Increasing Gain Evaluation of 2×1 and 2×2 MIMO Microstrip Antennas

Majed Omar Dwairi
Al-Balqa Applied University
Faculty of Engineering Technology
Department of Communication Technology
Jordan, Amman
majeddw@bau.edu.jo

Abstract-In this paper, a semi-circular ultra-wideband antenna has been modified according to the 2×1 and 2×2 MIMO scenarios. The proposed antennas were designed based on the FR-4 substrate material with dimensions of 36×50mm and 60×60mm for 2×1 and 2×2 scenarios respectively. Simulation results show that a gain improvement of the proposed MIMO antennas from 1 to 2.5dB has been achieved in comparison with the single patch antenna. The radiation pattern for the original and the proposed 2×2 MIMO antennas are exhibited. The main advantage of the proposed antennas is that the gain improves without the need to increase the operating power. This makes the proposed MIMO antennas suitable to be used for UWB antenna applications.

Keywords-microstrip antenna; ultra-wideband; multi-input multi-output; released gain

I. INTRODUCTION

The microstrip antenna has received the attention of microwave antenna designers in ultra-wideband (UWB) applications. This can be clearly seen after the adaption of the UWB range of 3.1-10.6GHz from the Federal Communication Community 2002 [1]. Microstrip antennas received much attention due to their advantages, when compared with other microwave antennas, such as low profile, light weight, low cost, capability of many frequency operations, and easy integration with the microwave integrated circuit. Many investigated directions in microstrip antenna design have been proposed. Also, many researchers analyzed and experimented with different microstrip antenna simulation tools [2]. The effect of inserting different slots in order to enhance the operating bandwidth has also been studied [3-8]. In [9-12], different patch and ground slots were inserted to eliminate unwanted narrow frequency bands, known as filters. Muli-Input Multi-Output (MIMO) microstrip antenna technology received more attention in microwave wireless communications, due to its ability to increase operating data throughput and its range without the need for additional transmitted power or bandwidth. MIMO technology received more attention as it does not need any additional spectrum while transmitting and receiving information from its multiple channels at the same time. By a proper design and layout of MIMO antenna, multipath characteristics of the wireless communications can be controlled. In [13], a high gain MIMO antenna was introduced for a mm-wave planar antenna by choosing the proper size of the feeding line and loading shorted vias on the particular location of the microstrip line. The mutual coupling effect is one of the most challenging in designing MIMO antennas. The authors minimized the mutual coupling for the designed MIMO by inserting ground slots orthogonally, receiving a triple band resonance frequency: 2.11GHz-2.23GHz, 3.24GHz-3.32GHz, and 3.92GHz-4.42GHz [13]. In [14-16], the authors produced insulation between the patch and the feeding element using aperture - coupling and DGS for isolating the ground. To achieve high isolation MIMO microstrip for WLAN applications, aperture-coupling was used to isolate the T-shaped and orthogonal microstrip around the squared patch. In [17], four pair conformal microstrip MIMO antennas of 35GHz have been investigated, resulting in increased bandwidth with reduced side lobes at the same time. Authors in [18, 19] experimented with a 2×2 MIMO microstrip antenna, suitable for new wireless communication demands. The investigated antenna was designed for 5G mobile handset. In [20], the authors designed a four-element MIMO patch antenna for the 5G mobile handset that increases the signal-to-noise ratio and reduces Bit Error Rate (BER). In [21], a high gain pentagonal microstrip antenna compatible for 5G applications, with a received gain of 6.17dB was proposed.

In this paper, an UWB antenna [7] has been modified to 2×1 and 2×2 MIMO scenarios. The modified antennas are compared in terms of increased released gain without any additional power. The expected gain increased from 1 to 2.5dB which shows the effectiveness of the proposed antennas.

II. MIMO ANTENNA DESIGN

Figures 1(b) and 2(b) show the front and back views of the investigated antenna respectively. This antenna was modified to the case of 2×1 structure by extending the width of the substrate to the double of its original value and adding P = 2mm to separate the ground of two antennas. The added patch antenna chased the same dimensions with inverting 180 degrees. The modified 2×2 microstrip antenna was designed by extending the dimension of the substrate width and length to the substrate width plus substrate length (Ws + Ls) as shown in

---

Corresponding author: Majed Omar Dwairi

Figures 1(c) and 2(c). All the investigated antenna materials are built on FR-4 with the substrate with dielectric tangent loss = tan δ= 0.02 and relative permittivity ε_r =4.4. The dimensions of the investigated antennas are listed in Table I. The microstrip patch antenna in [7] was investigated to enhance the bandwidth by adding different slots on the ground. In our work, the chosen slot is the dinking cup-shaped. The bandwidth of the investigated 2×1 and 2×2 antennas was limited to the maximum frequency of 15GHz in order to increase the released gain. In our work, the CST-2019 simulation tool was used.

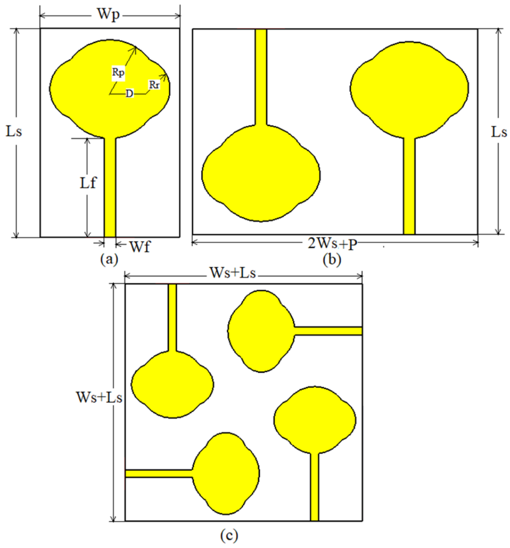


Fig. 1. Front view of (a) the single antenna, (b) the MIMO 2×1 antenna, (c) the MIMO2×2 antenna.

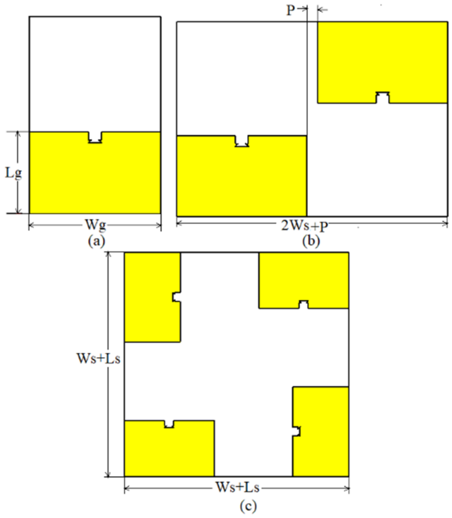


Fig. 2. Back view of (a) the single antenna, (b) the MIMO 2×1 antenna, (c) the MIMO2×2 antenna.

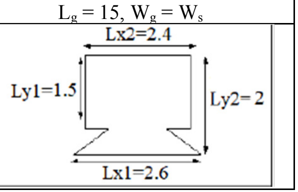
III. RESULTS AND DISCUSSION

All simulations were performed with the CST MW studio - 2019 simulation tool.

A. Return Loss Results

To investigate the return loss S_{11}, the bandwidth was limited to the frequency band from 2 to 15GHz. The results of S_{11}, S_{22} for the 2×1 antenna are shown in Figure 3 and the results of S_{11}, S_{22}, S_{33}, and S_{44} for the 2×2 antenna are presented in Figure 4. It is evident that all the parameter values are within the investigated UWB. The received return loss for each of the antennas and the coupling of the other antennas are presented in Figures 5-8. In these Figures, it can be seen that the reflection parameters have many resonant frequencies, from -17 to -32dB, while the coupling of the antenna reaches -50dB. The received return loss S_{11} for the original, 2×1, and 2×2 antennas at the operating band of frequencies is shown in Figure 9. Figure 3 shows the return loss versus frequencies for the original, 2×1, and 2×2 antennas. It can be noted that the two investigated antennas are still operating within the UWB range. Moreover, both of them start before the original antenna for return loss $S_{11} < -10$ dB, and the upper frequency ends after the original antenna for return loss $S_{11} < -10$ dB.

TABLE I. INVESTIGATED ANTENNA PARAMETERS

Antenna elements	Dimensions [mm]
Patch	R_p=8.5, Rr=5.36, t_p=0.009
Feed line	L_f = 16.82, W_f = 2
Substrate [mm]	L_s = 36, W_s= 24 , t_s= 1.6 ,P=2
Partial ground plane [mm]	L_g = 15, W_g = W_s
Drinking cup shape ground slots dimensions [mm]	

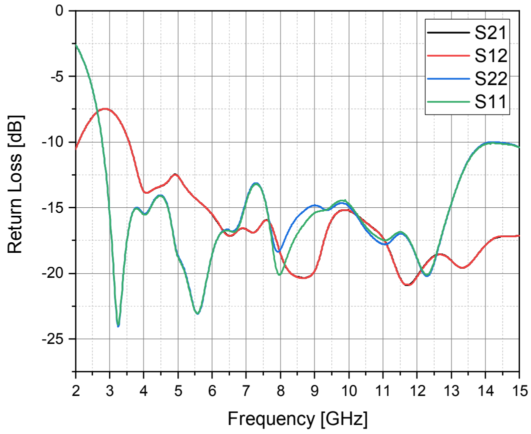


Fig. 3. Reflection and transmssion coefficients for the 2×1 MIMO antenna.

Table II lists the comparison between the original and the two investigated antennas in terms of bandwidth and relative bandwidth. The relative bandwidth for the 2×1 and 2×2 antennas are more than the original antenna by 7.5% and 11.4%, respectively. The relative bandwidth is given by:

$$FBW = 2\frac{f_H-f_L}{f_H+f_L} \quad (1)$$

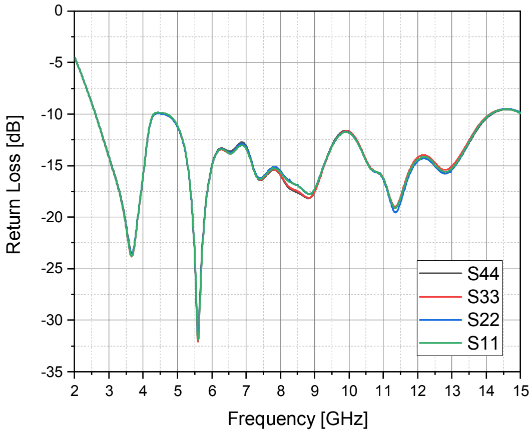


Fig. 4. Return loss for 2×2 MIMO antenna.

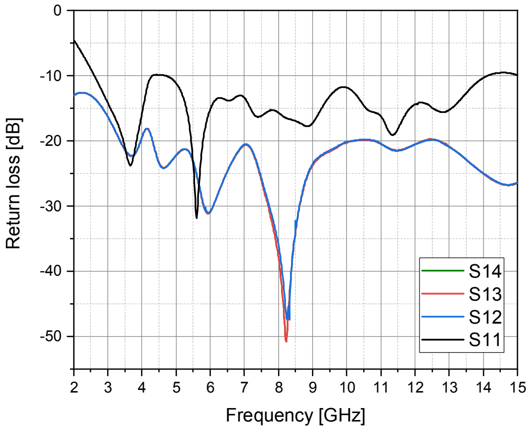


Fig. 5. The return loss for the first antenna and the coupling antenna for 2×2 MIMO.

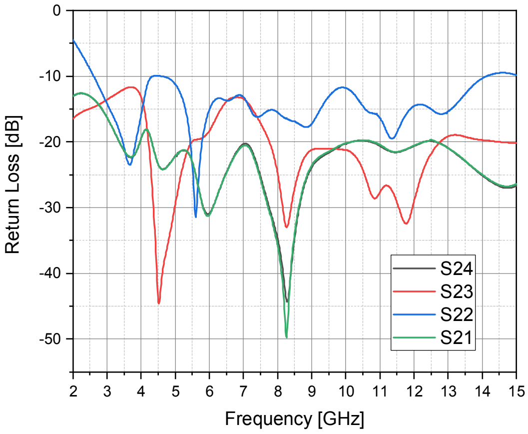


Fig. 6. The return loss for the second antenna and the coupling antenna for 2×2 MIMO.

TABLE II. RELATIVE BANDWIDTH

Antenna	Bandwidth [GHz]	Relative bandwidth %
Original	3.02-13.9	128.6
MIMO 2×1	2.85-15	136.1
MIMO 2×2	2.65-15	140

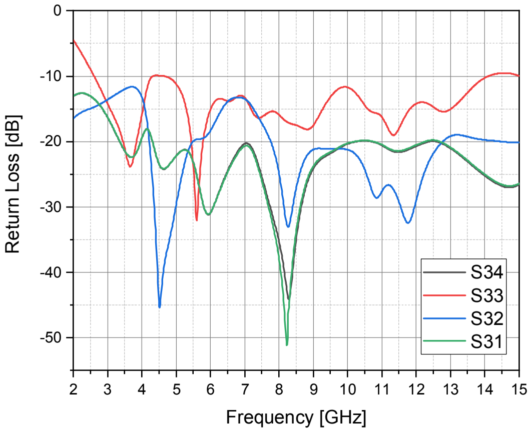


Fig. 7. The return loss for the third antenna and the coupling antenna for 2×2 MIMO.

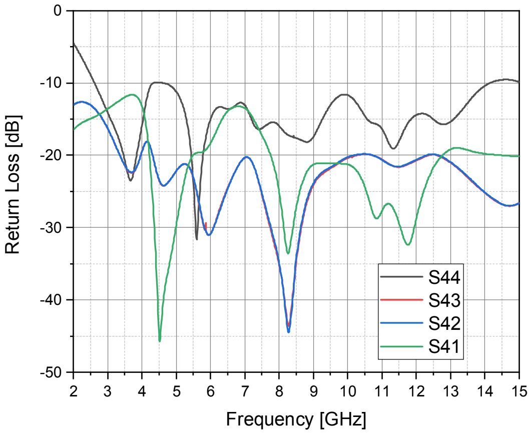


Fig. 8. The return loss for the fourth antenna and the coupling antenna for 2×2 MIMO.

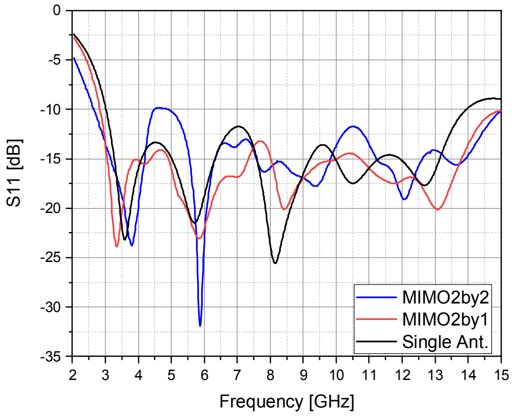


Fig. 9. Return loss versus frequencies for the original, 2×1, and 2×2 antennas.

B. Voltage Standing Wave Ratio (VSWR) Results

Figure 10 shows the results of the investigated antennas in terms of VSWR versus the operating UWB frequencies. As shown, the whole operating bandwidth lies below VSWR equal to 2, which is expected based on the return loss results.

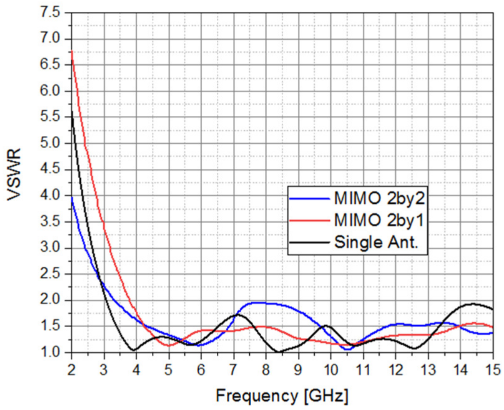


Fig. 10. VSWR versus operating frequencies for the original, 2×1, and 2×2 antennas.

C. *Released Gain Investigation*

The main purpose of this work was to evaluate the released gain for the original and the proposed 2×1 and 2×2 UWB microstrip antennas. Figure 11 represents the released gain for the original antenna and the proposed MIMO antennas. It can be seen from the Figure that the maximum gain for a single antenna was 5.2dB at 8.4GHz. For the 2×1 antenna, the maximum gain was 5.76dB at 9GHz and for the 2×2 antenna it was 6.03dB at 8.4GHz. Table III lists the gain for the three antennas at different frequencies.

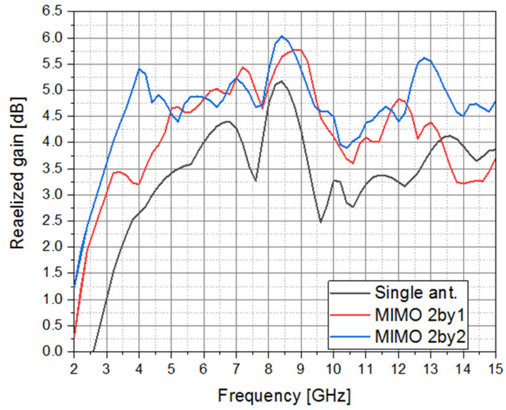


Fig. 11. The released gain for the original, 2×1, and 2×2 antennas.

TABLE III. COMPARISON OF THE GAIN VS FREQUENCY OF THE ORIGINAL AND THE PROPOSED ANTENNAS

Frequency (Hz)	**Antenna gain (dB)**		
	Original	**MIMO 2×1**	**MIMO 2×2**
3.2	1.52	3.42	4.01
5	3.42	4.66	4.55
7.6	3.26	4.98	4.68
8.6	5	5.72	5.94
9.6	2.47	4.47	4.6
11	3.2	4.1	4,4
12	3.25	4.93	4.4
14.6	3.73	3.26	4.65

Figure 12 gives a fast view of the comparison between the original and the two proposed MIMO antennas with different gain. The proposed antennas' gain exceed 2.5dB for some frequencies.

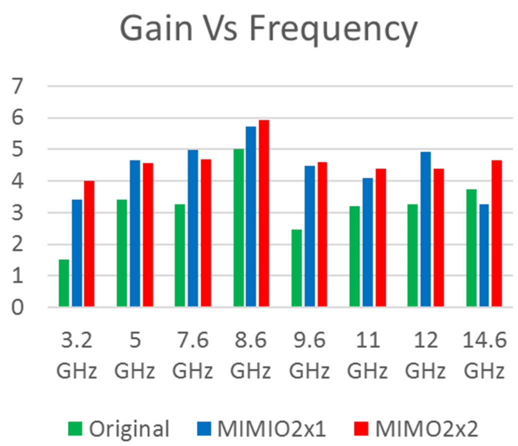


Fig. 12. Comparison between the gain and the frequency for the original, 2×1, and 2×2 antennas.

D. *Radiation Pattern*

Figure 13 shows the radiation pattern in the far-field for different selected resonance frequencies related to the original antenna and the proposed 2×2 antenna. It is evident that the radiation pattern for both antennas is omnidirectional for the lower resonant frequencies and directional for the higher ones.

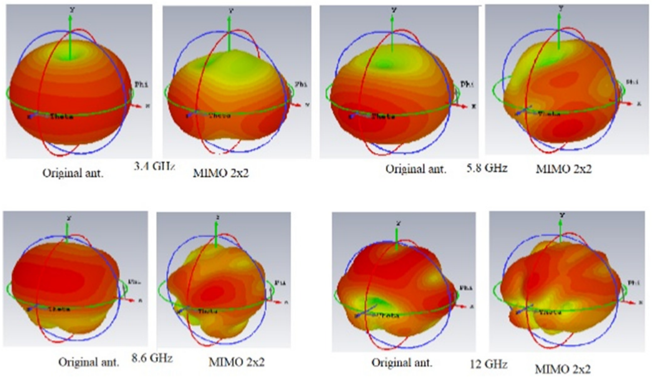


Fig. 13. Radiation pattern (far-field pattern) for the original antenna and the MIMO 2×2 antenna.

E. *A Comparison beween the Proposed 2×2 MIMO Antenna and other Reported Works*

In order to evaluate the proposed antenna, it is worth to compare it with the results of other reported works. The comparisons were conducted in terms of antenna dimensions, number of antennas in MIMO, operating frequencies, and peak gain. As shown in Table IV, it is clear that some of the reported works have peak gain more than that of the proposed antenna, but, in general, our peak gain is quite satisfying and similar or bigger than that of the other reported gains. The importance of the investigated antenna is that it operates in UWB frequencies which allows it to be implemented in different narrow bands.

TABLE IV. COMPARISON OF THE PROPOSED ANTENNAS AND OTHER REPORTED WORKS

Ref.	**Dimensions [mm]**	**MIMO antenna**	**Operating Frequency [GHz]**	**Peak gain for MIMO antenna [dB]**
[12]	20×24	2×1	42.0-49.0	>8
[14]	70×70	2×2	2-4.5	1.84-3.49
[18]	20×45	2×2	2.97-19.82	3.3-8.12
[19]	15×10.3	2×2	27-28.95	6.14
[20]	66×66	2×2	2.2-2.7	6.17
Proposed	60×60	2×2	2.65-15	>6

IV. CONCLUSION

Two UWB 2×2 and 2×1 microstrip patch antennas have been introduced in this paper. The released gain improvement, compared to the original one, for the two proposed MIMO antennas varied from 1 to 2.5dB without increasing the radiated power So, it is evident that the proposed MIMO antennas' release gain is bigger than that of the used UWB antenna for the same operating frequencies, which gives more stability. In the future, the simulated antenna will be fabricated for experimental validation.

REFERENCES

[1] "Revision of Part 15 of the Commission's Rules Regarding Ultra WideBand Transmission Systems," Federal Communications Commission, Washington DC, USA, ET Docket 98-153 FCC 02-48, 2002.

[2] M. S. Soliman, M. O. Dwairi, I. I. M. A. Sulayman, and S. H. A. Almalki, "A Comparative Study for Designing and Modeling Patch Antenna with Different Electromagnetic CAD Approaches," *International Journal on Communications Antenna and Propagation,* vol. 6, no. 2, pp. 90–95, Apr. 2016, https://doi.org/10.15866/irecap.v6i2.8718.

[3] M. O. Dwairi, M. S. Soliman, A. A. Alahmadi, S. H. A. Almalki, and I. I. M. Abu Sulayman, "Design and Performance Analysis of Fractal Regular Slotted-Patch Antennas for Ultra-Wideband Communication Systems," *Wireless Personal Communications*, vol. 105, no. 3, pp. 819–833, Apr. 2019, https://doi.org/10.1007/s11277-019-06123-5.

[4] M. A. S. Al-zahrani, O. I. S. Al-qahtani, F. D. M. Al-sheheri, A. S. M. Qarhosh, A. M. Al-zahrani, and M. S. Soliman, "Design and performance analysis of an ultra-wideband monopole microstrip patch antenna with enhanced bandwidth and single band-notched characteristics," in *2017 Progress in Electromagnetics Research Symposium - Fall (PIERS - FALL)*, Singapore, Nov. 2017, pp. 1488–1493, https://doi.org/10.1109/PIERS-FALL.2017.8293366.

[5] M. O. Dwairi, M. S. Soliman, and A. A. Al Ahmadi, "Enhanced Bandwidth Performance for Monopole Ultra-wideband Patch Antenna for Super-wideband Applications," presented at the The Second Irbid International Engineering Conference, Irbid, Jordan, 2017.

[6] M. O. Dwairi, M. S. Soliman, A. A. Alahmadi, I. I. M. A. Sulayman, and S. H. A. Almalki, "Design regular fractal slot-antennas for ultra-wideband applications," in *2017 Progress In Electromagnetics Research Symposium - Spring (PIERS)*, St. Petersburg, Russia, May 2017, pp. 3875–3880, https://doi.org/10.1109/PIERS.2017.8262435.

[7] M. O. Al-Dwairi, A. Y. Hendi, M. S. Soliman, and M. A. Nisirat, "Design of A Compact Ultra-Wideband Antenna for Super-Wideband Technology," in *2019 13th European Conference on Antennas and Propagation (EuCAP)*, Krakow, Poland, Mar. 2019, pp. 1–4.

[8] M. S. Soliman, M. O. Dwairi, and I. I. M. A. Sulayman, "The Effect of the Ground Slots Up on the Bandwidth Performance for UWB Antenna," in *2018 18th Mediterranean Microwave Symposium (MMS)*, Istanbul, Turkey, Oct. 2018, pp. 68–70, https://doi.org/10.1109/MMS.2018.8612063.

[9] M. S. Soliman, M. O. Al-Dwairi, A. Y. Hendi, and Z. Alqadi, "A Compact Ultra-Wideband Patch Antenna with Dual Band-Notch Performance for WiMAX / WLAN Services," in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, Amman, Jordan, Apr. 2019, pp. 831–834, https://doi.org/10.1109/JEEIT.2019.8717444.

[10] S. Alotaibi and A. A. Alotaibi, "Design of a Planar Tri-Band Notch UWB Antenna for X-Band, WLAN, and WiMAX," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6557–6562, Dec. 2020, https://doi.org/10.48084/etasr.3904.

[11] M. O. Al-Dwairi, "A planar UWB semicircular-shaped monopole antenna with quadruple band notch for WiMAX, ARN, WLAN, and X-Band," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 908–918, Feb. 2020, https://doi.org/10.11591/ijece.v10i1.pp908-918.

[12] M. O. Al-Dwairi, A. Y. Hindi, M. S. Soliman, and M. F. Aljafari, "A compact UWB monopole antenna with penta band notched characteristics," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 622–630, Apr. 2020, https://doi.org/10.12928/telkomnika.v18i2.14542.

[13] J.-T. Weng and Q.-X. Chu, "Wideband microstrip MIMO antenna for millimeter-wave applications," in *2017 Sixth Asia-Pacific Conference on Antennas and Propagation (APCAP)*, Xi'an, China, Oct. 2017, pp. 1–3, https://doi.org/10.1109/APCAP.2017.8420888.

[14] Banothu Y. V. N. R. Swamy and P. Siddaiah, "Design of a Compact 2×2 Multi Band MIMO Antenna for Wireless Applications," *International Journal of Recent Technology and Engineering*, vol. 7, no. 6S2, pp. 674–683, 2019.

[15] H. Li, T. Wei, J. Ding, and C. Guo, "A dual-band polarized diversity microstrip MIMO antenna with high isolation for WLAN application," in *2016 11th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, Guilin, China, Oct. 2016, pp. 88–91, https://doi.org/10.1109/ISAPE.2016.7833885.

[16] A. Khan, S. Geng, X. Zhao, Z. Shah, M. U. Jan, and M. A. Abdelbaky, "Design of MIMO Antenna with an Enhanced Isolation Technique," *Electronics*, vol. 9, no. 8, Aug. 2020, Art. no. 1217, https://doi.org/10.3390/electronics9081217.

[17] Q. Wang, N. Mu, L. Wang, S. Safavi-Naeini, and J. Liu, "5G MIMO Conformal Microstrip Antenna Design," *Wireless Communications and Mobile Computing*, vol. 2017, Dec. 2017, Art. no. e7616825, https://doi.org/10.1155/2017/7616825.

[18] H. Alsaif, "Extreme Wide Band MIMO Antenna System for Fifth Generation Wireless Systems," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5492–5495, Apr. 2020, https://doi.org/10.48084/etasr.3413.

[19] R. K. Goyal and U. S. Modani, "A Compact MIMO Microsrip Patch Antenna Design at 28 GHz for 5G Smart Phones," *International Journal of Engineering Research & Technology*, vol. 9, no. 4, Mar. 2021, Art. no. IJERTCONV9IS04001.

[20] R. S. Bhadade and S. P. Mahajan, "High Gain Circularly Polarized Pentagonal Microstrip for Massive MIMO Base Station," *Advanced Electromagnetics*, vol. 8, no. 3, pp. 83–91, Sep. 2019, https://doi.org/10.7716/aem.v8i3.764.

Denoising the ECG Signal Using Ensemble Empirical Mode Decomposition

Wahiba Mohguen
CCNS Laboratory
Department of Electronics
Faculty of Technology
University of Ferhat Abbas - Setif 1
Setif, Algeria
wasotel@yahoo.fr

Saad Bouguezel
CCNS Laboratory
Department of Electronics
Faculty of Technology
University of Ferhat Abbas - Setif 1
Setif, Algeria
bouguezel_saad@yahoo.com

Abstract-In this paper, a novel electrocardiogram (ECG) denoising method based on the Ensemble Empirical Mode Decomposition (EEMD) is proposed by introducing a modified customized thresholding function. The basic principle of this method is to decompose the noisy ECG signal into a series of Intrinsic Mode Functions (IMFs) using the EEMD algorithm. Moreover, a modified customized thresholding function was adopted for reducing the noise from the ECG signal and preserve the QRS complexes. The denoised signal was reconstructed using all thresholded IMFs. Real ECG signals having different Additive White Gaussian Noise (AWGN) levels were employed from the MIT-BIH database to evaluate the performance of the proposed method. For this purpose, output SNR (SNRout), Mean Square Error (MSE), and Percentage Root mean square Difference (PRD) parameters were used at different input SNRs (SNRin). The simulation results showed that the proposed method provided significant improvements over existing denoising methods.

Keywords-denoising; ECG; EMD; EEMD; customized thresholding

I. INTRODUCTION

Empirical Mode Decomposition (EMD) is a powerful algorithm for splitting non-stationary signals [1]. The goal of EMD is to represent the signals as sums of zero-mean oscillating components, named Intrinsic Mode Functions (IMFs) via a sifting process [1]. Signal reconstruction is achieved by summing all IMFs and the residual. EMD techniques have been used for signal denoising, and specifically, those based on thresholding were developed in [2-10]. A denoising technique can be based on signal estimation using all the previously thresholded IMFs [3-13]. Since the useful information of the signal is often concentrated on low-frequency IMFs (last IMFs) and the noise is primarily located in high-frequency IMFs (first IMFs), another approach is to perform denoising by partial construction of the signal with the IMFs that contain useful information [2, 14]. Authors in [2] proposed a method for estimating the energy of noisy IMFs from a theoretical model and IMFs' energies of the test signal, and the signal was reconstructed partially by using only the IMFs that contained useful information, eliminating those that essentially maintained noise. In [14], an EMD consecutive mean square error (EMD-CMSE) method was developed for IMF selection. Since Electrocardiogram (ECG) signals are nonstationary and nonlinear methods, a wavelet thresholding technique was proposed in [15, 16] without preserving ECG components such as QRS complexes [17]. A customized thresholding function was proposed in [18] to overcome the disadvantages of hard and soft thresholding functions [15, 16]. EMD combined with a customized thresholding function (EMD-Custom) can be useful for reducing noise and significantly improve the results of EMD soft and hard thresholding [3, 4, 8]. To overcome the drawbacks of EMD such as mode mixing (presence of oscillations of different amplitudes in one mode) [1], a variant of the EMD algorithm called Ensemble Empirical Mode Decomposition (EEMD) was proposed in [19]. EEMD was based on averaging the modes obtained from EMD applied to several trials of Additive Gaussian White Moise (AWGN) added to the signal. The EEMD decomposition resolved efficiently the mode mixing and has been widely used in noise reduction. Moreover, EEMD achieved better denoising performance than EMD with a reduced number of trials.

The main objective of this paper is to propose a denoising method for ECG signals using EEMD and a modified custom thresholding function. The basic principle of the proposed method is to decompose the noisy signal into a series of IMFs using the EEMD algorithm and then use the modified custom thresholding function. The denoised signal is reconstructed using all the thresholded IMFs. Denoising experiments were used on MIT-BIH ECG signals to assess the performance of the proposed method [20] with different AWGN levels. Three standard parameters were used at different input SNR (SNRin): output SNR (SNRout), Mean Square Error (MSE), and percentage Root Mean square Difference (PRD). The proposed method is compared to EMD-CMSE [14], EMD-Custom [8], and wavelet [15,16] denoising methods.

II. WAVELET DENOISING

Wavelet denoising is a powerful tool for removing the noisy component of a corrupted data sequence [15, 16]. Its

---

Corresponding author: Wahiba Mohguen

basic steps are:

- Decompose: Choose a wavelet and a level *N*. Compute the wavelet decomposition of the signal at level *N*.
- Threshold detail coefficients: For each level from 1 to *N*, select a threshold and apply soft thresholding to the detail coefficients.
- Reconstruct: Compute wavelet reconstruction using the original approximation coefficients of level *N* and the modified detail coefficients of levels 1 to *N*.

This work used the Symlet wavelet (sym8), while thresholding can be performed by using the soft or hard thresholding of the input signal. Wavelet soft thresholding-based denoising technique was applied, the signal was divided into a set of approximations, and detail coefficients were thresholded using soft thresholding. The universal threshold estimator proposed in [15, 16] was used.

III. EMD-CUSTOM THRESHOLDING

EMD-custom suggests the decomposition of a noisy signal to noisy IMFs via the EMD algorithm [1]. After that, the noisy denoted IMFs were thresholded using a customized thresholding function [18]. Finally, the denoised signal was reconstructed using all thresholded IMFs. The outline of the EMD-Custom [8] method is demonstrated in Figure 1.

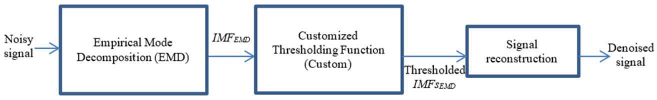


Fig. 1. Outline of the EMD-custom.

IV. EEMD ALGORITHM

The EEMD method [19] overcomes the "mode mixing" problem of the EMD method and consists of:

- Adding a white noise $w^j(t)$ to the original signal $x(t)$:

$$x^j(t) = x(t) + w^j(t), \ \ 1 \le j \le N_e \qquad (1)$$

where N_e is the ensemble number.

- Decomposing the noisy signal $x^j(t)$ into IMFs by the EMD method to obtain the corresponding IMF of each order denoted $z_i^j(t)$, where *i* is the IMF order, *j* is the trial index, *N* is the number of IMFs, and $1 \le i \le N$.
- Calculating the mean of the corresponding IMFs as the final signal IMF, by:

$$Z_{iEEMD}(t) = \frac{1}{N_e}\sum_{j=1}^{N_e} z_i^j(t), \ \ 1 \le i \le N \qquad (2)$$

V. PROPOSED DENOISING METHOD

The proposed method suggests the decomposition of the noisy signal to noisy IMFs via the EEMD algorithm [19]. Afterward, the noisy IMFs denoted as $f_i(t)$ are thresholded using a modified custom thresholding function. Let $s(t)$ be a noisy signal given as:

$$s(t) = x(t) + w(t) \qquad (3)$$

where $x(t)$ is the noiseless signal and $w(t)$ is an independent noise of finite amplitude. The proposed EEMD-Custom method consists of the following steps:

- Decompose the noisy signal $s(t)$ by the EEMD algorithm to extract the noisy IMFs $f_i(t)$.
- Apply a modified custom thresholding function on the noisy IMFs $f_i(t)$. A modification of the customized thresholding function [18] is introduced to define a new one as:

$$\hat{z}_i(t) = \begin{cases} f_i(t) - sgn(f_i(t))[1-\alpha]\tau_i, & if\ |f_i(t)| \ge \tau_i \\ 0, & if\ |f_i(t)| \le \gamma \end{cases} \qquad (4)$$

where $0 < \gamma < \tau_i$, $0 \le \alpha \le 1$ and τ_i is the universal threshold reported in [15, 16] defined as:

$$\tau_i = C\sqrt{E_i 2\ln(n)} \qquad (5)$$

where *C* is a constant depending on the type of the signal, *n* is the length of the signal, and E_i is given by:

$$E_i = \frac{E_1^2}{\theta}\rho^{-i}, \ i = 2,3,4,\ldots,N \qquad (6)$$

where E_1^2 is the energy of the first IMF, obtained as:

$$E_1^2 = \left(\frac{median(|f_1(t)|)}{0.6745}\right)^2 \qquad (7)$$

where $\theta = 0.719$ and $\rho = 2.01$ are empirically calculated constants [2].

- Reconstruct the signal using:

$$\hat{x}(t) = \sum_{i=1}^{N} \hat{z}_i(t) + r(t) \qquad (8)$$

where $r(t)$ is the residual signal.

VI. RESULTS AND DISCUSSION

In this section, the results of the proposed denoising method are assessed compared to three denoising methods: wavelet denoising [15, 16], EMD-CMSE [14], and EMD-Custom [8]. The proposed EEMD-Custom algorithm was applied to 8 real biomedical ECG signals using the MIT-BIH database [20], labeled 111m, 112m, 113m, 114m, 115m, 116m, 121m, and 122m. An AWGN was added to each clean ECG signal at different *SNRin* levels: -4dB, 0 dB, 4 dB, 8 dB, and 12 dB. The data length was 2048. At first, each noisy ECG signal was decomposed into a series of IMFs via the EEMD algorithm, and subsequently, the modified customized thresholding function (4) was utilized to threshold all IMFs for reducing noise and preserve QRS complexes. Thresholding can be used to detect QRS complexes [21]. So, the combination between EEMD and the modified customized thresholding function can be considered as an *R* peak preservation technique, as the IMFs containing high-frequency signal information (QRS complex) were thresholded by the modified customized thresholding function to preserve the QRS complexes.

Finally, the denoised signal was reconstructed using all thresholded IMFs. Three standard parameters, *SNRout*, *MSE*, and *PRD*, were used to evaluate the capabilities of the proposed method at different *SNRin*, which were respectively given as:

$$SNR_{out} = 10\log_{10}\frac{\sum_{t=1}^{n}(x(t))^2}{\sum_{t=1}^{n}(\hat{x}(t)-x(t))^2} \quad (9)$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(\hat{x}(t) - x(t) \quad (10)$$

$$PRD = 100 * \sqrt{\frac{\sum_{t=1}^{n}(\hat{x}(t)-x(t))^2}{\sum_{t=1}^{n}(x(t))^2}} \quad (11)$$

The performance of the proposed EEMD-Custom method was evaluated for different values of ensemble number *Ne* (10, 20, 30, 40, 50, 100, 150, 200, 250, 300). Figure 2 depicts the *SNRout* for different values of *Ne* at *SNRin*=4dB on ECG record 112m. Figure 3 displays the plot of *SNRout* for different values of *Ne* at *SNRin*=4dB on ECG records 114m, 116m, and 122m. As it can be observed, the SNRout increases as *Ne* increases. Moreover, the proposed method achieved a significant improvement when *Ne* was high. Based on the results, an ensemble number of 200 was selected as the best EEMD parameter. Furthermore, as the results of the proposed EEMD-Custom method were influenced by the α value in (4), an appropriate value of α should be determined.

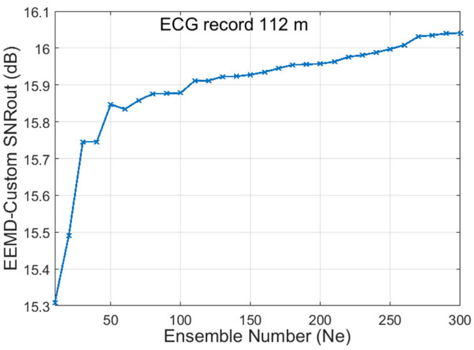


Fig. 2. *SNRout* of the proposed EEMD-Custom in terms of *Ne*, *SNRin*=4dB, for the ECG 112m signal.

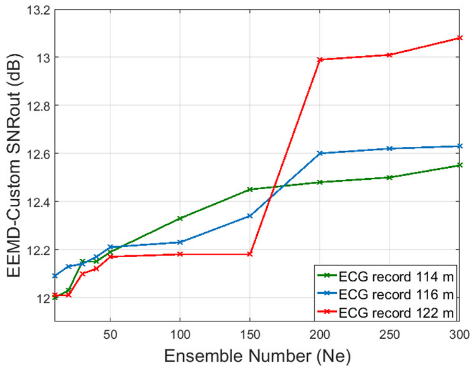


Fig. 3. *SNRout* of the proposed EEMD-Custom in terms of *Ne*, *SNRin*=4dB, for the ECG records 114m, 116m, and 122 m.

Figures 4 and 5 show the *SNRout* of the proposed EEMD-Custom method as a function of α for different ECG signals. The value of the ensemble number was *Ne*=200, the *SNRin* was 8dB, and the value of α was determined by trials between 0.1 and 1, with a fixed step of 0.1. The values of a where the SNRout was maximum, were 0.7, 0.3, 0.5, 0.5, 0.5, 0.5, 0.3, and 0.5 for ECG records 111m, 112m, 113m, 114m, 115m, 116m, 121m, and 122m respectively. The α parameter depended on SNRin and the type of ECG signal.

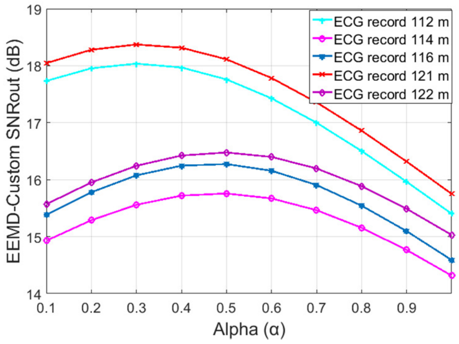


Fig. 4. *SNRout* of the proposed EEMD-Custom in terms of α, *SNRin*=8dB, for different ECG signals.

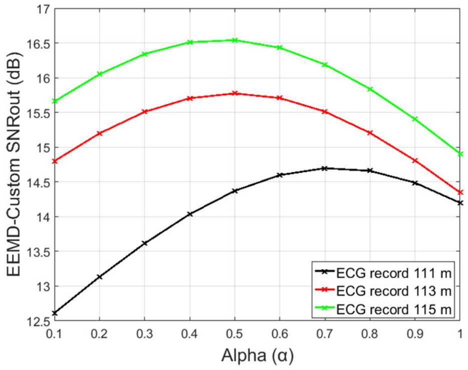


Fig. 5. *SNRout* of the proposed EEMD-Custom in terms of α, *SNRin*=8dB, for the ECG records 111m, 113m, and 115 m.

The SNRout of five ECG signals at various *SNRin* values is presented in Table I to provide quantitative analysis. The proposed EEMD-Custom denoising method on the ECG record 112m reached an improvement of SNR equal to 6.95dB at *SNRin*=-4dB compared to the wavelet denoising method. An improvement of 3.89dB was obtained at *SNRin*=0dB compared to EMD-CMSE, while an improvement of 4.46dB was obtained at *SNRin*=-4dB compared to the EMD-Custom. The proposed EEMD-Custom denoising method improved SNR by 6.36dB, 4.34dB, and 3.87dB on the ECG 121m at *SNRin*=-4dB compared to wavelet denoising [2], EMD-CMSE [5,6] and EMD-Custom [8], respectively. As it can be noted, the proposed method worked better for high and low values of *SNRin*. Moreover, the values of *MSE* and *PRD* of ECG signals for different SNRin values are presented in Tables II and III, respectively. The quality of the reconstructed ECG signal was evaluated in terms of *PRD*, *SNRout*, and *MSE*. The *MSE* and *PRD* values should be small for better denoising and preserving ECG signal details. Lower *PRD* and *MSE* values indicate better preservation of physiological information in ECG signal processing [17]. As it can be noted, the proposed EEMD-Custom method provided less *MSE* and *PRD* than the other methods.

TABLE I. SNROUT OBTAINED BY DIFFERENT METHODS

Signals	**Methods**				
	***SNRin* (dB)**	**Wavelet (Sym 8)**	**EMD-CMSE**	**EMD-Custom**	**Proposed EEMD-Custom**
ECG 111.M	-4	1.44	2.33	2.83	4.48
	0	5.38	5.87	6.51	7.79
	4	9.24	9.75	10.41	11.24
	8	12.90	12.79	14.26	14.69
	12	16.16	14.82	17.56	17.69
ECG 112.m	-4	4.61	7.89	7.10	11.56
	0	8.58	10.27	10.94	14.16
	4	12.44	13.78	14.31	15,95
	8	16.10	16.36	17.54	18.03
	12	19.36	18.71	20.77	20.99
ECG 113.M	-4	1.43	2.96	3.95	5.86
	0	5.35	6.12	7.55	9.20
	4	9.17	9.41	11.12	12.54
	8	12.75	12.75	14.89	15.77
	12	15.85	16.04	18.48	18.84
ECG 114.m	-4	4.21	6.37	6.13	8.44
	0	7.57	7.38	9.65	10.43
	4	10.30	10.54	12.00	12.48
	8	12.16	12.56	15.56	15.75
	12	13.20	13.64	18.00	18.33
ECG 115.M	-4	1.43	3.61	4.12	6.46
	0	5.36	5.87	7.72	10.16
	4	9.18	9.64	11.51	13.30
	8	12.77	12.87	15.02	16.54
	12	15.88	14.79	18.61	19.25
ECG 116. M	-4	4.26	5.64	5.96	7.54
	0	8.06	7.94	9.28	9.98
	4	11.23	10.30	12.10	12.60
	8	13.68	13.81	15.71	16.27
	12	15.24	16.88	19.02	19.32
ECG 121.M	-4	4.66	6.68	7.15	11.02
	0	8.64	8.52	10.36	12.98
	4	12.58	12.49	13.57	15.44
	8	16.42	16.66	17.56	18.37
	12	20.07	18.75	20.93	21.14
ECG 122.M	-4	4.57	6.95	6.45	8.60
	0	8.45	8.14	9.85	10.62
	4	12.14	11.81	13.02	12.99
	8	15.42	15.03	16.05	16.47
	12	18.03	18.32	19.17	19.68

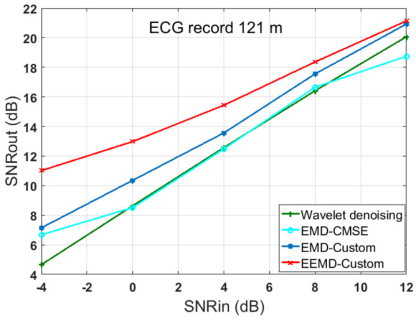


Fig. 6. *SNRout* versus *SNRin* for different denoising methods.

Figures 6, 7, and 8 present *SNRout*, *MSE*, and *PRD* versus *SNRin* of the denoised real ECG record 121m for wavelet denoising [2], EMD-CMSE [5, 6], EMD-Custom [8], and the proposed EEMD-Custom method. The proposed EEMD-Custom method gave better results in all cases.

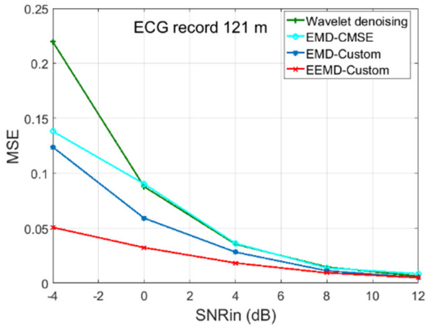


Fig. 7. *MSE* versus *SNRin* for different denoising methods.

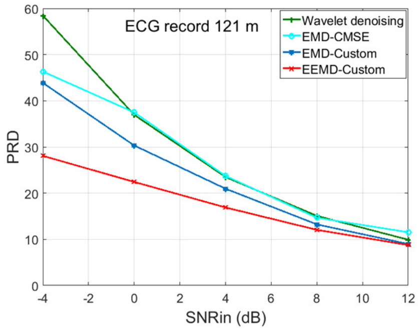


Fig. 8. *PRD* versus *SNRin* for different denoising methods.

The clean ECG records 112m and 122m, their noisy versions, and denoised ECG records using the proposed EEMD-Custom at *SNRin*=8dB with *Ne*=200 are depicted in Figures 9 and 10 respectively. It can be noted that the proposed method removes noise successfully. Figure 11 depicts the denoised ECG record 122m using the wavelet method at *SNRin*=8dB. A careful comparison of the denoised signals in Figures 10 and 11 shows that the proposed method preserves morphological information of ECG better than the wavelet denoising method. The results also indicate that the proposed method can remove noise from real ECG signals and provide significant improvements in denoising performance. The computational complexity of EEMD can be expressed as:

$$T_{EEMD} = N_e * T_{EMD} \qquad (12)$$

demonstrating that EEMD takes more time than EMD.

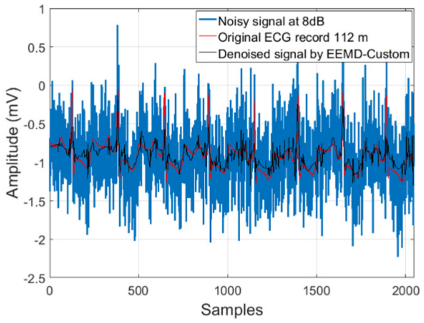


Fig. 9. ECG 112m signal denoised by the proposed EEMD-Custom method.

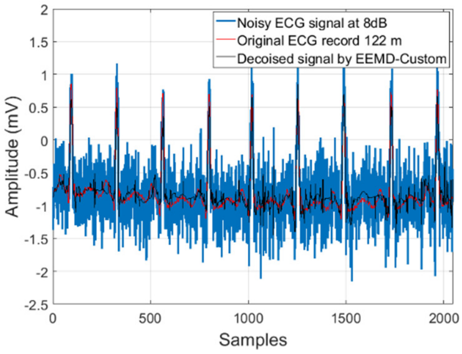


Fig. 10. ECG 122m signal denoised by the proposed EEMD-Custom method.

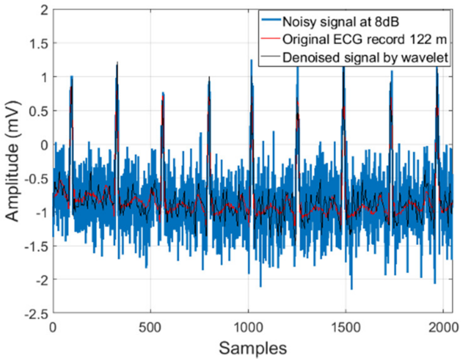


Fig. 11. ECG 122m signal denoised by the wavelet (sym8) method.

TABLE II. MSE OBTAINED BY DIFFERENT METHODS.

Signals	Methods				
	SNRin **(dB)**	**Wavelet (Sym 8)**	**EMD-CMSE**	**EMD-Custom**	**Proposed EEMD-Custom**
ECG 111.M	-4	0.0405	0.0330	0.0293	0.0201
	0	0.0163	0.0146	0.0125	0.0093
	4	0.0067	0.0059	0.0051	0.0042
	8	0.0028	0.0029	0.0021	0.0019
	12	0.0013	0.0018	0.0009	0.0009
ECG 112.m	-4	0.3053	0.1432	0.1719	0.0615
	0	0.1221	0.0828	0.0710	0.0338
	4	0.0502	0.0368	0.0327	0.0229
	8	0.0216	0.0203	0.0155	0.0138
	12	0.0102	0.0118	0.0073	0.0070
ECG 113.M	-4	0.1819	0.1278	0.1017	0.0655
	0	0.0736	0.0618	0.0443	0.0303
	4	0.0305	0.0289	0.0195	0.0140
	8	0.0134	0.0134	0.0081	0.0066
	12	0.0065	0.0062	0.0035	0.0033
ECG 114.m	-4	0.0201	0.0122	0.0129	0.0076
	0	0.0092	0.0097	0.0057	0.0048
	4	0.0049	0.0046	0.0030	0.0030
	8	0.0032	0.0029	0.0014	0.0014
	12	0.0025	0.0022	0.0007	0.0007
ECG 115.m	-4	0.2492	0.1506	0.1342	0.0782
	0	0.1008	0.0896	0.0585	0.0334
	4	0.0418	0.0376	0.0244	0.0161
	8	0.0183	0.0178	0.0109	0.0076
	12	0.0089	0.0114	0.0047	0.0041
ECG 116.m	-4	0.5085	0.3698	0.3437	0.2389
	0	0.2116	0.2176	0.1600	0.1361
	4	0.1019	0.1263	0.0835	0.0743
	8	0.0580	0.0564	0.0364	0.0319
	12	0.0405	0.0278	0.0169	0.0158
ECG 121.m	-4	0.2197	0.1381	0.1238	0.0507
	0	0.0878	0.090	0.0591	0.0323
	4	0.0354	0.0361	0.0282	0.0183
	8	0.0146	0.0138	0.0112	0.0093
	12	0.0063	0.0085	0.0051	0.0049
ECG 122.m	-4	0.2767	0.3997	0.1793	0.1092
	0	0.1131	0.1214	0.0819	0.0687
	4	0.0484	0.0522	0.0395	0.0398
	8	0.0227	0.0248	0.0196	0.0178
	12	0.0124	0.0116	0.0096	0.0085

TABLE III. PRD OBTAINED BY DIFFERENT METHODS.

Signals	Methods				
	SNRin **(dB)**	**Wavelet (Sym 8)**	**EMD-CMSE**	**EMD-Custom**	**Proposed EEMD-Custom**
ECG 111.M	-4	84.60	76.43	72.13	59.66
	0	53.79	50.86	47.21	40.77
	4	34.49	32.53	30.16	27.41
	8	22.62	22.91	19.34	18.41
	12	15.55	18.13	13.23	13.03
ECG 112.m	-4	58.81	40.28	44.13	26.41
	0	37.20	30.64	28.36	19.58
	4	23.86	20.44	19.24	16.12
	8	15.65	15.20	13.27	12.53
	12	10.76	20.63	9.14	8.92
ECG 113.M	-4	84.79	71.08	63.41	50.89
	0	53.95	66.91	41.88	34.65
	4	34.75	33.82	27.77	23.59
	8	23.02	23.02	17.99	16.26
	12	16.12	15.76	11.90	11.42
ECG 114.m	-4	61.58	47.99	49.35	37.81
	0	41.78	42.74	32.92	30.07
	4	30.52	29.69	23.74	23.75
	8	24.64	23.53	16.67	16.29
	12	21.86	20.78	12.10	12.10
ECG 115.m	-4	84.78	65.92	62.22	47.50
	0	53.94	50.85	41.07	31.03
	4	34.73	32.94	26.56	21.61
	8	22.98	22.70	17.73	14.89
	12	16.06	18.20	11.73	10.89
ECG 116.m	-4	61.22	52.21	50.34	41.97
	0	39.50	40.05	34.35	31.68
	4	27.41	30.52	24.81	23.41
	8	20.67	20.39	16.38	15.35
	12	17.27	14.31	11.18	10.80
ECG 121.m	-4	58.45	46.34	43.88	28.08
	0	36.96	37.49	30.32	22.43
	4	23.49	23.72	20.96	16.90
	8	15.08	14.68	13.23	12.05
	12	9.91	11.54	8.98	8.76
ECG 122.m	-4	59.05	44.87	47.54	37.11
	0	37.75	39.12	32.14	29.43
	4	24.71	25.66	22.32	22.40
	8	16.92	22.60	15.75	15.00
	12	12.54	12.12	11.00	10.36

VII. CONCLUSION

This paper presented a novel denoising method based on the EEMD algorithm for noise removal from ECG signals by introducing a modified custom thresholding function. Three standard parameters *SNRout*, *MSE*, and *PRD* were used for evaluating the capabilities of the proposed method at different values of *SNRin.* The simulation results on MIT-BIH ECG signals showed clearly that the proposed method provided better *SNRout* and lesser *MSE* and *PRD* compared to other well-known denoising methods. Therefore, the proposed method is characterized as highly suitable for denoising ECG signals.

REFERENCES

[1] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, Mar. 1998, https://doi.org/10.1098/rspa.1998.0193.

[2] P. Flandrin, P. Gonçalvès, and G. Rilling, "Emd equivalent filter banks, from interpretation to applications," in *Hilbert-Huang Transform and Its Applications*, vol. 5, Singapore: World Scientific, 2005, pp. 57–74.

[3] A. O. Boudraa, J. C. Cexus, and Z. Saidi, "EMD-Based Signal Noise Reduction," *International Journal of Signal Processing*, vol. 1, no. 1, pp. 33–37, 2004.

[4] A.-O. Boudraa and Jean-Christophe Cexus, "Denoising via Empirical Mode Decomposition," presented at the International Symposium on Communications, Control and Signal Processing (ISCCSP '06), Marrakech, Morocco, Mar. 2006.

[5] Y. Kopsinis and S. McLaughlin, "Empirical mode decomposition based soft-thresholding," in *2008 16th European Signal Processing Conference*, Aug. 2008, pp. 1–5.

[6] Y. Kopsinis and S. McLaughlin, "Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351–1362, Apr. 2009, https://doi.org/10.1109/TSP.2009.2013885.

[7] G. Yang, Y. Liu, Y. Wang, and Z. Zhu, "EMD interval thresholding denoising based on similarity measure to select relevant modes," *Signal Processing*, vol. 109, pp. 95–109, Apr. 2015, https://doi.org/10.1016/j.sigpro.2014.10.038.

[8] W. Mohguen and R. E. Bekka, "Empirical Mode Decomposition Based Denoising by Customized Thresholding," *International Journal of Electronics and Communication Engineering*, vol. 11, no. 5, pp. 519–524, Mar. 2017.

[9] W. Mohguen and R. E. Bekka, "New Denoising Method Based on Empirical Mode Decomposition and Improved Thresholding Function," *Journal of Physics: Conference Series*, vol. 787, Jan. 2017, Art. no. 012014, https://doi.org/10.1088/1742-6596/787/1/012014.

[10] W. Mohguen and R. E. Bekka, "An Empirical Mode Decomposition Signal Denoising Method Based on Novel Thresholding," presented at the 5th International Conference on Control & Signal Processing (CSP-2017), Kairouan, Tunisia, Oct. 2017.

[11] M. V. Sarode and P. R. Deshmukh, "Image Sequence Denoising with Motion Estimation in Color Image Sequences," *Engineering, Technology & Applied Science Research*, vol. 1, no. 6, pp. 139–143, Dec. 2011, https://doi.org/10.48084/etasr.54.

[12] W. Helali, Z. Hajaiej, and A. Cherif, "Real Time Speech Recognition based on PWP Thresholding and MFCC using SVM," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6204–6208, Oct. 2020, https://doi.org/10.48084/etasr.3759.

[13] M. Atif, Z. H. Khand, S. Khan, F. Akhtar, and A. Rajput, "Storage Optimization using Adaptive Thresholding Motion Detection," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6869–6872, Apr. 2021, https://doi.org/10.48084/etasr.3951.

[14] A. O. Boudraa and J.-C. Cexus, "EMD-Based Signal Filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, Dec. 2007, https://doi.org/10.1109/TIM.2007.907967.

[15] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994, https://doi.org/10.1093/biomet/81.3.425.

[16] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995, https://doi.org/10.1109/18.382009.

[17] M. Rakshit and S. Das, "An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter," *Biomedical Signal Processing and Control*, vol. 40, pp. 140–148, Feb. 2018, https://doi.org/10.1016/j.bspc.2017.09.020.

[18] B.-J. Yoon and P. P. Vaidyanathan, "Wavelet-based denoising by customized thresholding," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2004, vol. 2, https://doi.org/10.1109/ICASSP.2004.1326410.

[19] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 01, no. 01, pp. 1–41, Jan. 2009, https://doi.org/10.1142/S1793536909000047.

[20] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000, https://doi.org/10.1161/01.CIR.101.23.e215.

[21] M. Elgendi, M. Jonkman, and F. D. Boer, "Improved QRS Detection Algorithm using Dynamic Thresholds," *International Journal of Hybrid Information Technology*, vol. 2, no. 1, pp. 65–80, Jan. 2009.

A New Optimization Approach for a Solar Tracker Based on an Inertial Measurement Unit

Mohamed Redha Rezoug
Department of Electrical Engineering
University Kasdi Merbah Ouargla
Ouargla, Algeria
redha77.rmr@gmail.com

Mahdi Benaouadj
Department of Electrical Engineering
University Kasdi Merbah Ouargla
Ouargla, Algeria
benaouadjmahdi@gmail.com

Djamel Taibi
Department of Electrical Engineering
University Kasdi Merbah Ouargla
Ouargla, Algeria
taibi.djamel@yahoo.fr

Rachid Chenni
MoDERNa Laboratory
University of Constantine 1
Constantine, Algeria
rachid.chenni@gmx.fr

Abstract—Improvements and applications of Inertial Measurement Unit (IMU) sensors have increased in several areas. They are generally used in equipment that measures orientation, gravitational force, and speed. Therefore, in this paper, we worked on improving the performance of IMU in an application on solar trackers of the Kalman filter. This work illustrates the design of an autonomous device with astronomical control of a photovoltaic (PV) panel, allowing the optimization of the orientation/energy gain ratio. The device is based on two concepts at the same time, the modeling of the solar trajectory adopted by an algorithm which calculates continuously the solar angles (elevation and azimuth) and the approval of these by the IMU in order to sweep away any climatic fluctuations and thus allow an almost perfect adjustment relative to the perpendicular axis of the rays. The tracking system is based on two joints controlled by an Arduino control board. Experiments have shown a better performance of the two-axis device: the net energy gains can be around 34% with an additional 1.1% when the Kalman filter is applied.

Keywords-astronomical control; orientation; energy gain; optimization; inertial measurement unit

I. INTRODUCTION

Solar trackers are devices that allow the orientation of daylight reflectors regardless of their nature (photovoltaic (PV) solar panels, concentrated solar reflectors, or a lens focused towards the sun). The position of the sun in the sky varies with the seasons (elevation) and the time of day as the sun moves across the sky [1]. Most solar powered pieces of equipment serve us best when pointed at the sun [2, 3]. In the literature, we find many works focused on the knowledge of several astronomical and atmospheric quantities, whereas the use of these in algorithms makes the parameters difficult to implement when it comes to simple solar tracking systems [4]. A solar tracker can increase the efficiency of such equipment at any fixed position at the cost of additional system complexity [5-7]. There are many types of solar trackers of varying cost, sophistication, and performance [8, 9]. The accuracy of the solar tracker depends on the application [10-16]. Usually, PV systems do not work properly without monitoring, therefore at least one-axis tracking is mandatory [17].

In our work, we tried to structure a new two-axis tracking approach in which we aimed to orient the panel in real time towards the sun, in order to be placed in optimal position in relation to the incidence of solar radiation (perpendicular radiation if possible) because, throughout the day and depending on the season, the position of the sun varies constantly and in a different way, depending on latitude and longitude. First, we modeled the position of the sun by calculating each location on the earth at any time of day and determined the periods of daily sunshine as a function of time and place in order to get maximum benefit from the solar energy. Sunrise and sunset are defined as the instants when the upper limb of the sun's disk has just touched the horizon. This corresponds to an altitude of about zero degrees for the sun. The calculation of the position of the sun in our algorithm uses a certain coordinate system where time and location are given as input while the coordinates of the reference ground are given as output and this using a GPS sensor responsible for geolocation and timing. Real-time adaptation has the effect of substantially increasing the amount of reception and energy output of the PV panel as a second goal in our work by focusing on the layout, use, and angular correction of panel deflection due to the displacement of the sun by special Inertial Measurement Unit (IMU) sensors that empower the reading and inform us of all the surface angles of a solar panel in relation to the horizontal and vertical axes of the platform in a 3D aspect. For additional precision, correction by Kalman filter was carried out correcting the internal angular error. The correction was made with two actuators which are excited and

---

Corresponding author: Mohamed Redha Rezoug driven by a driver through an Arduino card. The experiments have shown better performance of the two-axis device by using IMU, especially through the Kalman filter, the net energy gains of which prove the efficiency of the proposed system.

II. MODELING THE PATH OF THE SUN

The position of any celestial object is described in two angles, the azimuth angle and the elevation angle. The elevation angle is the angle of the object above the horizon from the point of observation while the azimuth angle has diverse definitions. Figure 1 shows the concept of azimuth (γ) and elevation angles (α).

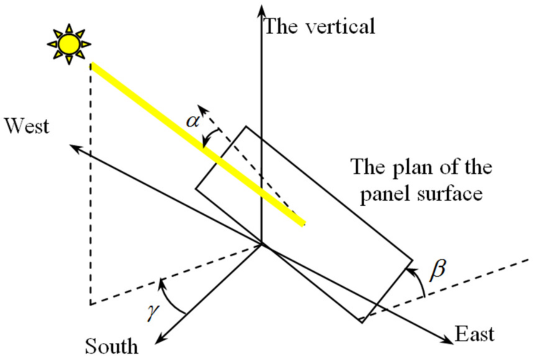


Fig. 1. Overview of tracking angles.

The calculation of the position of the sun in terms of azimuth and elevation angles and the sunshine duration requires the calculation of certain parameters such as Local Solar Time (LST) and Local Time (LT). The 12:00 hours LST is defined as the highest position of the sun in the sky. LT usually varies from LST due to the eccentricity of Earth's orbit and due to human adjustments such as time zones and daylight saving time.

A. Local Standard Time Meridian

The Local Standard Time Meridian (LSTM) is a reference meridian used for a particular time zone and is similar to the Prime Meridian which is used for Greenwich Mean Time (GMT) which can be calculated by the following equation:

$$LSTM = 15° \times \Delta T_{GMT} \quad (1)$$

B. The Equation of Time

The hour lines are the traces of the hourly plane which is perpendicular to the equator on the surface of the dial. The solar time measured and established by a dial is counted on the equator by the daily displacement of the hourly plane of the sun. One hour has passed when the direction of the sun's hourly plane has changed by 15°. The sundial does not take into account the declination of the sun, therefore, whatever the date of the year, the position of the time plane one hour after noon will be the same. Thus, 1 pm the time line will be identical for every day of the year. This is the principle of the polar style dials in which the duration of a solar hour is not constant throughout the year. Although the variation is minimal, on the order of a few seconds, the essential consequence is that these variations accumulate to form a larger gap. And during the year, the moment of the passage of the sun to the meridian (solar noon) slowly shifts day after day compared to noon which will be indicated by a perfect and regular clock. Our clock shows the average time. The annual variation in the difference between solar noon and mean noon is called the equation of time. By definition:

$$AverageTime = RealTime + EquationofTime \quad (2)$$

Equation (2) is periodic from year to year. It has a maximum of more than 14 minutes at the beginning of February and a minimum of -16 minutes at the beginning of November. It also has an intermediate maximum and minimum. It becomes zero 4 times a year: around April 15, June 15, September 1 and December 24. The exact value of the equation of time can be calculated at any moment with all the precision required by formulas presented in various works of astronomy. There are also simplified formulas and tables sufficient for the needs of gnomonics [18] but with the current power of PCs, it is better to use formulas as exact and complete as possible. It takes a little longer, but at least the calculated values are accurate within a few seconds. The equation of time (in minutes) is an empirical equation that corrects the eccentricity of the earth's orbit and the axial tilt of the earth is:

$$EoT = 9.87\sin(2B) - 7.53\cos(B) - 1.5\sin(B) \quad (3)$$

where $B = \frac{360}{365}(d-81)$, where d is the day of the year (i.e. d = 1 for January 1st, d = 32 for February 1st, etc.).

The precise Time Correction (TC) factor (in minutes) explains the variation in LST in a given time zone due to variations in the longitude of the location in the time zone and also incorporates the EoT as in (4). Number 4 in the equation shows the number of minutes the earth takes to rotate by one degree.

$$TC = 4 \times (Longitude - LSTM) + EoT \quad (4)$$

The LST can be calculated using (4) and LT as follows:

$$LST = LT + \frac{TC}{60} \quad (5)$$

C. The Angles of Solar Corrections

1) The Hour Angle (HRA)

As the first correction angle, the HRA presents the conversion of LST to the degrees with which the sun moves across the sky. By definition, the HRA is 0° at solar noon. Each hour from solar noon corresponds to an angular movement of the sun by 15°. This step of 15° presents the rotation of the earth during one hour. In the morning, the hour angle is negative.

$$HRA = 15° \times (LST - 12) \quad (6)$$

2) Declination Angle (δ)

The declination is the angle that the sun makes at the maximum of its course (solar noon) in relation to the equatorial plane. It varies between -23.45° on December 21 and +23.45° on June 21. For this, it is important to establish an equation relating the position of the sun in relation to the perpendicular to the plane of the PV module. The declination angle can be calculated by [7]:

$$\delta = 23.45 \times \sin\left(2\pi \frac{284+d}{365}\right) \quad (7)$$

3) Elevation Angle (α)

The elevation angle varies throughout the day. It also depends on latitude and date. The maximum elevation angle is an important parameter in the design of PV systems. This angle is measured at solar noon and depends on the angle of latitude and declination. The formula for the maximum elevation angle is:

$$\alpha = 90 + \phi - \delta \quad (8)$$

When (8) gives a number greater than 90°, the result is subtracted from 180°. This means that the sun at solar noon is coming from the south, as is typical of the northern hemisphere. ϕ is the latitude of the location of interest (with a positive sign for the northern hemisphere and negative for the southern hemisphere) and δ is the declination angle.

D. Sunrise and Sunset

In order to calculate the sunrise and sunset for a given location, altitude zeroing must be applied and the altitude equations are reformulated to give:

$$H_{sunrise} = 12 - \frac{1}{15°}\cos^{-1}\left(\frac{-\sin\varphi \times \sin\delta}{\cos\varphi \times \cos\delta}\right) - \frac{TC}{60} \quad (9)$$

$$H_{sunset} = 12 + \frac{1}{15°}\cos^{-1}\left(\frac{-\sin\varphi \times \sin\delta}{\cos\varphi \times \cos\delta}\right) - \frac{TC}{60} \quad (10)$$

All the calculations mentioned above were developed in the LabVIEW environment by the use of command blocks and micro programs. Figure 2 shows the overall code for modeling the path of the sun in the form of command frames designated as outputs from an Arduino board. This map takes over from correcting the angular deviation (elevation and azimuth) caused by the daily and seasonal variation of the position of the sun.

The selected IMU is a HMC5883L magnetic field sensor (gyroscope). This module encompasses in a single circuit all the functions necessary for the realization of a complete IMU. It integrates three axes for the two elements, the accelerometer and the magnetometer. It simultaneously measures the tracked movement and combines it with perfect synchronization with the clock. With the collaboration of the sensors, it performs an almost absolute position tracking without giving in to external transmission or emission elements. However, the majority of IMUs are exposed to negative distortion from external magnetic fields which affect the accuracy regardless of their nature. We therefore applied a correction by a Kalman filter in order to make the IMU more exact.

