Health Index*. <http://www.oceanhealthindex.org/region-scores/scores/philippines>
- [14] "Davao Gulf," *Wikipedia*. Feb. 02, 2021, Accessed: May 13, 2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Davao_Gulf&oldid=1004366756.
- [15] World Wildlife Foundation, "Protecting the Davao Gulf." <https://wwf.org.ph/what-we-do/food/davaogulf/>
- [16] World Wildlife Foundation, "17 Sightings Recorded in Recent SMARTSeas PH Survey in Davao Gulf." <https://wwf.org.ph/what-we-do/food/davaogulf/17-sightings-recorded-in-smartseas-ph-survey-in-davao-gulf/>
- [17] Permanent Service for Mean Sea Level, "Data and Station Information for DAVAO, DAVAO GULF," 2019. <https://www.psmsl.org/data/obtaining/stations/537.php>
- [18] AIS Marine Traffic, "Port of DAVAO (PH DVO) details - Departures, Expected Arrivals and Port Calls," *MarineTraffic.com*, 2020. http://www.marinetraffic.com/en/ais/details/ports/1657/Philippines_port:DAVAO
- [19] H. E. Hurlburt *et al.*, "Circulation in the Philippine Archipelago Simulated by 1/12° and 1/25° Global HYCOM and EAS NCOM," *Oceanography*, vol. 24, no. 1, pp. 28–47, 2011, <https://doi.org/10.5670/oceanog.2011.02>.
- [20] C. L. Villanoy *et al.*, "Monsoon-Driven Coastal Upwelling Off Zamboanga Peninsula, Philippines," *Oceanography*, vol. 24, no. 1, pp. 156–165, 2011, <https://doi.org/10.5670/oceanog.2011.12>.
- [21] A. L. Gordon, J. Sprintall, and A. Field, "Regional Oceanography of the Philippine Archipelago," *Oceanography*, vol. 24, no. 1, pp. 14–27, 2011, <https://doi.org/doi:10.5670/oceanog.2011.01>.
- [22] Department of Energy Philippines, "Biomass, Solar, Wind And Ocean." <https://www.doe.gov.ph/biomass-solar-wind-and-ocean?ckattempt=1>
- [23] M. A. J. R. Quirapas, H. Lin, M. L. S. Abundo, S. Brahim, and D. Santos, "Ocean renewable energy in Southeast Asia: A review," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 799–817, Jan. 2015, <https://doi.org/10.1016/j.rser.2014.08.016>.
- [24] Department of Energy Philippines, "National Renewable Energy Program," 2011. <https://www.doe.gov.ph/national-renewable-energy-program>
- [25] "Philippines plans 20GW of renewable energy capacity by 2040," *Offshore Energy*, Apr. 28, 2017. <https://www.offshore-energy.biz/philippines-plans-20gw-of-renewable-energy-capacity-by-2040/> (accessed May 13, 2021).
- [26] Department of Energy Philippines, "Renewable Energy Roadmap 2017-2040," 2017. <https://www.doe.gov.ph/pep/renewable-energy-roadmap-2017-2040>
- [27] "Renewable Energy-Awarded Ocean Projects," *Department of Energy Philippines*. <https://www.doe.gov.ph/renewable-energy?q=renewable-energy/awardedocean&ckattempt=1> (accessed May 13, 2021).
- [28] "Home - Delft3D - oss.deltares.nl." <https://oss.deltares.nl/web/delft3d/home> (accessed May 13, 2021).
- [29] C. Rangin, "MODEC 1 cruise,L'Atalante R/V," 1994, <https://doi.org/10.17600/94010020>.
- [30] L. Labeyrie, "ESTASE II cruise,Jean Charcot R/V," 1984, <https://doi.org/10.17600/84004711>.
- [31] C. Rangin, "DAVAPUS cruise,L'Atalante R/V," 1994, <https://doi.org/10.17600/94010030>.
- [32] "GEBCO_2020 Grid," *GEBCO*. https://www.gebco.net/data_and_products/gridded_bathymetry_data/GEBCO_2020.html (accessed May 13, 2021).

Facile Coating of HAP on Ti6Al4V for Osseointegration

Shafaq Asrar

Department of Metallurgical Engineering
NED University of Engineering and Technology
Karachi, Pakistan
shafaqfara@neduet.edu.pk

Muhammad Tufail

NED University of Engineering and Technology
Karachi, Pakistan
mtufail@neduet.edu.pk

Ali Dad Chandio

Department of Metallurgical Engineering
NED University of Engineering and Technology
Karachi, Pakistan
alidad@neduet.edu.pk

Abstract-Ti6Al4V alloy is a material with great strength, low-slung modulus, inferior density, and a virtuous blend of mechanical and exceptional corrosion resistance. However, it does not offer good osseointegration and bone development properties. Conversely, hydroxyapatite (HAP) is highly bioactive in nature to bind with the nearby bone tissues when implanted in the host body. In this work, we have extracted HAP from bovine bones by using the thermal decomposition method. This was followed by its deposition onto the Ti6Al4V alloy using the Electrophoretic Deposition (EPD) technique. TiO₂ is used as a bond coat layer to increase the adhesion between HAP and Ti6Al4V alloy substrates. The coated samples after sintering exhibited excellent adhesion. This was followed by characterization using Scanning Electron Microscopy (SEM) and Fourier Transformed Infrared Spectroscopy (FTIR). FTIR and SEM confirm the formation of HAP and its presence after the immersion in SBF. Vicker hardness tester confirms the increase in hardness value of coated samples up to 35%. Potentiostat tests were conducted to compare the corrosion rate of both samples. In addition, the particle sizes were also identified by a laser particle analyzer, whereas X-Ray Diffraction (XRD) technique was also used to determine the crystalline phases of alloy and HAP.

Keywords-corrosion; electrophoretic deposition; hydroxyapatite; simulated body fluid; Ti6Al4V alloy

I. INTRODUCTION

Materials having biotic nature or host tissue compatibility can be implanted into living organisms to augment or replace impaired parts. Generally, metals, ceramics, polymers, and composites are materials widely used in biomedical applications. Amongst them, titanium-based alloys are some of the most widely used as implant materials [1]. Titanium and its alloys bear enhanced properties like excellent strength [2], optimum elastic modulus [3], low density [4], blend of other mechanical properties [5], and excellent corrosion resistance [6-12], needed for various applications including orthopedic

[13], dental [14], and surgical implants [15], artificial joints [15], etc. It should be noted that the elastic modulus of titanium-based alloys is considerably close to bones' [16, 17] which makes it ideal for long term applications [18]. However, titanium based alloys do not possess good osseointegration [19] thereby requiring additional surface treatment. This means that there is a poor bond between titanium and bones causing implant loosening which is highly undesirable. Therefore, surface modification plays a vital role in optimum osseointegration [15, 20]. There are several techniques used in this regard such as: sand blasting [21, 22], etching [23], electrochemical treatment [24], and thermal spray coatings [25]. Amongst them, bio-ceramic coatings using, the hydroxyapatite (HAP - Ca₁₀(PO₄)₆(OH)₂) [26-28] are very promising for the modification of implant surfaces since they create strong bonding with bones [29-31]. One of the great advantages of the HAP is its great lifespan [32, 33]. The calcium phosphate (CaP) ratio of HAP is 1.67 which is highly stable at a normal temperature and its pH ranges from 4 to 12. However, the properties and resultant applications of HAP depend on morphology, size, chemical composition and crystallinity [31]. In addition, it provides speedy and durable fixation to the host bones and possesses osseointegrative properties [34], protecting the metal surfaces from the environmental effects and thereby reducing the discharge of metallic ions from the implant surface to the host body. The main advantage of CaP is that it is already present in the bones and teeth of the vertebrates [35]. Moreover, there are several methods to produce HAP such as: dry methods, wet methods, microwave (MW)-assisted methods, ball-milling or ultrasound, etc. [36]. Additionally, numerous coating strategies are available to coat HAP on metallic alloys, for instance: plasma spraying [37, 38], sol gel [39, 40], Electrophoretic Deposition (EPD) [11, 41-43], etc. Amongst them the most economically viable technique is EPD [44], offering a controlled coating composition with the process being highly pure [45], fast [41]

that could be used to coat complex shaped substrates [45]. In this study, HAP was produced from bovine bones since they are an abundant and economical source. This was followed by the deposition of HAP on chemically treated Ti6Al4V alloy.

II. EXPERIMENTAL PART

A. Materials

The materials utilized in this research are shown in Table I. All of them were of analytical grade and were used in as received condition.

TABLE I. MATERIALS/CHEMICALS USED

S.No.	Chemicals / materials	Purity (%)	Supplier
1	Ti-6Al-4V	99.0	Baoji North Hongsheng Industry & Trade Co., Ltd.
2	Sodium chloride	99.0	Sigma Aldrich
3	Sodium bicarbonate	99.0	Sigma Aldrich
4	Potassium chloride	99.0	Sigma Aldrich
5	Di sodium phosphate tri hydrate	99.0	Sigma Aldrich
6	Magnesium chloride hexa hydrate	99.0	Sigma Aldrich
7	Hydrochloric acid	99.0	Sigma Aldrich
8	Calcium chloride dehydrate	99.0	Sigma Aldrich
9	Di sodium sulphate	99.0	Sigma Aldrich
10	Cyano hydride tri methanol	99.0	Sigma Aldrich

B. Methods

1) Preparation of Hydroxyapatite

HAP was synthesized from bovine bones, purchased from the local market, by the thermal decomposition process. This was followed by their boiling in deionized water for about 3 hours to remove the unwanted fats. Additionally, second boiling was carried out for the deproteinization of bones. Thereafter, the bones were immersed in acetone for two hours in an ultrasonic bath for further cleaning. Subsequently, the bones were dried and cut into smaller sizes by mortar and pestle. The resultant powder was placed into a box furnace and heated at 1100°C at a heating rate of 5°C/min for 3 hours for the preparation of HAP. Later on, XRD and laser particle analysis were conducted to confirm the peaks and particle sizes of as-synthesized HAP powders respectively. The process details are shown in Figure 1.

2) Activation of Substrates

Titanium alloy Ti6Al4V was used as substrate in plate form, having a thickness of 2mm. The microstructure of the as-received Ti alloy was revealed by following the steps of grinding and polishing followed by etching for 20-25s in the Kroll's reagent (96ml distilled water + 6ml HNO₃ + 2ml HF). A wet chemical etching process was carried out to enhance the surface available for coating at room temperature. Several chemicals were used to immerse the electrode into the solution for the desired time and then it was ready to be used as a substrate for coating. Next, we took an equal quantity of 5ml each for hydrochloric acid, nitric acid, hydrogen peroxide, sulfuric acid and introduced them into 30ml of distilled water

followed by the immersion of substrates into the prepared solution for 3 hours. Then, titanium substrates were placed in 5 Molar NaOH solutions for 72 hours at room temperature. In the next step, the substrates were annealed in the air using a box furnace at 600°C followed by holding for 1 hour. This process was used to create the small nano-sized pits on the surface of the substrates used for the coating purpose. The samples were then considered ready for the EPD process.

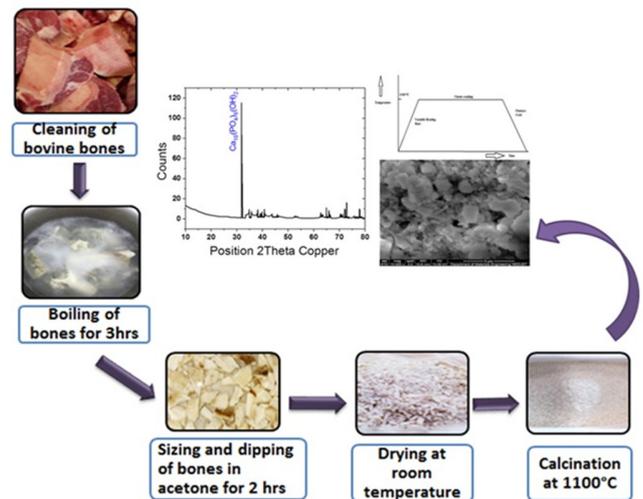


Fig. 1. Steps of producing HAP from bovine bones, calcination cycle, and XRD pattern of successfully produced HAP.

3) EPD of HAP

For the enhancement of adhesion properties of hydroxyapatite, Titania (TiO₂) was used as an intermediate layer between HAP and Titanium alloy substrate. The process flow chart is shown in Figure 2. In this process, an electrolytic solution containing 0.5g TiO₂ and 0.5g HAP was dissolved in 100ml ethanol solution followed by sonication for 15 minutes and was ultrasonically shook for 30 minutes. It was then left for settling for 20 minutes and was stirred again for 25 minutes at 40°C. The EPD process was done at 20V for 5 minutes while keeping the distance between electrodes at 2cm. The resultant coating was sintered in the tube furnace at 800°C. Finally, the samples were ready for characterization to further evaluate the properties of Ti-6Al-4V alloy coated with hydroxyapatite.

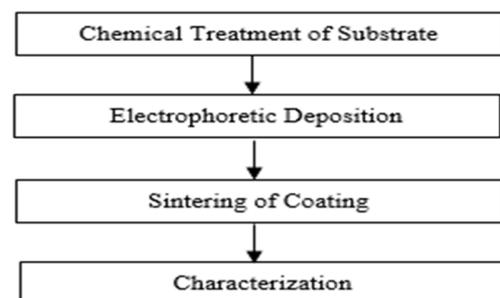


Fig. 2. Process flowchart of coating HAP on Ti-6Al-4V substrate.

III. RESULTS AND DISCUSSION

A. Particle Size Analysis

We used the BT-9300H laser particle analyzer for the determination of particle size of HAP powder. The results are shown in Figure 3. The average obtained particle size was 25.73 microns with respect to cumulative percentage.

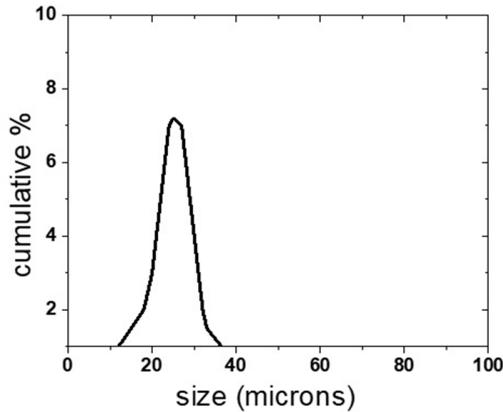


Fig. 3. Average particle size of HAP powder synthesized by bovine bones.

B. XRD Analysis

The X-pert Pro XRD DY3313 XRD machine operating at 40kV and 30mA using CuK α radiation was used to analyze the phase of the HAP. The diffraction pattern was recorded over 2 θ at the scanning rate of 0.1 and was performed over the angular range from 10° to 79°. Figure 4 reveals the XRD spectra of bovine bones calcinated at 1100°C which is in good agreement with the standard HAP pattern [46, 47]. The calcinated sample peak at 31.8 confirms the formation of hydroxyapatite powder. Moreover, the XRD spectra of the chemically treated Ti alloy surface show the presence of TiO $_2$ and coated alloy confirming the presence of the HAP layer on the chemically treated Ti6Al4V.

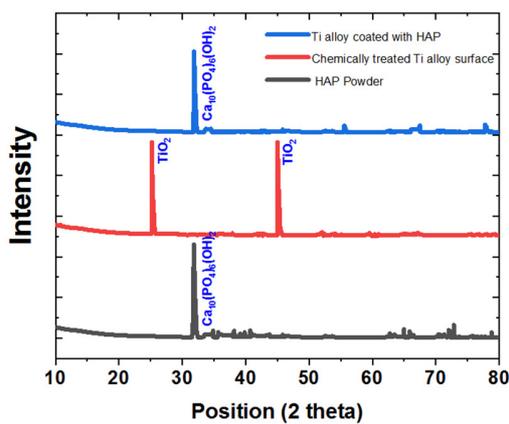


Fig. 4. XRD pattern of synthesised HAP, chemically treated surface and Titanium alloy coated with HAP powder after sintering.

C. Chemical Composition

The chemical composition of titanium alloy is found by using an X-ray fluorescence spectroscope (INNOV-X

SYSTEMS) which is comparable to ASTM F136 standard [48]. The results are presented in Table II.

TABLE II. CHEMICAL COMPOSITION USING XRF ANALYSIS

Alloy	Titanium	Vanadium	Aluminum
ASTMF136	Balance	3.5-4.5	5.5-6.75
Ti6Al4V	Balance	4.17	6.12

D. Microstructures

The polarized light microscope OLYMPUS GX51 was used to reveal the microstructure of the as-received samples. It was found that the titanium alloy consists of two different phases: equiaxed Alpha and transformed Beta phase [49] which are shown in Figure 5(a). Figure 5(b) shows the SEM image of the HAP powder. The powder comprises of agglomerated fine particles while the shape of the particles is angular and non-spherical [50]. Figure 5(c) shows Ti6Al4V alloy coated with HAP after sintering which revealed that there was an enhanced linkage and the interconnection of HAP powders that exist on the coated surface morphology [51]. Figure 5(d) shows the HAP coated on Ti6Al4V after the immersion in SBF [52].

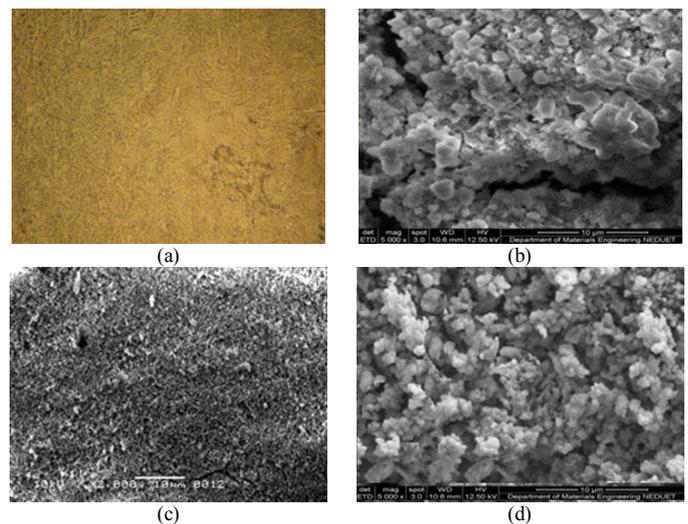


Fig. 5. (a) Optical structure of as received sample. (b) SEM image of the produced HAP powder. (c) Ti6Al4V coated with HAP and sintered at 800°C. (d) After immersion in SBF solution.

E. Hardness

Micro Vickers hardness tester (DIN EN 6507) was used to examine the hardness of bare and coated Ti6Al4V alloy by applying a load of 500g with a dwell time of 10s. The results are shown in Table III. A significant increase in hardness value of 468HV was obtained for a coated sample as compared to the pristine sample with a hardness value of 340HV. The Ti-based alloy produced a high hardness value which is a major requirement in being compatible with the body environment.

TABLE III. HARDNESS RESULTS

Alloy	Hardness (HV)
Ti6Al4V (before coating)	340
Ti6Al4V (after coating)	468

F. Corrosion Behavior

1) Immersion Test of Coated Samples in SBF

The coated samples of the HAP were immersed in Simulated Body Fluid (SBF) solution which was prepared according the procedure defined in [40]. The chemical composition of the SBF solution is shown in Table IV. The characterization was based upon the pH scale. For comparative analysis, the pH of the solution was measured before immersion and was found to be 7.4. The sample was then immersed in the solution enclosed in a glass beaker at 37°C for 7 days. Agitations and vibrations were provided to the beaker to create an artificial environment of fluid movement around the coating. After a passage of the prescribed time, the sample was removed from the solution. The solution pH was measured after the immersion and was found to be the same as before, i.e. 7.4. Therefore it was confirmed that no exchange of ions took place between the sample and SBF.

TABLE IV. CHEMICAL COMPOSITION OF THE SBF SOLUTION

S.No.	Chemicals	Amount
1	Sodium chloride	6.559g
2	Sodium bicarbonate	2.26g
3	Potassium chloride	0.3773g
4	Di potassium phosphate tri hydrate	0.1496g
5	Magnesium chloride hexa hydrate	0.3411g
6	Hydrochloric acid	10ml
7	Calcium chloride dehydrate	0.3635g
8	Di sodium sulphate	0.0731g
9	Cyano hydride tri methanol	6.0662g

2) FTIR

FTIR (Perkin Elmer spectrum one system) was used to identify the functional groups of HAP coated samples in the area of 400–4000 cm^{-1} . The FTIR spectra before and after the immersion in SBF solution is shown in Figures 6 and 7. The PO_4^{3-} group displays peaks at 560 and 600 cm^{-1} and at 1000–1100 cm^{-1} . The peak at about 2600–3600 cm^{-1} links to the hydrated OH^- ion of HA [53]. The peak from 2000 to 2200 cm^{-1} shows the stretching of P-O-H [54]. The peak at 1490 cm^{-1} matches CO_3^{2-} which specifies that HAP is a carbonated-apatite (HCA) [55]. The representative peaks of PO_4^{3-} appear at 1090, 1014, and 590 cm^{-1} . The peak at 1650 cm^{-1} shows the presence of H_2O [54].

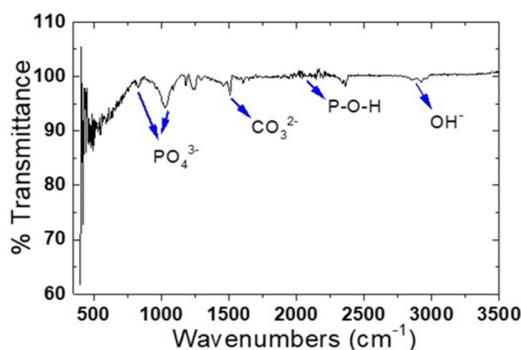


Fig. 6. FTIR spectra of Titanium alloy coated with HAP powder after the immersion in SBF solution.

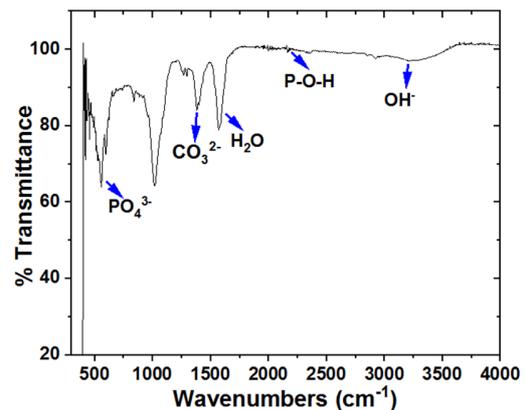


Fig. 7. FTIR spectra of the Titanium alloy coated with HAP powder before the immersion in SBF solution.

3) Potentiostat Test

Corrosion tests were conducted to find out the corrosion rate of the coated samples. The tests were conducted with a GAMRY potentiostat having a counter electrode (graphite) and a reference electrode (hydrogen) using Tafel. The corrosion rate was determined in an artificially created human body environment called SBF. The Tafel curve of the bare and coated sample is shown in Figure 8. The samples (i.e. the working electrode) were prepared according to the standard ASTM G108-94 analyzed at 37°C using voltage ranges between -0.3 and 0.3V. Tafel curves were plotted at a rate of 1mv/sec to find the current density and potential. Figure 8 also shows the potentiodynamic curve for bare Ti6Al4V alloy having a current potential of -240V and current density of $1.63 \times 10^{-4} \mu\text{A}/\text{cm}^2$. Hence, the corrosion current density (I_{corr}) in the passive zone shifted down from 252nA to 701nA and the corrosion potential (E_{corr}) from -0.24V to 0.25V. The corrosion rate for HAP coated substrate was decreased as $11.95 \times 10^{-3} \text{mpy}$ from $33.53 \times 10^{-3} \text{mpy}$ for the bare Ti6Al4V alloy. The HAP coating consequence on Ti6Al4V alloy infers that the coating shows minimum release of metallic ions in SBF solution [14].

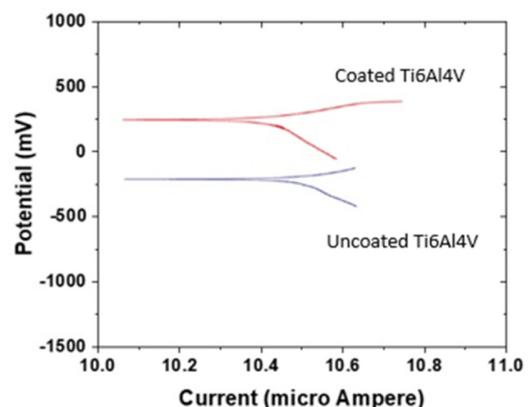


Fig. 8. Tafel curve of coated and uncoated Titanium alloy.

G. Adhesion

For evaluating the adhesion of the coated samples we used the scratch tester machine M-TGN80 having diamond indenter

landed on the coated substrate producing an indent on the substrate surface up to a maximum load of 60N as shown in Figure 9. The result shows an excellent strength for the coating [56]. The morphological measurements were taken in the stereo microscope under low magnifications (5X) which reveal the distances of the crack from various aspects.

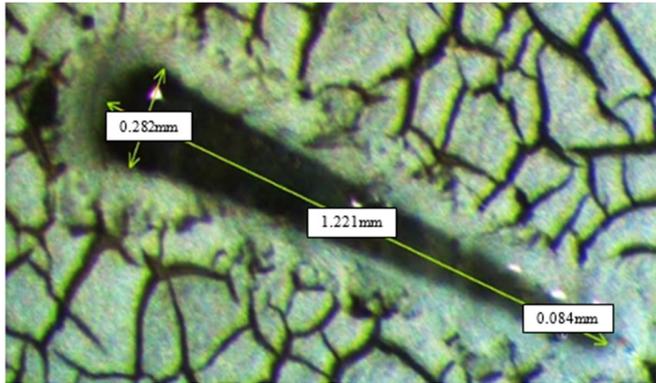


Fig. 9. Magnified scratch marked with the Adhesion tester.

IV. CONCLUSION

Hydroxyapatite is the furthestmost noticeable inorganic substance for biomedical applications. Research showed that natural HAP from bovine bones can be synthesized by using the thermal decomposition method. From this study, no significant difference was observed between bovine bone-derived HAP and naturally occurring HAP. HAP powder electrophoretically deposited on Ti-6Al-4V alloy was used as substrate. The characterization of the coated sample confirmed that the coating enhanced the osseointegration properties of the implant. Moreover, the hardness of the coated samples was observed to be increased up to 35% as compared to the pristine samples. The corrosion rate for coated samples was found to be decreasing from 33.53×10^{-3} mpy to 11.95×10^{-3} mpy. Therefore, the resultant structures show that our produced samples can be used for a wide variety of medical applications with minimal expenses.

ACKNOWLEDGMENT

The authors wish to express their gratitude to the Higher Education Commission of Pakistan for the funding under the NRP "National Research Program for Universities 2017-18" grand. We also thank the NED University of Engineering and Technology for providing us the platform to carry out this work.

REFERENCES

- [1] M. T. Mohammed, Z. A. Khan, and A. N. Siddiquee, "Surface Modifications of Titanium Materials for developing Corrosion Behavior in Human Body Environment: A Review," *Procedia Materials Science*, vol. 6, pp. 1610–1618, Jan. 2014, <https://doi.org/10.1016/j.mspro.2014.07.144>.
- [2] F. Trevisan *et al.*, "Additive manufacturing of titanium alloys in the biomedical field: processes, properties and applications," *Journal of Applied Biomaterials & Functional Materials*, vol. 16, no. 2, pp. 57–67, Apr. 2018, <https://doi.org/10.5301/jabfm.5000371>.
- [3] M. Lepicka and M. Gradzka-Dahlke, "Surface modification of Ti6Al4V titanium alloy for biomedical applications and its effect on tribological performance - A review," *Reviews on Advanced Materials Science*, vol. 46, no. 1, pp. 86–103, 2016.
- [4] J. W. Nicholson, "Titanium Alloys for Dental Implants: A Review," *Prosthesis*, vol. 2, no. 2, pp. 100–116, Jun. 2020, <https://doi.org/10.3390/prosthesis2020011>.
- [5] W. Liu, S. Liu, and L. Wang, "Surface Modification of Biomedical Titanium Alloy: Micromorphology, Microstructure Evolution and Biomedical Applications," *Coatings*, vol. 9, no. 4, Apr. 2019, Art. no. 249, <https://doi.org/10.3390/coatings9040249>.
- [6] M. L. Lourenço, G. C. Cardoso, K. dos S. J. Sousa, T. A. G. Donato, F. M. L. Pontes, and C. R. Grandini, "Development of novel Ti-Mo-Mn alloys for biomedical applications," *Scientific Reports*, vol. 10, no. 1, p. 6298, Apr. 2020, <https://doi.org/10.1038/s41598-020-62865-4>.
- [7] M. Kulkarni, A. Mazare, and P. Schmuki, "Biomaterial Surface Modification Of Titanium and Titanium Alloys for Medical Applications," in *Nanomedicine*, Cheshire, UK: One Central Press, 2014, pp. 111–136.
- [8] M. J. Jackson, J. Kopac, M. Balazic, D. Bombac, M. Brojan, and F. Kosei, "Titanium and Titanium Alloy Applications in Medicine," in *Surgical Tools and Medical Devices*, W. Ahmed and M. J. Jackson, Eds. New York, NY, USA: Springer, 2016, pp. 475–517.
- [9] Damisih, I. N. Jujur, J. Sah, Agustanhakri, and D. H. Prajitno, "Characteristics microstructure and microhardness of cast Ti-6Al-4V ELI for biomedical application submitted to solution treatment," *AIP Conference Proceedings*, vol. 1964, no. 1, May 2018, Art. no. 020037, <https://doi.org/10.1063/1.5038319>.
- [10] D. Aroussi, B. Aour, and A. S. Bouaziz, "A Comparative Study of 316L Stainless Steel and a Titanium Alloy in an Aggressive Biological Medium," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 5093–5098, Dec. 2019, <https://doi.org/10.48084/etasr.3208>.
- [11] S.-W. Lee *et al.*, "Hydroxyapatite and Collagen Combination-Coated Dental Implants Display Better Bone Formation in the Peri-Implant Area Than the Same Combination Plus Bone Morphogenetic Protein-2-Coated Implants, Hydroxyapatite Only Coated Implants, and Uncoated Implants," *Journal of Oral and Maxillofacial Surgery*, vol. 72, no. 1, pp. 53–60, Jan. 2014, <https://doi.org/10.1016/j.joms.2013.08.031>.
- [12] A. Boudjemline, M. Boujelbene, and E. Bayraktar, "Surface Quality of Ti-6Al-4V Titanium Alloy Parts Machined by Laser Cutting," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6062–6067, Aug. 2020, <https://doi.org/10.48084/etasr.3719>.
- [13] F. Zhang, A. Weidmann, J. B. Nebe, U. Beck, and E. Burkel, "Preparation, microstructures, mechanical properties, and cytocompatibility of TiMn alloys for biomedical applications," *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, vol. 94B, no. 2, pp. 406–413, 2010, <https://doi.org/10.1002/jbm.b.31668>.
- [14] P. Shanmugapriya, V. Srinivasan, B. Karthikeyan, and T.V.Rajamurugan, "Surface modification of nanocomposite Al2O3/Gr/HAP coating for improving wear and corrosion behaviour on Ti-6Al-4V alloy using sol-gel technique," *Multiscale and Multidisciplinary Modeling, Experiments and Design*, Jan. 2021, <https://doi.org/10.1007/s41939-021-00089-3>.
- [15] L. Xia, Y. Xie, B. Fang, X. Wang, and K. Lin, "In situ modulation of crystallinity and nano-structures to enhance the stability and osseointegration of hydroxyapatite coatings on Ti-6Al-4V implants," *Chemical Engineering Journal*, vol. 347, pp. 711–720, Sep. 2018, <https://doi.org/10.1016/j.cej.2018.04.045>.
- [16] X. Liu, S. Chen, J. K. H. Tsoi, and J. P. Matinlinna, "Binary titanium alloys as dental implant materials—a review," *Regenerative Biomaterials*, vol. 4, no. 5, pp. 315–323, Oct. 2017, <https://doi.org/10.1093/rb/rbx027>.
- [17] M. Alqattan, L. Peters, Y. Alshammari, F. Yang, and L. Bolzoni, "Antibacterial Ti-Mn-Cu alloys for biomedical applications," *Regenerative Biomaterials*, vol. 8, Feb. 2021, Art. no. rbaa050, <https://doi.org/10.1093/rb/rbaa050>.
- [18] T. Hryniewicz, K. Rokosz, J. Valicek, and R. Rokicki, "Effect of magnetoelectropolishing on nanohardness and Young's modulus of