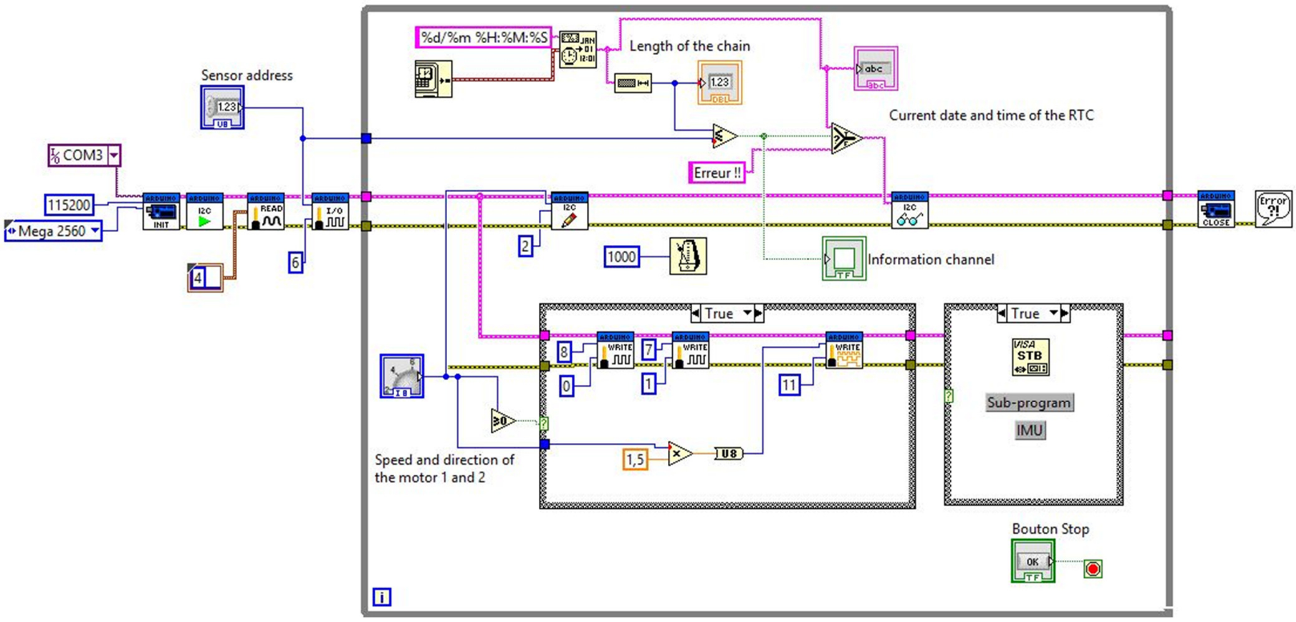


Fig. 2. Global diagram of the tracking system and the modeling of the trajectory in LabVIEW.

III. DESCRIPTION OF THE TRACKING SYSTEM

The tracking system corrects the angular deviation (elevation and azimuth) due to the apparent movement of the sun. It controls the movement of a PV module on two axes. The first has a 360° rotation capability, it is considered the main support of the whole mechanism. The second has a 90° rotation capability and is responsible for the rotation of the PV.

To keep the solar radiation perpendicular to the active surface of the PV module, the automation of the trajectory of the sun is defined thanks to the equations of time translated into programs in LabVIEW with the cooperation of the Arduino board. The control system is shown schematically in Figure 3. It consists of two parts: the operative part and the control part. The operative part has two joints actuated by two jacks.

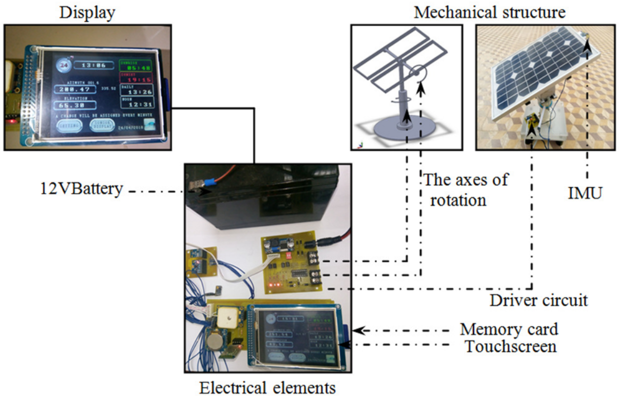


Fig. 3. General description of the system.

A. The Operative Part

The design of tracking systems is usually designed by one axis or by two axes. To amplify the energy gain of the PV module, two-axis trackers are the ideal solution. The choice of motors is essential, so we chose motors that have axial encoders in order to know at any time the exact angle of the motors after each movement. A powerful motor rotates 360° and presents the overall support of the module. It is responsible for correcting the azimuth deviation. The weight of the joint does not present any movement constraint on the motor because it is mounted axially on the latter using a bearing. The second motor allows the angular correction of the elevation. These motors are commonly used in the home automation field.

B. The Order Part

The control device is essentially based on multisensory technology in order to collect at any time any information that enters directly or indirectly into the calculations of the path of the sun and the angles of deviation which correct the deviation through the equations of time. The electronic IMU devices measure and report the specific strength of a body. The angular rate and sometimes the magnetic field surrounding the body and the use of a combination of accelerometers and gyroscopes or magnetometers as criteria of choice led us to use them to optimize the energy gain by orientation of the PV module. For a near-perfect PV initialization, a gyroscope is used.

IV. MAGNETOMETRIC SENSOR MODEL

Magnetic field measurement is subjected to two types of distortion: hard iron distortion and soft iron distortion. To obtain precise measurements, a mathematical correction of the scale factor, misalignment, and bias errors by a Kalman filter is applied to the model matrix. This process involves finding a transformation and bias matrix in order to obtain corrected measurements of the magnetic field. In this algorithm, a bias is applied to the vector of initial magnetometric data (*X*, *Y*, *Z* coordinates) which is then multiplied by the resulting vector. We then subtract the bias vector from the result as shown in (11). This was done by the intervention of an Arduino card in the LabVIEW environment.

$$\begin{bmatrix} X_{filtered} \\ Y_{filtered} \\ Z_{filtered} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \times \left(\begin{bmatrix} X_{initial} \\ Y_{initial} \\ Z_{initial} \end{bmatrix} - \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} \right) \quad (11)$$

We used intrinsic active rotations of the magnetometer module around the (*X*, *Y*, *Z*) axes in a three-dimensional plane. We created 3 angles which present respectively pitch, roll, and yaw. We thus obtained 12 probabilities as indicated in Figure 4.

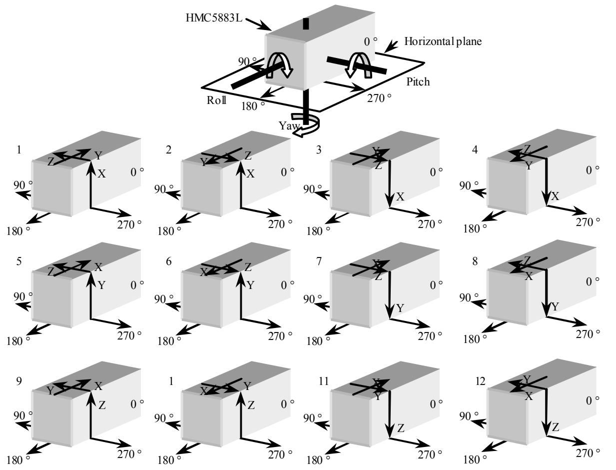


Fig. 4. Axial probabilities of the IMU on the three-dimensional plane.

In its initial position, the magnetometer module coordinate system and the panel coordinate system are aligned with each other. If we want to rotate the axis of the panel, we would perform the three elementary rotations as follows:

- Rotation by the roll angle around the *X* axis:

$$\begin{cases} X_1 = X_0 \\ Y_1 = Y_0 \times \cos(roll) - Z_0 \times \sin(roll) \\ Z_1 = Y_0 \times \sin(roll) + Z_0 \times \cos(roll) \end{cases} \quad (12)$$

- Rotation by the pitch angle around the *Y* axis:

$$\begin{cases} X_2 = X_1 \times \cos(pitch) - Z_1 \times \sin(pitch) \\ Y_2 = Y_1 \\ Z_2 = -X_1 \times \sin(pitch) + Y_1 \times \cos(pitch) \end{cases} \quad (13)$$

- Rotation by the yaw angle around the *Z* axis:

$$\begin{cases} X_3 = X_2 \times \cos(yaw) - Y_2 \times \sin(yaw) \\ Y_3 = X_2 \times \sin(yaw) + Y_2 \times \cos(yaw) \\ Z_3 = Z_2 \end{cases} \quad (14)$$

The numbering 0 on the *X*, *Y*, and *Z* coordinates presents the starting point. The numberings 1, 2 and 3 respectively present 3 elementary rotations (the first, the second, and the third). It is possible to return to Euler angles [19] from a rotation matrix. The equation varies in correspondence with the order of execution of the elementary rotations of Euler angles and the Euler angles of Tait-Bryan [20] which correspond to the order yaw-pitch-roll with corrections by a Kalman filter. We integrate (11) to obtain the angles as follows:

$$Yaw = \tan^{-1}\left(\frac{M_{21}}{M_{11}}\right) = atan2\left(M_{21}, M_{11}\right) \quad (15)$$

$$Pitch = -\sin^{-1}(M_{31}) = -a\sin(M_{31}) \quad (16)$$

$$Roll = \tan^{-1}\left(\frac{M_{32}}{M_{33}}\right) = atan2\left(M_{32}, M_{33}\right) \quad (17)$$

These were deduced using the fact that the magnetometer module rotates back to front and the local field vectors move inward. Using this information along with knowing the orientation of the IMU allows the formation of a direction cosine matrix which is used to convert the measurements from the IMU coordinate frame to the panel coordinate frame. We can present the order of transformations as follows:

- Turn the angle α +90° around the X axis to align the Z axis of the IMU with the Z axis of the panel support.

$$T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \quad (18)$$

- -Rotate β +90° around the Z axis to align the coordinate frames of the IMU and the panel support.

$$T_2 = \begin{bmatrix} \cos\beta & \sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (19)$$

Thus, the complete transformation of the IMU into coordinates of the panel support is given by the multiplication of the equation:

$$T = \begin{bmatrix} -\sin\beta & \sin\alpha.\cos\beta & \cos\alpha.\cos\beta \\ \cos\beta & \sin\alpha.\sin\beta & \cos\alpha.\sin\beta \\ 0 & \cos\alpha & -\sin\alpha \end{bmatrix} \quad (20)$$

The similarity of movement of the angles of our problematic which is the movement of a PV panel around its axes at the first two angles (yaw and pitch) led us to make a correspondence of the azimuth and elevation angles with the latter.

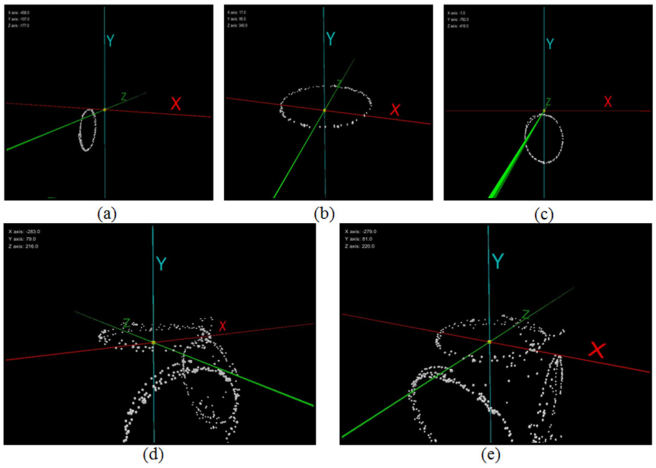

(a) (b) (c) (d) (e)

Fig. 5. Center of circle before calibration and filtering separately, (a) on the X axis, (b) on the Y axis, (c) on the Z axis, (d) and (e) on the XYZ kernel.

Figure 5 shows that the center of the circle after hard iron distortions is now at Y=100 and X=200. There is therefore a polarization of hard iron on the X axis of 200mG and a polarization of iron hard on the Y axis of 100mG. These distortions move the center of the circle away from its original position but preserve its original shape.

Soft iron distortions on their part distort existing magnetic fields. Soft iron distortions are easy to identify in a magnetic output graph because they distort the circular output. The soft iron distortions distort the circle and give it an elliptical shape. An illustration of the magnetometer output with the two distortions on the X and Y axes in order can be seen in Figure 6.

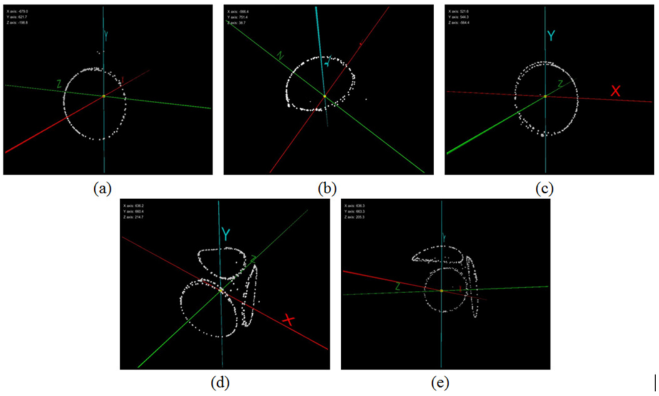

(a) (b) (c) (d) (e)

Fig. 6. Center of circle after calibration and filtering separately (a) on the X axis, (b) on the Y axis, (c) on the Z axis, (d) and (e) on the XYZ kernel.

A circle which does not undergo any distortion must be centered on X=0, Y=0 and its radius must be equal to the magnitude of the magnetic field.

Figure 7 shows the application of the Kalman filter with block-chained logic under the LabVIEW environment.

V. GENERAL STRUCTURE OF THE DEVICE

The produced system is an autonomous, automatic, and pre-programmed solar tracker based in its design on an Arduino Mega embedded system which is capable of astronomically resetting the local parameters of a PV installation by geolocation and instantly correcting with a new approach of the IMU the angular fluctuations (azimuth and elevation) caused by the displacement of the sun using a preprogrammed algorithm and pulse control in order to minimize the power consumed by the actuators and to increase the luminous flux on the active surface of the panel.

Usually, IMUs are used in the field of aeronautics, especially drones. They allow the collaboration with a GPS receiver when the GPS signals are absent as is the case in tunnels. In order to collect as much information as possible to geolocate a photovoltaic installation, we have chosen the GY-NEO6M type GPS compatible with the master Arduino Mega card. It can be synchronized with an RTC of our system or operate autonomously. The modeling of the above mentioned path of the sun requires knowledge of a few parameters in order to determine the angles of azimuth and elevation.

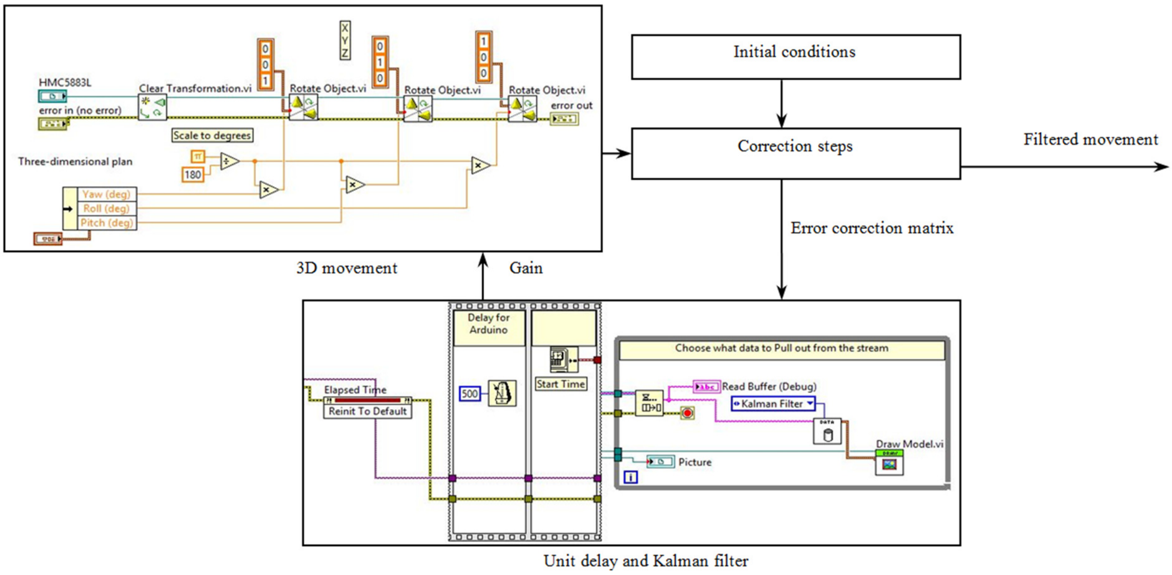


Fig. 7. Application of the Kalman filter in LabVIEW.

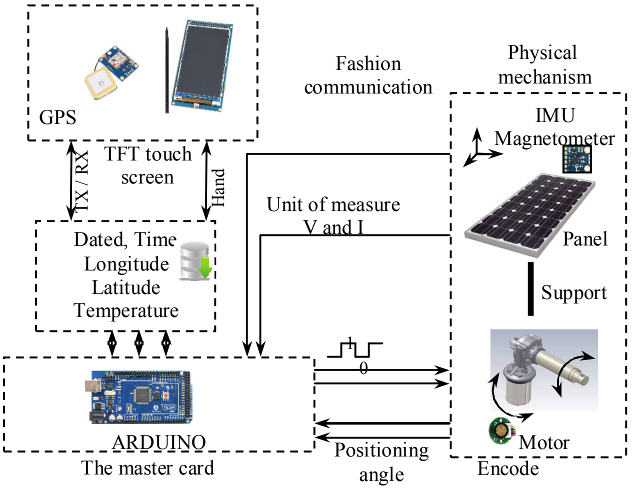


Fig. 8. General descriptive diagram of the system.

An internal clock updates the GPS in order to detect the local coordinates as well as the date and the time required in this calculation. The IMU works by sensing linear acceleration using the internal accelerometer and the rotation rate using the HMC5883L at each angular correction (azimuth or elevation). So, it is updated with the next angles and the current angles. These data are stored in the Arduino in an EEPROM to be compared to the angles of the motor encoders to confirm their values and to be used as correction steps through power stages that reposition the equipment. To access the internal parameters that are introduced into the initialization and configuration of the tracking system or the modeling of the trajectory, a touch screen can be used which does this through several actions. Figure 8 summarizes the general design of the produced device.

A. *GPS Communication*

Modeling the path of the sun requires the intervention of a few quantities. The GY-NEO6M uses geolocation by the World Geodetic System (WGS 84) which is used in mapping and navigation including satellite navigation systems. The geographic coordinates of the PV system are quantities (latitude, longitude, and time) which define its position on the installation surface. Latitude values collected by the system and longitude values are the considered entrances. Through two pins of the Arduino board (TXD_1 and RXD_1) the GPS communicates at a speed of 9600bps. More parameters are shown in Table I.

TABLE I. TECHNICAL PARAMETERS

Parameter	**Value**
Power supply	3V / 5V
Model	GPS-NEO-6M-001
Antenna	Ceramic antenna
Battery	Rechargeable battery back-up
Signal light	LED light
Antenna size	25×25mm
Model size	25.5mm×31.5mm
Mounting hole	2mm
Default baud rate	9600
Default output	Compatible with NMEA0

B. *Tactile Communication*

In order to simplify the communication between the user and the Arduino and to minimize the use of electronic components (push buttons), a touch screen type (320×240 TFT LCD) was used in the constitution of our device. Its role is also to display any action applied on the Arduino. A multi-screen menu has been created to organize the control and display of the various parameters which can be the measured quantities or different input and output data. The main purpose of this tactile communication is to access the internal parameters of the quantities recorded previously such as the initial variables stored on the Arduino EEPROM and change them according to our needs. This adjustment is considered as a manual mode (all the quantities are inserted manually without the use of the sensors). This choice is provided as an emergency solution if auto locate fails.

VI. EXPERIMENTAL TEST

In order to demonstrate the effectiveness of this new approach, we first carried out a comparison of the results obtained on a fixed module with respect to another module oriented by two axes based on the new approach while correcting the variation related to elevation and azimuth. These two angles are directly linked to the day, night, and seasonal variation under the same operating conditions. A further comparison was made between the results obtained by the technique without and with the Kalman filter. These experiments were carried out in a day where the illumination was relatively good with some disturbances in the afternoon in the eastern Algeria (Ain Beida) defined by 35.796 N altitude and 7.393 E longitude with a varied temperature. For a daily correction, a series of successive orders to the equipment at equal intervals during the period of sunshine carries the angular correction of deviation which varies from 1.5 to 2 degrees. The command information is extracted from the database which is calculated by the Arduino and stored in the EEPROM of its microcontroller. Voltage and current measurements were conducted at equal time intervals. The Copernicus equation gives us an initial declination of the PV module of 14.81°. The LabVIEW environment provides us with graphical interfaces to display the different quantities collected experimentally by the card (current, voltage, and power). The results are shown in Figures 9-13. In Figures 9-11, we obtained from the LabVIEW the responses represented in blue for the measured quantities of a tracking system without Kalman filter and in red for those of a fixed system.

In Figure 9, we see that the oriented system generates a current intensity that exceeds three times the current obtained by the fixed system at the start of the day because the panel has already been initialized in a perpendicular position with respect to the solar rays, so this system stabilizes at a maximum current value. We also notice disturbances caused at the end of the day by a cloud passage. During this time, the fixed system starts with a low current which gradually increases towards the optimal period of sunshine then decreases (knowing that the panel is fixed at a tilt angle of 32°). The current generated by the stationary panel always remains lower than that generated by the oriented panel. This validates the significant influence of irradiation on the current and more precisely that of the angles formed by the rays and the panels. It is always favorable to keep an angle of 90° with the panels for a total absorption of the rays without losses and without refractions which leads us to choose the most precise oriented systems.

In Figure 10, we see that the voltage supplied by the oriented module at the start of the day is high but decreases regularly due to the influence of the temperature on the voltage of the panel caused by the fact that the active surface of the panel remains perpendicular to solar radiation while the voltage supplied by the fixed module at the start of the day is low and remains constant with some variations and disturbances and exceeds the oriented module a few times.

Figure 11 shows the power obtained during this day by the two panels which represent the product of the voltage and the current. We notice the effect of the orientation technique on the efficiency which is a gain of 34% in comparison with the fixed system.

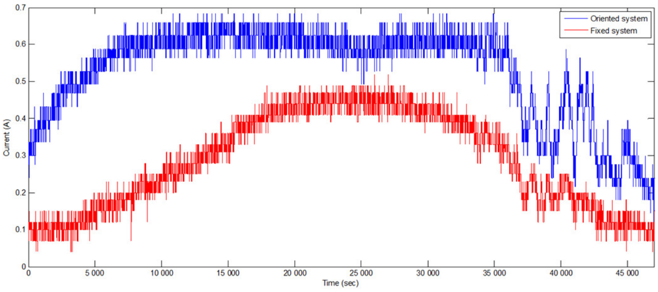


Fig. 9. Current-time curves for a fixed and an oriented module.

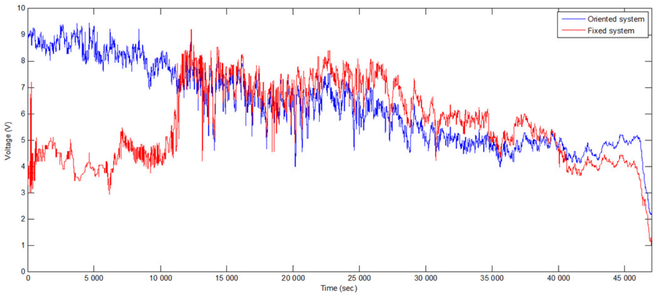


Fig. 10. Voltage-time curves for a fixed and an oriented module.

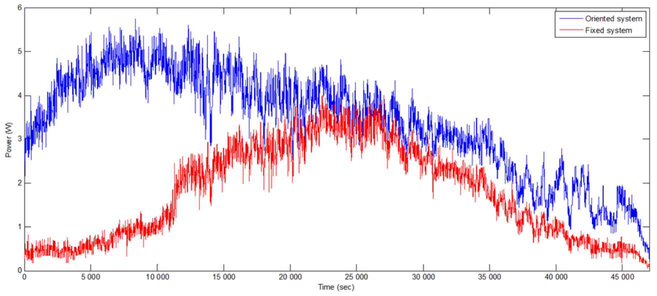


Fig. 11. Power-time curves for a fixed and an oriented module.

In Figure 9, we can see the gain obtained by the tracking system, manifested by the difference between the two schemes. This gain was obtained thanks to the angle detection precision of the IMU. This reflects the influence of this technique on the efficiency of a PV panel. To remedy the errors caused by the external magnetic effects on the IMU sensors, a Kalman filter is applied. In order to show the effect of the latter on the yield, two comparative experimental tests were carried out on a filtered and an unfiltered system during two different days to generalize the results. We plot the curves of the measured quantities (I and V) using LabVIEW. Figures 12 and 13 show the power plot obtained by the tracking system with (blue) and without (red) the Kalman filter. The difference between the plots in the same chart in Figures 12 and 13 is respectively a gain of 1.1% and 1.2%. This proves the advantage that this new technique gains with the application of the Kalman filter on the behavior of the IMU sensor.

In [13], a comparative study between a fixed system and another two-axis oriented system using a 120W PV panel showed an average daily energy gain of 112.5Wh. According to the energy aspect and over the same period, we arrived to obtain an average daily energy gain of 403.2Wh (since the PV panel used in [13] is 12 times greater than the one opted in the current work). So, the developed new optimization approach for a solar tracker based on an IMU is most suitable for two-axis tracking systems.

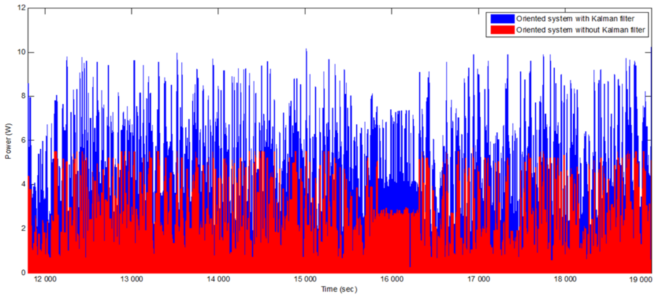


Fig. 12. Power curves during the first day.

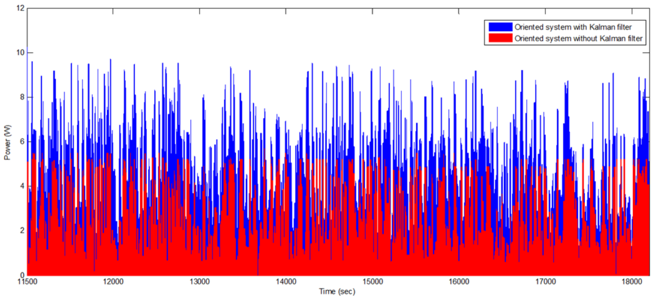


Fig. 13. Power curves during the second day.

VII. CONCLUSION

Due to the poor performance of fixed panels resulting from the movement of the sun and energy losses, the main goal of the researchers in this field is to track the variant trajectory of the sun and to optimize the techniques used while taking into account the materials. Various methods have been employed to remedy this. In the current work, a preprogrammed technique using different equations of time was adopted which includes the use of IMU modules. These new components are characterized by their precision of positioning identification along three axes. This helps us to collect more accurate data on the position of the panels and thus to better control them to follow the sun. However, these components operate using magnetic fields and are highly sensitive. They are therefore often influenced by neighboring magnetic fields and end up showing calculation errors, which cause energy losses. To solve this problem, we used an additional corrector, which is the Kalman filter. It corrects deviations caused by magnetic fields. According to the obtained results, a tracking technique supported by IMU is much more advantageous than a fixed PV system and the addition of a Kalman filter gives us more stable results and a higher efficiency than the filterless system (about 0.8%). This last technique is thus chosen for the considerable energy gain that it offers.

APPENDIX

ABBREVIATIONS

IMU	Inertial Measurement Unit
GPS	Global Positioning System
3D	Three Dimensional
PC	Personal Computer
LabVIEW	Laboratory Virtual Instrumentation Engineering Workbench
PV	Photovoltaic
RTC	Real Time Clock
EEPROM	Electrically Erasable Programmable Read-Only Memory
TXD or TX	Transmit Data
RXD or RX	Receive Data
LED	Light Emitting Diode
TFT	Thin Film Transistor
LCD	Liquid Crystal Display
I	Current intensity
V	Voltage

REFERENCES

[1] S. Abdallah and S. Nijmeh, "Two axes sun tracking system with PLC control," *Energy Conversion and Management*, vol. 45, no. 11, pp. 1931–1939, Jul. 2004, https://doi.org/10.1016/j.enconman.2003.10.007.

[2] A. A. Rizvi, K. Addoweesh, A. El-Leathy, and H. Al-Ansary, "Sun position algorithm for sun tracking applications," in *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, Dallas, TX, USA, Oct. 2014, pp. 5595–5598, https://doi.org/10.1109/IECON.2014.7049356.

[3] B. Krishna and K. Sinha, "Tracking of Sun for Solar Panels and Real Time Monitoring Using LabVIEW," *Journal of Automation and Control Engineering*, vol. 1, no. 4, pp. 312–315, 2013, https://doi.org/10.12720/joace.1.4.312-315.

[4] T.-C. Cheng, W.-C. Hung, and T.-H. Fang, "Two-Axis Solar Heat Collection Tracker System for Solar Thermal Applications," *International Journal of Photoenergy*, vol. 2013, Nov. 2013, Art. no. e803457, https://doi.org/10.1155/2013/803457.

[5] K.-K. Chong *et al.*, "Integration of an On-Axis General Sun-Tracking Formula in the Algorithm of an Open-Loop Sun-Tracking System," *Sensors*, vol. 9, no. 10, pp. 7849–7865, Oct. 2009, https://doi.org/10.3390/s91007849.

[6] S. M. Çinar, F. O. Hocaoğlu, and M. Orhun, "A remotely accessible solar tracker system design," *Journal of Renewable and Sustainable Energy*, vol. 6, no. 3, May 2014, Art. no. 033143, https://doi.org/10.1063/1.4885099.

[7] D. Ciobanu and C. Jaliu, "Step tracking program for concentrator solar collectors," *IOP Conference Series: Materials Science and Engineering*, vol. 147, Aug. 2016, Art. no. 012149, https://doi.org/10.1088/1757-899X/147/1/012149.

[8] R. Eke and A. Senturk, "Performance comparison of a double-axis sun tracking versus fixed PV system," *Solar Energy*, vol. 86, no. 9, pp. 2665–2672, Sep. 2012, https://doi.org/10.1016/j.solener.2012.06.006.

[9] M. Engin and D. Engin, "Optimization Controller for Mechatronic Sun Tracking System to Improve Performance," *Advances in Mechanical Engineering*, vol. 5, Jan. 2013, Art. no. 146352, https://doi.org/10.1155/2013/146352.

[10] H. Eric Tseng, L. Xu, and D. Hrovat, "Estimation of land vehicle roll and pitch angles," *Vehicle System Dynamics*, vol. 45, no. 5, pp. 433–443, May 2007, https://doi.org/10.1080/00423110601169713.

[11] F. M. Hoffmann, R. F. Molz, J. V. Kothe, E. O. B. Nara, and L. P. C. Tedesco, "Monthly profile analysis based on a two-axis solar tracker proposal for photovoltaic panels," *Renewable Energy*, vol. 115, pp. 750–759, Jan. 2018, https://doi.org/10.1016/j.renene.2017.08.079.

[12] R. Velicu and M. Lates, "Torques on Rotational Axes of PV Azimuthal Sun Tracking Systems," in *Sustainable Energy in the Built Environment - Steps Towards nZEB*, Cham, 2014, pp. 461–470, https://doi.org/10.1007/978-3-319-09707-7_34.

[13] Y. Rambhowan and V. Oree, "Improving the dual-axis solar tracking system efficiency via drive power consumption optimization," *Applied Solar Energy*, vol. 50, no. 2, pp. 74–80, Apr. 2014, https://doi.org/10.3103/S0003701X1402011X.

[14] M. R. Rezoug and R. Chenni, "A Real-Time Optimization Approach to the Efficiency of a Photovoltaic System by a Hybrid Tracking Technique using LabVIEW," *American Journal of Applied Sciences*, vol. 15, no. 5, pp. 267–277, May 2018, https://doi.org/10.3844/ajassp.2018.267.277.

[15] Y. Rizal, S. H. Wibowo, and Feriyadi, "Application of Solar Position Algorithm for Sun-Tracking System," *Energy Procedia*, vol. 32, pp. 160–165, Jan. 2013, https://doi.org/10.1016/j.egypro.2013.05.021.

[16] M. J. Ya'u, "A Review on Solar Tracking Systems and Their Classifications," *Journal of Energy, Environmental & Chemical Engineering*, vol. 2, no. 3, pp. 46–50, Aug. 2017, https://doi.org/10.11648/j.jeece.20170203.12.

[17] S. Yilmaz, H. Riza Ozcalik, O. Dogmus, F. Dincer, O. Akgol, and M. Karaaslan, "Design of two axes sun tracking controller with analytically solar radiation calculations," *Renewable and Sustainable Energy Reviews*, vol. 43, pp. 997–1005, Mar. 2015, https://doi.org/10.1016/j.rser.2014.11.090.

[18] S. Seme, G. Štumberger, and J. Voršič, "Maximum Efficiency Trajectories of a Two-Axis Sun Tracking System Determined Considering Tracking System Consumption," *IEEE Transactions on Power Electronics*, vol. 26, no. 4, pp. 1280–1290, Apr. 2011, https://doi.org/10.1109/TPEL.2011.2105506.

[19] B. Sun, Y. Wang, and J. Banda, "Gait Characteristic Analysis and Identification Based on the iPhone's Accelerometer and Gyrometer," *Sensors*, vol. 14, no. 9, pp. 17037–17054, Sep. 2014, https://doi.org/10.3390/s140917037.

[20] C.-C. Wei, Y.-C. Song, C.-C. Chang, and C.-B. Lin, "Design of a Solar Tracking System Using the Brightest Region in the Sky Image Sensor," *Sensors*, vol. 16, no. 12, Dec. 2016, Art. no. 1995, https://doi.org/10.3390/s16121995.

Parameter Selection to Ensure Multi-Criteria Optimization of the Taguchi Method Combined with the Data Envelopment Analysis-based Ranking Method when Milling SCM440 Steel

Nguyen Lam Khanh
Faculty of Mechanical Engineering
University of Transport and Communications
Hanoi, Vietnam
khanh_mxd@utc.edu.vn

Nguyen Van Cuong
Faculty of Mechanical Engineering
University of Transport and Communications
Hanoi, Vietnam
nguyencuong@utc.edu.vn

Abstract-SCM440 steel is a commonly used material for making plastic injection molds and components such as gears, transmission shafts, rolling pins, etc. Surface roughness has a direct influence on the workability and durability of the parts and/or components, while the Material Removal Rate (MRR) is a parameter that is used to evaluate the productivity of the machining process. Furnished products with small surface roughness and large MRR is the desired result by all milling processes. In this paper, the determination of the values of input parameters is studied in order to ensure that during the process of milling SCM440 steel, it will have the smallest surface roughness and the largest MRR. There are five parameters that are required to be determined, namely the cutting insert material, the tool nose radius, the cutting speed, the feed rate, and the cutting depth. The Taguchi method was applied to design the experimental matrix with a total of 27 experiments. Result analysis determined the influence of the input parameters on surface roughness and MRR. The Data Envelopment Analysis-based Ranking (DEAR) method was applied to determine the optimal value of the input parameters, which were used to conduct the milling experiments to re-evaluate their suitability.

Keywords-milling; SCM440 steel; surface roughness; MRR; optimization; Taguchi; DEAR

I. INTRODUCTION

Milling is a very common machining method in the mechanical engineering industry. It is considered to be the cutting method with the highest productivity. With the development of the cutting tool manufacturing technology as well as the emergence of the modern CNC machines, this method is capable of ensuring very high accuracy. Research of solutions that improve the accuracy of milling machines and milling cutters is continuously conducted [1]. Therefore, in some cases, this method is used as the final machining method for surfaces requiring high precision [2]. To make the most use of the achievements in the mechanical engineering technology and the cutting tool manufacturing technology, many researchers carried out experimental studies to determine the optimal value of parameters related to the machining process. The purpose of these studies is to determine the value of the machining process's parameters to ensure minimum surface roughness and maximum Material Removal Rate (MRR). This problem is known as milling operations optimization. When studying the milling operations optimization, many authors used the Taguchi method to design the experimental matrix. When comparing the Taguchi-based matrix design method with some other matrix design methods, it was found that the Taguchi method requires a smaller number of experiments. An advantage that only the Taguchi method possesses is that it can design a matrix with the input parameters being a qualitative (not a quantitative) parameter [3, 4].

In [5], the Taguchi method was applied to design the experimental matrix when milling the AA6082T6 material by a tungsten carbide cutting tool. Signal-to-Noise (S/N) ratio was analyzed to determine the optimal values of spindle speed, feed rate, and depth of cut to ensure the minimum value of surface roughness. The Taguchi method was also used to design the experimental matrix when milling AA6082T6 with PVD-coated and CVD-coated cutting tools. The determination of cutting speed, feed rate, and tool material type to ensure tool wear was similarly performed for surface roughness in [6]. In [7], the authors applied the Taguchi method to design the experimental matrix when milling D2 steel with the carbide inserts cutting tool. Cutting parameters including spindle speed, feed rate, and depth of cut were selected as input parameters for each experiment. The S/N ratio analysis method was applied to determine the optimal value of the cutting parameters to ensure the minimum value of surface roughness. When milling AISI P20 steel with the carbide inserts cutting tool, the authors in [8] used the Taguchi method to design an experimental matrix with spindle speed, feed rate, and depth of cut as input parameters [8]. They also used S/N ratio analysis to determine the optimal value of the input parameters to ensure the minimum value of surface roughness. To determine the optimal value of parameters including cutting speed, feed, radial depth, and axial depth to ensure the minimum value of surface roughness when milling 1.2738 steel with the WNHU 04T310 cutting tool

---

Corresponding author: Nguyen Van Cuong

(manufactured by Palbit), the authors in [9] also applied the Taguchi method to design the experimental matrix. The S/N ratio analysis method was applied to determine the optimal value of the cutting parameters. Authors in [10] also designed the experimental matrix according to the Taguchi method when milling 7075T6 aluminum alloy with an AlTiN PVD-coated cutting tool. The optimal values of cutting speed, feed rate, radial depth, and axial depth were also determined by the analysis of the S/N ratio. This study also aimed to ensure the minimum value of surface roughness. To determine the optimum value of cutting speed, feed rate, depth of cut, and coolant flow to ensure the minimum value of surface roughness when milling AISI 1040 MS steel with the carbide inserts cutting tool, the authors in [11] also designed the experimental matrix according to the Taguchi method and the S/N ratio analysis method was also used to determine the optimal value of the input parameters. Authors in [12] also designed the experimental matrix according to the Taguchi method when milling 7075 T6 aluminum alloy with a High-Speed Steel (HSS) cutting tool. They also applied the S/N ratio analysis method to determine two sets of optimal values of cutting speed, feed rate, and depth of cut, one set that ensures the smallest surface roughness and another set that ensures the largest MRR.

The experimental matrix design based on the Taguchi method has been successfully applied in a number of studies to ensure a certain criterion of the machining process. However, if only the Taguchi method is applied for the experimental design and the S/N ratio analysis to determine the optimal value of the machining process parameters, only one criterion of the machining process can be guaranteed. In order to resolve this shortcoming of the Taguchi method, many studies combined the Taguchi method with other methods to optimize the multi-objectives of the milling process: in [13], the Taguchi method has been combined with ANOVA to determine the values of the spindle speed, the feed rate and the cutting depth to ensure the minimum surface roughness and the maximum MRR when milling AISI 1005 steel with the TiN coated cutting tool. In [14], the Taguchi method and the Weighted Principal Component Analysis (WPCA) were combined to determine the milling type and the values of milling parameters to simultaneously ensure minimum surface roughness and maximum MRR when milling Al 6061 aluminum alloy with a high speed steel cutting tool. The Taguchi method and the Gray Relational Analysis (GRA) method were combined to determine the values of cutting speed, feed rate, and cutting depth to ensure simultaneously minimum surface roughness and maximum MRR when milling Inconel 718 super alloy by an uncoated tungsten carbide cutting tool in [15]. A combination of the Taguchi method, TOPSIS method, and ANOVA analysis was performed to determine the optimal values of cutting speed, feed rate, and depth of cut to ensure simultaneously the minimum value of surface roughness and maximum value of MRR when milling Ti-6Al-4V titanium alloy with TiN coated cutting tools in [16]. Through the above studies, it is shown that the cutting tool parameters are commonly selected as the input parameters of the milling experiment process. These parameters can be easily adjusted by the operators. However, to the best of our knowledge, there have been no studies that consider all the 5 parameters of the cutting tool material, i.e. the insert material, the tool nose radius, the cutting speed, the feed rate, and the depth of cut.

SCM440 steel is a type of steel used quite commonly to make plastic injection molds and components such as gears, transmission shafts, and rolling pins [17]. Due to the high content of Cr, Mo and Mn elements, this steel has a low thermal conductivity. When machining this steel, the tool wears out quickly, thus it is required to select the right cutting tool [18-20]. A number of studies on milling steels equivalent to this steel have been carried out [17-23]. However, to the best of our knowledge, there are no published studies on milling SCM440 (or equivalent) steel that consider all 5 parameters. DEAR is a method used for multi-criteria decision making that was introduced in 2002 [24]. This method has been used for the multi-objective optimization of the AISI 1055 steel turning process [25], the Ti-6Al-4V alloy turning process [26], the SAE420 steel grinding with a segmented grinding wheel [27], of the Electrical Discharge Machining (EDM) with the material type AA 6082 [26], etc. However, to the best of our knowledge, there have been no published studies on the application of this method in multi-criteria decision making for milling methods in general and for milling SCM440 steel in particular.

In this study, milling parameters such as the cutting tool material, the tool nose radius, the cutting speed, the feed rate, and the cutting depth will be determined to simultaneously ensure the optimization of two criteria, minimum surface roughness and maximum MRR when milling this type of steel. A combination of the Taguchi method and the DEAR method was used to solve this problem.

II. THE DEAR METHOD

The purpose of the experimental process of this study is to ensure that surface roughness (*Ra*) has the smallest value and the *MRR* reaches the maximum value. Thus, it is required to determine the values of the input parameters that ensure the set out objectives. The DEAR method will be applied in this study to carry out the above-stated work [24]. The DEAR method's steps are [24]:

- Determine the weight of each response for all experiments. This value is calculated as the ratio of the value of each response to the sum of all responses.
- Transfer the response data to the weight data by multiplying the observed data by their respective weight.
- Divide the inversed data by the sum of all inversed data.
- The Multi Response Performance Index (MRPI) is calculated by (1):

$$MRPI = W_{R_a} * R_a + W_{MRR} * MRR \quad (1)$$

The weights of the responses are calculated as:

$$W_{R_a} = \frac{R_a}{\sum R_a} \quad (2)$$

$$W_{MRR} = \frac{\frac{1}{MRR}}{\sum \frac{1}{MRR}} \quad (3)$$

III. EXPERIMENTS ON MILLING SCM400 STEEL

A. Experimental System

A CNC milling machine with the HAA5 serial was used to carry out the experiments (Figure 1). The experimental sample is SCM400 steel with length, width, and height of 80mm, 40mm, and 30mm respectively. The steel's chemical composition conducted with a spectrometer was: 0.43% C, 0.28% Si, 0.72% Mn, 1.05% Cr, 0.23% Mo, 0.024% P, and 0.026% S.

Fig. 1. The milling machine.

TABLE I. PARAMETERS OF THE CUTTING INSERTS

Parameter	**Cutting insert**		
	R390-11T303M-PM1025	**R390-11T305M-PM1025**	**R390-11T305M-PM1025**
Tool nose radius (mm)	0.3	0.5	0.8
Back edge length (mm)	0.8	0.9	1.2
Weight (kg)	0.0022	0.0026	0.003
Coating material	TiN; TiCN; TiAlN		
Cutting thickness (mm)	3.59		
Main cutting angle (degree)	90		
Maximum cutting depth (mm)	10		
Shape style of cutting piece	L		
Edge width (mm)	6.8		
Effective length of edge (mm)	10		

Three types of cutting pieces were used in the experiment namely the TiN-, TiCN-, and TiAlN-coated pieces. These cutting pieces have high thermal resistance, and have been proven to be very suitable for machining SCM440 steel. Each cutting piece was used with 3 tool nose radius values of 0.3mm, 0.5mm, and 0.8mm. On the tool shank with 12mm diameter, 2 symmetrical cutting pieces were installed. Each cutting insert was used only once for the purpose of eliminating the influence of tool wear on the output parameters of the milling process. In other words, the number of cutting inserts used in the experiment is twice the number of experiments to be carried out. The milling process has been carried out according to the method of symmetric milling, which means that the milling width was equal to the diameter of the milling cutter. Table I shows some basic parameters of the cutting inserts used.

The surface roughness was measured with a Mitutoyo - Japan SJ301 surface roughness tester of 0.8mm standard length. The surface roughness of each experimental sample was determined by averaging at least three consecutive measurements. The MRR was calculated according to:

$$MRR = V_f \cdot a_p \cdot b_w \ (\text{mm}^3/\text{min}) \qquad (4)$$

where V_f is the feed rate (mm/min), a_p is the cutting depth (mm), and b_w is the cutting width (mm). In this case the cutting width is just equal to the diameter of the milling cutter.

B. Experimental Design

The Taguchi method was applied to design the experimental process in this study. Five parameters were selected as the input parameters of the experimental process,. Each selected parameter has three levels of values (corresponding to three encoding degrees of 1, 2, and 3). The values of the experimental parameters, selected within their range as recommended by the cutting tool manufacturer [29], are shown in Table II.

TABLE II. INPUT PARAMETERS

Parameter	**Symbol**	**Unit**	**Value at level**		
			1	**2**	**3**
Insert material	***IM***	-	**TiN**	**TiCN**	**TiAlN**
Tool nose radius	r	mm	0.3	0.5	0.8
Cutting speed	V_c	m/min	80	120	160
Feed rate	V_f	mm/min	250	320	390
Depth of cut	a_p	mm	0.20	0.30	0.40

The Taguchi method was used to design the experimental matrix. When comparing the matrix design method by the Taguchi method with some other matrix design methods, it can be found that it requires a smaller number of experiments. For example, with 5 input parameters, in which each parameter has 3 levels of values, the Taguchi method only needs 27 experiments while the Box-Behnken design needs at least 46 experiments and the Central Composite Design (CCD) method needs a minimum of 43 experiments. An advantage that only the Taguchi method obtains is that it allows designing the experimental matrix with input parameters that are not quantitative parameters. In this case the qualitative parameter is just the cutting insert material type. So, the experimental matrix was designed according to the Taguchi method with a total of 27 experiments, as shown in Table III.

IV. RESULTS AND DISCUSSION

The experiments in Table III are given in accordance with the results shown in Table IV. Figure 2 shows the influence of the input parameters on surface roughness. The comparison of the difference at the lowest and highest levels, i.e. between level 1 and level 3 of the parameter line graph (red broken line) shows that the tool nose radius is the parameter that has the greatest influence on surface roughness, followed by the influence of the cutting insert material and the feed rate. The difference of the line graph of cutting speed and cutting depth is very small, showing that these two parameters have negligible influence on surface roughness.

TABLE III. EXPERIMENTAL MATRIX

No.	Code value					Actual value				
	IM	*r*	V_c	V_f	a_p	*IM*	*r (mm)*	V_c *(m/min)*	V_f *(mm/min)*	a_p *(mm)*
1	1	1	1	1	1	TiN	0.3	80	250	0.2
2	1	1	1	1	2	TiN	0.3	80	250	0.4
3	1	1	1	1	3	TiN	0.3	80	250	0.6
4	1	2	2	2	1	TiN	0.5	120	320	0.2
5	1	2	2	2	2	TiN	0.5	120	320	0.4
6	1	2	2	2	3	TiN	0.5	120	320	0.6
7	1	3	3	3	1	TiN	0.8	160	390	0.2
8	1	3	3	3	2	TiN	0.8	160	390	0.4
9	1	3	3	3	3	TiN	0.8	160	390	0.6
10	2	1	2	3	1	TiCN	0.3	120	390	0.2
11	2	1	2	3	2	TiCN	0.3	120	390	0.4
12	2	1	2	3	3	TiCN	0.3	120	390	0.6
13	2	2	3	1	1	TiCN	0.5	160	250	0.2
14	2	2	3	1	2	TiCN	0.5	160	250	0.4
15	2	2	3	1	3	TiCN	0.5	160	250	0.6
16	2	3	1	2	1	TiCN	0.8	80	320	0.2
17	2	3	1	2	2	TiCN	0.8	80	320	0.4
18	2	3	1	2	3	TiCN	0.8	80	320	0.6
19	3	1	3	2	1	TiAlN	0.3	160	320	0.2
20	3	1	3	2	2	TiAlN	0.3	160	320	0.4
21	3	1	3	2	3	TiAlN	0.3	160	320	0.6
22	3	2	1	3	1	TiAlN	0.5	80	390	0.2
23	3	2	1	3	2	TiAlN	0.5	80	390	0.4
24	3	2	1	3	3	TiAlN	0.5	80	390	0.6
25	3	3	2	1	1	TiAlN	0.8	120	250	0.2
26	3	3	2	1	2	TiAlN	0.8	120	250	0.4
27	3	3	2	1	3	TiAlN	0.8	120	250	0.6

TABLE IV. EXPERIMENTAL RESULTS

No	*IM*	*r* (mm)	V_c (m/min)	V_f (mm/min)	a_p (mm)	*Ra* (µm)	*MRR* (mm^3/min)
1	TiN	0.3	80	250	0.2	0.771	600
2	TiN	0.3	80	250	0.4	1.457	1200
3	TiN	0.3	80	250	0.6	1.697	1800
4	TiN	0.5	120	320	0.2	1.538	768
5	TiN	0.5	120	320	0.4	0.905	1536
6	TiN	0.5	120	320	0.6	0.986	2304
7	TiN	0.8	160	390	0.2	2.205	936
8	TiN	0.8	160	390	0.4	1.582	1872
9	TiN	0.8	160	390	0.6	0.863	2808
10	TiCN	0.3	120	390	0.2	1.024	936
11	TiCN	0.3	120	390	0.4	1.112	1872
12	TiCN	0.3	120	390	0.6	0.801	2808
13	TiCN	0.5	160	250	0.2	1.076	600
14	TiCN	0.5	160	250	0.4	2.908	1200
15	TiCN	0.5	160	250	0.6	1.014	1800
16	TiCN	0.8	80	320	0.2	0.985	768
17	TiCN	0.8	80	320	0.4	3.019	1536
18	TiCN	0.8	80	320	0.6	1.549	2304
19	TiAlN	0.3	160	320	0.2	0.927	768
20	TiAlN	0.3	160	320	0.4	0.877	1536
21	TiAlN	0.3	160	320	0.6	0.892	2304
22	TiAlN	0.5	80	390	0.2	1.234	936
23	TiAlN	0.5	80	390	0.4	1.929	1872
24	TiAlN	0.5	80	390	0.6	0.824	2808
25	TiAlN	0.8	120	250	0.2	0.545	600
26	TiAlN	0.8	120	250	0.4	1.549	1200
27	TiAlN	0.8	120	250	0.6	1.601	1800

Because the cutting insert material, the tool nose radius and the cutting speed do not exist in the MRR calculation formula (4), they have no influence on MRR. Figure 3 shows the influence of the feed rate and the cutting depth on MRR. The difference between level 1 and level 3 of cutting depth line graph is greater than the one of the feed rate line graph. This shows that the cutting depth has a greater influence on MRR than the feed rate. Thus, we see that the influence of the input parameters on surface roughness and MRR is different, even adverse, e.g. the tool nose radius has a great influence on surface roughness without any influence on MRR, the feed rate and the cutting depth are two parameters needed for calculating the MRR, but they do not significantly affect surface roughness, etc. Thus, it can be said that being based only on the two graphs of Figures 2 and 3, limits the determination of the values of the input parameters to ensure minimum surface roughness and maximum MRR. Table IV shows that surface roughness has the smallest value in experiment #25, while MRR has the maximum value in experiments #9, #12 and #24. Thus, if we only observe Table IV, it is not possible to determine the value of the input parameters to ensure both minimum surface roughness and maximum MRR. The DEAR method will be used to solve this problem in the next section.

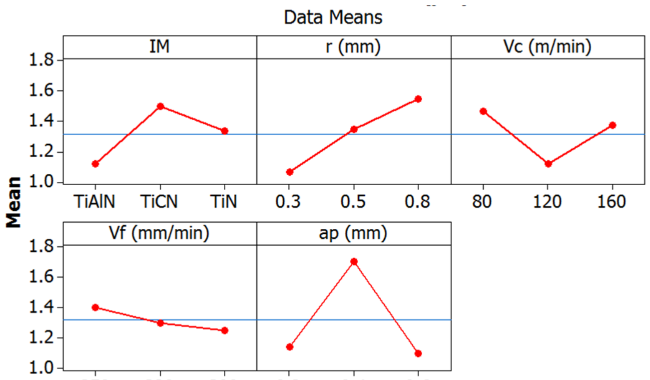


Fig. 2. Main effect plot for surface roughness.

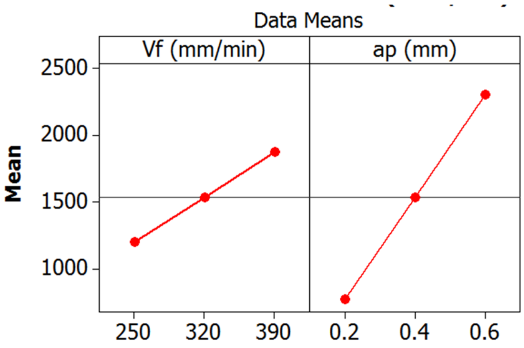


Fig. 3. Main effect plot for MRR.

V. SELECTION OF THE VALUE OF THE INPUT PARAMETERS

From the experimental data in Table IV, the weights of the responses and the MRPI value at each experiment are calculated according to (1) - (3), as shown in Table V. From the data in Table V, the MRPI values of all input parameters at all degrees were calculated. This value is calculated as the sum of the MRPI value of each parameter at the respective degree, as shown in Table VI. From the data in Table VI, it can be seen that the cutting insert material (*IM*) has the smallest value of MRPI corresponding to level 3, tool nose radius (*r*) has the smallest value of MRPI corresponding to level 1, cutting speed (V_c) corresponding to level 2, and feed rate (V_f) and cutting depth (a_p) corresponding to level 3. Thus, the optimal value of the parameters of the cutting insert material, the tool nose radius, the cutting speed, the feed rate, and the cutting depth are TiAlN, 0.3mm, 120m/min, 390mm/min, and 0.6mm respectively [24]. The MRPI value with the maximum Max-Min of 0.52972 is the cutting depth. Thus, the cutting depth is the parameter that has the greatest influence, followed by the tool nose radius, the cutting insert material, the cutting speed, and the feed rate [24].

TABLE V. EXPERIMENTAL RESPONSE WEIGHT AND MRPI

No.	W_{R_a}	W_{MRR}	**MRPI**
1	0.02168	0.07506	45.05364
2	0.04096	0.03753	45.09661
3	0.04771	0.02502	45.11789
4	0.04324	0.05864	45.10343
5	0.02544	0.02932	45.05996
6	0.02772	0.01955	45.06426
7	0.06199	0.04812	45.17362
8	0.04448	0.02406	45.10729
9	0.02426	0.01604	45.05787
10	0.02879	0.04812	45.06641
11	0.03126	0.02406	45.07169
12	0.02252	0.01604	45.05497
13	0.03025	0.07506	45.06948
14	0.08175	0.03753	45.27467
15	0.02851	0.02502	45.06584
16	0.02769	0.05864	45.06421
17	0.08487	0.02932	45.29317
18	0.04355	0.01955	45.10439
19	0.02606	0.05864	45.06109
20	0.02466	0.02932	45.05855
21	0.02508	0.01955	45.05930
22	0.03469	0.04812	45.07974
23	0.05423	0.02406	45.14154
24	0.01473	0.01604	45.04465
25	0.01532	0.07506	45.04528
26	0.04355	0.03753	45.10439
27	0.04501	0.02502	45.10899

TABLE VI. TOTAL MRPI

Parameter	**Level**			**Max - Min**
	1	**2**	**3**	
IM	405.83457	406.06482	405.70353	0.36129
r	405.64016	405.90357	406.05920	0.41904
V_c	405.99584	405.67938	405.92771	0.31646
V_f	405.93679	405.86835	405.79778	0.13901
a_p	405.71690	406.20787	405.67815	0.52972

VI. EXPERIMENTS WITH THE OPTIMAL VALUES OF THE PARAMETERS

The optimal set of the 5 input parameters defined above was used to experiment on the milling process with 3 steel samples. The surface roughness of each experimental sample is shown in Table VII. The MRR value at each experiment has also been calculated and is included in this Table. The average value of surface roughness in these cases is 0.724µm. If compared with the surface roughness values in Table IV, it can be seen that although 0.724µm is still larger than the value of surface roughness at experiment #25, this value is very small when compared to the total of 27 experiments that were carried out. For MRR, when calculated according to (4), in the three test samples, the MRR is equal to 2808mm^3/min, which is also larger than the data in Table IV. From that, it can be seen that when machining with the optimal values of the input parameters, MRR reaches its maximum value and surface roughness is also significantly improved. This result ensures the reliability when using the optimal value of the input parameters and proves the success in using the DEAR method in this study.

TABLE VII. OUTPUT PARAMETERS WHEN EXPERIMENTING WITH THE OPTIMAL VALUES OF THE INPUT PARAMETERS

No.	Optimization value					*Ra* (μm)	*MRR* (mm^3/min)
	IM	*r* (mm)	V_c (m/min)	V_f (mm/min)	a_p (mm)		
1	TiAlN	0.3	120	390	0.6	0.726	2808
2						0.721	2808
3						0.725	2808
Mean						0.724	2808

VII. CONCLUSION

An experimental process of milling SCM 440 steel was carried out in this study. Three types of cutting inserts were used, coated with TiN, TiCN, and TiAlN. The tool nose radius, the cutting speed, the feed rate, and the cutting depth were also determined as input parameters of the experimental process. The DEAR method was applied to determine the optimal value of the input parameters. Some of the conclusions drawn from this study are:

- The tool nose radius is the parameter that has the greatest influence on surface roughness, followed by the influence of the cutting insert material and the feed rate. The cutting speed and the cutting depth have no significant influence on surface roughness.
- Only the feed rate and the cutting depth have an influence on MRR, and the influence of the cutting depth on MRR is greater than the one of the feed rate.
- The parameter set that ensures simultaneously the two objectives is: TiAlN cutting insert material, 0.3mm tool nose radius, 120mm/min cutting speed, 390mm/min feed rate, and 0.6mm cutting depth.
- DEAR method was not only successful in determining the optimal values of the input parameters in this study as well as in [25-28] but it is also quite promising to being successful in the future when applied to determine the value of input parameters to simultaneously ensure multi-criteria optimization of the machining process.

ACKNOWLEDGMENT

This research is funded by University of Transport and Communications (UTC) under the grant number T2021-CK-003.

REFERENCES

[1] J. P. Hung, K. D. Wu, W. Z. Lin, and W. C. Shih, "Analyzing the Dynamic Characteristics of Milling Tool Using Finite Element Method and Receptance Coupling Method," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3918-3923, Apr. 2019, https://doi.org/10.48084/etasr.2463.

[2] D. D. Trung, "Influence of Cutting Parameters on Surface Roughness during Milling AISI 1045 Steel", *Tribology in Industry*, vol. 42, no. 4, pp. 658-665, Dec. 2020, https://doi.org/10.24874/ti.969.09.20.11.

[3] M. S. Phadke, *Quality Engineering Using Robust Design*, Printice Hall, 1989.

[4] S. K. Karna, R. V. Singh, and R. Sahai, "Application of Taguchi Method in Indian Industry", *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 11, pp. 387–391, 2012.

[5] M. Tamiloli and J. Venkatesan, "Investigation on Effect of Machining Parameters of End Milling on Surface Finish and Temperature Using Taguchi Technique", *Journal for Manufacturing Science and Production*, vol. 16, no. 4, pp. 255-261, Nov. 2016, https://doi.org/10.1515/jmsp-2016-0030.

[6] T. Kıvak, "Optimization of surface roughness and flank wear using the Taguchi method in milling of Hadfield steel with PVD and CVD coated inserts", *Measurement*, vol. 20, pp. 19-28, Apr. 2014, https://doi.org/10.1016/j.measurement.2013.12.017.

[7] M. H. F. Al-Hazza, N. A. bt Ibrahim, E. T. Y. Adesta, A. A. Khan, and A. B. A. Sidek, "Surface Roughness Optimization Using Taguchi Method of High Speed End Milling For Hardened Steel D2," *IOP Conference Series: Materials Science and Engineering*, vol. 184, Mar. Art. no. 012047, 2017, https://doi.org/10.1088/1757-899X/184/1/012047.

[8] I. Tlhabadira, I. A. Daniyan, R. Machaka, C. Machio, L. Masu, and L. R. VanStaden, "Modelling and optimization of surface roughness during AISI P20 milling process using Taguchi method," *The International Journal of Advanced Manufacturing Technology*, vol. 102, no. 9, pp. 3707–3718, Jun. 2019, https://doi.org/10.1007/s00170-019-03452-4.

[9] J. E. Ribeiro, M. B. César, and H. Lopes, "Optimization of machining parameters to improve the surface quanlity," *Procedia Structural Integrity*, vol. 5, pp. 355-362, Sep. 2017, https://doi.org/10.1016/j.prostr.2017.07.182.

[10] H. Öktem, T. Erzurumlu, and M. Çöl, "A study of the Taguchi optimization method for surface roughness in finish milling of mold surfaces", *The International Journal of Advanced Manufacturing Technology*, vol. 28, pp. 694–700, Oct. 2006, https://doi.org/10.1007/s00170-004-2435-6.

[11] A. A. Thakre, "Optimization of Milling Parameters for Minimizing Surface Roughness Using Taguchi's Approach", *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 6, pp. 226-230, Jun. 2013.

[12] S. Sakthivelu, T. Anandaraj, and M. Selwin, "Multi-Objective Optimization of Machining Conditions on Surface Roughness and MRR during CNC End Milling of Aluminium Alloy 7075 Using Taguchi Design of Experiments," *Mechanics and Mechanical Engineering*, vol. 21, no. 1, pp. 95–103, Jan. 2017.

[13] S. Kumar, I. Saravanan, and L. Patnaik, "Optimization of surface roughness and material removal rate in milling of AISI 1005 carbon steel using Taguchi approach," *Materials Today: Proceedings*, vol. 22, no. 3, pp. 654-658, Feb. 2020, https://doi.org/10.1016/j.matpr.2019.09.039.

[14] L. Ulfiyah, B. Pramujati, and B. O. P. Soepangkat, "Optimization of Tool Wear, Surface Roughness and Material Removal Rate in the Milling Process of Al 6061 using Taguchi and Weighted Principal Component Analysis (WPCA)," *Applied Mechanics and Materials*, vol. 493, pp. 535-540, Jan. 2014, https://doi.org/10.4028/www.scientific.net/AMM.493.535.

[15] L. M. Maiyar, R. Ramanujam, K. Venkatesan, and J. Jerald, "Optimization of Machining Parameters for End Milling of Inconel 718

Super Alloy Using Taguchi Based Grey Relational Analysis," *Procedia Engineering*, vol. 64, pp. 1276–1282, May. 2013, https://doi.org/10.1016/j.proeng.2013.09.208.

[16] V. C. Nguyen, T. D. Nguyen, and D. H. Tien, "Cutting Parameter Optimization in Finishing Milling of Ti-6Al-4V Titanium Alloy under MQL Condition using TOPSIS and ANOVA Analysis," *Engineering, Technology & Applied Science Research*, vol. 11, no. 11, pp. 6775-6780, Feb. 2021, https://doi.org/10.48084/etasr.4015.

[17] Q. Xu, J. Zhao, and X. Ai, "Cutting performance of tools made of different materials in the machining of 42CrMo4 high-strength steel: a comparative study," *The International Journal of Advanced Manufacturing Technology*, vol. 93, no. 5, pp. 2061–2069, Nov. 2017, https://doi.org/10.1007/s00170-017-0666-6.

[18] M. Elbah, M. A. Yallese, H. Aouici, T. Mabrouki, and J. Rigal, "Comparative assessment of wiper and conventional ceramic tools on surface roughness in hard turning AISI 4140 steel," *Measurement*, vol. 46, no. 9, pp. 3041–3056, Nov. 2013, https://doi.org/10.1016/j.measurement.2013.06.018.

[19] L. Chen, Y. Du, P. H. Mayrhofer, S. Q. Wang, and J. Li, "The influence of age-hardening on turning and milling performance of Ti–Al–N coated inserts," *Surface and Coatings Technology*, vol. 202, no. 21, pp. 5158–5161, Jul. 2008, https://doi.org/10.1016/j.surfcoat.2008.05.036.

[20] S. K. Khrais and Y. J. Lin, "Wear mechanisms and tool performance of TiAlN PVD coated inserts during machining of AISI 4140 steel," *Wear*, vol. 262, no. 1–2, pp. 64–69, Jan. 2007, https://doi.org/10.1016/j.wear.2006.03.052.

[21] Q. Xu, J. Zhao, X. Ai, W. Huang, and G. Wang, "Optimum selection of tool materials for machining of high-strength steels based on fuzzy comprehensive evaluation method," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 233, no. 1, pp. 145–153, Jan. 2019, https://doi.org/10.1177/0954405417713041.

[22] M. A. Stipkovic, E. C. Bordinassi, A. D. Farias, and S. Delijaicov, "Surface Integrity Analysis in Machining of Hardened AISI 4140 Steel," *Materials Research*, vol. 20, no. 2, pp. 387-394, Apr. 2017, https://doi.org/10.1590/1980-5373-MR-2016-0420.

[23] W. Sales, M. Becker, S. B. Clovis, and J. Landre, "Tribological behaviour when face milling AISI 4140 steel with minimum quantity fluid application," *Industrial Lubrication and Tribology*, vol. 61, no. 2, pp. 84–90, Mar. 2009, https://doi.org/ 10.1108/00368790910940400.

[24] H. Liao and Y. Chen, "Optimizing multi-response problem in the Taguchi method by DEA based ranking method," *International Journal of Quality & Reliability Management*, vol. 19, no. 7, pp. 825–837, Jan. 2002, https://doi.org/10.1108/02656710210434766.

[25] S. Nguyen Hong and U. Vo Thi Nhu, "Multi-objective Optimization in Turning Operation of AISI 1055 Steel Using DEAR Method," *Tribology in Industry*, vol. 43, no. 1, pp. 57–65, Mar. 2021, https://doi.org/10.24874/ti.1006.11.20.01.

[26] V. V. K. Lakshmi, K. V. Subbaiah, A. V. Kothapalli, and K. Suresh, "Parametric optimization while turning Ti-6Al-4V alloy in Mist-MQCL (Green environment) using the DEAR method," *Manufacturing Review*, vol. 7, 2020, Art. no. 38, https://doi.org/10.1051/mfreview/2020034.

[27] D. D. Trung, N.-T. Nguyen, D. H. Tien, and H. L. Dang, "A research on multi-objective optimization of the grinding process using segmented grinding wheel by Taguchi-Dear method," *Eureka: Physics and Engineering*, vol. 1, pp. 67-77, Jan. 2021, https://doi.org/10.21303/2461-4262.2021.001612.

[28] V. V. Reddy and C. S. Reddy, "Multi Response Optimization of EDM of AA6082 Material using Taguchi-DEAR Method," *International Journal of Scientific & Engineering Research*, vol. 7, no. 6, pp. 215-219, Jun. 2016.

[29] "Sandvik Coromant - manufacturing tools & machining solutions," *Sandvik Coromant*. https://www.sandvik.coromant.com/en-gb/pages/default.aspx (accessed Aug. 18, 2021).

Health Risk Assessment of Arsenic in the Drinking Water of Upper Sindh, Pakistan

Noor Ul Hassan Shar
Institute of Chemistry
Shah Abdul Latif University
Khairpur, Pakistan
noorulhassan4228@gmail.com

Ghulam Qadir Shar
Institute of Chemistry
Shah Abdul Latif University
Khairpur, Pakistan
gqadir.shar@salu.edu.pk

Abdul Raheem Shar
Institute of Chemistry
Shah Abdul Latif University
Khairpur, Pakistan
saifiraheem3863920@gmail.com

Shafi Muhammad Wassan
Department of Community Medicine
Ghulam Muhammad Mahar Medical College Sukkur
Sindh, Pakistan
wassanshafi1975@gmail.com

Zubeda Bhatti
Department of Physics
Shah Abdul Latif University
Khairpur, Pakistan
bhatti.zubeda@yahoo.com

Abida Ali
Department of Computer Science
Shah Abdul Latif University
Khairpur, Pakistan
abida.ali@mckru.edu.pk

Abstract-Water is a valuable compound for plants, animals, and humans. Various contaminating agents pollute it, with arsenic being one of them. Measurements of arsenic in potable water in Upper Sindh were conducted during this study. The samples were prepared by microwave-assisted digestion and analyzed by an atomic absorption spectrophotometer. A total of 240 potable water samples were collected from 8 Talukas of Upper Sindh. DMS coordinates were also recorded with the help of the Global Positioning System (GPS). The highest arsenic content of 50µg/L was observed in Garhi Khairo Taluka. The average arsenic content in water samples of all of the Talukas, except Miro Khan, was found higher than the WHO permissible limit. The 69.2% of samples were found to be contaminated by arsenic. Therefore, the water of the studied area is concluded to be in poor condition for cooking and drinking.

Keywords-arsenic; upper Sindh; atomic absorption spectrophotometer; water health risk assessment

I. INTRODUCTION

Arsenic is a metalloid with high toxicity, which can be found in the environment in geological substrata, whereas it can also be released by human activities [1]. In numerous areas, arsenic in freshwater atmosphere, due to direct release from anthropogenic and natural sources, is a severe concern [2]. The concentration of arsenic ranges from 0.5 to 5000µg/L in freshwater [3]. Four oxidation states of arsenic are observed in nature: (-3), (0), (+3) and (+5) whereas it can be found in various forms such as arsenate, arsenite, monomethylarsonic acid, dimethyl arsenic acid, trimethylarsine, arsenocholine, arsenobetaine, arsenosugras, etc. Organic arsenic is less poisonous than inorganic arsenic, while pentavalent arsenic is less toxic than its trivalent state [4]. Arsenic is accumulated in soil through human activities like fuel utilization, mining, smelting of arsenic ores, manufacture of arsenic–based compounds, and use of arsenic–based pesticides [5, 6]. Biomethylation mechanism of micro-organisms and use of organo-arsenical pesticides are responsible for organic arsenic in natural water systems [7]. Higher concentrations of arsenic in groundwater of countries like Italy, Vietnam, Bangladesh, West Bengal, Hungary, China, Mexico, Chile, and Argentina have been observed [8].

Water resources are exploited because of the continuous development of global population. Arsenic and heavy metal contamination of water sources are often found to be in extremely critical state [9]. These pollutants find their ways in surface and groundwater sources due to their high mobility. There are various groundwater and surface water sources [10]. When compared to surface water, groundwater is more secure regarding heavy metals and microbes [11]. Therefore, most people use groundwater for drinking purposes [12], although a lot of research has been conducted on groundwater containing heavy metals [13]. About 100 million people of Southeast Asia were found to be at high risk of arsenic toxicity [14]. There are two main ways of arsenic accumulation in the human body, taking arsenic polluted potable water and consuming arsenic contaminated food [15]. Stomach toxicity, brain, kidney, and liver cancers, and skin lesions may be the results of chronic

---

Corresponding author: Abdul Raheem Shar

effects of arsenic. Due to its toxicity, WHO and USEPA have reduced the threshold limit of arsenic from 50 to 10µg/L in 2001 [16].

To the best of our knowledge, in the 8 Talukas of upper Sindh that constitute the studied area, arsenic contamination and health risk assessment in potable water have not been conducted. The present work aimed to identify arsenic contamination and its potential health risk assessment in upper Sindh keeping in view the public, anthropogenic inputs and geology.

II. MATERIALS AND METHODS

A. Study Area

The weather of upper Sindh is moderate in winter and extremely hot in summer. The area is commonly considered as a hot arid zone. The lowest and the highest temperatures recorded are -3.9ºC and 52.8^{0}C respectively. In the season of monsoon from July to September the rainfall is not sufficient. The average recorded annual rainfall is 122.5mm and the air is generally dry. Upper Sindh lies from 27°56' to 28°27' N and from 68° to 69°44' E covering an area of about 6,790km^2. In its surroundings, Thar Desert is located on the eastern side, Bahurai range in west, Kherthar range in south west and Suleman range in the north. At the south eastern side of Jacobabad, Indus River flows from north to south and besides the Bolan River there are various canals and streams at the western and northern sides. There are parts of Baluchistan highlands at northern, western, and south western areas and as a result, non-perennial streams are found towards the province of Sindh [17].

Fig. 1. Hand pumps in the study area.

B. Chemicals and Reagents

Reagents bought from Merck (Darmstadt, Germany) along with de-ionized water were used for arsenic solution preparation.

C. Sampling

In order to get groundwater samples it is necessary to run hand pumps at least for 5 minutes to eliminate insoluble impurities and sand particles and to get depth water of required elemental amounts [18]. Water used for drinking was obtained from hand pumps whose depth was varied from 30 to 70 feet. Plastic 1500 mL bottles were used to collect the water samples. Eight Talukas of upper Sindh were selected to be sampled. The coordinates were recorded with the help of a GPS device. In total, 240 potable water samples were collected from 8 Talukas of upper Sindh. After collection, the water samples were transported immediately to the laboratory. Practical investigation of arsenic was conducted in the laboratory and its quantity was measured with the help of a calibration graph.

D. Instrumentation

To measure the arsenic contamination from potable water, the technique of Atomic Absorption Spectrophotometer (Perkin-A 700) was used along with Mercury Hydride System (MHS–15). It is a precise and simple method of analysis of metals found in various samples [19].

E. Microwave Digestion Method

PTFE flasks were taken and 500mL of water samples were kept in them. After closing tightly, the flasks were subjected to microwave radiation in stopped up vessel microwave digestion using Milestone Ethos D model (Sorisole-Bg, Italy). The digestion plan of microwave oven was 100W for 2min, 250W for 6min, 400W for 5min, 550W for 8min, and ventilation for 8min. After the cooling, the content of the flask was diluted with 0.2M nitric acid to 10mL. The same procedure was followed to prepare the reagent blank. The advantage of microwave digestion method is that it takes less time to digest water samples and the possibility of evaporation of elements is less, therefore more precise extraction of elements from samples can be accomplished as compared to conventional digestion methods. Less acid is used for digestion as well [20].

III. RESULTS AND DISCUSSION

More than 40% of the population of Pakistan suffers from arsenic contamination in potable water. More than 20% of the people of Punjab suffer from arsenic contamination in surface or groundwater sources, with the problem becoming bigger in industrial zones [21-24]. According to our results, the highest concentration of 50µg/L of arsenic was obtained in Garhi Khairo Taluka and the minimum of 3µg/L was found in samples from Ubauro, Tangwani, Kambar, and Miro Khan Talukas. The mean arsenic content in all Talukas except Miro Khan was higher than the WHO permissible limit. In Daharki and Kambar Talukas, 80% of groundwater samples were contaminated by arsenic, in Kashmore and Garhi Khairo 86.7%, and in Ubauro, Tangwani, Thul, and Miro Khan the contamination percentage was 70%, 73%, 56.7%, and 20% respectively (Table I). Water quality is worse in heavily populated regions of Pakistan such as Peshawar, Lahore, Karachi, Shaheed Benazirabad [49] and other cities and towns. The groundwater of upper Sindh is also contaminated and not safe for drinking due to its high arsenic contamination. Some reported cities of Pakistan with high arsenic content in groundwater are shown in Table II. Arsenic concentration in various countries is given in Table III. Descriptive statistics such as, minimum, maximum, mean and standard deviation of arsenic in potable water of upper Sindh are given in Table IV.

TABLE I. ARSENIC CONCENTRATION (µg/L) IN THE STUDIED EIGHT TALUKAS OF UPPER SINDH

Daharki	Ubauro	Kashmore	Tangwani	Garhi Khario	Thul	Kambar	Miro Khan
5	10	9	21	17	38	11	9
11	12	12	7	16	13	3	6
14	20	12	9	12	12	16	4
11	18	13	18	11	8	21	5
5	19	10	19	4	7	12	6
10	26	14	16	9	8	3	5
11	18	11	12	20	5	15	18
14	15	10	20	13	16	21	17
15	12	12	33	12	9	18	16
5	24	15	24	12	8	26	17
16	37	12	37	4	8	34	7
17	25	16	10	17	6	49	5
13	38	18	8	22	13	26	3
15	25	18	6	13	18	11	6
5	3	13	9	4	9	10	7
19	14	15	14	16	8	16	9
20	13	19	16	22	12	15	7
17	11	9	11	19	18	11	9
14	17	14	17	27	15	10	6
13	9	13	25	35	23	3	15
18	6	13	39	50	31	8	6
17	5	14	25	27	46	19	7
10	7	15	3	12	23	12	15
13	9	17	9	11	8	11	7
11	18	24	22	16	7	15	10
12	19	27	20	13	12	12	5
14	18	17	19	34	9	33	6
13	20	17	26	25	30	24	7
13	5	21	18	20	21	19	5
27	6	12	15	30	16	29	9

TABLE II. CONCENTRATION OF ARSENIC IN DIFFERENT CITIES OF PAKISTAN

Province	Area	As (µg/L)	References
Sindh	Sajawal	15.3	[25]
	Ghorabari	50	[26]
	MirpurSakro	80	[26]
	Ketibandar	5–25	[27]
	Khairpur	0.24−315.6	[28]
	Gambat	0.01–126	[29]
	Shaheed Benazirabad	10–200	[30]
	Dadu	8–67	[31]
	Thatta	10–200	[32]
	Jamshoro	13−106	[33]
Punjab	Sheikhupura	5–76	[34]
	Rahimyar Khan	20–500	[35]
	Muzffargarh	0.01−900	[36]
	Dera Gazi Khan	1–29	[37]
Balochistan	Sibi	0.3–3.5	[38]

TABLE III. ARSENIC CONCENTRATION IN VARIOUS COUNTRIES

Country	Concentration (µg/L)	Reference
India	44	[39]
China (Shanxi)	1932	[40]
Afghanistan	100	[41]
Taiwan	1800	[42]
Greece	> 20	[43]
Bangladesh	398	[44]
Veitnam	3100	[45]
Nepal	260	[46]
Cambodia	3500	[47]
Inner Mongolia, China	4000	[48]

TABLE IV. DESCRIPTIVE STATISTICS OF ARSENIC IN POTABLE WATER OF UPPER SINDH

Taluka	N	Min	Max	Mean	Std. deviation
Daharki	30	5	27	13.3	4.78
Ubauro	30	3	38	16.0	8.74
Kashmore	30	9	27	14.7	4.18
Tangwani	30	3	39	17.6	8.85
Garhi Khario	30	4	50	18.1	10.11
Thul	30	5	46	15.2	10.00
Kambar	30	3	49	17.1	10.11
Miro Khan	30	3	18	8.5	4.32

A. Pearson Correlation Coefficient

The Pearson correlation of the potable water of the 8 talukas of upper Sindh is shown in Table V. Positive correlation was observed among Garhi Khairo and Daharki (0.483). Arsenic in potable water of Thul and Garhi Khairo also showed positive correlation of 0.491, whereas the water of the Kambar and Ubauro Talukas also displayed positive correlation of 0.415. Negative correlation of arsenic in drinking water was found among Talukas Thul and Ubauro. All relations were observed significant at the level of 0.05, and 0.1 respectively.

B. Human health risk assessment

Equation (1) was used to estimate the total arsenic intake [33]:

$$\mathrm{ADI} = \frac{\mathrm{Cw \, x \, Di}}{\mathrm{BW}} \quad (1)$$

Body weight (BW) and daily water intake for common people were supposed to be 65kg and 3 to 3.5L respectively. The results of the Average Daily Intake (ADI) are given in Table VI. ADI values at alarming levels were found in the water of all Talukas of upper Sindh with the exception of Miro Khan. The values ranged from 0.46 to 0.97μg/day. Comparatively, Tangwani, Garhi Khairo, and Kambar Talukas showed higher ADI values. Various problems may be caused due to high ADI values such as diabetes, cardiovascular problems, hypertension, black foot disease, keratosis, bladder and lung cancer, and skin lesions [31] (Tables VI, VII).

TABLE V. CORRELATION COEFFICIENT OF ARSENIC IN GROUNDWATER OF DIFFERENT UPPER SINDH TALUKAS

	Daharki	**Ubauro**	**Kashmore**	**Tangwani**	**Garhi Khario**	**Thul**	**Kambar**	**Miro Khan**
Daharki	1							
Ubauro	-0.097	1						
Kashmore	0.057	0.150	1					
Tangwani	0.097	0.023	-0.063	1				
Garhi Khario	0.483^{**}	-0.308	0.040	0.319	1			
Thul	0.095	-0.432^{*}	-0.132	0.262	0.491^{**}	1		
Kambar	0.306	0.415^{*}	0.092	0.116	-0.004	-0.197	1	
Miro Khan	-0.136	-0.217	-0.300	0.159	-0.030	-0.029	-0.074	1

TABLE VI. AVERAGE DAILY INTAKE OF ARSENIC FROM DRINKING WATER OF DIFFERENT UPPER SINDH TALUKAS

Daharki	**Ubauro**	**Kashmore**	**Tangwani**	**Garhi Khario**	**Thul**	**Kambar**	**Miro Khan**
0.27	0.54	0.48	1.13	0.92	2.05	0.59	0.48
0.59	0.65	0.65	0.38	0.86	0.70	0.16	0.32
0.75	1.08	0.65	0.48	0.65	0.65	0.86	0.22
0.59	0.97	0.70	0.97	0.59	0.43	1.13	0.27
0.27	1.02	0.54	1.02	0.22	0.38	0.65	0.32
0.54	1.40	0.75	0.86	0.48	0.43	0.16	0.27
0.59	0.97	0.59	0.65	1.08	0.27	0.81	0.97
0.75	0.81	0.54	1.08	0.70	0.86	1.13	0.92
0.81	0.65	0.65	1.78	0.65	0.48	0.97	0.86
0.27	1.29	0.81	1.29	0.65	0.43	1.40	0.92
0.86	1.99	0.65	1.99	0.22	0.43	1.83	0.38
0.92	1.35	0.86	0.54	0.92	0.32	2.64	0.27
0.70	2.05	0.97	0.43	1.18	0.70	1.40	0.16
0.81	1.35	0.97	0.32	0.70	0.97	0.59	0.32
0.27	0.16	0.70	0.48	0.22	0.48	0.54	0.38
1.02	0.75	0.81	0.75	0.86	0.43	0.86	0.48
1.08	0.70	1.02	0.86	1.18	0.65	0.81	0.38
0.92	0.59	0.48	0.59	1.02	0.97	0.59	0.48
0.75	0.92	0.75	0.92	1.45	0.81	0.54	0.32
0.70	0.48	0.70	1.35	1.88	1.24	0.16	0.81
0.97	0.32	0.70	2.10	2.69	1.67	0.43	0.32
0.92	0.27	0.75	1.35	1.45	2.48	1.02	0.38
0.54	0.38	0.81	0.16	0.65	1.24	0.65	0.81
0.70	0.48	0.92	0.48	0.59	0.43	0.59	0.38
0.59	0.97	1.29	1.18	0.86	0.38	0.81	0.54
0.65	1.02	1.45	1.08	0.70	0.65	0.65	0.27
0.75	0.97	0.92	1.02	1.83	0.48	1.78	0.32
0.70	1.08	0.92	1.40	1.35	1.62	1.29	0.38
0.70	0.27	1.13	0.97	1.08	1.13	1.02	0.27
1.45	0.32	0.65	0.81	1.62	0.86	1.56	0.48

Note: Safe arsenic daily intake in water is 0.66μg/day

TABLE VII. MEAN ADI OF ARSENIC IN POTABLE WATER OF THE UPPER SINDH TALUKAS

Taluka	Mean (µg/L)	ADI Mean (µg/day)
Daharki	13.3	0.72
Ubauro	16	0.86
Kashmore	14.7	0.79
Tangwani	17.6	0.95
Garhi Khairo	18.1	0.97
Thul	15.2	0.82
Kambar	17.1	0.92
Miro Khan	8.5	0.46

IV. CONCLUSION

From the results obtained from the present work, it can be concluded that the groundwater of upper Sindh is unfit for drinking due to the presence of arsenic in high amounts. With the exception of Miro Khan where only 20% of water samples were found to be contaminated, the potable water of the Talukas was found highly contaminated due to the existence of arsenic concentration at alarming levels. Contaminated water may cause various health hazards. Therefore, it is suggested to the government of Sindh to pay special consideration to the matter. It is also recommended that awareness among farmers must be increased to use chemicals, particularly pesticides, carefully. Also, the government should provide safe drinking water by installing Reverse Osmosis (RO) plants in areas where people are using unsafe water for drinking and cooking.

REFERENCES

[1] S. Sher and A. Rehman, "Use of heavy metals resistant bacteria—a strategy for arsenic bioremediation," *Applied Microbiology and Biotechnology*, vol. 103, no. 15, pp. 6007–6021, Aug. 2019, https://doi.org/10.1007/s00253-019-09933-6.

[2] A. N. Chowdhury *et al.*, "Arsenic in freshwater ecosystems of the Bengal delta: status, sources and seasonal variability," *Toxicological & Environmental Chemistry*, vol. 97, no. 5, pp. 538–551, May 2015, https://doi.org/10.1080/02772248.2015.1053482.

[3] R. Foroutan *et al.*, "Efficient arsenic(V) removal from contaminated water using natural clay and clay composite adsorbents," *Environmental Science and Pollution Research*, vol. 26, no. 29, pp. 29748–29762, Oct. 2019, https://doi.org/10.1007/s11356-019-06070-5.

[4] J. Tapia, J. Murray, M. Ormachea, N. Tirado, and D. K. Nordstrom, "Origin, distribution, and geochemistry of arsenic in the Altiplano-Puna plateau of Argentina, Bolivia, Chile, and Perú," *Science of The Total Environment*, vol. 678, pp. 309–325, Aug. 2019, https://doi.org/10.1016/j.scitotenv.2019.04.084.

[5] X. Wan, M. Lei, and T. Chen, "Review on remediation technologies for arsenic-contaminated soil," *Frontiers of Environmental Science & Engineering*, vol. 14, no. 2, Dec. 2019, Art. no. 24, https://doi.org/10.1007/s11783-019-1203-7.

[6] S. Rai, D. Singh, and A. K. Singh, in *Mycoremediation of Arsenic from Contaminated Soil for Sustainable Agriculture*, New Delhi, India: Todya & Tomorrow's Printers and Publishers, 2020, pp. 575–591.

[7] E. Shaji, M. Santosh, K. V. Sarath, P. Prakash, V. Deepchand, and B. V. Divya, "Arsenic contamination of groundwater: A global synopsis with focus on the Indian Peninsula," *Geoscience Frontiers*, vol. 12, no. 3, May 2021, Art. no. 101079, https://doi.org/10.1016/j.gsf.2020.08.015.

[8] S. Bolisetty, M. Peydayesh, and R. Mezzenga, "Sustainable technologies for water purification from heavy metals: review and analysis," *Chemical Society Reviews*, vol. 48, no. 2, pp. 463–487, Jan. 2019, https://doi.org/10.1039/C8CS00493E.

[9] K. Khan *et al.*, "Prevalent fecal contamination in drinking water resources and potential health risks in Swat, Pakistan," *Journal of Environmental Sciences*, vol. 72, pp. 1–12, Oct. 2018, https://doi.org/10.1016/j.jes.2017.12.008.

[10] S. O. Olasoji, N. O. Oyewole, B. Abiola, and J. N. Edokpayi, "Water Quality Assessment of Surface and Groundwater Sources Using a Water Quality Index Method: A Case Study of a Peri-Urban Town in Southwest, Nigeria," *Environments*, vol. 6, no. 2, Feb. 2019, Art. no. 23, https://doi.org/10.3390/environments6020023.

[11] D. Wang, J. Wu, Y. Wang, and Y. Ji, "Finding High-Quality Groundwater Resources to Reduce the Hydatidosis Incidence in the Shiqu County of Sichuan Province, China: Analysis, Assessment, and Management," *Exposure and Health*, vol. 12, no. 2, pp. 307–322, Jun. 2020, https://doi.org/10.1007/s12403-019-00314-y.

[12] H. Rajkumar, P. K. Naik, and M. S. Rishi, "A new indexing approach for evaluating heavy metal contamination in groundwater," *Chemosphere*, vol. 245, Apr. 2020, Art. no. 125598, https://doi.org/10.1016/j.chemosphere.2019.125598.

[13] M. Shahid, C. Dumat, N. Khan Niazi, S. Khalid, and Natasha, "Global scale arsenic pollution: increase the scientific knowledge to reduce human exposure," *VertigO - la revue électronique en sciences de l'environnement*, no. Hors-série 31, Sep. 2018, https://doi.org/10.4000/vertigo.21331.

[14] M. Ozturk *et al.*, "Arsenic and Human Health: Genotoxicity, Epigenomic Effects, and Cancer Signaling," *Biological Trace Element Research*, Apr. 2021, https://doi.org/10.1007/s12011-021-02719-w.

[15] M. Costa, "Review of arsenic toxicity, speciation and polyadenylation of canonical histones," *Toxicology and Applied Pharmacology*, vol. 375, pp. 1–4, Jul. 2019, https://doi.org/10.1016/j.taap.2019.05.006.

[16] D. Mohan and C. U. Pittman, "Arsenic removal from water/wastewater using adsorbents—A critical review," *Journal of Hazardous Materials*, vol. 142, no. 1, pp. 1–53, Apr. 2007, https://doi.org/10.1016/j.jhazmat.2007.01.006.

[17] S. Ali, "Heavy Downpour Event over upper Sindh in September, 2012," *Pakistan Journal of Meteorology*, vol. 9, no. 18, pp. 59–72, 2013.

[18] A. R. Shar, G. Q. Shar, W. B. Jatoi, N.-U.-H. Shar, L. A. Shar, and W. M. Ghouri, "Assessment of the quality of drinking water of Thari Mirwah Town and surrounding villages, District Khairpur, Sindh, Pakistan," *Pakistan Journal of Analytical & Environmental Chemistry*, vol. 15, no. 2, pp. 39–59, 2014.

[19] J. B. Fisher, K. P. Tu, and D. D. Baldocchi, "Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 901–919, Mar. 2008, https://doi.org/10.1016/j.rse.2007.06.025.

[20] D. Guven and G. Akinci, "Comparison of acid digestion techniques to determine heavy metals in sediment and soil samples," *Gazi University Journal of Science*, vol. 24, no. 1, pp. 29–34, Jan. 2011.

[21] I. S. Khalid and A. A. Khaver, *Political Economy of Water Pollution in Pakistan: An Overview*. Islamabad, Pakistan: Sustainable Development Policy Institute, 2019.

[22] A. Kumar and C. K. Singh, "Arsenic enrichment in groundwater and associated health risk in Bari doab region of Indus basin, Punjab, India," *Environmental Pollution*, vol. 256, Jan. 2020, Art. no. 113324, https://doi.org/10.1016/j.envpol.2019.113324.

[23] F. Rehman *et al.*, "Groundwater quality and potential health risks caused by arsenic (As) in Bhakkar, Pakistan," *Environmental Earth Sciences*, vol. 79, no. 24, Nov. 2020, Art. no. 529, https://doi.org/10.1007/s12665-020-09270-2.

[24] A. Shahab, S. Qi, and M. Zaheer, "Arsenic contamination, subsequent water toxicity, and associated public health risks in the lower Indus plain, Sindh province, Pakistan," *Environmental Science and Pollution Research*, vol. 26, no. 30, pp. 30642–30662, Oct. 2019, https://doi.org/10.1007/s11356-018-2320-8.

[25] A. Shar *et al.*, "Quality Characteristics and Risk Assessment of Arsenic in Drinking Water of Different Villages of District Sujawal, Sindh, Pakistan," *International Journal of Biosciences (IJB)*, vol. 17, no. 3, pp. 287–294, Jun. 2020.

[26] A. R. Shar, G. Q. Shar, S. M. Wassan, Z. Bhatti, A. A. Shar, and N.-H. Shar, "Risk Assessment of Arsenic and Cadmium in Groundwater of

Talukas Ghorabari and Mirpur Sakro, Sindh, Pakistan," *Pakistan Journal of Analytical & Environmental Chemistry*, vol. 22, no. 1, pp. 100–114, Jun. 2021, https://doi.org/10.21743/pjaec/2021.06.11.

[27] A. Shar *et al.*, "Risk assessment of arsenic and heavy metals of drinking water in coastal area: Case for Taluka Keti Bandar, Sindh, Pakistan," *Journal of Biodiversity and Environmental Sciences*, vol. 16, no. 6, pp. 52–64, Aug. 2020.

[28] M. A. Jakhrani, A. J. Chaudhray, K. M. Malik, M. Q. Mazari, A. A. Jakhrani, and M. ul-Hassan, "Determination of Arsenic and Other Heavy Metals in Hand Pump and Tube-Well Ground Water of Khairpur, Sindh, Pakistan," in *Second International Conference on Environmental and Computer Science*, Dubai, United Arab Emirates, Dec. 2009, pp. 271–276, https://doi.org/10.1109/ICECS.2009.94.

[29] M. A. Jakhrani, K. M. Malik, S. Sahito, and A. A. Jakhrani, "Analytical Investigation of Arsenic and Iron in hand pump and tube-well groundwater of Gambat, Sindh, Pakistan," *Pakistan Journal Chemistry*, vol. 1, no. 3, pp. 140–144, 2011.

[30] A. J. Kandhro, N. A. Samoon, and J. H. Laghari, "Assessment of arsenic and essential metal ions in the quality of groundwater sources of Taluka Daur, District Shaheed Benazeer Abad, Sindh, Pakistan," *Journal of Peoples University of Medical & Health Sciences*, vol. 6, no. 1, pp. 1–8, 2016.

[31] A. Memon, G. Lund, N. Channa, S. Shah, M. Younis, and F. Buriro, "Contaminants Exposure and Impacts on Drinking Water of Johi Subdivision of Sindh, Pakistan," *Science Letters*, vol. 4, no. 1, pp. 78–83, Jan. 2016.

[32] G. Rubab, S. Naseem, A. Khan, V. Husain, and G. M. Arain, "Distribution and sources of arsenic contaminated groundwater in parts of Thatta district, Sindh," *Journal of Himalayan Earth Sciences*, vol. 47, no. 2, pp. 175–183, 2014.

[33] J. A. Baig *et al.*, "Evaluation of arsenic and other physico-chemical parameters of surface and ground water of Jamshoro, Pakistan," *Journal of Hazardous Materials*, vol. 166, no. 2, pp. 662–669, Jul. 2009, https://doi.org/10.1016/j.jhazmat.2008.11.069.

[34] M. Abbas and K. J. Cheema, "Arsenic levels in drinking water and associated health risk in district Sheikhupura, Pakistan," *The Journal of Animal & Plant Sciences*, vol. 25, no. 3, pp. 719–724, 2015.

[35] M. T. Mahar, M. Y. Khuhawar, T. M. Jahangir, and M. A. Baloch, "Determination of arsenic contents in groundwater of District Rahim Yar Khan Southern Punjab, Pakistan," *Arabian Journal of Geosciences*, vol. 8, no. 12, pp. 10983–10994, Dec. 2015, https://doi.org/10.1007/s12517-015-1979-0.

[36] R. T. Nickson, J. M. McArthur, B. Shrestha, T. O. Kyaw-Myint, and D. Lowry, "Arsenic and other drinking water quality issues, Muzaffargarh District, Pakistan," *Applied Geochemistry*, vol. 20, no. 1, pp. 55–68, Jan. 2005, https://doi.org/10.1016/j.apgeochem.2004.06.004.

[37] M. A. Malana and M. A. Khosa, "Groundwater pollution with special focus on arsenic, Dera Ghazi Khan-Pakistan," *Journal of Saudi Chemical Society*, vol. 15, no. 1, pp. 39–47, Jan. 2011, https://doi.org/10.1016/j.jscs.2010.09.009.

[38] T. A. Chandio, M. N. Khan, M. T. Muhammad, O. Yalcinkaya, A. A. Wasim, and A. F. Kayis, "Fluoride and arsenic contamination in drinking water due to mining activities and its impact on local area population," *Environmental Science and Pollution Research*, vol. 28, no. 2, pp. 2355–2368, Jan. 2021, https://doi.org/10.1007/s11356-020-10575-9.

[39] P. Bhattacharya *et al.*, "Health risk assessment of co-occurrence of toxic fluoride and arsenic in groundwater of Dharmanagar region, North Tripura (India)," *Groundwater for Sustainable Development*, vol. 11, Oct. 2020, Art. no. 100430, https://doi.org/10.1016/j.gsd.2020.100430.

[40] M. Sanjrani, B. Zhou, H. Zhao, S. Bhutto, A. Muneer, and S. Xia, "Arsenic contaminated groundwater in China and its treatment options, a review," *Applied Ecology and Environmental Research*, vol. 17, no. 2, pp. 1655–1683, Feb. 2019, https://doi.org/10.15666/aeer/1702_16551683.

[41] H. A. Jawadi, H. A. Malistani, M. A. Moheghy, and J. Sagin, "Essential Trace Elements and Arsenic in Thermal Springs, Afghanistan," *Water*, vol. 13, no. 2, Jan. 2021, Art. no. 134, https://doi.org/10.3390/w13020134.

[42] H.-J. Lin, T.-I. Sung, C.-Y. Chen, and H.-R. Guo, "Arsenic levels in drinking water and mortality of liver cancer in Taiwan," *Journal of Hazardous Materials*, vol. 262, pp. 1132–1138, Nov. 2013, https://doi.org/10.1016/j.jhazmat.2012.12.049.

[43] E. Zkeri, M. Aloupi, and P. Gaganis, "Natural Occurrence of Arsenic in Groundwater from Lesvos Island, Greece," *Water, Air, & Soil Pollution*, vol. 226, no. 9, Aug. 2015, Art. no. 294, https://doi.org/10.1007/s11270-015-2542-z.

[44] N. Saha, Md. Bodrud-Doza, A. R. M. T. Islam, B. A. Begum, and M. S. Rahman, "Hydrogeochemical evolution of shallow and deeper aquifers in central Bangladesh: arsenic mobilization process and health risk implications from the potable use of groundwater," *Environmental Earth Sciences*, vol. 79, no. 20, Oct. 2020, Art. no. 477, https://doi.org/10.1007/s12665-020-09228-4.

[45] M. Berg, H. C. Tran, T. C. Nguyen, H. V. Pham, R. Schertenleib, and W. Giger, "Arsenic Contamination of Groundwater and Drinking Water in Vietnam: A Human Health Threat," *Environmental Science & Technology*, vol. 35, no. 13, pp. 2621–2626, Jul. 2001, https://doi.org/10.1021/es010027y.

[46] S. Gwachha, B. N. Acharya, A. Dhakal, S. M. Shrestha, and T. P. Joshi, "Assessment of Arsenic Content in Deep Groundwater of Kathmandu Valley, Nepal," *Nepal Journal of Science and Technology*, vol. 19, no. 1, pp. 69–77, Jul. 2020, https://doi.org/10.3126/njst.v19i1.29785.

[47] K. Phan, S. Sthiannopkao, and K.-W. Kim, "Surveillance on chronic arsenic exposure in the Mekong River basin of Cambodia using different biomarkers," *International Journal of Hygiene and Environmental Health*, vol. 215, no. 1, pp. 51–58, Dec. 2011, https://doi.org/10.1016/j.ijheh.2011.07.002.

[48] X. J. Guo, Y. Fujino, S. Kaneko, K. Wu, Y. Xia, and T. Yoshimura, "Arsenic contamination of groundwater and prevalence of arsenical dermatosis in the Hetao plain area, Inner Mongolia, China," in *Molecular Mechanisms of Metal Toxicity and Carcinogenesis*, X. Shi, V. Castranova, V. Vallyathan, and W. G. Perry, Eds. Boston, MA, USA: Springer, 2001, pp. 137–140.

[49] N.-K. Bhatti, S. R. Samo, A. Saand, M. A. Keerio, and A. A. Bhuriro, "Ground Water Quality Assessment of Daur Taluka, Shaheed Benazir Abad," *Engineering, Technology and Applied Science Research*, vol. 8, no. 2, pp. 2785–2789, Apr. 2018, https://doi.org/10.48084/etasr.1925.

Engineering, Technology & Applied Science Research Vol. 11, No. 5, 2021, 7564-7570

Exploring the Relationship between Foreign Investments and Carbon Emission: A Bound Test Analysis for Pakistan

Hina Samreen
Department of Mathematics and Statistics
Institute of Business Management
Karachi, Pakistan
hina.samreen@iobm.edu.pk

Shahida Wizarat
Department of Economics
Institute of Business Management
Karachi, Pakistan
shahida.wizarat@iobm.edu.pk

Zafar Mehdi
Department of Management & HR
Institute of Business Management
Karachi, Pakistan
zafar.mehdi@iobm.edu.pk

Rizwan Ahmed
Department of Mathematics and Statistics
Institute of Business Management
Karachi, Pakistan
rizwan.ahmed@iobm.edu.pk

Abstract—**In this study, the theoretical and empirical effects of foreign investment, openness to trade, urbanization, and economic development on carbon dioxide emission levels (CO_2) in Pakistan were examined. Annual time series for 38 years 1970-2018 were utilized. The Auto Regressive Distributed Lag (ARDL) bound test has been used to estimate the model which indicates the presence of cointegration among the variables. Long-run and short-run relationships were observed between Foreign Direct Investment (FDI) and CO_2 emission levels revealing that increase in FDI results in increased CO_2 emission levels contributing to environmental degradation. This endorses the theory of Pollution Haven Hypothesis (PHH) in the case of Pakistan. To achieve the goal of sustainable economic development, it is imperative to consider the possible negative effects of foreign investment on the environment. The study suggests limiting CO_2 emissions by adopting more advanced low carbon technologies to ensure growth and a sustainable environment. Strict implementation of environmental laws for foreign and domestic firms both by defining CO_2 emission limits, installation of proper waste management plants, and imposing a pollution tax are crucial steps for the improvement of environmental quality.**

Keywords-FDI; Pakistan; CO2; emissions; environment; bound test; PHH

I. INTRODUCTION

Environmental pollution, climatic change, and global-warming are some of the most concerned topics in scientific research. The last century witnessed a significant change in climate as revealed by different indicators. It is reported that there is a 0.41° increase in average recorded global temperature from 1998 to 2007 as compared to the 1961 to 1990 period [1]. This weather change is a big ecological issue that has caught the attention of researchers, policy makers, international institutions, and environmentalists. According to the Intergovernmental Panel for Climate Change (IPCC), increased human activity due to the economic liberalization and industrialization has warmed the earth by 1° centigrade since the pre-industrialization era. Furthermore, the IPCC 2018 reported that the effects of global warming have been felt in many parts of the world. Nonetheless, they are not spread consistently over the globe, with different areas experiencing the impacts unexpectedly and unevenly. Among various effects, it is resulting in heavy rainfall in certain areas, especially in the high latitudes of the Northern Hemisphere, raising the risk of flooding. Further increase in global warming may aggravate the situation due to the rapid melting of glaciers, the liquefication of ice masses, and, consequently, the rising of the sea level [2].

Carbon dioxide (CO_2) emerges as a main contributor to global warming [3]. CO_2 diffuses into the environment through the extraction, refining, and consumption of petroleum products and fossil fuels. There are several factors which have contributed to the rise in temperature over the last century and energy demand is one of them. A rapidly growing economy has huge energy requirements for power generation, transportation, and industrial production. Because of the onset of globalization in the early 1990s, free capital inflows along with the increase in international trade intensified the movement of foreign investment in less developed countries in various sectors [4]. There is substantial evidence that Foreign Direct Investment (FDI) enhances the productivity of host countries and promotes economic growth. Foreign investment brings capital, technology, and skill development and provides access to new markets and greater opportunities to exports [5]. The industrial sector is a key sector of an economy, vastly contributing to economic growth. Foreign resource mobilization is crucial for industrial production and growth [6].

Corresponding author: Hina Samreen

The existing literature on FDI and CO_2 emissions adopts two different approaches. The Pollution Haven theory, which is the major focus of attention in this study, has a hypothesis referred to as "Pollution Haven Hypothesis" (PHH) stating that foreign investment moves towards countries with poor environmental laws and conditions and less environmental taxes. The hypothesis explains that Multi-National Corporations (MNCs) relocate the industries which are more pollution intensive in nature to Less Developed Countries (LDCs) with flexible environmental regulations to avoid high environmental costs in terms of environmental taxes. This results in increasing pollution level in the host country thus becoming a pollution haven [7]. The environmental quality is critical for human health. Poor environmental conditions pose a serious risk to human wellbeing in various ways [8]. According to the World Health Organization (WHO) Global Health Observatory, poor environmental quality causes roughly 200 deaths per 100,000 people in Pakistan. Indoor air pollution induced by residential biomass fuel combustion causes about 30 deaths per 100,000 people, whereas outdoor pollution causes around 25 deaths per 100,000 people. Rapid increase in population growth, number of vehicles, unplanned and rapid urbanization, and industrialization cause air pollution, particularly Particulate Matter (PM) pollution. Traffic congestion, open air combustion of untreated solid waste, and improper management of construction and demolition waste are the main sources of air pollution in urban areas [9].

Considering the importance of foreign investment in development, growth, and prosperity, the objective of the current study is to explore if the FDI is a possible major cause of environmental degradation, and if so, ways to design a policy framework that is sustainable for economic development with minimum hazards for the environment. Numerous empirical researches often aimed to study the direct relationship of economic growth and carbon emission [10-12] and FDI and Greenhouse Gas (GHG) emissions [13, 14], while in the current study, we analyze the indirect impact of foreign investment to GHG emissions through its direct relationship with economic growth, urbanization, and trade openness. The goal of this study is to look at the evidence of this indirect relationship and see if there is a significant effect, and if so, to quantify it. This paper contributes to earlier research on foreign investment inflow and its possible harmful effects on carbon emissions to help policy makers to design policies that combat carbon emissions but not at the cost of economic growth.

II. POLLUTION HAVEN HYPOTHESIS

Environmental pollution is the result of human activities which the environment cannot deal with naturally. FDI is considered as a main contributor to this environmental destruction as it accelerates economic activity that results in economic growth on one hand but it also causes an increase in GHG emissions, deforestation, and biodiversity loss on the other. Authors in [15] used a non-linear estimation method to check if there is any association between environmental devastation and foreign investment in the case of the Malaysian economy for the 1965-2010 period. The results confirm that FDI results in great environmental degradation. PHH asserts that a host country with weak environmental regulation attracts more foreign investments, as profit-oriented companies try to evade costly regulatory compliance in their home countries by relocating their dirty industries in developing countries [16]. The nexus between weak environmental laws and capital inflows of polluting industries is explored in [17]. The authors considered cross-sectional data for 7 developing economies and 15 developed countries for the period from 1985 to 1990 and found that poor environmental laws in the host country encourage FDI inflow in the industries that pollute environment, while for the less polluting sector, the results were insignificant. Authors in [18] tried to explore whether private sector involvement is beneficial to the environmental quality. The study considers CO_2 emission as a measure of environmental pollution, along with other variables like financial market scope, industrial sector composition, and foreign investment. The study found that greater involvement of the private sector in a developing economy resulted in environmental quality improvement. Earlier, authors in [19] analyzed the case of ASEAN economies to study the impact of capital inflow on the environmental degradation. The study used the ARDL model and time series data for 31 years (1970-2001). Thailand, Malaysia, and the Philippines are the countries where a significant relationship between FDI and pollution is observed. It has been observed that foreign capital inflows do not contribute to increasing pollution levels in Indonesia, while for Singapore no significant results were found. Authors in [20] studied the long run relationship between capital inflows, pollution level, and output for Malaysia using the Granger Causality test. The research employed annual per capita GDP as a proxy for income and annual CO_2 emission per capita as a proxy for pollution. The study found that FDI Granger-causes GDP growth and pollution level in the short run, while in the long run, causal relationship was observed for output and FDI inflows. Authors in [21] studied the association between foreign investment and environmental policy considering 33 countries. The study found that foreign investment can influence environmental policy only if there is a sincere effort from local government to design and implement environmental rules. The study concludes that the flexibility of environmental policy depends on the degree of corruption of government officials. The authors recommended the adoption of a political economic model with imperfect product market competition along with the application of pollution tax to combat pollution. Authors in [22] used data from 112 countries to study the relationship between foreign inflows and the level of pollution. Granger causality test was employed to test the results. The authors report that for lower income groups, CO_2-levels Granger-cause foreign investment inflows. Also, for middle income groups, FDI Granger-causes CO_2 emissions, whereas for higher income groups there is no evidence of Granger causality.

To confirm the PHH for ASEAN countries, authors in [23] used the ARDL method on time series data for the 1970-2008 period. In the case of the Philippines, the results indicated that all the regression coefficients except for per capita gross national income are significant. Mixed results were reported in the cases of Thailand and Indonesian economies. Authors in [24] found that economies with liberalized trade policies contribute more to pollution intensity than those following more restrictive policies. Authors in [25] tried to explore whether free trade agreements are affecting the environment of the countries that are bound by these agreements. The Gravity model framework was used to analyze the impact of free trade agreements on bilateral CO_2 emissions. The study found that weak environmental standards prevailing in low income countries result in greater pollution effect of free trade agreements, whereas, high income countries shift dirty industries to the countries where there is weak or no implementation of environmental policy, hence confirming the Pollution Haven theory. The export volume of China is a sign of greater economic activity, but also generates an enormous amount of CO_2 [26]. The author calculates the size of CO_2 emission by 18 exporting industries of China from 2001 to 2010 using the I-O framework. The study concludes that the increase in the volume of exports is a major contributor to the increment of the CO_2 level of a country. The study did not confirm China as a Pollution Haven for developed countries. Authors in [27] empirically explored the link between energy consumed by the transport sector, CO_2 emission, economic development, foreign investment, and urbanization. ARDL and VECM were used for estimation over the period from 1990 to 2015. The study found transport sector energy consumption as a significantly strong determinant of CO_2 emissions in Pakistan. The study also confirms the contribution of foreign investment in CO_2 emission, whereas economic growth and urbanization were found to be insignificant determinants of CO_2 emission. The study recommends that the government should play a more active role in improving environmental quality by providing energy efficient solutions in the transport sector with less adverse effects on economic growth. The author in [28] selected per capita GDP in real terms, energy consumption, urbanization, transport sector intensive-energy along with fuel prices as key determinants of pollution level in Tunisia during the period from 1980 to 2014. The author claimed that the reliance of the road transport system mainly on fossil fuel usage is a major cause of CO_2 emissions. The study supported the EKC theory by confirming the presence of an inverse U-shaped curve indicating the relationship between economic development and CO_2 level from the transport sector.

III. DATA SOURCES AND MODEL SPECIFICATION

Data on economic and other key indicators were collected from World Development Indicators (WDI) over the time span from 1970 to 2018 [29]. As a determinant of environmental quality, we have used CO_2 emission per capita (CO2PCAP), in metric tons. We have also considered other economic indicators due to the limited availability of carbon emission data. Gross Domestic Product Per Capita (GDPCAP) was used to define the level of economic development (in constant 2010 US \$). FDI is the net inflow of investments as a percentage of GDP (FDIPGDP). The urbanization variable (Urban) is calculated as a percentage of urban to total population. Trade to GDP ratio is a proxy for a liberal economy (TOPENESS) and was also taken from the WDI. To explore how capital inflows affect CO_2 emission levels in Pakistan, the following model has been specified using EViews.

$$\begin{gathered}\log(\mathrm{CO2PCAP}) = \\ \beta_0 + \beta_1\,(\log(\mathrm{GDPCAP}) + \beta_2(\log(\mathrm{TOPENNESS}) + \\ \beta_3\,(\log(\mathrm{FDIPGDP}) + \beta_4\,(\log(\mathrm{Urban}) + \phi \quad (1)\end{gathered}$$

ARDL-bound test estimation method, proposed in [30, 31] has been used to explore the relationship between CO_2 emissions with foreign investment, GDP per capita, urbanization, and trade openness. ARDL estimation generates results in two steps. The first is the ECM results and the second is the Bounds test results. When co-integration exists, it is recommended to also report the ECM results, in case of no co-integration, we conclude with the ARDL model.

IV. DATA ANALYSIS AND DIAGNOSTIC TEST

Tables I and II report the descriptive statistics and the correlation matrix for the variables considered in the current study. To confirm the robustness of the results, different diagnostic tests are employed which are discussed in the following section. To examine the stationarity of variables, both Augmented Dickey-Fuller (ADF) test and Phillips & Peron (PP test) were applied. Table III shows the findings. It is revealed from unit roots tests that initially, level following variables are non-stationary like per capita CO_2 emission, foreign investment, and per capita GDP, but become stationary after taking the first difference, whereas the remaining variables are stationary.

TABLE I. DESCRIPTIVE MEASURES

Variable	Mean	Std. Dev.	Min	Max
FDIPGDP	0.7606914	0.7968646	-0.063242	3.668323
Urban	31.16394	3.422316	24.81700	36.44200
GDPCAP	2.191195	2.162259	-2.218736	8.385496
TOPENESS	32.55893	4.043413	19.93229	38.90950
CO2PCAP	0.6512522	0.2209307	0.3083813	0.9878324

TABLE II. CORRELATION MATRIX

	FDIPGDP	Urban	GDPCAP	CO2PCAP	TOPENESS
FDIPGDP	1.0000				
Urban	0.5683	1.0000			
GDPCAP	0.0003	-0.0657	1.0000		
CO2PCAP	0.6382	0.9795	-0.0628	1.0000	
TOPENESS	0.2612	0.1392	-0.0446	0.0783	1.0000

TABLE III. AUGMENTED DICKEY FULLER UNIT ROOT TEST

Variable	ADF at level P-value	ADF at first difference (P-value)	Order of integration	PP test at level	PP test at first difference
LnCO2PCAP	0.8561	0.000	I(1)	0.9221	0.0000
Ln FDI	0.0816	0.0000	I(1)	0.1138	0.000
Ln Urban	0.0144		I(0)	0.0000	
LnGDPCAP	0.8550	0.0000	I(1)	0.9435	0.0000
LnTOPENSS	0.0112		I(0)	0.0097	

An optimal lag length selection is the first step in ARDL estimation. Table IV reports various lag selection criteria available in the literature. Except Schwartz Information Criterion (SIC) all three-information criteria suggest that we should consider a maximum lag length of two in this study. The * sign in Table IV defines the selected lag order by the chosen criterion. In the second step we proceed with the estimation of the ARDL model, which requires the computation and analysis of F-statistics using ARDL bound test methodology.

TABLE IV. LAG LENGTH SELECTION

Lag	LogL	LR	FPE	AIC	SIC	HQ
0	87.99601	NA	0.001465	-3.688711	-3.487971	-3.613878
1	100.7045	22.02806	0.000871	-4.209089	-3.968201	-4.119288
2	103.1279	4.092887*	0.000819*	-4.272352*	-3.991316*	-4.167585*
3	103.9750	1.393006	0.000825	-4.265557	-3.944372	-4.145822
4	103.9809	0.009329	0.000864	-4.221371	-3.860039	-4.086670

Table V reports the calculated F-statistic and its critical values. The F-statistic (17.773) is found to be greater than the upper bound critical value (5.06) at 1% significance level which helps us reject the null hypothesis of no long-run association at 1% level of significance. Long run elasticities for both the dependent variable (carbon emissions) and the independent variables (Foreign Investment, Per capita GDP, Urbanization and Trade openness) are summarized in Table VI. The coefficients for FDI, Urbanization and trade liberalization are found to be significant at the 1% level of significance in the long run, while the coefficient on per capita GDP is found to be significant at the 10% level of significance. As expected, the estimated coefficients of all the explanatory variables carry the expected positive sign. Foreign investment opens the door to infrastructure and technological and human development in the host country [32]. A significant and positive coefficient on FDFI describes investment opportunities available at the cost of environmental degradation. A 1% increase in foreign investments causes a 0.0418% increase in CO_2 emissions. The coefficient on urbanization is also found to be a positive and significant determinant of carbon emissions. High urbanization means that more people are settled in urban areas resulting in increase in human and economic activities. This result is in line with the findings in [33].

TABLE V. ARDL BOUND TEST

F-Statistic 17.773***		
Critical values bound		
Significance	**Lower bound**	**Upper bound**
1%	3.74	5.06
5%	2.86	4.01
10%	2.45	3.52

TABLE VI. LONG RUN ESTIMATES

Variable	Coefficient	Std. error	t-statistic	Prob.
LOG(FDIPGDP)	0.041811	0.007119	5.873148	0.0000
LOG(GDPCAP)	0.232906	0.133448	1.745300	0.0895
LOG(URBAN)	1.951081	0.329676	5.918186	0.0000
LOG(TOPENESS)	0.186990	0.041948	4.457692	0.0001
LOG(CO2PCAP)	-9.122096	0.480991	-18.965208	0.0000

A continuous growth in urbanization is responsible for the increase in the pollution level of a country causing damage to the environment, not only in industrialized nations, but also in developing economies [34]. People migrate to cities to find better health, education, and employment opportunities. This growth in urban population creates numerous challenges for urban areas due to the increase in demand and the limited resources. This unplanned population growth brings new challenges in various ways. Better infrastructure and transport facilities form a need for growing urban population. To provide transport facilities to a large urban population results in increase in the number of vehicles, resulting in more vehicular emissions. Rapid growth in urban population also exerts pressure on the availability of clean water and proper wastage and sewerage system [33]. As expected, we observed a positive and significant coefficient (0.2329) for per capita GDP growth and CO_2 emissions in the long run, namely 1% increase in per capita GDP growth contributes to 0.2329% increase in CO_2 emissions. This phenomenon may reflect in more economic activities and the number of vehicular emissions [35]. The coefficient (0.1869) of trade openness is also found to be positive and significant determinant of CO_2 emission levels. This is a significant result explaining why Pakistan experiences a comparative advantage in the production of goods that are harmful for the environment. This result helps to conclude that in the absence of strict environmental protection laws, greater trade openness does not lead to the path of sustainable growth in the long run. Since we found evidence for a long run relationship we further tested for short run relationship by adopting the ECM-ARDL approach. The results are given in Table VII. The adjustment coefficient (-1.0307) has the expected negative sign and is statistically significant. Table VII also provides the short run estimates of regressors and CO_2 emissions.

TABLE VII. SHORTRUN ESTIMATES

Variable	Coefficient	Std. error	t-statistic	Prob.
DLOG(CO2PCAP (-1))	0.060852	0.148354	0.410180	0.6842
DLOG(CO2PCAP (-2))	0.040098	0.106818	0.375383	0.7097
DLOG(FDIPGDP (-1))	0.017003	0.009797	1.735437	0.0917
DLOG(FDIPGDP (-2))	-0.000999	0.007147	-0.139768	0.8897
DLOG(GDPCAP (-1))	0.425215	0.261891	1.623635	0.1137
DLOG (GDPCAP (-2))	0.315363	0.308304	1.022896	0.3136
DLOG(URBAN (-1))	-5.003957	9.281719	-0.539120	0.5933
DLOG(URBAN (-2))	5.684971	8.876761	0.640433	0.5262
DLOG(TOPENESS(-1))	0.123089	0.050503	2.437276	0.0202
DLOG(TOPENESS(-2))	0.017732	0.037870	0.468230	0.6426
ECM(-1)	-1.030749	0.277675	-3.712069	0.0007
R-square	0.560200			

The optimal lag length is found to be 2 for every variable thus all the short run estimates given in the Table show 2 lags for every variable. All explanatory variables are found to be insignificant, except for the coefficient on trade openness, which means trade openness contributes significantly to carbon emission only in the short run. The Error Correction (ECM) term is found to be significant with the expected negative sign (-1.0307) which explains that whenever per capita carbon emission is above or below the equilibrium level, it automatically adjusts by 103% and the speed of adjustment occurs within the first year. The value of R square shows that the fitted model is a good fit and explains that the 56% variability in the model is due to the included explanatory variables. The high p-values of urbanization and trade openness coefficients are evident that no short run causal relationship exists between the explanatory variables and the dependent variable. Whereas a significant value of FDI at 1% and per capita GDP at 10% indicates that there is an evidence of short run causal relation from explanatory variables to the dependent variable. The highly significant ECM term (p-value 0.0007) in Table VII is evident of the causality among variables. Generally, the co-integration analysis is used to observe the possible causal relationship between the variables, at least uni-directional. But it does not provide the direction of causality. Thus, to examine the direction of causality among the variables the next step is to run a Granger causality test. Table IX provides the results. Two-way causality is observed between carbon emission and FDI. A very interesting finding of the Granger causality test is the uni-directional causality from FDI to CO_2 emission per capita GDP growth, urbanization, and trade openness. This validates the hypothesis of the study that foreign investment stimulates the economic activity reflected in terms of urbanization and per capita GDP growth but also explains the environmental degradation in terms of hazardous emissions.

TABLE VIII. WALD TEST

Variables	Probabilities
FDI	0.000
GDPCAP	0.0992
URBAN	0.5730
TOPENESS	0.1144

TABLE IX. VECM GRANGER CAUSILTY TEST

Hypothesis	Probability
LOG(FDIPGDP) does not Granger-cause LOG(CO2PCAP)	0.0058
LOG(CO2PCAP) does not Granger-cause LOG(FDIPGDP)	0.0981
LOG(GDPCAP) does not Granger-cause LOG(CO2PCAP)	0.0002
LOG(CO2PCAP) does not Granger-cause LOG(GDPCAP)	0.1487
LOG(URBAN) does not Granger-cause LOG(CO2PCAP)	1.E-05
LOG(CO2PCAP) does not Granger-cause LOG(URBAN)	0.0433
LOG(TOPENESS) does not Granger-cause LOG(CO2PCAP)	0.0273
LOG(CO2PCAP) does not Granger-cause LOG(TOPENESS)	0.4636
LOG(GDPCAP) does not Granger-cause LOG(FDIPGDP)	0.2950
LOG(FDIPGDP) does not Granger-cause LOG(GDPCAP)	0.0055
LOG(URBAN) does not Granger-cause LOG(FDIPGDP)	0.4332
LOG(FDIPGDP) does not Granger-cause LOG(URBAN)	0.0296
LOG(TOPENESS) does not Granger-cause LOG(FDIPGDP)	0.7558
LOG(FDIPGDP) does not Granger-cause LOG(TOPENESS)	0.0003
LOG(URBAN) does not Granger-cause LOG(GDPCAP)	0.0820
LOG(GDPCAP) does not Granger-cause LOG(URBAN)	0.1818
LOG(TOPENESS) does not Granger-cause LOG(GDPCAP)	0.4591
LOG(GDPCAP) does not Granger-cause LOG(TOPENESS)	0.5039
LOG(TOPENESS) does not Granger-cause LOG(URBAN)	0.0045
LOG(URBAN) does not Granger-cause LOG(TOPENESS)	0.7983

In Table IX we also observe that per capita GDP growth causes carbon emissions in the long run. This result also suggests that a stimulated economic activity may harm the environment if certain protective measures are not adopted. No long run causality was observed in the reverse direction which means that CO_2 emission does not Granger-cause economic growth, therefore a unidirectional causality is observed. This result is in contrast to the findings of [1], who asserted that any measure to control CO_2 emissions may have different implications for developed and developing countries. Measures to limit or control CO_2 emissions are more harmful for developing countries as compared to developed countries. To test the stability of the model, the Cumulative Sum of Recursive Residuals (CUSUSM) test was applied and the result is explained with the help of Figures 1 and 2. The plots of CUSUM test explain the stability of the ECM model. The straight lines in both Figures describe the critical values at 5% level of significance. The CUSUM and CUSUMSQ plots lie within the drawn critical values confirming the model is stable.

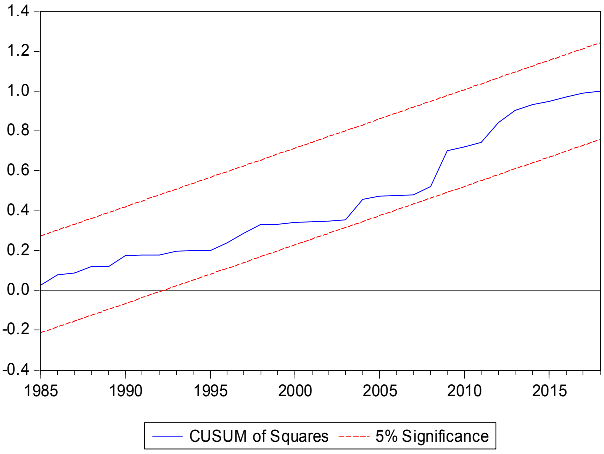


Fig. 1. CUSUM plot.

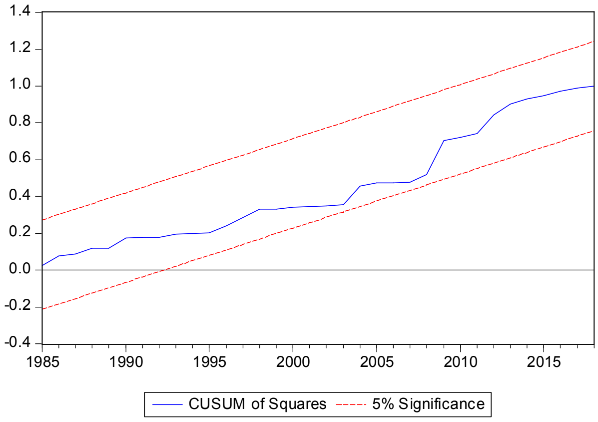


Fig. 2. CUSUMQ plot.

Further diagnostic tests are applied to the ECM model and the results are presented in Table X. The results suggest that the ARDL model fits well and passes the Breusch-Pagan-Godfrey heteroscedasticity test and Breusch-Godfrey LM Test for serial correlation.

TABLE X. DIAGNOSTIC TEST RESULTS

Breusch-Pagan-Godfrey for heteroscedasticity		
F-value	1.034710	Prob. F (14,30)
Breusch-Godfrey LM test of serial correlation		
F-value	1.178700	Prob. F (2,28)

V. CONCLUSION

In the neoliberal economic model, foreign investment is a major contributor to development through the introduction of new knowledge and technology and the creation of opportunities for employment and income growth. Therefore, the governments in host countries give lucrative offers to attract foreign investors. Our results confirm that increasing FDI causes an increase to CO_2 emissions, thus contributing to environmental degradation. A positive association between CO_2 emissions and FDI corroborates the PHH for Pakistan. The study suggests that before signing an agreement with the investor, host governments need to consider the possible adverse effects of FDI on the environment. Host governments should ensure that the investor company/country would adopt a clean energy technology that would reduce the pollution level. Thus, there is a need to set standards for industrial pollutants and emissions. To achieve the object of sustainable economic growth and development, governments should focus on strict implementation of policies that are more environmentally friendly and limit emissions to an ideal level. A sustainable approach is one that is low carbon in nature, such as wind and solar energy and aimed at reducing emissions but not at the cost of economic growth.

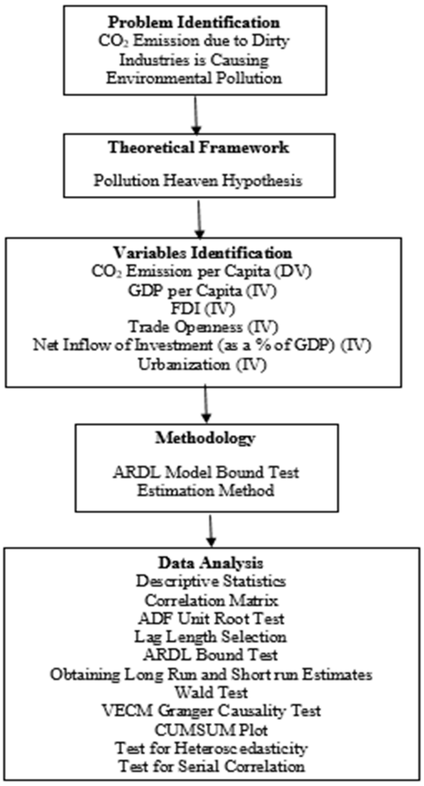


Fig. 3. Flow chart of the research.

This study establishes the understanding that a stimulated economic activity coupled with rapid urbanization results in more dependence on fossil fuel consumption and ultimately results in harmful emissions. Environmental degradation not only disturbs human life, but also puts pressure on economic activities in the long term. Urbanized activities are more energy intensive in nature. Dependence on fossil fuel consumption to cater for demands of the urban sector (electricity generation, heating, and transport) results in increased levels of pollution. With the growing population, it is difficult to limit the energy demand, hence awareness programs for the public are needed to be designed. To reduce the negative effects of urbanization, governments should monitor the technology transfer process and set standards to ensure environmental quality.

Introduction of more ecofriendly and energy efficient automobiles and provision of better public transport in the urban sector are some steps governments should adopt. There is a need to explore long term investment options that are cleaner in nature and provide sustainable solutions both for economic development and the environment. Moreover, to attract foreign investors with more ecofriendly transfer of technology, strict environmental policy, improved infrastructure, and human capital development are crucial. Figure 3 exhibits the flow chart of the research conducted to determine if there is a relationship between carbon emissions and FDI.

APPENDIX

LIST OF ACRONYMS

IPCC	Intergovernemental Panel for Climate Change
FDI	Forgien Direct Investment
CO2	Carbon Dioxide
PHH	Pollution Haven Hypothesis
MNCs	Multinational Corporations
LDC's	Less Developed Countries
WHO	World Health Organization
PM	Particulate Matter
GHG	Greenhouse Gas Emissions
ASEAN	Association for South East Asian Nations
ECM	Error Correction Model
ARDL	Autoregressive Distributed Lag Model
SIC	Schwartz Information Criteria
FDIPGDP	Net inflow of investement as a percentage of GDP
GDPCAP	Per Capita GDP
TOPENESS	Trade Openness
CO2PCAP	Carbon Dioxide Emissions per Capita
Urban	Urbanization
CUSUM	Cumulative Sum Control Chart

REFERENCES

[1] F. Bilgili, E. Kocak, and U. Bulut, "The dynamic impact of renewable energy consumption on CO2 emissions: A revisited Environmental Kuznets Curve approach," *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 838–845, Feb. 2016, https://doi.org/10.1016/j.rser.2015.10.080.

[2] O. Hoegh - Guldberg *et al.*, "Chapter 3: Impacts of 1.5°C global warming on natural and human systems," Intergovernmental Panel on Climate Change, 2018.

[3] Y. Liu, Y. Hao, and Y. Gao, "The environmental consequences of domestic and foreign investment: Evidence from China," *Energy Policy*, vol. 108, pp. 271–280, Sep. 2017, https://doi.org/10.1016/j.enpol.2017.05.055.

[4] L.-S. Lau, C.-K. Choong, and Y.-K. Eng, "Carbon dioxide emission, institutional quality, and economic growth: Empirical evidence in

Malaysia," *Renewable Energy*, vol. 68, pp. 276–281, Aug. 2014, https://doi.org/10.1016/j.renene.2014.02.013.

[5] G. Boluk and M. Mert, "The renewable energy, growth and environmental Kuznets curve in Turkey: An ARDL approach," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 587–595, Dec. 2015, https://doi.org/10.1016/j.rser.2015.07.138.

[6] A. Khan and Y.-H. Kim, "Foreign Direct Investment in Pakistan: Policy Issues and Operational Implications," Asian Development Bank, Philippines, Asia, 66, Jul. 1999. Accessed: Aug. 21, 2021. [Online]. Available: https://www.adb.org/publications/foreign-direct-investment-pakistan-policy-issues-and-operational-implications.

[7] M. Mert and A. E. Caglar, "Testing pollution haven and pollution halo hypotheses for Turkey: a new perspective," *Environmental Science and Pollution Research*, vol. 27, no. 26, pp. 32933–32943, Sep. 2020, https://doi.org/10.1007/s11356-020-09469-7.

[8] A. A. Mahessar, S. Qureshi, A. L. Qureshi, K. Ansari, and G. H. Dars, "Impact of the Effluents of Hyderabad City, Tando Muhammad Khan, and Matli on Phuleli Canal Water," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5281–5287, Feb. 2020, https://doi.org/10.48084/etasr.3269.

[9] A. A. Siyal, S. R. Samo, Z. A. Siyal, K. C. Mukwana, S. A. Jiskani, and A. Mengal, "Assessment of Air Pollution by PM10 and PM2.5 in Nawabshah City, Sindh, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3757–3761, Feb. 2019, https://doi.org/10.48084/etasr.2440.

[10] M. Bouznit and M. del P. Pablo-Romero, "CO2 emission and economic growth in Algeria," *Energy Policy*, vol. 96, pp. 93–104, Sep. 2016, https://doi.org/10.1016/j.enpol.2016.05.036.

[11] Y.-J. Joo, C. S. Kim, and S.-H. Yoo, "Energy Consumption, Co2 Emission, and Economic Growth: Evidence from Chile," *International Journal of Green Energy*, vol. 12, no. 5, pp. 543–550, May 2015, https://doi.org/10.1080/15435075.2013.834822.

[12] F. M. Mirza and A. Kanwal, "Energy consumption, carbon emissions and economic growth in Pakistan: Dynamic causality analysis," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 1233–1240, May 2017, https://doi.org/10.1016/j.rser.2016.10.081.

[13] N. A. Neequaye and R. Oladi, "Environment, growth, and FDI revisited," *International Review of Economics & Finance*, vol. 39, pp. 47–56, Sep. 2015, https://doi.org/10.1016/j.iref.2015.06.002.

[14] R. Kastratovic, "Impact of foreign direct investment on greenhouse gas emissions in agriculture of developing countries," *Australian Journal of Agricultural and Resource Economics*, vol. 63, no. 3, pp. 620–642, 2019, https://doi.org/10.1111/1467-8489.12309.

[15] M. B. Hitam and H. B. Borhan, "FDI, Growth and the Environment: Impact on Quality of Life in Malaysia," *Procedia - Social and Behavioral Sciences*, vol. 50, pp. 333–342, Jan. 2012, https://doi.org/10.1016/j.sbspro.2012.08.038.

[16] J. Hu, Z. Wang, Y. Lian, and Q. Huang, "Environmental Regulation, Foreign Direct Investment and Green Technological Progress—Evidence from Chinese Manufacturing Industries," *International Journal of Environmental Research and Public Health*, vol. 15, no. 2, Feb. 2018, Art. no. 221, https://doi.org/10.3390/ijerph15020221.

[17] Y. Xing and C. D. Kolstad, "Do Lax Environmental Regulations Attract Foreign Investment?," *Environmental and Resource Economics*, vol. 21, no. 1, pp. 1–22, Jan. 2002, https://doi.org/10.1023/A:1014537013353.

[18] D. Talukdar and C. M. Meisner, "Does the Private Sector Help or Hurt the Environment? Evidence from Carbon Dioxide Pollution in Developing Countries," *World Development*, vol. 29, no. 5, pp. 827–840, May 2001, https://doi.org/10.1016/S0305-750X(01)00008-0.

[19] Y. Merican, Z. Yusop, Z. Noor, and S. Law, "Foreign Direct Investment and the Pollution in Five ASEAN Nations," *International Journal of Economics and Management*, vol. 1, no. 2, pp. 245–261, Jun. 2007.

[20] C. G. Lee, "Foreign direct investment, pollution and economic growth: evidence from Malaysia," *Applied Economics*, vol. 41, no. 13, pp. 1709–1716, May 2009, https://doi.org/10.1080/00036840701564376.

[21] M. A. Cole, R. J. R. Elliott, and L. Zhang, "Foreign Direct Investment and the Environment," *Annual Review of Environment and Resources*, vol. 42, no. 1, pp. 465–487, 2017, https://doi.org/10.1146/annurev-environ-102016-060916.

[22] R. Hoffmann, C.-G. Lee, B. Ramasamy, and M. Yeung, "FDI and pollution: a granger causality test using panel data," *Journal of International Development*, vol. 17, no. 3, pp. 311–317, 2005, https://doi.org/10.1002/jid.1196.

[23] A. R. Ridzuan, A. H. M. Noor, and E. M. Ahmed, "The Impact of Foreign Direct Investment towards Carbon Dioxide Level: Pollution Havens Model for ASEAN5 Countries," *Asian Journal of Agricultural Extension, Economics & Sociology*, pp. 392–404, Jun. 2014, https://doi.org/10.9734/AJAEES/2014/10411.

[24] M. T. Rock, "Pollution intensity of GDP and trade policy: Can the World Bank be wrong?," *World Development*, vol. 24, no. 3, pp. 471–479, Mar. 1996, https://doi.org/10.1016/0305-750X(95)00152-3.

[25] X. Yao, R. Yasmeen, Y. Li, M. Hafeez, and I. U. H. Padda, "Free Trade Agreements and Environment for Sustainable Development: A Gravity Model Analysis," *Sustainability*, vol. 11, no. 3, Jan. 2019, Art. no. 597, https://doi.org/10.3390/su11030597.

[26] H. Zhang, "CO2 Emission Embodied in International Trade: Evidence for China," *International Journal of Economics and Finance*, vol. 7, no. 2, pp. 138-143, Jan. 2015, https://doi.org/10.5539/ijef.v7n2p138.

[27] Danish, M. A. Baloch, and S. Suad, "Modeling the impact of transport energy consumption on CO2 emission in Pakistan: Evidence from ARDL approach," *Environmental Science and Pollution Research*, vol. 25, no. 10, pp. 9461–9473, Apr. 2018, https://doi.org/10.1007/s11356-018-1230-0.

[28] B. Talbi, "CO2 emissions reduction in road transport sector in Tunisia," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 232–238, Mar. 2017, https://doi.org/10.1016/j.rser.2016.11.208.

[29] "Pakistan | Data," *The World Bank.* https://data.worldbank.org/country/pakistan?view=chart (accessed Aug. 21, 2021).

[30] H. Pesaran and Y. Shin, "An Autoregressive Distributed Lag Modeling Approach to Co-integration Analysis," *Econometncs and Economic Theory in the 20st Century: The Ragnar Frisch Centennial Symposium*, vol. 31, Feb. 1995, https://doi.org/10.1017/CCOL0521633230.011.

[31] M. H. Pesaran, Y. Shin, and R. J. Smith, "Bounds testing approaches to the analysis of level relationships," *Journal of Applied Econometrics*, vol. 16, no. 3, pp. 289–326, 2001, https://doi.org/10.1002/jae.616.

[32] H. Mahmood, "FDI, Population Density and Carbon Dioxide Emissions: A Case Study of Pakistan," *Iranica Journal of Energy & Environment*, vol. 3, no. 4, pp. 354–360, Jan. 2012, https://doi.org/10.5829/idosi.ijee.2012.03.04.09.

[33] R. Ali, K. Bakhsh, and M. A. Yasin, "Impact of urbanization on CO2 emissions in emerging economy: Evidence from Pakistan," *Sustainable Cities and Society*, vol. 48, Jul. 2019, Art. no. 101553, https://doi.org/10.1016/j.scs.2019.101553.