- titanium biomaterial," *Materials Letters*, vol. 83, pp. 69–72, Sep. 2012, <https://doi.org/10.1016/j.matlet.2012.06.010>.
- [19] L. Benea, E. Mardare-Danaila, M. Mardare, and J.-P. Celis, "Preparation of titanium oxide and hydroxyapatite on Ti-6Al-4V alloy surface and electrochemical behaviour in bio-simulated fluid solution," *Corrosion Science*, vol. 80, pp. 331–338, Mar. 2014, <https://doi.org/10.1016/j.corsci.2013.11.059>.
- [20] J. Jakubowicz, "Special Issue: Ti-Based Biomaterials: Synthesis, Properties and Applications," *Materials*, vol. 13, no. 7, p. 1696, Jan. 2020, <https://doi.org/10.3390/ma13071696>.
- [21] U. F. Gunpath and H. Le, "A Review of In-Situ Grown Nanocomposite Coatings for Titanium Alloy Implants," *Journal of Composites Science*, vol. 4, no. 2, Jun. 2020, Art. no. 41, <https://doi.org/10.3390/jcs4020041>.
- [22] U.-W. Jung *et al.*, "Surface characteristics of a novel hydroxyapatite-coated dental implant," *Journal of Periodontal & Implant Science*, vol. 42, no. 2, pp. 59–63, Apr. 2012, <https://doi.org/10.5051/jpis.2012.42.2.59>.
- [23] Y. Kirmanidou *et al.*, "New Ti-Alloys and Surface Modifications to Improve the Mechanical Properties and the Biological Response to Orthopedic and Dental Implants: A Review," *BioMed Research International*, vol. 2016, Jan. 2016, Art. no. e2908570, <https://doi.org/10.1155/2016/2908570>.
- [24] F. Findik, "Surface Treatment of Ti-Alloys," *Current Trends in Biomedical Engineering & Biosciences*, vol. 15, no. 3, Jun. 2018, Art. no. 555911, <https://doi.org/10.19080/CTBEB.2018.15.555911>.
- [25] J. Sharan, S. V. Lale, V. Koul, M. Mishra, and O. P. Kharbanda*, "An Overview of Surface Modifications of Titanium and its Alloys for Biomedical Applications," *Trends in Biomaterials & Artificial Organs*, vol. 29, no. 2, pp. 176–187, Jun. 2015.
- [26] A. Jaafar, C. Hecker, P. Arki, and Y. Joseph, "Sol-Gel Derived Hydroxyapatite Coatings for Titanium Implants: A Review," *Bioengineering*, vol. 7, no. 4, Dec. 2020, Art. no. 127, <https://doi.org/10.3390/bioengineering7040127>.
- [27] R. Family, M. Solati-Hashjin, S. Namjoy Nik, and A. Nemati, "Surface modification for titanium implants by hydroxyapatite nanocomposite," *Caspian Journal of Internal Medicine*, vol. 3, no. 3, pp. 460–465, 2012.
- [28] B.-O. Taranu, A. I. Bucur, and I. Sebarchievici, "Three-step procedure for the deposition of hydroxyapatite coatings," *Journal of Coatings Technology and Research*, vol. 17, no. 4, pp. 1075–1082, Jul. 2020, <https://doi.org/10.1007/s11998-020-00318-3>.
- [29] M. H. Fathi and F. Azam, "Novel hydroxyapatite/tantalum surface coating for metallic dental implant," *Materials Letters*, vol. 61, no. 4, pp. 1238–1241, Feb. 2007, <https://doi.org/10.1016/j.matlet.2006.07.013>.
- [30] Q. Chen and G. A. Thouas, "Metallic implant biomaterials," *Materials Science and Engineering: R: Reports*, vol. 87, pp. 1–57, Jan. 2015, <https://doi.org/10.1016/j.mser.2014.10.001>.
- [31] S. S. Bhasin, E. Perwez, S. Sachdeva, and R. Mallick, "Trends in prosthetic biomaterials in implant dentistry," *Journal of the International Clinical Dental Research Organization*, vol. 7, no. 3, pp. 148–159, Dec. 2015, <https://doi.org/10.4103/2231-0754.172936>.
- [32] M. Manoj, R. Subbiah, D. Mangalaraj, N. Ponpandian, C. Viswanathan, and K. Park, "Influence of Growth Parameters on the Formation of Hydroxyapatite (HAp) Nanostructures and Their Cell Viability Studies," *Nanobiomedicine*, vol. 2, Jan. 2015, Art. no. 2, <https://doi.org/10.5772/60116>.
- [33] M. Lukaszewska-Kuska, P. Krawczyk, A. Martyla, W. Hedzelek, and B. Dorocka-Bobkowska, "Hydroxyapatite coating on titanium endosseous implants for improved osseointegration: Physical and chemical considerations," *Advances in Clinical and Experimental Medicine: Official Organ Wrocław Medical University*, vol. 27, no. 8, pp. 1055–1059, Aug. 2018, <https://doi.org/10.17219/acem/69084>.
- [34] H. Daugaard, B. Elmengaard, J. E. Bechtold, T. Jensen, and K. Soballe, "The effect on bone growth enhancement of implant coatings with hydroxyapatite and collagen deposited electrochemically and by plasma spray," *Journal of Biomedical Materials Research Part A*, vol. 92A, no. 3, pp. 913–921, 2010, <https://doi.org/10.1002/jbma.a.32303>.
- [35] A. Szczes, L. Holysz, and E. Chibowski, "Synthesis of hydroxyapatite for biomedical applications," *Advances in Colloid and Interface Science*, vol. 249, pp. 321–330, Nov. 2017, <https://doi.org/10.1016/j.cis.2017.04.007>.
- [36] A. Fihri, C. Len, R. S. Varma, and A. Solhy, "Hydroxyapatite: A review of syntheses, structure and applications in heterogeneous catalysis," *Coordination Chemistry Reviews*, vol. 347, pp. 48–76, Sep. 2017, <https://doi.org/10.1016/j.ccr.2017.06.009>.
- [37] K. Kuroda and M. Okido, "Hydroxyapatite Coating of Titanium Implants Using Hydroprocessing and Evaluation of Their Osteoconductivity," *Bioinorganic Chemistry and Applications*, vol. 2012, Feb. 2012, Art. no. e730693, <https://doi.org/10.1155/2012/730693>.
- [38] P. Habibovic, F. Barrere, C. A. V. Blitterswijk, K. de Groot, and P. Layrolle, "Biomimetic Hydroxyapatite Coating on Metal Implants," *Journal of the American Ceramic Society*, vol. 85, no. 3, pp. 517–522, 2002, <https://doi.org/10.1111/j.1151-2916.2002.tb00126.x>.
- [39] V. Rattan, T. S. Sidhu, and M. Mittal, "An Overview of Hydroxyapatite Coated Titanium Implants," *Asian Journal of Engineering and Applied Technology*, vol. 1, no. 2, pp. 40–43, Jul. 2012.
- [40] G. Ciobanu and M. Harja, "Investigation on hydroxyapatite coatings formation on titanium surface," *IOP Conference Series: Materials Science and Engineering*, vol. 444, Nov. 2018, Art. no. 032007, <https://doi.org/10.1088/1757-899X/444/3/032007>.
- [41] M. Hadidi *et al.*, "Electrophoretic-deposited hydroxyapatite-copper nanocomposite as an antibacterial coating for biomedical applications," *Surface and Coatings Technology*, vol. 321, pp. 171–179, Jul. 2017, <https://doi.org/10.1016/j.surfcoat.2017.04.055>.
- [42] M. Komath, P. Rajesh, C. V. Muralidharan, H. K. Varma, R. Reshmi, and M. K. Jayaraj, "Formation of hydroxyapatite coating on titanium at 200°C through pulsed laser deposition followed by hydrothermal treatment," *Bulletin of Materials Science*, vol. 2, no. 34, pp. 389–399, 2011, <https://doi.org/10.1007/s12034-011-0069-5>.
- [43] M. Furko, K. Balazsi, and C. Balazsi, "Comparative study on preparation and characterization of bioactive coatings for biomedical applications—A review on recent patents and literature," *Reviews on Advanced Materials Science*, vol. 48, no. 1, pp. 25–51, 2017.
- [44] K. D. Patel, R. K. Singh, J.-H. Lee, and H.-W. Kim, "Electrophoretic coatings of hydroxyapatite with various nanocrystal shapes," *Materials Letters*, vol. 234, pp. 148–154, Jan. 2019, <https://doi.org/10.1016/j.matlet.2018.09.066>.
- [45] L. Sorkhi, M. Farrokhi-Rad, and T. Shahrabi, "Electrophoretic Deposition of Hydroxyapatite–Chitosan–Titania on Stainless Steel 316 L," *Surfaces*, vol. 2, no. 3, pp. 458–467, Sep. 2019, <https://doi.org/10.3390/surfaces2030034>.
- [46] M. Gardon, A. Concustell, S. Dosta, N. Cinca, I. G. Cano, and J. M. Guilemany, "Improved bonding strength of bioactive cermet Cold Gas Spray coatings," *Materials Science and Engineering: C*, vol. 45, pp. 117–121, Dec. 2014, <https://doi.org/10.1016/j.msec.2014.08.053>.
- [47] M. Rana, N. Akhtar, S. Rahman, H. Jamil, and S. Asaduzzaman, "Extraction of Hydroxyapatite from Bovine and Human Cortical Bone by Thermal Decomposition and Effect of Gamma Radiation: A Comparative Study," *International Journal of Complementary & Alternative Medicine*, vol. 8, no. 3, Aug. 2017, Art. no. 263, <https://doi.org/10.15406/ijcam.2017.08.00263>.
- [48] R. C. Rocha *et al.*, "Surface, microstructural, and adhesion strength investigations of a bioactive hydroxyapatite-titanium oxide ceramic coating applied to Ti-6Al-4V alloys by plasma thermal spraying," *Materials Research*, vol. 21, no. 4, 2018, Art. no. e20171144, <https://doi.org/10.1590/1980-5373-mr-2017-1144>.
- [49] V. Jordanovova, M. Losertova, M. Stencek, T. Lukasova, G. Simha Martynkova, and P. Peikertova, "Microstructure and Properties of Nanostructured Coating on Ti6Al4V," *Materials*, vol. 13, no. 3, Jan. 2020, Art. no. 708, <https://doi.org/10.3390/ma13030708>.
- [50] A. K. Khanra, H. C. Jung, S. H. Yu, K. S. Hong, and K. S. Shin, "Microstructure and mechanical properties of Mg-HAP composites," *Bulletin of Materials Science*, vol. 33, no. 1, pp. 43–47, Feb. 2010, <https://doi.org/10.1007/s12034-010-0006-z>.
- [51] H. Naseri, M. Ghatee, A. Yazdani, M. Mohammadi, and S. Manafi, "Characterization of the 3YSZ/CNT/HAP coating on the Ti6Al4V alloy by electrophoretic deposition," *Journal of Biomedical Materials*

- Research Part B: Applied Biomaterials, <https://doi.org/10.1002/jbm.b.34799>.
- [52] P. A. F. Sossa *et al.*, "Comparative study between natural and synthetic Hydroxyapatite: structural, morphological and bioactivity properties," *Materia (Rio de Janeiro)*, vol. 23, no. 4, 2018, <https://doi.org/10.1590/s1517-707620180004.0551>.
- [53] G. C. Gomes, F. F. Borghi, R. O. Ospina, E. O. López, F. O. Borges, and A. Mello, "Nd:YAG (532nm) pulsed laser deposition produces crystalline hydroxyapatite thin coatings at room temperature," *Surface and Coatings Technology*, vol. 329, pp. 174–183, Nov. 2017, <https://doi.org/10.1016/j.surfcoat.2017.09.008>.
- [54] A. A. Abdeltawab, M. A. Shoeib, and S. G. Mohamed, "Electrophoretic deposition of hydroxyapatite coatings on titanium from dimethylformamide suspensions," *Surface and Coatings Technology*, vol. 206, no. 1, pp. 43–50, Oct. 2011, <https://doi.org/10.1016/j.surfcoat.2011.06.034>.
- [55] H. Gheisari and E. Karamian, "Preparation and characterization of hydroxyapatite reinforced with hardystonite as a novel bio-nanocomposite for tissue engineering," *Nanomedicine Journal*, vol. 2, no. 2, pp. 141–152, Apr. 2015, <https://doi.org/10.7508/nmj.2015.02.006>.
- [56] L. Mohan, D. Durgalakshmi, M. Geetha, T. S. N. Sankara Narayanan, and R. Asokamani, "Electrophoretic deposition of nanocomposite (HAp+TiO₂) on titanium alloy for biomedical applications," *Ceramics International*, vol. 38, no. 4, pp. 3435–3443, May 2012, <https://doi.org/10.1016/j.ceramint.2011.12.056>.

AUTHORS PROFILE

Shafaq Asrar is currently serving as a Lecturer, and a Ph.D scholar in the department of Metallurgy Engineering, NED University.

Muhammad Tufail is currently serving as a Pro-Vice Chancellor of NED University. He has acquired his Ph.D from the University of Nottingham, UK

Ali Dad Chandio is currently serving as a Chairman of the Metallurgical Engineering Department, NED University. He has acquired his Ph.D from the University of Manchester, UK

An Ultra-Wideband Band Pass Filter using Metal Insulator Metal Waveguide for Nanoscale Applications

Surendra Kumar Bitra

Department of Electronics and Communication Engineering
Koneru Lakshamaiah Education Foundation
Vaddeswaram, Guntur District, Andhra Pradesh, India
bitrasurendrakumar@gmail.com

Sridhar Miriyala

Department of Electronics and Communication Engineering
Koneru Lakshamaiah Education Foundation
Vaddeswaram, Guntur District, Andhra Pradesh, India
Sridhar.m@kluniversity.in

Abstract—A T-stub Square Ring Resonator (SRR) based Ultra-Wide Band (UWB) Band Pass Filter (BPF) is studied and investigated in this paper. The proposed filter is based on coupled feed line connected to the T-stub SRR. Ultra-wideband characteristics can be realized by adjusting the T-stub lengths and coupling the gaps between both sides of waveguides and SRR. The characteristics of the T-stub SRR show that the miniaturized UWB BPF can be operated at THz frequencies. The proposed UWB filter is simulated and analyzed using the Finite Differential Time Domain (FDTD) solver-based Computer Simulation Technology (CST) studio suite. The resonance conditions are explained and the transmission performance of the filter agrees with the simulated and theoretical calculations. The proposed filter is best suitable for Electronic-Plasmonic Integrated Circuits (EPICs).

Keywords—Metal-Insulator-Metal (MIM); ring resonator; coupled lines; Finite Differential Time Domain (FDTD); Ultra-Wide Band (UWB); T-stub

I. INTRODUCTION

The light incident on a metal-insulator region produces a high-speed electromagnetic (EM) wave called Surface Plasmon Polarity (SPP) [1]. SPPs are known for overcoming the diffraction limit of the light at nanoscale wavelengths. Due to the low loss and high confinement of light at nanoscale range the Metal-Insulator-Metal (MIM) waveguide has been proposed by several researchers [2-4]. The subwavelength optical components that have been proposed when using MIM waveguide are reported in [5-9]. Several researchers worked on theories and experimentations of stub resonators like disk resonators [10], teeth-shaped resonators [11], rectangular [12], triangular [13], square [14], and circular [15] ring resonators. A Band Pass Filter (BPF) with symmetric side couple nano-disk resonators is analyzed for 1310nm and 1550nm in [16]. A concurrent dual band BPF is designed using slot waveguides and the performance is carried out at the resonant frequencies of 1300nm and 1600nm respectively in [17]. A tunable stepped impedance ring resonator for dual band BPF is designed and analysed at O and L bands in [18]. However, most of the investigations are carried out in single-, dual- and triple-band of operation for BPF. A limited work is carried out in Ultra-Wide

Band (UWB) BPF where the bandwidth is limited, and the size of UWB filters is more. The optical bands used for transmission are O-band (1260-1360nm), E-band (1360-1460nm), S-band (1460-1530nm), C-band (1530-1565nm), and the L-band (1565-1625nm). The ring resonators are a fundamental component used in the Electronic-Photonic Integrated Circuits (EPICs). MIM and Insulator Metal-Insulator (IMI) are important geometries used in plasmonic waveguides. IMI waveguides carry light in longer propagation distances than MIM waveguides due to their cladding layers, but the MIM structures have higher confinement. Due to the high confinement and moderate propagation lengths, the MIM structures are efficiently used by EPICs. Authors in [19] designed a two-dimensional structure consisting of a MIM waveguide with a ring resonator, and two resonant peaks observed. This work is a further extension of the SRR for UWB applications.

Most of the existing resonators have symmetrical shapes. In this work, a nanoplasmonic T-stub loaded Square Ring Resonator (SRR) at optical bands is proposed. Most of the existing work was carried out on the design of single and dual-operating bands. In our proposed work, more than one band of frequencies are considered simultaneously (Ultra-Wide Band-UWB).

II. T-STUB SRR FILTER DESIGN

SRR filter for dual band and SRR with stubs for triple band applications have been designed and their performance was investigated in our previous work [20]. In this paper, a T-stub SRR filter is designed which is suitable for UWB applications with compact dimensions. The methodology of the proposed filter design is illustrated with the flowchart shown in Figure 1. Firstly, the T-Stub SRR is designed using the Computer Simulation Technology (CST) studio suite. Simulations are carried out under Perfect Magnetic Layer (PML) boundary conditions for getting suitable transmission and reflection parameters. Figure 2 represents the proposed UWB filter designed using SRR. Four T-stubs are included at the four sides of the square ring representing the proposed T-stub SRR. The blue color in Figure 2 represents the metal and the gray color

Corresponding author: Sridhar Miriyala

indicates the insulator. Silver is used as metal and silica is used as an insulator. The silver metal permittivity is calculated using the Drude model [21], i.e. $\epsilon(m) = 1 - \omega_p^2 / (\omega(\omega + i\gamma_p^2))$, where $\omega_p (=1.38 \times 10^{16})$ is the plasma frequency and $\gamma_p (=2.73 \times 10^{13})$ is the collision frequency. SiO_2 ($\epsilon_i = 2.5$) is used as an insulator and its permittivity is considered constant [21].

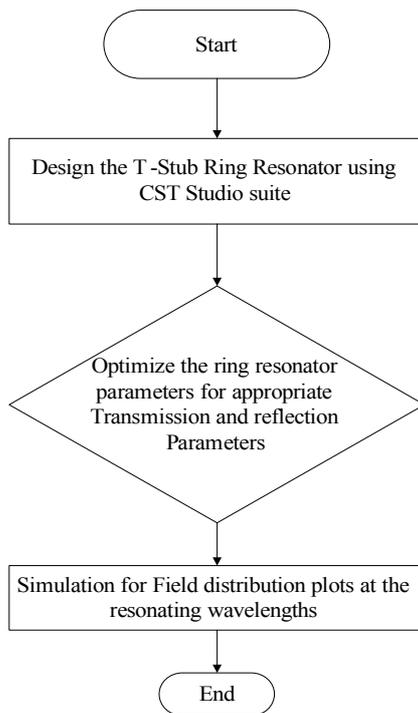


Fig. 1. Methodology of proposed T-stub ring resonator.

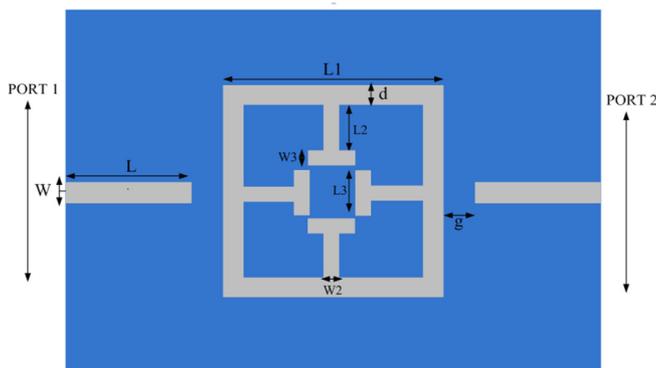


Fig. 2. SRR based UWB filter.

Two waveguides are situated on both sides of the T-stub SRR filter. The dimensions of the waveguide, T-stubs and SRR are tabulated in Table I. The basic properties of MIM waveguide like refractive index (N_{eff}), propagation length (PL) of SPP and calculated dispersion relation are defined in [17]. The proposed T-stub SRR filter is designed and analyzed using a commercially available Finite Differential Time Domain

(FDTD) solver-based CST studio suite. PML boundary conditions with mesh size of $5\text{nm} \times 5\text{nm}$ were used in simulations. Two power monitors were included at the two ports of the filter (port 1 and port 2) to measure the incident power (P_{in}) and transmitted power (P_{out}). At port 1 (input port), part of the EM waves is reflected at the interface of the MIM waveguide and the remaining part of the EM waves is coupled to the SRR cavity with a small gap (g). The transmitting and reflecting EM waves in the SRR form standing waves and these are again coupled to the output port (port 2). The rectangular ring resonance condition is given by [12]:

$$\lambda_m = \frac{\lambda_0}{(\beta_g/\beta_0)} = \frac{l_1 + l_2}{m} \quad (1)$$

where $m=0,1,2,3,\dots$, (β_g/β_0) is the normalized propagation constant, and λ_0 is the fundamental wavelength. The overall length of the SRR can be calculated from $L = N\lambda_g = N(\lambda_0/\text{Re}(n_{eff}))$ where $N=1,2,3,\dots$, λ_g is the guided wavelength and n_{eff} is the effective refractive index of the MIM waveguide.

TABLE I. DIMENSIONS OF SRR BASED UWB FILTER

S.No	Parameter	Dimensions (nm)
1	L	100
2	W	60
3	$L1$	1100
4	d	60
5	$L2$	295
6	$W2$	50
7	$L3$	275
8	$W3$	55
9	g	3

Simulations were carried out for rectangular ring resonator observed the dual band operation [11]. Including the T-stubs on the four sides of the SRR forms the proposed UWB structure and operated O, E, S, and C bands simultaneously. The proposed filter is very useful for EPICs. The fabrication of MIM waveguide can be done by film deposition and photolithography techniques.

III. RESULTS AND DISCUSSION

The reflection and transmission characteristics of the T-stub SRR are represented in Figure 3. The simulations were carried in the FDTD solver-based CST microwave studio suite under PML boundary conditions and mesh settings of $5\text{nm} \times 5\text{nm}$. The simulation results of Figure 3(a)-(b) clearly show the filter operating in the O, E, S, and C bands simultaneously. The transmission characteristics of the proposed T-stub SRR depends on the electromagnetic resonance of the SRR. The transmission and reflection characteristics are represented in Figure 3(a) for varying $L3$, from 290 to 300 with a step size of 5nm. For increasing $L3$ the covering UWB is moved right as shown in Figure 3(a). The optimum length $L3$ is 295nm. The bandwidth observed is $\sim 300\text{nm}$ by the T-stub SRR. By varying the optimum width $W3$ of the T-slot from 50nm to 60nm with a step size of 5nm, the UWB spectrum covers higher wavelengths i.e from 1240nm to 1570nm. However, the

reflection coefficient is near to the -10dB range bound. The optimum width for the T-stub is 55nm. The other parameters L , W , $W2$, d , and $L2$ have negligible effect on the UWB spectrum. The varied $W3$ reflection and Transmission coefficient is represented in the Figure 3(b). By varying the gap g between the coupled waveguide and SRR, the operating UWB band shifts to the right (Figure 3(c)). The value of g varies from 2 to 4nm with a step size of 1nm. For the 2nm gap, the reflection coefficient is above -10dB for wavelengths from 1360nm to 1450nm, for 3nm the RC is less than -10dB throughout the bands, and for 4nm gap the bands are decreasing. Due to that, 3nm is considered the optimum gap of the proposed UWB filter. The bandwidth of the proposed UWB observed from the simulation results is ~300nm. The symmetrical MIM waveguides are located on both sides of the T-stub SRR, the electromagnetic wave in the T-stub SRR will form the stable resonance observed at three wavelengths shown in Figure 3.

Extinction Ratio (ER), Bandwidth (BW), Area ($L \times W$) occupied by the filter, and application bands. The proposed filter provides an improved bandwidth of 310nm when compared with the bandwidth (150nm) of the filter presented in [23]. Even though the bandwidth is low for the proposed filter in comparison with the filter of [4] (320nm), it operates in four operating bands (O, S, C, and E).

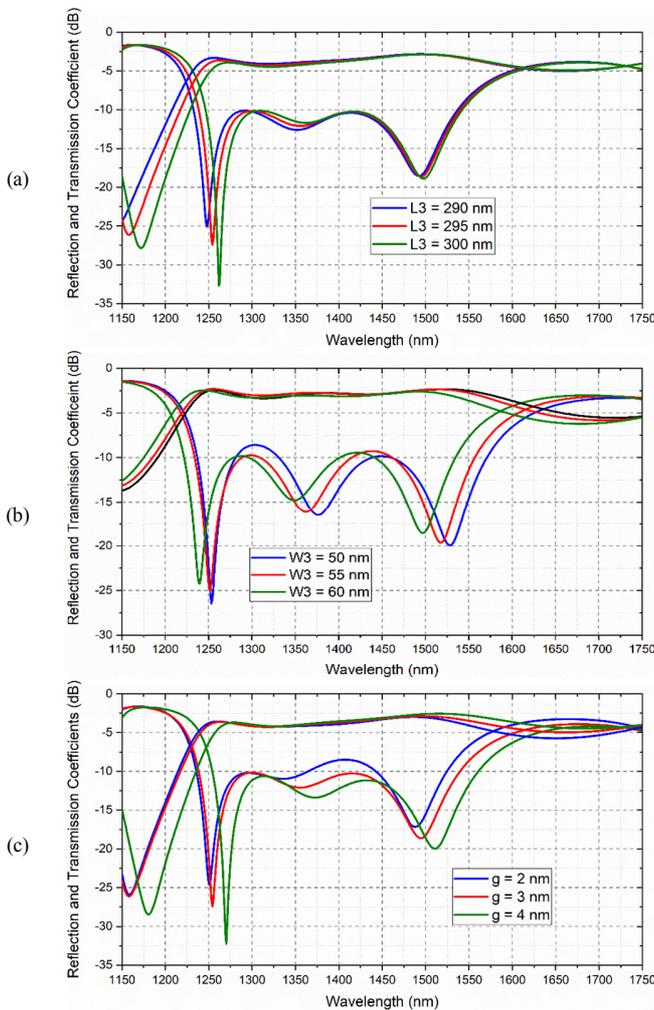


Fig. 3. Reflection and transmission coefficients for: (a) varying $L3$, (b) varying $W3$, and (c) varying gap g between the coupled feed and the SRR.

The proposed UWB filter is compared with existing filter models in terms of parameters like operating wavelength,

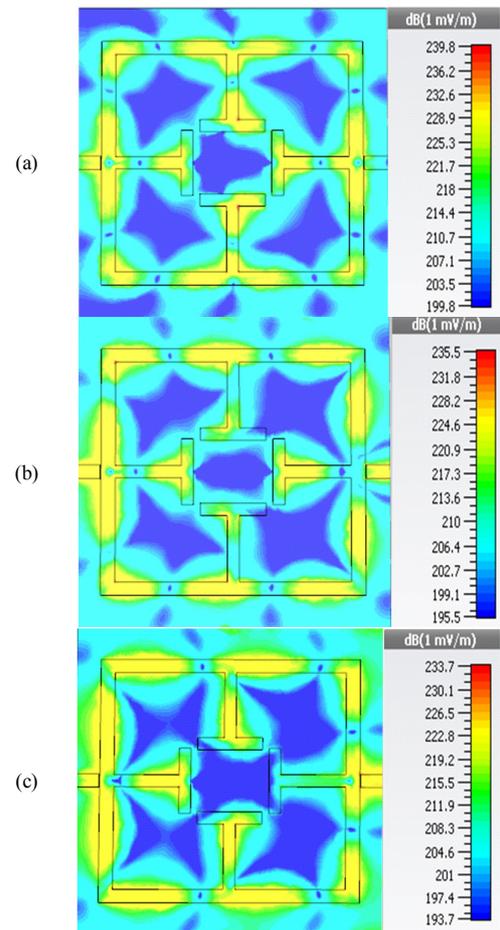


Fig. 4. Field distributions: (a) 1264.94nm (237THz), (b) 1362.69nm (220THz), and (c) 1495.9nm (200.4THz).

The proposed UWB T-stub SRR is giving better performance in terms of BW, filter size, and application bands as can be seen in Table II.

TABLE II. COMPARISON OF THE PROPOSED FILTER WITH EXISTING MODELS

Filter models	Operating wavelength (nm)	ER	BW (nm)	$L \times W$ (nm^2)	Application bands
BPF based on ARR [4]	1190–1510	~38	320	1100×1620	O, S, C and partial E
BPF based on U-shaped RR [23]	1280–1440	~26.5	150	1120×3232	O, S
Proposed filter based on SRR	1240–1550	~28.1	310	1200×1506	O, S, C, and E

The energy of the SPP wave at UWB wavelengths shows the stronger coupling of the T-stub SRR filter at 1264.94nm (237THz), 1362.69nm (220THz), and 1495.9nm (200.4THz) wavelengths as shown in Figure 3. According to the simulation, the odd mode analysis has given better results than the even mode analysis of the T-stub SRR filter.

IV. CONCLUSION

In this paper, the T-stub SRR filter for ultra-wide band applications based on transmission characteristics was studied. The coupled line MIM plasmonic structure is composed of two straight waveguides and the T-stub SRR. Simulations were carried out using the FDTD based CST studio suite. The simulation results show that the T-stub SRR filter gives better performance in O, E, S, and C optical bands. The transmission is slightly smaller at bigger wavelengths when comparing with the transmission at smaller wavelengths. The proposed T-stub SRR is optimized for the parameters of $L3$ (length of the stub) and gap g between the MIM waveguide and SRR. The ER for the proposed T-Stub UWB filter is ~ 28.1 at 1252nm and ~ 17.5 at 1492nm respectively. The bandwidth for the UWB band is 310nm. The optimized parameters also show that the proposed filter is best suitable for optical UWB applications. A strong EM coupling was observed at the MIM waveguides and SRR throughout the UWB range. The proposed UWB is best suitable for EPICs for transmitting high speed data. These structures are fabricated using lithographic techniques.

REFERENCES

- [1] W. Mu and J. B. Ketterson, "Long-range surface plasmon polaritons propagating on a dielectric waveguide support," *Optics Letters*, vol. 36, no. 23, pp. 4713–4715, Dec. 2011, <https://doi.org/10.1364/OL.36.004713>.
- [2] W. Liu, R. Zhong, J. Zhou, Y. Zhang, M. Hu, and S. Liu, "Investigations on a nano-scale periodical waveguide structure taking surface plasmon polaritons into consideration," *Journal of Physics D-applied Physics - J PHYS-D-APPL PHYS*, vol. 45, May 2012, <https://doi.org/10.1088/0022-3727/45/20/205104>.
- [3] K. M. Krishna, M. Z. U. Rahman, and S. Y. Fathima, "A New Plasmonic Broadband Branch-Line Coupler for Nanoscale Wireless Links," *IEEE Photonics Technology Letters*, vol. 29, no. 18, pp. 1568–1571, Sep. 2017, <https://doi.org/10.1109/LPT.2017.2737036>.
- [4] M. Z. U. Rahman, K. M. Krishna, K. K. Reddy, M. V. Babu, S. S. Mirza, and S. Y. Fathima, "Ultra-Wide-Band Band-Pass Filters Using Plasmonic MIM Waveguide-Based Ring Resonators," *IEEE Photonics Technology Letters*, vol. 30, no. 19, pp. 1715–1718, Oct. 2018, <https://doi.org/10.1109/LPT.2018.2866966>.
- [5] M. Vishwanath, H. Khan, and K. Thirupathiah, "Concurrent dual band filters using plasmonic MIM waveguide ring resonator," *ARNP Journal of Engineering and Applied Sciences*, vol. 13, no. 5, pp. 1813–1818, Mar. 2018.
- [6] M. Vishwanath and H. Khan, "Systematic analysis of T-junctions using plasmonic MIM waveguide," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 9, no. 14, pp. 1862–1868, Dec. 2017.
- [7] G. Veronis and S. Fan, "Modes of Subwavelength Plasmonic Slot Waveguides," *Journal of Lightwave Technology*, vol. 25, no. 9, pp. 2511–2521, Sep. 2007, <https://doi.org/10.1109/JLT.2007.903544>.
- [8] V. Mittapalli and H. Khan, "Excitation Schemes of Plasmonic Angular Ring Resonator-Based Band-Pass Filters Using a MIM Waveguide," *Photonics*, vol. 6, no. 2, Jun. 2019, Art. no. 41, <https://doi.org/10.3390/photronics6020041>.
- [9] M. Vishwanath and D. Habibullakhan, "Design and Simulation of Compact Band-Pass Filter Using Stub Loaded Plasmonic MIM Waveguide," *International Journal of Engineering & Technology*, vol. 7, no. 3.3, pp. 41–43, Jun. 2018, <https://doi.org/10.14419/ijet.v7i3.3.14482>.
- [10] H. Lu, X. Liu, D. Mao, L. Wang, and Y. Gong, "Tunable band-pass plasmonic waveguide filters with nanodisk resonators," *Optics Express*, vol. 18, no. 17, pp. 17922–17927, Aug. 2010, <https://doi.org/10.1364/OE.18.017922>.
- [11] X. Lin and X. Huang, "Numerical modeling of a teeth-shaped nanoplasmonic waveguide filter," *Journal of the Optical Society of America B*, vol. 26, no. 7, pp. 1263–1268, Jul. 2009, <https://doi.org/10.1364/JOSAB.26.001263>.
- [12] B. Yun, G. Hu, and Y. Cui, "Theoretical analysis of a nanoscale plasmonic filter based on a rectangular metal-insulator-metal waveguide," *Journal of Physics D: Applied Physics*, vol. 43, no. 38, p. 385102, Sep. 2010, <https://doi.org/10.1088/0022-3727/43/38/385102>.
- [13] G.-Y. Oh, D.-G. Kim, S. H. Kim, H. C. Ki, T. U. Kim, and Y.-W. Choi, "Integrated Refractometric Sensor Utilizing a Triangular Ring Resonator Combined With SPR," *IEEE Photonics Technology Letters*, vol. 26, no. 21, pp. 2189–2192, Nov. 2014, <https://doi.org/10.1109/LPT.2014.2349975>.
- [14] J. Liu, G. Fang, H. Zhao, Y. Zhang, and S. Liu, "Plasmon flow control at gap waveguide junctions using square ring resonators," *Journal of Physics D: Applied Physics*, vol. 43, no. 5, Jan. 2010, Art. no. 055103, <https://doi.org/10.1088/0022-3727/43/5/055103>.
- [15] A. Setayesh, S. R. Mirnaziry, and M. S. Abrishamian, "Numerical Investigation of Tunable Band-pass/band-stop Plasmonic Filters with Hollow-core Circular Ring Resonator," *Journal of the Optical Society of Korea*, vol. 15, no. 1, pp. 82–89, Mar. 2011.
- [16] S. Shi *et al.*, "Enhanced plasmonic band-pass filter with symmetric dual side-coupled nanodisk resonators," *Journal of Applied Physics*, vol. 118, no. 14, Oct. 2015, Art. no. 143103, <https://doi.org/10.1063/1.4932668>.
- [17] K. Thirupathiah, N. P. Pathak, and V. Rastogi, "Concurrent Dual Band Filters Using Plasmonic Slot Waveguide," *IEEE Photonics Technology Letters*, vol. 25, no. 22, pp. 2217–2220, Nov. 2013, <https://doi.org/10.1109/LPT.2013.2283880>.
- [18] H. Lu, G. Wang, and X. Liu, "Manipulation of light in MIM plasmonic waveguide systems," *Chinese Science Bulletin*, vol. 58, no. 30, pp. 3607–3616, Oct. 2013, <https://doi.org/10.1007/s11434-013-5989-6>.
- [19] S. Bitra and S. Miriyala, "Concurrent Dual and Triple Band Square Ring Resonator Base-Band Filter using Metal-Insulator-Metal for Plasmonic Applications," *Applied Computational Electromagnetics Society Journal*, vol. 35, no. 10, Dec. 2020, <https://doi.org/10.47037/2020.ACES.J.351010>.
- [20] S. A. Maier, *Plasmonics: Fundamentals and Applications*. New York, NY, USA: Springer, 2007.
- [21] P. B. Johnson and R. W. Christy, "Optical Constants of the Noble Metals," *Physical Review B*, vol. 6, no. 12, pp. 4370–4379, Dec. 1972, <https://doi.org/10.1103/PhysRevB.6.4370>.
- [22] S. K. Bitra and M. Sridhar, "Design of Nanoscale Square Ring Resonator Band-Pass Filter Using Metal-Insulator-Metal," in *Microelectronics, Electromagnetics and Telecommunications*, Singapore, 2021, pp. 659–664, https://doi.org/10.1007/978-981-15-3828-5_68.
- [23] P. Osman, P. V. Sridevi, and K. V. S. N. Raju, "Design of Metal-Insulator-Metal based Plasmonic Ultra-Wide Band U-Shaped resonators based Band-Pass Filter," *International Journal for Research in Applied Science and Engineering Technology*, vol. 7, no. 5, pp. 3737–3719, May 2019, <https://doi.org/10.22214/ijraset.2019.5615>.