[34] P. Sadorsky, "The effect of urbanization on CO2 emissions in emerging economies," *Energy Economics*, vol. 41, pp. 147–153, Jan. 2014, https://doi.org/10.1016/j.eneco.2013.11.007.

[35] K. Hamid, "Is Foreign Direct Investment a Cause of Environmental Degradation in Pakistan? An ARDL Approach to Cointegration," *Journal of Management and Research*, vol. 3, no. 2, pp. 60–76, Dec. 2016, https://doi.org/10.29145/jmr/32/0302003.

Moist Air Flow Analysis in an Open Enclosure. Part A: Parametric Study

Tounsi Chati
Laboratory of Development in Mechanics and Materials
Ziane Achour University of Djelfa
Djelfa, Algeria
t.chati@univ.djelfa.dz

Kouider Rahmani
Modeling, Simulation and Optimization of Real Complex Systems Research Laboratory,University of Djelfa
Djelfa, Algeria
kouiderrah2@yahoo.fr

Toufik Tayeb Naas
Gas Turbine Joint Research Team
Ziane Achour University of Djelfa
Djelfa, Algeria
toufiknaas@gmail.com

Abdelkader Rouibah
Modeling, Simulation and Optimization of Real Complex Systems Research Laboratory,University of Djelfa
Djelfa, Algeria
rouibah_a@yahoo.fr

Abstract-Heat and mass transfer in many systems is widely accomplished by applying natural convection processes due to their low cost, reliability, and easy support. Typical applications include different mechanisms in various fields such as solar energy, solar distillers, stream cooling, etc. This paper examines the turbulent natural convection and mass transfer in an open enclosure for different Aspect Ratios (AR = 0.5, 1, and 2) with humid air. Mass fraction and local Nusselt number were examined to investigate the heat and mass transfer. Heat flux boundary conditions were subjected to the lateral walls, the bottom one was an adiabatic wall while the top area was a free surface. The effects of Rayleigh numbers ($106 \leq Ra \leq 108$) on natural convection and mass flow behavior were analyzed. The governing equations were solved using the CFD Fluent code based on the SIMPLE algorithm. Results showed that a cavity with an aspect ratio of AR=2 enhanced the rates of both heat and mass transfer. Maximum heat transfer rates were observed when the Rayleigh number increased due to the more vigorous fluid flow. However, mass transfer improved when the Rayleigh number decreased.

Keywords-Rayleigh number; turbulent flow; moist air; CFD; free convection

I. INTRODUCTION

Thermal turbulent flow is used in multiple applications of different mechanisms and systems such as solar energy collectors and evaporation, drying, distillation, and cooling processes of electronic compounds [1-9]. Mass and heat transfers induced by temperature and concentration gradients, are generally applied in physical processes that involve the convective and diffusive transport of chemical substances. Thereby, these heat and concentration gradients are considered to assist or oppose the performance of a system, depending on the boundary conditions of a problem. Many researchers tried to analyze extensively solutions on the heat transfer flow within closed or open cavities [10-17]. Some studies used heat flux or temperature gradient fields to the cavity walls as boundary conditions [18-20]. Several investigations and experiments were conducted for various structural configurations, such as rectangular and square cavities [21-25], and non-structure enclosures such as sinusoidal and inclined protuberances [14, 16, 19, 26]. The effects of Rayleigh [17, 24] and Reynolds numbers on the behavior of free, forced, and mixed convection were studied in [24-29]. Furthermore, various studies were conducted on the conditions of mass transfer and heat fluid motion for square cavities and rectangular shapes in a closed system [7, 30-36]. The heat and mass transfer induced by natural convection in a steady-state square cavity was studied in [31], proposing a two-dimensional cavity with different heated walls. The effects of the solute momentum on the rates of heat and mass transfer and Lewis number on fluid motion were investigated, suggesting correlations between the properties of dimensional quantities and rates of thermal and mass transfer. The natural double diffusion convection in a square cavity with horizontal walls was examined in [32], characterizing the mass transfer parameters through the local and global Nusselt and Sherwood numbers on all walls. The 3D aspect was compared to the 2D in [37], finding a satisfactory agreement for low values of the Rayleigh number.

Several researchers chose the rectangular configuration area to study the heat and mass transfer in natural convection [38-40]. The effect of humidity on the natural heat flow and their characteristics between vertical walls in a differentially heated cavity was discussed in [4, 41, 42]. A CFD study of heat and mass transfer by natural convection in a 2D area with an air-CO_2 mixture was presented in [43-44]. The effect of the buoyancy ratio on mass and heat transfer was investigated in [44] for different contexts. The thermal air quality behavior of higher CO_2 concentration in the vertical walls was analyzed under assisted and opposite flow conditions, considering different values for the Rayleigh number. The temperature

---

Corresponding author: Tounsi Chati

gradient inside the cavity for turbulent flow was observed more homogeneous for all flow cases.

The effect of mass on heat transfer with radiation on the double diffusion problems has been examined for Newtonian and non-Newtonian fluids in [45] and [46] respectively concluding that the radiative heat transfer inside the enclosure gave better thermal results and did not affect the mass transfer process. The humid air transfer for natural convection in a 2D rectangular surface with a solid section was investigated in [35], dealing with a set of different variables: Rayleigh thermal number, partition thickness, inclination, and buoyancy ratio. In [47], the impact of a heated plate on double-diffusive free convection in a square closed cavity was examined. The vertical walls maintained low temperatures while the horizontal walls were adiabatic, and the vortex viscosity values reduced both heat and mass transfer rates. A new vision on the mass transfer behavior of moisture under the buoyancy-driven effect for low turbulent flows was carried out in [35], using different values of mass fraction and varying temperature gradients between 20 and 60°C. The study concluded that the vapor content in the humid air had an important impact on heat and current fields inside the enclosure, and the variation of the vertical wall temperature gradients showed the effect of moisture concentration inside the cavity. The percentage of the difference in heat transfer changed significantly depending on the mass fraction of air moisture and the heat gradient between the vertical sides. The air-steam mixture flow of the natural convection inside a square cavity with vertical wall temperature and concentration gradients was investigated in [48], proposing a Prandtl number equal to 1.32 and a Rayleigh number equal to 4.05×10^5. The influence of species spread was investigated with Soret and Dufour mixture effects, finding that these effects had little contribution for this binary mixture.

A parametric study of a differentially heated enclosure to treat the natural convection flows that contained a mixture of air and carbon dioxide was conducted in [49], analyzing the effect of change in buoyancy rate on the mass transfer for a laminar system. A new procedure for a solar system was proposed in [3], considering a latent heat source of evaporation condensation in the energy equation. The results proved that the properties were uniformly distributed. The relative humidity was almost 100% in the whole enclosure which implied that the air was saturated everywhere. However, the sudden change in the temperature or mass fraction was near the evaporating surface. The free convection of thermo-solute with mass lines and thermal flow in a rectangular area was investigated in [36], where the vertical sides were isotherms and the others adiabatic. The study examined Prandtl number, Rayleigh number, buoyancy ratio, and Lewis number, finding that these parameters significantly affect the overall rates of heat and solute transfer. In [50], an experimental study was conducted to examine the combined thermal flow caused by mass transport and buoyancy in an open enclosure for various aspect ratios. The results of retained condensate on mass movement areas due to water's thermal stability were also examined with no remarkable impacts.

In [51], the heat and mass transfer inside a ventilated cavity in laminar flow was examined for fluid of air-CO_2 mixtures. A comparative study was conducted to examine the thermal behavior and the air quality inside a cavity, considering the exit space location of the fluid mixture and three different values for the CO_2 contaminant (1000, 2000, 3000ppm). The air inlet space was located below the right vertical side of the cavity, and the location of the mixture outlet was considered in four different positions. In [33], double diffusion natural convection for open square enclosures was investigated with constant concentrations and temperatures imposed along the right and the left walls. Results showed that the top thermal flow was the most effective for heat transfer, while the middle heating was the most stable for mass transfer. The effects of buoyancy ratio, Rayleigh number, Soret and Dufour coefficients, and Lewis number on heat and mass transfer in an open cavity were simulated in [52]. The results showed that changing the Lewis number had little influence on Sherwood and Nusselt numbers, but when the Rayleigh number increased Lewis number increased gradually.

Despite the importance of mass transfer flow, there are only a few research efforts on the natural convection of open cavities, especially with moist air as a working fluid. This paper investigates the natural convection inside an open cavity through humid-air passes for turbulent flow to examine the impact of Rayleigh number and aspect ratios on the heat and mass transfer at heat flux subjected to the vertical walls.

II. MATHEMATICAL MODELING

A. Geometrical Description

The considered physical problem is shown schematically in Figure 1. Humid-air occupies a two-dimensional rectangular enclosure. The vertical parallel walls are subjected to uniform flux (*q*) and the bottom side is considered impermeable and adiabatic. The top domain is defined as a free surface. The dimension of the open fluid at the bottom left wall was equal to 20% of the enclosure's height. Different cases of aspect ratios (AR=*H*/*L*=0.5, 1, and 2) were proposed.

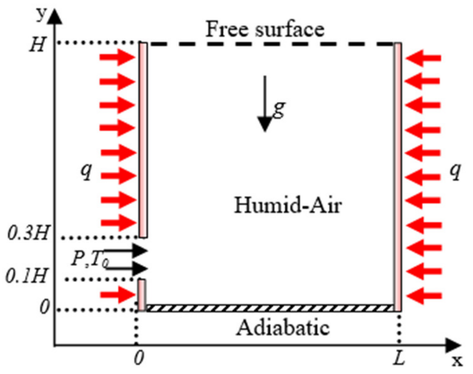


Fig. 1. A schematic representation of an open cavity.

B. Governing Equations

The governing equations for heat and mass transfer inside this open cavity are the continuity, momentum, energy, and concentration of humid-air liquid equation for natural convection [30]:

- Continuity equation:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \quad (1)$$

- Momentum equations:

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial x}+\upsilon\left(\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial x^2}\right) \quad (2)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial y}+\upsilon\left(\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial x^2}\right)+g\beta(T-T_0) \quad (3)$$

- Energy equation:

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=\alpha\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}\right) \quad (4)$$

- Concentration equation:

$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}+v\frac{\partial C}{\partial y}=D\left(\frac{\partial^2 C}{\partial x^2}+\frac{\partial^2 C}{\partial y^2}\right) \quad (5)$$

where υ is the kinematic viscosity, β is the thermal expansion coefficient, P is the fluid pressure, α and D are the thermal and mass diffusivity, respectively. All the governing equations of heat and mass transfer in the proposed open cavity were solved in a turbulent regime using the CFD code. A second-order upwind scheme was selected to solve the concentration and momentum equations. The computations were simulated to be converged at 10^{-7} of root mean square residual values.

C. Turbulence Model

A standard simple model of turbulence was proposed in [53], named k–ε, which is based on the turbulent kinetic energy equations k and the rate of dissipation of turbulent kinetic energy ε. This model was used to analyze the flow and heat transfer calculations due to its robustness, and consistent precision. The generic form of this model is [16]:

$$\frac{\partial}{\partial t}(\rho k)+\frac{\partial}{\partial x}(\rho k u)= \\ \frac{\partial}{\partial y}\left[\left(\mu+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial y}\right]+G_k+G_b-\rho\varepsilon-Y_M+S_k \quad (6)$$

$$\frac{\partial}{\partial t}(\rho\varepsilon)+\frac{\partial}{\partial x}(\rho\varepsilon u)=\frac{\partial}{\partial y}\left[\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial y}\right]+ \\ C_{\varepsilon 1}.\frac{\varepsilon}{k}(G_k+C_{\varepsilon 3}G_b)-C_{\varepsilon 2}\rho\frac{\varepsilon^2}{k}+S_\varepsilon \quad (7)$$

where μ_t is the Eddy viscosity, σ_k and σ_ε are the turbulent Prandtl numbers for k and ε, S_k and S_ε are user-defined source terms, G_k represents the kinetic energy production of turbulence due to velocity gradients, G_b is the kinetic energy production of turbulence due to buoyancy, Y_M represents the contribution of the expansion of fluctuations, C_μ=0.09, $C_{\varepsilon 1}$=1.44, $C_{\varepsilon 2}$=1.92, σ_k=1, and σ_ε=1.3.

D. Boundary Conditions

The boundary conditions associated with the problem were:

- $x=0;\ 0\le y\le 0.1H$: $u=v=0,\ q=5w,\ C=1$
- $x=0;\ 0.1H\le y\le 0.3H$: $u=v=0,\ q=0,\ C=0$

  $T=T_0,\quad P=P_0$
- $x=0;\ 0.3H\le y\le H$: $u=v=0,\ q=5w,\ C=1$
- $x=L;\ 0\le y\le H$: $u=v=0,\ q=5w,\ C=1$
- $y=0;\ 0\le x\le L$: $u=v=0, \partial T/\partial y=0, \partial C/\partial y=0$
- $y=H;\ 0\le x\le L$: $u=v=0,\ q=5w\ \ C=0$

  $T=T_0,\quad \Delta P=\rho g H$

The following variables are dimensionless:

$$X^*=x/L,\quad Y^*=y/L,\quad T^*=(T-T_C)/(T_H-T_C)$$

E. Meshing Sensitivity Analysis

The region was covered with a meshing grid having a further refinement in vertical sides. Different grids were used: 140×140, 160×160, 180×180, and 200×200 for a cavity with aspect ratio *AR=1*. From Table I, it is clear that the 180×180 mesh produced more satisfactory results than the others, for a maximum error of 0.02% and maximum values for the velocity components *u* and *v*.

TABLE I. MESH SENSITIVITY TEST FOR MAXIMUM X, Y-VELOCITIES, AND TEMPERATURE.

	140×140	**160×160**	**180×180**	**200×200**
***umax* (m/s)**	0.100652 (0.03%)	0.100668 (0.01%)	0.100673 (0.00%)	0.100678
***vmax* (m/s)**	0.087004 (0.05%)	0.086937 (0.03%)	0.0869814 (0.02%)	0.086962
***Tmax* (K)**	304.914 (0.005%)	304.92 (0.003%)	304.9246 (0.001%)	304.929

III. RESULTS AND DISCUSSION

The effects of aspect ratios (*AR*=0.5, 1 and 2) and Rayleigh number ($10^6\le Ra\le 10^8$) on heat and mass transfer of turbulent humid-air flow for an open enclosure were investigated parametrically. The thermophysical properties of the working fluid were [54]:

- $\rho_0=1.15\ [\mathrm{Kg.m^{-3}}]$
- $Cp_0=1005\ [\mathrm{J.Kg^{-1}.K^{-1}}]$
- $\lambda_0=2.65.10^{-2}\ [\mathrm{W.m^{-1}.s^{-1}}]$
- $\mu_0=1.85.10^{-5}\ [\mathrm{Kg.m^{-1}.s^{-1}}]$
- $D=2.59.10^{-5}\ [\mathrm{m^2.s^{-1}}]$.

A. Validation

The flow visualization of isotherms, streamlines, and iso-concentration results for the heat and mass transfer ($Ra=10^7$ with different buoyancy ratios) were compared with those obtained by [49], as shown in Figure 2. A second validation was investigated with results from [31, 37, 51], concerning the Nusselt number as a function of the buoyancy ratio (*N*). The validations were satisfactory exhibiting good agreement, as shown in Table II.

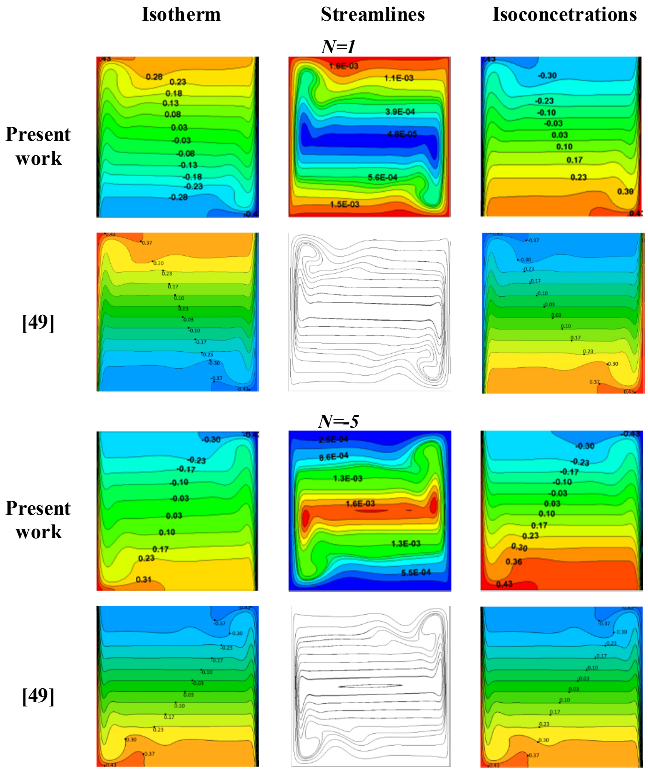


Fig. 2. Qualitative validation: Comparison of isotherm, streamline and iso-concentration contours for $Ra=10^7$ and $N=1$ and -5.

TABLE II. QUANTITATIVE VALIDATION; NUSSELT NUMBER AS FUNCTION OF BUOYANCY RATIO.

	Nu		
N	**-0.5**	**-0.8**	**-5.0**
Present work	13.6	10.3	23.2
[51]	13.6	10.6	23.7
[37]	13.5	10.5	23.5
[31]	13.6	10.6	23.7
Err % for [51]	0.7%	1.9%	1.2%

B. Discussion

Isotherms, streamlines, and mass fraction of H_2O visualizations are illustrated in Figures 3, 4, and 5 to examine the hydrodynamic behavior and the heat transfer characteristics for AR=0.5, 1, and 2, respectively. The main flow tended to cross the cavity in a semi-vertical direction, creating a strong vortex at the bottom right wall of each configuration and indicating an accelerating mass transfer. The heat transfer efficiency for the natural convection of humid-air inside open enclosures was influenced by the Rayleigh number at the lateral walls for all aspect ratios. For $AR=0.5$, the small cell found at the middle of the free surface was lost by the effect of enhanced buoyancy forces. Another clockwise small rotating vortex was created at the left bottom wall, where its size became increasingly powerful, while the high temperatures were localized in close spaces near the humid left wall, which corresponded to the concentration of the thermal boundary layer. For the square enclosure, three cells started to develop compared to $AR=2$ geometries. The contours of the H_2O mass fraction for all aspect ratios showed that the mass transfer dominated more as the Rayleigh number decreased. A close observation of the isotherms and hydrodynamics contours remarks that the mass and heat transfer processes were less concentrated at the top free surface. The concentration of H_2O increased significantly on the left wall of the enclosure, especially for $AR=2$.

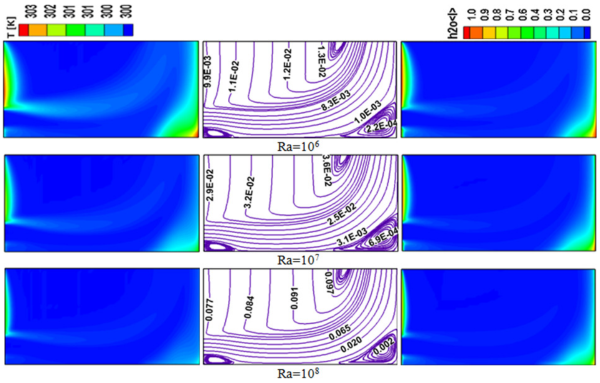


Fig. 3. Contours of isotherms (left), streamlines (middle), and mass fraction of H_2O liquid (right) at $AR = 0.5$ for different Rayleigh numbers ($10^6 \leq Ra \leq 10^8$).

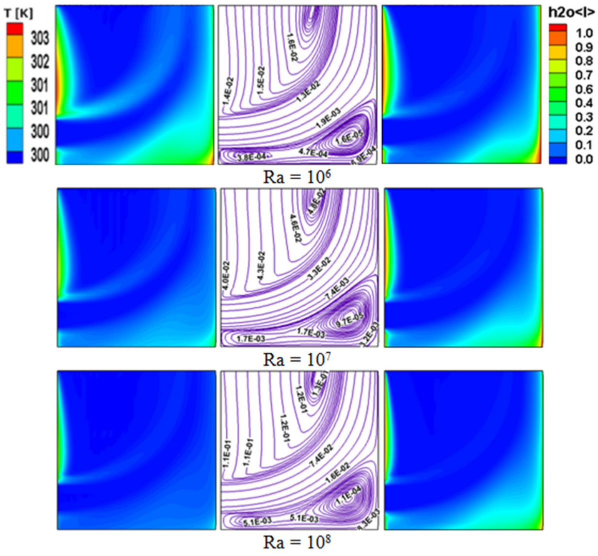


Fig. 4. Contours of isotherms (left), streamlines (middle), and mass fraction of H_2O liquid (right) at $AR = 1$ for different Rayleigh numbers ($10^6 \leq Ra \leq 10^8$).

The effect of the Rayleigh number with a fixed wall heat flux on the enhancement of thermal flow and mass transfer are presented in Figures 6 and 7 respectively. The local Nusselt number and mass fraction indicate the efficiency of the heat and mass transfer processes. As the height of the cavity has a significate role, the greater heat and mass transfers were found at AR=2. For low values of the Rayleigh number, the mass fraction of the humid-air increased while the Nusselt number decreased due to the slowly moving fluid. The heat transfer increased with the Rayleigh number, showing optimum thermal performance due to the vigorous turbulent flow.

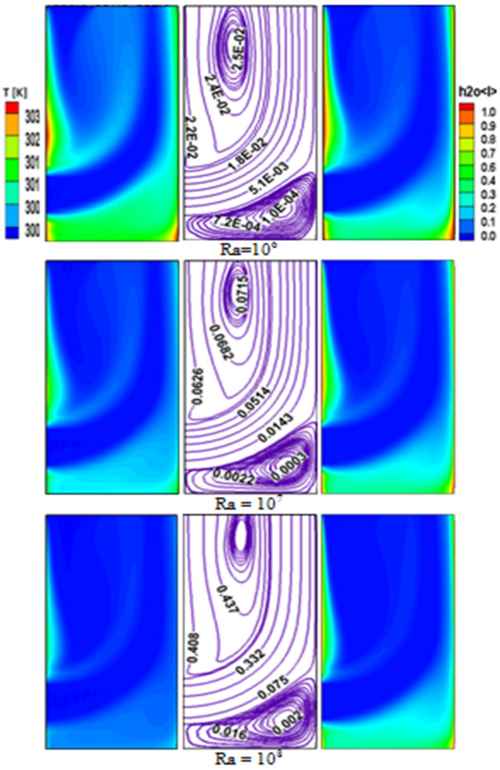


Fig. 5. Contours of isotherms (left), streamlines (middle), and mass fraction of H_2O liquid (right) at *AR*=2 for different Rayleigh numbers ($10^6 \leq Ra \leq 10^8$).

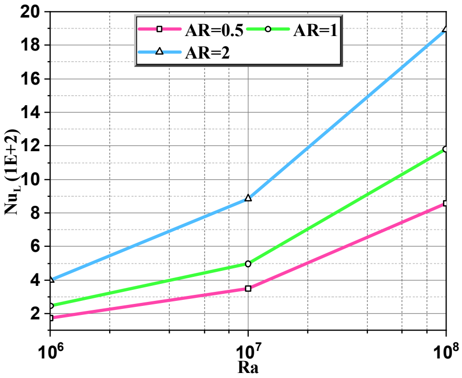


Fig. 6. Evolution of local Nusselt number as a function of Rayleigh number with different aspect ratios (*AR*= 0.5,1, and 2).

IV. CONCLUSIONS

This paper examined the effective mass and heat transfer efficiency for the natural convection of humid air inside open enclosures, concluding the following:

- Heat transfer efficiency for natural convection of humid air inside open enclosures is influenced by the Rayleigh number at the lateral walls for all aspect ratio cases. The mass transfer decreases because of vaporization. Moreover, as the Rayleigh number increased the flow visualization revealed that the vortex created in each enclosure had more vigorous intensity.

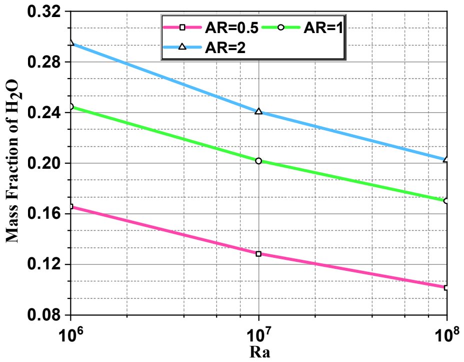


Fig. 7. Mass fraction of H_2O for various Rayleigh numbers with different aspect ratios (*AR* = 0.5, 1, and 2).

- The opening surface enhanced the mass transfer, while the bottom wall was the most stable for a few mass transfers.
- The mass transfer was reduced for the square cavity compared to the preferable enclosure at aspect ratio *AR*=2. Furthermore, higher rates of the Rayleigh number increased the thermal enhancement rate and decreased the mass fraction of humid air.
- As a consequence, the thermal and mass transfer performance in the cavity with *AR*=2 was greater than in the other configurations for moist air as a working fluid.

ACKNOWLEDGMENT

This research was funded by the Ministry of Higher Education and Scientific Research, Algeria

REFERENCES

[1] M. M. Rahman, H. F. Öztop, A. Ahsan, M. A. Kalam, and Y. Varol, "Double-diffusive natural convection in a triangular solar collector," *International Communications in Heat and Mass Transfer*, vol. 39, no. 2, pp. 264–269, Feb. 2012, https://doi.org/10.1016/j.icheatmasstransfer.2011.11.008.

[2] W. Aich, "3D Buoyancy Induced Heat Transfer in Triangular Solar Collector Having a Corrugated Bottom Wall," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2651–2655, Apr. 2018, https://doi.org/10.48084/etasr.1857.

[3] M. Keshtkar, M. Eslami, and K. Jafarpur, "A novel procedure for transient CFD modeling of basin solar stills: Coupling of species and energy equations," *Desalination*, vol. 481, May 2020, Art. no. 114350, https://doi.org/10.1016/j.desal.2020.114350.

[4] H. Sun, G. Lauriat, and X. Nicolas, "Natural convection and wall condensation or evaporation in humid air-filled cavities subjected to wall temperature variations," *International Journal of Thermal Sciences*, vol. 50, no. 5, pp. 663–679, May 2011, https://doi.org/10.1016/j.ijthermalsci.2010.12.010.

[5] N. Alleborn, H. Raszillier, and F. Durst, "Lid-driven cavity with heat and mass transport," *International Journal of Heat and Mass Transfer*, vol. 42, no. 5, pp. 833–853, Mar. 1999, https://doi.org/10.1016/S0017-9310(98)00224-5.

[6] L. B. Snoussi, R. Chouikh, and A. Guizani, "Numerical study of the natural convection flow resulting from the combined buoyancy effects of thermal and mass diffusion in a cavity with differentially heated side walls," *Desalination*, vol. 182, no. 1, pp. 143–150, Nov. 2005, https://doi.org/10.1016/j.desal.2005.03.014.

[7] R. Alvarado-Juárez, G. Álvarez, J. Xamán, and I. Hernández-López, "Numerical study of conjugate heat and mass transfer in a solar still device," *Desalination*, vol. 325, pp. 84–94, Sep. 2013, https://doi.org/10.1016/j.desal.2013.06.027.

[8] K. B. Saleem, L. Koufi, A. K. Alshara, and L. Kolsi, "Double-diffusive natural convection in a solar distiller with external fluid stream cooling," *International Journal of Mechanical Sciences*, vol. 181, Art. no. 105728, Sep. 2020, https://doi.org/10.1016/j.ijmecsci.2020.105728.

[9] T. Ikeshoji, F. N. B. de Nahui, S. Kimura, and M. Yoneya, "Computer analysis on natural convection in thin-layer thermocells with a soluble redox couple: Part 2. E-I relation, electric power, heat flux and electrochemical heat pump," *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, vol. 312, no. 1, pp. 43–56, Aug. 1991, https://doi.org/10.1016/0022-0728(91)85143-D.

[10] M. N. Hasan, S. C. Saha, and Y. T. Gu, "Unsteady natural convection within a differentially heated enclosure of sinusoidal corrugated side walls," *International Journal of Heat and Mass Transfer*, vol. 55, no. 21, pp. 5696–5708, Oct. 2012, https://doi.org/10.1016/j.ijheatmasstransfer.2012.05.065.

[11] M. M. El-Gendi and A. M. Aly, "Numerical simulation of natural convection using unsteady compressible Navier-stokes equations," *International Journal of Numerical Methods for Heat & Fluid Flow*, vol. 27, no. 11, pp. 2508–2527, Jan. 2017, https://doi.org/10.1108/HFF-10-2016-0376.

[12] I. Kouroudis, P. Saliakellis, and S. G. Yiantsios, "Direct numerical simulation of natural convection in a square cavity with uniform heat fluxes at the vertical sides: Flow structure and transition," *International Journal of Heat and Mass Transfer*, vol. 115, pp. 428–438, Dec. 2017, https://doi.org/10.1016/j.ijheatmasstransfer.2017.08.058.

[13] A. Fabregat and J. Pallarès, "Heat transfer and boundary layer analyses of laminar and turbulent natural convection in a cubical cavity with differently heated opposed walls," *International Journal of Heat and Mass Transfer*, vol. 151, Apr. 2020, Art. no. 119409, https://doi.org/10.1016/j.ijheatmasstransfer.2020.119409.

[14] M. Prakash, S. B. Kedare, and J. K. Nayak, "Numerical study of natural convection loss from open cavities," *International Journal of Thermal Sciences*, vol. 51, pp. 23–30, Jan. 2012, https://doi.org/10.1016/j.ijthermalsci.2011.08.012.

[15] B. Zamora and A. S. Kaiser, "Influence of the variable thermophysical properties on the turbulent buoyancy-driven airflow inside open square cavities," *Heat and Mass Transfer*, vol. 48, no. 1, pp. 35–53, Jan. 2012, https://doi.org/10.1007/s00231-011-0838-0.

[16] K. Rahmani, A. Bentebbiche, and B. Draoui, "Numerical Investigation of Turbulent Natural Convection for a Cavity Having Sinusoidal Protuberances on a Vertical Wall," *Journal of Applied Fluid Mechanics*, vol. 6, no. 04, pp. 491–199, Oct. 2013, https://doi.org/10.36884/jafm.6.04.21278.

[17] R. Choudhary, A. Saini, and S. Subudhi, "Oberbeck-Boussinesq approximations and geometrical confinement effects of free convection in open cavity," *Heat and Mass Transfer*, vol. 55, no. 8, pp. 2095–2102, Aug. 2019, https://doi.org/10.1007/s00231-019-02563-8.

[18] V. M. Maytorena, A. Piña-Ortiz, and J. F. Hinojosa, "Experimental and numerical study of turbulent natural convection in an open cubic cavity," *Heat and Mass Transfer*, vol. 51, no. 9, pp. 1205–1217, Sep. 2015, https://doi.org/10.1007/s00231-014-1492-0.

[19] T. T. Naas, Y. Lasbet, and C. Kezrane, "Entropy Generation Analyze Due to the Steady Natural Convection of Newtonian Fluid in a Square Enclosure," *International Journal of Mechanical and Mechatronics Engineering*, vol. 9, no. 4, pp. 582–586, Mar. 2015.

[20] D. Saury, N. Rouger, F. Djanna, and F. Penot, "Natural convection in an air-filled cavity: Experimental results at large Rayleigh numbers," *International Communications in Heat and Mass Transfer*, vol. 38, no. 6, pp. 679–687, Jul. 2011, https://doi.org/10.1016/j.icheatmasstransfer.2011.03.019.

[21] Y. S. Tian and T. G. Karayiannis, "Low turbulence natural convection in an air filled square cavity: Part I: the thermal and fluid flow fields," *International Journal of Heat and Mass Transfer*, vol. 43, no. 6, pp. 849–866, Mar. 2000, https://doi.org/10.1016/S0017-9310(99)00199-4.

[22] Y. S. Tian and T. G. Karayiannis, "Low turbulence natural convection in an air filled square cavity: Part II: the turbulence quantities," *International Journal of Heat and Mass Transfer*, vol. 43, no. 6, pp. 867–884, Mar. 2000, https://doi.org/10.1016/S0017-9310(99)00200-8.

[23] F. Ampofo and T. G. Karayiannis, "Experimental benchmark data for turbulent natural convection in an air filled square cavity," *International Journal of Heat and Mass Transfer*, vol. 46, no. 19, pp. 3551–3572, Sep. 2003, https://doi.org/10.1016/S0017-9310(03)00147-9.

[24] A. Vasiliev *et al.*, "High Rayleigh number convection in a cubic cell with adiabatic sidewalls," *International Journal of Heat and Mass Transfer*, vol. 102, pp. 201–212, Nov. 2016, https://doi.org/10.1016/j.ijheatmasstransfer.2016.06.015.

[25] Z. Altaç and N. Uğurlubilek, "Assessment of turbulence models in natural convection from two- and three-dimensional rectangular enclosures," *International Journal of Thermal Sciences*, vol. 107, pp. 237–246, Sep. 2016, https://doi.org/10.1016/j.ijthermalsci.2016.04.016.

[26] A. Latreche and M. Djezzar, "Numerical Study of Natural Convective Heat and Mass Transfer in an Inclined Porous Media," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3223–3227, Aug. 2018, https://doi.org/10.48084/etasr.2179.

[27] Y. Varol and H. F. Oztop, "Free convection in a shallow wavy enclosure," *International Communications in Heat and Mass Transfer*, vol. 33, no. 6, pp. 764–771, Jul. 2006, https://doi.org/10.1016/j.icheatmasstransfer.2006.02.004.

[28] M. El-Gendi, "Transient turbulent simulation of natural convection flows induced by a room heater," *International Journal of Thermal Sciences*, vol. 125, pp. 369–380, Mar. 2018, https://doi.org/10.1016/j.ijthermalsci.2017.12.012.

[29] L. Koufi, Z. Younsi, Y. Cherif, and H. Naji, "Numerical investigation of turbulent mixed convection in an open cavity: Effect of inlet and outlet openings," *International Journal of Thermal Sciences*, vol. 116, pp. 103–117, Jun. 2017, https://doi.org/10.1016/j.ijthermalsci.2017.02.007.

[30] V. Sévéléder and J. P. Petit, "Flow Structures Induced by Opposing Forces in Double-Diffusive Natural Convection in a Cavity," *Numerical Heat Transfer, Part A: Applications*, vol. 15, no. 4, pp. 431–444, Jun. 1989, https://doi.org/10.1080/10407788908944697.

[31] C. Béghein, F. Haghighat, and F. Allard, "Numerical study of double-diffusive natural convection in a square cavity," *International Journal of Heat and Mass Transfer*, vol. 35, no. 4, pp. 833–846, Apr. 1992, https://doi.org/10.1016/0017-9310(92)90251-M.

[32] V. A. F. Costa, "Double diffusive natural convection in a square enclosure with heat and mass diffusive walls," *International Journal of Heat and Mass Transfer*, vol. 40, no. 17, pp. 4061–4071, Oct. 1997, https://doi.org/10.1016/S0017-9310(97)00061-6.

[33] N. Arbin and I. Hashim, "Partial heating and partial salting on double-diffusive convection in an open cavity," *AIP Conference Proceedings*, vol. 1614, no. 1, pp. 891–897, Sep. 2014, https://doi.org/10.1063/1.4895320.

[34] J.-T. Hu, X.-H. Ren, D. Liu, F.-Y. Zhao, and H.-Q. Wang, "Natural convective heat and moisture transfer in an inclined building enclosure with one slender wall of finite thickness: Analytical investigation and non-unique steady flow solutions," *International Journal of Heat and Mass Transfer*, vol. 104, pp. 1160–1176, Jan. 2017, https://doi.org/10.1016/j.ijheatmasstransfer.2016.09.033.

[35] D. Iyi and R. Hasan, "Numerical investigation of the effect of moisture on buoyancy-driven low turbulence flow in an enclosed cavity," *International Journal of Heat and Mass Transfer*, vol. 136, pp. 543–554, Jun. 2019, https://doi.org/10.1016/j.ijheatmasstransfer.2019.02.102.

[36] D. Kushawaha, S. Yadav, and D. K. Singh, "Thermo-solute natural convection with heat and mass lines in a uniformly heated and soluted rectangular enclosure for low Prandtl number fluids," *International Journal of Thermal Sciences*, vol. 148, Art. no. 106160, Feb. 2020, https://doi.org/10.1016/j.ijthermalsci.2019.106160.

[37] I. Sezai and A. A. Mohamad, "Double diffusive convection in a cubic enclosure with opposing temperature and concentration gradients," *Physics of Fluids*, vol. 12, no. 9, pp. 2210–2223, Sep. 2000, https://doi.org/10.1063/1.1286422.

[38] Jae Min Hyun and J. W. Lee, "Double-diffusive convection in a rectangle with cooperating horizontal gradients of temperature and concentration," *International Journal of Heat and Mass Transfer*, vol. 33, no. 8, pp. 1605–1617, Aug. 1990, https://doi.org/10.1016/0017-9310(90)90017-O.

[39] J. W. Lee and Jae Min Hyun, "Double-diffusive convection in a rectangle with opposing horizontal temperature and concentration gradients," *International Journal of Heat and Mass Transfer*, vol. 33, no. 8, pp. 1619–1632, Aug. 1990, https://doi.org/10.1016/0017-9310(90)90018-P.

[40] Jin Wook Lee and Jae Min Hyun, "Time-dependent double diffusion in a stably stratified fluid under lateral heating," *International Journal of Heat and Mass Transfer*, vol. 34, no. 9, pp. 2409–2421, Sep. 1991, https://doi.org/10.1016/0017-9310(91)90065-M.

[41] O. Laguerre, S. Benamara, D. Remy, and D. Flick, "Experimental and numerical study of heat and moisture transfers by natural convection in a cavity filled with solid obstacles," *International Journal of Heat and Mass Transfer*, vol. 52, no. 25, pp. 5691–5700, Dec. 2009, https://doi.org/10.1016/j.ijheatmasstransfer.2009.07.028.

[42] O. Laguerre, S. Benamara, and D. Flick, "Numerical simulation of simultaneous heat and moisture transfer in a domestic refrigerator," *International Journal of Refrigeration*, vol. 33, no. 7, pp. 1425–1433, Nov. 2010, https://doi.org/10.1016/j.ijrefrig.2010.04.010.

[43] J. Serrano-Arellano, J. Xamán, G. Álvarez, and M. Gijón-Rivera, "Heat and mass transfer by natural convection in a square cavity filled with a mixture of Air–CO2," *International Journal of Heat and Mass Transfer*, vol. 64, pp. 725–734, Sep. 2013, https://doi.org/10.1016/j.ijheatmasstransfer.2013.05.038.

[44] R. Nikbakhti and A. B. Rahimi, "Double-diffusive natural convection in a rectangular cavity with partially thermally active side walls," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 43, no. 4, pp. 535–541, Jul. 2012, https://doi.org/10.1016/j.jtice.2012.02.010.

[45] J. Serrano-Arellano and M. Gijón-Rivera, "Conjugate heat and mass transfer by natural convection in a square cavity filled with a mixture of Air–CO2," *International Journal of Heat and Mass Transfer*, vol. 70, pp. 103–113, Mar. 2014, https://doi.org/10.1016/j.ijheatmasstransfer.2013.10.051.

[46] GH. R. Kefayati, "Simulation of double diffusive natural convection and entropy generation of power-law fluids in an inclined porous cavity with Soret and Dufour effects (Part I: Study of fluid flow, heat and mass transfer)," *International Journal of Heat and Mass Transfer*, vol. 94, pp. 539–581, Mar. 2016, https://doi.org/10.1016/j.ijheatmasstransfer.2015.11.044.

[47] M. Muthtamilselvan, K. Periyadurai, and D. H. Doh, "Impact of nonuniform heated plate on double-diffusive natural convection of micropolar fluid in a square cavity with Soret and Dufour effects," *Advanced Powder Technology*, vol. 29, no. 1, pp. 66–77, Jan. 2018, https://doi.org/10.1016/j.apt.2017.10.012.

[48] N. Jiang, E. Studer, and B. Podvin, "Physical modeling of simultaneous heat and mass transfer: species interdiffusion, Soret effect and Dufour effect," *International Journal of Heat and Mass Transfer*, vol. 156, Aug. 2020, Art. no. 119758, https://doi.org/10.1016/j.ijheatmasstransfer.2020.119758.

[49] L. Koufi, Y. Cherif, Z. Younsi, and H. Naji, "Double-Diffusive Natural Convection in a Mixture-Filled Cavity with Walls' Opposite Temperatures and Concentrations," *Heat Transfer Engineering*, vol. 40, no. 15, pp. 1268–1285, Sep. 2019, https://doi.org/10.1080/01457632.2018.1460928.

[50] W. Terrell and T. A. Newell, "Experimental techniques for determining heat and mass transfer due to condensation of humid air in cooled, open cavities," *Applied Thermal Engineering*, vol. 27, no. 8, pp. 1574–1584, Jun. 2007, https://doi.org/10.1016/j.applthermaleng.2006.09.022.

[51] J. Xamán, A. Ortiz, G. Álvarez, and Y. Chávez, "Effect of a contaminant source (CO2) on the air quality in a ventilated room," *Energy*, vol. 36, no. 5, pp. 3302–3318, May 2011, https://doi.org/10.1016/j.energy.2011.03.026.

[52] Z. Zhou, Z. Wang, and M. Yang, "Double Diffusive Natural Convection in Open Cavity Under the Soret and Dufour Effects," *Frontiers in Heat and Mass Transfer (FHMT)*, vol. 14, Feb. 2020.

[53] B. E. Launder and D. B. Spalding, *Lectures in mathematical models of turbulence.* New York, NY, USA: Academic Press, 1972.

[54] B. E. Poling, J. M. Prausnitz, and J. P. O'Connell, *Properties of Gases and Liquids, Fifth Edition*, 5th ed. New York, NY, USA: McGraw-Hill Education, 2001.

Spatiotemporal Rainfall Dynamics in Kosi Basin Using Wavelet Analysis

Aadil Towheed
Department of Civil Engineering
National Institute of Technology Patna
Patna, India
aadil.ce17@nitp.ac.in

Thendiyath Roshni
Department of Civil Engineering
National Institute of Technology Patna
Patna, India
roshni@nitp.ac.in

Abstract-Spatial and temporal analysis of rainfall data were carried out along with wavelet analysis for seven rain gauge sites of Kosi basin, India during the time period from 1985 to 2017. Wavelet spectrum analysis and wavelet coherence analysis were performed to fully characterize the time-frequency rainfall variability of the rain gauge data in these areas. For all the selected gauge stations during the study period, the peak value of the wavelet power spectrum was identified for the 8-16 month band. The results of wavelet spectrum analysis reveal a good correlation of rainfall data in the rain gauge sites lying in the southwest of the Kosi basin. The spectrum analysis also differentiates the wet and dry periods and it was observed that in the majority of the selected sites, a dry period occurred from the year 2005 onwards. This was again confirmed with breakpoint analysis. The wavelet coherence analysis explicit is a good correlation between the rain gauges in the study area. Overall, the variability of the rainfall parameters was more vivid with the wavelet analysis and this can be extended to other climatological parameters.

Keywords-wavelet spectrum; wavelet coherence; break point; rainfall analysis

I. INTRODUCTION

The extremes of precipitation and its patterns are affected by various climatic factors and the circulation patterns in the atmosphere [1-4]. Non-stationary techniques to analyze the variation of rainfall with respect to time and space were used in [5]. The climate is mathematically defined as the weather conditions for a period of 30 years in a particular region, however it is not defined for a specific duration. The weather is defined on daily basis and is a non-linear dynamic system of the atmosphere whereas climate is considered to be the average form of weather [6]. There is a huge impact of the rainfall variation to the agricultural output in India which thereby adversely affects food security [7], Gross Domestic Product (GDP), and inflation [8]. Authors in [9] analyzed the temporal variation of precipitation in Shaheed Benazir Abad district by using the linear regression method. Rainfall events may cause the intensification of frequent extreme events which make the area more prone to floods and droughts [10]. Precipitation plays an essential role in hydro-climatic studies and daily life activities prominently in developing countries [11, 12]. Authors in [13] focused on models to assess rainfall distribution in Kelantan, Malaysia. Authors in [14] focused on drought frequency and its intensity in the past decades over Asia and Africa. Higher changes in precipitation events may cause increase in the span of drought periods [15]. The pattern and trend of long term precipitation extreme events are essential in the Kosin basin area which is highly prone to hazardous floods [16]. The economy and the development are dependent on the availability of water resources in the area [16].

Wavelet Transforms (WTs) of the hydro-meteorological parameters have recently drawn the attention of water resource experts. They are commonly used in data analysis and modeling [17] of hydrological and meteorological parameters. For the modeling of hydrological processes, advanced Artificial Intelligence (AI) techniques can act as robust tools [18]. During the past decades, it was observed that different AI-based models have been applied for time series analysis of rainfall data [19]. Authors in [20] used Continuous Wavelet Transform (CWT) and the multiscale entropy concept for the time series analysis of monthly rainfall data. As similar patterns are involved in the rainfall data series and due to the high potential of WT in multi-scale analysis of signals, the concept of wavelet analysis in the hydro-meteorological field can be used for better data pre-processing [21]. The variables are well identified with the help of correlation analysis and wavelet coherence [22].

In this study, WT was used as it steadily maintains the time-frequency localization in signal analysis. It also alters the 1D time series data into a 2D time frequency image [23]. WT analysis is also applied to signals to get detail information, which cannot be readily read from the raw signal. The objectives of the current study are: (i) to investigate the rainfall data periodically and temporally over Kosi basin synoptically and (ii) to conduct wavelet coherence analysis between the rain gauge stations of Kosi basin in order to check the correlation between the rainfall data of the different rain gauge sites for the selected time period.

II. METHODOLOGY

A. *Wavelet Analysis*

This section discusses in detail the analysis of the wavelet power spectrum.

1) *Wavelet Transform*

Corresponding author: Aadil Towheed

The time series data can be analyzed at several frequencies using the WT [24]. Assuming a time series data set, x_n, where n varies from 0 to N-1, a wavelet function $\psi_o(\eta)$ being the function of time as parameter (η). The Morlet wavelet involves an exponential wave which is modulated by a Gaussian envelope as:

$$\psi_o(\eta) = \pi^{-1/4} e^{i\omega_o\eta} e^{-\eta^2/2} \quad (1)$$

where ω_o is non-dimensional frequency whereas its value is taken as 6 in the present work [25]. To be allowable as a basic wavelet, function (1) implies zero mean and is localized in both time and frequency spaces.

For a discrete sequence x_n, which is separated by a constant interval of time δt, the CWT is defined for x_n as the convolution with a translated and scaled form of $\psi_o(\eta)$. It is the product of the wavelet function with the original time series and is written as:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^* \left[\frac{(n'-n)\delta t}{s}\right] \quad (2)$$

where s indicates the wavelet scale which is translating along the localized time index n, (*) indicates the complex conjugate, and N is the number of points in time series data. Using (2) for time series analysis, this basic wavelet function is transformed to a new wavelet in time. The scaled wavelets s in (3) is used to change the scale [22, 25, 26]:

$$\psi\left[\frac{(n'-n)\delta t}{s}\right] = \left(\frac{\delta t}{s}\right)^{1/2} \psi_o\left[\frac{(n'-n)\delta t}{s}\right] \quad (3)$$

2) *Wavelet Power Spectrum*

Since wavelet function $\psi_o(\eta)$ and wavelet transform $W_n(s)$ are complex, the WT is divided into a real and an imaginary part. The wavelet power spectrum is now defined as $[W_n(s)]^2$ (detailed description is given in [25]). The parameters used for the analysis of rainfall data are δt =1 month, S_0=2 δt, δj =0.25, (shows 4 sub-octaves per octave), j_1=7/ δj (7 powers of two with δj sub-octaves).

3) *Global Wavelet Power Spectrum*

The global wavelet power spectrum is the square of the WT. The time-averaged wavelet spectrum over a period is given by:

$$W_n^2(s) = \frac{1}{N}\sum_{n=0}^{N} |W_n(s)|^2 \quad (4)$$

4) *Scale-Average Time Series*

The time series of scale average is the average variance in a certain band. In the present work the selected band is the 8-16 month band after the analysis.

B. *Wavelet Coherence Analysis (WCA)*

WCA is applied generally in catchment studies, space and time analysis, and for model validation. The WCA examines the correlation and phase lag between two series of time data as a function of both time and frequency [27]. Authors in [28] defined wavelet coherence for *X(n)* and *Y(n) as* two time series.

$$R_n^2(s) = \frac{\left|S(s^{-1}W_n^{XY}(s)\right|^2}{S(s^{-1}\left|W_n^X(s)\right|^2).S(s^{-1}\left|W_n^Y(s)\right|^2)} \quad (5)$$

where S is a smoothing operator and $W_n^x(s)$ and $W_n^y(s)$ are wavelet transforms of *X(n)* and *Y(n)* time series. $W_n^{XY}(s) = W_n^X(s) W_n^{Y*}(s)$ is the cross-wavelet transform, where (*) represents the complex conjugate and $W_n^x(s)$ is the cross wavelet power [17]. The statistical significance level of WCA is estimated with Monte Carlo methods. The value of R^2 ranges from 0 to 1, and is conceptualized as a localized correlation coefficient with time and frequency space. WCA for the signal comparison is carried out using the MATLAB functions provided by [28].

III. STUDY AREA AND AVAILABLE DATA

The Kosi river enters Bihar near Bhimnagar after draining a large area in Nepal. The drainage area (~52,731km^2) includes parts of Tibet, Nepal, and India [29]. The Kosi river is one of the most flood-prone rivers in India and is located in north Bihar. The study area extends between 25^{0}19'18"N - 26^{0}43'30"N latitude and 87^{0}4'35"E - 87^{0}12'32"E longitude. Some major tributaries of the Kosi river are Kamlabalan, Bagmati, and Bhutibalan. The Shuttle Radar Topography Mission (SRTM) based Digital Elevation Model (DEM) was used for the determination of elevation (maximum 109m) of the study area. There are 10 districts within the Kosi river basin while the degree of slope varies from 0 to 35^o. The spatial variation of rainfall and subsequently its effect in the soil erosion in the Kosi river was well studied in [30]. The main reason to choose this study area is that it is very vulnerable to the adverse effects of floods.

The teleconnections of the different regions within the study area are quite significant, as India has different climatic conditions and rainfall patterns [31]. Authors in [32] utilized the precipitation data to look into the effects of teleconnections on extreme precipitation for a period of 113 years. For the current study, the daily rainfall data were collected from India Meteorological Department (IMD), Pune for a duration of 33 years (1985 – 2017) for 6 raingauge stations (Kursela, Galgalia, Murliganj, Nirmali, Bhimnagar, and Bahadurganj) and 1985-2015 for Birpur station [30]. Concentration of rainfall occurs high in the monsoon season (from May to October) within the study area. The maximum average annual rainfall is 2,711.95mm obtained in Galgalia, and the least is 1,249.13mm in Bhimnagar [30].

IV. RESULTS AND DISCUSSION

A. *Wavelet Transform*

The WT is computed with the help of the MATLAB script developed in [25]. In Kosi basin, the CWT is performed for the monthly data of 7 gauging stations. For the computation of (3), δt and s used were 1 month and 2 months (as $s = 2\delta t$), 7 powers of 2 with 4 sub-octaves per octave and 'Morlet' as mother wavelet were used. The time series plots of monthly rainfall for the 7 gauging stations are shown in Figures 1(a)-7(a) for 395 months (1985-2017), except for the Birpur station (1985-2015). Figures 1(b)-7(b) show the power (squared absolute value) of the WT for the monthly rainfall data. This provides information on relative power at a certain scale and it displays the actual oscillations of the individual wavelets. Figures 1(c)-7(c) show the global wavelet spectrum and Figures 1(d)-7(d) show the variance of the power for a particular band of interest for the respective stations.

Figure 1(b) is evaluated for different frequency bands varying from 2 to 256 months for Bahadurganj data. The figures with color fill show the variation of rainfall in which the ranges of power are from weak (deep blue shades) to strong (dark red shades). Thus, the time series differences were mapped to wavelet region and to various scales. It is worth noting that the power is more for the 8-16 month band throughout the selected period and it is confirmed with the peak in the global wavelet power spectrum (Figure 1(c)). The scale average time series of average variance in the 8-16 month band is shown in Figure 1(d). The variance plot shows distinct dry and wet periods. The power decreases substantially during a dry year and maximum power means a wet year [34, 35]. In Bahadurganj, wet periods are observed more than dry periods throughout the studied duration. Similarly for Figures 2(b)-7(b), it is clearly visible that the power is more for the 8-16 month band throughout the selected time duration for all the selected sites and it is also confirmed with the peak in the global wavelet power spectrum (Figures 1(c)-7(c)). Hence, the scale average time series of average variance is plotted for the 8-16 month band only for the selected time period (Figures 1(d)-7(d)). In Bhimnagar, a significant wet period is observed during the 200-250 week period. However, wet periods are observed up to 240 months (from year 1985), with a few breakdowns during the months 60-100 in Birpur.

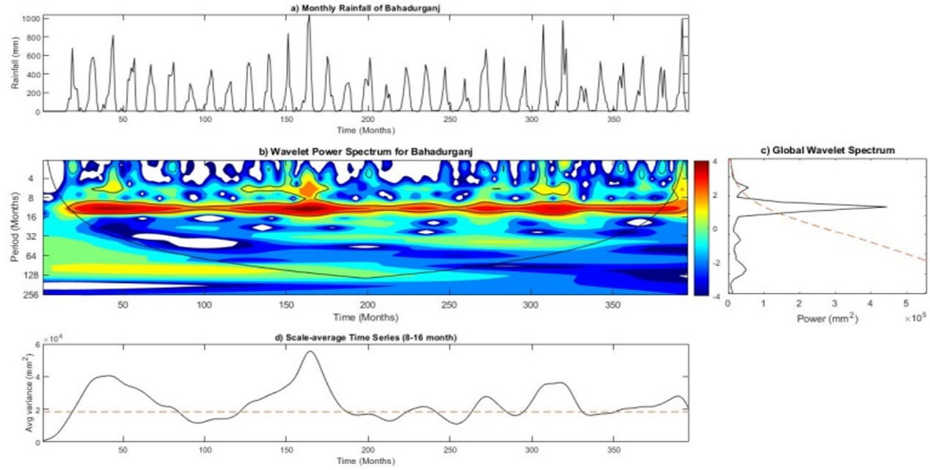


Fig. 1. Bahadurganj station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data with Morlet as mother-wavelet. (c) The global wavelet spectrum (black continuous line) and significance level 5% (broken red line). (d) 8-16 month band scale-average time series of wavelet power (black continuous line). The broken red line is the 95% confidence level.

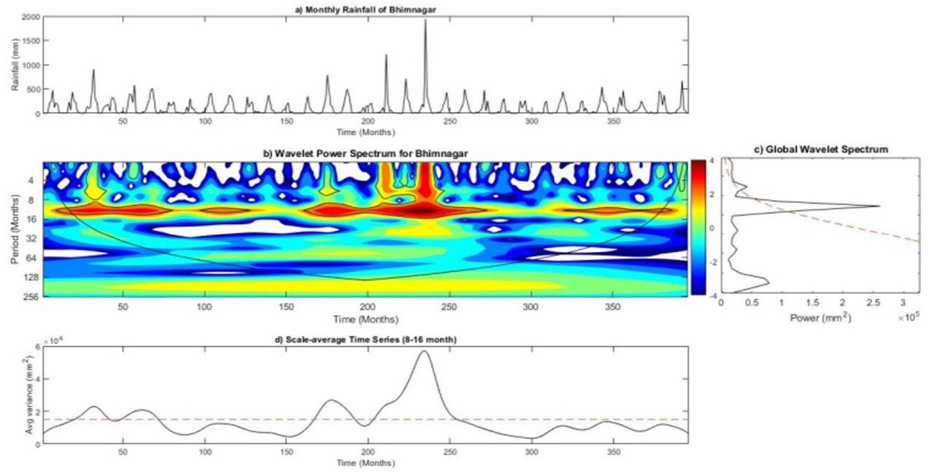


Fig. 2. Bhimnagar station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data with Morlet as mother-wavelet. (c) The global wavelet spectrum (black continuous line) and significance level 5% (broken red line). (d) 8-16 month band scale-average time series of wavelet power (black continuous line). The broken red line is the 95% confidence level.

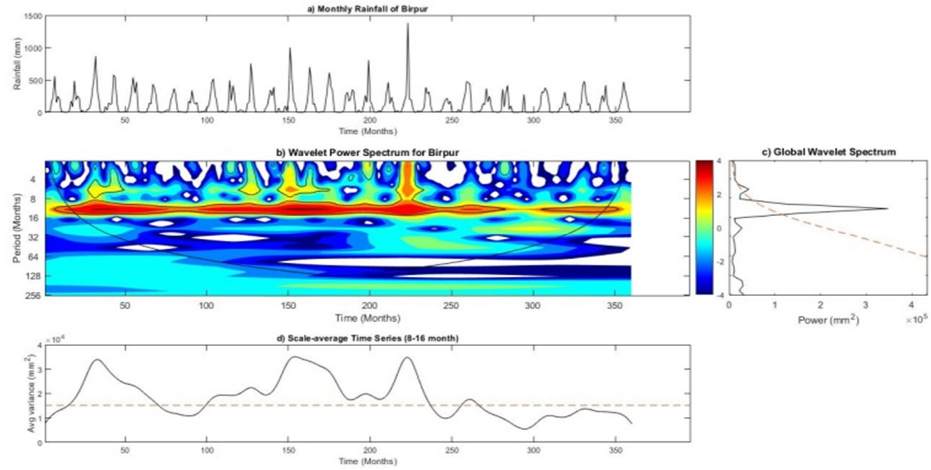


Fig. 3. Birpur station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data. (c) The global wavelet spectrum and significance level 5%. (d) 8-16 month band scale-average time series of wavelet power. The broken red line is the 95% confidence level.

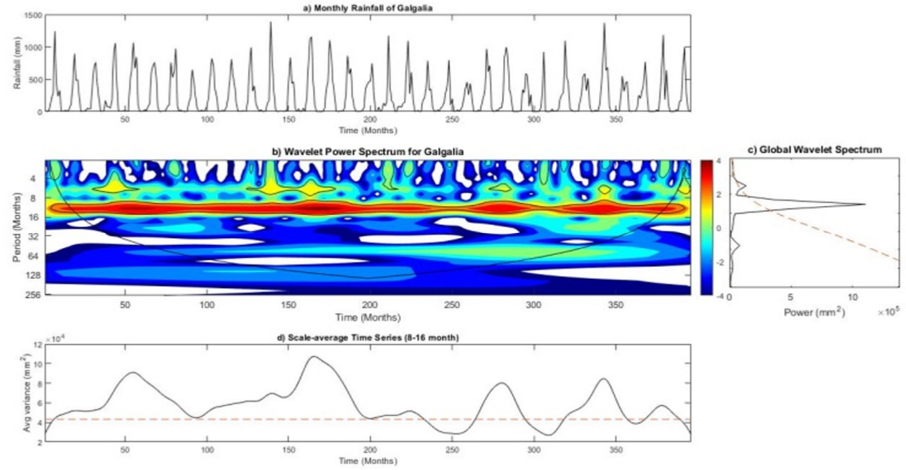


Fig. 4. Galgalia station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data. (c) The global wavelet spectrum and significance level 5%. (d) 8-16 month band scale-average time series of wavelet power. The broken red line is the 95% confidence level.

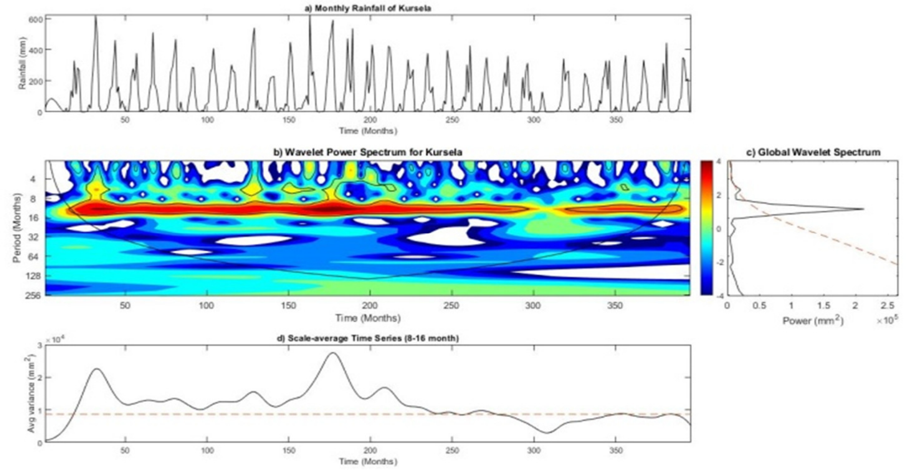


Fig. 5. Kursela station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data. (c) The global wavelet spectrum and significance level 5%. (d) 8-16 month band scale-average time series of wavelet power. The broken red line is the 95% confidence level.

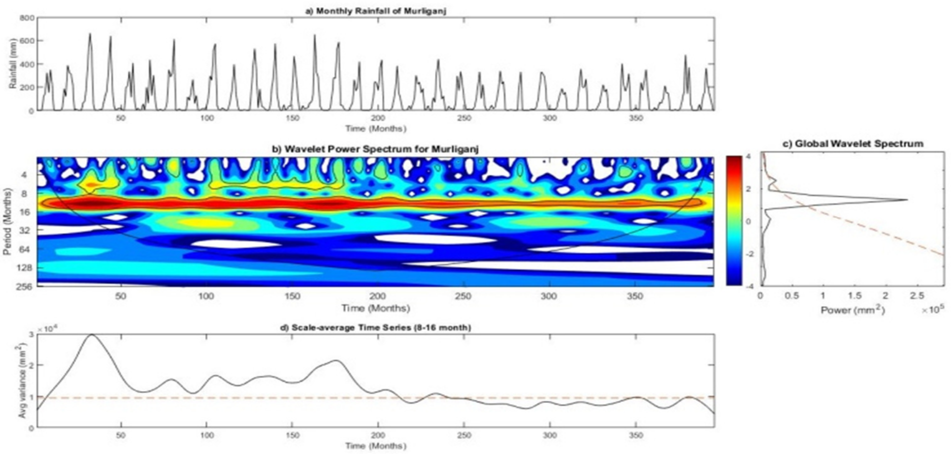


Fig. 6. Murliganj station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data. (c) The global wavelet spectrum and significance level 5%. (d) 8-16 month band scale-average time series of wavelet power. The broken red line is the 95% confidence level.

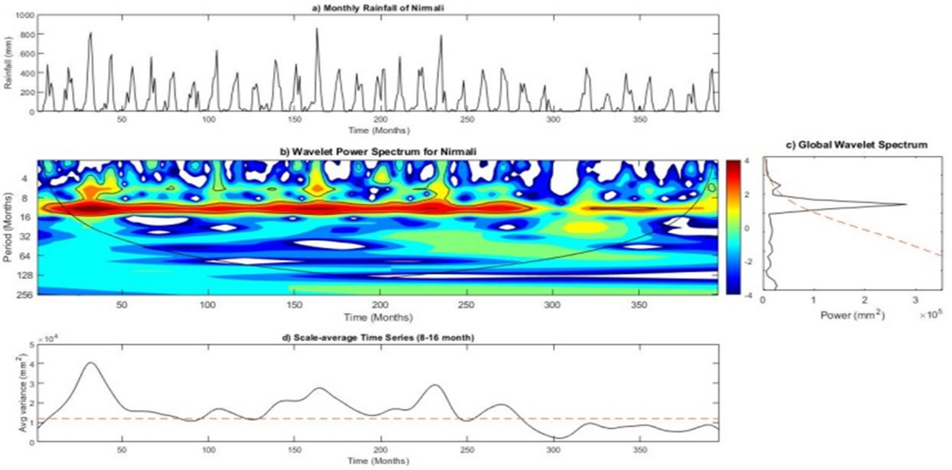


Fig. 7. Nirmali station: (a) Time series plot of monthly rainfall. (b) Analysis of the wavelet power spectrum for monthly rainfall data. (c) The global wavelet spectrum and significance level 5%. (d) 8-16 month band scale-average time series of wavelet power. The broken red line is the 95% confidence level.

In Galgalia, it is interesting to note that in the 8-16 band, the average power variance indicates a wet period throughout the time period except a few breakdowns during 240-260 and 300-320 months (Figure 4(d)). In Kursela (Figure 5(d)), a wet period is observed until 280 months and after, until the end of the studied duration, a continuous dry period is observed. Similarly in Murliganj (Figure 6(d)) and in Nirmali (Figure 7(d)) a continuous dry period is observed after 240 and 280 months. It is concluded from the wavelet power spectrum analysis for all the selected gauging stations for the selected time period shows that maximum power spectrum is observed in the 8-16 band. A similar observation in the variance also seen towards the southwest of Kosi basin (Kursela, Murliganj, Birpur and Nirmali, see Figure 1 in [30]).

B. Wavelet Coherence Analysis

Figure 8(a-c) shows the WCA between the Galgalia and Murliganj, Galgalia and Birpur, Kursela and Bhimnagar stations. It is clearly visible from the Figure that there is a good correlation between all the selected sites during the studied time period. Even though short breakdowns are visible for shorter time periods, the majority shows a very good coherence between the signals. This study is also proving the findings of [36] that WCA could be a useful tool to detect the relationship between rainfall in different sites of the same basin.

From the above results, it is clear that there is a good correlation between the rainfall of the selected gauging stations for the selected time period which is averaged and shown in Figure 9. It is found that the rainfall months are observed from the month of May to October throughout the studied duration with peaks during 1985-1990, 1995-2000, and 2000-2005. Maximum monthly rainfall reaches 700mm during the selected time duration. After 2005, a considerable decrease in the rainfall is observed. This is also in accordance with the average variance results of the wavelet spectrum analysis. This finding can be strengthened by the breakpoint analysis for the basin shown in Figure 10. It is found that after 237 months, there is a reduction in the mean rainfall from 147.7mm to 116.3mm.

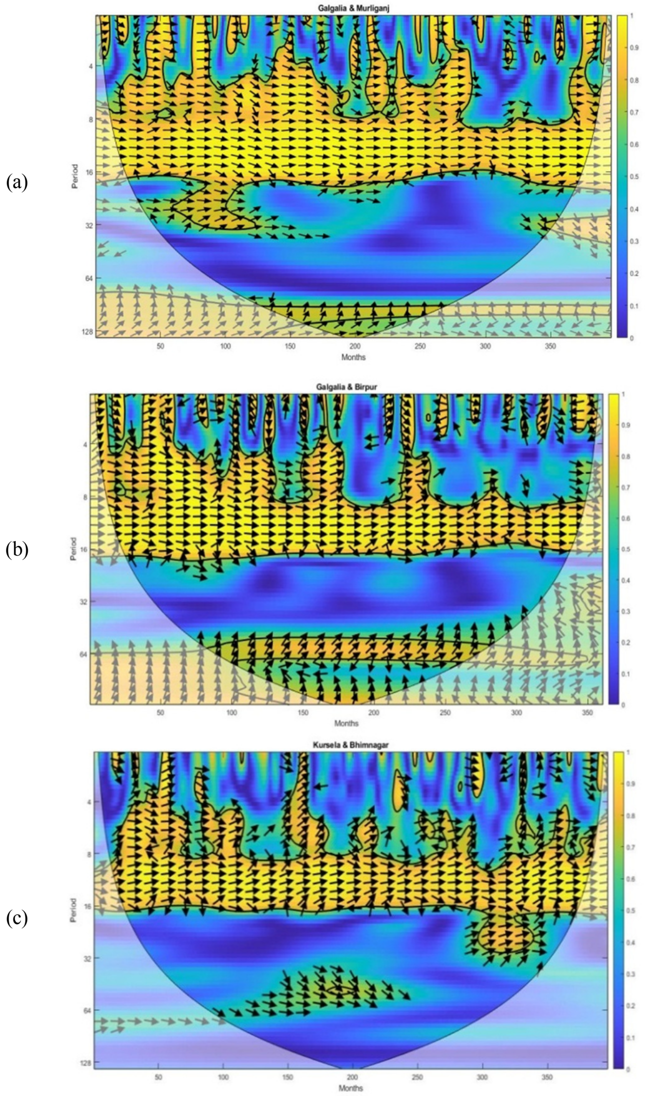


Fig. 8. Wavelet coherence analysis between: (a) Galgalia and Murliganj, (b) Galgalia and Birpur, and (c) Kursela and Bhimnagar.

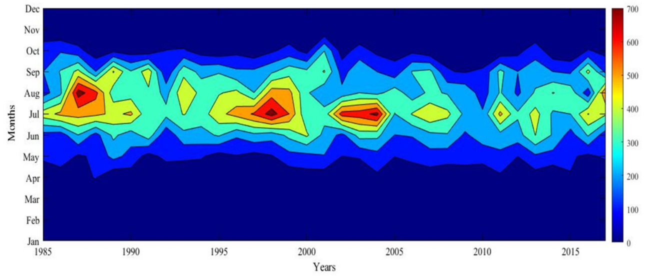


Fig. 9. Monthly average rainfall over the entire Kosi basin during the 1985-2017 period.

V. CONCLUSION

To study the variability of the rainfall in the Kosi basin, wavelet spectrum analysis and wavelet coherence analysis were performed. The wavelet spectrum analysis for the monthly time series data for all the selected 7 rain gauge stations during the 1985-2017 period show large power concentration in the 8-16 month band. This reveals an annual periodicity of events and is also confirmed by the peak in the global wavelet spectrum.