Manufacturing of Wear Resistant Iron-Steel: A Theoretical and Experimental Research on Wear Behavior

Oğuzhan Bildik

Industrial Design Engineering Department
Institute of Graduate Studies
Karabük University
Karabük, Turkey
oguzhanbildik@gmail.com

Mustafa Yaşar

Industrial Design Engineering Department
Technology Faculty
Karabük University
Karabük, Turkey
myasar@karabuk.edu.tr

Abstract—In this study, four different alloys of steel blocks with a thickness of 15mm were manufactured in order to develop an alternative to steel plates used in wear exposed areas of construction machines, trucks, and asphalt production plants. To further increase the wear resistance of the manufactured steel blocks, their thickness was reduced to 10mm by the hot-rolling method. Wear specimens were obtained from rolled blocks. These specimens were abraded at 20N, 40N, and 60N loads in reciprocating linear motion module ASTM G-33 standards to determine their wear resistance. SEM and EDX analyses were also conducted to see modifications on the worn surfaces. In addition, a theoretical model of wear behaviors was created, calculations were made with Archard wear equation and ANSYS software, and the theoretical and experimental results were compared.

Keywords—steel plates; wear; hot rolling; Archard wear analysis

I. INTRODUCTION

Steel powerfully impacts the way seven billion of us on the planet go about our daily lives, wherever we are [1]. This effect of steel in our lives brings along some engineering problems. A scientific, engineering and economic problem is the wear of mechanical parts caused by decrease of working surfaces' properties. The wear mechanisms are very complex, due to the interlinked factors. The intensity of interaction depends on the conditions of the environment, in which the mechanical parts are used as well as the type and parameters of the work. Tribological properties such as wear play a critical role in deciding the service life of components that form a moving system or in repetitive/regular contact with each other [2]. Wear resistant plates are a modern solution in the regeneration of worn machine parts and for producing new parts which bring together high wear and abrasion resistance with costs reduction. Authors in [3] examined wear resistant steel materials with different chemical ratios by different manufacturers used in construction machines and similar fields and found that wear behavior and wear amounts in materials vary significantly depending on the chemical content of the material.

In the mineral processing industry, the high stress conditions create a challenging environment for the wear protection steels. Great amounts of highly abrasive rocks are processed and transported in earth construction, excavation, mining, and mineral processing, inducing heavy abrasion, gouging and impact wear. Authors in [4] conducted wear tests at 20N, 40N, 60N, and 80N loads on steel plates with different C, Mn, and Si ratios. The C ratio varied between 0.15% and 2% and they found that the wear volume decreased by increasing C rate [4]. Wear is an important phenomenon that adversely affects the efficiency and lifespan of all machine components in contact under relative motion. In this regard, there is a need to develop new instruments to attain a better understanding of the wear phenomenon. Authors in [5] constructed a tribotester and used the topography of worn surfaces, to provide wear process and wear volume images along with more detailed correlation.

II. MODEL ANALYSIS

Prior to casting, Anycasting software was used for casting analysis of the model for molding, filling of the mold and prevention of internal cavities. In the Anycasting, casting shape, casting type, runner, and other parts were first defined with the Anypre module and the mold area was created by selecting the casting area and runner axis [6, 7]. Then, a theoretical analysis was carried out as shown in Figure 1 to get results such as casting filling time, solidification time, shrinkage cavities, which are of vital importance. After casting, corrections were made on the model by checking whether shrinkage cavities occurred on the piece. By reverse engineering the resulting model, downsprue, runner, nozzle, vent-riser design and calculations were made for casting [8-10].

III. MOLD MANUFACTURING

Most of the pieces manufactured through casting were shaped with the sand casting process. Sand mold casting was conducted by pouring molten metal into a sand mold cavity and after the solidification process, breaking away the sand mold to remove the cast piece [11]. In the model-mold design for sand

mold casting, the shrinkage allowance on the model and the values obtained from AnyCasting software were checked for relevance. The Model-Flask used in sand mold preparation is shown in Figure 2.

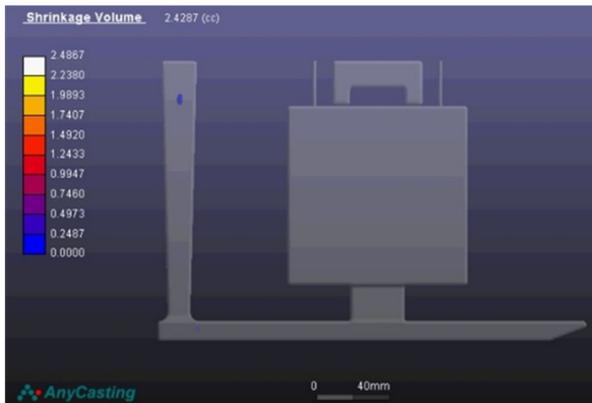


Fig. 1. Shrinkage volume display.

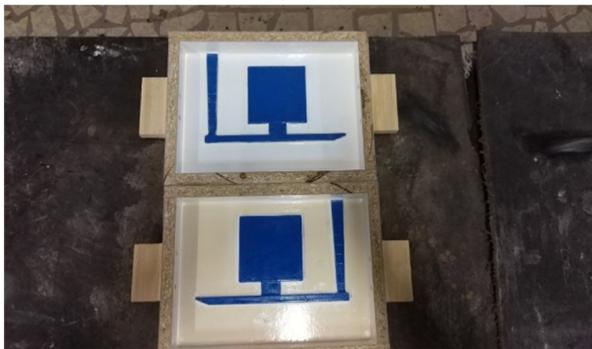


Fig. 2. Flask and model display prior to sand casting.

For the preparation of the flask, 2800g of sand, 80-85g of resin, and 20-25g of chemical hardener were weighed on a precision scale and then were mixed in a separate container manually or by using a machine with adding resin and chemical hardener for 1.5-2 minutes. The mixed sand was emptied on the model inside the flask with a foundry trowel. While emptying the mixture, special attention was paid to the sand squeezing process for proper mold formation. The molds prepared for the casting process were joined together with special adhesive-refractory material. The molds were put in a dryer furnace for dehumidification and made ready for casting before pouring.

IV. POURING PROCESS AND SPECTRAL ANALYSIS

A single material cannot meet all the high performance needs of long-lasting spare parts, so the properties of the material are improved by alloying through casting [12]. In pouring process, samples were taken from the furnace after burden materials fully melted and reached the desired temperature. The chemical analysis of the samples was carried out in a spectrometer to determine their chemical composition. Proper rates of ferroalloy were added to the first alloys in the furnace after composition values were calculated in the

analysis. After reaching the desired rate with calculations and adding ferroalloy, pouring process was completed. The mold was left for cooling and then the pieces were removed from the mold for cooling and cleaning process. The shake out after casting is shown in Figure 3. GNR Atlantis Optical Emission Spectrometer was used for spectral analysis. The rates of cast specimens with different chemical compositions are shown in Table I.



Fig. 3. The pouring moment-mold display after cooling.

TABLE I. CHEMICAL COMPOSITION OF SAMPLES AFTER CASTING

Cast specimen number	1	2	3	4
Specimen name	N1	N2	N3	N4
C	0.48	0.54	0.59	0.58
Mn	0.95-1.00	1.49	1.55	1.19
P	Max-0.016	Max-0.020	Max-0.020	Max-0.019
S	Max-0.009	Max-0.011	Max-0.010	Max-0.010
Si	0.58	0.60-0.70	0.49	0.60-0.70
Cr	1.30-1.40	1.26	1.27	1.24
Ni	1.26	1.245	1.25	1.24
Mo	0.54	0.62	0.63	0.61

V. ROLLING

Prior to the rolling process, cast specimens were cut at 100×30×15mm dimensions with a cutter in order to match the laboratory draw bench. Then the prepared specimens were heated up to 1200±10°C in a PID controlled furnace with digital temperature and program display for hot rolling process. The 100×30×15 steel cast specimens were rolled on a draw bench with 150mm press surface width and 150mm cylinder mill diameter as shown in Figure 4.



Fig. 4. Rolling process in laboratory conditions.

The rolling process was done in 3 passes. In the 1st pass the thickness was reduced to 14mm, in the 2nd pass to 12mm, and in the 3rd pass to 10mm. At the end of the rolling process, the average dimensions of the specimens were 140×37×10mm. The average, measured after rolling, Brinell hardness of the specimens is given in Table II.

TABLE III. AVERAGE BRINELL HARDNESS VALUES AFTER ROLLING

Specimen name	Brinell hardness(Hv)
N1	457
N2	612
N3	605
N4	670

VI. WEAR TEST

Tribometer T10/20 test equipment, manufactured by the UTS company, was used for wear experiments (Figure 5). Before the experiments, the specimens were cut and sandpapered and then mounted on the equipment. The tests were carried out in accordance with reciprocating linear motion module and ASTM G-133 (Standard Test Method for Linearly-Reciprocating Ball-on-Flat Sliding Wear) [11, 13]. 20N, 40N, and 60N loads were used in tests. Test parameters and load amounts are shown in Table III.



Fig. 5. Test set-up for wear testing.

TABLE II. WEAR TEST PARAMETERS

Load (F)	20N/40N/60N
Frequency	3Hz
Distance (X)	7mm
Sliding distance	250 m
Sliding velocity	42mm/s



Fig. 6. Images of wear testing and specimen after wear test.

Figure 7(a)-(d) shows the SEM views of the wear surfaces of steel specimens under 60N load. It can be seen that there are significant, superficial, groove-like scratches, along with deformation scratches that accumulate on the part that broke after wear and plastic deformation materials dispersed on its surface. It is observed that wear is less in areas with high manganese on their surfaces. Especially in the samples given in Figure 7(b) and (c), it is seen that the amount of abrasion is less

and the highest wear occurred in the sample in Figure 7(a). It has been observed that in all the figures, the materials show a serious resistance mechanism against abrasion that can be called abrasion resistant steel [14].

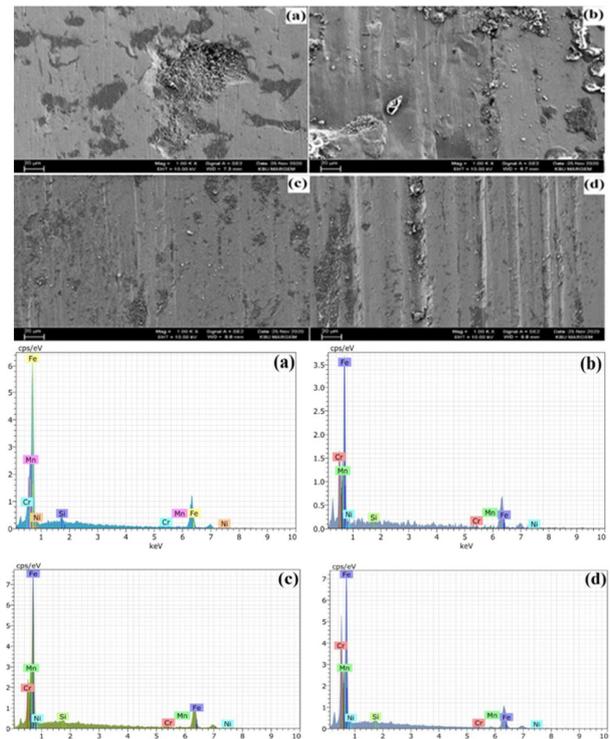


Fig. 7. SEM and EDX images of abrasion test with 60N load in 1.00Kx (a) Sample1, (b) Sample2, (c) Sample3, (d) Sample4.

VII. ARCHARD FEM ANALYSIS AND RESULT COMPARISON

The Archard wear equation [15-17] or Holm/Archard law, introduced by J.F. Archard in 1953, states the relationship among sliding wear volume (V), normal load (W), total sliding distance (L) and hardness of the softer of two contacting surfaces (H). The classical Archard's wear model is widely used in abrasive wear literature. The Archard's wear equation shows that wear rate is proportional to sliding velocity and normal contact force. The Archard wear model defines wear direction along with wear rate, contact pressure, sliding velocity, and material hardness:

$$w = \frac{K}{H} P^m v^n \quad (1)$$

Values such as wear coefficient, pressure exponent and velocity exponent were taken as constant: $C1$ = wear coefficient, $K = 1.05 \times 10^{-2}$, $C2$ = material hardness, H = variable as per specimen hardness [19], $C3$ = pressure exponent, $m = 1$, and $C4$ =velocity exponent, $n = 0$ [20]. Different hardness values were designated as variable for different specimens after casting [18]. Material hardness obtained after rolling were proportioned and hardness values were taken as 2400 for N1, 3250 for N2, 3120 for N3, and 3480 for N4.

ANSYS software was utilized for finite element analysis and the experiment set-up shown in Figure 8 was used for each specimen. The wear behavior at 250m distance was simulated for 20N, 40N, and 60N loads [21, 22].

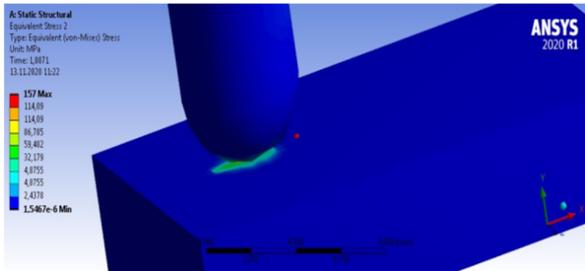


Fig. 8. Finite element analysis and wear image.

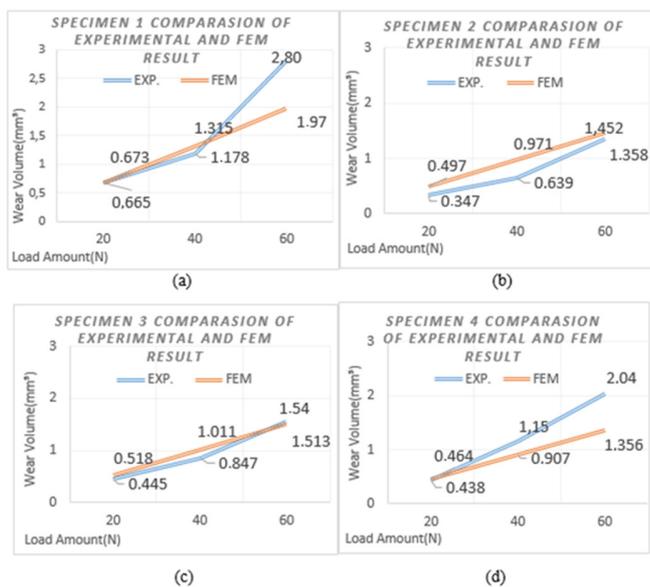


Fig. 9. Wear test applied to specimens 1,2,3, and 4 in Table I, and wear volume comparison of load amounts as a result of finite element analysis.

The result of wear experiment of Specimen 1 at 20N load was 0.665mm³ and the result from the FEM model was 0.673mm³ (Figure 9(a)). Both results shown in Figure 9(a) were quite close, suggesting that the simulation and experiment data were consistent with each other. When the load amount reached 40N, wear volume was observed to increase in accordance with the load both in the experiment and in the finite element analysis. At 60N load, the experimental wear volume was 2.80mm³, while in FEM analysis the wear volume was a bit lower with 1.97mm³. The result of wear experiment of Specimen 2 at 20N load was 0.347mm³ and the result from the FEM model was 0.497mm³ (Figure 9(b)). When the load reached 40N, the wear volume increased in both experimental and finite element analysis to 0.639mm³ and 0.971mm³ respectively. At 60N load, the experimental wear volume was 1.358mm³, while in FEM analysis the wear volume was a bit higher with 1.452mm³. In Figure 9(c) the result of wear experiment of Specimen 3 at 20N load was 0.445mm³ and the result from the FEM model was 0.518mm³. When the load

reached 40N, the experimental wear volume of the load was 0.847mm³ and 1.011mm³ for FEM. At 60N load, the experimental wear volume was 1.54mm³, while in FEM analysis it was 1.513mm³. In Figure 9(d), the result of wear experiment of Specimen 4 at 20N load was 0.438mm³ and the result from the FEM model was 0.464mm³. The results were quite close. When the load amount reached 40N, the wear volume was observed to increase in both experimental and finite element analysis with 1.15 mm³ and 0.907 mm³ respectively, in accordance with the load. At 60N load, the experimental wear volume was 2.04mm³, while in FEM analysis it was 1.356mm³.

As can be seen from the results of wear experiment at 20N load (Figure 10), the lowest wear volume was at specimen 2 with 0.347mm³ and the highest wear was 0.665mm³ at specimen 1 which had the lowest C amount. Experimental wear and finite element results were very close in all specimens. Wear experiment results and finite element simulation results at 40N load are shown in Figure 11. At 40N load, the lowest wear value was at specimen 2 with 0.639mm³ and the highest at specimen 1 with 1.178mm³. In all wear test results, except specimen 4, experimental wear volume values were lower than the FEM wear values. At 60N load, the highest wear was 2.80mm³ at specimen 1 which had the lowest C and Mn content. The wear volume was higher than the FEM value of 1.97mm³. The lowest wear was at specimen 2 and 3. These results were very close to the FEM values. The experimental wear volume for specimen 4 was higher than the FEM result.

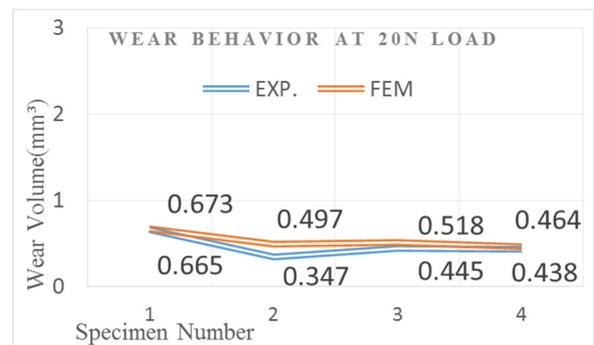


Fig. 10. Wear volume of specimens at 20 N load wear experiment.

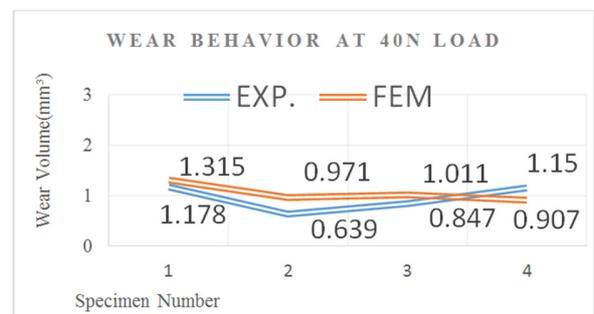


Fig. 11. Wear volume of specimens at 40N load wear experiment.

Figure 13 compares the wear values of the manufactured specimens to EN 10029 and EN 10051 standard steel materials, known as wear resistant steel plates in the market.

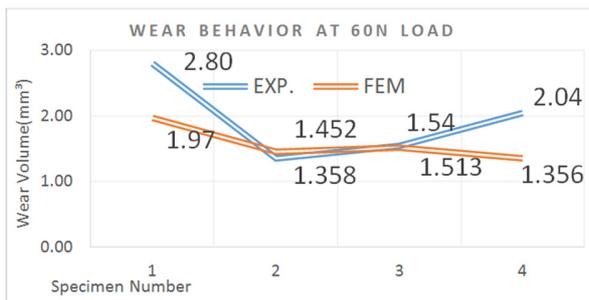


Fig. 12. Wear volume of specimens at 60N load wear experiment.

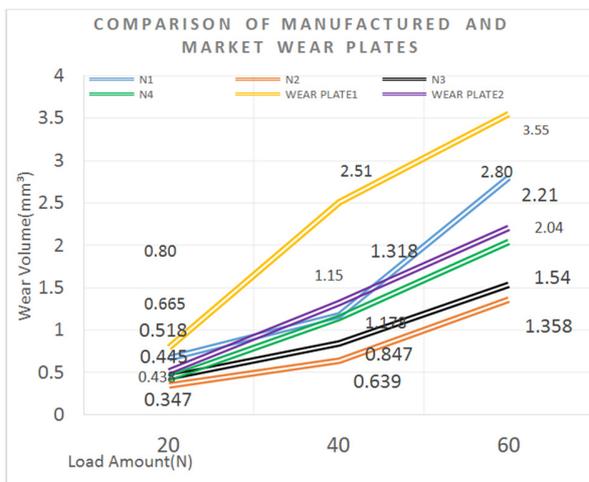


Fig. 13. Comparison of manufactured steel plates and market wear plates.

EN 10051 wear plate number 1 had the lowest wear resistance at all the tests. The EN 10051 wear plate number 2 had higher wear volume than specimens N2, N3, and N4. It had less wear volume than specimen N1 only at 20N and 60N tests. At 40N test, its wear volume with 1.318 mm³ was higher than specimen N1's.

VIII. RESULTS AND EVALUATION

This study aimed to manufacture different steel which has not yet entered the standards, and to examine its wear behavior. In particular, the aim was to provide an alternative to the wear parts in motorway vehicles which are subject to constant wear and that goal was achieved. Manufacturing conditions were simulated in virtual environment before specimens were cast. The runner and the full filling of the mold cavity with liquid metal were optimized. The mold manufacturing was done in compliance with the conditions and the plates were cast using an induction furnace. As a material standard, different alloy and rolling conditions were tried similar to EN 10029 and EN 10051 standards used in the market. In order to optimize wear conditions, finite element software and Archard wear theory were utilized. The experimental and theoretical results matched by 85%. The compliance rate was low for the results obtained from 60N wear load. The reasons are estimated to result from overheating and consequently increase in rupture at wear area. However, the heat effect was not taken into account in the theoretical model.

ACKNOWLEDGMENT

The work presented in this paper was supported by the University of Karabük Coordinatorship of Research Projects under Project No. KBUBAP-17-DR-271.

REFERENCES

- [1] N. Baluch, Z. M. Udin, and C. S. Abdullah, "Advanced High Strength Steel in Auto Industry: an Overview," *Engineering, Technology & Applied Science Research*, vol. 4, no. 4, pp. 686–689, Aug. 2014, <https://doi.org/10.48084/etasr.444>.
- [2] H. Ramadas, S. Sarkar, and A. K. Nath, "Three-body dry abrasive wear properties of 15–5 precipitation hardening stainless steel produced by laser powder bed fusion process," *Wear*, vol. 470–471, Apr. 2021, Art. no. 203623, <https://doi.org/10.1016/j.wear.2021.203623>.
- [3] M. Adamiak, G. Jacek, and T. Kik, "Comparison of abrasion resistance of selected constructional materials," *Journal of Achievements in Materials and Manufacturing Engineering*, vol. 37, no. 2, pp. 375–380, Dec. 2009.
- [4] M. Lindroos, K. Valtonen, A. Kemppainen, A. Laukkanen, K. Holmberg, and V.-T. Kuokkala, "Wear behavior and work hardening of high strength steels in high stress abrasion," *Wear*, vol. 322–323, pp. 32–40, Jan. 2015, <https://doi.org/10.1016/j.wear.2014.10.018>.
- [5] O. V. Penkov, M. Khadem, A. Nieto, T.-H. Kim, and D.-E. Kim, "Design and Construction of a Micro-Tribotester for Precise In-Situ Wear Measurements," *Micromachines*, vol. 8, no. 4, Mar. 2017, <https://doi.org/10.3390/mi8040103>.
- [6] L. Y. Ni and P. Liu, "Numerical Simulation and Production Trial of Aluminum Alloy Gear Box Housing," *Applied Mechanics and Materials*, vol. 703, pp. 232–236, 2015, <https://doi.org/10.4028/www.scientific.net/AMM.703.232>.
- [7] P.-H. Huang, J.-K. Kuo, T.-H. Fang, and W. Wu, "Numerical simulation and design of casting system for stainless steel exhaust manifold," *MATEC Web of Conferences*, vol. 185, 2018, Art. no. 00008, <https://doi.org/10.1051/mateconf/201818500008>.
- [8] H.-J. Kwon and H.-K. Kwon, "Computer aided engineering (CAE) simulation for the design optimization of gate system on high pressure die casting (HPDC) process," *Robotics and Computer-Integrated Manufacturing*, vol. 55, pp. 147–153, Feb. 2019, <https://doi.org/10.1016/j.rcim.2018.01.003>.
- [9] Z. Wenshan and L. Shuqin, "Design of drive motor of magnetic levitation artificial heart pump combined with magnetic circuit method and finite element method," *Electrical Machinery and Control Applications*, vol. 43, no. 4, pp. 71–76, 2016.
- [10] B. Dai, H. Gong, D. Kong, and Z. Zhou, "Simulation Analysis and Optimization of Die-Casting for Automobile Steering Valve Shell," *Journal of Chongqing University of Technology (Natural Science)*, vol. 11, 2016.
- [11] R. C. Morón *et al.*, "Wear Performance Under Dry and Lubricated Conditions of Post Boriding Heat Treatment in 4140 Steel," *Journal of Tribology*, vol. 143, no. 2, Aug. 2020, Art. no. 021702, <https://doi.org/10.1115/1.4047789>.
- [12] M. Ramadan, "Interface Structure and Elements Diffusion of As-Cast and Annealed Ductile Iron/Stainless Steel Bimetal Castings," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2709–2714, Apr. 2018, <https://doi.org/10.48084/etasr.1856>.
- [13] M. I. Pashechko, K. Dziedzic, E. Mendyk, and J. Jozwik, "Chemical and Phase Composition of the Friction Surfaces Fe–Mn–C–B–Si–Ni–Cr Hardfacing Coatings," *Journal of Tribology*, vol. 140, no. 2, Oct. 2017, Art. no. 021302, <https://doi.org/10.1115/1.4037953>.
- [14] M. P. Groover, *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*, 4th ed. Hoboken, NJ, USA: Wiley, 2010.
- [15] C. Dong, C. Yuan, X. Bai, J. Li, H. Qin, and X. Yan, "Coupling mechanism between wear and oxidation processes of 304 stainless steel in hydrogen peroxide environments," *Scientific Reports*, vol. 7, no. 1, May 2017, Art. no. 2327, <https://doi.org/10.1038/s41598-017-02530-5>.

- [16] J. F. Archard, "Contact and Rubbing of Flat Surfaces," *Journal of Applied Physics*, vol. 24, no. 8, pp. 981–988, Aug. 1953, <https://doi.org/10.1063/1.1721448>.
- [17] T. Dyck and A. Bund, "An adaption of the Archard equation for electrical contacts with thin coatings - PDF Free Download," *Tribology International*, no. 102, pp. 1–9, 2016.
- [18] K. Kato, "Classification of wear mechanisms/models," *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, vol. 216, no. 6, pp. 349–355, Jun. 2002, <https://doi.org/10.1243/135065002762355280>.
- [19] H. Kloß and M. Woydt, "Prediction of Tribological Limits in Sliding Contacts: Flash Temperature Calculations in Sliding Contacts and Material Behavior," *Journal of Tribology*, vol. 138, no. 3, May 2016, Art. no. 031403, <https://doi.org/10.1115/1.4033132>.
- [20] A. P. S. Arunachalam and S. Idapalapati, "Material removal analysis for compliant polishing tool using adaptive meshing technique and Archard wear model," *Wear*, vol. 418–419, pp. 140–150, Jan. 2019, <https://doi.org/10.1016/j.wear.2018.11.015>.
- [21] M. A. Mekicha, M. B. de Rooij, T. Mishra, D. T. A. Matthews, L. Jacobs, and D. J. Schipper, "Study of wear particles formation at single asperity contact: An experimental and numerical approach," *Wear*, vol. 470–471, Apr. 2021, Art. no. 203644, <https://doi.org/10.1016/j.wear.2021.203644>.
- [22] M. A. Ashraf, R. Ahmed, O. Ali, N. H. Faisal, A. M. El-Sherik, and M. F. A. Goosen, "Finite Element Modeling of Sliding Wear in a Composite Alloy Using a Free-Mesh," *Journal of Tribology*, vol. 137, no. 3, Jul. 2015, Art. no. 031605, <https://doi.org/10.1115/1.4029998>.

Harnessing Electrical Power from Hybrid Biomass-Solid Waste Energy Resources for Microgrids in Underdeveloped and Developing Countries

Muhammad Amir Raza
Department of Electrical Engineering
NED University of Engineering and Technology
Karachi, Pakistan
amir.eed.neduet@gmail.com

Krishan Lal Khatri
Department of Electrical Engineering
NED University of Engineering and Technology
Karachi, Pakistan
enrkrishan@yahoo.com

Khalid Rafique
Azad Jammu and Kashmir Information
Technology Board
Muzaffarabad, Pakistan
khalidrafiquepk@gmail.com

Abdul Sattar Saand
Electrical Engineering Department
Quaid-e-Awam University of Engineering, Science and
Technology, Nawabshah, Pakistan
as-saand@quest.edu.pk

Abstract-This paper presents an energy plan for harnessing electrical power from hybrid energy resources, including biomass and solid waste, through hybrid technologies for microgrid development to overcome the current energy crisis and provide future sustainable electricity pathways for a developing country. Biomass and solid waste were collected from different dumping sites in Pakistan and were tested for their calorific value, which was found to be 6519Kcal/Kg. The total power was calculated based on this calorific value and it was found that there is a potential of total 11,989.5GW of power generation. Thus, hybrid energy resources are suitable for harnessing electrical energy through hybrid technologies. Different hybrid systems were examined for these resources and the gasification technique is finally proposed as the most suitable method for producing energy from biomass and solid waste resources in Pakistan. This technique is also found to be economically feasible for processing all kinds of waste.

Keywords-biomass; solid waste; electrical power; sustainable development; developing countries; underdeveloped countries

I. INTRODUCTION

A developing country may face an exponential growth of population and industrialization, which is resulting in huge requirements of energy. This further leads to the exploitation of fossil fuels, such as domestic coal, natural gas, and oil which are not renewable. The use of fossil fuels poses a threat to the environment and public health in terms of harmful Greenhouse Gases (GHG) and huge production of waste [1]. The generated Municipal Solid Waste (MSW) will globally reach 2200 million tons annually by 2025 which would increase to 2600 million tons by 2030 [2, 3]. Sewage sludge produced from industrial and household wastewater treatment produces thousands of tons of solid waste daily [4]. Proper treatment of

biomass and solid waste is an alternative option for power generation which reduces ecological and social issues [5]. Facilitating biomass and solid waste renewable resources for power generation instead of fossil fuels provides better techno-economic opportunities for Pakistan [5]. If biomass and solid waste resources are exploited, the ratio of renewable power generation would be increased from 2% to 27% and would cater for the 56% of local energy needs [6]. The major advantage of biomass and solid waste is that they can be available at any location in contradiction with fossil fuels[6].

There are many biomass and solid waste resources including human waste, MSW, food waste, firewood, shrubbery waste, fabrics, paper products, latex, and plastics [7]. The quantity and composition of biomass and solid waste vary greatly depending upon the region and human living standards [7]. There are four major sectors involved in the production of biomass and solid waste in Pakistan: agriculture, domestic, industrial, and commercial. Easy accessibility to biomass and solid waste resources provides a sustainable option for harnessing electrical power while resolving the issue of inappropriate dumping of biomass and solid waste. Biomass and solid waste are considered as the most feasible option for biofuel production and biochemical energy generation [8]. Biomass and solid waste are converted into biofuel through thermochemical conversion processes including combustion, gasification, incineration, and pyrolysis [9]. Gasification is a suitable option for managing large quantities of biomass and solid waste in Pakistan because it produces multiple outputs including heat for commercial purposes, energy for residential purposes, and bio-oils for the chemical industries while it utilizes all types of biomass and solid waste products [10]. The

Corresponding author: Muhammad Amir Raza

incineration techniques, amongst all considered techniques, are responsible for large environmental pollution [10].

Economic activities in the developing countries lead to augmented generation of biomass and solid waste. These sources could be used for the generation of electrical energy, bio-oil, and biofuel [11]. The usage of biomass and solid waste is valuable because these sources are available in huge quantities [11]. The type of material and the quantity of generated waste are different in each region as shown in Table I. Pakistan generates over 64,000 tons/day of biomass and solid waste. This material is useful for power generation because it has good calorific value (6.9J/kg) [12]. In the main municipalities, the total production capacity of solid waste is around 712 million tons/year [12]. However, in India, the intermediate and massive populated cities generate a large quantity of waste which is increasing at a rate of 6% every year [13]. The production of biomass at a global level is also high [13].

TABLE I. PRODUCTION CAPACITY OF BIOMASS AND SOLID WASTE IN VARIOUS COUNTRIES

Ref	Country	Material type	Production capacity
[12]	Pakistan	Biomass	46886 million tons/year
		Solid waste	8536 million tons/year
[13]	India	Biomass	180 million tons/year
		Solid waste	600 million tons/year
[14]	Europe	Biomass	205 million tons/year
[15]	China	Biomass	850 million tons/year
[16]	Brazil	Biomass	597 million tons/year

The utilization of biomass and solid waste for power generation in Pakistan is the focus of this study. However, to overcome the environmental pollution and available energy issue in the country, the research is on the way for converting unused sources into useful sources by using hybrid technologies for the generation of energy and heat. The main aim of the current research is to develop an energy plan for the energy sector of Pakistan in order to overcome the energy crisis and identify a future path of electricity supply for sustainable development. This research identifies the quantity and quality of solid waste and biomass resources and it justifies the suitability of harnessing electrical power from them. An experiment was performed in the lab for the final testing of waste pallets and the waste calorific value was identified in order to calculate its total power capacity. Furthermore, hybrid technologies were examined and the most suitable technology for hybrid resources was selected.

II. MATERIALS AND METHODS

A. Study Area

Pakistan is located in the north western part of South Asia and covers an area of 881,913Km² with an overall population of 207 million with a growth rate of 2.4% [17]. Its neighboring countries are China from the northeastern side, India from the eastern side, Iran, and Afghanistan from the western side and the Arabian Sea from the southern side. The growth rate of the Gross Domestic Product (GDP) in Pakistan is 5.8% [18] whereas the per capita income is \$1641 [18]. If the growth rate of the population continues to increase at a pace of 2.4% then

by 2050 the country will become the 4th largest in the world [18]. The power consumption per capita is around 500kWh which is quite low compared with the global per-capita of power consumption which is around 2603kWh [19].

B. Composition of Biomass and Solid Waste

Table II shows the chemical composition criteria of biomass resources through ultimate and proximate analysis on the physicochemical characteristics [20]. These characteristics would help to identify the selectivity and suitability of biomass resources for power generation. The chemical composition of solid waste is quite different from the biomass in the sense that the quality and capacity of solid waste is affected by various factors including the living standards, weather conditions, surrounding region, and financial status. Solid waste normally comprises of organic waste, inorganic waste, hazardous waste, paper waste, plastic waste, and textile waste. The waste composition type and percentage is shown in Figure 1 [21]. The physicochemical characteristics of solid waste are obtained from ultimate and proximate analysis tests. The ultimate analysis test is used to find the proportion of oxygen, nitrogen, sulphur, carbon, and hydrogen in the total solid waste whereas the proximate analysis test is used to find the fixed carbon, ash, moistness, and volatile matter. The physicochemical characteristics of solid waste are shown in Figure 2 [21].

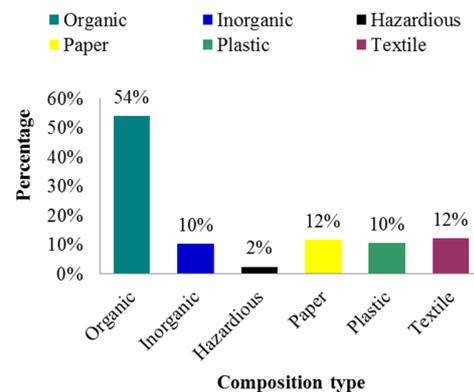


Fig. 1. Composition type of solid waste in Pakistan.

C. Experimental Study of Biomass and Solid Waste Samples

An experimental setup is designed based on the quartering method as shown in Figure 3. Initially, 25kg of biomass and 25kg of solid waste were collected from different regions of Pakistan. Then, manual mixing and cutting were performed and all the waste was gathered at a single place. Further, this total waste was divided into 8 sections namely I, II, III, IV, V, VI, VII, and VIII. These sections were separated into even (II, IV, VI, and VIII) and odd sections (I, III, V, and VII). Even sections were mixed and again separated into two sections namely M and N. Similarly, odd section were mixed and separated into two sections namely O and P. Then these four sections (M, N, O and P) were mixed diagonally like M & P and N & O and separated into two sections like M & P into Y and N & O into Z. Finally, Y and Z were mixed for the final sample of analysis. The manual mixing and cutting were performed many times until the weight became 15kg.

TABLE II. BIOMASS CHEMICAL COMPOSITION [22, 23]

	Sulphur	Carbon	Oxygen	Hydrogen	Fixed carbon	Ash	Volatile matter	Moisture
Rice husk	0.02	35.6	59.7	4.5	14.6	26.2	59.2	8.8
Straw	0.21	39	35.46	5.73	17.5	10.1	64.43	8.32
Bamboo	0.09	50.52	42.80	6.04	16.05	1.95	83.95	6.14
Wheat straw	-	47.9	45.2	6.1	16.5	6.9	76.5	0.1
Sugarcane bagasse	0.09	46.6	43.3	5.92	8.30	3.90	82.5	5.25
Pine sawdust	0.36	49.79	44.12	6.36	13.91	1.29	82.03	2.77
Oat	-	42.3	40.9	6.3	-	1.5	-	7.1
Kenaf	0.05	46.71	54.32	6.71	17.18	5.45	73.64	3.73

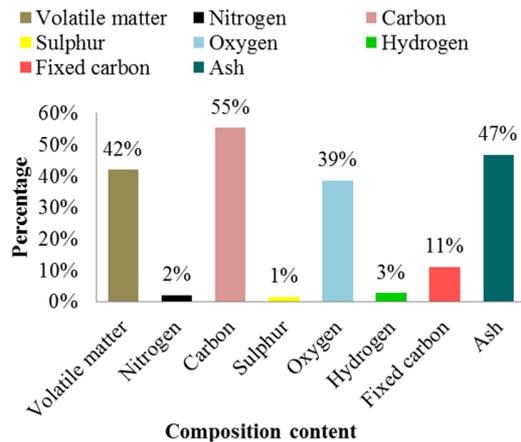


Fig. 2. Composition content of solid waste in Pakistan.

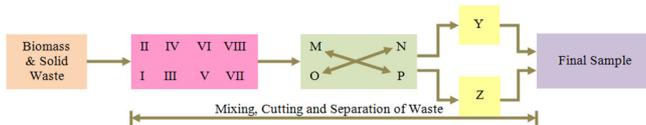


Fig. 3. Quartering diagram method for waste sampling.

D. Determination of Heating or Calorific Value

The calorific value is highly dependent on the amount of heat produced during the combustion process. The calorific value of the final sample of waste pellets was determined in the laboratory using the Gallen Kamp Ballistic Bomb (GKBB) Calorimeter. The net calorific value is determined as 6519kcal/kg. However, from other research data, the calorific value ranges from 9MJ/kg to 44MJ/kg respectively [24]. Equations (1) and (2) were used for finding the higher and lower values of calorific of mixed biomass and solid waste pallets [25].

$$\text{Higher calorific value} = \frac{\sum Q_p (C.V)_H}{T_p} \quad (1)$$

$$\text{Lower calorific value} = \frac{\sum Q_p (C.V)_L}{T_p} \quad (2)$$

where (C.V)L is the lower calorific value in kcal/kg, (C.V)H is the higher calorific value in kcal/kg, Qp is the quantity of the specific material in the total waste pallets in kg, and Tp is the total waste pallets in kg.

E. Theoretical Power Potential of Biomass and Solid Waste

The power potential of mixed biomass and solid waste pallets can be calculated by:

$$E_p = (C.V)L \times A_w \times 1.16 \quad (3)$$

where Ep is the energy potential in kWh and Aw is the aggregate waste in kg.

III. RESULTS AND DISCUSSION

A. Net Power Calculations

Net power of biomass and solid waste pallets were calculated as follows:

The Heat Of Combustion (HOC) was calculated using the total Quantity of Waste Pallets (QWP) and Calorific Value (CV):

$$HOC = QWP \times CV \quad (4)$$

$$HOC = 55,422,000,000 \text{ ton} \times 6,519,000 \text{ kcal/ton}$$

$$HOC = 3.61296018 \times 10^{17} \text{ kcal}$$

The Heat Output (HO), considering efficiency of 25% [25], was:

$$HO = \text{overall efficiency} \times HOC \quad (5)$$

$$HO = 0.25 \times 3.61296018 \times 10^{17}$$

$$HO = 9.03240045 \times 10^{16} \text{ kcal}$$

With 1kWh equal to 860kcal the Units Generated/Annum (UGA) are:

$$UGA = HO/860 \quad (6)$$

$$UGA = 9.03240045 \times 10^{16}/860 \text{ kcal}$$

$$UGA = 105,027,912,209,302.3 \text{ kWh}$$

$$UGA = 105,027,912,209.3023 \text{ MWh}$$

The Average Load on the System (ALS) is:

$$ALS = UGA/\text{total hours in a year} \quad (7)$$

$$ALS = 105,027,912,209.3023 \text{ MWh} / 8760 \text{ h}$$

$$ALS = 11,989,487.69512584 \text{ MW}$$

This mathematical framework calculated the total power production capacity from the hybrid energy resources based on authentic and realistic data. The available capacity and quality of waste pallets can generate a power of 11,989.5GW. This power capacity can drive the economy of the country at a great extent but there is a need to develop a proper energy harnessing system with sufficient financing and resources.

TABLE III. TECHNOLOGY SELECTION PARAMETERS FOR POWER GENERATION

Parameters	Pyrolysis	Plasma	Gasification	Incineration
Feedstock	Solid waste	Solid waste	Biomass and solid waste	Biomass and solid waste
Variable composition	Did not accept variable waste	Did not accept variable waste	Accept heterogeneous waste of all types	Little difficult to proceed variable waste
Product	Oil, Syngas and energy	Syngas and energy	Syngas and energy	Heat and energy
Operating cost	150\$ per tone	120\$ per tone	60\$ per tone	60\$ per tone
Annual capital cost	1500\$ per tone	1300\$ per tone	850\$ per tone	775\$ per tone
Disposal	Ash 0.3 tone per feedstock tone	Bottom ash around 10%	Less than 1% bottom ash	Around 5% bottom ash
Energy production capability	800 kWh per tone	600kWh per tone	800kWh per tone	850 Kwh per tone
Efficiency	Less than 18%	Less than 10%	18% to 30%	18% to 25%

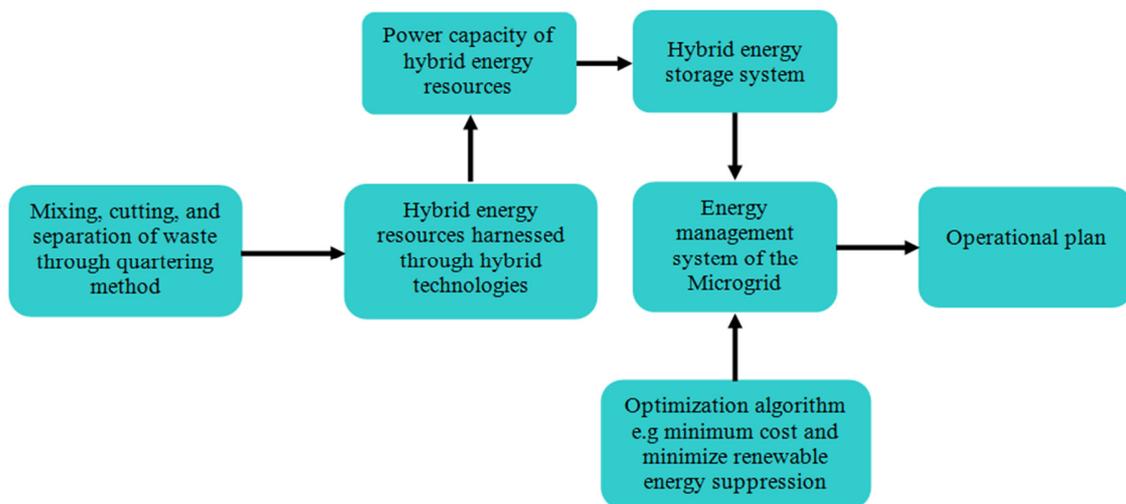


Fig. 4. Hybrid energy storage systems for renewable source integration in the microgrid energy management system.

B. Power Generation Technology for Hybrid Resources

The gasification technique for power generation is more suitable than the other techniques mentioned in Table III [26-29]. The gasification technique accepts all types of waste (biomass and solid waste) for power generation with greater power production efficiency and has the capability of producing less ash. Gasification with its hybrid system opens the door for new development in the country because hybrid technologies based on other resources, including wind, solar, and nuclear, coupled with biomass and solid waste resources can provide more energy benefits to the community.

C. Integration of Hybrid Biomass-Solid Waste Energy Resources in a Microgrid

Most renewable sources are used in microgrids with lower power levels (around 200kW) and are connected with the main bus through power converters [28-31]. These systems are installed near the commercial and industrial sites for meeting the power demand. These systems produce low noise and emissions which ultimately provide reliable power on low cost. The selection of a suitable renewable source is a complex process. This research paper has described a step-wise feasibility study of hybrid biomass and solid waste resource for microgrid development. Figure 4 shows the hybrid energy

storage systems for hybrid renewable source (biomass/solid waste) integration in the energy management system of a microgrid. The integration of a hybrid energy storage system, distributed energy resources, and distributed loads with a renewable energy ecosystem is called microgrid. A microgrid helps to setup the smart and active electrical grid with the potential to increase the efficiency, reliability, and safety of the system.

IV. CONCLUSION

This study aims to alleviate the current energy crisis and suggests future electricity pathways to drive the economy of a developing country such as Pakistan. An experiment was performed in the lab in order to identify the suitability for power generation of waste pallets collected from different sites of Pakistan. The results of this study were examined quantitatively to justify the feasibility of the energy plan. The projected power generation is 11,989.5GW which is sufficient to alleviate the ongoing energy crisis and facilitate the future industrial development in the country. Pakistan therefore needs to develop a new policy that is economically feasible and environmentally friendly for longer tenure. Hybrid energy resources provide a suitable option through hybrid technologies for microgrid development.

REFERENCES

- [1] S. Hansen, A. Mirkouei, and L. A. Diaz, "A comprehensive state-of-technology review for upgrading bio-oil to renewable or blended hydrocarbon fuels," *Renewable and Sustainable Energy Reviews*, vol. 118, Feb. 2020, Art. no. 109548, <https://doi.org/10.1016/j.rser.2019.109548>.
- [2] F. Campuzano, R. C. Brown, and J. D. Martinez, "Auger reactors for pyrolysis of biomass and wastes," *Renewable and Sustainable Energy Reviews*, vol. 102, pp. 372–409, Mar. 2019, <https://doi.org/10.1016/j.rser.2018.12.014>.
- [3] O. K. M. Ouda, S. A. Raza, A. S. Nizami, M. Rehan, R. Al-Waked, and N. E. Korres, "Waste to energy potential: A case study of Saudi Arabia," *Renewable and Sustainable Energy Reviews*, vol. 61, pp. 328–340, Aug. 2016, <https://doi.org/10.1016/j.rser.2016.04.005>.
- [4] W. Rulkens, "Sewage Sludge as a Biomass Resource for the Production of Energy: Overview and Assessment of the Various Options," *Energy & Fuels*, vol. 22, no. 1, pp. 9–15, Jan. 2008, <https://doi.org/10.1021/ef700267m>.
- [5] A. R. K. Gollakota, N. Kishore, and S. Gu, "A review on hydrothermal liquefaction of biomass," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1378–1392, Jan. 2018, <https://doi.org/10.1016/j.rser.2017.05.178>.
- [6] F. Manzano-Agugliaro, A. Alcayde, F. G. Montoya, A. Zapata-Sierra, and C. Gil, "Scientific production of renewable energies worldwide: An overview," *Renewable and Sustainable Energy Reviews*, vol. 18, pp. 134–143, Feb. 2013, <https://doi.org/10.1016/j.rser.2012.10.020>.
- [7] K. L. Khatri, A. R. Muhammad, S. A. Soomro, N. A. Tunio, and M. M. Ali, "Investigation of possible solid waste power potential for distributed generation development to overcome the power crises of Karachi city," *Renewable and Sustainable Energy Reviews*, vol. 143, Jun. 2021, Art. no. 110882, <https://doi.org/10.1016/j.rser.2021.110882>.
- [8] G. W. Huber, S. Iborra, and A. Corma, "Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering," *Chemical Reviews*, vol. 106, no. 9, pp. 4044–4098, Sep. 2006, <https://doi.org/10.1021/cr068360d>.
- [9] J. Ahrenfeldt, T. P. Thomsen, U. Henriksen, and L. R. Clausen, "Biomass gasification cogeneration – A review of state of the art technology and near future perspectives," *Applied Thermal Engineering*, vol. 50, no. 2, pp. 1407–1417, Feb. 2013, <https://doi.org/10.1016/j.applthermaleng.2011.12.040>.
- [10] G. Kumar *et al.*, "A review of thermochemical conversion of microalgal biomass for biofuels: chemistry and processes," *Green Chemistry*, vol. 19, no. 1, pp. 44–67, 2017, <https://doi.org/10.1039/C6GC01937D>.
- [11] M. Rehfeldt, E. Worrell, W. Eichhammer, and T. Fleiter, "A review of the emission reduction potential of fuel switch towards biomass and electricity in European basic materials industry until 2030," *Renewable and Sustainable Energy Reviews*, vol. 120, p. 109672, Mar. 2020, <https://doi.org/10.1016/j.rser.2019.109672>.
- [12] M. Saghir, S. Zafar, A. Tahir, M. Ouadi, B. Siddique, and A. Hornung, "Unlocking the Potential of Biomass Energy in Pakistan," *Frontiers in Energy Research*, vol. 7, 2019, Art. no. 24, <https://doi.org/10.3389/fenrg.2019.00024>.
- [13] K. Kapadia and A. Agrawal, "Municipal Solid Waste to Energy Options - A Review," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3361532, Mar. 2019. <https://doi.org/10.2139/ssrn.3361532>.
- [14] P. J. Verkerk *et al.*, "Spatial distribution of the potential forest biomass availability in Europe," *Forest Ecosystems*, vol. 6, no. 1, Feb. 2019, Art. no. 5, <https://doi.org/10.1186/s40663-019-0163-5>.
- [15] Z. Qin *et al.*, "Biomass and biofuels in China: Toward bioenergy resource potentials and their impacts on the environment," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 2387–2400, Feb. 2018, <https://doi.org/10.1016/j.rser.2017.08.073>.
- [16] V. Ferreira-Leitao, L. M. F. Gottschalk, M. A. Ferrara, A. L. Nepomuceno, H. B. C. Molinari, and E. P. S. Bon, "Biomass Residues in Brazil: Availability and Potential Uses," *Waste and Biomass Valorization*, vol. 1, no. 1, pp. 65–76, Mar. 2010, <https://doi.org/10.1007/s12649-010-9008-8>.
- [17] M. A. Wazir and A. Goujon, *Assessing the 2017 Census of Pakistan Using Demographic Analysis: A Sub-National Perspective*. Vienna, Austria: Vienna Institute of Demography, 2019.
- [18] "GDP growth (annual %)," The World Bank. <https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG> (accessed May 26, 2021).
- [19] *SAARC Energy Outlook 2030*. Mumbai, India: CRISIL, 2018.
- [20] A. A. Yusuf and F. L. Inambao, "Effect of low bioethanol fraction on emissions, performance, and combustion behavior in a modernized electronic fuel injection engine," *Biomass Conversion and Biorefinery*, vol. 11, no. 3, pp. 885–893, Jun. 2021, <https://doi.org/10.1007/s13399-019-00519-w>.
- [21] A. Shehzad, M. J. K. Bashir, and S. Sethupathi, "System analysis for synthesis gas (syngas) production in Pakistan from municipal solid waste gasification using a circulating fluidized bed gasifier," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1302–1311, Jul. 2016, <https://doi.org/10.1016/j.rser.2016.03.042>.
- [22] Y. Yu *et al.*, "Pyrolysis of Rice Husk and Corn Stalk in Auger Reactor. 1. Characterization of Char and Gas at Various Temperatures," *Energy & Fuels*, vol. 30, no. 12, pp. 10568–10574, Dec. 2016, <https://doi.org/10.1021/acs.energyfuels.6b02276>.
- [23] T. M. Ismail, A. Ramos, E. Monteiro, M. A. El-Salam, and A. Rouboa, "Parametric studies in the gasification agent and fluidization velocity during oxygen-enriched gasification of biomass in a pilot-scale fluidized bed: Experimental and numerical assessment," *Renewable Energy*, vol. 147, pp. 2429–2439, Mar. 2020, <https://doi.org/10.1016/j.renene.2019.10.029>.
- [24] C. Sheng and J. L. T. Azevedo, "Estimating the higher heating value of biomass fuels from basic analysis data," *Biomass and Bioenergy*, vol. 28, no. 5, pp. 499–507, May 2005, <https://doi.org/10.1016/j.biombioe.2004.11.008>.
- [25] *(Draft) Guideline for Solid Waste Management*. Islamabad, Pakistan: Pakistan Environmental Protection Agency, 2005.
- [26] A. Gomez-Barea, P. Ollero, and B. Leckner, "Optimization of char and tar conversion in fluidized bed biomass gasifiers," *Fuel*, vol. 103, pp. 42–52, Jan. 2013, <https://doi.org/10.1016/j.fuel.2011.04.042>.
- [27] S. H. Solangi, A. Q. Jakhriani, K. C. Mukwana, A. R. Jatoi, and M. R. Luhur, "Investigation of Quantity, Quality and Energy Content of Indigenous Sugarcane Trash in Naoshehro Feroze District, Sindh," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3609–3613, Dec. 2018, <https://doi.org/10.48084/etasr.2411>.
- [28] K. Sookramoon, "80 kW Updraft Gasifier Performance Test using Biomass Residue Waste from Thailand Rural Areas," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6349–6355, Oct. 2020, <https://doi.org/10.48084/etasr.3820>.
- [29] N. M. Kumar, S. S. Chopra, A. A. Chand, R. M. Elavarasan, and G. M. Shafiqullah, "Hybrid Renewable Energy Microgrid for a Residential Community: A Techno-Economic and Environmental Perspective in the Context of the SDG7," *Sustainability*, vol. 12, no. 10, Jan. 2020, Art. no. 3944, <https://doi.org/10.3390/su12103944>.
- [30] F. Aghae, N. Mahdian Dehkordi, N. Bayati, and A. Hajizadeh, "Distributed Control Methods and Impact of Communication Failure in AC Microgrids: A Comparative Review," *Electronics*, vol. 8, no. 11, Nov. 2019, Art. no. 1265, <https://doi.org/10.3390/electronics8111265>.
- [31] Bayati, N. Bayati, H. R. Baghaee, A. Hajizadeh, and M. Soltani, "Localized Protection of Radial DC Microgrids With High Penetration of Constant Power Loads," *IEEE Systems Journal*, pp. 1–12, 2020, <https://doi.org/10.1109/JSYST.2020.2998059>.

Utilization of Machine Learning in Supporting Occupational Safety and Health Decisions in Hospital Workplace

Kyriakos Koklonis

Biomedical Engineering Laboratory
National Technical University of Athens
Athens, Greece
kkoklonis@biomed.ntua.gr

Maria Vastardi

Metaxa Cancer Hospital of Piraeus
Piraeus, Greece
mariavastardi@yahoo.gr

Michail Sarafidis

Biomedical Engineering Laboratory
National Technical University of Athens
Athens, Greece
msarafidis@biomed.ntua.gr

Dimitrios Koutsouris

Biomedical Engineering Laboratory
National Technical University of Athens
Athens, Greece
dkoutsou@biomed.ntua.gr

Abstract-The prediction of possible future incidents or accidents and the efficiency assessment of the Occupational Safety and Health (OSH) interventions are essential for the effective protection of healthcare workers, as the occupational risks in their workplace are multiple and diverse. Machine learning algorithms have been utilized for classifying post-incident and post-accident data into the following 5 classes of events: Needlestick/Cut, Falling, Incident, Accident, and Safety. 476 event reports from Metaxa Cancer Hospital (Greece), during 2014-2019, were used to train the machine learning models. The developed models showed high predictive performance, with area under the curve range 0.950-0.990 and average accuracy of 93% on the 10-fold cross set, compared to the safety engineer's study reports. The proposed DSS model can contribute to the prediction of incidents or accidents and efficiency evaluation of OSH interventions.

Keywords-occupational health and safety; OSH; machine learning; hospital workplace

I. INTRODUCTION

Occupational accidents and diseases result in an additional economic burden on public social security agencies. These losses are estimated at 3.9% of the global Gross Domestic Product (GDP) and 3.3% of the European GDP, as it has been reported from the European Agency for Safety and Health at Work (EU-OSHA) (with variance according to working fields, legislative context and prevention incentives) [1]. According to the EU-OSHA, proper management of Occupational Safety and Health (OSH) leads to numerous benefits, such as reduced absenteeism, lessened costs and improved efficiency of working methods and technologies [2]. This necessity becomes even more imperative during periods of economic recession, since poor OSH outlays valuable resources. Particularly,

regarding the healthcare sector, the hospital workplace is characterized by increased level and diversity of occupational risks. In addition, the coronavirus disease 2019 (COVID-19) pandemic and its immediate aftermath pose a significant burden of workload and a major strain on mental health of healthcare workers [3].

Health professionals constitute a significant part of the total European and Greek workforce. According to a relevant study of the Greek Labor Inspectorate Body in 2015, 10% of workers in the European Union (EU) belong to the healthcare sector [4]. More specifically, according to the Hellenic Statistical Authority, the number of Greek hospitals amounted up to 283 during the 2015-2018 period. Moreover, their staff, during 2017, consisted of 38,952 nursing, 23,354 medical, 6,044 auxiliary nursing, and 7,752 paramedical staff [5]. In Greek healthcare sector, a high percentage of work accidents (34% higher than the corresponding average in EU, according to the EU-OSHA) and an increased incidence of musculoskeletal disorders (second largest after construction sector) are recorded. According to the data from the Hellenic Statistical Authority, in 2018, 3.2% of all work accidents that have been officially recorded are related to the health and social care sector [6]. Concurrently, there is a dramatic public health funding reduction of up to 60% for the last three years [4].

Computational intelligence and Machine Learning (ML) approaches can be utilized towards an improved management of OSH. Their potential in occupational accident analysis has been highlighted from several studies, as they have been reviewed in [7]. Indicatively, statistical methods have been used for analyzing health and safety issues of women in industry [8], the OSH risk in fuel stations has been assessed through Failure Mode and Effects Analysis (FMEA) method

[9], Naive Bayesian (NB) model has been used for coding causation of workers' compensation claims [10], Bayesian Networks (BNs) have been used for analyzing data on occupational accidents [11–13], and decision trees have been used in the industrial mining sector for predicting the type of accident [14]. Artificial Neural Networks (ANNs) have been used to correlate causes and OSH conditions [15–19] and they have been combined with Support Vector Machines (SVMs) for the prediction of occupational accidents [20]. In addition, Analytical Hierarchical Process (AHP) has been used in order to analyze, control, and provide the occupational risk in the industrial mining sector [21], and fuzzy logic has been used for occupational risk assessment [22]. The self-organizing map and k-means clustering (SOM k-means) methods have been used for identifying the dynamics of critical accidents [23]. In [24], body-mounted sensors were combined with ML techniques to evaluate the ergonomic risk levels caused by overexertion. The performance of ML algorithms in classifying post-incident outcomes of occupational injuries in agro-manufacturing operations has been evaluated in [25]. The minimization of the undesirable adverse effects in the development and implementation of ML-based Decision Support Systems (DSSs) was the objective in [26].

Particularly in the healthcare sector, ML has been utilized for targeting specific hazards. For instance, ANNs have been used to approach the burnout process [27] and statistical analysis has been employed to identify factors of musculoskeletal disorders in nursing staff [28] and to improve the management of indoor air quality in the hospital workplace [29]. Correlation methods have been used to assess the Quality of Work-Life (QWL) for the nursing profession in China [30]. In a more generalized approach, a tool for estimating OSH conditions through factor analysis has been evaluated in [31]. Factor analysis has also been used for the assessment of workplace violence towards healthcare staff in public hospitals in Turkey [32]. In addition, a fuzzy logic model for OSH risk assessment has been proposed in [33] and a fuzzy cognitive map approach for the impact assessment of processes on the OSH management system's effectiveness has been used in [34]. Ambulatory monitoring devices, such as wearable and environmental sensors, have been combined with ML algorithms for a better understanding of stress evolution in healthcare workers [35].

Based on the previous concise literature review, there is a promising research activity in capitalizing ML methods for better implementing effective preventive measures and policies in the OSH framework. Most of the published studies are based on questionnaires to OSH experts [12, 13, 36], instead of supporting immediate incident or accident investigations and following up the efficiency of the proposed interventions. To the best of our knowledge, a DSS which utilizes the legally established safety engineer investigations and their audit reports has not been reported in the literature. The present study proposes a DSS model which integrates all incident or accident data along with the safety engineer investigations and audit reports in the hospital workplace. Through the input data utilization, our DSS aims at supporting OSH decisions by predicting future possible incidents or accidents and by assessing the efficiency of OSH interventions.

II. MATERIALS AND METHODS

A. Data

In order to train and test the ML models, 476 event reports from Metaxa Cancer Hospital (reference period from January 1, 2014 to December 31, 2019) were collected and utilized. The reports were categorized into 5 classes depending on the occurred events: 136 reports of "Needlestick /Cut" injuries, 59 reports of various "Incidents", 20 reports of "Falling" injuries, 23 reports of "Accidents" and 238 reports of "Safety" conditions. One event report consists of multiple variables derived from various reports, as detailed below. Reports including undone corrective actions or incomplete data, during the reference period, were not included in the dataset. The worker population of the 476 records included nursing staff (52.52%), medical staff (13.45%), cleaning staff (10.50%), laboratory staff (9.24%), administrative staff (5.46%), technical staff and workers (4.62%), and auxiliary staff (4.20%). The mean age of the worker population of the records was 46.61 years and the mean experience was 16.98 years. Our study has been approved by the Scientific and Administrative Council of the Metaxa Hospital. Under the Regulation (EU) 2016/679 of the European Parliament (General Data Protection Regulation – GDPR) on the protection of natural persons [37], pseudonymization was used during the coding and data entry process.

B. Software

Waikato Environment for Knowledge Analysis version 3.8.3 has been used for the data analysis and predictive modeling of the system [38]. This software has been chosen because of its wide applications in data mining and ML. Moreover, it is freeware and publicly available under the General Public License, it includes several data mining and ML methods, libraries, and relative supporting tools–procedures, it is compatible with other hardware or software platforms, and it includes a graphical user interface, making its usage easier even without programming knowledge.

C. Risk Assessment Method

A very common risk assessment method is the one which is consistent with the Memorandum on Occupational Risk Assessment by Directorate-General for Employment in Labor Relations and Social Affairs (DG V) of the EU (1997) [39]. This methodology has also been recommended by the Technical Chamber of Greece in a relevant conference [40]. In addition, it is widely used by safety engineers and OSH auditors. This risk assessment method was used in the proposed model. Its approach is based on defining a risk value as a product of three factors: the seriousness of the consequences of the potential hazard (values 1, 4, 8, 16), the frequency of exposure to the hazard (values 1, 2, 3, 4), and the probability of occurrence of this hazard (values 1, 2, 3, 4), according to the principle of reasonable ambiguity. Therefore, the risk value (R) lies within the range of 1 to 256. The usage of these three factors is in accordance with relevant legal requirements according to Greek Law (Law 3850/2010, Article 43). These factors are presented in Table I. The risk values and their correlation with the necessity of taking actions are presented in Table II.

TABLE I. RISK – RATING FACTORS

Factor	Rating	Interpretation
Seriousness	1	Negligible
	4	Middle
	8	Critical
	16	Disastrous
Frequency	1	Zero
	2	Limited
	3	Often
	4	Continuous
Probability	1	Zero
	2	Low
	3	Middle
	4	High

TABLE II. RISK RATING

Risk value R	Risk description	Measures - actions
R<16	Negligible: Risk is very low and may not increase during the near future without changing working conditions	Measures – corrective actions are not necessary
16<R<32	Low: Risk is under control, but the incident is possible	Need for monitoring and actions for risk reduction. Measures – corrective actions are not an urgent need
32<R<64	Middle: Risk is not effectively controlled. An incident cannot be excluded	Need for programming measures – corrective actions for risk reduction
64<R<128	High: Risk is not effectively controlled. A serious incident is possible	Need for programming measures – corrective actions for risk elimination and urgent measures – corrective actions for risk reduction
R>128	Critical: Possibility of human loss. A serious incident is about to happen	Need for urgent measures – corrective actions for risk elimination

D. DSS Model Information Flow

The OSH related events are categorized into two main categories: incidents and accidents. The accidents, according to Greek Law, are all the events that result in damage, injury or harm and require, at least a two-day, medical leave and relevant notification to the Labor Inspectorate Body. The incidents (near-misses) are less serious events that have unintentionally happened and are linked to absence from work only on the specific day of their occurrence. The procedure that is followed after an occupational incident or accident in the hospital workplace is described below: the supervisor of the relevant division and the responsible desk officer are alerted. The update of the associated data files follows, according to the internal process of the hospital and the relevant legal requirements (Law 3850/2010, Article 2, Paragraph 2). The issue is registered in the book of incidents, the book of accidents, the incident – accident report, the record of exposure to biological agents at work, the report to the Social Insurance Institute, the police and the electronic notification to the Labor Inspectorate Body. Then, the safety engineer in charge investigates the incident or accident in order to identify its

causes and propose interventions, by completing risk assessment tables manually in printed versions (Law 3850/2010, Articles 14, 15). The safety engineer's actions aim at ensuring that repetition of the incident or the accident could be avoided in the future. The manually filled tables are in accordance with the Methodological guide for the assessment and prevention of occupational risks of the Hellenic Institute for Occupational Health and Safety [41] regarding the coding of hazards, and with the method described in the previous section [39] regarding the risk values. The safety engineer cooperates with the occupational doctor, when necessary.

The above-mentioned data for every incident and accident are used as an input to the proposed DSS. For the efficient management of this information, the basic idea was to organize and incorporate the data into a database, which makes the information more usable and manageable in comparison with the printed forms. The information flow diagram of the proposed DSS is depicted in Figure 1. There are three major blocks in this diagram. Firstly, the block of recording and notifying incidents – accidents (upper left) which describes the process that the responsible personnel follow to record every incident and accident in the corresponding books, to fill in the incident – accident report and notify the authorities. Secondly, the block of risk assessment and investigation (lower left) which describes the process that the safety engineer and occupational doctor follow to identify OSH hazards, to express the corresponding risk values and to associate these hazards with every incident and accident through investigation and estimation of OSH interventions. These two blocks are the inputs to the proposed DSS model. The output is the third block which describes the results of the classification process, according to the trained models. The decision is based on the risk values of the different variables. The input of the model includes two sets of variables. The first set consists of 6 variables (F1 – F6) from the first block which are related to the employees. Their values come up from the incident or accident relative paper report. The second set consists of 77 variables (F7 – F83) from the second block. All 83 variables are summarized in Table III. The selection of these variables was based on [41]. The main categories of this set are workplace related variables relevant to the infrastructural building, machinery - equipment, electrical installation, hazardous substances, fire, chemical - physical - biological factors, and transversal - organizational risks.

Each occupational event corresponds to one instance and was categorized into one of the following 4 classes: "Needlestick/Cut", "Falling", "Incident" and "Accident" according to their type. The class of "Incident" was divided into three different subclasses because of the frequency distribution of incidents in the hospital workplace. Needlesticks/cuts and falling events present the highest frequency, which is in accordance with [4]. For this reason, two separate classes were used for this kind of incidents, "Needlestick/Cut" and "Falling", respectively. The class named "Incident" was used to include all other kinds of incidents in the hospital workplace. The class "Accident" was used to include all kinds of accidents. The values for these variables express the OSH conditions and hazards that are related to the incident or accident, as they have been identified by the safety engineer. These values were

recorded according to the safety engineer’s reports and risk assessment tables after investigation of the corresponding incident or accident. Finally, after the completion of the corrective actions, all the above instances were assigned with updated risk values and were categorized into the class "Safety". Each new instance of "Safety" originated from one instance of "Needlestick/Cut" or "Falling" or "Incident" or "Accident". In this way, our dataset consists of 238 incident/accident instances and 238 safety instances (as described in section "Data"). More specifically, the values for the instances of "Safety" express the proposed measures or interventions to reduce risk at acceptable levels ($R \leq 32$ for accidents and $R \leq 24$ for incidents, according to the relative methodology). These values were recorded according to the post-investigation safety engineer's updated reports and risk assessment tables after up to two months for every incident or accident. During this period, the corrective actions had been completed and confirmed according to the safety engineer’s proposed measures. The output of the system is one of the above-mentioned 5 classes (Y1 – Y5) and expresses the type of the event that certain occupational conditions could lead to (Table III). As the predicted classes were a priori known, our model appertains to supervising learning algorithms.