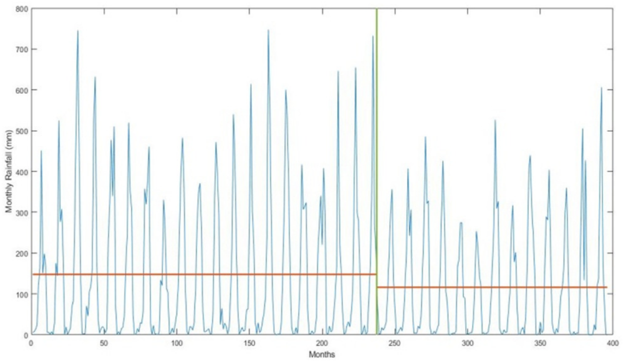


Fig. 10. Breakpoint analysis of the monthly average rainfall data of the Kosi basin during the 1985-2017 period.

Scale averaged time series of the selected 8-16 band reveals that the regions in the southwest of the Kosi basin show similar patterns of dry and wet periods during the studied time duration. Wavelet coherence analysis of the rainfall data reveals a good correlation between the selected stations. The average rainfall data also confirm the considerable reduction in the rainfall towards the beginning of the year 2005. Break point analyses were also in accordance with the results of the average rainfall data and found a breakdown of average rainfall from 147.7mm to 116.3mm for the selected time duration.

REFERENCES

[1] C. P. Khedun, A. K. Mishra, V. P. Singh, and J. R. Giardino, "A copula-based precipitation forecasting model: Investigating the interdecadal modulation of ENSO's impacts on monthly precipitation," *Water Resources Research*, vol. 50, no. 1, pp. 580–600, 2014, https://doi.org/10.1002/2013WR013763.

[2] D. Himayoun and T. Roshni, "Spatio-temporal variation of drought characteristics, water resource availability and the relation of drought with large scale climate indices: A case study of Jhelum basin, India," *Quaternary International*, vol. 525, pp. 140–150, Aug. 2019, https://doi.org/10.1016/j.quaint.2019.07.018.

[3] C. D. Ahrens and R. Henson, *Meteorology Today*, 11th ed. Toronto, Canada: Cengage Learning, 2015.

[4] C. Sireesha, T. Roshni, and M. K. Jha, "Insight into the precipitation behavior of gridded precipitation data in the Sina basin," *Environmental Monitoring and Assessment*, vol. 192, no. 11, Oct. 2020, Art. no. 729, https://doi.org/10.1007/s10661-020-08687-3.

[5] M. H. Elsanabary, T. Y. Gan, and D. Mwale, "Application of wavelet empirical orthogonal function analysis to investigate the nonstationary character of Ethiopian rainfall and its teleconnection to nonstationary global sea surface temperature variations for 1900–1998," *International Journal of Climatology*, vol. 34, no. 6, pp. 1798–1813, 2014, https://doi.org/10.1002/joc.3802.

[6] M. Suman and R. Maity, "Hybrid Wavelet-ARX approach for modeling association between rainfall and meteorological forcings at river basin scale," *Journal of Hydrology*, vol. 577, Oct. 2019, Art. no. 123918, https://doi.org/10.1016/j.jhydrol.2019.123918.

[7] S. Kumar, T. Roshni, E. Kahya, and M. A. Ghorbani, "Climate change projections of rainfall and its impact on the cropland suitability for rice and wheat crops in the Sone river command, Bihar," *Theoretical and Applied Climatology*, vol. 142, no. 1, pp. 433–451, Oct. 2020, https://doi.org/10.1007/s00704-020-03319-9.

[8] S. Gadgil and S. Gadgil, "The Indian Monsoon, GDP and Agriculture," *Economic and Political Weekly*, vol. 41, no. 47, pp. 4887–4895, 2006.

[9] S. R. Samo, N. Bhatti, A. Saand, M. A. Keerio, and D. K. Bangwar, "Temporal Analysis of Temperature and Precipitation Trends in Shaheed Benazir Abad Sindh, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2171–2176, Dec. 2017, https://doi.org/10.48084/etasr.1388.

[10] A. J. McMichael *et al.*, Eds., *Climate change and human health: risks and responses*. Geneva, Switzerland: World Health Organization, 2003.

[11] K. Ahmed, S. Shahid, and S. B. Harun, "Spatial interpolation of climatic variables in a predominantly arid region with complex topography," *Environment Systems and Decisions*, vol. 34, no. 4, pp. 555–563, Dec. 2014, https://doi.org/10.1007/s10669-014-9519-0.

[12] W. Wang, S. Wang, X. Ma, and J. Gong, "Recent advances in catalytic hydrogenation of carbon dioxide," *Chemical Society Reviews*, vol. 40, no. 7, pp. 3703–3727, Jun. 2011, https://doi.org/10.1039/C1CS15008A.

[13] A. H. Syafrina, A. Norzaida, and O. N. Shazwani, "Stochastic Modeling of Rainfall Series in Kelantan Using an Advanced Weather Generator," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2537–2541, Feb. 2018, https://doi.org/10.48084/etasr.1709.

[14] K. L. Ebi, R. S. Kovats, and B. Menne, "An Approach for Assessing Human Health Vulnerability and Public Health Interventions to Adapt to Climate Change," *Environmental Health Perspectives*, vol. 114, no. 12, pp. 1930–1934, Dec. 2006, https://doi.org/10.1289/ehp.8430.

[15] P. Högy, C. Poll, S. Marhan, E. Kandeler, and A. Fangmeier, "Impacts of temperature increase and change in precipitation pattern on crop yield and yield quality of barley," *Food Chemistry*, vol. 136, no. 3, pp. 1470–1477, Feb. 2013, https://doi.org/10.1016/j.foodchem.2012.09.056.

[16] P. K. Srivastava *et al.*, "Long-Term Trend Analysis of Precipitation and Extreme Events over Kosi River Basin in India," *Water*, vol. 13, no. 12, Jan. 2021, Art. no. 1695, https://doi.org/10.3390/w13121695.

[17] T. Roshni, M. K. Jha, R. C. Deo, and A. Vandana, "Development and Evaluation of Hybrid Artificial Neural Network Architectures for Modeling Spatio-Temporal Groundwater Fluctuations in a Complex Aquifer System," *Water Resources Management*, vol. 33, no. 7, pp. 2381–2397, May 2019, https://doi.org/10.1007/s11269-019-02253-4.

[18] B. S. N. Reddy, S. K. Pramada, and T. Roshni, "Monthly surface runoff prediction using artificial intelligence: A study from a tropical climate river basin," *Journal of Earth System Science*, vol. 130, no. 1, Feb. 2021, Art. no. 35 (2021), https://doi.org/10.1007/s12040-020-01508-8.

[19] S. J. Burian, S. J. Nix, R. E. Pitt, and S. R. Durrans, "Urban Wastewater Management in the United States: Past, Present, and Future," *Journal of Urban Technology*, vol. 7, no. 3, pp. 33–62, Dec. 2000, https://doi.org/10.1080/713684134.

[20] K. Roushangar, F. Alizadeh, and J. Adamowski, "Exploring the effects of climatic variables on monthly precipitation variation using a continuous wavelet-based multiscale entropy approach," *Environmental Research*, vol. 165, pp. 176–192, Aug. 2018, https://doi.org/10.1016/j.envres.2018.04.017.

[21] V. Nourani and N. Farboudfam, "Rainfall time series disaggregation in mountainous regions using hybrid wavelet-artificial intelligence methods," *Environmental Research*, vol. 168, pp. 306–318, Jan. 2019, https://doi.org/10.1016/j.envres.2018.10.012.

[22] S. Sithara, S. K. Pramada, and S. G. Thampi, "Sea level prediction using climatic variables: a comparative study of SVM and hybrid wavelet SVM approaches," *Acta Geophysica*, vol. 68, no. 6, pp. 1779–1790, Dec. 2020, https://doi.org/10.1007/s11600-020-00484-3.

[23] C. A. G. Santos, I. C. Guerra-Gomes, B. M. Gois, R. F. Peixoto, T. S. L. Keesen, and R. M. da Silva, "Correlation of dengue incidence and rainfall occurrence using wavelet transform for João Pessoa city," *Science of The Total Environment*, vol. 647, pp. 794–805, Jan. 2019, https://doi.org/10.1016/j.scitotenv.2018.08.019.

[24] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990, https://doi.org/10.1109/18.57199.

[25] C. Torrence and G. P. Compo, "A Practical Guide to Wavelet Analysis.," *Bulletin of the American Meteorological Society*, vol. 79, pp. 61–78, Jan. 1998, https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

[26] J. O. Adepitan and E. O. Falayi, "Variability changes of some climatology parameters of Nigeria using wavelet analysis," *Scientific African*, vol. 2, Mar. 2019, Art. no. e00017, https://doi.org/10.1016/j.sciaf.2018.e00017.

[27] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI," *NeuroImage*, vol. 50, no. 1, pp. 81–98, Mar. 2010, https://doi.org/10.1016/j.neuroimage.2009.12.011.

[28] A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," *Nonlinear Processes in Geophysics*, vol. 11, no. 5/6, pp. 561–566, Nov. 2004, https://doi.org/10.5194/npg-11-561-2004.

[29] K. Mishra and R. Sinha, "Flood risk assessment in the Kosi megafan using multi-criteria decision analysis: A hydro-geomorphic approach," *Geomorphology*, vol. 350, Feb. 2020, Art. no. 106861, https://doi.org/10.1016/j.geomorph.2019.106861.

[30] A. Towheed and T. Roshni, "Linking climate change to soil loss estimation in the Kosi river basin, India," *Journal of Water and Climate Change*, Mar. 2021, Art. no. jwc2021259, https://doi.org/10.2166/wcc.2021.259.

[31] M. Rathinasamy, A. Agarwal, B. Sivakumar, N. Marwan, and J. Kurths, "Wavelet analysis of precipitation extremes over India and teleconnections to climate indices," *Stochastic Environmental Research and Risk Assessment*, vol. 33, no. 11, pp. 2053–2069, Dec. 2019, https://doi.org/10.1007/s00477-019-01738-3.

[32] A. Agarwal, "Unraveling spatio-temporal climatic patterns via multi-scale complex networks," Ph.D. dissertation, University of Potsdam, Potsdam, Germany, 2019.

[33] N. Saada, M. R. Abdullah, A. Hamaideh, and A. Abu-Romman, "Application of Stochastic Analysis, Modeling and Simulation (SAMS) to Selected Hydrologic Data in the Middle East," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4261–4264, Jun. 2019, https://doi.org/10.48084/etasr.2750.

[34] J. C. N. dos Santos, E. M. de Andrade, M. J. S. Guerreiro, P. H. A. Medeiros, H. A. de Queiroz Palácio, and J. R. de Araújo Neto, "Effect of dry spells and soil cracking on runoff generation in a semiarid micro watershed under land use change," *Journal of Hydrology*, vol. 541, pp. 1057–1066, Oct. 2016, https://doi.org/10.1016/j.jhydrol.2016.08.016.

[35] V. A. H. F. D. Santos *et al.*, "Causes of reduced leaf-level photosynthesis during strong El Niño drought in a Central Amazon forest," *Global Change Biology*, vol. 24, no. 9, pp. 4266–4279, Sep. 2018, https://doi.org/10.1111/gcb.14293.

[36] V. Nourani and S. Mousavi, "Spatiotemporal groundwater level modeling using hybrid artificial intelligence-meshless method," *Journal of Hydrology*, vol. 536, pp. 10–25, May 2016, https://doi.org/10.1016/j.jhydrol.2016.02.030.

A New Particle Swarm Optimization Based Strategy for the Economic Emission Dispatch Problem Including Wind Energy Sources

Gharbi A. Alshammari
Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
gharbi.a.alshammari@gmail.com

Fahd A. Alshammari
Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
fahad.a.s.alshammari@gmail.com

Tawfik Guesmi
Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
tawfik.guesmi@istmt.rnu.tn

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

Naif A. Alshammari
Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
naif.a.s.alshammari@gmail.com

Abstract-Power dispatch has become an important issue due to the high integration of Wind Power (WP) in power grids. Within this context, this paper presents a new Particle Swarm Optimization (PSO) based strategy for solving the stochastic Economic Emission Dispatch Problem (EEDP). This problem was solved considering several constraints such as power balance, generation limits, and Valve Point Loading Effects (VPLEs). The power balance constraint is described by a chance constraint to consider the impact of WP intermittency on the EEDP solution. In this study, the chance constraint represents the tolerance that the power balance constraint cannot meet. The suggested framework was successfully evaluated on a ten-unit system. The problem was solved for various threshold tolerances to study further the impact of WP penetration.

Keywords-economic emission dispatch; wind energy; stochastic optimization; particle swarm optimization

I. INTRODUCTION

Wind energy has expanded rapidly the recent years at a global level. Wind power is becoming more and more economically competitive compared to conventional energy production methods due to improvements in turbine efficiency and rising fuel prices [1]. In addition, wind energy sources are growing at a rapid pace reaching a technical maturity that allows them to become important components of the energy industry. On the other hand, the inclusion of wind energy in power grids introduced new challenges. The high penetration of wind energy has a significant impact on system security due to its intermittent characteristics [2]. One of these challenges is the power dispatch problem. In general, the dispatch problem aims to find the optimal generation of all generators and sources minimizing energy production cost and system losses. In addition, global warming and increased initiatives to protect the environment are forcing producers to reduce the gas emissions produced by fossil fuel combustion in power stations. The fuels used in thermal power stations (coal, fuel oil, natural gas, etc.) produce harmful gases like carbon dioxide (CO_2), sulfur dioxide (SO_2), and nitrogen oxides (NO_x) which are toxic and cause the greenhouse effect. Thus, the reduction of the emission of these gases during electricity production has become a primordial task [3].

Several studies combined the economic and environmental aspects in one problem called Economic Emission Dispatch Problem (EEDP) [4-5], considering several constraints such as generation capacity, power balance, and Valve Point Loading Effects (VPLE) [4-5]. Various methods have been suggested in the past two decades to solve this nonlinear and nonconvex problem. For instance, classical techniques such as dynamic programming [6], linear programming [7], lambda iteration [8], and interior-point [9] have been widely used for solving the dispatch problem. However, in these techniques, the fuel cost was approximated by a quadratic, and VPLE constraints were neglected. In addition, these conventional methods were iterative and required an initial solution which may affect the convergence of the employed method and produce only local solutions. Various intelligent optimization methods were presented to overcome the limitations of classical methods, like the Genetic Algorithm (GA) [10], Artificial Bee Colony (ABC) [11], Bacterial Foraging Algorithm (BFA) [12], Particle Swarm Optimization (PSO) [13], Differential Evolution (DE) [14], and Simulated Annealing (SA) [15]. In general, these meta-heuristic techniques have achieve good results in solving

---

Corresponding author: Tawfik Guesmi

various engineering problems. However, the aforementioned techniques minimized fuel cost and emissions by seeking the optimal production of the existing thermal units. At the moment, wind energy has attracted much attention in the power sector due to its zero fuel cost and emissions. Hence, the inclusion of wind power in the EEDP formulation has gained wide attention.

In [16], a new mathematical formulation was developed based on the here-and-now approach for the stochastic EEDP integrating WP sources. The intermittency of wind power was described by the Weibull distribution function. The same approach was extended for the dynamic EEDP in [17]. Various fuzzy membership functions were suggested in [18], taking into account that system security may be affected by the randomness of wind power, to describe the dispatcher's attitude regarding WP penetration. Two objective functions, based on operational cost and risk level, were considered and minimized using a PSO-based method, but emissions were not included in the problem formulation. The risk level of WP uncertainty was considered in [19], incorporating VPLE in the cost function. Fuzzy quadratic functions that described dispatcher's attitudes were investigated in [20] to determine the quantity of additional WP to minimize generation cost without affecting system security. The effect of fluctuations of WP on the EEDP was modeled in [21] by over- and under-estimation costs of available WP, where a hybrid algorithm combining PSO and gravitational search was used to minimize the objective functions. In [22], the under- and over-estimation costs of uncertain WP were also included in the total production cost, using an improved fireworks algorithm to find the optimal generation. The randomness of WP was modeled by a chance constraint in the dispatch problem formulation to avoid the over- and under-estimation costs in [23], where WP was represented by a Weibull distribution function, and the impact of WP penetration on the total fuel cost and emissions was studied and analyzed.

In recent years, PSO-based techniques have been favored by researchers due to their low parameter number, convergence rate, and easy implementation. PSO was introduced in [24] as an efficient optimization tool for complex optimization problems. This study presents a new PSO-based strategy for solving the stochastic EEDP incorporating a wind farm. At first, the problem is formulated as a stochastic optimization problem. Then, the stochastic constraint, which describes power balance, was converted to a deterministic constraint. The Weibull distribution function was used to describe the randomness of WP. The PSO algorithm was used to solve the obtained deterministic problem. The effectiveness of the proposed method was tested on a 10-unit system, investigating the cases with and without WP sources. Moreover, the impact of WP penetration rate was studied.

II. PROBLEM FORMULATION

The EEDP is treated as a multi-objective mathematical programming problem that attempts to minimize both cost and emissions simultaneously while satisfying equality and inequality constraints. The following objectives and constraints were taken into account in the EEDP problem formulation:

A. Objective Functions

The thermal units with multi-steam admission valves that work sequentially to cover the ever-increasing generation increase the nonlinearity order of the total fuel cost due to the VPLE, as illustrated in Figure 1.

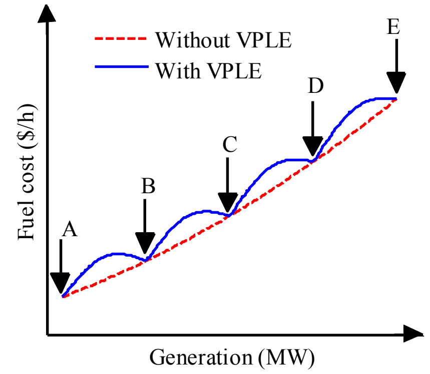


Fig. 1. Fuel cost function with five valves (A, B, C, D, E).

The fuel cost function of a thermal generator, considering the VPLE, is expressed as the sum of a quadratic and a sinusoidal function. Thus, the total fuel cost in terms of real power output can be expressed as [23]:

$$C_T = \sum_{i=1}^{N} a_i + b_i P_i + c_i P_i^2 + \left| d_i \, sin\left\{ e_i \left(P_i^{min} - P_i \right) \right\} \right| \quad (1)$$

where a_i, b_i, c_i, d_i, and e_i are the cost coefficients of the i-th unit, P_i is the output power in MW, and the total cost C_T is in \$/h. The second objective function considered is the atmospheric pollutants such as sulfur (SO_x) and nitrogen oxides (NO_x) caused by fossil-fueled generator units. This can be modeled as the summation of a quadratic polynomial and an exponential function [23]:

$$E_T = \sum_{i=1}^{N} \alpha_i + \beta_i P_i + \gamma_i (P_i)^2 + \eta_i \, exp(\xi_i P_i) \quad (2)$$

where a_i, β_i, γ_i, η_i, and ξ_i are the emission coefficients, and the total emission is in ton/h. In several works, the bi-objective EEDPs were converted into a mono-objective optimization problem [3], and the Price Penalty Factor (PPF) based method was adopted. Thus, the combined economic-emission objective function can be described by:

$$F_T = \mu C_T + (1 - \mu) \lambda E_T \quad (3)$$

where, $\mu = \text{rand}(0,1)$, F_T will be minimized for each generated value of μ to obtain the optimal solution that can be a nominee solution in the Pareto front, and λ is the average of the PPF thermal units. As shown in (4), the PPF of the i-th unit is the rate between its fuel cost and its emission for maximum generation capacity, and (5) gives the expression of λ.

$$\lambda_i = \frac{C_{i max}}{E_{i max}} \quad (4)$$

$$\lambda = \frac{1}{N} \sum_{i=1}^{N} \lambda i \quad (5)$$

B. Problem Constraints

The EEDP can be solved by minimizing the F_T defined in (3) for the following constraints [23]:

- Generation Capacity: Because of the unit design, the real power output of each unit i should be within its minimum P_i^{min} and maximum limit P_i^{max}:

$$P_i^{min} \leq P_i \leq P_i^{max} \quad i = 1, \dots, N \quad (6)$$

- Real power balance constraints: The total of real power generation must balance the predicted power demand P_d plus the real power losses P_L in the transmission lines, at each time interval over the scheduling horizon:

$$\sum_{i=1}^{N} P_i^t - P_D^t - P_L^t = 0 \quad t = 1, \dots, T \quad (7)$$

P_L can be calculated using a constant loss formula [4]:

$$P_L^t = \sum_{i=1}^{N} \sum_{j=1}^{N} P_i B_{ij} P_j + \sum_{i=1}^{N} B_{oi} P_i + B_{oo} \quad (8)$$

where B_{ij}, B_{oi}, and B_{oo} are the loss parameters also called *B*-coefficients.

- Prohibited Operating Zones (POZ) constraints: The POZ constraints are described as:

$$P_i^t \in \begin{cases} P_i^{min} \leq P_i \leq P_{i,1}^{down} \\ P_{i,k-1}^{up} \leq P_i \leq P_{i,k}^{down} & , k = 2, \dots, z_i \\ P_{i,z_i}^{up} \leq P_i \leq P_i^{max} \end{cases} \quad (9)$$

where $P_{i,k}^{down}$ and $P_{i,k}^{up}$ are the down and up bounds of POZ number k, and z_i is the number of POZ for the *i*-th unit due to the vibrations in the shaft or other mechanical faults. Therefore, the machine has discontinuous input-output characteristics [4].

C. Description of WP Randomness

A major challenge in integrating wind power output into a power network is its uncertainty, fluctuation, and intermittent nature. Hence, WP output should be expressed as a stochastic variable utilizing a transformation from wind speed to power output. A simplified linear piecewise function can describe the actual relationship between them when ignoring some minor nonlinear factors. This study adopts the two-factor Weibull distribution [16]. The main advantage of this distribution type is that if its parameters are specified at a given altitude, they can be found for another one. The Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) of wind speed are described by (10) and (11), respectively:

$$f_V(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} exp\left[-\left(\frac{v}{c}\right)^k\right] \quad (10)$$

$$F_V(v) = \int_0^v f_V(\tau)\, d\tau = 1 - exp\left(-\left(\frac{v}{c}\right)^k\right), \quad v \geq 0 \quad (11)$$

where, k and c are positive parameters called shape and scale factors for a given location, respectively. The speed-power characteristic of the wind turbine can be described by:

$$W = \phi(V) = 0, \text{ if } V < v_{in} \text{ or } V > v_{out} \quad (12)$$

$$W = \phi(V) = \frac{(V - v_{in}) w_r}{v_r - v_{in}} \text{ if } v_{in} \leq V < v_r \quad (13)$$

$$W = \phi(V) = w_r, \text{ if } v_r \leq V < v_{out} \quad (14)$$

Based on probability theories, the CDF corresponding to the WP can be described by:

$$\begin{aligned} F_W(w) &= Pr(W \leq w) = \\ 1 - exp&\left\{-\left(\frac{\left(1 + \frac{hw}{w_r}\right) v_{in}}{c}\right)^k\right\} \\ &+ exp\left(-\left(\frac{v_{out}}{c}\right)^k\right), \\ &0 \leq w < w_r \end{aligned} \quad (15)$$

where, $h = \frac{v_r - v_{in}}{v_{in}}$. Taking into account the intermittency characteristic of WP, the power balance constraint given by (7) can be modified as:

$$Pr\{\sum_{i=1}^{N} P_i + W \leq P_D + P_L\} \leq \sigma\} \quad (16)$$

where $P_r(x)$ is the probability of event x, W is the WP output of the wind farm, and σ is the tolerance that power balance between total generation, load, and total system losses cannot meet.

III. THE PSO ALGORITHM

PSO is considered an efficient and robust method that can be applied to nonlinear optimization problems and more particularly on electrical systems [25-26]. This algorithm ignores several conditions, such as differentiability and continuity regardless of the objective functions and the constraints to be optimized or respected. For an optimization problem with n decision variables, the *i*-th particle at iteration k is presented by its position $X_i^k = (X_{i1}^k, \dots, X_{in}^k)$ that is considered as a candidate solution and velocity $V_i^k = (V_{i1}^k, \dots, V_{in}^k)$. At the next generation $k+1$, the velocity and the position of this particle will be updated according to:

$$\begin{aligned} V_i^{k+1} = wV_i^k + c_1 r_1 (pbest_i^k - X_i^k) \\ + c_2 r_2 (gbest^k - X_i^k) \end{aligned} \quad (17)$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \quad (18)$$

where, w, c_1, and c_2 are the PSO parameters, r_1 and r_2 are random numbers in the range [0,1], and $pbest_i^k$ and $gbest^k$ are the best solution of the *i*-th particle and the overall population at the *k*-th iteration respectively. At each iteration k, the inertia weight w used for balancing between local and global searches can be calculated as:

$$w = w^{max} - \frac{w^{max} - w^{min}}{k^{max}} * k \quad (19)$$

where k^{max} is the maximum number of iterations, and w^{max} and w^{min} are the upper and lower bounds of w. From (19), it is clear that w^{max} is the initial value of the inertia weight while w^{min} is its final value.

IV. SIMULATION AND RESULTS

Two cases were studied to verify the effectiveness of the suggested strategy for solving the EEDP including a wind farm. Simulations were carried out on MATLAB R2009a installed on a PC with an i7-4510U@2.60GHz CPU. The studied cases were: A ten-unit system without a wind farm

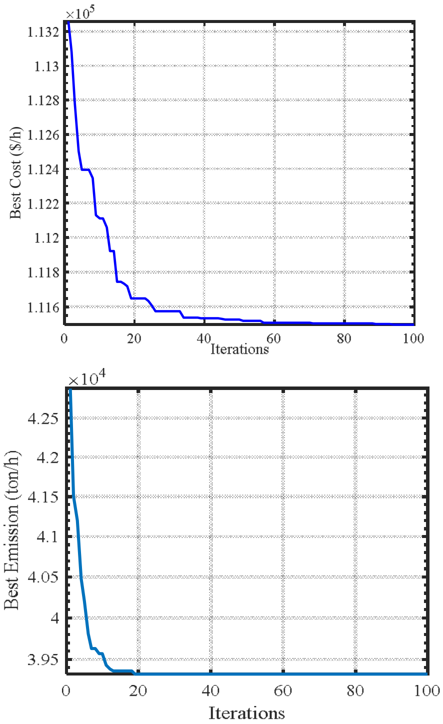
(Case 1) and a ten-unit system with a wind farm (Case 2). All data of both systems were taken from [3, 23]. The wind parameters are shown in Table I.

TABLE I. WIND PARAMETERS

K	C	v_{in}	v_{out}	v_r
1.7	15	5	45	15

A. Case 1

Since the EEDP is a multi-objective optimization problem, a set of non-dominated solutions is required. Table II shows a list of non-dominated solutions obtained for various values of μ ranging from 0 to 1. From Table II, it can be noted that as μ increases, the total production cost decreases and the total emissions increase. The convergence characteristics of the proposed PSO-based technique for the economic ($\mu=1$) and the emission ($\mu=0$) dispatch problems are shown in Figure 2. The Pareto-front resulted from the PSO-based strategy is depicted in Figure 3. The best economic dispatch solution correspond to 111498.49$/h fuel cost and 4567.27ton/h total emissions, while the best emission dispatch solution corresponds to 3932.24ton/h total emissions and 116412.49 $/h total fuel cost. To further test the effectiveness of the proposed method, the simulation results obtained using the proposed PSO-based method were compared with various algorithms. From Table III, it is clear that the proposed PSO method outperforms the others in solving power dispatch problems.


(a) for μ=1

(b) for μ=0

Fig. 2. Convergence characteristics of the proposed method (case 1).

TABLE II. PARETO SOLUTIONS FOR VARIOUS VALUES OF μ (CASE 1).

λ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
P_1	55.0000	55.0000	55.0000	55.0000	55.0000	55.0000	55.0000	55.0000	55.0000	55.0000	54.9736
P_2	80.0000	80.0000	79.9999	80.0000	80.0000	80.0000	80.0000	80.0000	80.0000	79.9980	80.0000
P_3	81.1292	81.1079	81.0693	81.8332	83.0377	84.7423	86.7683	88.4989	90.57678	98.1322	106.2337
P_4	81.3701	81.1322	80.8085	81.2860	82.1286	83.4244	84.9239	85.9500	87.2575	93.1849	100.3274
P_5	160.0000	160.0000	160.0000	160.0000	160.0000	143.7728	126.1284	109.9550	96.7236	88.4957	82.5885
P_6	240.0000	240.0000	240.0000	219.5599	189.0966	164.2697	142.7271	121.8599	103.5178	92.2169	82.98739
P_7	294.4776	292.2409	289.7346	291.3277	294.5846	299.5123	300.0000	300.0000	300.0000	299.9786	299.9923
P_8	297.2982	296.9563	296.5578	300.8168	307.3015	315.4370	321.2987	327.2378	333.8038	340.0000	340.0000
P_9	396.7566	398.0034	399.4279	406.0273	415.3302	427.8233	442.3925	456.2269	469.9842	470.0000	469.9574
P_{10}	395.5627	397.2015	399.1011	406.2881	416.3488	429.8128	445.6171	461.2196	469.9878	469.9907	469.9736
C_T	116412.49	116399.01	116384.25	115599.76	114608.47	113504.92	112644.77	112023.28	111650.66	111530.31	111498.49
E_T	3932.2432	3932.3162	3932.5799	3961.3722	4014.4321	4105.6762	4210.6645	4325.7406	4434.2593	4501.6670	4567.2691
P_L	81.5947	81.6424	81.6993	82.1394	82.8283	83.7950	84.8563	85.9483	86.8517	86.9972	87.0343

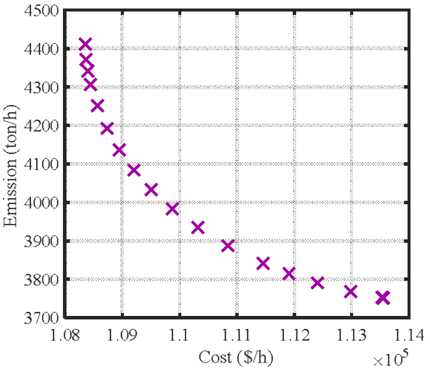


Fig. 3. Pareto-front (case 1).

TABLE III. SIMULATION RESULTS OBTAINED FOR CASE 1.

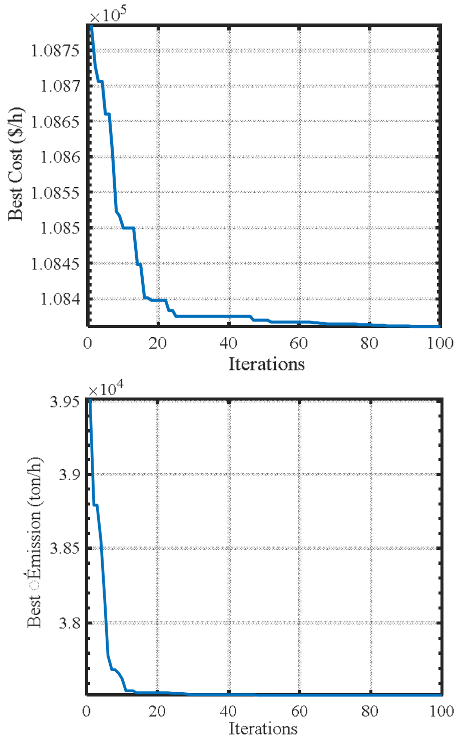
	Best cost		Best emission	
	Cost ($/h)	**Emission (ton/h)**	**Cost ($/h)**	**Emission (ton/h)**
PSO	111498.49	4567.27	116412.49	3932.24
DE	111565.71	4572.68	116418.34	3946.24
FA	111500.79	4581.00	116443.05	3932.62

B. Case 2

In this case, a wind farm with a rated power of $w_r=1.0$pu on a 100MVA base was incorporated in the ten-unit system. The problem was solved for various values of the tolerance σ to investigate the impact of the penetration level of WP on the EEDP solutions. Figure 4 shows the convergence characteristics of production cost ($\mu=1$) and emissions ($\mu=0$) for $\sigma=0.3$.

TABLE IV. PARETO SOLUTIONS FOR VARIOUS VALUES OF μ (CASE 2 – σ = 0.3).

λ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
P_1	55.0000	55.0000	55.0000	54.9927	54.9975	55.0000	55.0000	55.0000	55.0000	55.0000	55.0000
P_2	79.3793	79.0572	79.2443	80.0000	79.9935	80.0000	80.0000	80.0000	80.0000	80.0000	79.9671
P_3	79.2368	79.1410	79.1944	95.6866	89.4404	87.6390	86.3276	84.6486	83.0527	81.3265	80.1881
P_4	79.4393	78.8639	79.1835	87.6092	85.3617	84.0826	83.7210	82.7481	81.6908	80.3916	79.6187
P_5	160.0000	160.0000	160.0000	71.2762	80.4537	91.7389	105.7879	120.9868	138.6410	158.8935	160.0000
P_6	240.0000	240.0000	240.0000	70.2414	82.0695	97.1504	116.1033	135.7649	157.5176	180.8750	210.0817
P_7	283.2762	278.7705	281.1587	299.6879	296.3502	294.8372	296.4263	2.9430731	290.9290	285.5700	282.3278
P_8	285.6298	285.0466	285.3679	337.1601	327.7143	319.2092	316.2059	314.8193	306.9338	297.6583	291.2683
P_9	384.6910	387.3684	385.9387	470.0000	469.9934	459.1521	444.9970	431.0097	418.3983	405.4202	396.0864
P_{10}	383.5466	387.0554	385.1589	469.9973	469.9908	466.8807	450.0184	434.2659	420.3902	406.4300	396.3520
C_T	113553.68	113527.38	113541.18	108361.08	108398.65	108566.61	108947.91	109504.42	110317.96	111457.33	112401.88
E_T	3752.5080	3752.8219	3752.5756	4411.1741	4341.4299	4251.0979	4136.8332	4033.7302	3934.7089	3841.9554	3791.0480
P_L	77.4441	77.5478	77.4915	83.8963	83.6103	82.9351	81.8324	80.7956	79.7983	78.8100	78.1351



(a) μ=1

(b) μ=0

Fig. 4. Convergence characteristics for case 2 (σ=0.3).

The Pareto solutions for various values of the weight factor, ranging from 0 to 1, are presented in Table IV. Meanwhile, the Pareto-front for this case is shown in Figure 5. Figure 6 illustrates the impact of the variation of the tolerance on the minimum fuel cost and the total emission functions. From this Figure, it is obvious that the more the tolerance that power balance constraint cannot meet is, the less the cost and emissions are because the more the tolerance is, the more the WP penetration is.

V. CONCLUSION

This study presented a PSO-based strategy for solving the multi-objective EEDP incorporating wind energy sources. The power balance constraint was converted into a chance constraint and the intermittency of WP was described by the Weibull distribution to consider the stochastic characteristic of WP. This chance constraint represents the probability that the power balance constraint cannot meet.

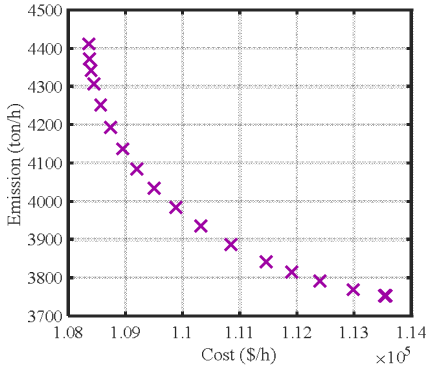


Fig. 5. Pareto-front for case 2 (σ=0.3).

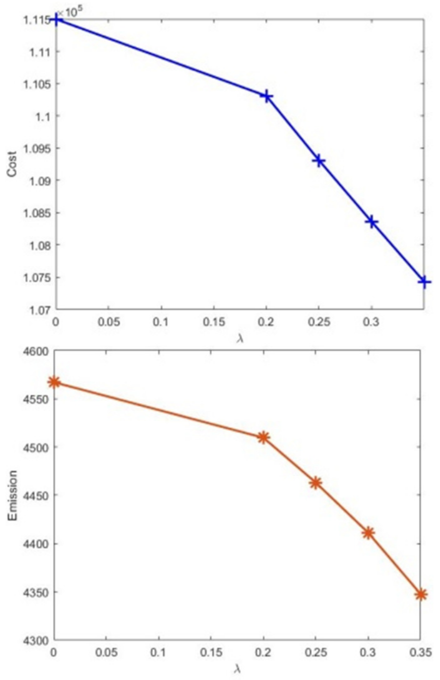


Fig. 6. Impact of the tolerance on the EEDP solutions (case 2).

The EEDP was solved using a PSO-based method depending on several operating constraints such as generators, limits, valve point loading effects, and real power losses. Simulation results, performed on a 69-bus ten-unit system, showed that the level of available wind power (WP) was highly dependent on the threshold tolerance. The results also showed the effectiveness of the proposed optimization method for solving the non-convex EEDP.

REFERENCES

[1] P. D. Chung, "Evaluation of Reactive Power Support Capability of Wind Turbines," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5211–5216, Feb. 2020, https://doi.org/10.48084/etasr.3260.

[2] J. Jin, D. Zhou, P. Zhou, and Z. Miao, "Environmental/economic power dispatch with wind power," *Renewable Energy*, vol. 71, pp. 234–242, Nov. 2014, https://doi.org/10.1016/j.renene.2014.05.045.

[3] R. Dong and S. Wang, "New Optimization Algorithm Inspired by Kernel Tricks for the Economic Emission Dispatch Problem With Valve Point," *IEEE Access*, vol. 8, pp. 16584–16594, 2020, https://doi.org/10.1109/ACCESS.2020.2965725.

[4] M. Basu, "Dynamic economic emission dispatch using nondominated sorting genetic algorithm-II," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 2, pp. 140–149, Feb. 2008, https://doi.org/10.1016/j.ijepes.2007.06.009.

[5] D. Zou, S. Li, Z. Li, and X. Kong, "A new global particle swarm optimization for the economic emission dispatch with or without transmission losses," *Energy Conversion and Management*, vol. 139, pp. 45–70, May 2017, https://doi.org/10.1016/j.enconman.2017.02.035.

[6] Z. X. Liang and J. D. Glover, "A zoom feature for a dynamic programming solution to economic dispatch including transmission losses," *IEEE Transactions on Power Systems*, vol. 7, no. 2, pp. 544–550, May 1992, https://doi.org/10.1109/59.141757.

[7] G. W. Chang *et al.*, "Experiences with mixed integer linear programming based approaches on short-term hydro scheduling," *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 743–749, Nov. 2001, https://doi.org/10.1109/59.962421.

[8] J. B. Park, K. S. Lee, J. R. Shin, and K. Y. Lee, "A particle swarm optimization for economic dispatch with nonsmooth cost functions," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 34–42, Feb. 2005, https://doi.org/10.1109/TPWRS.2004.831275.

[9] G. L. Torres and V. H. Quintana, "On a nonlinear multiple-centrality-corrections interior-point method for optimal power flow," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 222–228, May 2001, https://doi.org/10.1109/59.918290.

[10] P. Jain, K. K. Swarnkar, and S. Wadhwani, "Prohibited Operating Zones Constraint with Economic Load Dispatch Using Genetic Algorithm," *International Journal of Engineering and Innovative Technology*, vol. 1, no. 3, pp. 179–183, Jan. 2012.

[11] 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.

[12] N. Pandit, A. Tripathi, S. Tapaswi, and M. Pandit, "An improved bacterial foraging algorithm for combined static/dynamic environmental economic dispatch," *Applied Soft Computing*, vol. 12, no. 11, pp. 3500–3513, Nov. 2012, https://doi.org/10.1016/j.asoc.2012.06.011.

[13] K. Tlijani, T. Guesmi, and H. H. Abdallah, "Extended Dynamic Economic Environmental Dispatch using Multi-Objective Particle Swarm Optimization," *International Journal on Electrical Engineering and Informatics*, vol. 8, no. 1, pp. 117–131, Mar. 2016, https://doi.org/10.15676/ijeei.2016.8.1.9.

[14] A. Srinivasa Reddy and K. Vaisakh, "Shuffled differential evolution for economic dispatch with valve point loading effects," *International Journal of Electrical Power & Energy Systems*, vol. 46, pp. 342–352, Mar. 2013, https://doi.org/10.1016/j.ijepes.2012.10.012.

[15] I. Ziane, F. Benhamida, A. Graa, "Simulated annealing algorithm for combined economic and emission power dispatch using max/max price penalty factor," *Neural Comput Appl*, vol. 28, no. 1, pp. 197–205, 2017.

[16] X. Liu and W. Xu, "Economic Load Dispatch Constrained by Wind Power Availability: A Here-and-Now Approach," *IEEE Transactions on Sustainable Energy*, vol. 1, no. 1, pp. 2–9, Apr. 2010, https://doi.org/10.1109/TSTE.2010.2044817.

[17] M. H. Alham, M. Elshahed, D. K. Ibrahim, and E. E. D. Abo El Zahab, "A dynamic economic emission dispatch considering wind power uncertainty incorporating energy storage system and demand side management," *Renewable Energy*, vol. 96, pp. 800–811, Oct. 2016, https://doi.org/10.1016/j.renene.2016.05.012.

[18] L. Wang and C. Singh, "Balancing risk and cost in fuzzy economic dispatch including wind power penetration based on particle swarm optimization," *Electric Power Systems Research*, vol. 78, no. 8, pp. 1361–1368, Aug. 2008, https://doi.org/10.1016/j.epsr.2007.12.005.

[19] F. Yao, Z. Y. Dong, K. Meng, Z. Xu, H. H.-C. Iu, and K. P. Wong, "Quantum-Inspired Particle Swarm Optimization for Power System Operations Considering Wind Power Uncertainty and Carbon Tax in Australia," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 4, pp. 880–888, Nov. 2012, https://doi.org/10.1109/TII.2012.2210431.

[20] M. H. Aliabadi, S. Javadi, E. Tafehi, and A. Bakhtiari, "Risk Analysis and Economic Load Dispatch Evaluation of Network with High Wind Power Penetration," *International Journal of Smart Electrical Engineering*, vol. 3, no. 02, pp. 79–85, Apr. 2014.

[21] S. Jiang, C. Zhang, W. Wu, and S. Chen, "Combined Economic and Emission Dispatch Problem of Wind-Thermal Power System Using Gravitational Particle Swarm Optimization Algorithm," *Mathematical Problems in Engineering*, vol. 2019, Nov. 2019, Art. no. e5679361, https://doi.org/10.1155/2019/5679361.

[22] V. K. Jadoun, V. C. Pandey, N. Gupta, K. R. Niazi, and A. Swarnkar, "Integration of renewable energy sources in dynamic economic load dispatch problem using an improved fireworks algorithm," *IET Renewable Power Generation*, vol. 12, no. 9, pp. 1004–1011, 2018, https://doi.org/10.1049/iet-rpg.2017.0744.

[23] T. Guesmi, A. Farah, I. Marouani, B. Alshammari, and H. H. Abdallah, "Chaotic sine–cosine algorithm for chance-constrained economic emission dispatch problem including wind energy," *IET Renewable Power Generation*, vol. 14, no. 10, pp. 1808–1821, 2020, https://doi.org/10.1049/iet-rpg.2019.1081.

[24] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Nov. 1995, vol. 4, pp. 1942–1948 vol.4, https://doi.org/10.1109/ICNN.1995.488968.

[25] Z. Bo and C. Yi-jia, "Multiple objective particle swarm optimization technique for economic load dispatch," *Journal of Zhejiang University-SCIENCE A*, vol. 6, no. 5, pp. 420–427, May 2005, https://doi.org/10.1631/jzus.2005.A0420.

[26] A. Mahor, V. Prasad, and S. Rangnekar, "Economic dispatch using particle swarm optimization: A review," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8, pp. 2134–2141, Oct. 2009, https://doi.org/10.1016/j.rser.2009.03.007.

Evaluation of Passive Islanding Detection Methods for Line to Ground Unsymmetrical Fault in Three Phase Microgrid Systems

Microgrid Islanding Detection Method

L. Bangar Raju
Electrical and Electronics Engineering Department
Koneru Lakshmaiah Educational Foundation
Guntur, India
lsmlbr@yahoo.in

K. Subba Rao
Electrical and Electronics Engineering Department
Koneru Lakshmaiah Educational Foundation
Guntur, India
principal.coe@kluniversity.in

Abstract-Distributed Generators (DGs) are incorporated in the power distribution systems to develop green energies in microgrids. Islanding is a challenging task in a microgrid. Different types of islanding methods, e.g. local and remote methods, have been developed for handling this task, with local methods being easier to implement, while remote methods are communication-based and costly. The local methods are classified as passive, active, and hybrid, out of which the passive methods are more simple and economical. In this paper, a passive islanding detection method is proposed to detect single line to ground fault. This fault is considered to represent the 60 to 70% of the total un-intentional faults of this category. The available passive methods cannot detect islanding at lower power mismatches as the variations in voltage and frequency fall within thresholding values. In this method, the voltage signals are first retrieved at the targeted DG output and then the phase angle is estimated. Finally, the phase angle is differentiated to get Rate Of Change Of Voltage Phase Angle (ROCOVPA) to detect islanding, and then it is compared with the Rate Of Change Of Frequency (ROCOF) at zero percent power mismatch. Simulation results depict that the ROCOVPA is more effective than ROCOF. The proposed method not only reduces detection time and Non-Detection Zone (NDZ) but is also stable during non-islanding cases like load connection and disconnection to avoid nuisance tripping.

Keywords-distributed generation; non detection zone; point of common coupling; ROCOVPA; ROCOF; line to ground fault

I. INTRODUCTION

During the normal mode, a microgrid is in synchronism with the main grid feeding the connected loads and importing the mismatch power of the load from the grid. Based on load demand, the mismatched power is supplied by the grid. But during unintentional unsymmetrical faults on the system, the microgrid has to detect the fault condition and isolate itself from the main grid within 2s as per IEEE-1547-2018 regulations. There are symmetrical and unsymmetrical faults, but in this paper only unsymmetrical single line to ground fault was tested as the majority of faults are of this nature. The different islanding detection methods have their own advantages and disadvantages. The active methods affect the power quality due to perturbations during testing. The passive methods fail to detect islanding at lower power mismatches due to the smaller variation in voltage and frequency above threshold. The hybrid methods are good to detect islanding but power quality problems arise due to injections of perturbations. Communication methods are good with a smaller Non Detection Zone (NDZ) but are costly. The implementation of such methods depends on the size of the microgrid and the criticality of loads.

To mitigate these problems, in this paper the Rate Of Change Of Voltage Phase Angle (ROCOVPA) method is proposed which detects islanding at zero percent mismatch power and also avoids nuisance tripping during non-islanding operations like load connection and disconnection. A simple methodology is used. At first, the voltage phase angle is calculated at the output of DG and then differentiated to get the ROCOVPA to detect islanding. In normal conditions, when there is no fault, the variations are within the threshold values but during fault conditions the variations exceed the threshold values to island the microgrid safely from the main grid within 2s. The DG feeds the connected load in droop control in islanded mode, without any power interruption to load.

The proposed ROCOVPA methodology can be extended to detect symmetrical and unsymmetrical faults. In this paper, the ROCOVPA is tested and compared with the Rate Of Change Of Frequency (ROCOF) method in Matlab/Simulink, in which ROCOVPA is proved to perform better. As per the result analysis, the proposed ROCOVPA method detects islanding at even zero percent mismatch power (0% NDZ). The testing is extended for non-islanding cases during load connection and disconnection for stability. The ROCOVPA method is proved to be stable during non-islanding cases and also avoids nuisance tripping.

The comparison of the proposed ROCOVPA method with ROCOF proves that it is better in discriminating islanding and non-islanding operations with less NDZ and detection time.

Corresponding author: L. Bangar Raju

II. NETWORK AND MATHEMATICAL MODEL

The network model is shown in Figure.1. The ROCOVPA islanding detection method is tested on a DG with 2.5KW with an interfaced inverter. A parallel RLC load is connected to the DG with a quality factor of 1.8 at the Point of Common Coupling (PCC). The DG inverter is connected to the main grid to the PCC through a 3-phase circuit breaker. The inverter is connected to the PCC with filter values: Cf=2μF ,Lf=2mH, and Rf=10Ω.

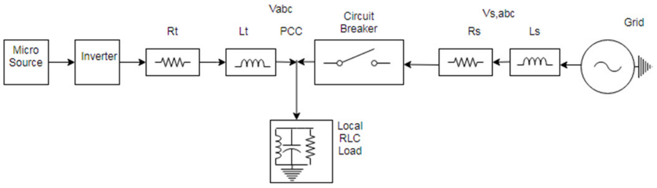


Fig. 1. Network model of the islanding detection ROCOVPA method.

The mathematical model of the islanded microgrid in the *abc* frame is given by the following equations:

$$V_{t,abc} = L_t d \frac{i_{t,abc}}{dt} + R_t i_{t,abc} + V_{abc} \quad (1)$$

$$i_{t,abc} = \; = Vabc/\mathrm{R} + i_{\mathrm{L},abc} + C\frac{d}{dt} Vabc \quad (2)$$

$$V_{abc} == L \; {}^{d}/_{dt} + i_{L,abc} + R_L i_{L,abc} \quad (3)$$

where $V_{t,abc}$, V_{abc}, $i_{t,abc}$, and $i_{L,abc}$ are the terminal's phase voltages and currents respectively of the inverter at the PCC .

These three-phase instantaneous voltages and currents are to be transformed to the synchronous rotating frame *dq0* in order to have control of active power (d-axis) and reactive power (q-axis), to keep mutual inductance constant, to achieve the desired output, to have infinite gain control on PI, PID, by adjusting integrators, and to make steady state error to zero to make computations easy.

$$X(t) = AX(t) + Bu(t) \quad (4)$$

$$y(t) = CX(t) \quad (5)$$

$$u(t) = V_{td} \quad (6)$$

A, B, C, and D are constants given by:

$$A = \begin{bmatrix} \frac{-Rt}{L_t} & \omega_0 & 0 & \frac{-1}{L_t} & \\ \omega_0 & \frac{-R_t}{L} & -2\omega_0 & \frac{RlC\omega_0}{L} & -\frac{\omega_0}{R} \\ 0 & \omega_0 & \frac{-Rl}{L} & \frac{1}{L} - {\omega_0}^2 C & \\ \frac{1}{C} & 0 & \frac{-1}{C} & -\frac{1}{RC} & \end{bmatrix} \quad (7)$$

$$B^T = \begin{bmatrix} \frac{1}{L_t} & 0 & 0 & 0 \end{bmatrix} \quad (8)$$

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} \quad (9)$$

$$D = [0] \quad (10)$$

$$X^T = [i_{td} \quad i_{tq} \quad i_{Ld} \quad V_d] \quad (11)$$

These equations give the transfer functions of ${}^{V_d}\!/_{V_{td}}$, where V_d and V_{td} are the input and output components of the d axis.

III. NON DETECTION ZONE

The efficacy of islanding detection depends on minimizing the NDZ [4, 5]. The method depends on the percentage of active power mismatch, which as per IEEE-1547 has to be smaller than 15% and the detection time has to be less than 2s. The network for the study of NDZ is shown in Figure 2. The DG is connected to the grid through an interfacing inverter, the PCC, and a utility switch [6]. The 3-phase parallel RLC load is connected at the PCC [7].

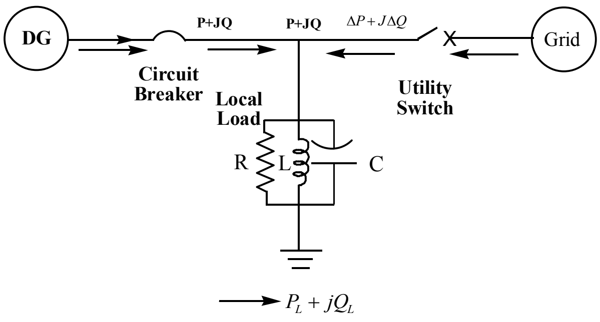


Fig. 2. DG network connected with the grid system.

Equations (12) and (13) give the voltage and frequency at the PCC:

$$P + \Delta P = \frac{v^2}{R} \quad (12)$$

$$Q + \Delta Q = \frac{V^2}{2\pi f L} \quad (13)$$

$$V = \sqrt{R(P + \Delta P)} \quad (14)$$

$$f = \frac{V^2}{2\pi L(Q + \Delta Q)} \quad (15)$$

In islanding conditions, ΔP and ΔQ become zero, as there is no main grid. The under/over voltage protection (UVP/OVP) and under/over frequency protection (UFP/OFP) methods are very simple and incorporated in all grid connected inverters and relays connected with DG system protection. In these methods, the voltage and frequency are both constantly monitored at the PCC [1, 8], where:

$$\Delta P = P_{LOAD} - P_{DG} \quad (16)$$

$$\Delta Q = Q_{LOAD} - Q_{DG} \quad (17)$$

If $\Delta P \neq 0$, then the voltage at PCC will fluctuate from the normal level, which is an indication of islanding condition. If $\Delta Q \neq 0$, then the frequency at PCC will fluctuate from the normal level, which is an indication of islanding condition. Both these islanding detection methods leave behind a large NDZ. The islanding detection may fail, when ΔP and ΔQ are close to zero [8].

Voltage V' and frequency f' under islanding mode are given by:

$$V' = \sqrt{R(P)} \quad (18)$$

$$f' = \frac{V'^2}{2\pi L(Q)} = \frac{RP}{2\pi L(Q)} \quad (19)$$

With these, the voltage and frequency deviations due to power mismatches are given by:

$$\Delta V = V' - V = \sqrt{R(P)} - \sqrt{R(P + \Delta P)} \quad (20)$$

$$\Delta f = f' - f = \frac{V'^2}{L(Q)} - \frac{V'^2}{L(Q+\Delta Q)} = \frac{R\times P}{L\times Q} - \frac{R\times(P+\Delta P)}{L\times(Q+\Delta Q)} \quad (21)$$

Equations (20) and (21) show the variations in voltage and frequency due to power mismatch [2]. If the power mismatch is substantial, the variations in voltage and frequency can be identifiable. If the mismatch is too small (less than 15%) the islanding cannot be detected and hence the formation of the NDZ. Figure 4 shows the NDZ for different power mismatch percentages [8, 9]. When the reference value is 88% and 110% (as per Table I), under / over voltage and grid RMS Voltage V_g = 415V, then *Vmax* = 456V and *Vmin* = 365V. The schematic diagram of the grid-connected microgrid is shown in Figure 3. The inverter gate pulses take the feedback grid voltage and the set reference. The trigger pulses are given to the inverter accordingly and the output voltage is controlled. The PV is connected to the buck/boost converter to have control on DC reference voltage. The DC/DC converter is connected to the interfacing inverter. The inverter is connected to the PCC via a low pass filter LCL. The microgrid is connected to the grid via a circuit breaker, so that the microgrid can be islanded during faults. A local parallel RLC load is connected to the microgrid at the PCC and the microgrid feeds this local load from the DG, even in the absence of grid.

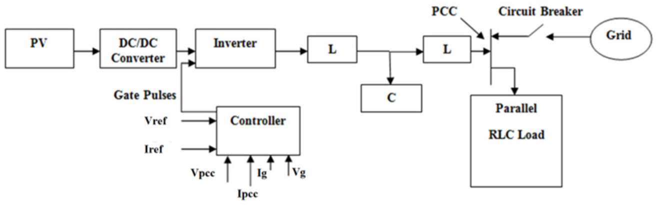


Fig. 3. Schematic diagram of the grid-connected micorgrid.

NDZ is the operating region in which islanding detection methods cannot detect islanding as specified by IEEE-1547. It is expressed in terms of % power mismatch or in terms of parameters of the load like *R*, *L*, and *C*. The NDZ of OVP/UVP and OFP/UVP islanding schemes are shown in Figure 4 [10, 11]. These techniques fail to detect islanding in mismatched power less than 15%. In a distribution network, voltage values, as per standards are between 0.88p.u. and 1.1p.u. for voltage relays. These voltage levels are equivalent to DV = 0.12p.u. and DV = 0.1p.u. respectively. The calculated power mismatch for our test network (the inverter rated output power is 2.5KW) are 0.3KW and 0.25KW respectively.

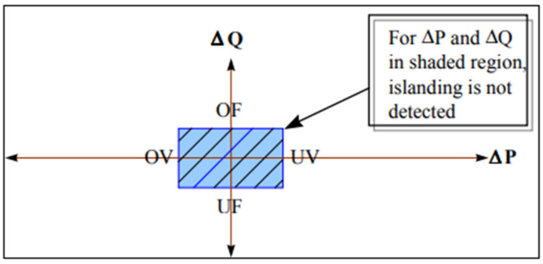


Fig. 4. Mapping of the NDZ in ΔP versus ΔQ for OVP-UFP relays.

In grid mode, the load consumes the reactive power [12]. But in islanding mode, DGs cannot inject reactive power to the load, as DGs operate at unity power factor, because the load behaves like a resistance [13], since the load resonance frequency is equal to the system frequency at the PCC. Hence, to find more deviations in frequency, the load selected is parallel RLC with a high quality factor of 1.8 in islanding mode [14, 15]. The quality factor is given by:

$$Q_f = \omega_0 RC = \frac{R}{\omega_0 L} = R\sqrt{\frac{C}{L}} \quad (22)$$

where $\omega_0 = 2\pi f_0 = \frac{1}{\sqrt{LC}}$.

Equation (20) gives the energy stored in the RLC circuit. High quality factor loads have high capacitance and small inductance with or without a high parallel resistance [16, 17]. The islanding detection is complex with resonant frequency loads of higher quality factor [18, 19]. The percentage mismatch is not the criterion for load parameters [20, 21]. The load reactive power is given by:

$$Q_{Load} = Vrms^2 \left[\frac{1}{\omega L} - \omega c\right] = \Delta Q \quad (23)$$

Equation (21) depicts the variation in reactive power for different values of *L* and *C*. The % mismatch power for OVP/UVP and OFP/UFP relays is shown in Figure 3 and is given by the equations for active power imbalance:

$$\Delta P = 3V\times I - 3(V + \Delta V) \times I = -3 \times \Delta \mathrm{V} \times I \quad (24)$$

$$\Delta Q = 3\frac{V^2}{\omega_n L}(1 - \omega^2 LC) = 3\frac{V^2}{\omega_n L}\left(1 - \frac{\omega_n^2}{\omega_r^2}\right) \quad (25)$$

where ω_n and ω_r are system and resonance frequencies [22, 23]. The system frequency varies until it reaches the resonant frequency of the load in islanding mode and is given by:

$$\omega_{r=}\ \frac{1}{\sqrt{LC}} \quad (26)$$

and the reactive power imbalance is given by:

$$\Delta Q = 3\frac{V^2}{\omega_n\ L}\left(1-\frac{f_n^2}{(f_{n\pm\Delta f}{}^2)}\right) \quad (27)$$

where f_n is the nominal frequency.

The different islanding standards are shown in Table I. The frequency range as per IEEE-1547-2018 is between 49 and 51Hz, the voltage is between 365 and 456V and the range of the parameter $\Delta V = V' - V$ is not considered as the case of zero power mismatch is considered [24]. In this case, the voltage fluctuations will not cross the limits. But in other cases, when load active power $P_L >$ generated active power P_G,(-15%) or $P_L < P_G$ (+15%) the voltage suddenly goes down or up respectively.

TABLE I. ISLANDING STANDARDS

Standard	Detection time (s)	Quality factor	Trip frequency range, nominal frequency f_0 (Hz)	Trip voltage range (V)
IEC 62116	t < 2	1	$f_0 - 15 \le f \le f_0 + 1.5$	$0.88 \le V \le 1.15$
Korean	t < 0.5	1	$59.3 \le f_0 \le 60.5$	$0.88 \le V \le 1.10$
IEEE-1547-2018	t < 2	1	$58.8 \le f_0 \le 61.2$	$0.88 \le V \le 1.10$
IEEE-929-2000	t < 2	2.5	$59.3 \le f_0 \le 60.5$	$0.88 \le V \le 1.10$

IV. THE PROPOSED METHOD OF ISLANDING DETECTION

The proposed method can detect islanding at zero % power mismatch and the detection time is less than in ROCOF. In this method, at first the voltage phase angle is measured at the targeted DG and then the rate of change of voltage phase angle is calculated to detect the islanding phenomenon. In non-islanding condition, the rate of change of the phase angle is negligible after certain time but in islanding this becomes substantial and the islanding is detected. ROCOVPA also avoids nuisance tripping, thus protecting the stability of the microgrid.

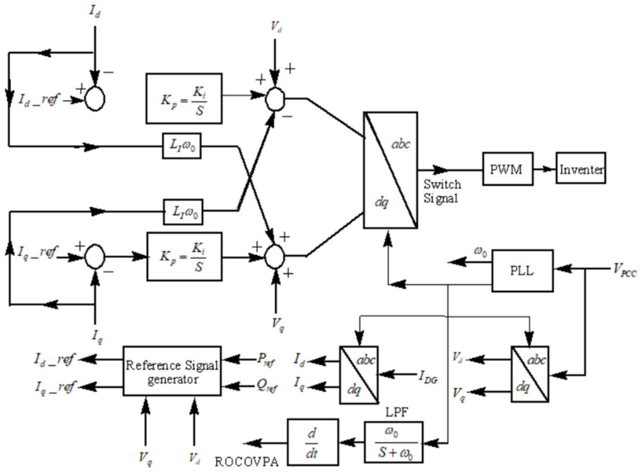


Fig. 5. Current controller block diagram.

The ROCOVPA method is tested with the 2.5KW DG with a current control mode inverter connected to an RLC load with 1.8 quality factor. Figure 5 shows the current control mode to control the active and the reactive power of the load. In the proposed method, the variation of voltage phase angle is monitored at the specified DG. If there is a change in the voltage phase angle, the rate with respect to time is calculated. During the islanding, the deviations of the rate of change of the phase angle exceed the threshold in order to detect the islanding condition. If the relay threshold is fixed, then the trip command to circuit breaker can be initiated.

The algorithm of the proposed ROCOVPA method of islanding detection is shown in Figure 6. The voltage phase angle at the DG is measured first. After the measurement of the phase angle of the voltage, the rate of change of voltage phase angle is calculated. In a normal situation this value is smaller than 1deg/sec (fixed threshold value) but during islanding, the value suddenly crosses the threshold, depending upon the fault severity, by means of which the islanding is detected. During the non-islanding mode this value is in limit, hence nuisance tripping is avoided.

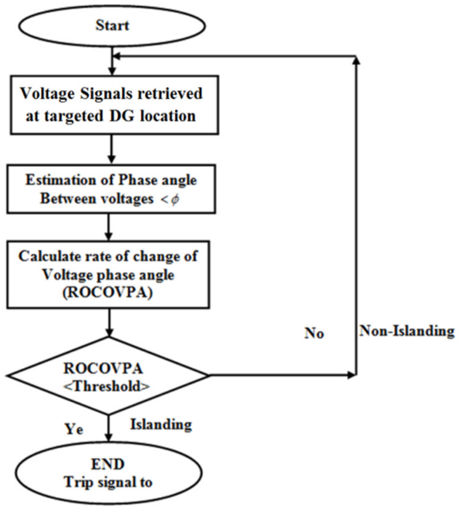


Fig. 6. Flow chart of the proposed ROCOVPA.

V. DESIGN PARAMETERS OF THE INVERTER

The proposed ROCOVPA method was tested on the network shown in Figure 1 with the parameters given in Table II. The DG capacity with interfaced inverter is 2.5KW. The interfaced inverter is connected to the main grid through a breaker via the PCC. A 3-phase parallel RLC load is connected at the PCC. The input DC Voltage to the inverter is 500V. The output line to line voltage of the inverter is 380V. The line resistance and inductance are 1.5mΩ and 2mH respectively. The nominal grid frequency is 50Hz. The inverter carrier frequency is taken as 1KHz. The load parameters with a quality factor of 1.8 are: R_L = 1.76mΩ, L_L =3.2mH, C_L = 3.2μF. The load resonant frequency is 50Hz. Current controller gains are K_p = 0.4 and K_i = 500.

TABLE II. SIMULATION INVERTER PARAMETERS

Component	Value and units
DG power	2.5KW
Switching frequency	10KHz
DC input voltage	500
Line voltage	440V
Filter capacitance C_f	2μF
Filter inductance L_f	5mH
Damping resistance R_f	10hms
Nominal frequency	50Hz
Load resistance R_L	1.76mΩ
Load inductance L_L	3.2 mH
Load capacitance C_L	3.2μF
Load quality factor Q	1.8
Load resonant frequency f_r	50Hz
Current controller proportional gain, k_p	0.4
Current controller integral gain, k_i	500

VI. RESULTS, ANALYSIS, AND DISCUSSION

The designed network was tested in Matlab/Simulink for the islanding case of unintentional unsymmetrical L-G fault and non-islanding cases of connection and disconnection of non-linear load. The Matlab Simulation results of ROCOVPA and ROCOF were compared. It was proved that ROCOVPA is better than ROCOF. The proposed method was tested and compared with ROCOF for the islanding case of unintentional unsymmetrical L-G fault at 0% power mismatch. In this section the simulations are discussed for both ROCOVPA and ROCOF.

A. Islanding Case for Unsymmetrical Fault

An L-G unsymmetrical fault was initiated on the system at the PCC at 0.4s at 0% power mismatch. $P_L = P_G$ is the condition for 0% power mismatch and at that load, a single line to ground fault was initiated on the grid side at 0.4s. The simulation graph is shown in Figure 7.

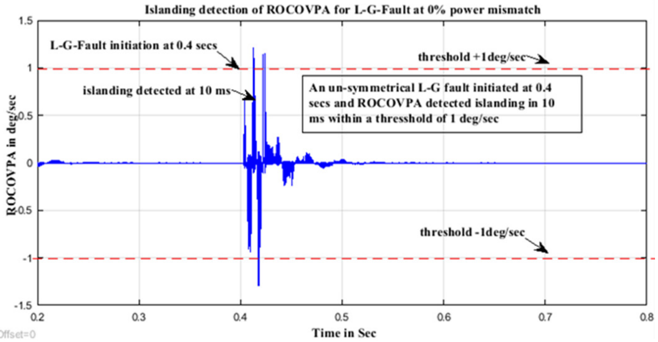


Fig. 7. Islanding detection of ROCOVPA for a L-G fault on the system.

The proposed ROCOVPA detected islanding in 10ms within a fixed threshold of 1deg/s and the relay exactly detected and sent command to trip the circuit breaker to bring the microgrid to islanding mode from the grid mode. The total time is the sum of relay time and breaker time. Any type of fault is to be cleared within 4 cycles (2 cycles, i.e.0.5s of relay operation plus 2 cycles, 0.5s). Hence the ROCOVPA can detect the fault condition within less than a cycle and island the microgrid in around 1s by tripping the circuit breaker, which is less than the 2s of the IEEE-1547 and UL-1741 standards. The same fault conditions were applied and tested with ROCOF in Matlab as shown in Figure 8 and the islanding was detected in 30ms. If the threshold value is fixed at 0.02Hz/sec, the tripping of the circuit breaker can be actuated in around 1s which is below the 2s of the standards. The detection time of ROCOF is more than that of ROCOVPA. As the ROCOF is dependent on frequency, at lower % power mismatches, the threshold value cannot be fixed exactly. Hence, the detection time varies inversely with % power mismatch. To mitigate these issues, ROCOVPA was proposed and proved to be a better islanding detection method for unsymmetrical faults.

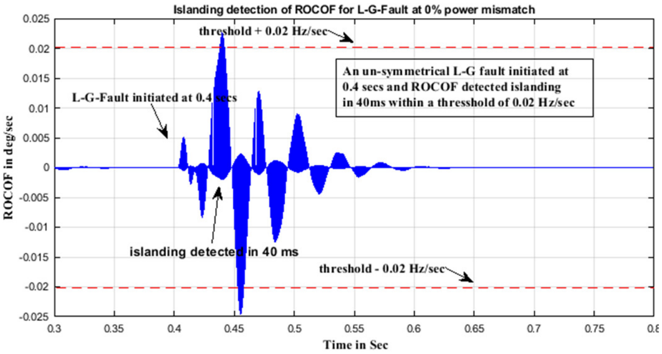


Fig. 8. Islanding detection of ROCOF for a L-G fault on the system.

B. Non-Islanding Case for Non-Linear Load Connection and Disconnection at the PCC on the System

System stability has been studied for different transient conditions during load connection and disconnection at the PCC with non-linear load for both ROCOVPA and ROCOF in Matlab/Simulink. Both methods proved their stability by avoiding nuisance tripping within the threshold values. The ROCOVPA threshold value is 1deg/s and that of ROCOF is 0.02Hz/sec.

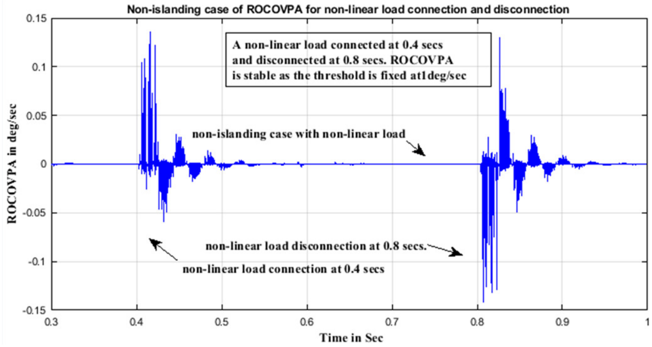


Fig. 9. Non-islanding case of ROCOVPA for non-linear load connection and disconnection.

A non-linear load was connected at the PCC at 0.4s and was disconnected at 0.8s. The results of non-islanding scenerios of ROCOVPA and ROCOF are shown in Figures 9 and 10. The readings of ROCOVPA and ROCOF show that the thresholds are much higher. Hence, the system is stable without any nuisance tripping of the circuit breaker.