TABLE III. VARIABLES OF THE PROPOSED MODEL

	Variables	Code	Type	Values	
Employee variables	Age	F1	Numeric	Years	
	Experience	F2	Numeric	Years	
	Training	F3	Nominal	Yes, No	
	Repetitive training every six months at least	F4	Nominal	Yes, No	
	Relative PPE during the incident/accident	F5	Nominal	Yes, No	
	Satisfaction for OSH climate	F6	Numeric	1-5	
Workplace variables relatively to infrastructural building	Floors/stairs	F7	Numeric	1-256	
	Workplace space	F8	Numeric	1-256	
	Workplace height	F9	Numeric	1-256	
	Workplace volume	F10	Numeric	1-256	
	Doors/windows	F11	Numeric	1-256	
	Space illumination	F12	Numeric	1-256	
	Storage/lofts	F13	Numeric	1-256	
	Uncovered gaps	F14	Numeric	1-256	
	Obstacles	F15	Numeric	1-256	
	Exits/escape routes	F16	Numeric	1-256	
	Walls	F17	Numeric	1-256	
	Shelves/dexions	F18	Numeric	1-256	
	Ceilings	F19	Numeric	1-256	
	Basements	F20	Numeric	1-256	
	Walkways	F21	Numeric	1-256	
	Roof	F22	Numeric	1-256	
	Cleanliness	F23	Numeric	1-256	
	Lack of access to exits or fire-extinguishing system	F24	Numeric	1-256	
	Safety/escape signing	F25	Numeric	1-256	
	Workplace variables relatively to machinery – equipment	Maintenance protocol	F26	Numeric	1-256
		Lack of safety during usage	F27	Numeric	1-256
		Guards for avoiding the accidental start	F28	Numeric	1-256
		Guards from moving parts	F29	Numeric	1-256
Ejectable particles		F30	Numeric	1-256	
CE signing		F31	Numeric	1-256	
Cutting works		F32	Numeric	1-256	
Lifting machinery		F33	Numeric	1-256	
Transport vehicles		F34	Numeric	1-256	

	Ladders	F35	Numeric	1-256	
	Pneumatic tools	F36	Numeric	1-256	
	Elevators	F37	Numeric	1-256	
	Other machinery	F38	Numeric	1-256	
	Non-usage of PPE	F39	Numeric	1-256	
	Pressure devices	F40	Numeric	1-256	
	Access to facilities or equipment	F41	Numeric	1-256	
	Workplace variables relatively to electrical installation	Electrical installation	F42	Numeric	1-256
		Inappropriate usage of electrical installation	F43	Numeric	1-256
		Inappropriateness in explosive atmospheres	F44	Numeric	1-256
		Lack of safety during usage of electrical installation	F45	Numeric	1-256
		Lack of safety during maintenance of the electrical installation	F46	Numeric	1-256
Hazardous substances related to equipment (generators, batteries, etc.)		F47	Numeric	1-256	
Workplace variables relatively to hazardous substances	Toxic substances	F48	Numeric	1-256	
	Caustic substances	F49	Numeric	1-256	
	Corrosive substances	F50	Numeric	1-256	
	Irritant substances	F51	Numeric	1-256	
	Oxidizing substances	F52	Numeric	1-256	
	Explosive substances	F53	Numeric	1-256	
	Flammable material	F54	Numeric	1-256	
	Appropriate cabinets	F55	Numeric	1-256	
Workplace variables relatively to fire	Fire signing	F56	Numeric	1-256	
	Banning smoking	F57	Numeric	1-256	
	Banning flame	F58	Numeric	1-256	
	Storage of flammable material	F59	Numeric	1-256	
	Lack of fire protection systems	F60	Numeric	1-256	
	Training in fire emergency plan	F61	Numeric	1-256	
	Fire rxtinguishers	F62	Numeric	1-256	
	Lack of training in fire security	F63	Numeric	1-256	
Workplace variables relatively to chemical factors	Dust	F64	Numeric	1-256	
	Fibres asbestos	F65	Numeric	1-256	
	Smoke/steam	F66	Numeric	1-256	
	Particles	F67	Numeric	1-256	
	Other substances	F68	Numeric	1-256	
	Dipping/splashing	F69	Numeric	1-256	
Workplace variables relatively to physical factors	Noise	F70	Numeric	1-256	
	Vibrations	F71	Numeric	1-256	
	Radiation	F72	Numeric	1-256	
	Illumination	F73	Numeric	1-256	
	Microclimate	F74	Numeric	1-256	
	Workplace variables relatively to biological factors	Bacteria	F75	Numeric	1-256
Fungi		F76	Numeric	1-256	
Viruses		F77	Numeric	1-256	
Other factors		F78	Numeric	1-256	
Workplace variables relatively to transversal risks	Settlement - manual handling	F79	Numeric	1-256	
	Psychological factors	F80	Numeric	1-256	
	Provision - intervention programs	F81	Numeric	1-256	
	Ergonomy	F82	Numeric	1-256	
	Hardship conditions	F83	Numeric	1-256	
OUTPUT	Falling	Y1	Nominal	Yes, No	
	Needlestick/cut	Y2	Nominal	Yes, No	
	Incident	Y3	Nominal	Yes, No	
	Accident	Y4	Nominal	Yes, No	
	Safety	Y5	Nominal	Yes, No	

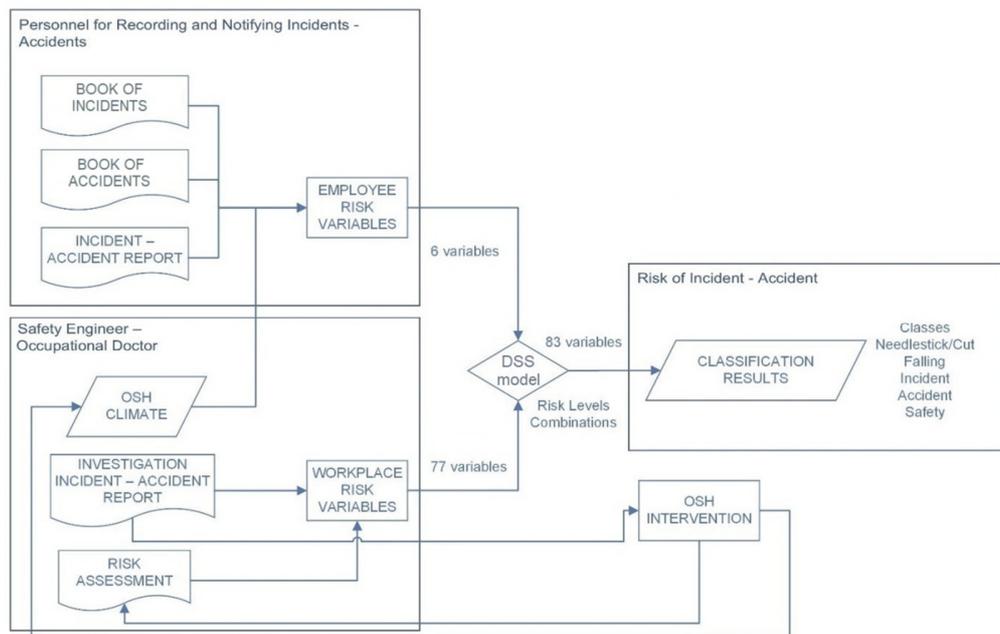


Fig. 1. The information flow diagram of the proposed model.

E. Prediction Models and Classification

ML is concerned with the development of intelligent systems that can learn from data. These systems can be used to classify an object into a correct class based on features characterizing the object. Actually, the algorithms are searching for patterns in data. The ML algorithms that implement pattern classification are known as classifiers. There are several classifiers, each one of them having certain limitations and advantages [42]. In this study, in order to develop the proposed DSS model, we employed and tested 4 classifiers: The Naive Bayesian (NB) classifier [43], the Bayesian Network (BN), the k-Nearest Neighbors (k-NN) classifier, and the Multilayer Perceptron (MLP) network. The choice was based on the fact that the use of conditional probabilities offers a practical metric as it can model occupational health in a class conditional way. This formulation helps us elucidate class conditional variables of interest.

NB classification is one of the most popular data mining algorithms. Its efficiency comes from the assumption of attribute independence [40, 41]. It is a simple probabilistic classifier based on Bayes' theorem and it uses the method of maximum similarity. It can work in complex real-world situations and it requires a small amount of training data [42, 43]. BN is an effective tool to demonstrate complicated relationships and a powerful tool for knowledge representation and reasoning. It is also suitable for small and incomplete datasets. It is particularly convenient for modeling causal relationships among variables (with a combination of different sources of information) and handling of uncertainty for decision analysis [13]. It is a model of probabilistic inference defined by a triplet (X, G, P) , where $X = (X_1, X_2, \dots, X_n)$ is the set of factors, G is a directed acyclic graph, and P is a joint probability distribution. The nodes for G are labeled with the

elements of X and the arcs of G represent the probabilistically conditional dependence relationship between nodes. Each factor is associated with a conditional probabilistic table that defines the probability of each state for that factor [10, 44]. The k-NN is a simple and intuitive classifier. It classifies a sample based on the k-closest training samples in the feature space, by computing the distances between the new sample and the samples of the training set. According to these distances it classifies the sample to the class of the closest training samples (neighbors) [39, 45]. An MLP is a class of feedforward ANN and one of the most widely used. It consists of the input layer, which receives external inputs, one or more hidden layers, and an output layer. Each layer includes one or more neurons linked between successive layers. The Back-Propagation (BP) algorithm was used for the training of the model. In this technique, information about the errors of the network on known data is propagated backwards, the connections are adjusted, and the error is minimized [49]. All Weka parameters for each of the above models are summarized in the Appendix.

The classifiers were designed to classify the cases into one of the afore-mentioned 5 classes. Each classifier takes as inputs the data from the reports of each event and outputs its classification group, providing in this way a prediction regarding the actual occupational risk of hazards of each hospital worker in their certain hospital workplace, and under certain OSH conditions.

F. Training, Validation, and Performance Evaluation

The train/test split method [50] was used as a validation technique for training. The available dataset was divided into 2 subsets: the training set (80% of cases - 379 cases) which was used to train the classifiers, and the test set (20% of cases - 97 cases) which was used to evaluate their predictive performance (previously unseen data). The 2 sets were properly stratified so that the distribution of classes in each subset is approximately

the same as that in the initial dataset. Thus, each subset contains representative samples of the same larger population. To minimize sampling bias, k-fold cross-validation was also used for the whole dataset for training and test (k=10). In this technique, k dataset divisions are used for equivalent separation of the dataset into folds, k-1 folds are used for training and one-fold is used for test. The process is repeated k times. Through this process, each one of the available data is used one time as a member of the test set and k-1 times as a member of the training set [48, 49]. Data entry was done in a spreadsheet file, since this format can be easily used in daily OSH practice by the safety engineer and hospital staff. Files were saved in .csv type. Inserting and processing files to Weka needed their conversion to the .arff type. Considering that incidents and accidents are stochastic events, that are very differentiated in detail, a certain methodology has been used for risk assessment (including a certain set of risk variables that should be investigated in every case). Feature selection method was not applied.

In order to evaluate the performance of the proposed DSS model compared to the safety engineer's reports involved in this study, we calculated the recall (sensitivity), and the Area Under the Receiver Operating Characteristic (ROC) Curve of the models on the basis of detecting the 5 output classes. The ROC curve is a standard validation approach for probabilistic classifiers [50, 51]. A ROC curve is a two-dimensional graph in which the True Positive Rate (TPR) (sensitivity or recall) is plotted on the y-axis and the False Positive Rate (FPR or 1 - specificity) is plotted on the x-axis, as a function of the threshold u , ranging from 0 to 1 [55]. The Area Under the Curve (AUC) can be interpreted as a measure of overall accuracy, varying from 0.5 (random guess) to 1 (perfect performance) [56]. The recall metric answers the question of what proportion of actual positives is correctly classified. Its choice was based on the fact that recall measure penalizes false negatives but not false positives [57]. In an OSH approach, we want to capture as many positives as possible (we want to capture any contingent incident or accident even if we are not totally sure). These metrics can be expressed according to the number of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) classifications through the following equations [58]:

$$\text{Recall} = \text{Sensitivity} = \text{TPR} = \text{TP}/(\text{TP}+\text{FN}) \quad (1)$$

$$1 - \text{Specificity} = \text{FPR} = \text{FP}/(\text{TN}+\text{FP}) \quad (2)$$

The confusion matrices obtained through testing the classifiers on the training and the test sets were also used for the predictive performance of the models. Each column in a confusion matrix represents the instances that are classified in one class and each row represents the instances in their real class. The error rate of the classified instances was calculated for each class.

III. RESULTS

A. Performance Comparison of the Algorithms

Figure 2 presents the performance of the different models compared to the safety engineer's reports involved in this study, in terms of mean recall, AUC, error rate, and the overall

accuracy of each model (percentage of the correctly classified instances). The detailed table with the predictive performance of the models for each of the 5 classes is included in the Appendix.

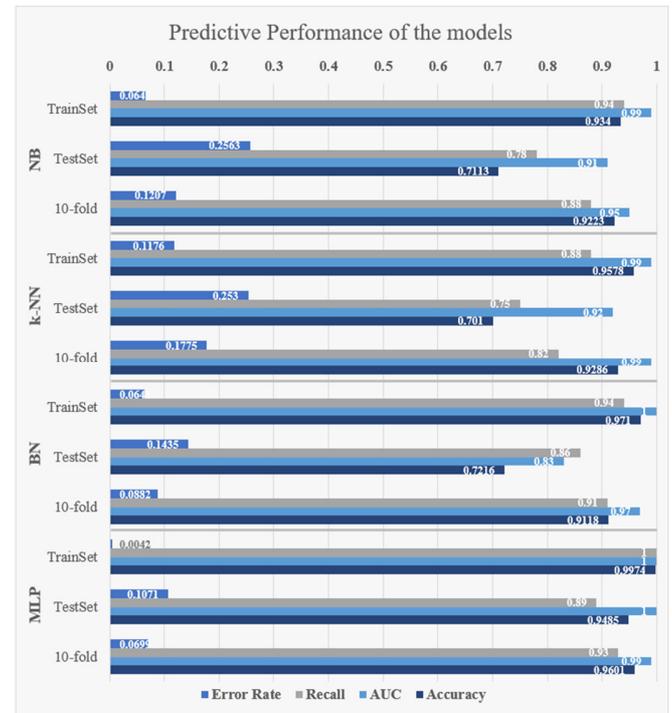


Fig. 2. The predictive performance of the models.

The overall classification accuracies of the developed models on the training set, test set, and 10-fold cross set were 97%, 77%, and 93% respectively. The MLP showed the highest accuracy for the test set and the 10-fold cross set (94.85% and 96.01%, respectively). The MLP and the BN showed the lower average Error Rate (ER) on the test set (10.71% and 14.35%, respectively) and on the 10-fold cross set (6.99% and 8.82%, respectively). The MLP showed the highest sensitivity (recall) for the test set and the 10-fold cross set (89% and 93%, respectively). In addition, this model showed the most highly balanced combination of sensitivity results on the 10-fold cross set (see Figure 2). The average AUC was high for both the test and the 10-fold cross sets (0.91 and 0.98, respectively). The MLP showed the highest AUC value for the test set and the 10-fold cross set (1.00 and 0.99, respectively). It must be noted that all of the developed models showed high AUC (range: 0.83-1.00 on the test set and range: 0.95-0.99 on the 10-fold cross set).

B. Practical Implications

Safety engineers carry out inspections of workplaces, equipment, and OSH conditions, during their visits. The safety engineers complete risk assessment tables manually in printed versions to record the results of their inspections. This process is repeated after every incident or accident. In this case, the assessment focuses on causes that are related to the specific incident or accident. The developed DSS model aims to

support the safety engineer with evidence for the risk values of hazards and recommended actions. The MLP and the BN models (10-fold cross set) were used in the DSS, as they were the models with the best predictive performance. A further elaboration of the solution methodology including the proposed model is depicted in Figure 3. The utilization of the DSS model can start in every case of incident, accident or audit. The

associated data files should be then updated by the responsible personnel and the safety engineer. Coding these event reports to instances in risk assessment tables should follow in order to update the database. Afterwards, the re-trained or the existing models can be used to support with evidence any OSH decision (efficiency of OSH intervention, employee's movement, appropriateness of workplace or equipment etc.).

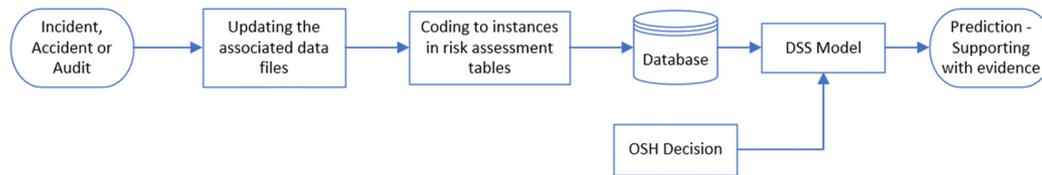


Fig. 3. A solution methodology including the proposed model.

In order to evaluate the proposed DSS model, a case study was carried out. The new case study test set resulted from a regular process. A typical instance of this process is the following: The safety engineers of the hospital, during the audit in specific departments and workplaces (01/2020), proceeded in identifying hazards of possible incidents. In parallel, they proposed specific corrective actions for the limitation or the elimination of these hazards by inserting the desired risk values for the model's variables (assuming that corrective actions have been taken). The basic safety engineers' observations, during the one-month period, and the corresponding incidents were:

- In the archive room, an employee used a chair instead of an appropriate ladder for catching files from shelves. Moreover, there was a lack of tidiness in the workplace and several obstacles prevented the free movement (possible incident: "Falling").
- When the safety engineer came out of the archive room, he observed that even though the floor had just been wiped out, the sign "Caution: slippery floor" had not been placed (possible incident: "Falling").
- During their visit to the pathologist laboratory, the safety engineer found out that there was no preventive nor corrective maintenance file for one microtome according to its operation manual. In a conversation with an employee in this laboratory, the latter expressed complains about OSH conditions (possible incident: "Needlestick/Cut").
- During the visit in the surgery room, a surgical nurse complained about excessive workload and lack of training (possible incident: "Needlestick/Cut").
- In the laundry room, and specifically in the bedsheet folder machine, the proper guard of the moving part was missing (possible incident: "Incident").
- A member of the waiting staff used an unfunctional trolley and lifted many food trays simultaneously to carry them to each patient room (possible incident: "Incident").

For each one of these instances, a "Safety" class was also recorded, relatively to the corresponding measures – corrective actions. The dataset that came up from this scenario was used

as a new case study test set for the proposed DSS model. The total 12 instances were correctly classified by the DSS model. Only one instance was incorrectly classified by BN model. Even in this case, the result was not classified as a "Safety" instance but one "Falling" was classified as "Incident". This result was expected since this kind of incident (falling from chair) was not included in the training set of the used model. To evaluate further the usability of the proposed system, the DSS model was used to support human resource management for an internal employee movement decision approached from an OSH perspective. In January 2020, 10 hospital workers were moved in different departments to cover the respective demands. The safety engineer inspected the destination workplaces and completed the relevant risk assessment tables manually in printed versions. These tables were classified to 9 instances of "Safety" and 1 instance of "Incident". The proposed interventions for the case of "Incident" included training on the work practice (waiting staff duties) and corrective actions in equipment (lighter trays and functional trolleys). The movement was done after the completion of these measures. The "Safety" state was confirmed by the safety engineer one week after the movement.

IV. DISCUSSION AND CONCLUSIONS

The aim of this study was to create a DSS model for the prediction of incidents and accidents in the hospital workplace. This model is based on the incident and accident data from reports of personnel and safety engineer investigations and audits in the hospital workplace. The preliminary results suggest that the proposed system may support OSH decisions by predicting incidents and accidents, and also by assessing the effectiveness of the OSH interventions. Through the evaluation of the different classification models, we concluded that MLP on the 10-fold cross set had the most balanced predictive performance in terms of overall accuracy (96.01%), ER (6.99%), AUC (0.990), and recall (93%), rendering this model reliable as a tool in OSH analysis in accordance with [13–17, 25]. Furthermore, the high predictive performance of all the developed models is notable (AUC range: 0.95-0.99 and average accuracy: 93% on the 10-fold cross set).

A major benefit of the proposed DSS model is that it utilizes processes and resources that already exist and are

applied in the daily hospital working life. The exploitation of these printed data records for incidents and accidents, and safety engineer audits and investigation reports are essential for OSH management. The utilization is related with identifying hazards that can lead to incident or accident, reducing the occupational risk, evaluating the efficacy of OSH measures – interventions, supporting management decisions (employee’s recruitment or movement to a different department, etc.) from the OSH perspective, improving OSH conditions and reducing the overall cost of injuries [23, 34]. The proposed change is related to the transfer of coded paper records into a DSS which is promising for OSH decisions. This process could result in faster and more accurate predictions according to OSH injuries and interventions, in accordance with [9, 34, 62], emphasizing the significance of quantitative analysis of empirical injury data in the safety science [25]. Another advantage of the proposed model is the fact that it can be adjusted to various working sectors, depending on the training sets and the existing records of incidents and accidents in each of them. For its application in other hospitals, the system can be easily updated with new datasets, according to the status of each hospital (incidents and accidents that take place and OSH conditions). The frequency of updating or retraining the classification models depends on the type and the frequency of incidents and accidents.

To maximize the usability and the practicality of the DSS model, a certain approach for the risk values of various factors should be followed. For instance, the classes "Incident" and "Accident" may be confused during initial workplace risk assessment of a random scenario, because of the lack of knowledge about the necessity of the required medical leave and its length. The dataset for accidents needs to be improved, however, any kind of incident should be interpreted as an alert which can anytime lead to an accident. Incidents are very important for OSH, in the sense of "near miss", as a situation that could lead to an accident, particularly in the case of repeated incidents of the same kind [59, 60]. In addition, certain incidents can evolve in occupational diseases (like blood-borne diseases because of needlestick/cuts and chronic musculoskeletal disorders) [61]. An occupational disease according to the domestic legislative framework is equivalent to an occupational accident [62]. The experience of the safety engineer and the accordance with the corresponding methodology are very important key factors for this assessment.

According to the proposed DSS model, an alert for taking corrective actions is related to all the classes that are different from the "Safety" one. The proposed measures and corrective actions target the elimination of hazards. The elimination is not always possible in all cases. Instead of elimination, limitation of hazards to acceptable levels should be the next step forward. The relation between cost and optimal level of risk is formed according to the ALARP (As Low As Reasonably Practicable) theory [60, 61]. In other words, entering the minimum values of risk for "Safety" class has no practical meaning since total lack of risk cannot happen in real conditions, both in terms of cost and inherent risk (a risk that is related to the nature of work). On the contrary, "armor logic" should not be applied during risk assessment by inserting maximum values during identifying and highlighting a hazard. In any case, the

interesting point related to the values of risk variables is the threshold where the value of a specific risk factor leads to the transition from one class to another.

A field for further research is the extensibility of the proposed DSS model and its development as a user-friendly web application platform. The development of the proposed DSS as a web app could make it more usable and accessible. Furthermore, the directions of this extensibility can include the facilitation of data entry to the system through Natural Language Process (NLP) (because there are several handwritten records) and the interconnection between the proposed DSS model and other tools/applications that will improve the efficiency of the proposed measures and corrective actions or/and the validity of risk assessment (scales for musculoskeletal disorders like key item method [66], working stress questionnaires, optimal allocation of OSH capital investment, etc.).

ACKNOWLEDGMENT

We would like to thank the Administration, the Scientific Council, and the personnel of Metaxa Cancer Hospital of Piraeus, Greece for their approval and participation in the implementation of the proposed model.

APPENDIX

WEKA PARAMETERS FOR CLASSIFICATION ALGORITHMS

Model	Weka parameters
NB	batch size = 100 debug = False kernel estimator = False doNotCheckCapabilities = False useSupervisedDiscretization = False
BN	batch size = 100 debug = False doNotCheckCapabilities = False estimator = Sample Estimator – A 0.5 searchAlgorithm = K2-P1-S BAYES useADTree = False
k-NN	k = 9 batch size = 100 crossValidate = False debug = False distanceWeighting = No doNotCheckCapabilities = False meanSquared = False nearestNeighbourSearchAlgorithm = LinearNNSearch - Euclidean distance First-Last
MLP	autoBuild = True batch size = 100 debug = False decay = False doNotCheckCapabilities = False hiddenLayers = a, where a = (number of attributes + number of classes) / 2 = 44 learning rate = 0.3 momentum = 0.2 nominalToBinaryFilter = True normalizeAttributes = True normalizeNumericClass = True reset = True seed = 0 trainingTime (number of epochs) = 500 validationSetSize = 0 validation threshold = 20

PREDICTIVE PERFORMANCE OF THE DIFFERENT MODELS

VM ^a	MA ^a (%)	Performance Evaluation Metrics					MV ^a	
		Accident	Falling	Incident	Cutting	Safety		
NB								
Train set (80%) (379)	93.40	ER (%)	0.00	0.00	23.40	3.70	5.26	6.47
		Recall	1.000	1.000	0.766	0.963	0.947	0.94
		AUC	1.000	0.995	0.968	0.982	0.986	0.99
Test set (20%) (97)	71.13	ER (%)	60.00	0.00	16.67	3.57	47.92	25.63
		Recall	0.600	1.000	0.833	0.964	0.521	0.78
		AUC	0.753	0.992	0.862	0.964	0.991	0.91
10-fold cross (476)	92.23	ER (%)	8.69	15.00	28.81	3.67	4.20	12.07
		Recall	0.913	0.850	0.712	0.963	0.958	0.88
		AUC	0.995	0.944	0.890	0.969	0.966	0.95
k-NN								
Train set (80%) (379)	95.78	ER (%)	0.00	37.5	21.28	0.00	0.00	11.76
		Recall	1.000	0.625	0.787	1.000	1.000	0.88
		AUC	1.000	0.986	0.990	0.997	1.000	0.99
Test set (20%) (97)	70.10	ER (%)	0.00	50.00	25.00	3.57	47.92	25.30
		Recall	1.000	0.500	0.750	0.964	0.521	0.75
		AUC	0.875	0.985	0.983	0.997	0.751	0.92
10-fold cross (476)	92.86	ER (%)	0.00	50.00	37.29	1.47	0.00	17.75
		Recall	1.000	0.500	0.627	0.985	1.000	0.82
		AUC	1.000	0.981	0.966	0.994	1.000	0.99
BN								
Train set (80%) (379)	97.10	ER (%)	0.00	18.75	10.64	2.78	0.00	6.43
		Recall	1.000	0.813	0.894	0.972	1.000	0.94
		AUC	1.000	0.993	0.987	0.996	1.000	1.00
Test set (20%) (97)	72.16	ER (%)	0.00	0.00	16.67	7.14	47.92	14.35
		Recall	1.000	1.000	0.833	0.929	0.521	0.86
		AUC	0.875	1.000	0.810	0.900	0.570	0.83
10-fold cross (476)	91.18	ER (%)	0.00	15.00	13.56	5.88	9.66	8.82
		Recall	1.000	0.850	0.864	0.941	0.903	0.91
		AUC	0.975	0.985	0.979	0.994	0.913	0.97
MLP								
Train set (80%) (379)	99.74	ER (%)	0.00	0.00	2.12	0.00	0.00	0.42
		Recall	1.000	1.000	0.979	1.000	1.000	1.00
		AUC	1.000	1.000	0.984	0.998	1.000	1.00
Test set (20%) (97)	94.85	ER (%)	0.00	25.00	25.00	3.57	0.00	10.71
		Recall	1.000	0.750	0.750	0.964	1.000	0.89
		AUC	1.000	1.000	0.992	0.996	1.000	1.00
10-fold cross (476)	96.01	ER (%)	0.00	10.00	22.03	2.94	0.00	6.99
		Recall	1.000	0.900	0.780	0.971	1.000	0.93
		AUC	0.999	0.996	0.971	0.991	1.000	0.99

^a VM: Validation Model, MA: Model Accuracy, MV: Mean Value.

REFERENCES

- [1] D. Elsler, J. Takala, and J. Remes, "An international comparison of the cost of work-related accidents and illnesses," *European Agency for Safety and Health at Work*, 2017.
- [2] "Good OSH is good for business," *EU-OSHA*. <https://osha.europa.eu/el/themes/good-osh-is-good-for-business> (accessed May 26, 2021).
- [3] K. Cosic, S. Popovic, M. Sarlija, I. Kesedzic, and T. Jovanovic, "Artificial intelligence in prediction of mental health disorders induced by the COVID-19 pandemic among health care workers," *Croatian Medical Journal*, vol. 61, no. 3, pp. 279–288, Jun. 2020, <https://doi.org/10.3325/cmj.2020.61.279>.
- [4] K. Dimoulas, G. Kollias, C. Bagavos, and T. Ganetaki, *Work and health problems in Greece*. Athens, Greece: INE-GSEE Work Institute, 2015.
- [5] *Hospital Inventory 2018*. Athens, Greece: Hellenic Statistical Authority, 2020.
- [6] *Survey on Accidents at Work, 2018*. Athens, Greece: Hellenic Statistical Authority, 2020.
- [7] S. Sarkar and J. Maiti, "Machine learning in occupational accident analysis: A review using science mapping approach with citation network analysis," *Safety Science*, vol. 131, p. 104900, Nov. 2020, <https://doi.org/10.1016/j.ssci.2020.104900>.
- [8] F. Siddiqui, M. A. Akhund, A. H. Memon, A. R. Khoso, and H. U. Imad, "Health and Safety Issues of Industry Workmen," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3184–3188, Aug. 2018, <https://doi.org/10.48084/etasr.2138>.
- [9] S. Y. Far, R. Mirzaei, M. B. Kattini, M. Haghshenas, and Z. Sayahi, "Assessment of Health, Safety and Environmental Risks of Zahedan City Gasoline Stations," *Engineering Technology & Applied Science Research*, vol. 8, no. 2, pp. 2689–2692, 2018.
- [10] S. J. Bertke, A. R. Meyers, S. J. Wurzelbacher, J. Bell, M. L. Lampl, and D. Robins, "Development and evaluation of a Naïve Bayesian model for coding causation of workers' compensation claims," *Journal of Safety Research*, vol. 43, no. 5, pp. 327–332, Dec. 2012, <https://doi.org/10.1016/j.jsr.2012.10.012>.
- [11] G. Nanda, K. M. Grattan, M. T. Chu, L. K. Davis, and M. R. Lehto, "Bayesian decision support for coding occupational injury data," *Journal of Safety Research*, vol. 57, pp. 71–82, Jun. 2016, <https://doi.org/10.1016/j.jsr.2016.03.001>.
- [12] J. E. M. E. Martin, J. T.-G. Taboada-Garcia, S. G. Gerassis, A. S. Saavedra, and R. Martinez-Alegria, "Bayesian network analysis of accident risk in information-deficient scenarios," *Revista de la Construcción. Journal of Construction*, vol. 16, no. 3, pp. 439–446, 2017, <https://doi.org/10.7764/RDLC.16.3.439>.
- [13] A. P. C. Chan, F. K. W. Wong, C. K. H. Hon, and T. N. Y. Choi, "A Bayesian Network Model for Reducing Accident Rates of Electrical and Mechanical (E&M) Work," *International Journal of Environmental Research and Public Health*, vol. 15, no. 11, Nov. 2018, Art. no. 2496, <https://doi.org/10.3390/ijerph15112496>.
- [14] L. Sanmiquel, M. Bascompta, J. M. Rossell, H. F. Anticoi, and E. Guash, "Analysis of Occupational Accidents in Underground and Surface Mining in Spain Using Data-Mining Techniques," *International Journal of Environmental Research and Public Health*, vol. 15, no. 3, Mar. 2018, Art. no. 462, <https://doi.org/10.3390/ijerph15030462>.
- [15] A. Soltanzadeh, I. Mohammadfam, S. Mahmoudi, B. A. Savareh, and A. M. Arani, "Analysis and forecasting the severity of construction accidents using artificial neural network," *Safety promotion and injury prevention (Tehran)*, vol. 4, no. 3, pp. 185–192, 2016.
- [16] D. A. Patel and K. N. Jha, "Neural Network Approach for Safety Climate Prediction," *Journal of Management in Engineering*, vol. 31, no. 6, Nov. 2015, Art. no. 05014027, [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000348](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000348).
- [17] A. M. Abubakar, H. Karadal, S. W. Bayighomog, and E. Merdan, "Workplace injuries, safety climate and behaviors: application of an artificial neural network," *International Journal of Occupational Safety and Ergonomics*, vol. 26, no. 4, pp. 651–661, Oct. 2020, <https://doi.org/10.1080/10803548.2018.1454635>.
- [18] F. A. Moayed and R. L. Shell, "Application of Artificial Neural Network Models in Occupational Safety and Health Utilizing Ordinal Variables," *The Annals of Occupational Hygiene*, vol. 55, no. 2, pp. 132–142, Mar. 2011, <https://doi.org/10.1093/annhyg/meq079>.
- [19] I. Mohammadfam, A. Soltanzadeh, A. Moghimbeigi, and B. A. Savareh, "Use of Artificial Neural Networks (ANNs) for the Analysis and Modeling of Factors That Affect Occupational Injuries in Large Construction Industries," *Electronic Physician*, vol. 7, no. 7, pp. 1515–1522, Nov. 2015, <https://doi.org/10.19082/1515>.
- [20] S. Sarkar, S. Vinay, R. Raj, J. Maiti, and P. Mitra, "Application of optimized machine learning techniques for prediction of occupational accidents," *Computers & Operations Research*, vol. 106, pp. 210–224, Jun. 2019, <https://doi.org/10.1016/j.cor.2018.02.021>.
- [21] J. Bao, J. Johansson, and J. Zhang, "An Occupational Disease Assessment of the Mining Industry's Occupational Health and Safety

- Management System Based on FMEA and an Improved AHP Model," *Sustainability*, vol. 9, no. 1, Jan. 2017, Art. no. 94, <https://doi.org/10.3390/su9010094>.
- [22] H. R. S. A. Mard, A. Estiri, P. Hadadi, and M. S. A. Mard, "Occupational risk assessment in the construction industry in Iran," *International Journal of Occupational Safety and Ergonomics*, vol. 23, no. 4, pp. 570–577, Oct. 2017, <https://doi.org/10.1080/10803548.2016.1264715>.
- [23] L. Comberti, M. Demichela, G. Baldissoni, G. Fois, and R. Luzzi, "Large Occupational Accidents Data Analysis with a Coupled Unsupervised Algorithm: The S.O.M. K-Means Method. An Application to the Wood Industry," *Safety*, vol. 4, no. 4, Dec. 2018, Art. no. 51, <https://doi.org/10.3390/safety4040051>.
- [24] N. D. Nath, T. Chaspari, and A. H. Behzadan, "Automated ergonomic risk monitoring using body-mounted sensors and machine learning," *Advanced Engineering Informatics*, vol. 38, pp. 514–526, Oct. 2018, <https://doi.org/10.1016/j.aei.2018.08.020>.
- [25] F. Davoudi Kakhki, S. A. Freeman, and G. A. Mosher, "Utilization of Machine Learning in Analyzing Post-incident State of Occupational Injuries in Agro-Manufacturing Industries," in *Advances in Safety Management and Human Performance*, P. M. Arzes and R. L. Boring, Eds. New York, NY, USA: Springer, 2020, pp. 3–9.
- [26] S. D. Mwmc *et al.*, "Ethical Considerations of Using Machine Learning for Decision Support in Occupational Health: An Example Involving Periodic Workers' Health Assessments," *Journal of Occupational Rehabilitation*, vol. 30, no. 3, pp. 343–353, Sep. 2020, <https://doi.org/10.1007/s10926-020-09895-x>.
- [27] F. Ladstatter, E. Garrosa, B. Moreno-Jimenez, V. Ponsoda, J. M. R. Aviles, and J. Dai, "Expanding the occupational health methodology: A concatenated artificial neural network approach to model the burnout process in Chinese nurses," *Ergonomics*, vol. 59, no. 2, pp. 207–221, Feb. 2016, <https://doi.org/10.1080/00140139.2015.1061141>.
- [28] Y.-H. Kim and M.-H. Jung, "Effect of occupational health nursing practice on musculoskeletal pains among hospital nursing staff in South Korea," *International Journal of Occupational Safety and Ergonomics*, vol. 22, no. 2, pp. 199–206, Apr. 2016, <https://doi.org/10.1080/10803548.2015.1078046>.
- [29] A. Fonseca, I. Abreu, M. J. Guerreiro, C. Abreu, R. Silva, and N. Barros, "Indoor Air Quality and Sustainability Management—Case Study in Three Portuguese Healthcare Units," *Sustainability*, vol. 11, no. 1, Jan. 2019, Art. no. 101, <https://doi.org/10.3390/su11010101>.
- [30] S. Lin, N. Chaiear, J. Khiewyoo, B. Wu, and N. P. Johns, "Preliminary Psychometric Properties of the Chinese Version of the Work-Related Quality of Life Scale-2 in the Nursing Profession," *Safety and Health at Work*, vol. 4, no. 1, pp. 37–45, Mar. 2013, <https://doi.org/10.5491/SHAW.2013.4.1.37>.
- [31] W. Turnberg and W. Daniell, "Evaluation of a healthcare safety climate measurement tool," *Journal of Safety Research*, vol. 39, no. 6, pp. 563–568, Jan. 2008, <https://doi.org/10.1016/j.jsr.2008.09.004>.
- [32] A. K. Celik, E. Oktay, and K. Cebi, "Analysing workplace violence towards health care staff in public hospitals using alternative ordered response models: the case of north-eastern Turkey," *International Journal of Occupational Safety and Ergonomics*, vol. 23, no. 3, pp. 328–339, Jul. 2017, <https://doi.org/10.1080/10803548.2017.1316612>.
- [33] M. Stefanovic, D. Tadic, M. Djapan, and I. Macuzic, "Software for Occupational Health and Safety Risk Analysis Based on a Fuzzy Model," *International Journal of Occupational Safety and Ergonomics*, vol. 18, no. 2, pp. 127–136, Jan. 2012, <https://doi.org/10.1080/10803548.2012.11076923>.
- [34] A. Sklad, "Assessing the impact of processes on the Occupational Safety and Health Management System's effectiveness using the fuzzy cognitive maps approach," *Safety Science*, vol. 117, pp. 71–80, Aug. 2019, <https://doi.org/10.1016/j.ssci.2019.03.021>.
- [35] V. Ravuri *et al.*, "Group-specific models of healthcare workers' well-being using iterative participant clustering," in *Second International Conference on Transdisciplinary AI*, Irvine, CA, USA, Sep. 2020, pp. 115–118, <https://doi.org/10.1109/TransAI49837.2020.00026>.
- [36] K. Vallmuur, "Machine learning approaches to analysing textual injury surveillance data: A systematic review," *Accident Analysis & Prevention*, vol. 79, pp. 41–49, Jun. 2015, <https://doi.org/10.1016/j.aap.2015.03.018>.
- [37] "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)." Publications Office of the European Union, Apr. 27, 2016.
- [38] "Home - Weka Wiki," *The University of Waikato*. <https://waikato.github.io/weka-wiki/> (accessed May 27, 2021).
- [39] "Memorandum on Occupational Risk Assessment." Directorate-General for Employment in Labor Relations and Social Affairs (DG V) of the European Union, 1997.
- [40] "Occupational Risk Assessment." Technical Chamber of Greece, 2001.
- [41] S. Drivas, K. Zorba, and T. Koukoulaki, *Methodological guide for the assessment and prevention of occupational risk*. Athens, Greece: Hellenic Institute of Occupational Health and Safety, 2000.
- [42] P. Bountris *et al.*, "An Intelligent Clinical Decision Support System for Patient-Specific Predictions to Improve Cervical Intraepithelial Neoplasia Detection," *BioMed Research International*, vol. 2014, 2014, <https://doi.org/10.1155/2014/341483>.
- [43] S. Chen, G. I. Webb, L. Liu, and X. Ma, "A novel selective naïve Bayes algorithm," *Knowledge-Based Systems*, vol. 192, Mar. 2020, Art. no. 105361, <https://doi.org/10.1016/j.knsys.2019.105361>.
- [44] K. Koutroumbas and S. Theodoridis, *Pattern Recognition*, 4th ed. London, UK: Elsevier, 2008.
- [45] M. A. Burhanuddin, R. Ismail, N. Izzaimah, A. A.-J. Mohammed, and N. Zainol, "Analysis of Mobile Service Providers Performance Using Naive Bayes Data Mining Technique," *International Journal of Electrical & Computer Engineering*, vol. 8, no. 6, pp. 5153–5161, 2018.
- [46] R. Shinde, S. Arjun, P. Patil, and J. Waghmare, "An Intelligent Heart Disease Prediction System Using K-Means Clustering and Naive Bayes Algorithm," *International Journal of Computer Science and Information Technologies*, vol. 6, no. 1, pp. 637–639, 2015.
- [47] S. J. Russell, P. Norvig, S. Russell, and Russell, *Artificial intelligence: A Modern Approach*. New Jersey, USA: Prentice Hall, 2010.
- [48] D. Michie, D. J. Spiegelhalter, C. C. Taylor, and J. Campbell, Eds., *Machine learning, neural and statistical classification*. New York, NY, USA: Ellis Horwood, 1995.
- [49] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Hoboken New Jersey, USA: Wiley, 2001.
- [50] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2011.
- [51] S. M. Weiss and C. A. Kulikowski, *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning and Expert Systems*. San Mateo, CA, USA: Morgan Kaufmann, 1990.
- [52] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge, MA, USA: Cambridge University Press, 2008.
- [53] A. Nola *et al.*, "Occupational accidents in temporary work," *La Medicina Del Lavoro*, vol. 92, no. 4, pp. 281–285, Aug. 2001.
- [54] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006, <https://doi.org/10.1016/j.patrec.2005.10.010>.
- [55] J. López-García, M. Saldaña, S. Herrero, and J. Gutiérrez, "Bayesian network analysis of the influence of labour market variables on accident rates of workers in Spain," in *Risk, Reliability and Safety: Innovating Theory and Practice: Proceedings of ESREL 2016*, Glasgow, UK, Sep. 2016, pp. 1660–1667, <https://doi.org/10.1201/9781315374987-250>.
- [56] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, Apr. 1982, <https://doi.org/10.1148/radiology.143.1.7063747>.
- [57] S. Alvarez, "An exact analytical relation among recall, precision, and classification accuracy in information retrieval," Boston College,

- Boston, MA, USA, Technical Report BCCS-02-01 (2002): 1-22, Jan. 2002.
- [58] R. Burduk, "Classification Performance Metric for Imbalance Data Based on Recall and Selectivity Normalized in Class Labels," *arXiv:2006.13319 [cs, stat]*, Jun. 2020, Accessed: May 26, 2021. [Online]. Available: <http://arxiv.org/abs/2006.13319>.
- [59] O. Ug, S. Wd, S. M, and P. A, "Improve Process Safety with Near-Miss Analysis," *Chemical Engineering Progress*, vol. 109, no. 5, pp. 20–27, 2013.
- [60] M. G. Gnani, S. Andriulo, G. Maggio, and P. Nardone, "'Lean occupational' safety: An application for a Near-miss Management System design," *Safety Science*, vol. 53, pp. 96–104, Mar. 2013, <https://doi.org/10.1016/j.ssci.2012.09.012>.
- [61] E. Alexopoulos, *Greek and International experience of accidents at work and occupational diseases of hospital employees. Guide to Occupational Risk Assessment and Prevention*. Athens, Greece: EL.Y.A., 2007.
- [62] "Circular 45/24-06-2010: Occupational Accident 2010." Social Security Institution, 2010.
- [63] G. Reniers and T. Brijs, "An Overview of Cost-benefit Models/Tools for Investigating Occupational Accidents," *Chemical Engineering Transactions*, vol. 36, pp. 43–48, Apr. 2014, <https://doi.org/10.3303/CET1436008>.
- [64] Health and Safety Executive, "Risk management: Expert guidance - ALARP at a glance." <https://www.hse.gov.uk/managing/theory/alarplance.htm> (accessed May 26, 2021).
- [65] S. J. Bertke, A. R. Meyers, S. J. Wurzelbacher, J. Bell, M. L. Lampl, and D. Robins, "Development and evaluation of a Naïve Bayesian model for coding causation of workers' compensation claims," *Journal of Safety Research*, vol. 43, no. 5, pp. 327–332, Dec. 2012, <https://doi.org/10.1016/j.jsr.2012.10.012>.
- [66] K. Koklonis, A. Anastasiou, O. Petropoulou, S. Pitoglou, D. Iliopoulou, and D. Koutsouris, "Utilizing Key Item Method to Manage Musculoskeletal Disorders in a Hospital Workplace," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Berlin, Germany, Jul. 2019, pp. 3420–3423, <https://doi.org/10.1109/EMBC.2019.8857649>.

Attack Detection in IoT using Machine Learning

Maryam Anwer

Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
maryam.anwer2@gmail.com

Muhammad Umer Farooq

Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
umer@neduet.edu.pk

Shariq Mahmood Khan

Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
shariq@neduet.edu.pk

Waseemullah

Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
waseemu@neduet.edu.pk

Abstract—Many researchers have examined the risks imposed by the Internet of Things (IoT) devices on big companies and smart towns. Due to the high adoption of IoT, their character, inherent mobility, and standardization limitations, smart mechanisms, capable of automatically detecting suspicious movement on IoT devices connected to the local networks are needed. With the increase of IoT devices connected through internet, the capacity of web traffic increased. Due to this change, attack detection through common methods and old data processing techniques is now obsolete. Detection of attacks in IoT and detecting malicious traffic in the early stages is a very challenging problem due to the increase in the size of network traffic. In this paper, a framework is recommended for the detection of malicious network traffic. The framework uses three popular classification-based malicious network traffic detection methods, namely Support Vector Machine (SVM), Gradient Boosted Decision Trees (GBDT), and Random Forest (RF), with RF supervised machine learning algorithm achieving far better accuracy (85.34%). The dataset NSL KDD was used in the recommended framework and the performances in terms of training, predicting time, specificity, and accuracy were compared.

Keywords—cyber security; artificial intelligence; IoT; machine learning

I. INTRODUCTION

The Internet of Things (IoT) is probably the greatest modern advancement, considering its effect on our daily life, while the zones of its utilization are quickly expanding. In 2018, the quantity of IoT devices was roughly 28 billion. This amount is expected to touch 49.1 billion by 2022 and the showcase size of IoT is estimated to reach around \$10 trillion by 2022. IoT is recognized as a method regarding suitable mechanisms that interconnect by servers, sensors, and various software. A city structure, is shown in Figure 1 which comprises of three main layers: fog, cloud, and terminal layer.

The data obtained from the IoT are saved on the Cloud Computing (CC) ecosystem which has progressively high-level processors and sufficient memory. The cloud layer has grown

fast by the modern developments in IoT. Fog-to-things is created with a feasible clarification of those difficulties. In the fog layer, devices can experience some larger values of data basically given to the cloud layer, which decreases power damage, bandwidth, network traffic, and eliminates the data storage and communication challenges. In addition, it tries to accelerate the estimated method near the endpoint, facilitating some fast reply to the IoT-based urban use. There are two advantages of attack detection in the fog-to-things layer. Either the internet service provider or the network administrator can practice certain measures which can stop extensive destruction if these network attacks are recognized in the fog layer. Besides, this strategy does not prevent the regular daily experience for the people.

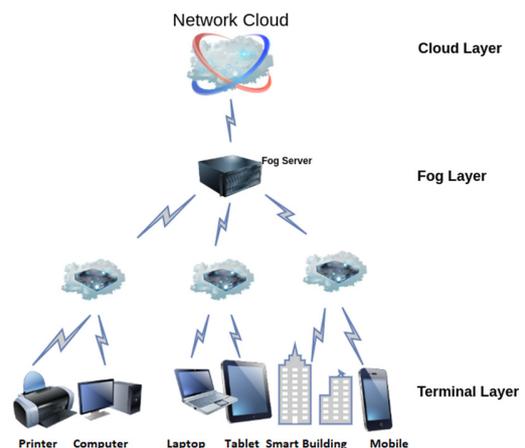


Fig. 1. Framework of a smart IoT-based city.

The model traces the web traffic which passes by every fog-to-things node. As fog-to-things connections resemble IoT devices, it will be more efficient to recognize these network attacks at the fog-to-things connections rather than at the cloud

layer. Immediate attack detection can inform the network controllers of the IoT devices of those attacks, which will then support them to evaluate and improve their systems. Artificial intelligence technology like Machine Learning (ML) will do the whole evaluation and send video pictures to people who can react speedily to solve troubles and maintain residents' safety. There are two types of attack detection: primarily signature-based or primarily anomaly-based. In the former, a primarily based solution fits the in-coming traffic closer to acknowledge attack/crime kinds in database whilst the latter checks the behavior of everyday traffic.

II. RELATED WORK

Many research studies in the application of ML have presently been presented in the domain, like object identification/recognition, pattern recognition, text processing, and image processing. In addition, much security research had been done using the Deep Learning (DL) approach [1]. Authors in [2] describe the expansion of big data and the evolution of IoT in a smart city. The author in [3] explains the evolution of CC and how big data have been engaged in the advancement of smart cities. He proposed a framework for managing big data for smart city purposes. The framework concentrates on difficulties related to smart cities for real-time decision planning. Many aspects and components of a smart city for upgrading the standard of the people are described in [4]. Authors in [5] suggest a platform design to secure a smart city facing cyber attackers. The structure is giving a warning DL model to identify attackers based on the user's data performance. In [6], resource administration methods of fog-computing are analyzed, well-systematic research in taxonomy is presented, and various features of resource administration, i.e. mass balancing, resource/device scheduling and allocation, job/task allocation, device/resource provisioning, and task offloading, are highlighted. The given resource management procedures are analyzed by estimating factors such as: Qos metrics, different researches, and applied methods. The benefits and hindrances of these approaches are compared.

Authors in [7] used the idea of an unknown and secure total plan (ASAS) in mist-based open distributed computing. In ASAS, the cloud gives advanced information about open cloud servers. When the ASAS is used, the fog gives devices to exchange information with PCS. Authors in [8] reported the advancements of remote sensor organization (WSN), correspondence innovation, and IoT innovation. Authors in [9] used ML techniques such as KNN, SVM, DT, Naïve Bayes, neural networks, and RF which can be applied in IDS. The authors compared ML models for multi and binary class combinations on the data set of Bot-IoT. Depending on these models they calculated the F1 score, recall, precision, and accuracy. The detection of attacks in FOG design was examined in [10], in which ML is compared with deep-learning neural networks working on an internet-available dataset.

Authors in [11] examined TCP SYN network attacks and authors in [12] introduced deep neural networks for attack detection in IoT systems. The self-adaptive identification method of the security index of the network was studied, performed risk assessment was conducted, and the system was mapped. Authors in [13] developed network NIDS based on

the conception of DL. For attack detection, they implemented network intrusion detection system on fog node. Authors in [14] used a novel method that combines isolation forest and One Class Support Vector Machine (OCSVM) with an active learning method to detect attacks with no prior information. Authors in [15] used a two-stage approach combining a fast preprocessing or filtering method with a variation auto encoder using reconstruction probability. Authors in [16] performed a Distributed Denial of Service (DDoS) attack using the ping of death technique and detected it using RF algorithm by using the WEKA tool with classification accuracy of 99.76%. Authors in [17] proposed the detection of network dictionary attacks using a data set collected as flows based on a clustered graph. The results of the mentioned methods on the CAIDA 2007 data set give high accuracy for the model.

III. GAP ANALYSIS

These are some prefaced problems taken from earlier research.

- Worst performance of the detection of attacks on the fog layer.
- Feature selection decreasing the accuracy.
- Low accuracy of DoS, R2L, and U2R attack types.
- Execution of multiple classifier algorithms on reduced data sets
- False positive rate and false negatives rate is still in doubt.

IV. A FRAMEWORK TO SOLVE ATTACK DETECTION IN IOT USING MACHINE LEARNING

The proposed model for this research work is an ordinary huge organization or a smart city going through an increasing variety of IoT-associated cyber threats, such as heavy-obligation DDoS attacks, achieved with an enormous botnet, e.g. Mirai, which exploit default or weak passwords. The current research specializes in advanced attacks which can be primarily based on violations of organizational protection guidelines. Once completed, an attacker is permitted to take advantage of individuals who connect unauthorized styles of IoT devices to the smart town. The previous approaches have been used broadly because of their excessive detection accuracy and low fake alarms. However, they lack the capability of seizing novel attacks. On the other hand, anomaly detection detects new attacks, although it lacks accuracy. In both procedures, classical ML analysis has been used prominently. Popular devices gaining knowledge of algorithms are incapable to detect complex data breaches [18]. In this research, we examined different algorithms for the different sub-processes of the framework shown in Figure 2 [19].

A. Approaches to Solve Attack Detection using ML

There are six main approaches in ML:

- Supervised learning: In this, the data should be labeled like feeding a model with multiple examples of files and decide whether they are malware or not. Based on this data labeling [20], the model could decide on extra data. It is also called the task driven approach.

- Ensemble learning: It is an addition of label data like supervised learning while combining multiple models to solve the task.
- Unsupervised learning: In this learning, unlabeled data are used and the model marks them by itself based on the data properties. It is considered to be the more powerful and it usually finds anomalies in the data set [21]. This is also called the data-driven approach.
- Semi-supervised learning: It tries to combine both supervised and unsupervised approaches when there is a data set with some labeled data [22].
- Reinforcement learning: This behavior should be used in a changing environment. It is also called the environment driven approach [23].
- Active learning: It works like a teacher who can help in correcting error and behavior in environmental changes [24]. It is a subclass of reinforcement learning.

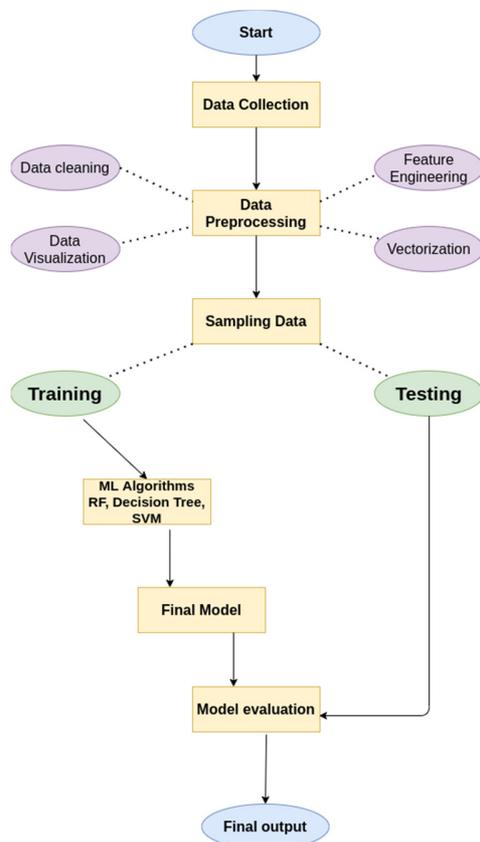


Fig. 2. A framework to solve attack detection in IoT using ML.

B. Attack Detection using ML Methods

In this part, the attack detection problems are studied by statistical classification of measurements using the implementation of ML. The spam filter in cyber security separates spam from different communications services. Spam is apparently the leading ML method applied in information security. The supervised learning labeled data method is

usually used for classification. In our research, we used the Gradient Boosted DT, SVM, and RF classifications and the results were compared.

1) Support Vector Machine

It is the most popular and widely recognized technique. It can be used for regression, but mostly it is used in classification algorithms. In SVM, we sketch data items by the point in an n -dimensional area where n represents the considered features [25]. It creates a hyper plane and separates the data into classes [26].

2) Gradient Boosted Decision Tree

GBDT is an ensemble of DTs. GBDT is an ML algorithm which constructs vulnerable DTs through the boosting technique. For building the tree ensemble, we need to train over the algorithms on different samples. Unfortunately, we cannot train them on a single set. GBDT uses the present-day ensemble to predict the label of every instance, after which the results are compared with the accurate labeled data. It works on large datasets and has high predicting power [27].

3) Random Forest

RF [28] is based on random subspace, bagging, and uses CART DTs as base algorithm. It works on both regression and classification. The education is achieved in parallel. It injects randomness within the learning (testing and training), a process in which each tree isn't the same with the others. In predictions, each tree is combined, which reduces the variance of prediction and hence improves performance [29].

V. EXPERIMENTS AND RESULTS

In this section, the dataset, which is applied for the experiment and for testing results, is described along with the performance metrics used for result comparison and the recommended model is reviewed by applying various selections and classifications. Three ML algorithms were applied for the evaluation of the given proposed model [30].

A. Dataset

The NSL KDD dataset was used in this research. This dataset is available in CSV and JSON files. We can use this for the model and the evaluation phase. The dataset is modifiable, extensible, and reproducible [31].

B. Proposed Method

Our research is a novel combination of several independent ML algorithms. In our framework, the first step is the dataset collection and analysis. In this process, the data were collected and observed deeply to analyze the types of data. In the data preprocessing step, the data were cleaned, visualized, and feature engineering was applied along with implemented vectorizations. Hence, the data were converted into feature vectors [32]. After the analysis of the NSL-KDD dataset, the attacks can be categorized into four principal classes:

- Unauthorized to remote (R2L)
- Denial-of-Service (DoS)
- Unauthorized to root super user privileges (U2R attack)

- Port scanning attack (Probe)

The details of each attack are shown in Figure 3.

TABLE I. TRAIN SET OF NSL-KDD

Type	Original records	Distinct records	Reduction rate
Attacks	3,925,640	262,178	93.3
Normal	972,782	812,814	16.44
Total	4,798,431	1,074,992	78.05

TABLE II. TEST SET OF NSL-KDD

Type	Original records	Distinct records	Reduction rate
Attacks	250,436	29,378	88.26
Normal	60,591	47,911	20.92
Total	311,027	77,289	75.15

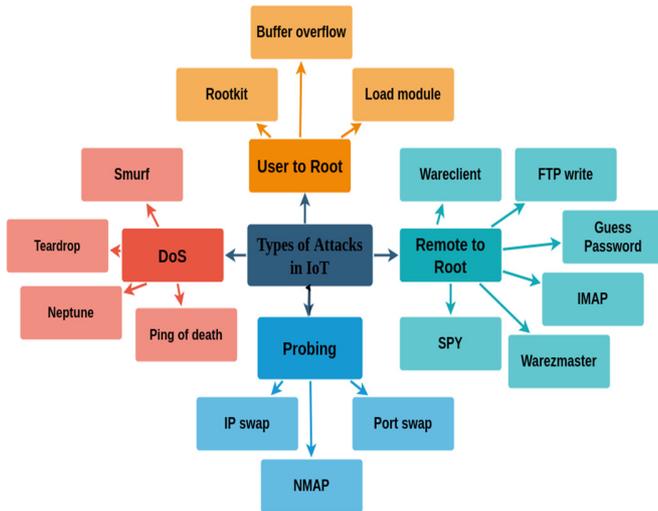


Fig. 3. Attack categories of the IoT ecosystem.

Our data is converted into feature vectors. The dataset is then split into 80% for training and 20% for testing sets (Tables I-II). For the learning algorithm, the training data set was utilized and our final model was deployed using a boosting technique. Figure 4 shows the data distribution in testing and training subsets.

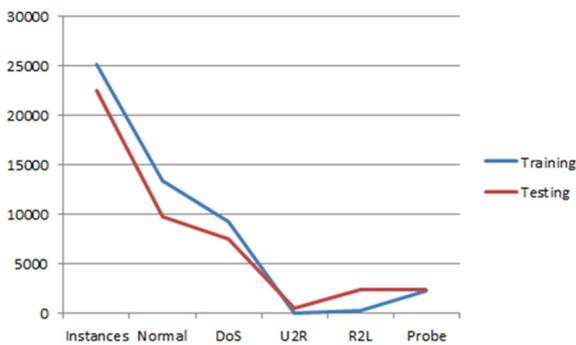


Fig. 4. NSL-KDD dataset distribution.

C. Algorithm

The algorithmic steps are mentioned below.

- Load the NSL-KDD data set.
- Apply the pre-processing technique.
- Divide into 80-20 ratios of testing and training datasets.
- Select feature selection vectors.
- The training dataset is given to the classifiers.
- The test data set is fed to the three selected classifiers for classification.
- Calculate accuracy, specificity, FPR, and TPR.

D. Classifiers and Training

For the model training, RF supervised ML algorithm was selected. The algorithms which combine DT with ensemble learning have several advantages, such as their need of only a few input parameters and their resistance to overfitting. The number of tree parameters is set to be 500. When the number of branches increases, the variance would be decreased without ensuing in bias. RF has changed into applied to traffic data sets, which include in-network traffics misuse detection and Command and Control (C&C) IoT attack detection from traffic flow-base [33].

E. Performance Metrics

In the suggested framework, four performance metrics were considered: Accuracy (A), Training Time (TT), which is the total time to train a classifier, Specificity (S), and Prediction Time (PT), which is the total time which an algorithm takes to predict all the data. TP (true positive) represents the correct identification of an attack, FP (false positive) represents the incorrect identified attacks, TN (true negative) represents the correctly identified normal connections, and FN (false negative) represents the number of attacks that were not correctly identified [34]:

$$\text{Accuracy} = A$$

$$\text{True Positive} = \Theta$$

$$\text{False Positive} = \xi$$

$$\text{True Negative} = \omega$$

$$\text{False Negative} = \Pi$$

Accuracy shows how accurately the algorithm can detect the normal and attack connections:

$$A = \frac{\Theta + \omega}{\Theta + \xi + \omega + \Pi} \quad (1)$$

Specificity is used for measuring the negatives which are correctly identified:

$$S = \frac{\omega}{\xi + \omega} \quad (2)$$

Roc gives a graphical representation that compiles the review of a classifier's overall thresholds on a diagnostic criterion. That is created on mapping the True Positive Rate (TPR) against the False Positive Rate (FPR) as the use of the

threshold is different for selecting algorithms for a provided class:

$$FPR = \frac{\xi}{\xi + \omega} \quad (3)$$

$$TPR = \frac{\Theta}{\xi + \Theta} \quad (4)$$

A threshold is the expected value for all the predicted classes. The ROC curve can be drawn using binary classes. The values of the TPR and FPR range from 0 to 1.

F. Experimental Setup

The experiments were conducted on a Lenovo Thinkpad system with Ubuntu 20.04 operating system, 4500U Processor, 8GB memory, integrated AMD (attached NVIDIA) graphic card which was used for training the dataset. During data preprocessing, cleaning, and feature selection, Numpy and Pandas libraries were used.

G. Result Analysis

As mentioned above, three ML algorithms were applied to the NSL-KDD dataset, namely RF, GDBT, and SVM. From the cross-validation, RF has performed best in terms of testing and training accuracy. The results show that the RF obtained the highest accuracy on fog layer which is 85.34%. The obtained accuracy of SVM and GDBT was 32.38% and 78.01% respectively, as shown in Table III. In terms of specificity, GDBT algorithm performed best with 97.02%. The specificity achieved by SVM and RF was 2.02% and 95.09% respectively. Table III shows the result of the performance evaluation of the mentioned algorithms including *A*, *TT*, *PT*, and *S*.

TABLE III. RESULT ANALYSIS TABLE

Method	<i>A</i>	<i>S</i>	<i>TT</i>	<i>PT</i>
SVM, RF	32.38	2.02	10.87	1.056
GDBT	78.01	97.02	7.78	1.6
RF	85.34	95.09	6.10	1.345

VI. CONCLUSION

Through the obtained results, it can be confirmed that supervised ML can be used to analyze traffic data and accurately expose the data that are maliciously traveling over IoT devices. To identify that traffic accurately, NSL KDD dataset is critically evaluated by making use of ML techniques. This dataset is used for the comparison of the given framework by employing functions such as selection and classification. Overall, the RF algorithm provided the best accuracy of 85.34% on the fog layer in comparison with the other two learning algorithms. In the future, it is planned to analyze different IoT devices, explore further technologies and, testing with different data of IoT devices infected by malware and cyber-attacks.

ACKNOWLEDGEMENT

The authors would like to thank the Department of Computer Science and Information technology, NED

University of Engineer and Technology for giving permission for this research project.