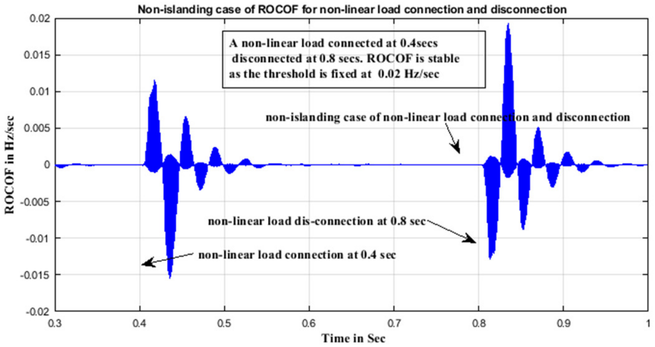


Fig. 10. Non-islanding case of ROCOF for non-linear load connection and disconnection.

VII. CONCLUSIONS

Islanding detection is a major challenge when considering microgrids. The most common unintentional faults on the system are supposed to the be unsymmetrical L-G faults, in which one phase snapping and making phase to ground is involved. The proposed ROCOVPA passive islanding detection method was tested for islanding detection at 0% power mismatch (0% NDZ) along with the widely used ROCOF and proved to be a better alternative having less detection time. The ROCOVPA method was also tested for its stability to avoid nuisance tripping during load connection and disconnection and found to be quite effective. It is also simple and faster in discriminating between islanding and non-islanding operations of the microgrid. The detection is more accurate as the phase angle does not depend on mismatched power like voltage and frequency relays. In the future, the proposed ROCOVPA method can be extended for the detection of symmetrical faults with hybrid DGs.

ACKNOWLEDGMENT

The authors wish to thank the faculty of Electrical and Electronics Department of Koneru Lakshmaiah Educational Foundation (Deemed to be University).

FINANCIAL DISCLOSURE

The authors declare that this research work received no financial support from anywhere.

AUTHORS CONFLICT/COMPETING INTEREST

The authors declare that there is no conflict/competing interest.

REFERENCES

[1] D. Milošević and Ž. Đurišić, "Technique for stability enhancement of microgrids during unsymmetrical disturbances using battery connected by single-phase converters," *IET Renewable Power Generation*, vol. 14, no. 9, pp. 1529–1540, 2020, https://doi.org/10.1049/iet-rpg.2019.0625.

[2] D. Kumar, "A Survey on Recent Developments of Islanding Detection Techniques," *Turkish Journal of Electrical Power and Energy Systems*, vol. 1, pp. 42–53, 2021, https://doi.org/10.5152/tepes.2021.21012.

[3] C. R. Reddy and K. H. Reddy, "Islanding detection for inverter based distributed generation with Low frequency current harmonic injection through Q controller and ROCOF analysis," *Journal of Electrical Systems*, vol. 14, no. 2, pp. 179–191, Jun. 2018.

[4] Y. Elshrief, A. Asham, S. Abdelhaleem, and B. Abozalam, "Methods for protecting Network from Islanding Danger," *Journal of Engineering Research*, vol. 9, no. 2, May 2021, https://doi.org/10.36909/jer.v9i2.9695.

[5] M.-S. Kim, R. Haider, G.-J. Cho, C.-H. Kim, C.-Y. Won, and J.-S. Chai, "Comprehensive Review of Islanding Detection Methods for Distributed Generation Systems," *Energies*, vol. 12, no. 5, Jan. 2019, Art. no. 837, https://doi.org/10.3390/en12050837.

[6] M. Hosseinzadeh and F. Rajaei Salmasi, "Islanding Fault Detection in Microgrids—A Survey," *Energies*, vol. 13, no. 13, Jan. 2020, Art. no. 3479, https://doi.org/10.3390/en13133479.

[7] F. Namdari, M. Parvizi, and E. Rokrok, "A Novel Passive Method for Islanding Detection in Microgrids," *Iranian Journal of Electrical and Electronic Engineering*, vol. 12, no. 1, pp. 82–90, Mar. 2016.

[8] O. Tshenyego, R. Samikannu, and B. Mtengi, "Wide area monitoring, protection, and control application in islanding detection for grid integrated distributed generation: A review," *Measurement and Control*, Apr. 2021, Art. no. 0020294021989768, https://doi.org/10.1177/0020294021989768.

[9] P. Kumar, V. Kumar, and R. Pratap, "Design and verification of hardcore reconfigurable relay for islanding detection and subsequent mode adaptation of microgrid," *International Transactions on Electrical Energy Systems*, vol. 29, no. 4, 2019, Art. no. e2807, https://doi.org/10.1002/etep.2807.

[10] K. Rayudua, A. J. laxmi, P. Soumya, R. Pradeep, and T. Kochito, "Design of Controller for Transition of Grid Connected Microgrid to Island Mode," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 2, pp. 845–854, Apr. 2021, https://doi.org/10.17762/turcomat.v12i2.1093.

[11] F. Chanaa *et al.*, "Islanding Detection Method of a Photovoltaic Installation Destined to Power a RLC Load and Integrated to LV Network," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 4, pp. 299–311, Aug. 2021, https://doi.org/10.22266/ijies2021.0831.27.

[12] E. Pathan, S. A. Zulkifli, U. B. Tayab, and R. Jackson, "Small Signal Modeling of Inverter-based Grid-Connected Microgrid to Determine the Zero-Pole Drift Control with Dynamic Power Sharing Controller," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3790–3795, Feb. 2019, https://doi.org/10.48084/etasr.2465.

[13] 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.

[14] M. Karimi, M. Farshad, Q. Hong, H. Laaksonen, and K. Kauhaniemi, "An Islanding Detection Technique for Inverter-Based Distributed Generation in Microgrids," *Energies*, vol. 14, no. 1, Jan. 2021, Art. no. 130, https://doi.org/10.3390/en14010130.

[15] R. S. Somalwar, D. A. Shinde, and S. G. Kadwane, "Performance Evaluation of Voltage Ripple-Based Passive Islanding Detection Method for a Single-Phase Utility-Connected Micro-Grid System," *IETE Journal of Research*, Nov. 2020, https://doi.org/10.1080/03772063.2020.1844066.

[16] S. Dutta, P. K. Sadhu, M. Jaya Bharata Reddy, and D. K. Mohanta, "Shifting of research trends in islanding detection method - a comprehensive survey," *Protection and Control of Modern Power Systems*, vol. 3, no. 1, Jan. 2018, Art. no. 1 (2018), https://doi.org/10.1186/s41601-017-0075-8.

[17] N. C. Munukutla, V. S. K. Rao Gadi, and R. Mylavarapu, "A Simplified Approach to Controlled Islanding of Power System," in *2019 8th International Conference on Power Systems (ICPS)*, Jaipur, India, Dec. 2019, https://doi.org/10.1109/ICPS48983.2019.9067725.

[18] J. Sridevi, V. U. Rani, and B. L. Rao, "Integration of Renewable DGs to Radial Distribution System for Loss Reduction and Voltage Profile Improvement," in *2019 1st International Conference on Electrical, Control and Instrumentation Engineering (ICECIE)*, Kuala Lumpur, Malaysia, Nov. 2019, https://doi.org/10.1109/ICECIE47765.2019.8974670.

[19] S. Salee and P. Wirasanti, "Optimal siting and sizing of battery energy storage systems for grid-supporting in electrical distribution network," in

2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON), Chiang Rai, Thailand, Feb. 2018, pp. 100–105, https://doi.org/10.1109/ECTI-NCON.2018.8378290.

[20] B. Srinivasarao, S. Lalitha, and Y. Sreenivasarao, "Evaluation of Closed-Loop-P.I.D, Fractional-Order-P.I.D and Proportional Resonant Controlled Micro-Grid-Schemes," *Journal of Computational and Theoretical Nanoscience*, vol. 16, no. 5–6, pp. 2479–2487, May 2019, https://doi.org/10.1166/jctn.2019.7919.

[21] K. Subramanian and A. K. Loganathan, "Islanding Detection Using a Micro-Synchrophasor for Distribution Systems with Distributed Generation," *Energies*, vol. 13, no. 19, Jan. 2020, Art. no. 5180, https://doi.org/10.3390/en13195180.

[22] B. V. Rajanna and K. S. Srikanth, "Grid Connected Inverter for Current Control by Using Anti-Islanding Technique," *International Journal of Power Electronics and Drive Systems (IJPEDS)*, vol. 9, no. 2, pp. 926–932, Jun. 2018, https://doi.org/10.11591/ijpeds.v9.i2.pp926-932.

[23] M. K. Kumar, C. Veeranjaneyulu, and P. S. Nikhil, "Interfacing of distributed generation for micro grid operation," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, no. 4, pp. 427–477, 2018.

[24] K. Jyothi, K. S. Srikanth, and H. S. Jain, "Control of stand alone and grid connected distributed energy resource: A review," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, no. 12, pp. 1603–1615, 2018.

AUTHOR PROFILES

L. Bangar Raju is currently working as an Associate Professor in Swarna Bharati Institute of Technology, Khammam, Telangana State, India. He received his B.E. from the Jawaharlal Nehru Technological University, Kakinada, A.P. India and M.Sc. (Engineering) in Power Systems from the Calicut University. He has got 30 years of industrial experience in power, oil and gas, petrochemical, transmission, distribution, and utilization plants. He has published 10 international Scopus indexed papers. He is a Life Member of ISTE (India) and IE (India). He is currently pursuing his Ph.D. from the Koneru Lakshmaiah Educational Foundation (Deemed to be University), Guntur, A.P., India. His areas of interest are power system protection, distributed generation, and microgrids.

K. Subba Rao is presently the Principal at the College of Engineering and a Professor of Electrical and Electronics Engineering, Department of Koneru Lakshmaiah Educational Foundation (Deemed to be University), Guntur, A.P., India. He received his Ph.D. from Koneru Lakshmaiah Educational Foundation, M.Tech. from CIT, Coimbatore and B.Tech. from Dayalbagh University, Agra. He has earned the best teacher award10 times. He is a life member of ISTE and of the Indian Society for Quality. He has published national and international papers and a book. His areas of interest are power electronics, power systems, microgrids, and Security Systems

English-Vietnamese Cross-Lingual Paraphrase Identification Using MT-DNN

Hung Vo Tran Chi
Faculty of Information Technology
VNU-HCM - University of Science
Ho Chi Minh City, Vietnam
hung.votranchi@gmail.com

Duy Lu Anh
Faculty of Information Technology
VNU-HCM - University of Science
Ho Chi Minh City, Vietnam
lu.duy.pro@gmail.com

Nguyen Le Thanh
Faculty of Information Technology
VNU-HCM - University of Science
Ho Chi Minh City, Vietnam
lethanhnguyen.vn@gmail.com

Dien Dinh
Faculty of Information Technology
VNU-HCM - University of Science
Ho Chi Minh City, Vietnam
ddien@fit.hcmus.edu.vn

Abstract-Paraphrase identification is a crucial task in natural language understanding, especially in cross-language information retrieval. Nowadays, Multi-Task Deep Neural Network (MT-DNN) has become a state-of-the-art method that brings outstanding results in paraphrase identification [1]. In this paper, our proposed method based on MT-DNN [2] to detect similarities between English and Vietnamese sentences, is proposed. We changed the shared layers of the original MT-DNN from original the BERT [3] to other pre-trained multi-language models such as M-BERT [3] or XLM-R [4] so that our model could work on cross-language (in our case, English and Vietnamese) information retrieval. We also added some tasks as improvements to gain better results. As a result, we gained 2.3% and 2.5% increase in evaluated accuracy and F1. The proposed method was also implemented on other language pairs such as English – German and English – French. With those implementations, we got a 1.0%/0.7% improvement for English – German and a 0.7%/0.5% increase for English – French.

Keywords-MT-DNN; BERT; XLM-R; English; Vietnamese; cross-language; paraphrase identification

I. INTRODUCTION

Paraphrase Identification (PI) is a task in Natural Language Processing (NLP) that concerns detecting a pair of text fragments that has the same meaning at different textual levels [1]. PI has a relation with the way we would quantify the number of mutual semantics between two text fragments. Measuring how two text fragments are semantically related is essential. The example in Table I was taken from the MRPC (Microsoft Research Paraphrase Corpus) dataset and was translated to Vietnamese. PI applications can be found in search engines, legal matters, or, especially, plagiarism check. A few solutions have been proposed for the mentioned problem such as the Fuzzy-based method and the BabelNet semantic network [5] or Siamese LSTM (Long-Short Term Memory) [6] but they are still limited.

TABLE I. AN EXAMPLE OF TWO ENGLISH PARAPHRASED SENTENCES AND THEIR TRANSLATION INTO VIETNAMESE

Sentence A	Singapore is already the United States' 12th-largest trading partner, with two-way trade totaling more than $ 34 billion. Singapore đã là đối tác thương mại lớn thứ 12 của Hoa Kỳ, với tổng kim ngạch thương mại hai chiều hơn 34 tỷ USD.
Sentence B	Although a small city-state, Singapore is the 12th-largest trading partner of the United States, with trade volume of $ 33.4 billion last year. *Mặc dù là một thành phố nhỏ, Singapore là đối tác thương mại lớn thứ 12 của Hoa Kỳ, với kim ngạch thương mại đạt 33,4 tỷ USD vào năm ngoái.*

With the use of transfer learning by applying the pre-trained model in machine learning models, NLP tasks in general or the PI task in particular had significant improvements in their results. Among many pre-trained models, Bidirectional Encoder Representation Transformer (BERT) [3], used to be considered as a state-of-the-art model, with impressive results in many NLP tasks. In this paper, the MT-DNN model was applied which is a combination of pre-trained models like BERT [3], Multilingual BERT (M-BERT) [3], or Cross-Language Model RoBERTa (XLM-R) [4] with Multi-Task Learning (MTL). The objective of this study is to improve the PI task between pairs of multilingual documents (namely English and Vietnamese) through applying transitional learning from a better pre-trained language model, with an MTL approach, including adding new improved tasks.

II. LITERATURE REVIEW

A. Pre-Trained Model and Transfer Learning

Transfer learning is a method of using pre-trained models and then optimize them for our purposes. A pre-trained model is a model that has been previously trained with a large dataset or with advanced methods to reduce the effort of training from scratch [8]. The model can then be further trained to fit the

---

Corresponding author: Hung Vo Tran Chi

actual data set or used directly in a machine learning problem. Between the pre-trained models, BERT [3] which was a state-of-the-art model in 2019, is still a solution worth considering. Specifically, BERT was constructed based on transformer (or attention's mechanism), which is a deep learning model having combined inputs and outputs where the model calculates their weights. BERT was trained on 2 main tasks: MLM (Masked Language Modeling) and NSP (Next Sentence Prediction). BERT has been applied in many applications [9, 10]. M-BERT is a single language model pre-trained from 104 languages (including English and Vietnamese). It showed abilities to not only generalize cross-lingual but also to transform scripts between many languages without having lexical overlap [3]. XLM-R was trained with MLM only, but with much more data and was based on Large-BERT. The solution used in XLM-R to deal with the burden of multilinguality was to increase model capacity by learning much more data than before. Authors in [4] showed that for the first time we can have a single large model for all languages without losing performance for any language.

B. Paraphrase Identification Methods

Before the rising of deep learning, the most common solutions were lexical, syntactic, semantic, or hybrid techniques. After that, some sophisticated approaches have been applied by supervised or unsupervised learning.

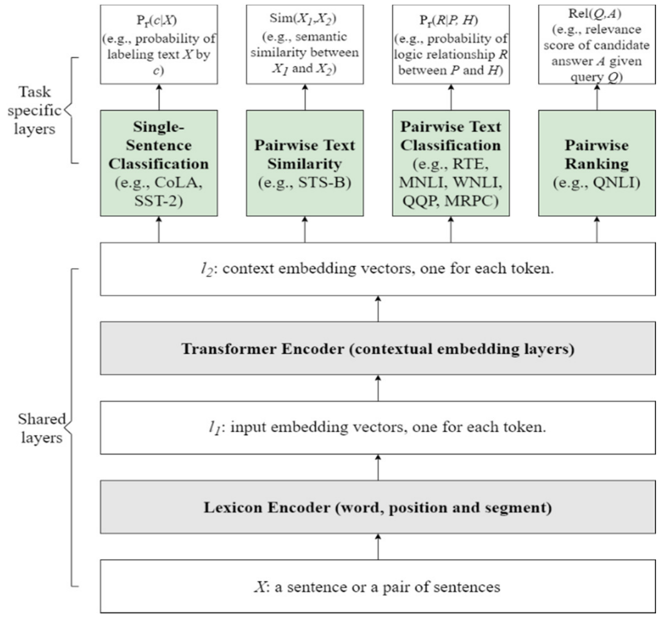


Fig. 1. The original architecture of the MT-DNN.

By using supervised learning, many traditional techniques were applied [10-13] or more advanced, using pre-trained models, specifically BERT and its improved models such as M-BERT, XLM-R, etc. MTL is an approach that achieves the generalization of results by using the inductive transfer method [14]. With MTL, we assume that knowledge gained from previous tasks could help us achieve better results when learning a new task. The main benefits from MTL are: It helps us build a platform from the previous task in order to gain better results and it does not need a big dataset with labeled data (which is hard to get) as the Deep Neural Networks (DNNs) do [15]. Furthermore, this approach creates a universal presentation without getting bias in any particular tasks. It could be achieved by regularized effects thus minimize overfitting error [15]. In the original MT-DNN architecture, we have two main layers: Shared and Task-specific. The first layer type takes the input and encodes them using Lexicon Encoder and Transformer Encoder (through the Attention mechanism) to create embedding context vectors. These vectors are shared through all tasks. After that, the Task-specific layer (depended on the tasks we use) will take these embedding vectors and process them to get the final results (e.g. with the PI problem, MT-DNN's task-specific layers will use the formula $Sim(X_1, X_2) = w_{STS}^T.x$ with X_1, X_2 being the input pair and w_{STS} the weight matrix)

III. THE PROPOSED METHOD

A. Schematic Overview

The overview of the proposed method can be seen in Figure 2.

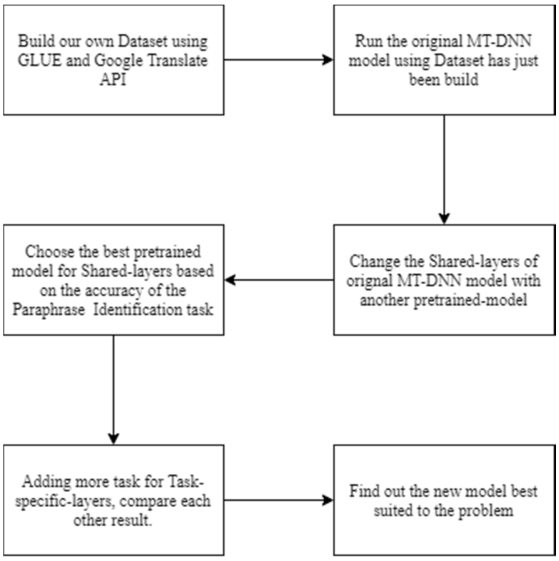


Fig. 2. The proposed solution.

B. Updating the Pretrained Model

With the change of subject in our model from a single language to cross-language, we need to change the shared layers of the original MT-DNN in Figure 1 due to two main reasons:

- The original BERT used in MT-DNN only can be used for English but for no other languages.
- With many other multi-language pre-trained models, we can achieve not only better embedding but also have a positive effect on our results from the transfer learning method. We replaced BERT with M-BERT and XLM-R to be able to work on the English-Vietnamese language pair. M-BERT was not trained specifically for having shared presentations through languages.

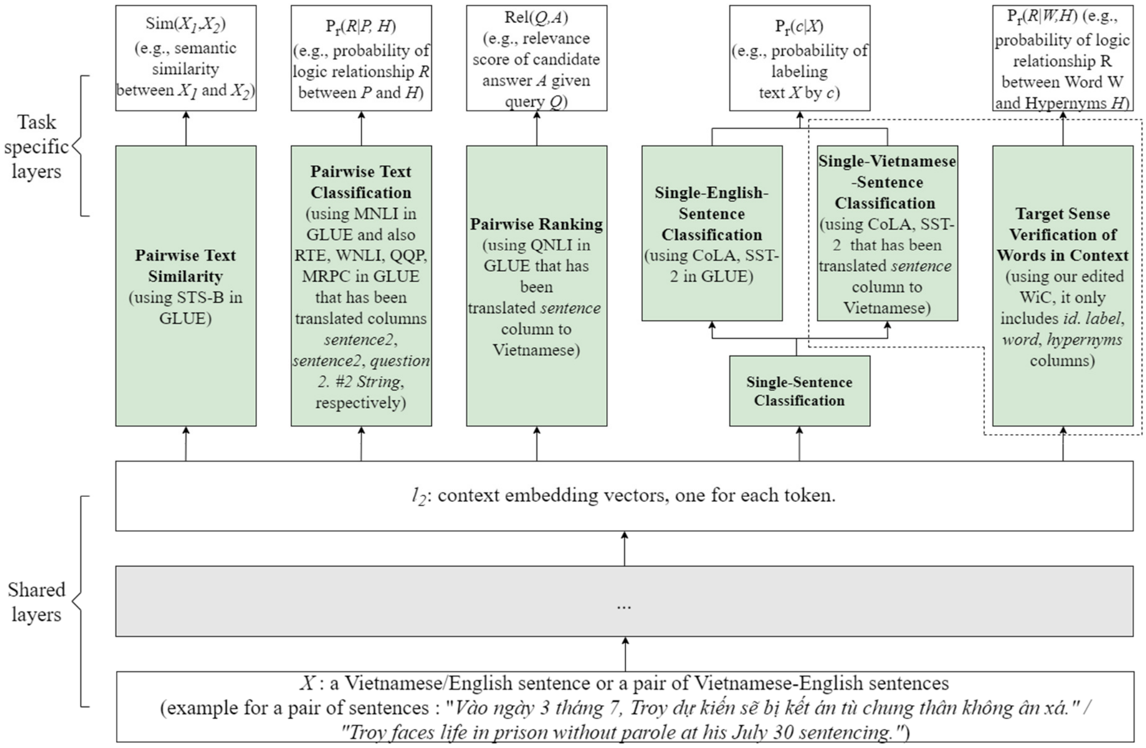


Fig. 3. The proposed model.

C. Adding More Tasks

1) CoLA-Vie (Corpus of Linguistic Acceptability in Vietnamese)

The existing CoLA [16] dataset was added to our model. Translating the original dataset to Vietnamese and training our model on this dataset could create a mapping in grammaticality between Vietnamese and English. The CoLA dataset in the original language (English) helped the model to learn and get many unacceptable sentences for various reasons such as pragmatical anomalies, unavailable meanings, syntactic or semantic violations, etc. [16]. By translating the original CoLA dataset to Vietnamese, we want our model to be able to learn when to accept a sentence in Vietnamese with the same reasoning, which creates mappings in how we include and exclude sentences between Vietnamese and English. With our evaluation task, which is MRPC, if two sentences in Vietnamese and English are both acceptable, they will be more likely to be similar or paraphrased.

2) SST-Vie (Stanford Sentiment Treebank in Vietnamese)

Contrary to the original SST, this dataset instead of using 5 classes (very negative, negative, neutral, positive, very positive), only uses 2 classes (positive and negative) which allows us to create easier and more accurate mapping. It does not only look for specific words such as love or hate (which can be wrong in cases of sarcasm or negative words being positive) but checks the whole sentence. With this, we can make use of pre-trained models presenting the context much better than only considering some words. The dataset was translated to Vietnamese because:

- It can help the model learn the presentation of context and vocabulary in Vietnamese in order to improve accuracy when doing tasks that involved Vietnamese.
- To create class mapping in order to make our model able to differentiate negative or positive in Vietnamese. Thus, if two sentences in Vietnamese and English are both in the same class, the chance they are paraphrased is higher.

3) NER-Vie (Named Entity Recognition in Vietnamese)

We used the datasets from CoNLL (English) [17] and VLSP (Vietnamese) [18] in order to recognize the named entities – which can help us extract information and get tags for words in a sentence - in both languages and create the mapping between them. When the classes between CoNLL and VLSP datasets were compared, it was found that they had the same labels. With that, we added this task to:

- Help our model to learn how to recognize the named entities in both Vietnamese and English so it can extract information more efficiently (which is the original purpose of this task).
- Create mappings of the named entities between two languages. If two sentences in cross-language have several same-named entities, they are more likely to be relevant than those that do not.

4) PoS-Vie (Part-of-Speech Tagging in Vietnamese)

With the same idea of NER, we tried to make our model recognize PoS tags. When checking the dataset of CoNLL and VLSP, we saw differences in quantity (49 to 35) and the number of similar tags (approximately 80%). The reasons behind this utilization are:

- Learning how to tag PoS for sentences in both Vietnamese and English could not only help our model to extract information but also to check the grammar in both languages.
- If we have more identical part-of-speech tags in both sentences, the chance of being relevant will be higher. Because the number of similar tags is little, we want our model to be able to create mappings between different tags in both languages.

5) WiC (Word in Corpus) [19]

The dataset of this task included 3,832 samples with 4 attributes. Originally, this task is a binary classification problem that wants to define whether some words are the hypernyms of another word or not. We only used the labels combined with the hypernyms and the original words and removed the definitions of that word. As well as extending our vocabulary in English, we assume that if our model can learn the connection and the hierarchy between words, so it will not only be able to detect the link between a word in English and Vietnamese but also to expand to its hypernyms in English as well. Consequently, when we compare two sentences in Vietnamese and English, not only words having the same meaning are checked but also any hypernyms, if they exist in English sentences, will influence the result.

6) SemEval Task 8 – 2010 [20]

This dataset includes 8,000 samples with 3 attributes. By adding this dataset, we tried to expand MNLI (Multi-Genre Natural Language Inference) and SNLI (Single Natural Language Inference). While MNLI only has some simple relationships (contradiction, neutral, and entailment), this task will specify the relationships with more details by increasing the number of relationships (19 classes in the dataset which includes 9 pairs of active or passive relationships and 1 class for others) and the positions of relationship's subjects in the sentences. By using this task in the proposed model, we are able:

- To use the benefit from NER and PoS tasks if they improve our result, especially the specified pair of subjects in sentences so we can specify the relationship between them.
- To expand the benefits from MNLI and SNLI for a more specific classification of relationships. If two words belong to the same classification, they will be more likely to be paraphrased.

TABLE II. THE CONFIGURATION OF MT-DNN ARCHITECTURE OF BOTH XLM-R AND M-BERT

		num_embeddings	**embedding_dim**	**padding_idx**	
Embeddings	**Word_embeddings**	250002	768	1	adding NormLayer: size(768,) and Dropout(p=0.1)
	Position_embeddings	514	768	1	
	Token_type_embeddings	1	768	1	
		in_features	**out_features**	**bias**	
Encoder: Include 11 linear BERT layers	**BertSelfAttention**	768	768	TRUE	3 linear layers for query + key + value and dropout layers (p=0.1)
	BertSelfOutput	768	768	TRUE	adding NormLayer: size(768,) and Dropout(p=0.1)
	BertIntermediate	768	3072	TRUE	
	BertOutput	3072	768	TRUE	adding NormLayer: size(768,) and Dropout(p=0.1)
Pooler	**Dense layer**	768	768	TRUE	With activation function = Tanh
Scoring list: Number of layers = number of tasks	**Linear**	768	number of classes of this task	TRUE	

IV. IMPLEMENTATION

A. Preparing the Dataset

We started with the original GLUE (General Language Understanding Evaluation) dataset [21]. We used 9 out of 10 tasks in our dataset (excluding the AX dataset for format reasons). These 9 datasets represent 4 tasks that will be learned in our model. To make it more appropriate with our problems, we translated the dataset as follows:

- Single-Sentence Classification (CoLA and SST-2): We translated all to Vietnamese and used both versions (Vietnamese and English) to create Single-Vietnamese-Sentence Classification (SVSC).
- Pairwise Text Classification (MNLI, RTE, WNLI, QQP, and MRPC dataset) and Text Similarity (STS-B - Question-Answering Natural Language Inference): We kept the question in English and translated the answer to Vietnamese. The AX (diagnostic in GLUE) dataset was excluded because we should create labels by using a model trained on MNLI.
- Relevance Ranking (QLNI - Question-Answering Natural Language Inference): We kept the question in English and translated the answer to Vietnamese.

Then, NER and POS from CoNLL for English and from VLSP for Vietnamese were added to make a mapping in entities from English to Vietnamese and backwards.

In the end, we added the WiC-TSV and SemEval-Task8 2010 datasets to enhance the connection of words and relationships in English. We also translated SemEval-Task 8 2010 to Vietnamese for learning relationships in this language.

After translation, we used preprocessing scripts to encode our data with XLM-R (or M-BERT) to make input for our model.

So, our model inputs and outputs were:

- Inputs: A pair of sentences in English - Vietnamese
- Output: Whether they are paraphrased or not

Example: Inputs: There are 103 Democrats in the Assembly and 47 Republicans - Đảng Dân chủ chiếm ưu thế trong Quốc hội trong khi Đảng Cộng hòa kiểm soát Thượng viện. Output: Not paraphrased (True – in original methods they are marked as paraphrased)

B. Training

For each time of traning, we selected the task we wanted to add to our original model and the hyperparameters. After many epochs, we got the checkpoint model which had best results on our MPRC dataset

C. Fine-Tuning

In the fine-tuning stage, we striped the task-specific layers (included other tasks which are not MRPC) and trained our model (which has the best checkpoint from training) with the MRPC dataset again to achieve as best results as possible. The reason for our fine-tuning step is that it improves the result significantly with a little consumption of time and resources. As in the training step, we got the checkpoint model which had the best results on the MPRC dataset after many epochs.

D. Configuration

We used Google Colaboratory Pro configuration with: Intel Xeon R (2 cores) 2.20 Ghz CPU, NVIDIA V100-SXM2 (16GB VRAM HBM2) GPU, 12.72GB RAM, and 150GB Hard Disk.

TABLE III. RESULTS OF VIETNAMESE-ENGLISH PAIRS

Task	**No finetuning**		**With finetuning**	
	Accuracy	**F1 Score**	**Accuracy**	**F1 Score**
Original Architecture (XLM-R) – Standard result	82.8	87.6	84.3	88.5
Original Architecture (XLM-R) + SVSC	80.1	86.2	85.7	89.6
Original Architecture (XLM-R) + SVSC, NER (English), POS (English)	79.1	85.2	84.8	88.8
Original Architecture (XLM-R) + SVSC, NER (English)	77.9	84.3	81.6	86.3
Original Architecture (XLM-R) + SVSC, POS (English)	80.6	86.4	83.8	88.3
Original Architecture (XLM-R) + SVSC, NER (English, Vietnamese), POS (English, Vietnamese)	80.6	86.3	85.5	89.6
Original Architecture (XLM-R) + SVSC, POS (English, Vietnamese)	82.1	87.3	83.3	87.3
Original Architecture (XLM-R) + SVSC, NER (English, Vietnamese)	79.4	85.4	85.2	89.2
Original Architecture (XLM-R) + SVSC, WiC-TSV	**83**	**88.1**	**87**	**91**
Original Architecture (XLM-R) + SVSC, WiC-TSV, NER (English, Vietnamese)	80.6	86.7	86	89.7

When configuring the pre-trained models, we used M-BERT and XLM-R with the following hyperparameter specifications:

- Batch-size: 8
- Learning rate: 5×10^{-5}
- Gradient accumulation step: 4 (for making training more stable and faster)
- While they have a different number of parameters, both XLM-R and M-BERT have the same structure as presented in Table II.

TABLE IV. SPECIFICATIONS OF M-BERT AND XLM-R

Name	**Number of layers**	**Detail name (in Huggingface)**
M-BERT	177865744	bert-base-multilingual-cased
XLM-R	278060566	xlm-roberta-base

V. RESULT AND DISCUSSION

The results from Table V indicate that XLM-R shows an outstanding performance when compared to M-BERT with a 9% and 6% increase in Accuracy and F1 Score respectively before the finetuning stage. After the finetuning stage, that gap expanded to 11% and 7%. That result could be explained by the way M-BERT was trained, since it did not have shared presentations through languages specifically.

TABLE V. RESULT OF THE ORIGINAL MODEL WITH M-BERT AND XLM-R

Dataset	**Pretrained model**	**No finetuning**		**With finetuning**	
		Accuracy	**F1 Score**	**Accuracy**	**F1 Score**
All Dataset in Glue	M-BERT	73.2	81.4	73.7	81.3
All Dataset in Glue	XLM-R	82.8	87.6	**84.3**	**88.5**

On the other hand, XLM-R was trained with Common Crawl [4] which helps in gaining data from languages with low resources. With the result of 84.3% in Accuracy and 88.5% in F1 score, we chose that result as our standard to compare when adding the improvement tasks. In Table III, when we look at the Accuracy and F1 score in adding CoLA-vie and SST-vie, the results were all decreased by approximately 1-4% (Accuracy) and 0.5-3% (F1 score). Only the model with the additional WiC task had a slightly higher result. After the finetuning stage, the results improved. The accuracy increased by 0.5-2.7% and the F1 score by 0.3-2.5% (with the highest result firmly belonging to the model which the WiC task). In the final model, we had 83% and 88.1% before finetuning (Accuracy and F1 score respectively). After the finetuning stage, those numbers increased to 87% and 91%. Besides this increase, adding NER-en and POS-en tasks made our model's performance decrease.

While the change of the evaluation's result after the finetuning stage was proved in the paper, adding PoS-en and PoS-vie did not improve the result. When looking at the data from CoNLL (English) and VLSP (Vietnamese), they have differences in the number of total and identical tags. Due to this, we must transfer those data to a universal PoS tag. Because of the differences in sentence structure, accurate mappings could not be created between the two languages and the context of learning was different, which incommodated our model. When we added NER-en and NER-vie, we also found that the result did not improved because the units in the vocabulary of English and Vietnamese have differences. In English, a word is a single unit while in Vietnamese, a word can be formed by many units (e.g in English we have "toothbrush" whereas in Vietnamese will be "bàn-chải"). Adding the WiC task made the difference in terms of Accuracy and F1 score. Because this task did not care about the way words are built like NER or POS but concerned more about the meaning of the sentence or the connection between words, which improved our model significantly. In Table VI, by applying the above improvements (SVSC + WiC-TSV) in German and French, we proved that adding this task not only worked for the original pair but also for other language pairs. More specifically, the result for English – German increased by 1.0% in terms of Accuracy (from 86.2% to 87.2%) and 0.7% in F1 score (from 90.1% to 90.8%). In English – French, those numbers were 0.7% (from 85.5% to 86.2%) and 0.5% (from 89.7% to 90.2%).

TABLE VI. RESULTS OF VIETNAMESE-ENGLISH PAIRS

Language pair	**No finetuning**		**With finetuning**	
	Accuracy	**F1 Score**	**Accuracy**	**F1 Score**
English - French	83.5	88.4	85.5	89.7
	80.3	86.4	**86.2**	**90.2**
English - German	81.8	87	86.2	90.1
	82.1	87.3	**87.2**	**90.8**

For further improvements, many pre-trained models with better results could be applied to the model. Another improvement is to standardize the dataset in the translation step so it could make the mappings more accurate.

VI. CONCLUSION

In this paper, the application of MT-DNN with transfer learning (by using a pre-trained model improved from BERT, M-BERT, and XLM-R), combined with modified MTL for the cross-language English - Vietnamese pair to achieve competitive performance in paraphrase identification task, was studied and presented. The result evaluation stage confirmed the suitability of the proposed model which includes XLM-R, SVSP, and TSV of WiC which helped obtaining better results, such as 87% and 91% in Accuracy and F1 Score. Besides the original language pair, our proposed method also had a good performance for English – German and English – French pairs with Accuracy and F1 Score of 86.2% and 90.2% and 87.2% and 90.8% respectively. Our model can be improved by changing the pre-trained models with a state-of-the-art model in the future.

ACKNOWLEDGEMENT

This research is partially funded by the Faculty of Information Technology, University of Science, Ho Chi Minh City, Vietnam.

REFERENCES

[1] A. Amaral, "Paraphrase Identification and Applications in Finding Answers in FAQ Databases." 2013, [Online]. Available: https://fenix.tecnico.ulisboa.pt/downloadFile/395145918749/resumo.pdf.

[2] X. Liu, P. He, W. Chen, and J. Gao, "Multi-Task Deep Neural Networks for Natural Language Understanding," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy, Jul. 2019, pp. 4487–4496, https://doi.org/10.18653/v1/P19-1441.

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv:1810.04805 [cs]*, May 2019, Accessed: Aug. 26, 2021. [Online]. Available: http://arxiv.org/abs/1810.04805.

[4] A. Conneau *et al.*, "Unsupervised Cross-lingual Representation Learning at Scale," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, Jul. 2020, pp. 8440–8451, https://doi.org/10.18653/v1/2020.acl-main.747.

[5] L. T. Nguyen and D. Dien, "English- Vietnamese Cross-Language Paraphrase Identification Method," in *Proceedings of the Eighth International Symposium on Information and Communication Technology*, New York, NY, USA, Dec. 2017, pp. 42–49, https://doi.org/10.1145/3155133.3155187.

[6] D. Dinh and N. Le Thanh, "English–Vietnamese cross-language paraphrase identification using hybrid feature classes," *Journal of Heuristics*, Apr. 2019, https://doi.org/10.1007/s10732-019-09411-2.

[7] M. Mohamed and M. Oussalah, "A hybrid approach for paraphrase identification based on knowledge-enriched semantic heuristics," *Language Resources and Evaluation*, vol. 54, no. 2, pp. 457–485, Jun. 2020, https://doi.org/10.1007/s10579-019-09466-4.

[8] 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.

[9] S. Mandava, S. Migacz, and A. F. Florea, "Pay Attention when Required," *arXiv:2009.04534 [cs]*, May 2021, Accessed: Aug. 26, 2021. [Online]. Available: http://arxiv.org/abs/2009.04534.

[10] B. Ahmed, G. Ali, A. Hussain, A. Baseer, and J. Ahmed, "Analysis of Text Feature Extractors using Deep Learning on Fake News," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 7001–7005, Apr. 2021, https://doi.org/10.48084/etasr.4069.

[11] R. Mihalcea, C. Corley, and C. Strapparava, "Corpus-based and knowledge-based measures of text semantic similarity," in *Proceedings of the 21st national conference on Artificial intelligence*, Boston, MA, USA, Jul. 2006, vol. 1, pp. 775–780.

[12] W. Yin and H. Schütze, "Convolutional Neural Network for Paraphrase Identification," in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, CO, USA, May 2015, pp. 901–911, https://doi.org/10.3115/v1/N15-1091.

[13] H. Shahmohammadi, M. Dezfoulian, and M. Mansoorizadeh, "Paraphrase detection using LSTM networks and handcrafted features," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 6479–6492, Feb. 2021, https://doi.org/10.1007/s11042-020-09996-y.

[14] R. Caruana, "Multitask Learning," *Machine Learning*, vol. 28, no. 1, pp. 41–75, Jul. 1997, https://doi.org/10.1023/A:1007379606734.

[15] M. Crawshaw, "Multi-Task Learning with Deep Neural Networks: A Survey," *arXiv:2009.09796 [cs, stat]*, Sep. 2020, Accessed: Aug. 26, 2021. [Online]. Available: http://arxiv.org/abs/2009.09796.

[16] A. Warstadt, A. Singh, and S. R. Bowman, "Neural Network Acceptability Judgments," *Transactions of the Association for Computational Linguistics*, vol. 7, pp. 625–641, Mar. 2019, https://doi.org/10.1162/tacl_a_00290.

[17] E. F. Tjong Kim Sang and F. De Meulder, "Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition," in *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, 2003, pp. 142–147.

[18] H. T. M. Nguyen, Q. T. Ngo, L. X. Vu, V. M. Tran, and H. T. T. Nguyen, "VLSP Shared Task: Named Entity Recognition," *Journal of Computer Science and Cybernetics*, vol. 34, no. 4, pp. 283–294, 2018, https://doi.org/10.15625/1813-9663/34/4/13161.

[19] A. Breit, A. Revenko, K. Rezaee, M. T. Pilehvar, and J. Camacho-Collados, "WiC-TSV: An Evaluation Benchmark for Target Sense Verification of Words in Context," in *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, Online, Apr. 2021, pp. 1635–1645.

[20] I. Hendrickx *et al.*, "SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations between Pairs of Nominals," in *Proceedings of the 5th International Workshop on Semantic Evaluation*, Uppsala, Sweden, Jul. 2010, pp. 33–38.

[21] A. Wang, A. Singh, J. Michael, F. Hill, O. Levy, and S. Bowman, "GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding," in *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, Brussels, Belgium, Nov. 2018, pp. 353–355, https://doi.org/10.18653/v1/W18-5446.

TV Ad Detection Using the Base64 Encoding Technique

Waseemullah
Department of Computer Science and IT
NED University of Engineering & Technology
Karachi, Pakistan
waseemu@neduet.edu.pk

Muhammad Faraz Hyder
Department of Software Engineering
NED University of Engineering & Technology
Karachi, Pakistan
farazh@neduet.edu.pk

Maria Andleeb Siddiqui
Department of Software Engineering
NED University of Engineering & Technology
Karachi, Pakistan
mandleeb@cloud.neduet.edu.pk

Muhammad Mukarram
Department of Computer Science and IT
NED University of Engineering & Technology
Karachi, Pakistan
mukki0303@gmail.com

Abstract-Automatic TV ad detection is a challenging task in computer vision. Manual ad detection is considered a tedious job. Detecting advertisements automatically saves time and human effort. In this paper, a method is proposed for detecting repeated video segments automatically, since generally, ads appear in TV transmissions frequently. At first, the user is allowed to browse the advertisements needed to be detected, and the video in which they are to be detected. The videos are then converted into a text file using the Base64 encodings. In the third step, the advertisements are detected using string comparison methods. In the end, a report, with the names of the advertisements is shown against the total time and the number of times these advertisements appeared in the stream. The implementation was carried out in python.

Keywords-TV ads; ad detection; base64

I. INTRODUCTION

Advertisements displayed in TV broadcasts are a very important part of a transmission as most of the revenue of a broadcaster is generated by advertising. Fast and accurate advertisement discovery is an important issue in the computer vision field. The main challenge of ad detection is the lack of information about the TV transmission structure and the unpredictable appearance of advertisements in the transmissions. In this paper, a novel approach is developed for TV ad detection. An algorithm has been developed for TV commercial (ad) detection. The proposed algorithm breaks the videos into frames. The sub-regions of frames take part in the comparison. The pixel values of the middle region of the frames of the TV transmited video and the ad segment are compared. This approach results in TV ad detection with a 60% precision. Meanwhile, it suffers from the problem that the results rely heavily on the size of the video file. Another factor that significantly influences ad detection is the quality of the video. The proposed algorithm does not provide good results on grainy and poor-quality videos. The developed system based on the novel approach for TV ad detection can be used by the PEMRA (Pakistan Electronic Media Regulatory Authority) to identify particular ads and their statistics. The framework is also usable for advertisement's agencies to compute and visualize the air time used for different ads. The experimental results confirm the importance of the proposed framework. This study may be helpful to (Figure 1):

- TV ad detection and identification: Many companies in marketing research and advertisement are keen on identifying commercial segments from TV broadcasts to verify the number of times a particular ad has been aired.
- TV advertiser concerns: A product owner has to keep a close eye on their competitors' marketing tactics by knowing the number of advertisements being aired per day and the total amount spent for advertising the products of a particular vendor.
- TV consumer woes: To give viewers an uninterrupted service is an expensive luxury nowadays as the bombardment of ads in every program is taking a toll on audience viewing. A viewer may find it easier to avoid any such occurrence in advance if the system is equipped with a consumer-friendly tool that reduces the ads' nuisance in the future.
- Market analysts inspection: In relation to the revenue aspect, a media monitoring agency such as PEMRA may need to know the broadcasted time ratio for ad and non-ad transmission of a TV channel. The resulting information may help the agency answer questions such as how much revenue a TV channel is generating through advertisements and which advertiser is paying higher for advertisements.
- Government regulatory body surveillance: Government agencies have to ensure that there should be no breaching in regulationg of enforced laws during channel transmission.

Corresponding author: Waseemullah

One such agency would find useful a video segment identification application. On the other hand, many media observatory and regulatory bodies, want to know the real revenue generated by the various media channels.

Fig. 1. The possible applications for TV commercials detection.

Many approaches based on different features have been discussed [1, 2] for video detection and classification tasks. Authors in [3] investigated the problem of retrieving ads depending on their salient semantics. Semiotic is the knowledge of signs, associating signs with their meaning by communal conventions and particular cultural context. Identifying cuts, dissolve, and rhythm is used for video segmentation. Hough transform is used for calculating significant line slopes for describing shot content cluster analysis. Practical, playful, topic, and critical were recognized as semiotic categories of commercials. Evaluation was conducted on 150 ads from different Italian channels in compliance with the one conducted by human experts. The system showed best results on playful and worst results on practical ads. Authors in [4] explored the problem of real-time commercial detection using MPEG features in MPEG compressed videos. Black frame, unicolor frame, and change in aspect ratio were used for ad detection. It was shown that the least duration for any ad is one minute. Also, it was found that the strongest ad detecting parameter is the presence of black frames in commercial breaks. Authors in [5] proposed a new learning-based approach for ad detection. The approach uses several audio and visual features for classification based on Support Vector Machine (SVM). Average of edge change ratio, the variance of edge change ratio, average of frame difference, and variance of frame difference were used as visual features. NBC, ESPN2, and CNN TV channel transmissions of 10.75 hours was recorded for the evaluation, including different genres, i.e. movies, sports, and news. Recall values of 88.21% and 91.77% were observed without and with post-processing respectively. The respective precision values were 89.39% and 91.65%. Authors in [6] investigated the problem of identification and categorization of ads from TV transmission. Boundary detection was performed through a multi-modal approach. The ads were separated through black frame and silence features. Classification was conducted through text detection. In addition, the absence of a channel logo was also used for ad segmentation from the normal transmission. Hidden Markov model based on audiovisual features was used for training. The results showed a precision value of 90% and a recall value of 80%.

Authors in [7] proposed a commercial detection that is depending on cookery/cooking programs. Audio-visual features were used for ommercial boundary detection, including zero crossing rate, short time energy, edge detection, and corner detection. Furthermore, the logo of the program name was matched with a commercial break. Authors in [8] proposed a signal-based approach that uses an automatic unsupervised method for the segmentation of TV transmission. Using the general likelihood ratio and the Bayesian information criterion ,this approach can be applied to audio signals, visual signals, or their combination. The evaluation was performed on recordings from French TV and the TRECVid dataset with recall values of 93% and 89% respectively. The respective precision was 93% and 91%. An approach based on Optical Character Recognition (OCR) and Automatic Speech Recognition (ASR) techniques for the classification of TV commercials was proposed in [9]. The output transcripts generated by these technologies normally produce a small number of keywords. These keywords are used to perform searching which results in semantically relevant information from the web. After that, retrieved information from the web was used to develop a feature vector based on text, and the commercial classification was performed based on text. The experimental results show that combining external resources improved the classification accuracy and helped avoiding speech recognition issues in TV commercial videos. A methodology based on subtitle detection to detect TV commercials in videos was presented in [10]. Many constraints were used to distinguish subtitles and other text that appeared in a frame. After detecting subtitles, a scheme was presented that decides whether a TV commercial exists or not, based on the appearance of subtitles. Then, the genetic algorithm was used to point out the mark-in and mark-out points. The reported precision and recall values were more than 90%. Authors in [11] evaluated an algorithm for the detection of text that normally cannot be detected easily in a video. This algorithm performs text detection in cases like a scene with high texture background. The reported results were 96% and 82% for precision and recall respectively.

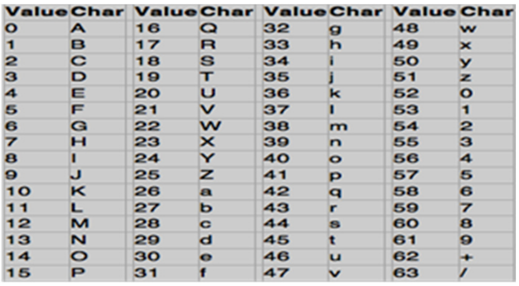
Authors in [12] introduced a framework for the automatic semantic annotation of unconstrained videos. The framework helps to minimize the semantic gap, i.e. the difference between the low-level visual information and the corresponding human perception. It was also proposed that integrating visual similarity matching with common sense semantic relationships is a highly effective approach to automated video annotation. Authors in [13] investigated the problem of automatic management of videos taking into account syntactic and semantic features for TV NEWS programs. Authors in [14] presented an innovative solution of a smart emotional system for impaired people. They aimed to accompany the cognitive information contained in a movie, with the affective content for emotion recognition. The author in [15] described the TV stream as a collection of programs (P) and breaks (advertisements) (B) [15]. A TV stream is a collection of program segments and break segments where C.1, C.2, and C.3 represent the different type of program segments having an aim of information or entertainment (Figure 2 in [15]). Each category of program segments is separated by breaks represented through C.

II. METHODOLOGY

In this research work, Base64 Encoding and Decoding technique has been used to detect TV advertisements. The following section defines the technique first and the steps carried out to perform the advertisement detection.

A. Base64 Encoding and Decoding

It is an encoding style that uses 64 numbers of characters to represent the binary data. Initially, data were sent over the network, i.e. via email through only a set of text characters but after the advancement in multimedia technology, it was expected to send other data types, i.e. images (binary data), attachments, and executables as well. Thus the need emerged for some encoding techniques for binary data to transport over the network. The encoding of binary data is needed to avoid the problem caused by the existence of null characters in binary data. The encoding Table of Base64 is shown in Figure 2.



Value	Char	Value	Char	Value	Char	Value	Char
0	A	16	Q	32	g	48	w
1	B	17	R	33	h	49	x
2	C	18	S	34	i	50	y
3	D	19	T	35	j	51	z
4	E	20	U	36	k	52	0
5	F	21	V	37	l	53	1
6	G	22	W	38	m	54	2
7	H	23	X	39	n	55	3
8	I	24	Y	40	o	56	4
9	J	25	Z	41	p	57	5
10	K	26	a	42	q	58	6
11	L	27	b	43	r	59	7
12	M	28	c	44	s	60	8
13	N	29	d	45	t	61	9
14	O	30	e	46	u	62	+
15	P	31	f	47	v	63	/

Fig. 2. Base64 encoding and decoding values.

B. The Followed Approach for TV Advertisement Detection.

The first technique that we tried was to compare and count the times the frames of an ad appear in the main broadcast video. Drawbacks such as high storage utilization, due to the splitting of videos into frames and more time, occurrred. The second technique that we tried to use was to detect the advertisements during a live broadcast. The main problem that occurred was more time utilization although less storage was needed. The third technique was to encode all the pixels of all the frames, of all the videos, and then perform the comparison. This required more storage utilization as the size of the produced file was large, and more time in performing the comparison. The final approach chosen for this project was to encode only selected pixels belonging to sub-regions of the frames of the videos and then perform the comparison in order to reduce the computation cost. The time and the storage capacity required in this approach were comparatively lesser than the approaches mentioned above. The proposed methodology is shown in Figure 3. In Figure 3, the "Main File" is the video file that contains both advertisements and other program segments whereas the "Ad File" denotes the separate categories of advertisements that may exist in the input video file. Both advertisement files and the Main File are converted into encoded text files using the Base64 encoding. The resulting text files are compared line by line to find the match of any advertisement file in the Main File. Each line of the text file corresponds to a set of pixel positions in a frame. These pixel positions are picked from the middle region of the frame only.

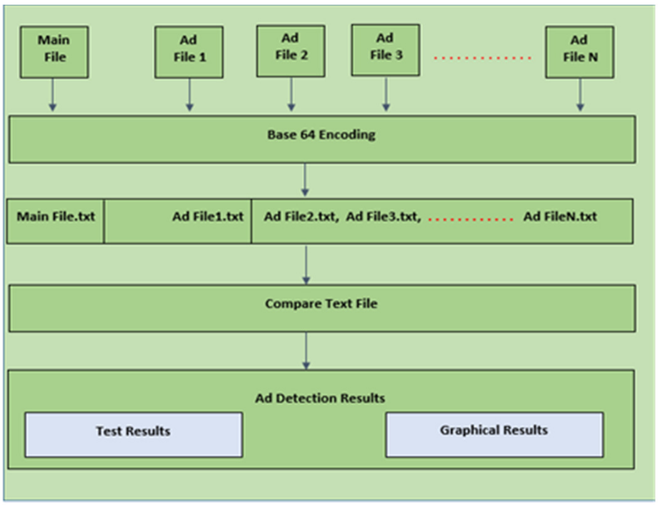


Fig. 3. The proposed approach for TV advertisement detection.

The obtained results were further presented in both textual and graphical forms. When the user selects the advertisements to be detected and the video they are to be detected in, instead of encoding each pixel of every frame, only selected pixels are encoded and saved into a text file and the file size remains small even for long videos. When the user hits the detect button, each pixel in each advertisement file is compared to each pixel in the Main video File until a match is found. When each pixel in the advertisement file is found in the main video file in sequence, the count of the comparison for that specific advertisement increases. Once the comparison is done for all selected ads, a table consisting of the ads against the time (seconds) they have appeared in the main video is shown to the user, with a graphical donut chart representation of the obtained results. Details are given in the result section of this paper. The steps of the algorithm are:

Step 01: Break the input videos (Main video and ad videos) into frames.

Step 02: Select the set of pixel values from each frame belonging to the sub-region of the selected frames.

Step 03: Convert the set of pixel values into base64 type.

Step 04: Compare all pixel values of an advertisement frame with the frame of the Main video

If the pixel points match then

Take the frame as an Ad frame and move to the next frame

Repeat the process until the Ad frame matches with the frames of the Main Video

Return the Total Number of Frames matched

Else

Select the next frame in the sequence

C. End-User System

A system has been developed for the end-user to evaluate the performance of the devised algorithm. The developed system is mainly comprised of four steps.

Step 01: In this step, the end-user is provided with the Ad Detection screen. The Select Main File button allows the user to select the video file which contains advertisements and non-advertisement content. The Select Ad Files button allows the user to either select a single advertisement segment or set of advertisement videos to be searched in the Main File (Figure 4).

Step 02: In this step, after providing the Main File and the Ad file, the end-user is provided with the Upload button. The button converts the videos into frames and then converts these frames into Base64 encoded strings (Figure 5).

Step 03: In this step each file is converted into a Base64 encoded text string (Figure 6).

Step 04: The screen in Figure 7 appears when the binary images are successfully converted into text files using the Base64 encoding. The end-user is provided with the Detect button to start the advertisement detection process with the developed algorithm. The system will produce two different forms of detection results, one in text form and another in graphical form.

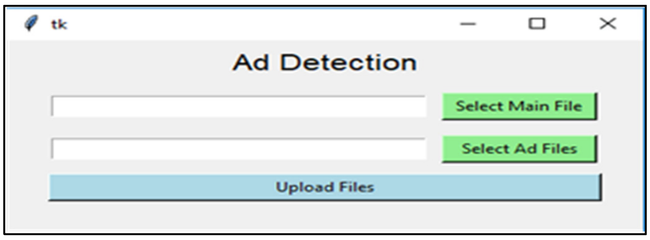
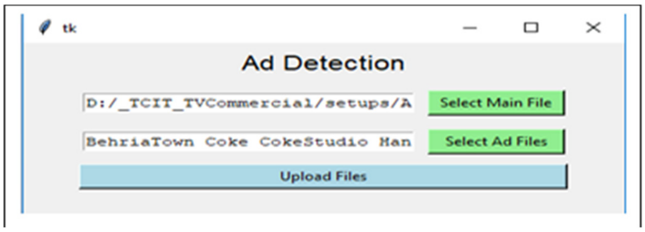


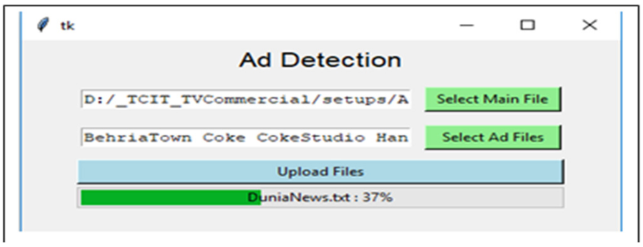
Fig. 4. Main screen of the system.



D:/_TCIT_TVCommercial/setups/A

BehriaTown Coke CokeStudio Han

Fig. 5. Selecting the main video file and the advertisement file.

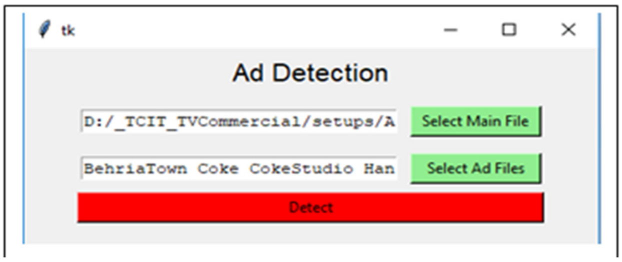


D:/_TCIT_TVCommercial/setups/A

BehriaTown Coke CokeStudio Han

DuniaNews.txt : 37%

Fig. 6. Converting the main video file and the advertisement file into text files.



D:/_TCIT_TVCommercial/setups/A

BehriaTown Coke CokeStudio Han

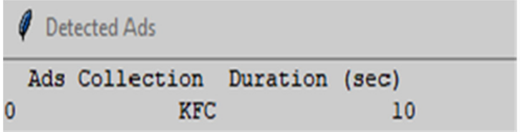
Fig. 7. After the successful conversion, the Detect button appears.

D. Dataset

The Video dataset was recorded from National TV channel transmissions, including ARY Digital 1, and Hum TV. The test data composed of three segments of transmission that were recorded from ARY Digital and Hum TV. The two segments from ARY had 2.5 hours length and the HUM TV segment had a length of 1 hour 40 minutes. All three segments were recorded from a local cable provider and were sampled at 25 frames per second.

III. RESULTS AND DISCUSSION

The detection results were obtained in two ways: single ad and multiple ad detection as shown in Figures 8-11. Figure 8 shows the text file output of the results and Figure 9 the produced graphical representation. They shows the detection of a KFC advertisement which lasts for 10 seconds in a given 152 second video file.



```
 Ads Collection  Duration (sec)
0          KFC              10
```

Fig. 8. Single ad detection text results.

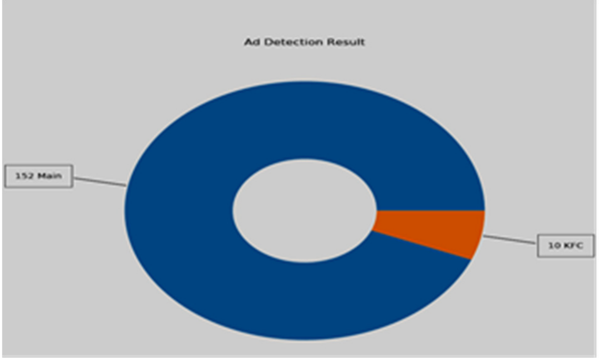
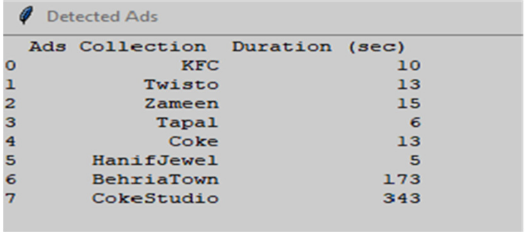


Fig. 9. Single ad detection graphical results.

Figure 10 shows the text file output of the results of multiple ad detection and Figure 12 their graphical representation. They show the number of detected advertisements. Moreover, the name of each detected advertisement is given in the adjacent column along with their duration in seconds in another adjacent column. To the best of our knowledge, this is the first time the Base64 encoding technique is used for TV advertisements detection. The technique results in TV advertisement detection with 60% precision. On the other hand, it suffers from the problem that the results rely heavily on the size of the video file. One more factor that influences advertisement detection is video quality. The proposed algorithm does not provide good ad detection results on grainy and poor-quality videos.



```
  Ads Collection  Duration (sec)
0            KFC              10
1         Twisto              13
2         Zameen              15
3          Tapal               6
4           Coke              13
5     HanifJewel               5
6     BehriaTown             173
7    CokeStudio              343
```

Fig. 10. Multiple ad detection text results.

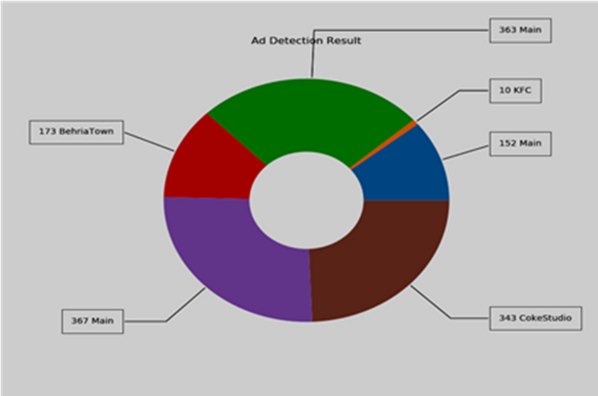


Fig. 11. Multiple ad detection graphical results.

IV. CONCLUSION

Many attempts have been made to find suitable features for TV commercial detection. It had been observed that some features that have been used by other researchers, such as silence, blank frame, and absence of TV channel logo during advertisement breaks have no use in Pakistani TV transmissions. It had been noticed that silence may appear not only at the beginning and the end of an advertisement but also at several intervals during the shot transitions in an advertisement. The blank frame that exists in TV transmission in Australia and European countries has been used by some researchers to detect the advertisement breaks, but Pakistani TV transmission does not contain such information due to the lack of media industry legislation in Pakistan. Another technique that is widely used in American TV channels is to detect the TV advertisement by finding the absence of the TV channel logo during the advertisements. Also, this is not true in Pakistani TV channel transmission, since the channel logo remains during the advertisement. In this paper, TV advertisement detection with Base64 encoding of the video segments was used. This saves computations as the ad detection is performed on text rather than on each pixel of an image. It also saves memory as only a very small set of pixels is used each time. The results can be improved by further reducing the sub-regions of images while converting them into text files. This can save memory and further reduce the computation cost but it also causes the loss of important information needed for ad detection.

REFERENCES

[1] 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.

[2] P. Matlani and M. Shrivastava, "An Efficient Algorithm Proposed For Smoke Detection in Video Using Hybrid Feature Selection Techniques," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3939–3944, Apr. 2019, https://doi.org/10.48084/etasr.2571.

[3] C. Colombo, A. D. Bimbo, and P. Pala, "Retrieval of Commercials by Semantic Content: The Semiotic Perspective," *Multimedia Tools and Applications*, vol. 13, no. 1, pp. 93–118, Jan. 2001, https://doi.org/10.1023/A:1009681324605.

[4] N. Dimitrova *et al.*, "Real time commercial detection using MPEG features," in *Proceedings of the 9th International Conference on Information Processing and Management of Uncertainty in Knowlwdge-based Systems (IPMU2002)*, 2002.

[5] X.-S. Hua, L. Lu, and H.-J. Zhang, "Robust learning-based TV commercial detection," in *2005 IEEE International Conference on Multimedia and Expo*, Amsterdam, Netherlands, Jul. 2005, https://doi.org/10.1109/ICME.2005.1521382.

[6] L.-Y. Duan, J. Wang, Y. Zheng, J. S. Jin, H. Lu, and C. Xu, "Segmentation, categorization, and identification of commercial clips from TV streams using multimodal analysis," in *Proceedings of the 14th ACM international conference on Multimedia*, New York, NY, USA, Oct. 2006, pp. 201–210, https://doi.org/10.1145/1180639.1180697.

[7] N. Venkatesh, B. Rajeev, and M. G. Chandra, "Novel TV Commercial Detection in Cookery Program Videos - PDF Free Download," in *Proceedings of the World Congress on Engineering and Computer Science 2009*, San Francisco, CA, USA, Oct. 2009, vol. 2.

[8] E. El-Khoury, C. Sénac, and P. Joly, "Unsupervised Segmentation Methods of TV Contents," *International Journal of Digital Multimedia Broadcasting*, vol. 2010, Jun. 2010, Art. no. e539796, https://doi.org/10.1155/2010/539796.

[9] Y. Zheng, L. Duan, Q. Tian, and J. S. Jin, "TV Commercial Classification by using Multi-Modal Textual Information," in *2006 IEEE International Conference on Multimedia and Expo*, Toronto, Canada, Jul. 2006, pp. 497–500, https://doi.org/10.1109/ICME.2006.262434.

[10] Y.-P. Huang, L.-W. Hsu, and F.-E. Sandnes, "An Intelligent Subtitle Detection Model for Locating Television Commercials," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 2, pp. 485–492, Apr. 2007, https://doi.org/10.1109/TSMCB.2006.883428.

[11] L. Meng, Y. Cai, M. Wang, and Y. Li, "TV Commercial Detection Based on Shot Change and Text Extraction," in *2009 2nd International Congress on Image and Signal Processing*, Tianjin, China, Oct. 2009, https://doi.org/10.1109/CISP.2009.5302320.

[12] A. Altadmri and A. Ahmed, "A framework for automatic semantic video annotation," *Multimedia Tools and Applications*, vol. 72, no. 2, pp. 1167–1191, Sep. 2014, https://doi.org/10.1007/s11042-013-1363-6.

[13] J. Wang, M. Xu, H. Lu, and I. Burnett, "ActiveAd: A novel framework of linking ad videos to online products," *Neurocomputing*, vol. 185, pp. 82–92, Apr. 2016, https://doi.org/10.1016/j.neucom.2015.12.038.

[14] D. Affi, J. Dumoulin, M. Bertini, E. Mugellini, O. Abou Khaled, and A. Del Bimbo, "SensiTV: Smart EmotioNal System for Impaired People's TV," in *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, New York, NY, USA, Jun. 2015, pp. 125–130, https://doi.org/10.1145/2745197.2755512.

[15] Z. A. A. Ibrahim, "TV Stream Table of Content: A New Level in the Hierarchical Video Representation," *Journal of Computer Sciences and Applications*, vol. 7, no. 1, pp. 1–9, Dec. 2018, https://doi.org/10.12691/jcsa-7-1-1.

An Enhanced Electroporator Design for Pulse Generation

Sadasivam Pachamuthu
Electrical and Electronics Engineering
B. S. Abdur Rahman Crescent Institute of Science and Technology
Vandalur, Chennai, India
psvijay92@gmail.com

D. Najumnissa
Electronics and Instrumentation Engineering
B. S. Abdur Rahman Crescent Institute of Science and Technology
Vandalur, Chennai, India
najumnissa.d@crescent.education

Kavitha Sankaranarayanan
AU - KBC Research Centre MIT
Campus of Anna University
Chennai, India
kavitham@yahoo.com

Raja Prabu Ramachandran
Gandhi Institute of Technology and Management
Deemed to be University
Vishakapatnam, Andhra Pradesh, India
director_ar@gitam.edu

Abstract-An electroporator is an instrument used for delivering electrical pulses to a tumor. In this work, an electroporator consisting of three main system blocks, namely High Voltage (HV) source, nanosecond (ns) switching, and pulse generation, was designed, developed, and evaluated to generate high voltage ns pulses to treat tumors. The high-voltage source block was used to convert the 9.6V DC from the battery to a variable HV output and store this charge for later use. The ns switching block contained a MOSFET-based low-side switch which applies short ns pulses to the load. The pulse generation block generates short ns pulses and supplies the adequate current to turn on the MOSFET at a quicker rate aiding the application of these pulses to the load. This process was simulated using PSpice software and the results are presented.

Keywords-electroporator; electrochemotherapy; boost converter microcontroller;

I. INTRODUCTION

Cancer is the second most lethal disease, as it is responsible for almost the 13% of all human deaths. The extra economic burden related to cancer complicates the administration of presently adopted remedy strategies [1]. Reversible electroporation is the process where a short electric field on a cell retains the membrane's ability to recover at a quicker rate, while the cell remains viable. On the other hand, irreversible electroporation is the process where longer and high-intensity pulses are applied breaking down the cell membrane irreversibly [2]. Reversible electroporation can be further improved or optimized [3]. It is also feasible to electroporate using very brief and very strong electric pulses [4]. In this study, a new electroporator to produce nanosecond (ns) high voltage pulses for treating abnormalities of skin cells is designed, developed, and evaluated. The electroporator instrument consists of a pulse generator and an electrode for delivering the pulse to the tumor's location. The pulse generator produces an electric pulse of predetermined magnitude, width, and frequency. The basic principles used in almost all pulse generators are the capacitor charging and discharging mechanisms. Three system blocks make up the miniature high voltage ns pulse: The High-Voltage (HV) source block converts the battery's 9.6V DC into a variable high voltage output, and this charge is stored to use when creating the pulse. The ns switching block is a low-side switch with a MOSFET that sends brief ns pulses to the load. The pulse generation block generates the ns pulses and provides sufficient current to turn the MOSFET on quickly enough to deliver them to the load.

II. MATERIALS AND METHODS

The following materials were used:

- Pulse generator: A logic-based PIC controller was used for creating the ns pulses using a program written in mikroC, converted into hex code, and loaded to the microcontroller.
- Programmable Integrated Circuit (PIC) Controller: The PIC controller featured program memory, SRAM, EEPROM, analog comparators, and Pulse With Modulation (PWM). All these characteristics make it appropriate for more advanced analog to digital applications. Using a PIC as a master controller to generate the PWM and interface with the LCD and the keypad reduced the required components and made the electroporator more compact.
- HV source block: 0-12V DC was converted to 0-1000V DC using the HV module. The 230V AC supply was connected to an adjustable voltage regulator, which fed the HV module. A potentiometer was used to set the regulator voltage around 9VDC and obtain a 1000V output. Two diodes were utilized to safeguard the HV module's output from transient spikes and damage. As the HV module was

---

Corresponding author: Sadasivam Pachamuthu rated at 10W, a 150k resistor was used to limit the current to 6.6mA and not exceed the module's rated output power. When the module was pulsed, it charged a 1uF capacitor which gave energy to the load.

- Ns switch block: The source of the MOSFET was connected to the ground, and the drain was connected to the load's negative side. A fast switching diode was placed across it to reduce ringing on the load. A 150Ω resistor was connected in parallel with the load to aid in measurements using an oscilloscope.
- LCD: An LCD 4×16 display was used (Figure 1).

Fig. 1. 4×16 LCD.

- Keypad: The keypad interfaced with the PIC for obtaining input from the user. A 4×4 keypad was used, having the 0-9 digits and four keys for other operations. The keypad was used to select the parameters and protocols according to the program in the PIC microcontroller.

III. PROPOSED ELECTROPORATOR DESIGN

In this work, a programmable electroporator using the PIC16F877A microcontroller was designed and developed [5]. Figure 2 shows the schematic arrangement of the programmable electroporator. The designed electroporator consists of a microcontroller, a power circuit as an electroporator, and an LCD. The electric supply is given to the electroporator through a boost converter. A logic-based PIC controller is used for programming the pulses using embedded C. The program was written in MikroC Integrated Development Environment (IDE), converted into hex code, and loaded to the microcontroller. The pulses were generated at the output PORTB pin. The microcontroller can store up to 16 programs for different applications, and the user can select which fits its needs. Figure 3 shows the boost converter's control diagram used for generating the high voltage supply for electroporation. The AC input is converted to a high voltage AC supply using the HV transformer and rectifier [6-8]. There are many ways to achieve a 1000V potential. One could be to transform 120VAC and then rectify it to 1000V. This was not chosen because of the large size and weight of the needed transformer and rectifier. This study's motive was to keep the pulser small and portable using a high voltage DC-DC converter module (0-12VDC to 0-1500VDC). The battery voltage was applied to the variable voltage regulator and the high voltage module. A potentiometer was used to adjust the regulator voltage up to 8VDC to obtain an output voltage of around 1000V. Figure 4 shows the booster voltage display and Table I presents a comparison between an existing and the designed electroporator.

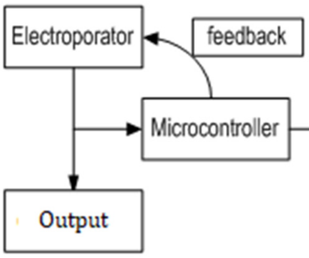


Fig. 2. Schematic diagram of the programmable electroporator.

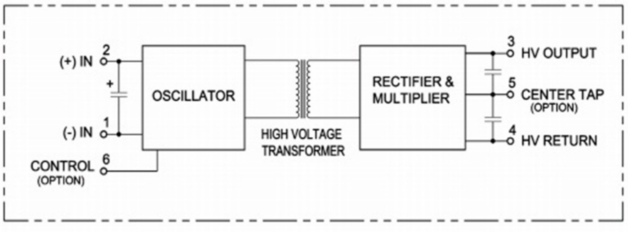


Fig. 3. Boost converter control diagram.

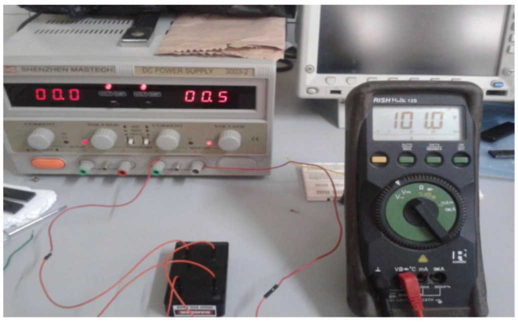


Fig. 4. Boosted voltage displayed in voltmeter.

TABLE I. COMPARISON BETWEEN EXISTING AND DEVELOPED MODELS OF ELECTROPORATOR DESIGN

Existing model ECM 830 [9]	Developed model
Has more components for pulse generation.	Uses a PIC controller
Uses an RC network and flip flops before sending the pulses to the IC driver	Uses an isolator between the pulse and the driver
Costs about 800,000INR	Costs about 50,000INR
Cannot use display and input device without having a separate processor which further increases its cost.	Already uses a controller so LCD and keypad can be easily attached.
Must be changed manually.	Logically controlled so multiple programs can be executed in a single device.

A. Switch Block Description

Pulses having a duration as low as 20ns were required in the design specifications; therefore an extremely fast RF MOSFET was used, as shown in Figure 5. A 5ns rise time 1200V MOSFET was chosen as a switch. A high current driver was used to switch the MOSFET. The 1960pF capacitance of the MOSFET was charged through an initial high current draw using this driver. The 4ns rise time MOSFET driver was used to make the driver act very fast. The battery was used to charge the MOSFET driver with no specific regulations. The Complementary Metal-Oxide-Semiconductor (CMOS) and the Transistor-Transistor Logic (TTL) were the inputs of the MOSFET driver, which were connected to the pulse generation circuit output.

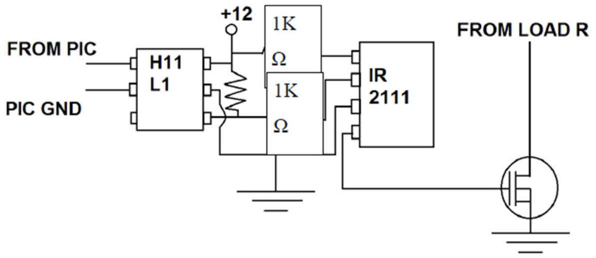


Fig. 5. ns switch design.

Two CMOS digital Integrated Circuits (ICs) were used to construct a ns pulse generator circuit. Those two ICs were re-triggerable monostable multivibrator and D flip flop. The pulses were produced by the re-triggerable monostable multivibrator, determined by the value of one external capacitor and resistor. The multivibrator was a dual IC, so using one multivibrator two pulses can be generated, and the pulses can be shifted by increasing the external resistor of the multivibrator. A variable resistance was connected in series with the external resistor for allowing pulse duration adjustments. These pulses were then used to clock and preset the D flip flop. Separating the pulses by 20ns allows the D flip flop to be clocked high and then preset low 20ns later. The re-triggerable monostable multivibrator needs a clock signal to repeat the pulses. A 555 timer was used to create a CMOS clock of 1.5kHz. The 555 timer was also built with potentiometers to allow adjustment of frequency and duty cycle. All three ICs were powered using a simple 5VDC regulator.

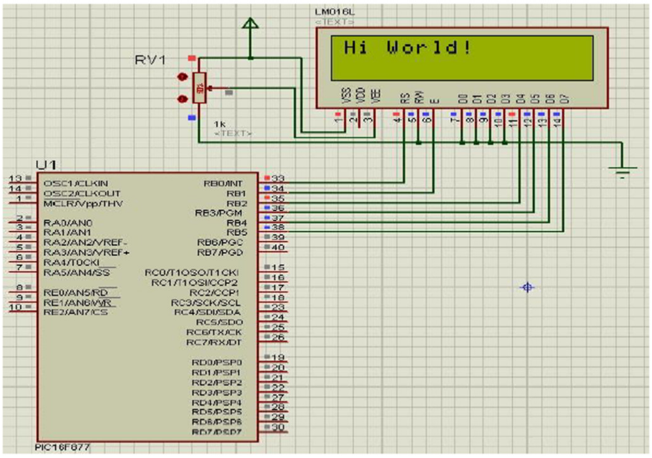


Fig. 6. Interfacing PIC with LCD.

B. *Interfacing PIC with LCD*

The PIC was interfaced with the LCD, as shown in Figure 6, for displaying the output after the pulse. Simple code was used for the interface, as the interfacing of an LCD and a keypad to a PIC is simple. The number of delivered pulses was displayed in the LCD through software.

C. *Hardware Circuit Testing and Modeling of Electroporator*

All hardware components (pulse generator, PIC controller, HV source block, ns switch block, LCD, and keypad) were connected and tested in the laboratory, as shown in Figure 7, and the results were evaluated.

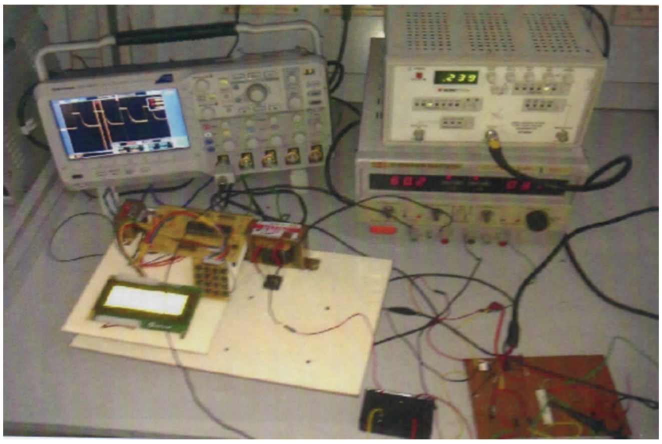


Fig. 7. Testing and modeling of the electroporator.

D. *Steps for the Electroporator Using Various Applications*

- Display the parameters to control the output of the electroporator.
- Choose a parameter, give an in-range value, and press OK.
- Pulses are applied through the electrodes.
- The number of pulses applied will be displayed.
- Similarly, repeat for all the parameters and press OK.

IV. RESULTS AND DISCUSSION

A. *Output Waveforms*

The results depicted in Figures 8-12 show the effect of electric pulses of 50mV magnitude, in both positive and negative, in cancer cell HEP using the patch-clamp technique and obtained a single ion channel output which proves the electroporation effect. This technique provides knowledge on the permeability and the changes that occur while electric pulses are delivered to the cell membrane [10]. The electroporator consists of electrodes delivering high voltage pulses to the affected site, and other electronic components such as power converters to convert the low voltage to high [11, 12].

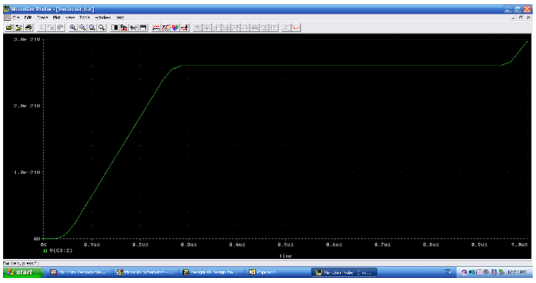


Fig. 8. Transient Voltage graph of the HV block.

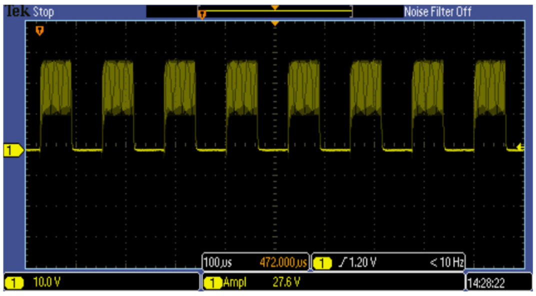


Fig. 9. The obtained output without the capacitor to reduce noise.

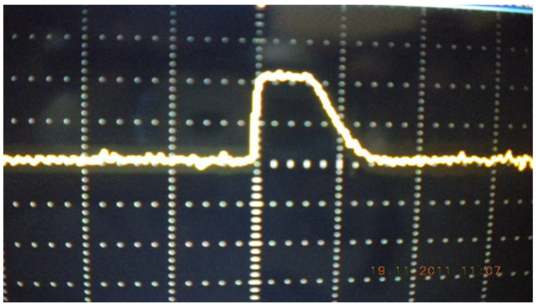


Fig. 10. The output of BTX ECM 830.

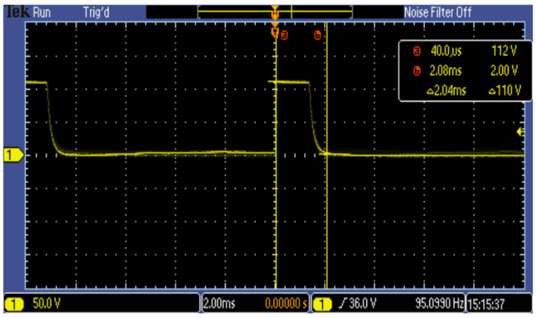


Fig. 11. Single pulse output of RAPKVISS.

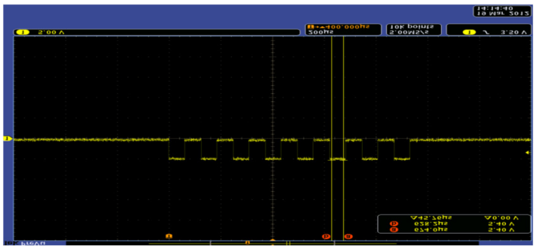


Fig. 12. The output of RAPKVISS for skin cancer.

V. CONCLUSION

Electrochemotherapy is an effective technique that uses high voltage electrical pulses to build up the drug transport across the plasma cell membranes. Moreover, it controls the drug proliferation rate by cancer cell apoptosis, as their morphological properties are identical to that of a human stem cell. In this paper, an electroporator was designed and developed to enhance the drug delivery in electrochemotherapy, using a PIC microcontroller and a high voltage boost converter. The proposed electroporator had an LCD to display the parameters of the pulses delivered to the cancer cells. The effect of electric pulses was analyzed and a single ion channel output that proves the electroporation effect was obtained. The main advantages of the proposed electroporator are its low cost and its controllability, as multiple programs can be executed in a single electroporator. Thus, the proposed electroporator was proved to be more advantageous over existing models and had additional features, like displaying the applied pulses with their parameters and a 4×4 keypad for data input to control its output. Furthermore, it produced a more squarely output than any other electroporator existing to date. Thus, the proposed design can be a cost-effective and efficient pulse generator for electroporation in many applications such as gene transplantation, immunotherapy, and food processing for protection against microbes.

REFERENCES

[1] J. C. Weaver, "Electroporation of cells and tissues," *IEEE Transactions on Plasma Science*, vol. 28, no. 1, pp. 24–33, Feb. 2000, https://doi.org/10.1109/27.842820.

[2] L. M. Mir, O. Tounekti, and S. Orlowski, "Bleomycin: Revival of an old drug," *General Pharmacology: The Vascular System*, vol. 27, no. 5, pp. 745–748, Jul. 1996, https://doi.org/10.1016/0306-3623(95)02101-9.

[3] L. M. Mir, "Therapeutic perspectives of in vivo cell electropermeabilization," *Bioelectrochemistry*, vol. 53, no. 1, pp. 1–10, Jan. 2001, https://doi.org/10.1016/S0302-4598(00)00112-4.

[4] T. B. Napotnik, M. Reberšek, T. Kotnik, E. Lebrasseur, G. Cabodevila, and D. Miklavčič, "Electropermeabilization of endocytotic vesicles in B16 F1 mouse melanoma cells," *Medical & Biological Engineering & Computing*, vol. 48, no. 5, pp. 407–413, May 2010, https://doi.org/10.1007/s11517-010-0599-9.

[5] A. Gothelf, L. M. Mir, and J. Gehl, "Electrochemotherapy: results of cancer treatment using enhanced delivery of bleomycin by electroporation," *Cancer Treatment Reviews*, vol. 29, no. 5, pp. 371–387, Oct. 2003, https://doi.org/10.1016/S0305-7372(03)00073-2.

[6] C. T. S. Ching, T. P. Sun, W. T. Huang, S. H. Huang, C. S. Hsiao, and K. M. Chang, "A circuit design of a low-cost, portable and programmable electroporation device for biomedical applications," *Sensors and Actuators B: Chemical*, vol. 166–167, pp. 292–300, May 2012, https://doi.org/10.1016/j.snb.2012.02.065.

[7] A. Al-Ateeq and A. J. Alateeq, "Soft-Charging Effects on a High Gain DC-to-DC Step-up Converter with PSC Voltage Multipliers," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6323–6329, Oct. 2020, https://doi.org/10.48084/etasr.3773.

[8] M. Taherian, M. Allahbakhshi, E. Farjah, and H. Givi, "A Modular Topology of Marx Generator Using Buck–Boost Converter," *IEEE Transactions on Plasma Science*, vol. 47, no. 1, pp. 549–558, Jan. 2019, https://doi.org/10.1109/TPS.2018.2876474.

[9] BTX Harvard Apparatus, Holliston MA, USA, *ECM 830 Electroporation System User's Manual*, Accessed: Aug. 29, 2021. [Online]. Available: https://www.btxonline.com/media/wysiwyg/tab_content/ECM_830.pdf

[10] J. Cutrera, M. Torrero, K. Shiomitsu, N. Mauldin, and S. Li, "Intratumoral Bleomycin and IL-12 Electrochemogenetherapy for Treating Head and Neck Tumors in Dogs," in *Electroporation Protocols: Preclinical and Clinical Gene Medicine*, S. Li, Ed. Totowa, NJ, USA: Humana Press, 2008, pp. 319–325.

[11] P. Vu, D. T. Anh, and H. D. Chinh, "A Novel Modeling and Control Design of the Current-Fed Dual Active Bridge Converter under DPDPS Modulation," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 7054–7059, Apr. 2021, https://doi.org/10.48084/etasr.4067.

[12] B. Zhao, Q. Song, W. Liu, and Y. Sun, "Overview of Dual-Active-Bridge Isolated Bidirectional DC–DC Converter for High-Frequency-Link Power-Conversion System," *IEEE Transactions on Power Electronics*, vol. 29, no. 8, pp. 4091–4106, Aug. 2014, https://doi.org/10.1109/TPEL.2013.2289913.

Efficiency Improvement of Permanent Magnet BLDC Motors for Electric Vehicles

Dinh Bui Minh
School of Electrical Engineering
Hanoi University of Science and Technology
Vietnam
dinh.buiminh@hust.edu.vn

Vuong Dang Quoc
School of Electrical Engineering
Hanoi University of Science and Technology
Vietnam
vuong.dangquoc@hust.edu.vn

Phuong Nguyen Huy
School of Electrical Engineering
Hanoi University of Science and Technology
Vietnam
phuong.nguyenhuy@hust.edu.vn

Abstract-A permanent magnet Brushless DC (BLDC) motor has been designed with different rotor configurations based on the arrangement of the permanent magnets. Rotor configurations strongly affect the torque and efficiency performance of permanent magnet electric motors. In this paper, different rotor configurations of the permanent magnet BLDC motor with parallel the Halbach array permanent magnet were compared and evaluated. Many applications of electric drives or air-crafts have recently preferred the surface-mounted permanent magnet design due to its ease of construction and maintenance. The finite element technique has been used for the analysis and comparison of different geometry parameters and rotor magnet configurations to improve efficiency and torque performance. A comprehensive design of a three-phase permanent magnet BLDC 35kW motor is presented and simulations were conducted to evaluate its design. The skewing rotor and Halbach magnet array are applied to the permanent surface-mounted magnet on the BLDC motor for eliminating torque ripples. In order to observe the skewing rotor effect, the rotor lamination layers were skewed with different angles and Halbach sinusoidal arrays. The determined skewing angle, the eliminated theoretically cogging torque, and the back electromotive force harmonics were also analyzed.

Keywords-BLDC; permanent magnet; finite element method; Ansys Maxwell; SPEED; magnetic flux density

I. INTRODUCTION

Permanent magnet (PM) Brushless DC (BLDC) motors are widely used due to features such as compactness, low weight, high efficiency, and easy assembly [1, 2]. The reliability of BLDC motor is high since it is easy to mount the PM and has a robust structure. Different rotor configurations are available for the PM BLDC motor, e.g. the surface mounted PM design with the interior/exterior rotor, the interior PM design with buried magnets due to specific strengths and weaknesses [3-5]. Among these, there are the radial-flux motors, the surface mounted types with different magnet arrangement that have been used for electrical drives. This paper introduces a novel design of the BLDC35kW-Z36P12 rotor with high efficiency and low torque ripple. The electromagnetic performance of the PM BLDC with 36 stator slots and 12 rotor poles is discussed in this paper.

II. ELECTROMAGNETIC TORQUE ANALYSIS

The aim is to lower the cost with high overload efficiency and reliability. High efficiency and torque density for electric vehicle applications are the first priorities of this design program. The calculation process of the proposed BLDC motor conducted with the SPEED software is presented in Figure 1. The geometry specifications of the motor used for the analysis are given in Table I.

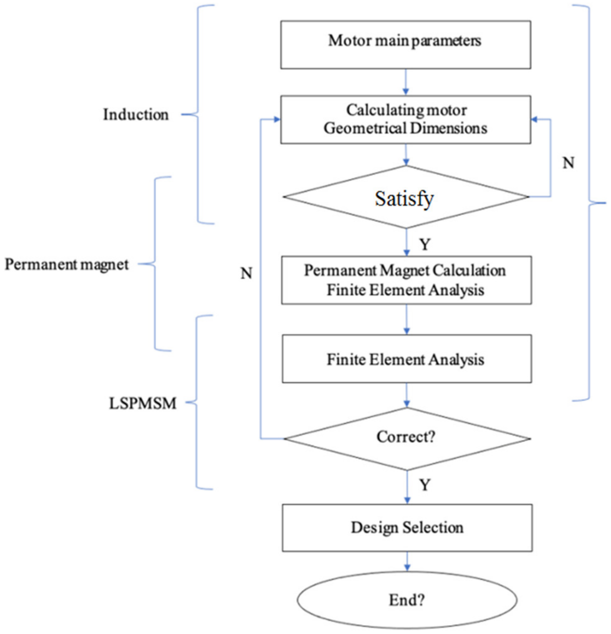


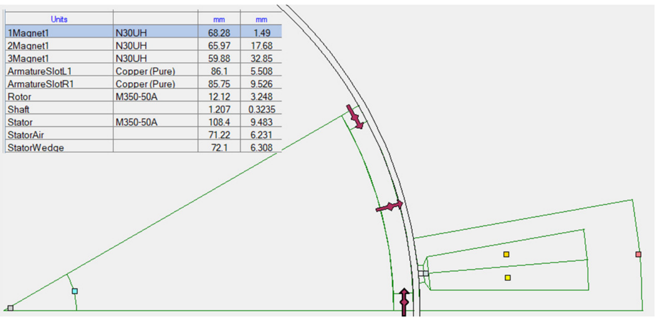
Fig. 1. Design calculation process.

TABLE I. GEOMETRY SPECIFICATIONS OF THE PM BLDC MOTOR

No	Parameters	Unit
1	Outer diameter	218 mm
2	Rotor diameter	142 mm
3	Slot length	112 mm
4	Normal torque	200 Nm
5	Maximum torque	350 Nm
6	Speed	3600 rpm

Corresponding author: Vuong Dang Quoc

The geometry dimensions of the PM BLDC motor are saved in a database in matrix form. When the export command is generated, the drawing process will be executed. The program is developed with the MATLAB DXF library shown in Figure 2. For the drawing circle line, rotating object of stator slots, 2D modeling is implemented to get geometrical formulas of the circle lines. The algorithm needs to satisfy two requirements: the shape of these lines must be similar to the desired curves and the least possible points must be used. Using the minimum number of lines will help the system because it will not have to store a lot of data, which will slow down speed and present difficulties at exporting the drawings. On the other hand, rotation and mirroring are difficult tasks in programming. The strategy is to use a loop function to redraw several times and trigonometric function with angle steps was applied returning good results.



Units		mm	mm
1Magnet1	N30UH	68.28	1.49
2Magnet1	N30UH	65.97	17.68
3Magnet1	N30UH	59.88	32.85
ArmatureSlotL1	Copper (Pure)	86.1	5.508
ArmatureSlotR1	Copper (Pure)	85.75	9.526
Rotor	M350-50A	12.12	3.248
Shaft		1.207	0.3235
Stator	M350-50A	108.4	9.483
StatorAir		71.22	6.231
StatorWedge		72.1	6.308

Fig. 2. MATLAB DXF drawing.

The system will export separate drawings of the motor, the rotor, and the stator. These drawings can be used in several simulation programs and in the design and manufacturing progress. The detail parameters are shown in Table II.

TABLE II. DESIGN PARAMETERS OF PM BLDC MOTOR

Parameters	**Values**	**Unit**
Stator bore	142	mm
Tooth width	7	mm
Slot depth	29	mm
Tooth tip depth	1	
Slot opening	3	
Tooth tip angle	30	degrees
Pole number	12	
Magnet thickness	4	mm
Housing dia	228	mm
Stator lam dia	218	mm
Stator bore	142	mm
Airgap	1	mm

Programs can support the material weight and volume calculation because the power and torque density are also important factors in each motor design. From the geometry parameters, the material weights of two PM BLDC motors were obtained and are shown in Table III. The two PM BLDC motors with the parallel magnet and the Halbach magnet array are presented in Figure 3. The magnet thickness of the conventional design is 4 mm and the electric angle is 150^0 while the Halbach magnet array has a thickness of 3.5 mm and 175 electric embraces. The total weight of a novel design of the Halbach magnet is lower with the same magnet weight.

TABLE III. WEIGHT DETAILS OF THE PM BLDC MOTORS

Parameters	**Material**	**Parallel**	**Halbach**
Stator lam (back iron)	M350-50A	3.947	3.947
S tator lam (yooth)	M350-50A	4.967	4.317
Stator lamination [total]		8.913	8.263
Armature winding [active]	Copper (Pure)	3.23	3.621
Armature EWdg [front]	Copper (Pure)	1.273	0.4767
Armature EWdg [rear]	Copper (Pure)	1.273	0.4767
Armature winding [total]		5.777	4.575
Rotor lam (back iron)	M350-50A	8.811	8.811
Rot inter lam (back iron)		4.02E-05	4.02E-05
Rotor lamination [total]		8.812	8.812
Magnet	N30UH	0.9613	0.9613
Total		25.42	23.57

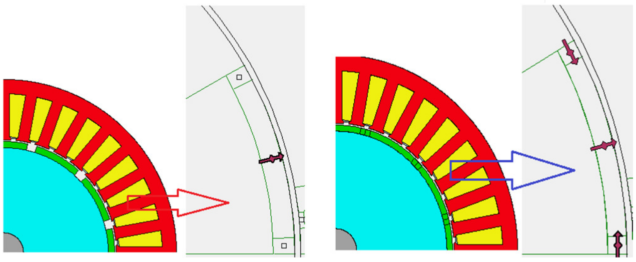


Fig. 3. BLDC motor *p*=12, *Z*=36 parrallel magnet (left) and Halbach magnet (right).

Electro Magnetic Force (EMF) and torque, calculated by a MATLAB analytical coupling program with the Finite Element Method (FEM) are expressed as [11, 12]:

$$\Phi_g = B_g \times \frac{\pi D L_{stk}}{p} \quad (1)$$

$$E_a = \frac{2p}{\pi} w_m T_{ph} \Phi_g \quad (2)$$

$$T_e = \frac{\sqrt{3}}{2} p k_w T_{ph} \Phi_g i_{Lpk} \quad (3)$$

where T_{ph} is the turn per phase term, T_e is the electromagnetic torque and E_a is the back EMF.

The electromagnetic performances based on this design are shown in Table IV. The most important parameter is the efficiency of 93.792%. The efficiency of Halbach magnet and skewed rotor with 5 slices is optimized by the control current of 250A with 200VDC.

TABLE IV. ELECTROMAGNETIC PERFORMANCE COMPARISON

Parameters	**Parallel Magnet**	**Halbach Magnet**	**Unit**
Average torque	226.42	243.41	Nm
Torque ripple	24.635	13.081	Nm
Torque ripple [%]	10.815	5.3608	%
Cogging torque ripple	20.709	4.4287	Nm
Input power	38555	96326	W
Output power	35389	90347	W
Total losses	3165.9	5979.7	W
System efficiency	91.789	93.792	%
Shaft torque	225.3	239.65	Nm

In order to evaluate the maximum torque of the motor, a maximum current is applied to determine when the permanent magnetic is irrecoverable. The maximum torque is 350 Nm at the speed of 1500 rpm with a current of 500 A. Efficiency is calculated based on copper and iron losses. Those losses depend on the stator and rotor teeth dimensions. An efficiency map of the Halbach magnet array with rotor skewed slices is shown in Figure 4.

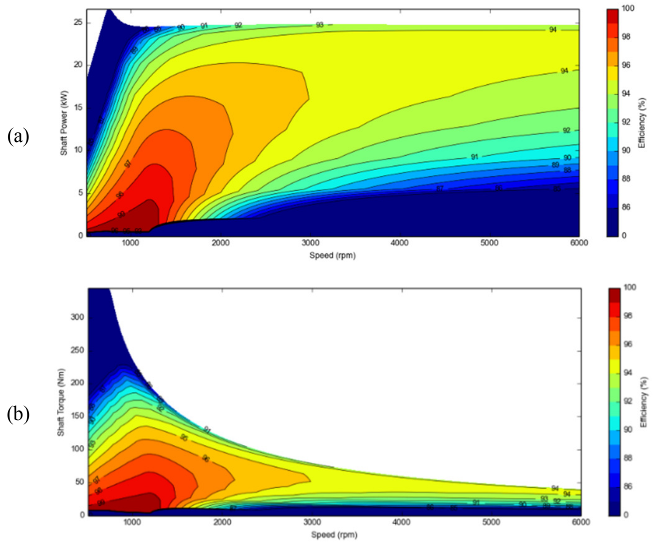


Fig. 4. Efficiency map of (a) Power and (b) Torque.

The torque and speed operation areas have been plotted with different efficiency values. The maximum torque acquired is 350 Nm at the speed of 1000 rpm with low efficiency of 79%. A 2D BLDC motor model is solved and simulated by FEM [13-15]. After meshing the geometry model including the magnetic, the silicon steel, and the insolation materials, the flux density distribution of rotor and stator, is shown in Figure 5.

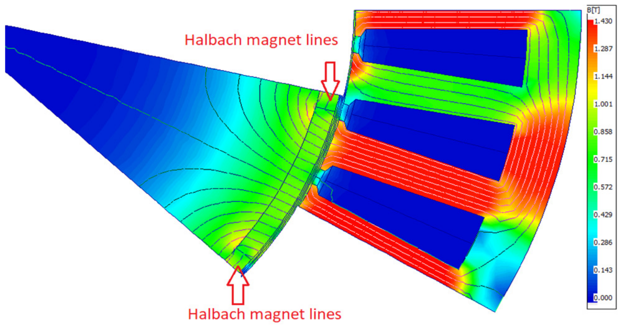
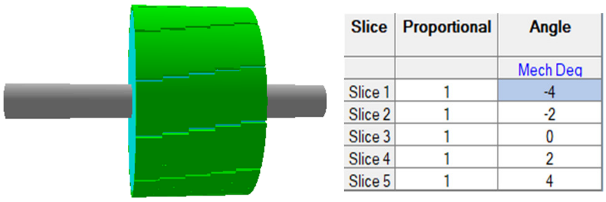


Fig. 5. Flux density results.

The flux density of the air gaps is investigated in one pole. Many steps of rotor position, currents, torque, and flux density were recorded and saved in Matlab files to plot those characteristics. The leakage flux lines are enclosed to air gap areas, which also helps to increase the flux density and reduce the rotor yoke iron loss.

III. TORQUE RIPPLE ANALYSIS OF ROTOR SKEWED SLICES

Skewing slices of rotor are frequently used in PM BLDC motors for eliminating the cogging torque. For the optimum skew angle of those slices, the cogging torque can be eliminated theoretically. The skewed slots for the rotor slices are illustrated in Figure 6.



Slice	Proportional	Angle
		Mech Deg
Slice 1	1	-4
Slice 2	1	-2
Slice 3	1	0
Slice 4	1	2
Slice 5	1	4

Fig. 6. Rotor skewing slices.

The torque ripple results are shown in Figure 7. The cogging torque can be calculated by the stored energy in the air gap. The variation of the co-energy given the cogging torque is expressed as [6-8]:

$$T_c = \frac{dW}{d\theta} \quad (4)$$

where T_c is the cogging torque, $\partial\theta$ is the displacement with mechanical degree, and ∂W is the stored co-energy in the air gap.

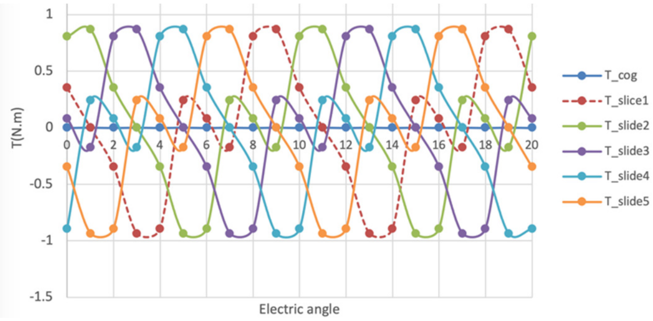


Fig. 7. Torque ripple analysis.

The cogging torque is periodic along the air gap. By using this periodicity feature, the Fourier series of the cogging torque can be obtained [6, 7]:

$$T_{skew}(\theta) = \sum_{i=1}^{\infty} K_{sk}\,.\,T_i.\sin(iC_p\theta_m + \theta_i) \quad (5)$$

where K_{sk} is the skew factor that is 1 for non-skewed motor laminations, C_p is the least common multiple of the number of poles and number of stator slots, T_i is the absolute value of the harmonic i, θ_m is the mechanical angle between the stator and the rotor axis while the motor is rotating, and θ_i represents the phase angle. The skew factor K_{sk} is defined by:

$$K_{sk} = \frac{Sin(\frac{iC\pi\alpha_{sk}}{N_s})}{iC\pi\alpha_{sk}/N_s} \quad (6)$$

where α_{sk} is the skew angle and N_s is the number of the slide. The average values of load torques are nearly the same values for even one slot pitch skewed motor result in terms of average load torque that are coherent with the non-skewed motor model. The relative torque ripples can be defined as:

$$T_{ripple} = \frac{(T_{max} - T_{min})}{T_{avg}} \quad (7)$$

If the skew angle is increased, the torque ripple is reduced along with the average torque, so for optimal torque performance the ratio of magnetic pole/stator slots needs to be increased.

IV. CONCLUSION

In this paper, a comprehensive design of the PM BLDC motor for electric vehicles has been presented. The design was computed by the analytical method, it was optimized by the SPEED software and the electromagnetic characteristics were evaluated by FEM. In particular, the parallel and Halbach PM array rotor with skewed slides have been compared in terms of efficiency, torque, and cogging torque. The torque ripple is minimum with a skew angle of 8 degrees. The best skewing angle is determined by a stator slot angle of 10 degrees. For the electrical drive, the PM surface mounted motor is easy to arrange Halbach and skewing structures.

ACKNOWLEDGEMENT

The authors wish to thank the Technical University of Berlin for providing the means and conditions needed to carry out this work.

REFERENCES

[1] Y. Park, H. Kim, H. Jang, S.-H. Ham, J. Lee, and D.-H. Jung, "Efficiency Improvement of Permanent Magnet BLDC With Halbach Magnet Array for Drone," *IEEE Transactions on Applied Superconductivity*, vol. 30, no. 4, pp. 1–5, Jun. 2020, https://doi.org/10.1109/TASC.2020.2971672.

[2] P. Ji, W. Song, and Y. Yang, "Overview on application of permanent magnet brushless DC motor," *Electrical Machinery Technology*, vol. 40, pp. 32–36, 2003.

[3] B. V. R. Kumar and K. S. Kumar, "Design of a new Dual Rotor Radial Flux BLDC motor with Halbach array magnets for an electric vehicle," in *2016 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Trivandrum, India, Dec. 2016, https://doi.org/10.1109/PEDES.2016.7914552.

[4] A. Vadde and S. Sachin, "Influence of Rotor Design in BLDC Motor for Two -Wheeler Electric Vehicle," in *2021 1st International Conference on Power Electronics and Energy (ICPEE)*, Bhubaneswar, India, Jan. 2021, https://doi.org/10.1109/ICPEE50452.2021.9358520.

[5] V. R. Bommadevara, "Design of a High Power density Halbach BLDC Motor for Electric Vehicle Propulsion.," in *2018 IEEE International Magnetics Conference (INTERMAG)*, Singapore, Apr. 2018, https://doi.org/10.1109/INTMAG.2018.8508741.

[6] L. Dosiek and P. Pillay, "Cogging Torque Reduction in Permanent Magnet Machines," *IEEE Transactions on Industry Applications*, vol. 43, no. 6, pp. 1565–1571, Nov. 2007, https://doi.org/10.1109/TIA.2007.908160.

[7] R. Krishnan, *Switched Reluctance Motor Drives: Modeling, Simulation, Analysis, Design, and Applications*, 1st ed. Boca Raton, FL, USA: CRC Press, 2001.

[8] 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.

[9] 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.

[10] R. Islam, I. Husain, A. Fardoun, and K. McLaughlin, "Permanent-Magnet Synchronous Motor Magnet Designs With Skewing for Torque Ripple and Cogging Torque Reduction," *IEEE Transactions on Industry Applications*, vol. 45, no. 1, pp. 152–160, Jan. 2009, https://doi.org/10.1109/TIA.2008.2009653.

[11] M. D. Bui, S. Schneider, S. Arnaout, and U. Schaefer, "Torque maximization of a high-speed switched reluctance starter in acceleration test," in *2013 15th European Conference on Power Electronics and Applications (EPE)*, Lille, France, Sep. 2013, https://doi.org/10.1109/EPE.2013.6631908.

[12] D. B. Minh, L. D. Hai, T. L. Anh, and V. D. Quoc, "Electromagnetic Torque Analysis of SRM 12/8 by Rotor/Stator Pole Angle," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7187–7190, Jun. 2021, https://doi.org/10.48084/etasr.4168.

[13] 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.

[14] 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 PMaSynRM 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.

[15] V. Q. Dang and C. Geuzaine, "Two-way coupling of thin shell finite element magnetic models via an iterative subproblem method," *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering*, vol. 39, no. 5, pp. 1085–1097, Jan. 2020, https://doi.org/10.1108/COMPEL-01-2020-0035.

The Effect of Nanomaterials on the Properties of Limestone Dust Green Concrete

Sajjad Monther Alsaedy
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
s.naeem1901m@coeng.uobaghdad.edu.iq

Nada Aljalawi
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
nada.aljalawi@coeng.uobaghdad.edu.iq

Abstract- Portland cement is considered the most involved product in environmental pollution. It is responsible for about 10% of global CO_2 emissions [1]. Limestone dust is a by-product of limestone plants and it is produced in thousands of tons annually as waste material. To fulfill sustainability requirements, concrete production is recommended to reduce Portland cement usage with the use of alternative or waste materials. The production of sustainable high strength concrete by using nanomaterials is one of the aims of this study. Limestone dust in 12, 16, and 20% by weight of cement replaced cement in this study. The study was divided into two parts: the first was devoted to the investigation of the best percentage of replacement of waste lime. The second part of the study evaluated the performance of concrete when adding nanomaterials. Three percentages of cement replacement 0.5%, 1%, and 1.5% with nano-Al_2O_3 were used. The most efficient content of hydrated lime used in this study which achieves sustainability and maintains the quality of concrete was (16%). On the other hand, it was found that the best percentage of nano-Al_2O_3 as a partial replacement of cement is 1.5%.

Keywords-green concrete; limestone dust; nano Al_2O_3

I. INTRODUCTION

Concrete is a widely used construction material, with an annual production that exceeds 10 billion tons [2]. Durability, fire resistance, water impermeability, cost efficiency, energy efficiency, and the ability to produce on-site are some of its advantages. However, cement manufacturing, which is the primary binder in concrete, necessitates a significant quantity of natural resources and energy. Approximately 1.5 tons of raw ingredients are required to produce one ton of cement [3]. The need to reduce CO_2 emissions makes the search for alternative binders necessary. The necessity for more economical and environmentally friendly cement materials expanded the interest in other materials which can be used as substitutes to partially replace Ordinary Portland Cement (OPC) [4-6]. The cost of natural resources is increasing constantly, leading to the search for alternatives, such as recycled materials, rice husk ash, sawdust ash, tile powder, wood waste ash, silica fume, fly ash, coal bottom ash, limestone dust, Porcelinite, etc. [7-12]. Also, OPC is related to several diseases [13-17]. Limestone is the most prevalent type of calcium carbonate, often used in cement production. Calcium carbonate ($CaCO_3$), magnesium carbonate ($MgCO_3$), silica (SiO_2), alumina (Al_2O_3), iron oxide (Fe_2O_3), and sulphate (SO_3), are the principal components of limestone, with $CaCO_3$ and $MgCO_3$ being its major components [18]. Cements are primarily created by calcining a mixture of about 75% limestone and 25% clay to produce a calcium silicate clinker, which is then crushed and combined with a tiny amount of gypsum [19]. Nanomaterials are important due to properties such as the high surface to volume ratio. As the surface area per mass of a material increases, a larger amount of the material can come to contact with neighboring particles, therefore nanomaterials have high reactivity [20]. The inclusion of ultrafine nanomaterial particles fills the holes in the concrete microstructure. The nanopowder increases the surface area of the pozzolanic reaction, resulting in a more cementitious product [21, 22]. In the present research, the effects of nano-oxide (alumina) and limestone dust on concrete microstructure and mechanical properties are investigated.

II. MATERIAL CHARACTERIZATION

OPC produced by the Tassloja Cement Factory, confirming to Iraqi Specification No.5/2019 was used in this research [23]. Natural fine sand zone 2 (Table I), which is within the limits of the Iraqi Specification No. 45/1984 [24], was used as fine aggregates. Coarse aggregates with size of 5-14mm were used to prepare concrete samples. The physical and chemical properties of fine and coarse aggregates are shown in Tables II and III respectively. All concrete samples were produced with the same water to cement ratio (w/c) of 0.3. The concrete ingredients were mixed using 1:1.4:1.8 mixing ratio for cement, fine aggregates, and coarse aggregates respectively according to the British method for concrete mix design. Limestone dust was obtained from a limestone quarry in Karbala. It was finely ground in the form of dust, most of it passing the No.100 sieve. The chemical composition of the dust is given in Table IV. It was used to partially replace cement in concrete. Three different percentages (12%, 16%, and 20%) by weight were used to replace cement.

Corresponding author: Sajjad Monther Alsaedy

TABLE I. NATURAL FINE SAND GRADING

Sieve size (mm)	Accumulative passing (%)	Accumulative passing (%) according to the limits of I.Q.S No.45/1984
10	100	100
4.75	100	90-100
2.36	88.7	75-100
1.18	66.9	55-90
0.6	51.8	35-59
0.3	18.8	8-30
0.15	3.4	0-10

TABLE II. PHYSICAL AND CHEMICAL PROPERTIES OF FINE AGGREGATES

Properties of sand	Test results of sand	Limits of IQS No. 45/1984
Fineness modulus	2.76	-
Specific gravity	2.63	-
Absorption	0.8%	-
SO_3	0.23%	≤ 0.5%
Dry rodded density	1604kg/m³	-

TABLE III. PHYSICAL AND CHEMICAL PROPERTIES OF COARSE AGGREGATES

Properties of gravel	Test results of gravel	Limits of IQS No. 45/1984
Specific gravity	2.652	-
Absorption	0.7%	-
SO_3	0.05%	≤ 0.1%
Dry rodded density	1640kg/m³	-

TABLE IV. CHEMICAL ANALYSIS OF LIMESTONE DUST

Chemical composition	Constituent %
Silicon dioxide (SiO_2)	5.15
Calcium oxide (CAO)	52.1
Aluminium oxide (Al_2O_3)	0.66
Ferric oxide (Fe_2O_3)	0.08
Magnesium oxide (MgO)	0.51
Sulphur oxide (SO_3)	0.0
Loss on ignition	41.49

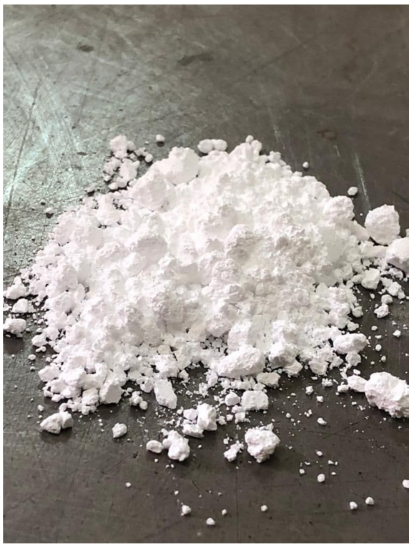


Fig. 1. Nano Al_2O_3 used in the research.

The nanomaterial Al_2O_3 shown in Figure 1 was used as the partial replacement of cement. It was purchased from Skyspring Nanomaterials Ltd and its physical analysis is shown in Table V. Three different percentages of 0.5, 1, and 1.5wt% of cement were used.

TABLE V. PHYSICAL PROPERTIES OF NANO Al_2O_3

Property	Specifications
Appearance	Milky white
Crystal structure and type	Alpha
Purity %	99
Form	Powder
Density	3.5-3.9g/cm^3

III. EXPERIMENTAL PROGRAM

The experimental part consisted of 7 concrete mixes as shown in Table VI (the concrete ingredients were mixed using 1:1.4:1.8 mixing ratio for cement, fine aggregates, and coarse aggregates respectively, according to the British method for concrete mix design).

TABLE VI. MIX DESIGN DETAILS FOR CONCRETE MIXES

Mix symbol	w/b	HRWRA (lt/100kg cement)	Cement (kg/m^3)	Limestone dust (kg/m^3)	Nano Al_2O_3 (%)
MR	0.3	2.5	500	0	0
ML1			440	60	0
ML2			420	80	0
ML3			400	100	0
MN1			417.9	80	0.5
MN2			415.8	80	1
MN3			413.7	80	1.5

IV. SPECIMEN PREPARATION

Specimens in the shape of cube, prism, and cylinder were prepared to test compressive strength, flexural strength, and splitting tensile strength respectively. The steel molds were cleaned and their internal surfaces were lubricated with oil to prevent adhesion with concrete after hardening. The molds were filled with concrete in layers, and each layer was compacted by a vibrating table according to ASTM C-192/C192M [25], which is sufficient to remove any entrapped air. After compaction, the specimens were leveled by hand troweling, then left for 24 hours. The specimens were then removed from the molds, and were cured in water until the time of test. All specimens were cured in laboratory environment conditions according to ACI 308R-01 [26].

V. RESULTS

The tests performed on hardened concrete were the slump test, the compressive strength test, the flexural strength test, the splitting tensile strength test, and the SEM test.

A. Workability (Slump Test)

The workability of concrete mixes was measured directly after mixing, according to ASTM C143 [27]. The results are shown in Table VII.

B. Compressive Strength Test

The compressive strength of the concrete mixes was determined according to British Standard BS 12390-3:2019 [28] on 100×100×100mm cubes. The cubes were tested as shown in Figure 2 at the age of 7, 28, and 90 days. The average result of three tested cubes was taken for each mix. The compressive strength for each cube was calculated as shown in (1):

$$F_{cu} = P/A \quad (1)$$

where F_{cu} is the compressive strength (MPa), A the face area of the cube (mm^2), and P the compressive load at failure (N). The results are shown in Table VIII.

TABLE VII. SLUMP TEST RESULTS

Mix symbol	Slump (mm)
MR	102
ML1	116
ML2	122
ML3	125
MN1	117
MN2	111
MN3	107

Fig. 2. Compressive strength test.

C. Flexural Strength Test

Flexural strength was determined according to ASTM C 293-07 [29] by the center point method. The prism specimens with dimensions of 100×100×400mm were simply supported with 300mm span and were tested at the age of 7 and 28 days. The average of three prisms was taken for each mix. Modulus of rupture was calculated with:

$$F_{r.} = 3PL/2bd^2 \quad (2)$$

where $F_{r.}$ is the flexural strength (MPa), P the maximum applied load indicated by the test machine (N), L the average length of the specimen (mm), b the average width of the specimen (mm), and d the average depth of the specimen (mm). The results are shown in Table VIII.

D. Splitting Tensile Strength Test

Concrete cylindrical specimens with dimensions of 150×300mm were used in this test to determine the splitting tensile strength according to ASTM C 496-04 [30] using a compressive machine. In this test method, a diametric compressive force is applied along the side of a concrete cylindrical specimen until tensile failure occurs. The steel plate of the compressive machine is used to distribute uniformly the load applied along the length of the cylinder. The cylinders were tested at the age of 7 and 28 days, and the average of three cylinders was taken as the final result. Splitting tensile strength was calculated by:

$$\tau = 2P / \pi LD \quad (3)$$

where τ the splitting tensile strength (MPa), P the maximum. applied load (N), L the length of the specimen (mm), and D its diameter (mm). The results are shown in Table VIII.

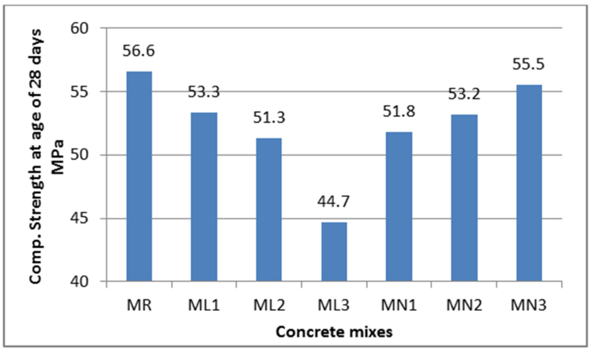


Fig. 3. Compressive strength for all concrete mixes at the age of 28 days.

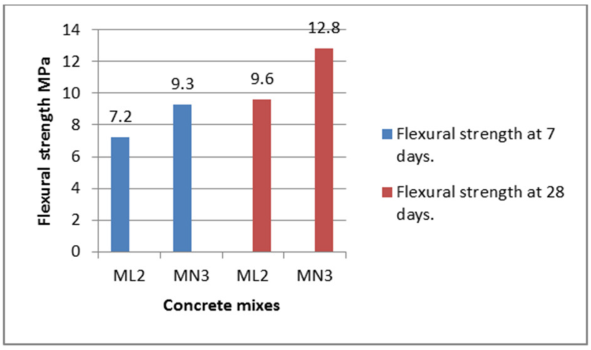


Fig. 4. Flexural strength for ML2 and MN3 mixes at 7 and 28 days.

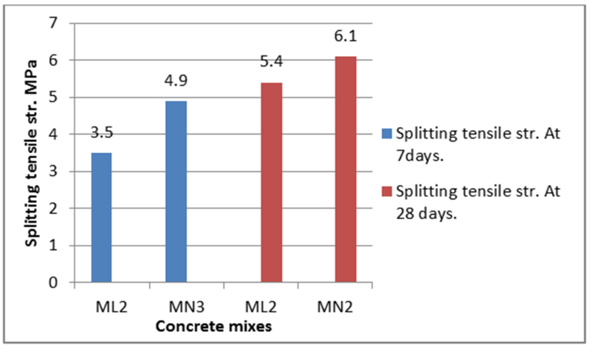


Fig. 5. Splitting tensile strength for ML2 and MN3 at 7 and 28 days.

E. Dry Density

The specimens were tested according to the Iraqi Guide 274 [31]. The average result of three samples was calculated at 7, 28, and 90 days for each mix after picking the cubes out of the curing water and drying their surfaces with a cloth. Density (ρ) is the mass of a unit volume of hardened concrete expressed in kg/m^3. The density was calculated from the specimen mass with volume of 100×100×100 mm obtained after air dry by:

$$\rho = M / V \quad (4)$$

where ρ is the density(kg/m³), M the mass of specimen at the time of test (kg), and V the volume of the specimen calculated from its dimensions (m^3). The results are shown in Table IX.

TABLE VIII. SUMMARY OF COMPRESSIVE, FLEXURAL, AND TENSILE STRENGTH TEST RESULTS

Mix symbol	**Compressive strength (MPa)**			**Flexural strength (MPa)**		**Splitting tensile strength (MPa)**	
	7 days	28 days	90 days	7 days	28 days	7 days	28 days
MR	43.8	56.6	68.5	-	-	-	-
ML1	40.6	53.3	66.0	-	-	-	-
ML2	39.1	51.3	63.6	6.2	8.6	3.5	5.4
ML3	34.2	44.7	60.1	-	-	-	-
MN1	40.4	51.8	64.9	-	-	-	-
MN2	44.0	53.2	65.8	-	-	-	-
MN3	46.1	55.5	69.0	8.3	11.8	4.9	6.1

TABLE IX. DRY DENSITY RESULTS

Mix symbol	**Dry density(kg/m^3)**		
	7 days	28 days	90 days
MR	2511	2525	2536
ML1	2452	2476	2484
ML2	2449	2472	2481
ML3	2437	2459	2470
MN1	2523	2534	2541
MN2	2540	2548	2557
MN3	2556	2569	2580

F. Scanning Electron Microscopy (SEM)

SEM was used to observe the products of nanomaterials. Figures 6 and 7 show the SEM results of a sample from the MLN concrete mix. For this observation, samples were taken from the cracked surface of the specimens of the concrete mixes. In Figures 6, 7 the samples prepared with nano Al_2O_3 show more density at C-S-H gel.

G. Result Summary

When adding limestone dust to the green concrete production, there was an improvement in workability and a slight decrease in the mechanical properties of concrete, but after adding nanomaterials, the properties begun to increase again, so green concrete containing waste material with high properties can be produced. The importance of this research lies in the possibility of replacing a good percentage of cement with waste material while maintaining the same properties. The results of this study are similar with the findings in [21, 32].

VI. CONCLUSIONS

Based on the results obtained from the experimental investigation, the following conclusions can be drawn:

- Within certain ranges of strength and workability criteria, finely ground limestone dust can be recommended as a partial alternative for cement to produce green concrete.
- The dry density of concrete with limestone dust is always lower than that of normal concrete made with OPC only. When adding nano Al_2O_3, the dry density increases with the increase in nano Al_2O_3.
- From all the tested mixes, 16% dust by weight was found to be most satisfactory as it exhibits optimum compressive, tensile, and flexural strength, and a remarkable increase in workability.

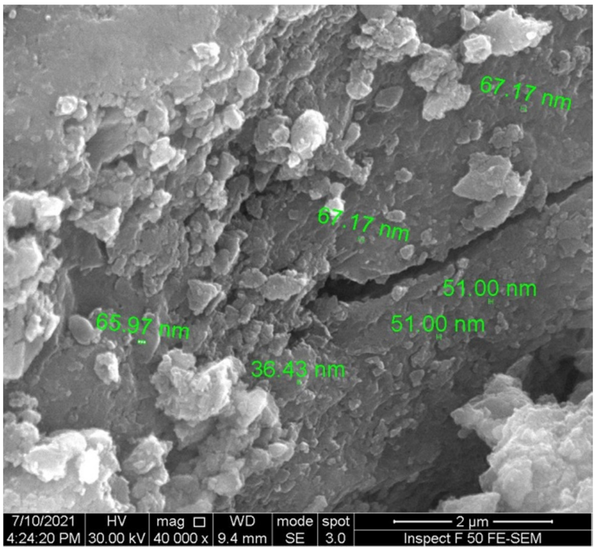


Fig. 6. SEM image of concrete hydrated for 28 days with 1.5% nano-alumina.

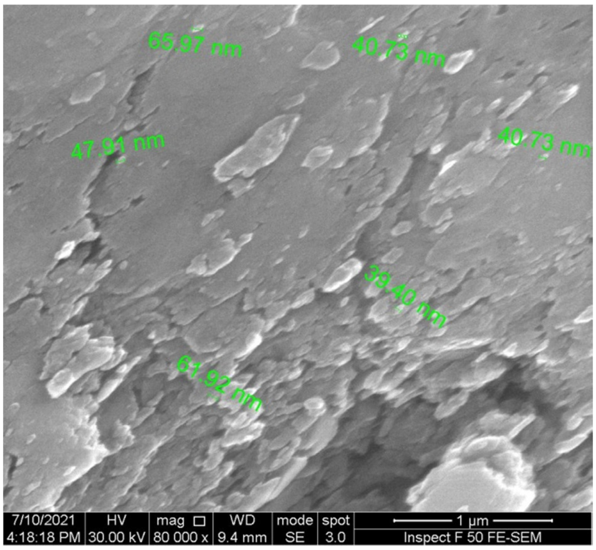


Fig. 7. SEM image of concrete hydrated for 28 days with 1.5% nano-alumina with different scale.

- The compressive strength result of samples with nano alumina showed that the combination of nano-particles and micro-particles of limestone dust increased the compressive strength of concrete samples compared with the compressive strength of concrete with only limestone dust.

This observation indicates that nano alumina has a good filler effect and further distribution of the particles in the remaining voids lead to a more homogeneous concrete matrix which increases the compressive strength of the produced concrete.

- Homogeneity was found to be increased with nano particle addition. This was due to the smaller size of the particles leading to an improvement in the surface characteristics of concrete by smoothening the grain and closing the pores thus increasing concrete strength.
- The addition of nano-alumina with limestone dust to the concrete mixture improved concrete properties and allows the production of high strength green concrete.

REFERENCES

[1] M. J. Memon, A. A. Jhatial, A. Murtaza, M. S. Raza, and K. B. Phulpoto, "Production of eco-friendly concrete incorporating rice husk ash and polypropylene fibres," *Environmental Science and Pollution Research*, vol. 28, no. 29, pp. 39168–39184, Aug. 2021, https://doi.org/10.1007/s11356-021-13418-3.

[2] A. M. Neville, *Properties of Concrete*, 5th ed. Philadelphia,USA: Trans-Atlantic Publications, 2012.

[3] A. M. Rashad, "An exploratory study on high-volume fly ash concrete incorporating silica fume subjected to thermal loads," *Journal of Cleaner Production*, vol. 87, pp. 735–744, Jan. 2015, https://doi.org/10.1016/j.jclepro.2014.09.018.

[4] A. Goyal, A. M. Anwar, H. Kunio, and O. Hidehiko, "Properties of sugarcane bagasse ash and Its Potential as Cement-Pozzolana Binder," in *Twelfth International Colloquium on Structural and Geotechnical Engineering*, Dec. 2007, pp. 10–12.

[5] N. Bheel, A. Memon, S. Meghwar, A. Abro, and I. Shar, "Millet Husk Ash as Environmental Friendly Material in Cement Concrete," in *5th International Conference on Energy, Environment and Sustainable Development*, Jamshoro, Pakistan, Nov. 2018.

[6] N. Bheel, S. Abbasi, S. Meghwar, and F. Shaikh, "Effect of Human hair as Fibers in Cement Concrete," in *International Conference on Sustainable Development in Civil Engineering*, Jamshoro, Pakistan, Nov. 2017.

[7] S. A. Mangi, N. Jamaluddin, M. H. Wan Ibrahim, N. Mohamad, and S. Sohu, "Utilization of Sawdust Ash as Cement Replacement for the Concrete Production: A Review," *Engineering Science and Technology International Research Journal*, vol. 1, no. 3, pp. 11–15, 2017.

[8] N. Bheel, R. A. Abbasi, S. Sohu, S. A. Abbasi, A. W. Abro, and Z. H. Shaikh, "Effect of Tile Powder Used as a Cementitious Material on the Mechanical Properties of Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4596–4599, Oct. 2019, https://doi.org/10.48084/etasr.2994.

[9] M. S. Raza, K. Rai, D. Kumar, and M. Ali, "Experimental Study of Physical, Fresh-State and Strength Parameters of Concrete incorporating Wood Waste Ash as a Cementitious Material," *Journal of Materials and Engineering Structures*, vol. 7, no. 2, pp. 267–276, Jul. 2020.

[10] M. Kepniak, P. Woyciechowski, P. Lukowski, J. Kuziak, and R. Kobylka, "The Durability of Concrete Modified by Waste Limestone Powder in the Chemically Aggressive Environment," *Materials*, vol. 12, no. 10, Jan. 2019, Art. no. 1693, https://doi.org/10.3390/ma12101693.

[11] S. A. Mangi *et al.*, "A Review on Potential use of Coal Bottom Ash as a Supplementary Cementing Material in Sustainable Concrete Construction," *International Journal of Integrated Engineering*, vol. 10, no. 9, pp. 28–36, 2018.

[12] Q. J.-A. Hachim and N. M. Fawzi, "The Effect of Different Types of Aggregate and Additives on the Properties of Self-Compacting Lightweight Concrete," *Journal of Engineering*, vol. 18, no. 8, pp. 875–888, 2012.

[13] M. J. Memon, A. A. Jhatial, Z. A. Rid, T. A. Rind, and A. R. Sandhu, "Marble Powder As Fine Aggregates in Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4105–4107, Jun. 2019, https://doi.org/10.48084/etasr.2698.

[14] N. Bheel, S. L. Meghwar, S. Sohu, A. R. Khoso, A. Kumar, and Z. H. Shaikh, "Experimental Study on Recycled Concrete Aggregates with Rice Husk Ash as Partial Cement Replacement," *Civil Engineering Journal*, vol. 4, no. 10, pp. 2305–2314, Oct. 2018, https://doi.org/10.28991/cej-03091160.

[15] N. Bheel, A. W. Abro, I. A. Shar, A. A. Dayo, S. Shaikh, and Z. H. Shaikh, "Use of Rice Husk Ash as Cementitious Material in Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4209–4212, Jun. 2019, https://doi.org/10.48084/etasr.2746.

[16] N. M. Fawzi and A. Y. E. AL-Awadi, "Enhancing Performance of Self–Compacting Concrete with Internal Curing Using Thermostone Chips," *Journal of Engineering*, vol. 23, no. 7, pp. 1–13, Jun. 2017.

[17] A. A. Jhatial, W. I. Goh, S. Sohu, S. A. Mangi, and A. K. Mastoi, "Preliminary Investigation of Thermal Behavior of Lightweight Foamed Concrete Incorporating Palm Oil Fuel Ash and Eggshell Powder," *Periodica Polytechnica Civil Engineering*, vol. 65, no. 1, pp. 168–180, 2021, https://doi.org/10.3311/PPci.16498.

[18] P. M. Harris, *Limestone & Dolomite : Mineral Resources Consultative Committee*. London, UK: HMSO, 1982.

[19] A. Dinku, "The Influence of Limestone Aggregates and Limestone Fillers on the Properties of Concrete," M.S. thesis, University of Leeds, Leeds, UK, 1991.

[20] P. van Broekhuizen, F. van Broekhuizen, R. Cornelissen, and L. Reijnders, "Use of nanomaterials in the European construction industry and some occupational health aspects thereof," *Journal of Nanoparticle Research*, vol. 13, no. 2, pp. 447–462, Feb. 2011, https://doi.org/10.1007/s11051-010-0195-9.

[21] R. K. Mohammed, A. A. Abdel-Hamead, and F. M. Othman, "Effect of Nano-Alumina on Microstructure and Mechanical properties of Recycled Concrete," *Journal of Engineering and Sustainable Development*, vol. 22, no. 2 (Part–1), pp. 90–103, 2018.

[22] S. Hallad *et al.*, "Effect of Multiwalled Carbon Nanotubes and Nano Aluminum Oxide On Flexural and Compressive Strength of Cement Composites," *International Journal of Advance Research In Science And Engineering*, vol. 3, no. 8, pp. 215–223, Aug. 2014.

[23] *Iraqi Standard No. 5: Portland Cement*. Baghdad, Iraq: Central Organizationfor Standardization and Quality Control, 2019.

[24] *Iraqi Specifications No. 45 for Aggregates of Natural Resources used for Concrete and Construction. Baghdad, Iraq, 1984.*

[25] *ASTM C192/C192M-16a (2017), Standard Practice for Making and Curing Concrete Test Specimens in the Laboratory*. West Conshohocken, PA, USA: ASTM International, 2017.

[26] ACI Committee 308, *308R-01: Guide to Curing Concrete*. Technical Documents, 2001.

[27] *ASTM C143/C143M-20(2020), Standard Test Method for Slump of Hydraulic-Cement Concrete*. West Conshohocken, PA, USA: ASTM International, 2020.

[28] *BS EN 12390-3(2019), Testing hardened concrete. Compressive strength of test specimens*. London, UK: British Standards Institution, 2019.

[29] *ASTM C293/C293M-16(2016), Standard Test Method for Flexural Strength of Concrete*. West Conshohocken, PA, USA: ASTM International, 2016.

[30] *ASTM C496/C496M-17(2017), Standard Test Method for Splitting Tensile Strength of Cylindrical Concrete Specimens*. West Conshohocken, PA, USA: ASTM International, 2017.

[31] *Iraqi Guide No. 274 Methods for determining the density of hardened concrete*. Baghdad, Iraq: Central Organizationfor Standardization and Quality Control, 1992.

[32] H. Binici, H. Kaplan, and S. Yilmaz, "Influence of marble and limestone dusts as additives on some mechanical properties of concrete," *Scientific Research and Essays*, vol. 2, no. 9, pp. 372–379, Sep. 2007, https://doi.org/10.5897/SRE.9000594.

A Study of State Parameters for Road Construction of MSWI Bottom Ash

Thi Kieu Vinh Nguyen
Faculty of Construction
Vinh University
Vinh City, Vietnam
kieuvinhkxd@vinhuni.edu.vn

Thi Thuy Huong Doan
Faculty of Construction
Vinh University
Vinh City, Vietnam
doanhuongkxd@vinhuni.edu.vn

Thi Hao Cao
Faculty of Construction
Vinh University
Vinh City, Vietnam
caohaokxd@gmail.com

Hong Son Pham
Faculty of Construction
Vinh University
Vinh City, Vietnam
hongsonkxd@vinhuni.edu.vn

Abstract-The current work tries to valorize the bottom ash from Municipal Solid Waste Incineration (MSWI). The bottom ash from waste incineration consists of atypical granular materials. They are industrial by-products resulting from the incineration of domestic waste and the way of the considered valorization is road gravel. In this paper, we present the state parameters of bottom ash taken from a recycling company in the North of France. These features can help us evaluate our bottom ash according to the technical guide of realization of embankments and subgrades.

Keywords-bottom ash; compaction characteristics; valorization; incineration; road gravel

I. INTRODUCTION

Municipal solid waste management technologies include landfilling, recycling/recovery, and incineration. In many countries such as France, Sweden, Denmark, and Taiwan, Municipal Solid Waste Incineration (MSWI) for energy recovery represents the most common waste management technology [1]. The incineration reduces the mass and volume of the solid waste dramatically, thus the requirement for landfilling is decreased [2-7]. However, there is still a considerable amount of solid incineration residue, generated after the combustion, typically bottom ash, fly ash, boiler ash, etc. of which bottom ash accounts for about 80% [8]. In the past, MSWI bottom ash was mostly treated by sanitary landfilling. Possibilities other than landfilling have been sought since MSWI started, and reutilization of incinerator bottom ash has already been considered many years ago. In civil engineering, the road field consumes a significant quantity of aggregates [9]. However, the aggregate reserves are not always exploitable for various reasons: they are often inaccessible, they must be integrated into an urban area, in classified or protected sites, and they have too expensive exploitation and risks of environmental impact. In this context, the valorization of the bottom ash in road field is an interesting alternative. Bottom ash is mainly used in civil engineering for constructing embankments, road layers, and parking areas [9, 10]. In France, about 3 million tons of bottom ash are produced annually [11]. The use of bottom ash began during the late '50s. The expansion of its use throughout the country occurred in the late '80s-'90s [12]. This article presents the state parameters by particle size distribution, moisture content, absolute density, Proctor compaction and bearing capacity index tests. These tests results are evaluated according to the technical guide of realization of embankments and subgrades.

II. THE MATERIAL

The MSWI bottom ash used in this study originated from the platform of recycling of the PréFerNord Company located in Fretin, France. PréFerNord recovers "slag" resulting from the combustion of 5 incineration plants. To calibrate the materials, pre-treatment of this bottom ash like sifting, removal of ferrous, and non-ferrous elements, was carried out on site. After, the bottom ash was left to mature for 3 months (Figure 1). The particle sizes range from 0 to 20mm and approache the size range of natural aggregates which are usually used in the road field.

Fig. 1. Bottom ash.

Corresponding author: n Thi Hao Cao

III. THE EXPERIMENTAL METHOD

The particle size distribution was obtained by sieving at 80μm of diameter three samples that were taken after quartering and washing. Moisture content, absolute density and compaction characteristics are the state parameters of bottom ash. The moisture content is equal to the ratio of the mass of water contained in the sample and the dry mass of the sample and is expressed in %. The samples were dried in an oven for 3 days at a temperature of 105°C. The moisture content corresponds to the natural moisture content of the collected bottom ash. The absolute density of bottom ash was determined by using a helium pycnometer of AccuPyc 1330 type. This test is to measure the volume of solid grains from the change of helium pressure by applying the perfect gas law: $PV = nRT$. By knowing the mass of the sample, the absolute density is determined by the ratio between the mass of the solid grain and volume. The measurement was carried out on ground and dried bottom ash.

The compaction is the densification of the soil by applying pressure to it. It contributes in particular to translate or eliminate compaction risk, increase the resistance of soil and slope stability, improve the bearing capacity of road, limit unwanted volume changes, for example, by the action of frost, swelling, or shrinkage [13]. The compaction capacity of the material is assessed through the Normal Proctor and the Modified Proctor tests. Both trials are identical in principle, apart from the difference of the parameters defining the applied compaction energy. The principle of the test is to compact the material at varying moisture content and process energy. For every moisture content, wet and dry masses were determined and compaction characteristics (dry density and optimum moisture content) were determined. The Bearing Capacity Index is the quantity used to assess the capacity of a material to bear directly on its surface the movement of construction equipment. In conjunction with Modified Proctor testing, punching action on compacted specimens was performed to estimate the Bearing Capacity Index.

IV. RESULTS

A. Particle Size Distribution

The results of the particle size distribution presented in Figure 2 show that the bottom ash is characterized by a spread out size distribution (coefficient of uniformity $C_u = 35.5$) along with the too many coarse elements which generate much vacuum (coefficient of curvature $C_c = 2{,}3$). The C_u and C_c are defined by:

$$C_u = \frac{D_{30}^2}{D_{10} * D_{60}} \quad (1)$$

$$C_c = \frac{D_{60}}{D_{10}} \quad (2)$$

where D_x is the diameter of particles for x% of cumulative passing.

Table I represents the useful parameters for the classification of the acquired bottom ash. The tested bottom ash is a granular material with continued grain size distribution and low proportions of non-plastic fine (< 63μm,) and coarse (> 20mm) fractions. Therefore, this may be easily compacted to obtain a higher resistance. Bottom ash could be considered a well-graded material. The obtained remarks are similar to those obtained by [14, 15].