REFERENCES

- [1] S. Mendhurwar and R. Mishra, "Integration of social and IoT technologies: architectural framework for digital transformation and cyber security challenges," *Enterprise Information Systems*, vol. 15, no. 4, pp. 565–584, Apr. 2021, <https://doi.org/10.1080/17517575.2019.1600041>.
- [2] Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities," *Cities*, vol. 89, pp. 80–91, Jun. 2019, <https://doi.org/10.1016/j.cities.2019.01.032>.
- [3] K. K. Mohbey, "An Efficient Framework for Smart City Using Big Data Technologies and Internet of Things," in *Progress in Advanced Computing and Intelligent Engineering*, Singapore, 2019, pp. 319–328, https://doi.org/10.1007/978-981-13-0224-4_29.
- [4] N. T. Archibald, "Cybersecurity and Critical Infrastructure: An Analysis of Securitization Theory," *Undergraduate Journal of Politics, Policy and Society*, vol. 3, no. 1, pp. 39–54, 2020.
- [5] A. Elsaedy, I. Elgendi, K. S. Munasinghe, D. Sharma, and A. Jamalipour, "A smart city cyber security platform for narrowband networks," in *27th International Telecommunication Networks and Applications Conference*, Melbourne, VIC, Australia, Nov. 2017, pp. 1–6, <https://doi.org/10.1109/ATNAC.2017.8215388>.
- [6] M. Ghobaei-Arani, A. Souri, and A. A. Rahmanian, "Resource Management Approaches in Fog Computing: a Comprehensive Review," *Journal of Grid Computing*, vol. 18, no. 1, pp. 1–42, Mar. 2020, <https://doi.org/10.1007/s10723-019-09491-1>.
- [7] H. Wang, Z. Wang, and J. Domingo-Ferrer, "Anonymous and secure aggregation scheme in fog-based public cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 712–719, Jan. 2018, <https://doi.org/10.1016/j.future.2017.02.032>.
- [8] D. Li, L. Deng, W. Liu, and Q. Su, "Improving communication precision of IoT through behavior-based learning in smart city environment," *Future Generation Computer Systems*, vol. 108, pp. 512–520, Jul. 2020, <https://doi.org/10.1016/j.future.2020.02.053>.
- [9] A. Churcher *et al.*, "An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks," *Sensors*, vol. 21, no. 2, Jan. 2021, Art. no. 446, <https://doi.org/10.3390/s21020446>.
- [10] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, May 2018, <https://doi.org/10.1016/j.future.2017.08.043>.
- [11] B. K. Mohanta, U. Satapathy, and D. Jena, "Addressing Security and Computation Challenges in IoT Using Machine Learning," in *Advances in Distributed Computing and Machine Learning*, Singapore, Asia, 2021, pp. 67–74, https://doi.org/10.1007/978-981-15-4218-3_7.
- [12] J. Li and B. Sun, "A Network Attack Detection Method Using SDA and Deep Neural Network Based on Internet of Things," *International Journal of Wireless Information Networks*, vol. 27, no. 2, pp. 209–214, Jun. 2020, <https://doi.org/10.1007/s10776-019-00462-7>.
- [13] N. Sahar, R. Mishra, and S. Kalam, "Deep Learning Approach-Based Network Intrusion Detection System for Fog-Assisted IoT," in *Proceedings of International Conference on Big Data, Machine Learning and their Applications*, Singapore, 2021, pp. 39–50, https://doi.org/10.1007/978-981-15-8377-3_4.
- [14] S. Kavitha, U. Maheswari, and R. Venkatesh, "Network Anomaly Detection for NSL-KDD Dataset Using Deep Learning," *Information Technology in Industry*, vol. 9, no. 2, pp. 821–827, Mar. 2021, <https://doi.org/10.17762/itii.v9i2.419>.
- [15] H. Neuschmied, M. Winter, K. Hofer-Schmitz, and B. Stojanovic, "Two Stage Anomaly Detection for Network Intrusion Detection," in *7th International Conference on Information Systems Security and Privacy*, Vienna, Austria, Feb. 2021, pp. 450–457.
- [16] S. Pande, A. Khamparia, D. Gupta, and D. N. H. Thanh, "DDoS Detection Using Machine Learning Technique," in *Recent Studies on Computational Intelligence: Doctoral Symposium on Computational*

- Intelligence*, A. Khanna, A. K. Singh, and A. Swaroop, Eds. Singapore, Asia: Springer, 2021, pp. 59–68.
- [17] A. T. Siahmarzkooh, J. Karimpour, and S. Lotfi, "A Cluster-based Approach Towards Detecting and Modeling Network Dictionary Attacks," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1227–1234, Dec. 2016, <https://doi.org/10.48084/etasr.937>.
- [18] A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng, "Fog Computing for the Internet of Things: Security and Privacy Issues," *IEEE Internet Computing*, vol. 21, no. 2, pp. 34–42, Mar. 2017, <https://doi.org/10.1109/MIC.2017.37>.
- [19] I. Kotenko, I. Saenko, A. Kushnerevich, and A. Branitskiy, "Attack Detection in IoT Critical Infrastructures: A Machine Learning and Big Data Processing Approach," in *27th Euromicro International Conference on Parallel, Distributed and Network-Based Processing*, Pavia, Italy, Feb. 2019, pp. 340–347, <https://doi.org/10.1109/EMDP.2019.8671571>.
- [20] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. M. A. Hashem, "Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches," *Internet of Things*, vol. 7, Sep. 2019, Art. no. 100059, <https://doi.org/10.1016/j.iot.2019.100059>.
- [21] A. Haldorai, A. Ramu, and M. Suriya, "Organization Internet of Things (IoTs): Supervised, Unsupervised, and Reinforcement Learning," in *Business Intelligence for Enterprise Internet of Things*, A. Haldorai, A. Ramu, and S. A. R. Khan, Eds. Cambridge, UK: Springer, 2020, pp. 27–53.
- [22] S. Rathore and J. H. Park, "Semi-supervised learning based distributed attack detection framework for IoT," *Applied Soft Computing*, vol. 72, pp. 79–89, Nov. 2018, <https://doi.org/10.1016/j.asoc.2018.05.049>.
- [23] M. Lopez-Martin, B. Carro, and A. Sanchez-Esguevillas, "Application of deep reinforcement learning to intrusion detection for supervised problems," *Expert Systems with Applications*, vol. 141, Mar. 2020, Art. no. 112963, <https://doi.org/10.1016/j.eswa.2019.112963>.
- [24] K. Yang, J. Ren, Y. Zhu, and W. Zhang, "Active Learning for Wireless IoT Intrusion Detection," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 19–25, Dec. 2018, <https://doi.org/10.1109/MWC.2017.1800079>.
- [25] B. S. Bhati and C. S. Rai, "Analysis of Support Vector Machine-based Intrusion Detection Techniques," *Arabian Journal for Science and Engineering*, vol. 45, no. 4, pp. 2371–2383, Apr. 2020, <https://doi.org/10.1007/s13369-019-03970-z>.
- [26] M. M. N. Aboelwafa, K. G. Seddik, M. H. Eldefrawy, Y. Gadallah, and M. Gidlund, "A Machine-Learning-Based Technique for False Data Injection Attacks Detection in Industrial IoT," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8462–8471, Sep. 2020, <https://doi.org/10.1109/JIOT.2020.2991693>.
- [27] L. Liu, J. Yang, and W. Meng, "Detecting malicious nodes via gradient descent and support vector machine in Internet of Things," *Computers & Electrical Engineering*, vol. 77, pp. 339–353, Jul. 2019, <https://doi.org/10.1016/j.compeleceng.2019.06.013>.
- [28] M. B. Farukee, M. S. Z. Shabit, Md. R. Haque, and A. H. M. S. Sattar, "DDoS Attack Detection in IoT Networks Using Deep Learning Models Combined with Random Forest as Feature Selector," in *Advances in Cyber Security*, Singapore, Asia, 2021, pp. 118–134, https://doi.org/10.1007/978-981-33-6835-4_8.
- [29] P. A. A. Resende and A. C. Drummond, "A Survey of Random Forest Based Methods for Intrusion Detection Systems," *ACM Computing Surveys*, vol. 51, no. 3, pp. 48:1–48:36, May 2018, <https://doi.org/10.1145/3178582>.
- [30] R. Kozik, M. Choras, M. Ficco, and F. Palmieri, "A scalable distributed machine learning approach for attack detection in edge computing environments," *Journal of Parallel and Distributed Computing*, vol. 119, pp. 18–26, Sep. 2018, <https://doi.org/10.1016/j.jpdc.2018.03.006>.
- [31] G. Meena and R. R. Choudhary, "A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA," in *2017 International Conference on Computer, Communications and Electronics (Comptelx)*, Jaipur, India, Jul. 2017, pp. 553–558, <https://doi.org/10.1109/COMPTelx.2017.8004032>.
- [32] J. Liu, B. Kantarci, and C. Adams, "Machine learning-driven intrusion detection for Contiki-NG-based IoT networks exposed to NSL-KDD dataset," in *2nd ACM Workshop on Wireless Security and Machine Learning*, New York, NY, USA, Jul. 2020, pp. 25–30, <https://doi.org/10.1145/3395352.3402621>.
- [33] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, "Rule Generation for Signature Based Detection Systems of Cyber Attacks in IoT Environments," *Bulletin of Networking, Computing, Systems, and Software*, vol. 8, no. 2, pp. 93–97, Jul. 2019.
- [34] M. Aamir and S. M. A. Zaidi, "DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation," *International Journal of Information Security*, vol. 18, no. 6, pp. 761–785, Dec. 2019, <https://doi.org/10.1007/s10207-019-00434-1>.
- [35] 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>.
- [36] Z. A. Shaikh, "Keyword Detection Techniques: A Comprehensive Study," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2590–2594, Feb. 2018, <https://doi.org/10.48084/etasr.1813>.
- [37] N. F. Syed, Z. Baig, A. Ibrahim, and C. Valli, "Denial of service attack detection through machine learning for the IoT," *Journal of Information and Telecommunication*, vol. 4, no. 4, pp. 482–503, Oct. 2020, <https://doi.org/10.1080/24751839.2020.1767484>.

Study and Order of a Medium Power Wind Turbine Conversion Chain Connected to Medium Voltage Grid

Abdelhafid Guediri
Faculty of Technology
VTRS Laboratory
The University of El Oued
El Oued, Algeria
abdelhafid-guediri@univ-eloued.dz

Abdelkarim Guediri
Faculty of Technology
VTRS Laboratory
The University of El Oued
El Oued, Algeria
karim_elect@yahoo.fr

Abstract—In this article, we will study a system consisting of a wind turbine operating at a variable wind speed and a two-feed asynchronous machine (DFIG) connected to the grid by the stator and fed by a transducer at the rotor side. The conductors are separately controlled for active and reactive power flow between the stator (DFIG) and the network, which is achieved using conventional PI and fuzzy logic. The proposed controllers generate reference voltages for the rotor to ensure that the active and reactive powers reach the required reference values, in order to ensure effective tracking of the optimum operating point and to obtain the maximum electrical power output. System modeling and simulation were examined with Matlab. Dynamic analysis of the system is performed under variable wind speed. This analysis is based on active and reactive energy control. The results obtained show the advantages of the proposed intelligent control unit.

Keywords—Doubly Fed Induction Generator (DFIG); active and reactive control; fuzzy management; fuzzy PI controller; variable speed wind turbine

I. INTRODUCTION

Variable speed wind generators are more attractive than fixed speed systems due to their efficiency, their quality in the power output, and their dynamic performance in the occurrence of network faults. The majority of wind diversion systems are equipped with Doubly Fed Induction Generators (DFIGs) [1]. Under ideal control conditions, at different wind speeds, the system can extract maximum energy. Wind speed measuring instruments are used and the command is derived for the required optimum speed turbine [2]. However, anemometers can increase the cost of the wind transmission system and reduce its reliability [3]. Several turbine control technologies have been developed to improve the energy output for a specific wind speed. Some variable speed techniques require a strategy for estimating wind velocity, which is very difficult under highly variable wind conditions [4]. Three types of PI regulators are offered. In this paper, a fuzzy logic controller is proposed to control the DFIG speed for the maximum operating point to track force for a wide range of wind speeds.

II. WIND TURBINE MODEL

A DFIG wind turbine consists of a variable speed turbine, a generator and a gearbox [5, 6]. The main objective of the turbine is to convert wind power to mechanical power on its shaft. The relationship between mechanical power, P_{mech} and rotor speed, w is described as [7]:

$$P_{mech} = \frac{1}{2} C_p \rho \pi R^2 V^3 \quad (1)$$

where ρ corresponds to the density of the air and R to the radius of the turbine fan respectively. The force modulus can be described because the fraction of the mechanical strength extracted from the total force to be obtained from the wind is specific to each turbine [8]. The power factor C_p is generally known as the pause ratio characteristic [9].

III. MODELING OF THE DFIG

The DFIG electrical state can be modeled using the Park Transformation as follows [10]:

$$\begin{cases} V_{sd} = R_s I_{sd} + \frac{d\phi_{sd}}{dt} \\ V_{sq} = R_s I_{sq} + \omega_s \phi_{sd} \\ V_{rd} = R_r I_{rd} + \frac{d\phi_{rd}}{dt} - (\omega_s - \omega) \phi_{rq} \\ V_{rq} = R_r I_{rq} + \frac{d\phi_{rq}}{dt} + (\omega_s - \omega) \phi_{rd} \end{cases} \quad (2)$$

To obtain separate control over the stator reactive and active forces, the DFIG model is required to express all quantities. This is in accordance with the concept of stator flow direction and the assumption that the stator resistance is small compared to the stator reactance of a DFIG of medium and high power volume, where the stator flux can be computed as [11]:

$$\begin{cases} P_s = -V_s \frac{M}{L_s} I_{rq} \\ Q_s = \frac{V_s \phi_s}{L_s} - \frac{V_s M}{L_s} I_{rd} \end{cases} \quad (3)$$

The electromagnetic torque is expressed as:

$$C_{em} = -p \frac{M}{L_s} \varphi_s I_{rq} \quad (4)$$

IV. SYSTEM DESCRIPTION

The first DFIG configuration used in this paper is shown in Figure 1. By means of a gearbox and a coupling shaft mechanism the wind turbine is mechanically connected with the DFIG [12]. The wound-rotor induction generator is fed to the stator and rotor from each element. The stator is directly connected to the grid, while the rotor is fed via lower back to another four-quadrant PWM power converter (RSC and GSC) connected with the use of a battery inside the direct current-link condenser [13].

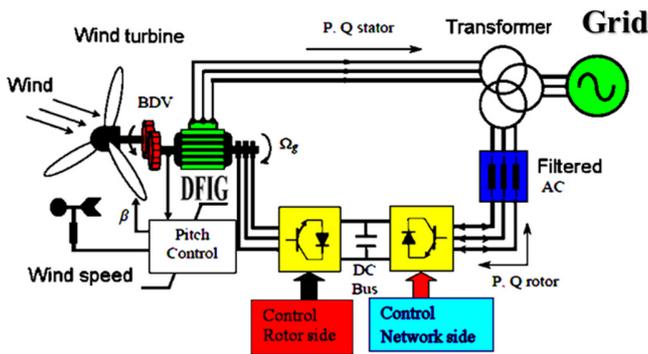


Fig. 1. DFIG-based wind conversion system.

Figure 2 displays a schematic diagram of an ordinary conversion [14]. Figure 3 shows the forces between DFIG, converters, and grid. The rotor aspect converter guarantees a decoupled lively and reactive stator power control, P_s and Q_s , consistent with the reference torque introduced via the Maximum Power Point Tracking (MPPT) control. The grid facet converter manipulates the electricity drift trade with the grid through. The rotor maintains the direct current bus at a steady voltage state and enforces the reactive energy QL at zero [15].

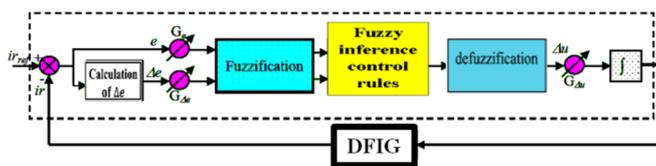


Fig. 2. Block diagram of the fuzzy controller.

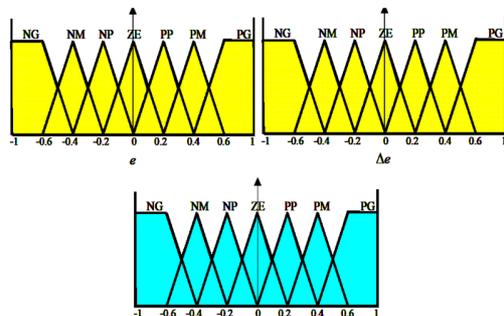


Fig. 3. Membership functions used by the controller.

TABLE I. RULE TABLE OF THE FUZZY CONTROLLER

ΔU		E						
		NG	NM	NP	EZ	PP	PM	PG
Δe	NG	NG	NG	NG	NM	NP	NP	EZ
	NM	NG	NM	NM	NM	NP	NP	EZ
	NP	NG	NM	NP	NP	EZ	PP	PM
	EZ	NG	NM	NP	EZ	PP	PM	PG
	PP	NM	NP	EZ	PP	PP	PM	PG
	PM	NP	EZ	PP	PM	PM	PM	PG
	PG	EZ	PP	PP	PM	PG	PG	PG

V. ADJUSTMENT OF THE ROTOR CURRENTS OF THE DFIG

We will now use the same diagram of the vector control except that this time the rotor current regulators are fuzzy regulators (Figure 4).

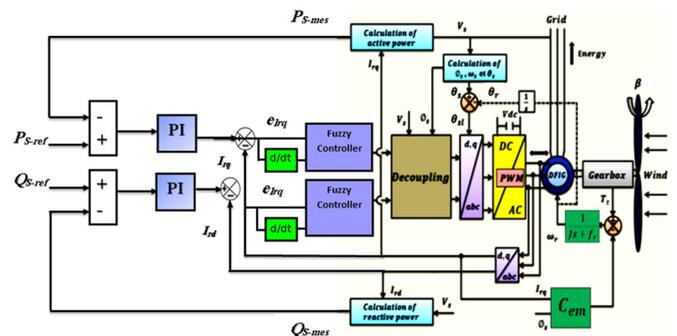


Fig. 4. Block diagram of the DFIG fuzzy-PI regulator of speed for supply of power to the electrical grid.

The two current regulators are of the same type (seven-class Mamdani-type regulator), and have the same membership functions.

VI. RESULTS AND DISCUSSION

For the robustness tests of the control by fuzzy regulators, we studied the influence of the variation of the rotor resistance and own and mutual inductance on the performance of control. Our wind power system (turbine+DFIG) is controlled by fuzzy regulators. The starting is no-load then a reference active power is applied:

- Between $t=0.2s$ and $t=0.6s$ negative scale ($P_{ref}=-20000W$).
- Between $t=0.6s$ and $t=1s$ ($P_{ref}=-10000W$).
- Reactive power: Between $t=0s$ and $t=1s$ step ($Q_{ref}=0VAR$)

Figures 5-13 show the performance of the reactive and active stator power PI- fuzzy control applied to the DFIG. In Figures 6 and 7 we observe that the direct and quadrature components of the stator and rotor currents follow their reference values. In view of these results, better tracking of the fuzzy regulator compared to that of the PI regulator is achieved. Fuzzy regulators do not generate any overshoot, particularly at transient. For the other performances, they are almost similar to that of the PI regulator. Figure 12 shows the waveform of the stator voltage and current.

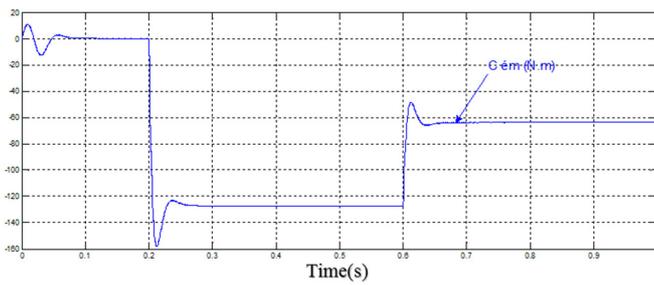


Fig. 5. Electromagnetic torque.

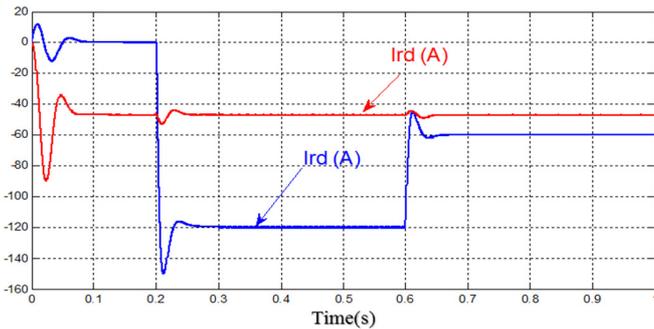


Fig. 6. Direct and quadrature rotor currents.

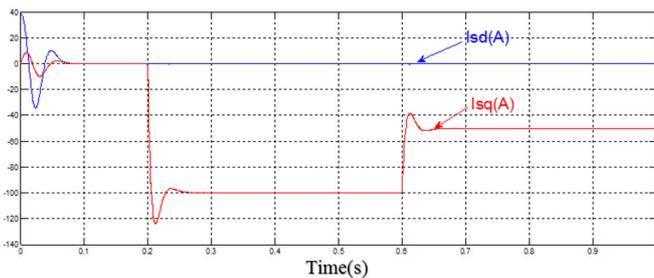


Fig. 7. Direct and quadrature stator currents.

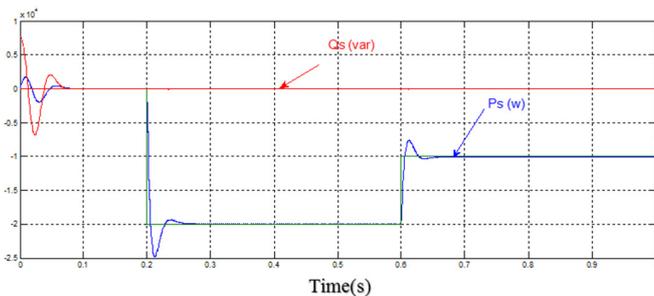


Fig. 8. Stator active power (W) and reactive power (VAR).

We can notice that the stator voltage is equal to that of the network, while the waveform of the current is related to that of the active power and the reactive power. The stator current and voltage are in phase opposition, as shown in Figure 12. This means that the stator active power is sent from the generator to the network. Note that the stator voltage and current waveforms are independent of the wind speed profile. Tuning by fuzzy logic may override tuning by PI with respect to the quality of the dynamic response of the system. Indeed, the latter further reduces the response time by producing a limited

overshoot accompanied by weak oscillations around the setpoint in steady state. The precision is not as good as that of a PI regulator where the integral action eliminates the static error. This suggests the combination of two types of regulators: a fuzzy regulator for the transient regime and a PI regulator for the steady state.

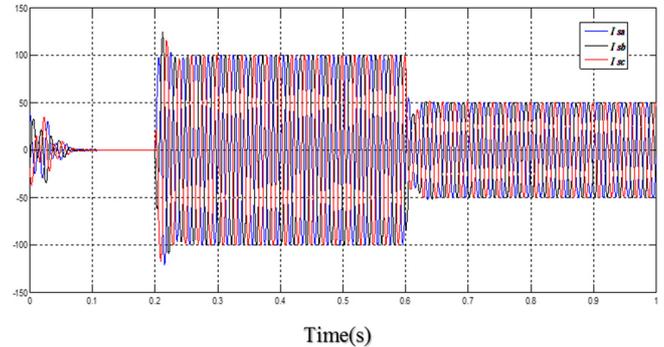


Fig. 9. Three-phase stator current.

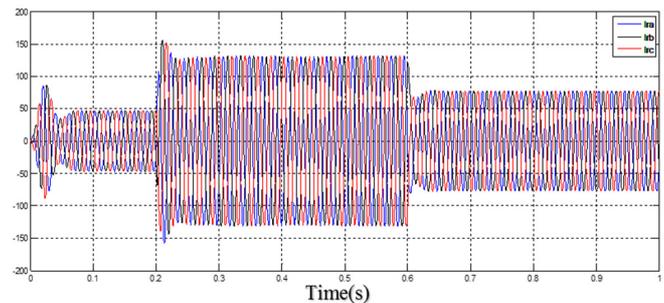


Fig. 10. Three-phase rotor current.

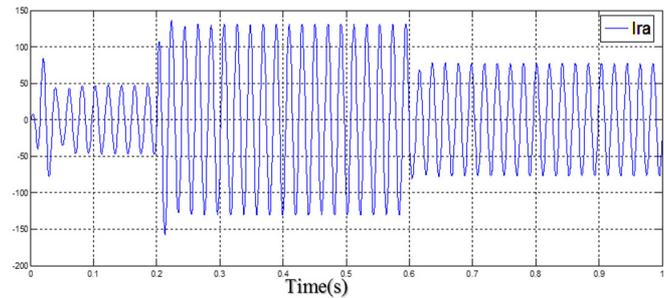


Fig. 11. Current per rotor phase.

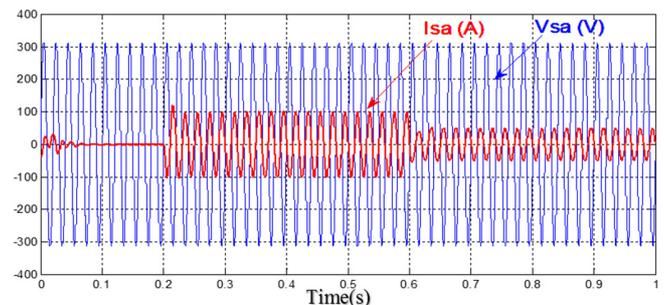


Fig. 12. The stator current and voltage.

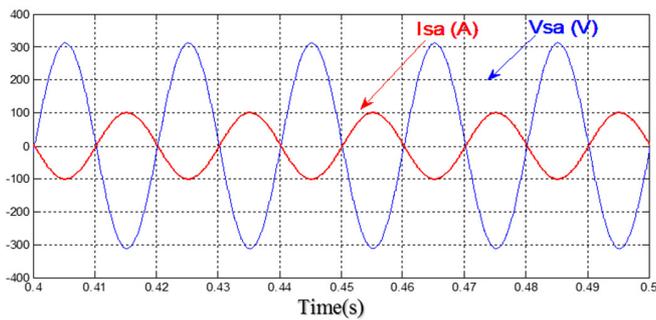


Fig. 13. Focus on stator current and voltage.

VII. CONCLUSION

In this paper, the technique of fuzzy logic has been explored. A fuzzy logic controller using the concept of offline decision table was installed in the vector control for a asynchronous dual power machine (DFIG). This choice of command was justified by the ability of fuzzy logic to handle the imprecise, the uncertain, and the vague. The obtained results show that the FLC has very satisfactory tracking performance. It has improved the dynamics of the rotor currents compared to that of the adjustment by PI. The major problem in designing an FLC is the choice of membership functions for the input and output variables, which is generally done thanks to the expertise of the process. However, a fuzzy system is difficult to grasp. Its control and adjustment can be relatively long. Sometimes it is a lot more of a trial and error procedure than a real reflection. Fuzzy logic therefore lacks a high-performance learning medium for solving a fuzzy system.

APPENDIX

PARAMETERS OF THE 1.5MW DFIG

Symbol	Parameters	Value
Pn	Rated Power	1.5MW
Vs	Stator Voltage	300V
Fs	Stator Frequency	50Hz
Rs	Stator Resistance	0.012Ω
Ls	Stator Leakage Inductance	0.0205H
Rr	Rotor Resistance	0.021Ω
Lr	Rotor Leakage Inductance	0.0204H
Lm	Mutual Inductance	0.0169H
P	Poles Pairs Number	2
J	Rotor Inertia	1000Kg.m ²

PARAMETERS OF THE TURBINE

Symbol	Parameters	Value
R	Blade Radius	35.25m
N	Number of Blades	3
G	Gearbox ratio	90
J	Moment of inertia	Kg.m ²
f _v	Viscous friction coefficient	0.0024N.m.s ⁻¹
V	Nominal Wind Speed	16m/s
V _d	Cut-in Wind Speed	4m/s
V _m	Cut-out Wind Speed	25m/s

REFERENCES

[1] A. K. Guediri and D. Ben Attous, "Fuzzy control of a doubly fed asynchronous machine (DFAM) generator driven by a wind turbine

modeling and simulation," *International Journal of System Assurance Engineering and Management*, vol. 8, no. 1, pp. 8–17, Jan. 2017, <https://doi.org/10.1007/s13198-014-0256-z>.

[2] W. Guo, F. Liu, D. He, J. Si, R. Harley, and S. Mei, "Reactive power control of DFIG wind farm using online supplementary learning controller based on approximate dynamic programming," in *2014 International Joint Conference on Neural Networks (IJCNN)*, Beijing, China, Jul. 2014, pp. 1453–1460, <https://doi.org/10.1109/IJCNN.2014.6889871>.

[3] S. A. E. M. Ardjoun and M. Abid, "Fuzzy sliding mode control applied to a doubly fed induction generator for wind turbines," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 23, no. 6, pp. 1673–1686, Nov. 2015, <https://doi.org/10.3906/elk-1404-64>.

[4] J. Yang, S. Dongran, H. Han, P. Tong, and L. Zhou, "The integrated control of fuzzy logic and model-based approach for variable-speed wind turbine," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 23, no. 6, pp. 1715–1734, Jan. 2015, <https://doi.org/10.3906/elk-1403-295>.

[5] O. P. Bharti, R. K. Saket, and S. K. Nagar, "Controller Design For DFIG Driven By Variable Speed Wind Turbine Using Static Output Feedback Technique," *Engineering, Technology & Applied Science Research*, vol. 6, no. 4, pp. 1056–1061, Aug. 2016, <https://doi.org/10.48084/etasr.697>.

[6] P. D. Chung, "Retaining of Frequency in Micro-grid with Wind Turbine and Diesel Generator," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3646–3651, Dec. 2018, <https://doi.org/10.48084/etasr.2413>.

[7] H. Bassi and Y. A. Mobarak, "State-Space Modeling and Performance Analysis of Variable-Speed Wind Turbine Based on a Model Predictive Control Approach," *Engineering, Technology & Applied Science Research*, vol. 7, no. 2, pp. 1436–1443, Apr. 2017, <https://doi.org/10.48084/etasr.1015>.

[8] "Wind in Numbers," *Global Wind Energy Council*. <https://gwec.net/global-figures/wind-in-numbers/> (accessed Jun. 01, 2021).

[9] K. Siraj, H. Siraj, and M. Nasir, "Modeling and control of a doubly fed induction generator for grid integrated wind turbine," in *2014 16th International Power Electronics and Motion Control Conference and Exposition*, Antalya, Turkey, Sep. 2014, pp. 901–906, <https://doi.org/10.1109/EPEPMC.2014.6980613>.

[10] I. Yaichi, A. Semmah, P. Wira, and Y. Djeriri, "Super-twisting Sliding Mode Control of a Doubly-fed Induction Generator Based on the SVM Strategy," *Periodica Polytechnica Electrical Engineering and Computer Science*, vol. 63, no. 3, pp. 178–190, Jun. 2019, <https://doi.org/10.3311/PPee.13726>.

[11] Z. Kara and K. Barra, "Hybrid Controller for Variable Speed Wind Energy Conversion System with Slip Energy Recovery Using Matrix Converter Topology," *Periodica Polytechnica Electrical Engineering and Computer Science*, vol. 59, no. 4, pp. 160–174, Dec. 2015, <https://doi.org/10.3311/PPee.8507>.

[12] S. Chikha, "Active and Reactive Power Management of Wind Farm Based on a Six Leg Tow Stage Matrix Converter Controlled by a Predictive Direct Power Controller," *Iranian Journal of Electrical and Electronic Engineering*, vol. 14, no. 3, pp. 245–258, Sep. 2018, <https://doi.org/10.22068/IJEEE.14.3.245>.

[13] S. Massoum, A. Meroufel, A. Massoum, and P. Wira, "A direct power control of the doubly-fed induction generator based on the SVM Strategy," *Elektrotehniski Vestnik/Electrotechnical Review*, vol. 84, no. 5, pp. 235–240, Jan. 2017.

[14] H. Benbouhenni, A. Belaidi, and Z. Boudjema, "Sensorless Twelve Sectors Implementation of Neural DPC Controlled DFIG for Reactive and Active Powers Ripples Reduction," *Majlesi Journal of Energy Management*, vol. 7, no. 2, pp. 13–21, Jun. 2018.

[15] H. Benbouhenni, Z. Boudjema, and A. Belaidi, "A Direct Power Control of the Doubly Fed Induction Generator Based on the Three-Level NSVPWM Technique," *International Journal of Smart Grid - ijSmartGrid*, vol. 3, no. 4, pp. 216–225, Dec. 2019.

Local Search-based Non-dominated Sorting Genetic Algorithm for Optimal Design of Multimachine Power System Stabilizers

Fahad A. Alshammari

Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
fahad.a.s.alshammari@gmail.com

Gharbi A. Alshammari

Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
gharbi.a.alshammari@gmail.com

Tawfik Guesmi

Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
tawfik.guesmi@istmt.mu.tn

Ahmed A. Alzamil

Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
aa.alzamil@uoh.edu.sa

Badr M. Alshammari

Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
bms.alshammari@uoh.edu.sa

Ahmed S. Alshammari

Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
ahm.alshammari@uoh.edu.sa

Abstract-This study presents a metaheuristic method for the optimum design of multimachine Power System Stabilizers (PSSs). In the proposed method, referred to as Local Search-based Non-dominated Sorting Genetic Algorithm (LSNSGA), a local search mechanism is incorporated at the end of the second version of the non-dominated sorting genetic algorithm in order to improve its convergence rate and avoid the convergence to local optima. The parameters of PSSs are tuned using LSNSGA over a wide range of operating conditions, in order to provide the best damping of critical electromechanical oscillations. Eigenvalue-based objective functions are employed in the PSS design process. Simulation results based on eigenvalue analysis and nonlinear time-domain simulation proved that the proposed controller provided competitive results compared to other metaheuristic techniques.

Keywords-power system stabilizer; non-dominated sorting genetic algorithm; local search; eigenvalue analysis; nonlinear time domain simulation

I. INTRODUCTION

The complexity of electricity networks due to several economic, ecological, and technical exigencies has obliged electric companies to operate at full network capacity in order to achieve a balance between the increased consumption and the production, under severe conditions increasingly close to the stability limits. Under these drastic conditions and operating limits, the occurrence of any contingencies or disturbances such as short-circuits, sudden variations in loads, and line outage, can lead to a critical situation often starting with poorly damped Electromechanical Oscillations (EMOs) followed by loss of synchronism and system instability [1]. For instance, these low frequency oscillations may limit the transfer capacity of the power system and continue to grow up resulting

in the separation of the system if no adequate response is quickly taken. To overcome the problem of EMOs and improve the system damping, additional stabilizing signals are usually added to the excitation system via the voltage regulator [2]. The Conventional Power System Stabilizer (CPSS) has long been employed to damp out EMOs. Generally, CPSSs are based on lead-lag compensators with fixed parameters determined at the nominal operating condition. However, power systems are strongly nonlinear with configurations and load changing over time, which implies that these fixed parameters of the stabilizers are no longer adapted to the new operating conditions [3]. Therefore, the fundamental problem is not only to determine the optimal parameters of these stabilizers, but also to make them adapt to the modification of the operating points and system configurations. Within this context, diverse research works have been directed towards the design of Power System Stabilizers (PSSs) with optimal performance for a wide range of system parameters, configurations, and operating conditions [4-8]. From the literature review, it was found that several ideas and methods have been suggested for the optimal setting of PSS parameters. The most used PSS design methods are summarized in [9-10] and are divided into three main categories which are adaptive control [11-12], linear approximation [13], and nonlinear models [14]. In [12], Model Reference Adaptive System-based PSS (MRAS-PSS) design has been addressed. The performance of the MRAS-PSS has been assessed through Nonlinear Time Domain Simulation (NTDS). Despite the fact that the adaptive control-based PSS may mitigate the shortcomings of CPSS, it appears complex in nature and costly.

Several numerical techniques have been suggested for the enhancement of small signal stability of interconnected multimachine power systems, such as linear programming and

gradient methods [15]. Unfortunately, these classical methods require initialization of the problem solutions and they are iterative. Thus, they may fail to converge to the global optima. To avoid the limitations of classical techniques, nonconventional optimization techniques have been used for solving several complex problems. In particular, these techniques have demonstrated high performance when applied to power system problems, such as the problem of the enhancement of power system stability. For instance, a Genetic Algorithm (GA)-based method for the optimal setting of lead-lag PSSs parameters was suggested in [16]. In [4], the same regulators have been optimally designed using Simulated Annealing (SA) where the objective function has been optimized in order to shift all electromechanical modes to the left side in the s-plane. An artificial bee colony-based method has been suggested in [6] for dynamic stability enhancement and its performance was compared with other techniques. In [17], a time domain response based function has been minimized using Particle Swarm Optimization (PSO) in order to tune the parameters of PSS regulators. These regulators are employed for damping inter-area oscillations and local modes. In [18], the whale optimization algorithm was used for tuning the PSS regulators in order to suppress power system oscillations and maintain system stability after the occurrence of faults. The authors have considered an eigenvalue-based objective function in the design process. Other metaheuristic techniques have also been employed for the enhancement of power system stability, such as the Bacteria Foraging Optimization Algorithm (BFOA) [19] and Fuzzy Gravitational Search Algorithm (FGSA) [20]. Unfortunately, these random-based methods have been criticized for their low convergence rate and the fact that they can be trapped in local minima when complex multimodal problems are considered [21]. Within this context, a modified version of the Non-dominated Sorting Genetic Algorithm (NSGAI) for the optimal design of PSS regulators is presented in this study. The main contributions of this work are:

- Two eigenvalue-based objective functions are optimized simultaneously in order to provide optimum PSS design. The first one is related to the real part of the electromechanical modes whilst the second one corresponds to the damping ratios of the same modes. The optimization of these two objective functions aims to shift all electromechanical modes as much as possible in the left side of the s-plane. In order to make the proposed controller more robust, the PSS parameters are optimized over a wide range of operating conditions.
- The aforementioned objective functions are minimized simultaneously using an improved version of the NSGAI. To do this, a local search procedure is embedded at the end of all iterations of the NSGAI in order to increase its convergence rate and avoid the convergence to local optima. Decision variables of the problem are the PSS parameters and the problem constraints are the bounds of these parameters.
- The simulation results based on eigenvalue analysis and nonlinear time domain simulation demonstrated that the proposed controller provided results competitive with the

other metaheuristic techniques implemented recently for the resolution of the PSS design problem, such as NSGAI, SA [4] and Fuzzy Gravitational Search Algorithm (FGSA) [20].

II. PROBLEM FORMULATION OF PSS DESIGN

A. Power Network Model

For stability studies, power network is mostly modeled by a set of nonlinear Differential-Algebraic Equations (DAE) as given in (1)-(3):

$$\dot{X} = f(X, Y, U) \quad (1)$$

$$0 = g(X, Y) \quad (2)$$

$$W = h(X, Y, U) \quad (3)$$

The state variables vector and the algebraic variables vector are represented by X and Y respectively. U and W express the input variables vector and the output variables set respectively. The equations of network power flow are represented by a nonlinear algebraic set defined with the symbol g . The dynamics of the system and controller are expressed by the first-order nonlinear differential equations described by the function f . The output variables are described by function h .

In this study, the state vector is $X = [\delta \ \omega \ E'_q \ E_{fd}]^T$, where δ is the rotor angle, and ω is the rotor speed. E'_q is the internal voltage and E_{fd} is the field voltage. The bus voltage magnitudes and phase angles constitute the vector Y . The PSS output signals constitute the control vector U . The PSS design is based on a linearized incremental model of the power system. At an equilibrium point of the power system, the state equations of the system can easily be written if Y is removed and it is assumed that the power flow Jacobian is non-singular, as follows.

$$\Delta \dot{X} = AX + BU \quad (4)$$

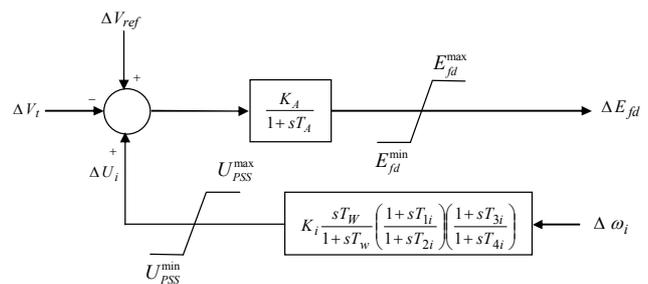


Fig. 1. IEEE type-ST1 excitation system with PSS.

It is worth noting that if the power network is with n machines and m PSSs, the state matrix A is a $4n \times 4n$ matrix and the control matrix B is a $4n \times m$ matrix.

B. Excitation System with PSS Structure

The PSS gives a control effect, by acting via the exciter, to the power system under consideration. This study considers the IEEE type-ST1 excitation system with PSS [6] as shown in Figure 1. V_t is the generator terminal voltage and V_{ref} is the

reference terminal voltage. T_A is the amplifier time constant, K_A is the amplifier gain constant, U_i is the output that is the supplementary stabilizing signal, and $\Delta\omega_i$ is the input signal of the i^{th} PSS that is the normalized speed deviation. The PSS's transfer function is shown in the following expression:

$$U_i(s) = K_i \frac{sT_W}{1+sT_W} \left[\frac{(1+sT_1)(1+sT_3)}{1+(sT_2)(1+sT_4)} \right] \quad (5)$$

In (5), the washout block with time constant T_{Wi} is utilized as a high-pass filter, allowing signals in the range of 0.2-2Hz associated with rotor oscillation to pass unchanged, and is generally between 1 and 20s [4]. Compensating for the phase lag between the PSS output and the control operation, which is the electrical torque, is conducted by the two first-order lead-lag transfer functions. Thus, the representation of PSS comprises of two lead-lag blocks, a gain K_i , and a washout bloc with time constant T_{Wi} .

C. Damping Controller Design

The system closed-loop eigenvalues are measured after linearizing the power system around the operating point. Using only the unstable or lightly damped electromechanical forms that need to be moved, the objective functions can be formulated. The issue of the parameter tuning of the PSS controllers that stabilize the system is transformed into a multiobjective minimization problem in this study. One of the considered two eigenvalue-based objective functions stated in [6] aims to transfer the closed-loop eigenvalues to the left side of the line expressed by $\sigma_{ij} = \sigma_0$, as shown in Figure 2(a). J_1 in (12) represents this function and the second objective function is defined by J_2 . In fact, and as presented in Figure 2(b), minimizing J_2 equals to raising the damping ratios of all electromechanical modes and place the closed-loop eigenvalues in a D-shape sector corresponding to $\xi_{ij} \geq \xi_0$.

$$\begin{cases} \text{if } \max\{\sigma_{ij}\} \leq \sigma_0 \text{ and } \min\{\xi_{ij}\} \geq \xi_0, \\ \quad J_1 = \max\{\sigma_{ij}\} \\ \quad J_2 = -\min\{\xi_{ij}\} \\ \text{else} \\ \quad J_1 = J_{1max} \\ \quad J_2 = J_{2max} \end{cases} \quad (6)$$

where the real part and damping ratio of the i^{th} electromechanical modes corresponding to the j^{th} operating point are expressed by σ_{ij} and ξ_{ij} respectively. The fitness functions J_1 and J_2 are equal to their upper limits J_{1max} and J_{2max} if one or more electromechanical modes are outside the D-shape sector shown in Figure 2(c). It is worth noting that functions J_1 and J_2 have to be minimized subject to the following inequality constraints that describe the bounds of the adjustable parameters of the regulators:

$$K_i^{\min} \leq K_i \leq K_i^{\max} \quad (7)$$

$$T_{1i}^{\min} \leq T_{1i} \leq T_{1i}^{\max} \quad (8)$$

$$T_{2i}^{\min} \leq T_{2i} \leq T_{2i}^{\max} \quad (9)$$

$$T_{3i}^{\min} \leq T_{3i} \leq T_{3i}^{\max} \quad (10)$$

$$T_{4i}^{\min} \leq T_{4i} \leq T_{4i}^{\max} \quad (11)$$

The washout time constant T_w is fixed to 5s in this paper, and typical ranges of the decision variables are [0.01–1.5] for T_{1i} to T_{4i} and [0.1–50] for K_i .

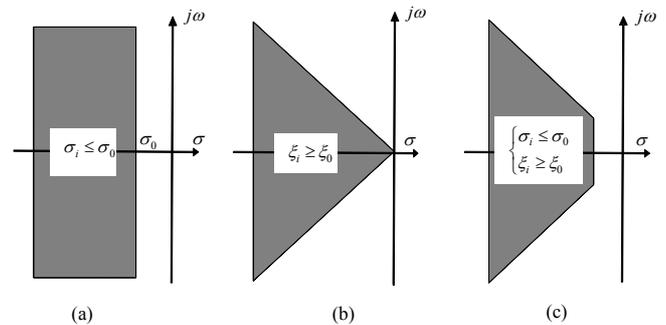


Fig. 2. Electromechanical modes location for different objective functions.

III. DESCRIPTION OF THE PROPOSED OPTIMIZATION TECHNIQUE

Due to its high computational complexity and absence of elitism, the first version of NSGA has been criticized frequently in the literature. To overcome these drawbacks, authors in [22] proposed a new version of NSGA, called NSGAI. The NSGAI algorithm comprises mainly of two parts, which are the non-dominated sorting of solutions and the preservation of the population diversity. In NSGAI, offspring population Q_t is combined with its parent population P_t at each iteration t . The combined population is sorted based on non-dominated sorting mechanism into fronts. Then, solutions of one front are sorted using their crowding distances.

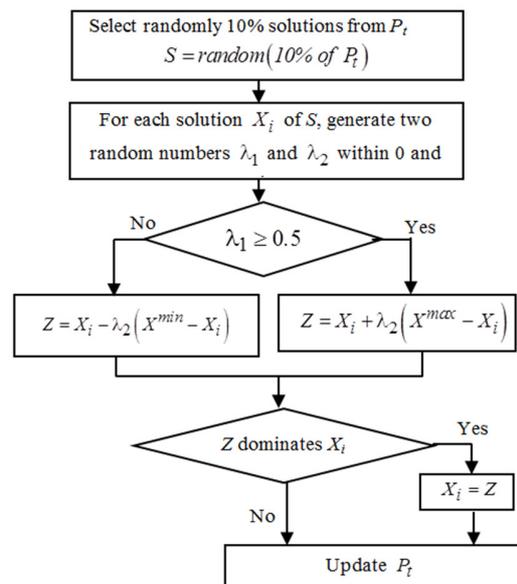


Fig. 3. Flowchart of the local search algorithm.

Despite the modifications applied in the NSGAIL, it suffers from its low convergence rate due to its random parameters. In addition, NSGAIL may fail to converge to the global optima. In order to overcome these limitations, in this study, a local search mechanism is included at the end of NSGAIL iterations. This mechanism explores the less-crowded zone in the current archive in order to obtain more non-dominated solutions nearby. The flowchart of the local search algorithm applied for an iteration k is shown in Figure 3.

IV. SIMULATION AND DISCUSSION

In this section, the robustness and effectiveness of the LSNSGA technique, proposed for the determination of the optimal PSS parameters, is evaluated on the 3-machine 9-bus WSCC (Western System Coordinating Council). As shown in Figure 4, this system comprises of 9 buses and 3 generators. All system data are extracted from [6].

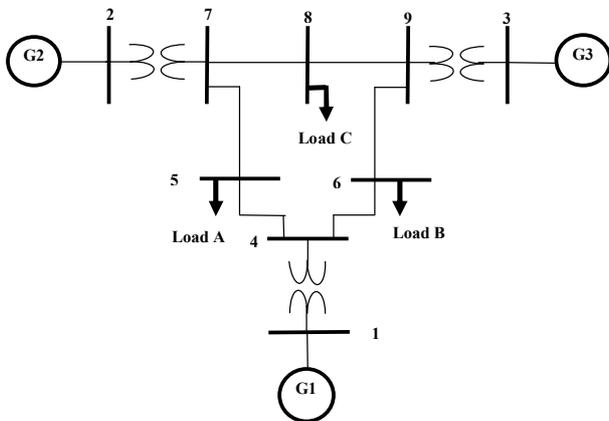


Fig. 4. Single line diagram of the WSCC system.

For economic reasons, only machines 2 and 3 are equipped with PSSs. In order to increase the robustness of the proposed LSNSGA-based controller (LSNSGA-PSS), 4 different operating conditions were used. These conditions are shown in Table I. All experiments are implemented on Matlab environment using a machine with 8GB RAM, Intel core i7 1.8GHz, and Windows 7. Maximum number of iterations and population size are selected to be 100 and 200 respectively. The mutation and crossover probabilities are 0.2 and 0.9.

TABLE I. SYSTEM OPERATING CONDITIONS (IN PU)

		Base case		Case 1		Case 2		Case 3	
		P	Q	P	Q	P	Q	P	Q
Gen	G1	0.72	0.27	2.21	1.09	0.36	0.16	0.33	1.12
	G2	1.63	0.07	1.92	0.56	0.8	-0.11	2.00	0.57
	G3	0.85	-0.11	1.28	0.36	0.45	-0.20	1.50	0.38
Load	A	1.25	0.50	2.00	0.80	0.65	0.55	1.50	0.90
	B	0.90	0.30	1.80	0.60	0.45	0.35	1.20	0.80
	C	1.00	0.35	1.50	0.60	0.50	0.25	1.00	0.50

A. Implementation of the LSNSGA for Optimum PSS Design

To investigate the performance of the proposed method, the results obtained using LSNSGA-PSS are compared with those

of other metaheuristic techniques such as NSGAIL, SA [4], and FGSA [20]. The convergence characteristic of the proposed optimization technique for the optimum tuning of PSS parameters is illustrated in Figure 5. Optimum PSSs parameters for the proposed method and for the SA and FSGA techniques are given in Table II. System eigenvalues and damping ratios corresponding to the optimal PSS parameters, obtained using these techniques are given in Table III.

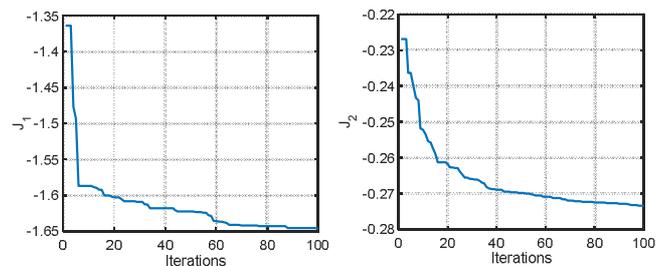


Fig. 5. Convergence characteristics of the LSNSGA.

TABLE II. OPTIMAL PSS PARAMETERS

Method	Gen	K	T ₁	T ₂	T ₃	T ₄
LSNSGAIL-PSS based on J ₁	G2	2.4238	0.8075	0.1611	1.0973	0.7417
	G3	9.9970	0.9867	1.2183	0.2652	0.4252
LSNSGAIL-PSS based on J ₂	G2	2.4529	0.8034	0.1642	1.1028	0.7448
	G3	9.9996	0.9840	1.2069	0.2252	0.4313
NSGAIL-PSS	G2	2.4530	0.8035	0.1642	1.1027	0.7447
	G3	9.9995	0.9834	1.2069	0.2267	0.4327
SA-PSS	G2	11.008	0.216	0.05	0.104	0.05
	G3	0.319	0.410	0.05	0.233	0.05
FGSA-PSS	G2	30.911	0.222	0.012	0.34	0.054
	G3	30.931	0.229	0.015	0.24	0.034

From Table III, it is obvious that the system is poorly damped when no controller is used for the base case, case 1, and case 2 and it is unstable for case 3. In addition, it can be seen that the proposed controller provides better damping of all electromechanical modes than the other controllers. It is worth noting that all electromechanical controllers obtained using LSNSGA-PSS are shifted in D-shape zone defined by $\sigma_0 = -1$ and $\xi_0 = 20\%$.

B. Nonlinear Time Domain Simulation

To assess the effectiveness and robustness of the proposed controller, a 6-cycle fault disturbance in the line 5-7, close to bus 5 is applied. The fault is cleared by tripping the line 5-7 with successful reclosure after 1.0s. Nonlinear simulation results obtained using LSNSGA are compared with the results from SA-PSS, FGSA-PSS, and without controller. System responses at the operating conditions specified above are shown in Figures 6 and 7. Figure 6 depicts the speed deviations in per unit and Figure 7 illustrates the internal voltages. From Figures 6 and 7, it can be clearly seen that the suggested LSNSGA-PSS controller improved greatly the system stability and it provided better damping of the electromechanical oscillations than the other techniques, at all operating conditions.

TABLE III. EIGENVALUES AND DAMPING RATIOS OF THE ELECTROMECHANICAL MODES

Method	Base case	Case 1	Case 2	Case 3
Without PSS	-0.1124±j7.7400, 0.0145	-0.0374±j7.8347, 0.0048	-0.2142±j6.3226, 0.0339	+0.0181±j8.0903, -0.0022
	-1.3346±j9.1096, 0.1450	-0.7023±j10.5832, 0.0662	-0.8227±j6.9390, 0.1177	-0.4515±j11.3794, 0.0396
LSNSGAIL-PSS based on J_1	-2.6031±j6.2412, 0.3849	-1.7361±j6.3024, 0.2656	-1.6459±j5.2350, 0.2999	-1.9885±j6.4392, 0.2951
	-2.9611±j6.8626, 0.3962	-4.3102±j8.9358, 0.4345	-2.0874±j6.3268, 0.3133	-4.5721±j9.4259, 0.4364
LSNSGAIL-PSS based on J_2	-2.6448±j6.1179, 0.3968	-1.7774±j6.3792, 0.2684	-1.6318±j5.2411, 0.2973	-2.0258±j6.5302, 0.2963
	-2.7138±j7.3837, 0.3450	-3.9494±j9.3519, 0.3890	-1.9142±j6.5763, 0.2795	-4.2547±j9.9153, 0.3943
LSNSGAIL-PSS (best compromise solution)	-2.5491±j5.9300, 0.3949	-1.7847±j6.2809, 0.2733	-1.6283±j5.1460, 0.3017	-2.0326±j6.4097, 0.3023
	-2.7394±j7.5113, 0.3426	-3.8660±j9.3531, 0.3820	-1.8925±j6.6077, 0.2753	-4.1434±j9.9371, 0.3849
SA-PSS	-1.6530±j5.1835, 0.3038	-1.3442±j5.6104, 0.2330	-1.3510±j4.5402, 0.2852	-1.4203±j5.6820, 0.2425
	-0.7082±j7.6085, 0.0927	-0.6774±j8.6128, 0.0784	-0.4869±j6.4701, 0.0750	-0.6073±j9.1726, 0.0661
FGSA-PSS	-1.0692±j1.9971, 0.4719	-0.9256±j2.4987, 0.3473	-1.0438±j1.8555, 0.4903	-0.9196±j2.5427, 0.3401
	-0.2328±j4.1174, 0.0564	-0.1382±j4.5924, 0.0301	-0.3301±j3.6252, 0.0907	-0.1696±j4.4911, 0.0377

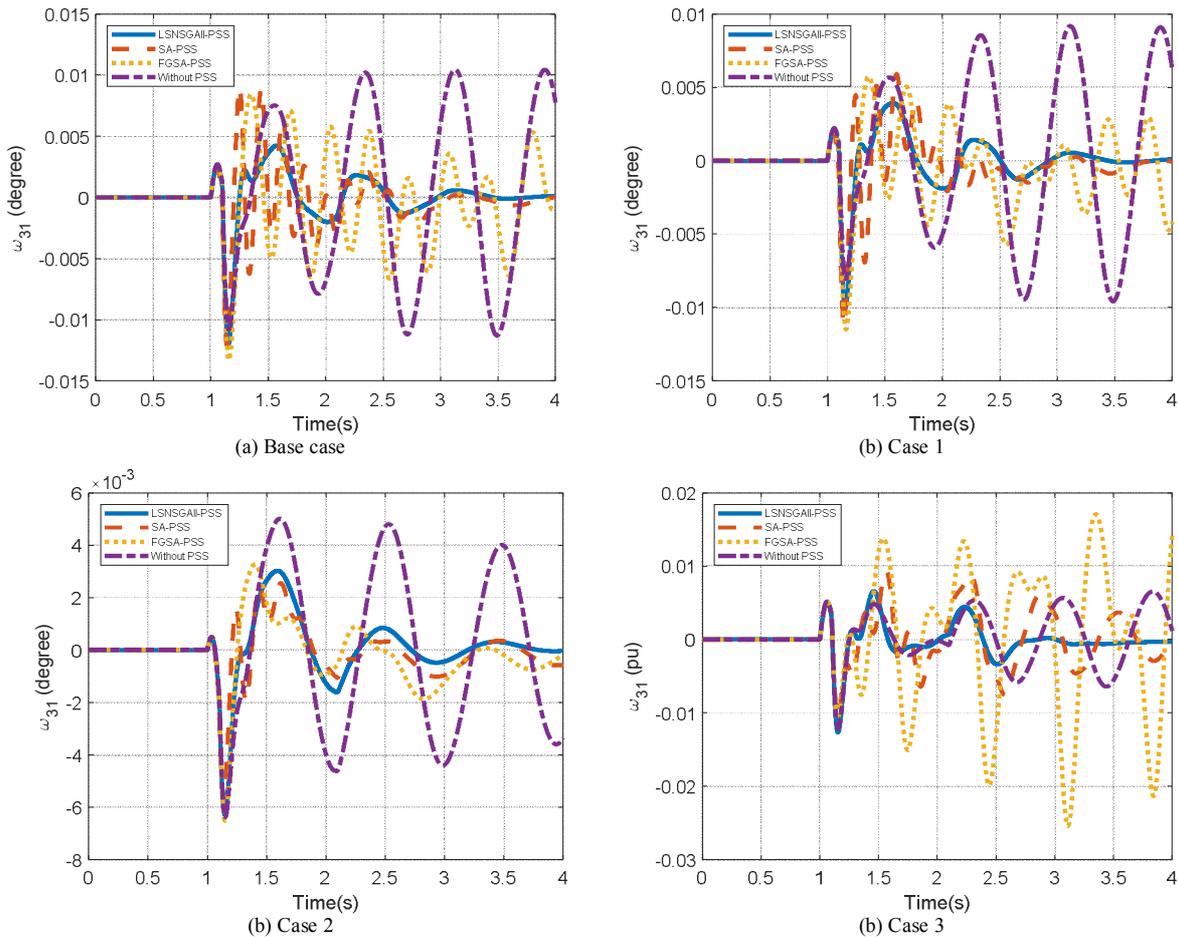


Fig. 6. System responses for speed deviations in pu.

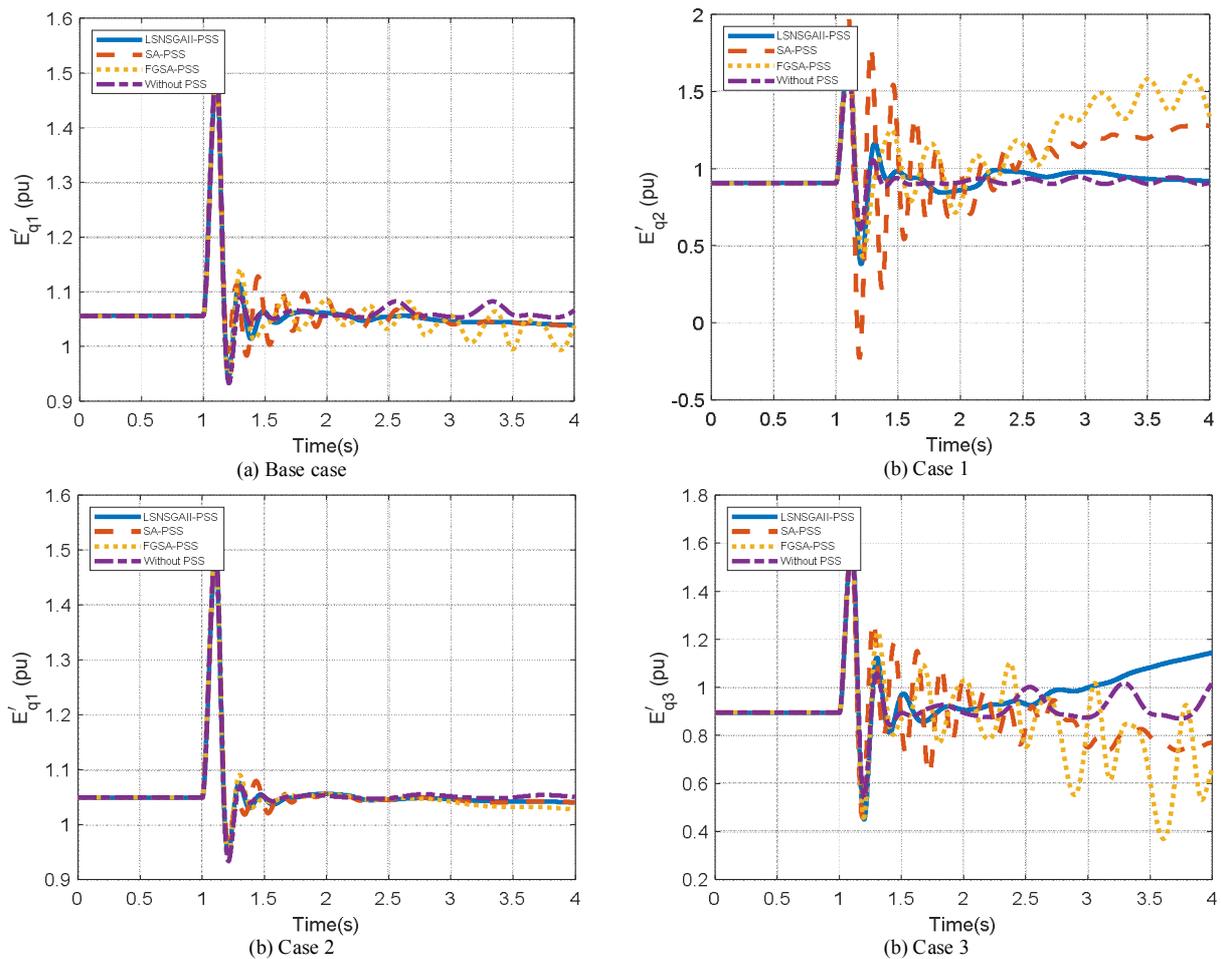


Fig. 7. System responses for internal voltages in pu.

V. CONCLUSION

PSSs represent an effective means for damping electromechanical modes. However, current power networks are becoming strongly nonlinear with configurations and load changing over time. Thus, it is mandatory to adjust the PSS parameters in order to guarantee the optimal damping of system oscillations at any operating condition, system configuration, or disturbance. In order to do this, this study presents an improved version of NSGAI, referred to as LSNSGA, for robust PSS design over a wide range of operating conditions. In the proposed technique, a local search procedure is added to the original NSGAI in order to improve its convergence characteristics. The LSNSGA is used to tune the PSSs parameters in a way that the system stability is optimally improved after the occurrence of fault. In the design process, two eigenvalue-based objective functions are considered. The robustness and performance of the proposed controllers (LSNSGA-PSSs) are tested on a 3-machine 9-bus system. The comparison with other metaheuristic techniques showed that LSNSGA-PSSs controllers provide the best results.

REFERENCES

- [1] A. Farah, T. Guesmi, H. Hadj Abdallah, and A. Ouali, "A novel chaotic teaching-learning-based optimization algorithm for multi-machine power system stabilizers design problem," *International Journal of Electrical Power & Energy Systems*, vol. 77, pp. 197–209, May 2016, <https://doi.org/10.1016/j.ijepes.2015.11.050>.
- [2] K. Tang and G. K. Venayagamoorthy, "Damping inter-area oscillations using virtual generator based power system stabilizer," *Electric Power Systems Research*, vol. 129, pp. 126–141, Dec. 2015, <https://doi.org/10.1016/j.epsr.2015.08.004>.
- [3] Y. Welhazi, T. Guesmi, and H. H. Abdallah, "Coordinated Tuning of SVC and PSSs in Power System using Teaching Learning Based Algorithm," in *2019 10th International Renewable Energy Congress (IREC)*, Sousse, Tunisia, Mar. 2019, pp. 1–6, <https://doi.org/10.1109/IREC.2019.8754644>.
- [4] M. A. Abido, "Robust design of multimachine power system stabilizers using simulated annealing," *IEEE Transactions on Energy Conversion*, vol. 15, no. 3, pp. 297–304, Sep. 2000, <https://doi.org/10.1109/60.875496>.
- [5] E. Nechadi, "Adaptive Fuzzy Type-2 Synergetic Control Based on Bat Optimization for Multi-Machine Power System Stabilizers," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4673–4678, Oct. 2019, <https://doi.org/10.48084/etasr.2970>.
- [6] T. Guesmi and B. M. Alshammari, "An improved artificial bee colony algorithm for robust design of power system stabilizers," *Engineering Computations*, vol. 34, no. 7, pp. 2131–2153, Jan. 2017, <https://doi.org/10.1108/EC-12-2016-0459>.

- [7] D. K. Sambariya, "Power System Stabilizer Design Using Compressed Rule Base of Fuzzy Logic Controller," *Journal of Electrical and Electronic Engineering*, vol. 3, no. 3, pp. 52–64, Jul. 2015, <https://doi.org/10.11648/j.jeece.20150303.16>.
- [8] 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>.
- [9] A. Sabo, N. I. A. Wahab, M. L. Othman, M. Z. A. M. Jaffar, and H. Beiranvand, "Optimal design of power system stabilizer for multimachine power system using farmland fertility algorithm," *International Transactions on Electrical Energy Systems*, vol. 30, no. 12, 2020, Art. no. e12657, <https://doi.org/10.1002/2050-7038.12657>.
- [10] X. Li, Z. Wang, J. Xu, and B. Chen, "Power System Stabilizer Parameters Designing Based on Genetic Simulated Annealing Algorithm," *Journal of Clean Energy Technologies*, vol. 4, no. 3, pp. 178–182, Jan. 2015, <https://doi.org/10.7763/JOCET.2016.V4.275>.
- [11] S. Zhang and F. L. Luo, "An Improved Simple Adaptive Control Applied to Power System Stabilizer," *IEEE Transactions on Power Electronics*, vol. 24, no. 2, pp. 369–375, Feb. 2009, <https://doi.org/10.1109/TPEL.2008.2007490>.
- [12] R. Hemmati, "Power system stabilizer design based on optimal model reference adaptive system," *Ain Shams Engineering Journal*, vol. 9, no. 2, pp. 311–318, Jun. 2018, <https://doi.org/10.1016/j.asej.2016.03.002>.
- [13] D. K. Sambariya and R. Prasad, "Design of Optimal Proportional Integral Derivative Based Power System Stabilizer Using Bat Algorithm," *Applied Computational Intelligence and Soft Computing*, vol. 2016, Mar. 2016, Art. no. e8546108, <https://doi.org/10.1155/2016/8546108>.
- [14] T. Guesmi, B. M. Alshammari, Y. Almalaq, A. Alateeq, and K. Alqunun, "New Coordinated Tuning of SVC and PSSs in Multimachine Power System Using Coyote Optimization Algorithm," *Sustainability*, vol. 13, no. 6, Jan. 2021, Art. no. 3131, <https://doi.org/10.3390/su13063131>.
- [15] S. M. Abd-Elazim and E. S. Ali, "A hybrid Particle Swarm Optimization and Bacterial Foraging for optimal Power System Stabilizers design," *International Journal of Electrical Power & Energy Systems*, vol. 46, pp. 334–341, Mar. 2013, <https://doi.org/10.1016/j.ijepes.2012.10.047>.
- [16] Y. L. Abdel-Magid and M. A. Abido, "Optimal multiobjective design of robust power system stabilizers using genetic algorithms," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1125–1132, Aug. 2003, <https://doi.org/10.1109/TPWRS.2003.814848>.
- [17] D. Butti, M. Siva Kumar, and R. Srinivasa Rao, "Design of robust modified power system stabilizer for dynamic stability improvement using Particle Swarm Optimization technique," *Ain Shams Engineering Journal*, vol. 10, no. 4, pp. 769–783, Dec. 2019, <https://doi.org/10.1016/j.asej.2019.07.002>.
- [18] D. Butti, M. Sivakumar, and R. Srinivasarao, "Interconnected multimachine power system stabilizer design using whale optimization algorithm," *Protection and Control of Modern Power Systems*, vol. 4, no. 1, Feb. 2019, Art. no. 2, <https://doi.org/10.1186/s41601-019-0116-6>.
- [19] S. M. Abd-Elazim and E. S. Ali, "Power System Stability Enhancement via Bacteria Foraging Optimization Algorithm," *Arabian Journal for Science and Engineering*, vol. 38, no. 3, pp. 599–611, Mar. 2013, <https://doi.org/10.1007/s13369-012-0423-y>.
- [20] A. Ghasemi, H. Shayeghi, and H. Alkhatib, "Robust design of multimachine power system stabilizers using fuzzy gravitational search algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 51, pp. 190–200, Oct. 2013, <https://doi.org/10.1016/j.ijepes.2013.02.022>.
- [21] I. Marouani, A. Boudjemline, T. Guesmi, and H. H. Abdallah, "A Modified Artificial Bee Colony for the Non-Smooth Dynamic Economic/Environmental Dispatch," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3321–3328, Oct. 2018, <https://doi.org/10.48084/etasr.2098>.
- [22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002, <https://doi.org/10.1109/4235.996017>.

Erratum and Addendum: “Contractor’s Attitude towards Risk and Risk Management in Construction in Two Western Provinces of Vietnam”

Van Tien Phan
Department of Construction
Vinh City, Viet Nam
vantienkxd@vinhuni.edu.vn

Since the publication of our original paper in vol. 10, no. 6 of this journal [1], the author has detected several misprints and errors that are corrected here, in the order in which they appear in the original paper.

A new reference ([2]) should be added in the reference list of the original paper as reference number 2. The last sentence of the abstract should be rewritten as follows: “The importance of applying an effective risk management has been investigated, which is shared between the planning and production phase, whereas risk identification is the most important in the risk management process.” Similarly, the conclusions of [1] should be rewritten as above. In part II: the first sentence should be modified as “The questionnaire investigation method applied in [2] to study the risk management in construction projects of Swedish contractor, is also applied in this paper.” The next sentence: “Therefore, this form of survey integrates two types of data and the core assumption of this approach is that a combination of qualitative and quantitative methods leads to a better understanding of the problem [10]” should be removed.

The last paragraph of part II should include a citation of [2] in the first part and update the research data as following: “An invitation email was sent to 215 contractors in two western provinces of Vietnam, which received 120 responses. The response rate is about 55%.” The last part of the paragraph should be removed as redundant. The text to be removed is “About 70% of the respondents had more than 15 years of experience within the construction industry and the majority (88%) were contractors, 24% were developers and 2,38% were consultants. The size of the companies were equally represented, approximately 48% had more than 1000 employees while 52% had less than 1000 employees.”

In part III, the following part should also be removed as redundant: “Contractors were the 78%, developers (clients) the 19%, and consultants the 13% of the participants. Interviews were conducted only with contractors. An equal distribution among company sizes was attained in the data collection, as stated above. A difference of opinion related to the company size will merely be mentioned when a significant differentiation can be observed between them, otherwise an

overall picture of the industry will be presented due to similar answers to the questions.”

The following sentence in part III should also be removed: “The overwhelming majority of the respondents in both the questionnaire and the interviews described themselves as being risk-neutral rather than risk averse or risk-seekers, which coincides with previous studies.” In the discussion of question 1) How do you perceive risk within the construction industry, the collected data should be adjusted as following: “The results indicate that most of the respondents consider risk as a combination of threats and opportunities. About 94% (113 respondents) have chosen the above option. The rest (7 respondents) perceive risk as a natural hazard.” Figure 1 may be removed as redundant. In the discussion of question 2) What is your attitude in relation to risk, the collected data should be adjusted as following: “Most of the respondents dislike risks (92% or 110 respondents) and some selected “Risk seeking” or “Risk avoiding” as their answers. This result is almost contrary to a result published in [2], in which the majority of the respondents choose a neutral approach balanced between avoiding and risks. This hints that in Vietnam, contractors seem to be more prone to avoid and limit risks in construction”

Finally, the interview part should be perceived as a subjective view of the matter that cannot offer objectivity.

REFERENCES

- [1] V. T. Phan, "Contractor's attitude towards risk and risk management in construction in two western provinces of Vietnam," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6418-6421, 2020. <https://doi.org/10.48084/etasr.3339>.
- [2] D. Petrovic, "Risk Management in Construction Projects: A Knowledge Management Perspective from Swedish Contractors," M.S. Thesis, Department of Real Estate and Construction Management, Royal Institute of Technology, Stockholm, Sweden, 2017.