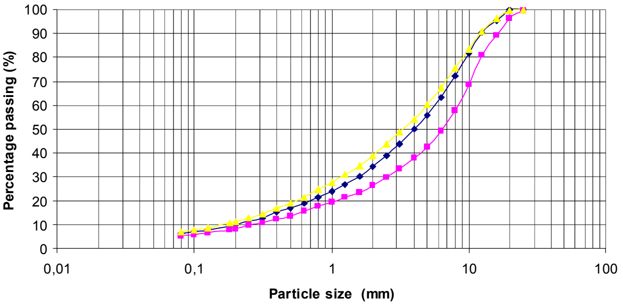


Fig. 2. Particle size distribution by sieving.

TABLE I. PARTICLE SIZE NECESSARY FOR CLASSIFICATION

D_{max}	**20mm**
Passing to 80μm	6.3%
Passing to 2mm	33.2%

B. Moisture Content

Table II represents the moisture content of the studied bottom ash. The main cause of this high moisture content is the influence of rain before samples were taken from storage.

TABLE II. MOISTURE CONTENT VALUES

	Wet mass (g)	**Dry mass (g)**	**Moisture content (%)**
Sample 1	643	542	18.6
Sample 2	697	602	15.8
Sample 3	810	675	20
Average			18.1

C. Absolute Density

The obtained value of the absolute density of the studied bottom ash is about 2.70t/m^3 (Table III). This value is similar to that of quartz-based sand. The obtained density lies in the range of values measured by [16, 17]. This value classifies bottom ash as an aggregate that is lighter than natural ones like sand and gravel. The density is an added benefit that may reduce settlement in use, due to the lower normal stresses caused by the self-weight [18].

TABLE III. ABSOLUTE DENSITY

	Sample 1	Sample 2	Sample 3	Average
Absolute density (g/cm3)	2.70	2.70	2.69	2.70

D. Proctor Compaction

Figures 3 and 4 show the Normal and Modified Proctor compaction curves. Table IV shows the obtained characteristics with Optimal Normal Proctor (OPN) and Optimal Modified

Proctor (OPM). The obtained values of compaction characteristics of bottom ash fall within the variability of the field defined by information note of SETRA on the use of bottom ash in road construction. Based on Proctor test results, bottom ash may be considered as a highly compactable material, which is desirable for preventing future settlements and for increased strength and stability of the layer.

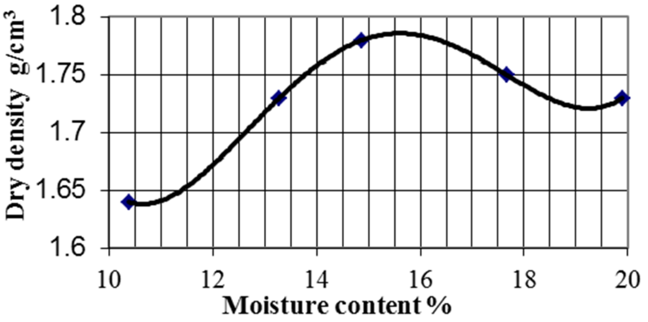


Fig. 3. Normal Proctor compaction curve.

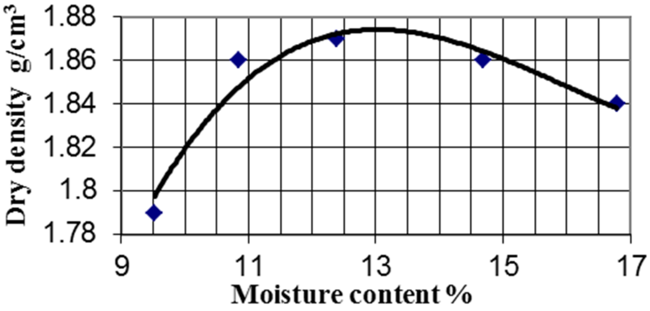


Fig. 4. Modified Proctor compaction curve.

TABLE IV. COMPACTION CHARACTERISTICS

	OPN	OPM
Optimal moisture content (%)	15.0	12.5
Optimal dry density (g/cm3)	1.78	1.87

E. Bearing Capacity Index

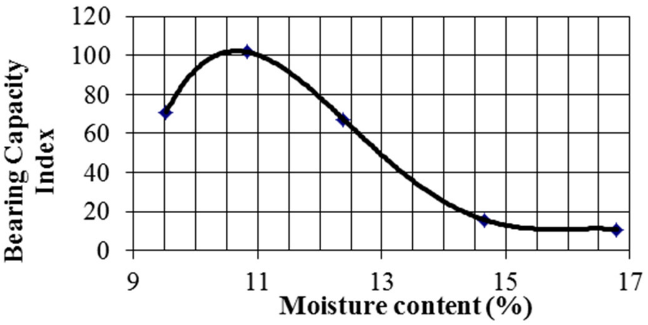


Fig. 5. Bearing Capacity Index curve.

Figure 5 shows the variation of the Bearing Capacity Index with moisture content. According to the recommendations of the French standard (NF P 98 115) [19], to ensure the normal circulation of the machines on a construction site, the desirable values of the Bearing Capacity Index are at least 45 for the base layers and 35 for the foundation layers. This standard also defines the minimum values that must not be less than 35 for the base layers and 25 for the foundation layers. With Bearing Capacity Index of 102 (Figure 5), bottom ash can be considered stable.

V. CONCLUSIONS

The state parameters of the studied bottom ash were analyzed and presented in this paper. From the analysis, the following conclusions can be drawn:

- The bottom ash is characterized by spread out size distribution. Bottom ash can be considered a well-graded material.
- The main cause of its high moisture content is the influence of rain.
- Bottom ash is lighter than natural ashes like sand and gravel. The density is an added benefit that may reduce settlement in use, due to the lower normal stresses caused by the self-weight.
- The obtained values of compaction characteristics of bottom ash fall within the variability of the field defined by information note of SETRA on the use of bottom ash in road construction. Based on Proctor test results, bottom ash may be considered as a highly compactable material, which is desirable in order to prevent future settlements and to increase the strength and the stability of the layer.
- With Bearing Capacity Index of 102, bottom ash can be considered stable.

REFERENCES

[1] C.-L. Lin, M.-C. Weng, and C.-H. Chang, "Effect of incinerator bottom-ash composition on the mechanical behavior of backfill material," *Journal of Environmental Management*, vol. 113, pp. 377–382, Dec. 2012, https://doi.org/10.1016/j.jenvman.2012.09.013.

[2] F. Becquart, "Comportement mécanique au triaxial d'un mâchefer d'incinération d'ordures ménagères," *European Journal of Environmental and Civil Engineering*, vol. 12, no. 6, pp. 673–686, Jun. 2008, https://doi.org/10.1080/19648189.2008.9693038.

[3] N. H. Le, N. E. Abriak, C. Binetruy, M. Benzerzour, and S. Chaki, "The study of behavior of bottom ash under homogeneous stresses. Determination of parameters for Nova behavior models," in *Euromediterranean Symposium on Advances in Geomaterials and Structures*, Djerba, Tunisia, 2010, pp. 653–660.

[4] L. Ngoc, N.-E. Abriak, C. Binetruy, M. Benzerzour, I. Shahrour, and I. Shahrour, "Finite element modeling of the mechanical behavior of municipal solid waste incineration bottom ash with the Mohr-Coulomb model," *International Journal of Environmental, Chemical, Ecological, Geological and Geophysical Engineering*, vol. 9, no. 12, pp. 1315–1321, Dec. 2015.

[5] 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.

[6] A. A. Mahessar, S. Qureshi, A. L. Qureshi, K. Ansari, and G. H. Dars, "Impact of the Effluents of Hyderabad City, Tando Muhammad Khan, and Matli on Phuleli Canal Water," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5281–5287, Feb. 2020, https://doi.org/10.48084/etasr.3269.

[7] N. Radwan and S. A. Mangi, "Municipal Solid Waste Management Practices and Opportunities in Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4516–4519, Aug. 2019, https://doi.org/10.48084/etasr.2870.

[8] J. M. Chimenos, M. Segarra, M. A. Fernández, and F. Espiell, "Characterization of the bottom ash in municipal solid waste incinerator," *Journal of Hazardous Materials*, vol. 64, no. 3, pp. 211–222, Feb. 1999, https://doi.org/10.1016/S0304-3894(98)00246-5.

[9] ADEME, BRGM, *Mâchefers d'incinérations des ordures ménagères - États de l'art et perspectives: États de l'art et perspectives*. Paris, France: Dunod, 2008.

[10] R. Forteza, M. Far, C. Seguí, and V. Cerdá, "Characterization of bottom ash in municipal solid waste incinerators for its use in road base," *Waste Management*, vol. 24, no. 9, pp. 899–909, Jan. 2004, https://doi.org/10.1016/j.wasman.2004.07.004.

[11] *ITOM: Les installations de traitement des ordures ménagères en France*. Paris, France: ADEME, 2008.

[12] R. Badreddine and D. François, "Assessment of the PCDD/F fate from MSWI residue used in road construction in France," *Chemosphere*, vol. 74, no. 3, pp. 363–369, Jan. 2009, https://doi.org/10.1016/j.chemosphere.2008.09.028.

[13] F. Bernard and N.-E. Abriak, "Etude physique, géotechnique et mécanique d'un mâchefer d'incinération d'ordures ménagères," *Journal of Catalytic Materials and Environment*, vol. 1, pp. 9–18, Jan. 2003.

[14] M. Izquierdo, Á. López-Soler, E. V. Ramonich, M. Barra, and X. Querol, "Characterisation of bottom ash from municipal solid waste incineration in Catalonia," *Journal of Chemical Technology & Biotechnology*, vol. 77, no. 5, pp. 576–583, 2002, https://doi.org/10.1002/jctb.605.

[15] I. Maria, E. Vázquez, X. Querol, M. Barra Bizinotto, L. Ángel, and P. Felicià, "Use of bottom ash from municipal solid waste incineration as road material," in *International Ash Utilization Symposium*, 2001.

[16] C. J. Lynn, G. S. Ghataora, and R. K. Dhir Obe, "Municipal incinerated bottom ash (MIBA) characteristics and potential for use in road pavements," *International Journal of Pavement Research and Technology*, vol. 10, no. 2, pp. 185–201, Mar. 2017, https://doi.org/10.1016/j.ijprt.2016.12.003.

[17] B. Muhunthan, R. Taha, and J. Said, "Geotechnical engineering properties of incinerator ash mixes," *Journal of the Air & Waste Management Association (1995)*, vol. 54, no. 8, pp. 985–991, Aug. 2004, https://doi.org/10.1080/10473289.2004.10470959.

[18] M. Arm, "Variation in deformation properties of processed MSWI bottom ash: results from triaxial tests," *Waste Management*, vol. 24, no. 10, pp. 1035–1042, Jan. 2004, https://doi.org/10.1016/j.wasman.2004.07.013.

[19] M. Izquierdo, X. Querol, and E. Vazquez, "Procedural uncertainties of Proctor compaction tests applied on MSWI bottom ash," *Journal of Hazardous Materials*, vol. 186, no. 2, pp. 1639–1644, Feb. 2011, https://doi.org/10.1016/j.jhazmat.2010.12.045.

Investigation of the Effects of Distributed Generation on Protection Coordination of a Power System

Ayesha Tariq
Department of Electrical Engineering
NED University of Engineering and Technology
Karachi, Sindh, Pakistan
ayesha_0405@hotmail.com

Krishan Lal Khatri
Department of Electrical Engineering
NED University of Engineering and Technology
Karachi, Sindh, Pakistan
engrkrishan@yahoo.com

Muhammad Ibrar ul Haque
Department of Electrical Engineering
Sir Syed University of Engineering and Technology
Karachi, Sindh, Pakistan
mihaque@ssuet.edu.pk

Muhammad Amir Raza
Department of Electrical Engineering
NED University of Engineering and Technology
Karachi, Sindh, Pakistan
amir.eed.neduet@gmail.com

Sadia Ahmed
Department of Electrical Engineering
Sir Syed University of Engineering and Technology
Karachi, Sindh, Pakistan
sadiaa@ssuet.edu.pk

Muhammad Muzammil
Department of Electrical Engineering
Sir Syed University of Engineering and Technology
Karachi, Sindh, Pakistan
muzammil@ssuet.edu.pk

Abstract-The rapid increase of the electrical power demand gave rise to many challenging situations for power system control engineers as the transmission lines are operating at their maximum capacity in most developing economies. To solve this, Distributed Generation (DG), i.e. the generation of electrical power in a distribution network that provides clean energy, is gaining popularity. There are several challenges the protection of distribution networks faces after DG installation, such as variations in short circuit levels, protection blinding, reverse power flow, protection coordination, change in fault impedance, recloser-fuse coordination, selectivity, unsynchronized reclosing, false tripping, etc. In this paper, an IEEE 13-Bus System Radial Distribution System is simulated using Electrical Transient Analyzer Program (ETAP), various scenarios of DG placement are considered, their impact on the protection system is analyzed, and different techniques are proposed to minimize the effect on protection coordination. The use of directional relays, current limiting reactors, and small magnitude DGs is tested and analyzed. The way this effect varies by changing the location of DG is also analyzed.

Keywords-distributed generation; protection coordination; relays; radial distribution network

I. INTRODUCTION

In traditional power generation, large central generators at a remote end control the flow of power in transmission and distribution networks. Duting the '90s, a revolution caused by the connection of power generators in the distribution network as Distributed Energy Resources (DERs) or Distributed Generators (DGs). DGs can be connected to voltages from 230V to 150kV. Small power generators are coupled at smaller voltage networks. Larger generators of capacity ranging around 100MW are connected to high voltage buses. With the insertion of DGs, the electricity supply by central generators reduces. DGs need additional facilities for the power system to function with fewer central generators being utilized [1]. There are several types of DGs, e.g. solar, wind, fuel cell, small hydro, and rotating machines [2].

The addition of Distributed Generation in a distribution network effects the power system protection coordination in several ways, e.g. fault currents and short circuit currents from the DG increase, but there are changes in power quality, increase of reliability, and changes in losses and the voltage profile. The contribution of short circuit current depends on the DG location and the location of the fault. The fault current supplied by the substation is decreased and relays are not able or take longer time to detect the fault, which reduces relay sensitivity. Fault currents caused by rotating machine DGs are higher but those by inverter-based DGs are lower. To maintain power quality, Line Drop Compensators (LDCs) or Programmed Timer Distribution Transformers (PTDTs) are used to control voltage of distribution lines. Step Voltage Regulators (SVRs) are used to compensate the line voltage for dense power flow. These voltage regulators work properly if the feeder load is balanced. However, if many DGs are located on a particular line, the power flow gap widens between the feeder lines due to back power flow and results in voltage

---

Corresponding author: Krishan Lal Khatri fluctuations. The reliability of system increases due to the addition of DG sources as they locally inject power into the system. However, DGs' power flows in reverse direction and affects usual power flow, which can cause under- or over-voltage problems in the network. On the other hand, reduction in losses is one of the major positive impacts. The location of Distributed Generation units is a major issue that needs consideration in order to result in performance enhancement of the system with fewer losses. In general, the placement of DGs in the network is like placing capacitors, the only difference is that DGs provide both active and reactive power unlike capacitors that provide reactive power only [3].

In this work, the impact of the addition of DGs to a distribution network is analyzed mainly with regard to protection coordination of relays and the sizing and placement of DGs in IEEE 13-bus system is suggested with regard to the demand that there is minimum impact on protection coordination.

II. LITERATURE REVIEW

Authors in [4] proposed an expert system, tested it on a 22kV distribution network, and proposed multiple coordination settings to prevent the ill effects of DG placement. The system operates on Microsoft Windows, the interface modules and knowledge base are in Microsoft Access via Open Data Base Connectivity (ODBC). The expert system is used for the modification of protection settings. The method of "Adaptive Inverse Time Over Current Protection" was proposed in [5]. The authors suggested the calculation of newer values of Plug Setting Multiplier (PSM) by taking information from over current, fault current, voltage drop, and zero sequence fault current in the network. However, their results are not valid in changing network conditions. Authors in [6] simulated the Tunisian distribution network using directional relays and changing coordination settings on ETAP. Directional over current relays were used for correct tripping and to ensure time grading between primary and backup relay and the settings of over current relays were revised. To avoid the miss coordination problem of relays after the DG insertion, a directional feature was employed in relays. Relays and recloser settings were updated. The 11kV distribution feeder of El-Gharbia was simulated on ETAP in [7]. The directional protection scheme with relays of microprocessor type offered in [8] was used and the islanded mode with Distributed Generation was discussed. For achieving least-disturbance optimal protection settings and failure timings, the Time Coordination Method was introduced. With the new DGs, a new relay setting is generated and proposed every time on the ring system for achieving zero interruption frequency and supply failure rate [9]. Voltage sag coordination technique is based on disconnecting the DG when the voltage falls in the system to 0.5pu in less than 6 cycles. This technique is limited to small DGs [10, 11]. A method based on neutral networks and fuzzy logic was proposed in [12]. The neuro-fuzzy protection scheme was based on finding the optimal settings of Inverse Definite Minimum Time (IDMT) relays constructed on the position of DG and also on power system network conditions. By this technique, the results were improved through a current and voltage phasor-based Discrete Fourier Transform (DFT) algorithm. A Genetic Algorithm (GA) method based on finding the optimum Time Multiplier Setting (TMS) values of over current relays and circuit breaker's operating time for proper coordination of relays after DG insertion was proposed in [13]. This method is not suitable if the relays are conventional and they should be replaced by new relays which are costly. In [14], sizes and locations of DGs were found on MATLAB by using the GA on a 33-node system by allowing the relays to work on their conventional settings. The effect of three different sources (induction machines, wind turbine, and synchronous machines) on distribution network's fault current was checked on ETAP. By adding DGs, "Blind Zones" are created and different fault currents by different DGs were observed on a 4-bus system in [15]. Some major challenges in distribution systems were discussed in [16], including protection blinding, false tripping, fuse-recloser coordination, impedance variations during fault, unsynchronized reclosers, change in power flow's direction, loss of mains, deviations in levels of short-circuit currents, and the selection of a protection device. Fault Current Limiters (FCLs) were added to the system to limit the level of fault current and to avoid miss coordination issues [17].

In a nutshell, some methods that have been proposed work on changing protection device's settings after the DGs' insertion in the system. Methods that are used for solving coordination issues like Directional Relays, and FCL work with small DGs and will be discussed in detail below.

III. RESEARCH METHODOLOGY

We used the ETAP 12.6 simulation software to analyze the impact of the addition of DG units on the protection coordination of the network. ETAP is specialized software used for the analysis, monitoring, simulation, control, automation, and optimization of electrical systems. ETAP software in electrical engineering has the most complete collection of combined power system enterprise solutions. An IEEE 13-bus radial test system, as shown in Figure 1, is used in this study. The system is first modeled on ETAP and Load Flow Analysis is executed on the network to find out the power ratings of system. Short Circuit Analysis is done for finding the systems' behavior during fault conditions. Standard IEEE 13 bus is a radial feeder system with a small number of components, however it contains numerous features and can be used as an exemplary to find out the performance of power system. This system contains underground and overhead lines, distributed and spotted loads, a 500 KVA transformer, and capacitor banks. The following assumptions are considered:

- The Voltage regulator at bus 650 is not considered.
- The loads that were distributed between bus 671 and 632 are not considered.
- All the loads are three phase and constant impedance loads differentiate from one motor load.

For the protection of IEEE 13 bus system, fuses and over current relay devices are used. Current Transformers (CTs) are used for current measurement. High Voltage Circuit Breakers (HVCBs) and Low Voltage Circuit Breakers (LVCBs) are used as interrupting devices at 4.16kV and 0.4kV respectively. To find the overall behavior of the system, and for the selection of ratings of fuses, relays, and circuit breakers, load flow analysis and short-circuit analysis is performed. Maximum 3-phase and minimum single line-to-ground short circuit currents are considered and further studied because maximum short circuit current is used for protection device's selection (devices should be capable to withstand short circuit current without damage). Minimum values are used so that protection devices should not mal-operate at minimum short circuit current. The pickup current of relays should be more than the overload current in the system and less than Single Line-to-Ground (SLG) fault current. For the selection of protection devices SLG and 3-phase fault current values are calculated on the buses as shown in Figure 2 [17].

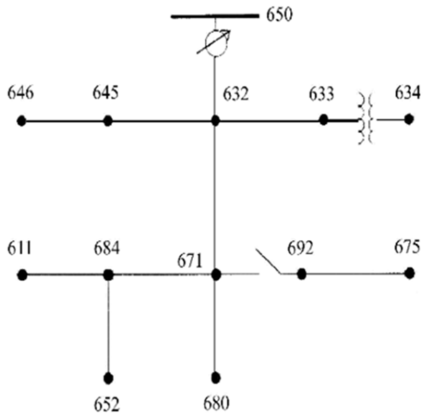


Fig. 1. Standard IEEE 13 bus system.

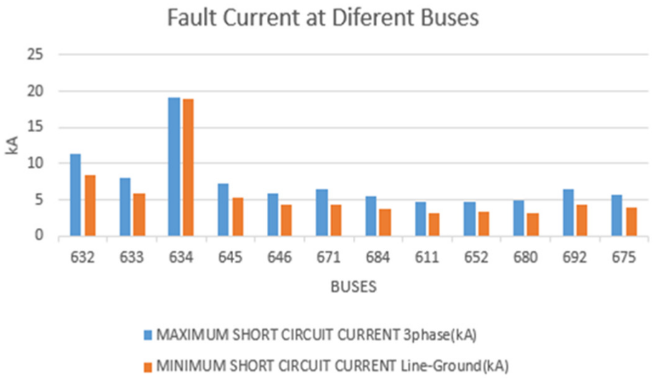


Fig. 2. Fault current at different bus.

As per calculations, maximum short circuit current is seen at Bus 634. Bus 634 is the bus with the 500KVA transformer 4.16/0.48 and is rated at 0.48kV while all the remaining buses are operating at higher voltage of 4.16kV. Fault current level is low at higher voltage levels. Inverse time over-current relays are used for feeder protection. Fuses are used for load protection. These relays operational time is inversely proportional to the fault current as shown (1)-(3). The different types of IEC curves are shown in (1), (2), and (3). Standard Inverse (SI) relays are used for the coordination of relays and the time of operation can be found with (1) [17]. ALSTOM P125 relay is used in the network. As PSM is required in (1), it can be calculated by (4) whereas the pickup current in (4) is calculated by (5). Relays should operate when current in the CTs exceeds the 25% of normal current. Time Dial Setting (TMS) varies from 0.025 to 1 in steps of 0.025. The operating time of the circuit breaker is considered as 0.4s and 10% overshoot is considered for proper coordination between the primary and the backup relay. The time of operation of backup relays should be greater than the time of operation of the primary relay and the operating time of circuit breaker [18, 19].

TABLE I. IEC 60255 CHARACTERISTIC CURVES

Relay types	Equation
Standard Inverse (SI)	$\mathrm{top} = \frac{0.14 * \mathrm{TMS}}{\mathrm{PSM}^{0.02} - 1}$ (1)
Very Inverse (VI)	$\mathrm{top} = \frac{13.5 * \mathrm{TMS}}{\mathrm{PSM} - 1}$ (2)
Extremely Inverse (EI)	$\mathrm{top} = \frac{80 * \mathrm{TMS}}{\mathrm{PSM}^2 - 1}$ (3)

$$\mathrm{PSM} = \frac{\text{Fault current in relay coil}}{\text{Pickup current}} \quad (4)$$

$$\text{Pickup current} = \text{Rated secondary current of CT} * \mathrm{PS} \ (5)$$

$$\mathrm{PSM} = \mathrm{top(p)} + 0.4 + 10\%\mathrm{top(b)} + 0.4 \quad (6)$$

In (6), top(p) is the primary relay's operating time and top(b) is backup relay's time of operation. An arrangement of thirteen over-current relays is used on ETAP. The TMS of relays that are closest to the load is 0.1s and the TMS of other relays is found by (1) and (6), making sure that the difference of 0.4s is maintained between the primary and the backup relay. In order to find out the impact of DG on the power system, different scenarios will be considered.

- Scenario 1: The impact on protection coordination of distribution network with the addition of DGs in the system.
- Scenario 2: The impact on protection coordination of distribution network with non-directional relays replaced by directional relays.
- Scenario 3: The addition of small DGs in the system instead of larger DGs.
- Scenario 4: The use of current limiting reactors or fault current limiters to limit the impact on protection coordination.

IV. RESULTS AND DISCUSSION

The standard system with supply from the grid only is considered. The protection of the IEEE 13-bus system is done by fuses and over current relays. Setting of fuses and relays is done on ETAP. System is perfectly coordinated before the placement of DGs in the system as verified on ETAP.

A. Scenario 1: The Effect on Protection Coordination when Adding DGs

Directly connected motors and generators are spinning as usual and supply a high amount of fault current. This extremely high short circuit current is dangerous because it may over stress equipment that is being used in the network to break the current, such as circuit breakers. If the value of short circuit current is very small, then it is also of concern, because fault will not be detected if this is too low and there is a possibility that protection devices will not operate correctly. Reliable fault detection from GDs can be difficult due to their small ratings. Designing the scheme of DG protection should be done carefully [1]. When the size of DG increases, the contribution of upstream fault current from grid will decrease which can cause a decrease in protection device sensitivity. Short circuit current increases in the downstream network as the size of DG increases or location of the DG is moved closer to the fault location. This may cause false tripping of relays. Some such examples are considered below. The relay settings are the same.

1) 3 MW DG at Bus 611

In this case, when a fault occurs at Bus 692, Relay 12 should operate primarily to isolate the fault. However, what happens in the system is that Relay 16 is operated first as shown in Figure 3 [16] because of the DG interconnection at Bus 611. Maloperation of relays has occurred, and protection coordination is lost. The Coordination Time Interval (CTI) between relays is below 0.3s.

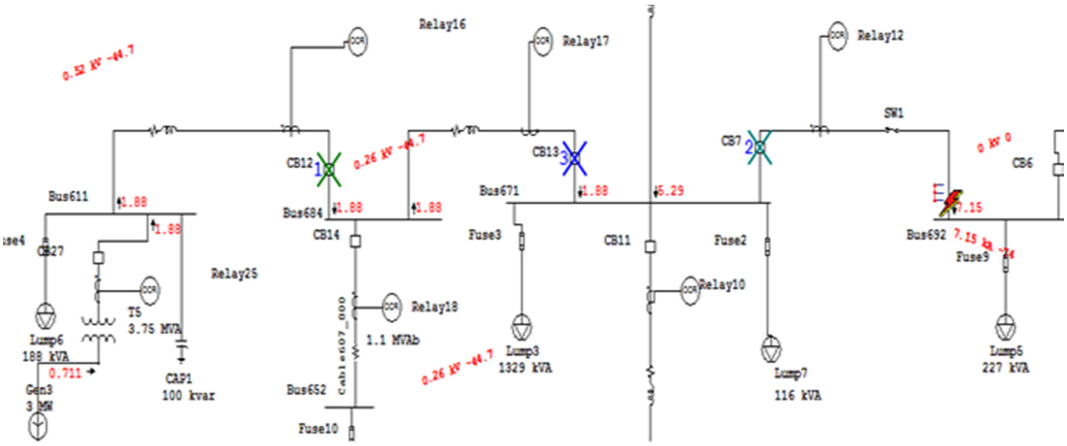


Fig. 3. 3MW DG at Bus 611.

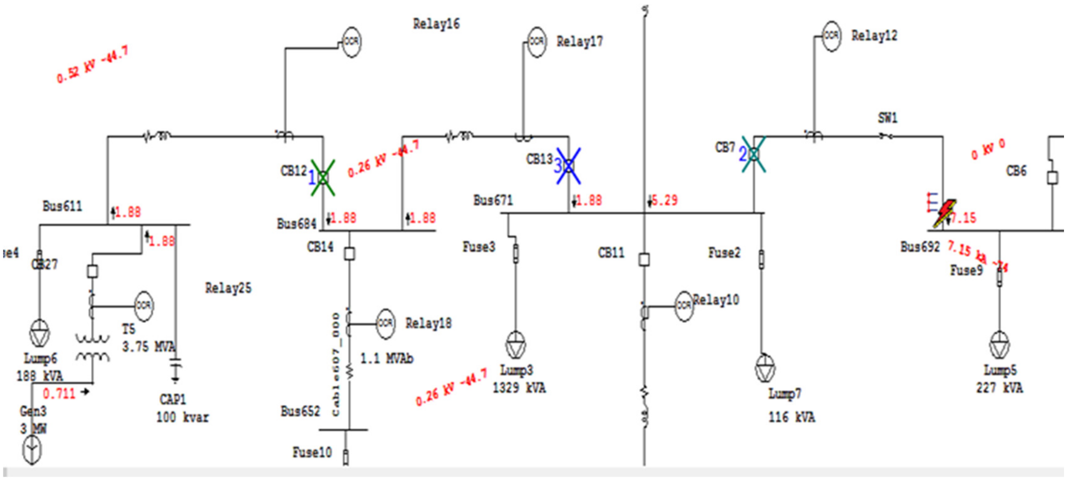


Fig. 4. 4 DGs at three different buses.

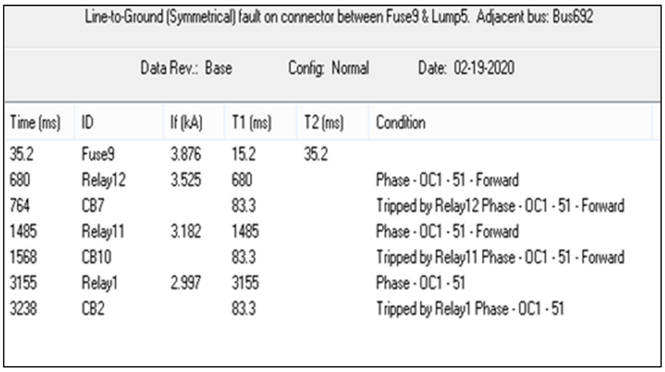
2) Three 1.5MW DGs at Bus 611, 675, and 646

In this case, 3 different DGs, each of 1.5MW are placed at 3 different points, 611, 675, and 646. A fault near Bus 692 is shown in Figure 4. False tripping occurs in this case as fault current is coming from 4 different points (the grid and the 3 DGs). The condition worsens by increasing the number of DGs in the system. Selectivity and sensitivity between the relays is disturbed. It is concluded that by the addition of DG in the system, the protection system is no more perfectly coordinated and some changes in relay settings are required. The following three techniques can be used for solving the protection coordination issues.

- Conventional (non-directional over current) relays should be replaced by directional relays.
- Adding small DGs in the network in order to have minimum effect on the protection.
- Use of fault current limiter.

B. Scenario 2: Conventional (Non Directional Over Current) Relays Replaced by Directional Relays

Electromechanical relays were used for protection in the above scenario. With the advancement of technology, microprocessor (numerical) based devices are utilized today but there are still a lot of conventional power systems utilizing electro-mechanical relays. A directional feature can be added in modern microprocessor-based relays, but our simulation system doesn't have this capability. Hence, we used digital relays: three 1.5 MW Synchronous (diesel) DGs at Buses 611, 675, and 646. The rest of the scenario is the same with the one considered above. Conventional non-directional relays were replaced by directional relays. The system is perfectly coordinated during the fault. For the 3-phase fault in the system between fuse 9 and lump 5 as shown in Figure 5, fault current is as high as 7.5kA in Fuse 9 which operated primarily at 15.8ms, then Relay 12 at 680ms, and finally Relay 11 at 1138ms. This can make relays insensitive for fault detection or to distinguish between load and fault current. By the addition of the directional feature in over current relays, the selectivity and all other protection issues are solved. There are certain situations where coordination interval can be disturbed. To correct this, relay settings need to be changed if there is any such coordination issue.



Line-to-Ground (Symmetrical) fault on connector between Fuse9 & Lump5. Adjacent bus: Bus692

Data Rev.: Base Config: Normal Date: 02-19-2020

Time (ms)	ID	If (kA)	T1 (ms)	T2 (ms)	Condition
35.2	Fuse9	3.876	15.2	35.2	
680	Relay12	3.525	680		Phase - OC1 - 51 - Forward
764	CB7		83.3		Tripped by Relay12 Phase - OC1 - 51 - Forward
1485	Relay11	3.182	1485		Phase - OC1 - 51 - Forward
1568	CB10		83.3		Tripped by Relay11 Phase - OC1 - 51 - Forward
3155	Relay1	2.997	3155		Phase - OC1 - 51
3238	CB2		83.3		Tripped by Relay1 Phase - OC1 - 51

Fig. 5. Fault placed at Bus 692.

C. Scenario 3: Addition of Small DGs in the Network

Replacement of conventional relays by the addition of each DG in the network is not possible sometimes due to cost and system redesign issues. Another solution is adding small DGs that have overall very small impact on the fault currents of the system. In order to penetrate Distributed Generation into the network for this scenario, power flow is performed on the network so that the voltage range is not violated. As it is discussed above, over-voltages occur on the nodes by the addition of DGs in the network. Thermal limits should not be violated, any of the equipment should not be overloaded by the addition of DGs. Protection coordination and CTI should not be violated. We analyzed DGs in the range of 100kW~10MW with a step size of 100kW. Bus 632 cannot be considered for the addition of DG because of it is connected to the grid. Maximum increase, of the order of 100%, in short circuit current is observed on Bus 634, hence it is also not considered for the addition of DG. The required conditions to maintain power quality and system integrity are:

- Voltage range should be within limits, i.e., 95%-105%.
- The thermal limit of equipment should not be exceeded.
- The protection devices should be perfectly coordinated.

Small synchronous DGs are installed in the system at Buses 611, 633, 645, 692, and 671 simultaneously. The total load connected in the distribution system is of 3.5MW. To meet the total load demand, 3.5MW of generation which consists of small DGs are added in the system and the impact on protection coordination is checked and verified as shown in Figures 6 and 7. Small DGs are installed at distant buses having low TMS values as the fault current of DG increases. Increased fault current has maximum impact on the relay that has low TMS because the time of operation of the relays is proportional to TMS as shown in (7).

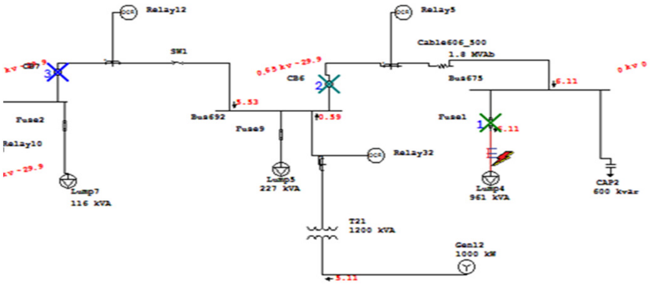
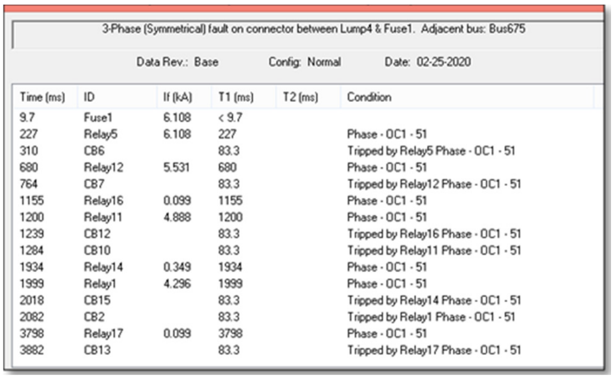


Fig. 6. Fault near Bus 675.



3-Phase (Symmetrical) fault on connector between Lump4 & Fuse1. Adjacent bus: Bus675

Data Rev.: Base Config: Normal Date: 02-25-2020

Time (ms)	ID	If (kA)	T1 (ms)	T2 (ms)	Condition
9.7	Fuse1	6.108	< 9.7		
227	Relay5	6.108	227		Phase - OC1 - 51
310	CB6		83.3		Tripped by Relay5 Phase - OC1 - 51
680	Relay12	5.531	680		Phase - OC1 - 51
764	CB7		83.3		Tripped by Relay12 Phase - OC1 - 51
1155	Relay16	0.099	1155		Phase - OC1 - 51
1200	Relay11	4.888	1200		Phase - OC1 - 51
1239	CB12		83.3		Tripped by Relay16 Phase - OC1 - 51
1284	CB10		83.3		Tripped by Relay11 Phase - OC1 - 51
1934	Relay14	0.349	1934		Phase - OC1 - 51
1999	Relay1	4.296	1999		Phase - OC1 - 51
2018	CB15		83.3		Tripped by Relay14 Phase - OC1 - 51
2082	CB2		83.3		Tripped by Relay1 Phase - OC1 - 51
3798	Relay17	0.099	3798		Phase - OC1 - 51
3882	CB13		83.3		Tripped by Relay17 Phase - OC1 - 51

Fig. 7. Sequence of operation for the fault near Bus 675.

$$\text{Time of Operation (TOP)} = \text{TMS} * \text{PSM} \quad (7)$$

DGs are installed at 4.16kV Buses and should not be installed at 0.48kV Bus 634 as it shows the highest impact on the system. This is because of the placement of the step down transformer between Buses 633 and 634. The over-voltages on buses are within the allowed range. The system is coordinated as observed in the two different examples discussed above and the operating time of protection devices is checked.

D. Scenario 4: Use of Fault Current Limiter (FCL)

Working with small DGs is costly and technically challenging. Large DGs can be placed in the system in the presence of FCLs. We used Current Limiting Reactors (CLRs) that limit the fault current. A CLR is used for limiting current because it produces small loss in real (active) power. However, the use of FCL has limitations because of the power loss it causes in normal conditions. An example is considered, and the results are shown below with and without the presence of FCL in the system. In Figure 8, a DG of 3.5MW at Bus 684 with a reactor is shown. For a fault near Bus 611, the time of operation without FCL is shown in Figure 9. Relay 17 operates as a backup of Relay 16. After Relay 16 any other relay should operate after a time interval of 0.4s but in this case Relay 12 is operating simultaneously with Relay 17. This can be restricted by using FCL in series with the DG. When we did this, the relays operated sequentially as shown in Figure 10.

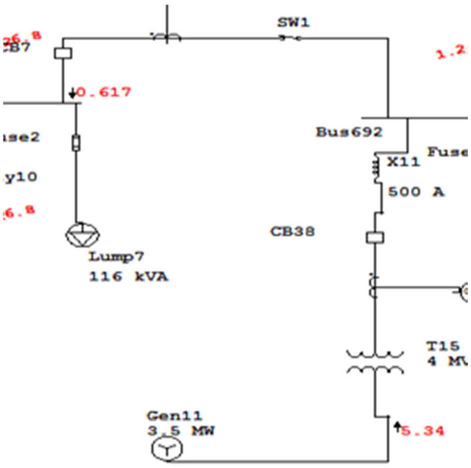
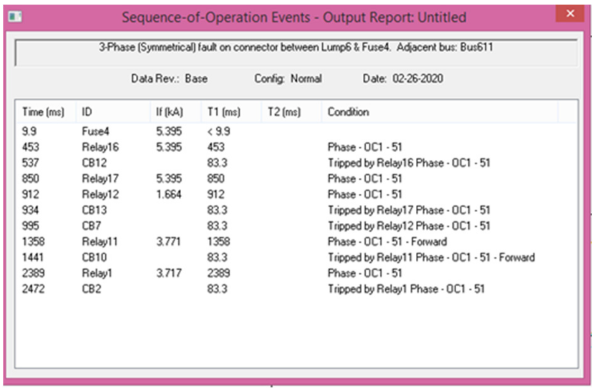


Fig. 8. 8 DGs of 3.5MW with CLR.

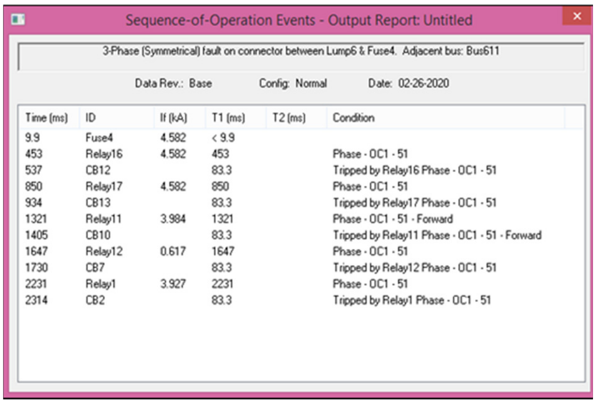


```
3-Phase (Symmetrical) fault on connector between Lump6 & Fuse4.  Adjacent bus: Bus611

Data Rev.:  Base        Config:  Normal        Date:  02-26-2020

Time (ms)  ID        If (kA)  T1 (ms)  T2 (ms)  Condition
9.9        Fuse4     5.395    < 9.9
453        Relay16   5.395    453               Phase - OC1 - 51
537        CB12               83.3              Tripped by Relay16 Phase - OC1 - 51
850        Relay17   5.395    850               Phase - OC1 - 51
912        Relay12   1.664    912               Phase - OC1 - 51
934        CB13               83.3              Tripped by Relay17 Phase - OC1 - 51
995        CB7                83.3              Tripped by Relay12 Phase - OC1 - 51
1358       Relay11   3.771    1358              Phase - OC1 - 51 - Forward
1441       CB10               83.3              Tripped by Relay11 Phase - OC1 - 51 - Forward
2389       Relay1    3.717    2389              Phase - OC1 - 51
2472       CB2                83.3              Tripped by Relay1 Phase - OC1 - 51
```

Fig. 9. Sequence of operations without the CLR.

V. CONCLUSION AND RECOMMENDATIONS

In this paper, the effect caused by DGs on the power system's protection network is analyzed and possible techniques for the addition of DGs to the distribution network have been proposed that result in minimum impact on protection coordination in the IEEE 13-bus system. Since it is costly and technically challenging to improve the settings of original relay systems or to upgrade the relays, the proposed techniques would help network operators install DGs with minimum impact on the distribution network. Different scenarios have been established in order to investigate the effects of DG and to find solutions to the emerging issues.



```
3-Phase (Symmetrical) fault on connector between Lump6 & Fuse4.  Adjacent bus: Bus611

Data Rev.:  Base        Config:  Normal        Date:  02-26-2020

Time (ms)  ID        If (kA)  T1 (ms)  T2 (ms)  Condition
9.9        Fuse4     4.582    < 9.9
453        Relay16   4.582    453               Phase - OC1 - 51
537        CB12               83.3              Tripped by Relay16 Phase - OC1 - 51
850        Relay17   4.582    850               Phase - OC1 - 51
934        CB13               83.3              Tripped by Relay17 Phase - OC1 - 51
1321       Relay11   3.984    1321              Phase - OC1 - 51 - Forward
1405       CB10               83.3              Tripped by Relay11 Phase - OC1 - 51 - Forward
1647       Relay12   0.617    1647              Phase - OC1 - 51
1730       CB7                83.3              Tripped by Relay12 Phase - OC1 - 51
2231       Relay1    3.927    2231              Phase - OC1 - 51
2314       CB2                83.3              Tripped by Relay1 Phase - OC1 - 51
```

Fig. 10. Sequence of operations with the CLR.

Coordination issues and other problems related to protection can be solved by the use of directional over current relays in the network. The drawback is the difficulty of replacement of the conventional electromechanical relays that are already installed in the system. This will not be the case for numerical relays. Another solution is using small DGs with very low power ratings and low fault currents. This will have minimum impact on protection system coordination and very few protection coordination issues will arise. However, it can be economically and technically challenging due to the issue of protection blinding. The third method proposed is using FCLs to limit the current from the DG. FCLs can be used in the network for limiting the fault current from the DGs. For this optimal value, CLRs need to be calculated as they produce power loss in normal conditions also. By using FCLs in the network, DG selection is based on the trade between minimum impact on protection coordination with low real power loss in normal conditions.

VI. FUTURE WORK

Simple CLRs were analyzed in this work and their effects on protection coordination of relays were studied on ETAP software. By the use of limiters in the network, working with DGs of different types and sizes will have minimum impact on system in terms of current injection from the devices. However, it results in a tradeoff in terms of power loss in usual network operation. With the recent advances in conductor materials, advanced technologies have been introduced for limiting currents from DGs such as the use of superconducting materials. Super conducting FCLs have very small resistance and have almost no impact on normal real power flows. However, in order to use such current limiters in the system, a comparison between the cost of power loss and the cost of changing of relays should be analyzed before implementing. In future work, we aim to study and analyze limiters that operate in fault conditions only and do not cause any power loss during normal operating conditions.

REFERENCES

[1] N. Jenkins, J. Ekanayake, and G. Strbac, *Distributed Generation*. London, UK: The Institution of Engineering and Technology, 2010.

[2] X. Huang, Z. Zhang, and J. Jiang, "Fuel Cell Technology for Distributed Generation: An Overview," in *2006 IEEE International Symposium on*

Industrial Electronics, Montreal, Canada, Jul. 2006, vol. 2, pp. 1613–1618, https://doi.org/10.1109/ISIE.2006.295713.

[3] C. L. T. Borges and D. M. Falcão, "Optimal distributed generation allocation for reliability, losses, and voltage improvement," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 6, pp. 413–420, Jul. 2006, https://doi.org/10.1016/j.ijepes.2006.02.003.

[4] K. Tuitemwong and S. Premrudeepreechacharn, "Expert system for protection coordination of distribution system with distributed generators," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 3, pp. 466–471, Mar. 2011, https://doi.org/10.1016/j.ijepes.2010.10.009.

[5] E. Purwar, D. N. Vishwakarma, and S. P. Singh, "A new adaptive inverse-time protection scheme for modern distribution systems with distributed generation," in *2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington, DC, USA, Apr. 2017, https://doi.org/10.1109/ISGT.2017.8086049.

[6] A. Meddeb, N. Sahbeni, H. Jmii, and S. Chebbi, "Impact of Distributed Generation on the Protection System in Tunisian Distribution Network," in *2018 15th International Multi-Conference on Systems, Signals Devices (SSD)*, Yasmine Hammamet, Tunisia, Mar. 2018, pp. 514–520, https://doi.org/10.1109/SSD.2018.8570584.

[7] A. Kamel, M. Alaam, A. Azmy, and A. Abdelaziz, "Protection Coordination Of Distribution Systems Equipped With Distributed Generations," *Electrical and Electronics Engineering: An International Journal*, vol. 2, pp. 1–13, May 2013.

[8] M. A. Zamani, T. S. Sidhu, and A. Yazdani, "A Protection Strategy and Microprocessor-Based Relay for Low-Voltage Microgrids," *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1873–1883, Jul. 2011, https://doi.org/10.1109/TPWRD.2011.2120628.

[9] C. W. So and K. K. Li, "Protection relay coordination on ring-fed distribution network with distributed generations," in *2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. TENCOM '02 Proceedings*, Beijing, China, Oct. 2002, vol. 3, pp. 1885–1888 vol.3, https://doi.org/10.1109/TENCON.2002.1182705.

[10] F. A. Viawan and M. Reza, "The impact of synchronous distributed generation on voltage dip and overcurrent protection coordination," in *2005 International Conference on Future Power Systems*, Amsterdam, Netherlands, Nov. 2005, https://doi.org/10.1109/FPS.2005.204298.

[11] J. C. Gomez and M. M. Morcos, "Coordination of voltage sag and overcurrent protection in DG systems," *IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. 214–218, Jan. 2005, https://doi.org/10.1109/TPWRD.2004.839216.

[12] D. S. Kumar, D. Srinivasan, and T. Reindl, "A Fast and Scalable Protection Scheme for Distribution Networks With Distributed Generation," *IEEE Transactions on Power Delivery*, vol. 31, no. 1, pp. 67–75, Feb. 2016, https://doi.org/10.1109/TPWRD.2015.2464107.

[13] M. Dahal, A. K. Jha, and R. Ghimire, "Impact of Renewable Distributed Generation on Protection Coordination of Distribution System," *International Journal of Engineering and Applied Sciences*, vol. 6, no. 4, pp. 36–42, Apr. 2019, https://doi.org/10.31873/IJEAS/6.4.2019.17.

[14] H. Zhan *et al.*, "Relay Protection Coordination Integrated Optimal Placement and Sizing of Distributed Generation Sources in Distribution Networks," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 55–65, Jan. 2016, https://doi.org/10.1109/TSG.2015.2420667.

[15] A. Girgis and S. Brahma, "Effect of distributed generation on protective device coordination in distribution system," in *LESCOPE 01. 2001 Large Engineering Systems Conference on Power Engineering. Conference Proceedings. Theme: Powering Beyond 2001*, Jul. 2001, pp. 115–119, https://doi.org/10.1109/LESCPE.2001.941636.

[16] U. Shahzad, S. Kahrobaee, and S. Asgarpoor, "Protection of Distributed Generation: Challenges and Solutions," *Energy and Power Engineering*, vol. 9, no. 10, pp. 614–653, Sep. 2017, https://doi.org/10.4236/epe.2017.910042.

[17] M. Barzegari, A. N. Fard, M. M. Hamidi, and A. J. Shahrood, "Optimal coordination of directional overcurrent relays in the presence of distributed generation using FCLs," in *2010 IEEE International Energy Conference*, Manama, Bahrain, Dec. 2010, pp. 826–829, https://doi.org/10.1109/ENERGYCON.2010.5771796.

[18] M. A. Aman, S. Ahmad, B. Noor, and F. W. Karam, "Mitigating the Adverse Impact of Un-Deterministic Distributed Generation on a Distribution System Considering Voltage Profile," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2998–3003, Jun. 2018, https://doi.org/10.48084/etasr.2033.

[19] M. A. Raza, K. L. Khatri, K. Rafique, and A. S. Saand, "Harnessing Electrical Power from Hybrid Biomass-Solid Waste Energy Resources for Microgrids in Underdeveloped and Developing Countries," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7257–7261, Jun. 2021, https://doi.org/10.48084/etasr.4177.

A Multipath Cluster-Based Routing Protocol For Mobile Ad Hoc Networks

Mohammed A. Mahdi
Computer Science and Information Department
College of Computer Science and Engineering
University of Ha'il
Ha'il, Saudi Arabia
wsabi3@gmail.com

Tat-Chee Wan
School of Computer Science
Universiti Sains Malaysia
Penang, Malaysia
tcwan@usm.my

Adnan Mahdi
School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Malaysia and College of Computer Science and Engineering, University of Hafr Al Batin, Saudi Arabia
adnanmhdi@gmail.com

Mohamed A.G. Hazber
Computer Science and Information Department, College of Computer Science and Engineering
University of Ha'il
Ha'il, Saudi Arabia
m.hazber@uoh.edu.sa

Badiea Abdulkarem Mohammed
Computer Engineering Department
College of Computer Science and Engineering
University of Ha'il
Ha'il, Saudi Arabia
b.alshaibani@uoh.edu.sa

Abstract-A MANET (Mobile Ad-hoc Network) is a group of mobile network nodes dynamically forming a network without any pre-existing infrastructure. Multi-path routing protocols in MANETs try to discover and use multiple routes between source and destination nodes. Multipath routing is typically used to reduce average delay, increase transmission reliability, provide load balancing among multiple routes, and improve security and overall QoS (Quality of Service). In this paper, the Cluster Based Routing Protocol (CBRP), which is a single path MANET protocol is enhanced to use multiple paths. The traffic will be distributed among multiple paths to reduce network traffic congestion and decrease delay. An analytical model is used for multipath and single path CBRP routing protocols in MANETs to estimate the end-to-end delay and queue length. The analytical results show that the average delay and average queue length in multipath CBRP are less than the average delay and queue length in single path CBRP.

Keywords-MANET; routing protocols; single path; multi path; CBRP

I. INTRODUCTION

A Mobile Ad-hoc Network (MANET) is a collection of wireless mobile nodes/routers. The nodes dynamically form a temporary network without the use of any existing network infrastructure or centralized administration. The mobile nodes are equipped with wireless interfaces that utilize radio channels to communicate with each other, without the need for centralized management [1]. MANETs can be used in classrooms, battlefields, and disaster recovery situations [1, 2]. Routing protocols are an important part of any network. They play the main role for any communication in the network and they are used to establish and correct efficient routes among a pair of network nodes and to deliver messages in a timely manner [3-5]. Routing protocols in MANETs are differentiated in terms of hop-by-hop or source routing, reactive or proactive approach, single path or multipath, distance vector or link state, and unicast or multicast [6]. Multipath routing is a technique to overcome the problems of frequent topological changes and link instability because the use of multiple paths could diminish the effect of possible link failures between the nodes [7, 8]. Thus, multipath routing protocols of wireless ad-hoc networks are superior over conventional single-path ad-hoc routing protocols [9] because the multipath routing protocols reduce end-to-end delay, provide more reliable network services, allow load balancing among multiple routes, provide higher aggregate bandwidth, improve security, and overall they improve the Quality of Service (QoS) [10, 11]. Various multipath routing protocols have been proposed for wireless ad hoc networks. A multipath routing protocol based on the source routing protocol DSR is the Salvage capable opportunistic Node-disjoint Multipath Routing (SNMR) [12]. A multipath routing protocol based on the Ad hoc On-demand routing scheme is the Ad hoc On-demand Multipath Distance Vector (AOMDV) [3, 10].

II. MANET ROUTING PROTOCOLS

MANET routing protocols based on the schemes of discovering and maintaining paths are classified into three classes: reactive, proactive, and hybrid [13, 14] (Figure 1). In terms of reactivity, they are classified as reactive or proactive, and regarding the number of paths as single or multi-path. The reactive method is more efficient than the proactive method. The most common-known reactive routing protocols in MANETs (single-path and multipath) are described below.

Corresponding author: Mohammed A. Mahdi

A. MANET Single-Path Routing Protocols

MANET single path routing protocols mainly intend to maintain a single route between a source and a destination node. The most popular routing protocols of this type are AODV, DSR, and CBRP.

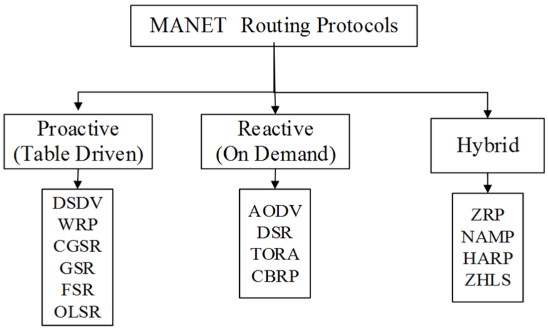


Fig. 1. MANET routing protocol taxonomy.

1) Dynamic Source Routing (DSR)

DSR [15, 16] is an on-demand routing protocol that is based on the concept of source routing. It is a simple and an efficient routing protocol, specifically designed for use in multi-hop wireless ad hoc networks. It consists of two mechanisms, Route Discovery and Route Maintenance, that work together to allow nodes to discover and maintain routes from the sources to arbitrary destinations in network. DSR computes the paths when necessary and then maintains them. It applies on-demand schemes for both path discovery and path maintenance. Thus, it makes the routing overhead traffic to scale automatically to the actual needed size, which is considered as its main advantage.

2) Ad hoc On-demand Distance Vector (AODV)

AODV [17] is a reactive on-demand routing protocol that allows dynamic, self-starting, multi-hop routing among mobile nodes desiring to establish and maintain an MANET. It is essentially a combination of DSR and DSDV (Destination Sequenced Distance Vector) protocol. AODV uses the basic on-demand mechanisms of Route Discovery and Route Maintenance as in DSR and a next hop routing model with sequence numbers and periodic beacons to discover and maintain routes of DSDV. Also, when the ad-hoc network topology changes, an AODV sequence number method is used to avoid long-term loops. Moreover, AODV allows the mobile nodes to obtain routes quickly and does not require maintaining routes that are not in active communication.

3) Cluster-Based Routing Protocol (CBRP)

CBRP [18-20] is a hierarchical on-demand routing protocol that uses source routing, like DSR, to avoid forming loops. CBRP is an efficient, scalable, and robust routing protocol for MANETs. It features more advantageous metrics than existing routing protocols, its overhead is less and its throughput is higher than AODV's. The CBRP is designed to be used in medium-to-large MANETs. CBRP groups the nodes of a network into several clusters. Each cluster has a cluster head which coordinates data transmission within the cluster and with other clusters. There are four possible states of the node in the cluster: NORMAL, ISOLATED, CLUSTERHEAD, and GATEWAY. Figure 2 shows these states. One of the advantages of CBRP is that only cluster heads exchange routing information, therefore the number of control overhead transmitted through the network is far less than the traditional flooding methods. Other features of CBRP [20] are:

- Distributed operation is fully used.
- During the dynamic route discovery process there is less flooding traffic.
- Explicit exploitation of uni-directional links that would otherwise be unused.
- Broken paths can be repaired locally without route rediscovery.
- Sub-optimal routes can be shortened as they are used.

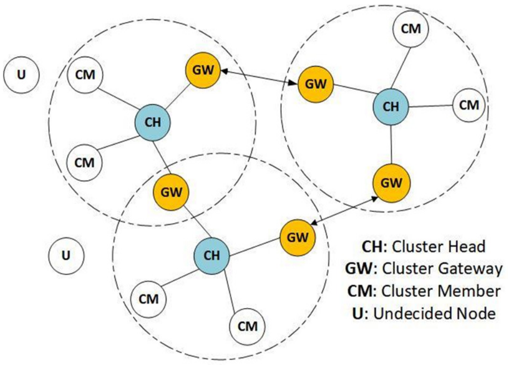


Fig. 2. Cluster Based Routing Protocol (CBRP).

B. MANET Multi-Path Routing Protocols

Multi-path routing protocols try to discover a multiple route from the source node to the destination node. Multipath routing can effectively improve reliability, end-to-end delay, and achieve load balancing, but it is more complicated than single path routing [11]. Most multipath routing protocols proposed for wireless ad hoc networks are based on reactive routing. They are mostly created on the basis of DSR and AODV.

1) Ad hoc On-demand Multipath Distance Vector routing (AOMDV)

AOMDV [3, 10] extends the AODV protocol to compute multiple paths between the source and the destination during route discovery. Multiple paths are computed to guarantee that the route is loop-free and disjoint. The AOMDV protocol calculates multiple loop-free paths, where every node maintains an advertised hop count for each destination node. Therefore, loop freedom is accomplished, and the advertised hop count represents the maximum hop count for all multiple paths. AOMDV has three novel aspects compared to other on-demand multipath protocols. First, it does not have high intermodal coordination overhead. Second, it ensures disjointness of alternate routes via distributed computations without the use of source routing. Finally, it computes alternate paths with minimal additional overhead.

2) *Salvage Capable Opportunistic Node-Disjoint Multipath Routing (SNMR)*

SNMR [12] is based on DSR. In SNMR, a route discovery attempt, a primary path, and a node disjoint route are created between the source and the destination node along with alternative routes for the intermediate nodes of the primary route. SNMR reduces the overhead caused by duplicate RREQ broadcasting in creating multiple routes and recovers quickly from connection failures due to its packet salvaging ability. Minimum overhead, node disjoint multipath, non-duplicate RREQ rebroadcasting, and packet salvaging are the main advantages of SNMR protocol. A bloom filter is used in SNMR for space saving. Within the primary route, the intermediate nodes are compressed in a bloom filter containing their hashed addresses. SNMR uses the hashed node address to check the presence of a node in the bloom filter. Therefore, an intermediate node can check if a neighboring node is included in the primary path or not.

III. THE PROPOSED MULTIPATH CBRP

This paper focuses on the enhancement of CBRP routing protocol to work as a multipath routing protocol with improved QoS.

A. *Overview*

Multipath CBRP routing protocol extends the single path CBRP to use multiple paths in distributing traffic from the source node (S) to destination node (D). CBRP is selected due to its above mentioned features. For example, CBRP has less routing overhead and higher throughput than the AODV protocol. CBRP repaires broken links locally without rediscovery, which leads to the reduction of broken links in the network and increased QoS. On the other hand, CBRP has higher average delay than AODV. Therefore, it should be enhanced to reduce the average delay. The proposed Multi Path CBRP (MP-CBRP) improves the QoS in MANETs. MP-CBRP uses the idea of disjoint paths to distribute data packets over multiple routes. This distribution increases packet delivery ratio, data throughput, and reduces the average delay and queue length because it decreases congestion. An analytical model is used for multipath CBRP routing protocol and single path CBRP routing protocol to estimate the end-to-end delay and queue length.

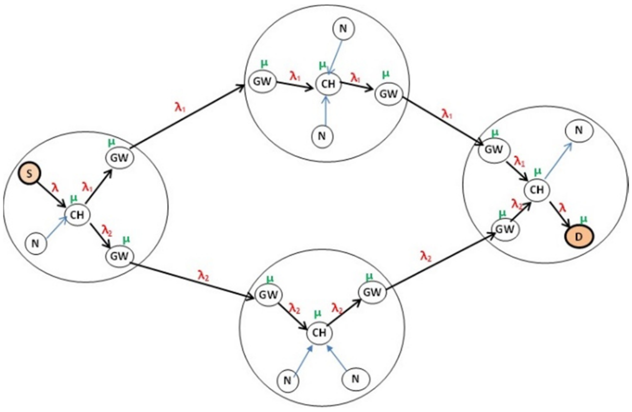


Fig. 3. The proposed Multipath CBRP Routing Protocol (MP-CBRP).

As shown in Figure 3 two paths are discovered from the source node (S) to destination node (D). Data traffic λ is distributed through these two paths, divided to $\lambda 1$ and $\lambda 2$. The $\lambda 1$ is sent through path1 and $\lambda 2$ is sent through path2. S is the source node and D the destination node. CH is the cluster head, RREQ is the Route Request, RREP the Route Reply and RC is the Route Cache. The pseudo code of the proposed MP-CBRP is illustrated in Figure 4 and the flow diagram of the proposed MP-CBRP is shown in Figure 5.

```
BEGIN
Source node (S) sends a RREQ) to its CH
IF D is found in its Cluster
    S Sends RREQ to D
ELSE
      Loop: CH discovers multiple paths by broadcasting RREQ to all
      adjacent clusters using its GWs.
          When GWs in adjacent clusters receive RREQ, it forwards
                              RREQ to its CH.
        IF CH finds D in its cluster
            CH sends RREQ to D
            D sends an RREP message to S
                      IF RREP reaches to S
                            S Stores the Multiple paths which is recoded
                            in RREP in its RC
                            S selects the disjoint multiple paths from RC
                            S Sends Data Packets through the disjoint
                            multiple paths
                      ELSE go to Loop:
                      END IF
            ELSE go to Loop:
          END IF
END IF
END
```

Fig. 4. Pseudo code of the proposed MP-CBRP.

B. *Network Model*

To obtain the Average Delay and Queue Length for multipath CBRP, let us characterize the disjoint multipath network as follows:

- The network with N disjoint path resources is characterized by the set of values $\mu = (\mu 1, \mu 2, \mu 3, \mu i, \dots, \mu n)$ for the individual resources. A resource could be the data transmission capacity of the given communication channel.
- The set $\lambda = (\lambda 1, \lambda 2, \lambda 3, .., \lambda i, \dots, \lambda n)$ represents the average arrival rate distribution across the set of the resources (μ).
- The set $D = (D1, D2, D3, \dots, Di, \dots, Dn)$ represents the delay of values that applied on each of the n paths when the distribution of the traffic along the paths is λ, and the distribution of the resources along the paths is μ.
- The set $Q = (Q1, Q2, Q3, \dots, Qi, \dots, Qn)$ represents the queue length values experienced on each of the n paths

when the distribution of the traffic along the paths is λ, and the distribution of the resources along the paths is μ.

According to the above assumption, and using the queuing model *M/M/1* queue [21] the delay (Di) and queue size (Qi) can be calculated as:

$$Di = \frac{1}{\mu i - \lambda i} \quad (1)$$

$$Qi = \frac{pi}{1-pi} \text{ and } pi = \frac{\lambda i}{\mu i} \quad (2)$$

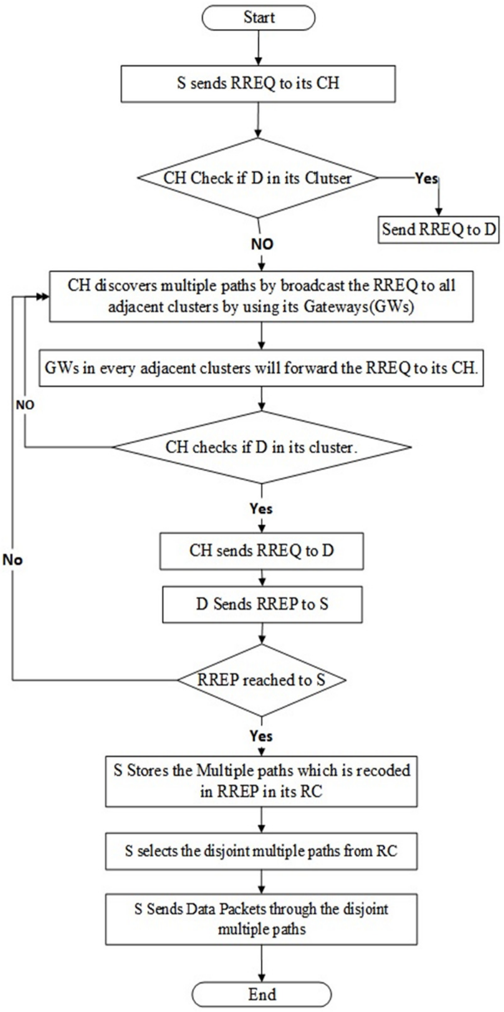


Fig. 5. Flow diagram of the proposed MP-CBRP.

IV. RESULTS AND DISCUSSION

The results are shown in the form of graphs. Graphs show the comparison between single path CBRP and MP-CBRP with two paths, based on the terms of average delay and queue length. Values of average arrival rate (λ): 0.1, 0.3, 0.5, 0.7, 0.9 are used for all the nodes in the network model of MP-CBRP and service rate μ=1.

A. *Average Delay*

Average delay is calculated by (1). Figures 6 and 7 illustrate the average delay for single path and MP CBRP routing protocol. Two paths have been considered for the MP CBRP, and two different scenarios have been considered.

1) *Average Delay Scenario 1*

Traffic is split equally between the two paths in the MP CBRP model. Figure 6 shows that the average delay for MP CBRP is lesser than the average delay of single path CBRP using different arrival rates.

2) *Average Delay Scenario 2*

Traffic is not split equally between the two paths in MP CBRP model. The traffic for path 1 is 30% and for path 2 is 70%. Figure 7 shows that the average delay of MP CBRP is lesser than the average delay of single path CBRP using different arrival rates.

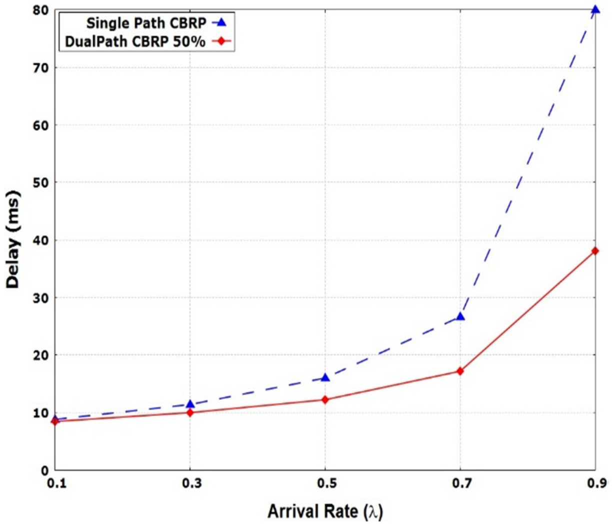


Fig. 6. Average delay for single path CBRP and equally divided traffic MP CBRP.

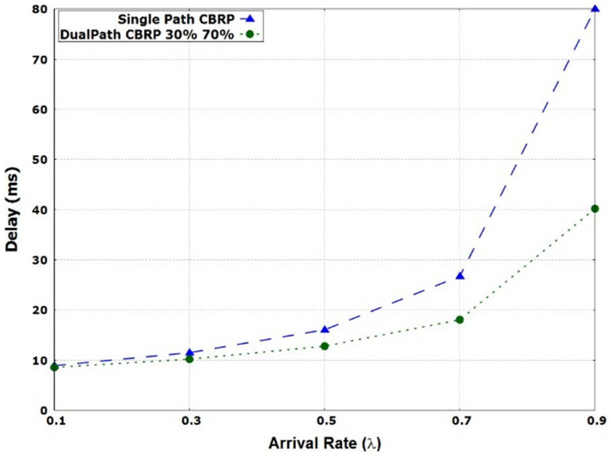


Fig. 7. Average delay for single path CBRP and MP CBRP using 30% traffic for path1 and 70% traffic for path 2

Figure 8 shows the results of average delay for scenario 1 and scenario 2. The average delay for MP-CBRP in scenario 1 is better than the average delay for MP-CBRP in scenario 2 due to the high network congestion in scenario 2.

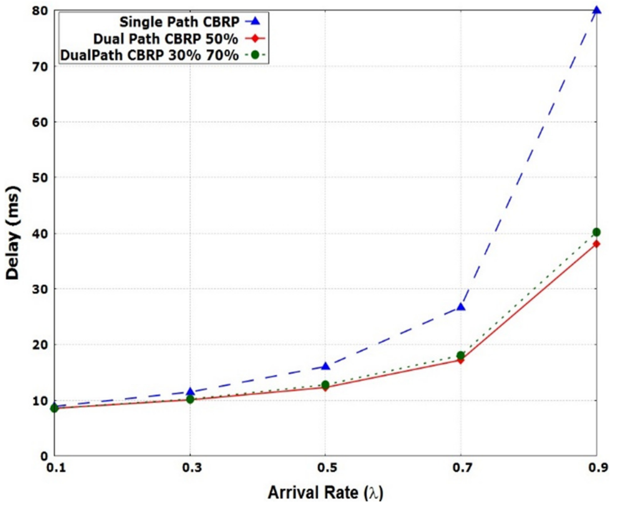


Fig. 8. Average delay for single-path CBRP and MP-CBRP (scenarios 1 and 2).

B. *Average Queue Length*

The average queue length is calculated by (2). Figures 9 and 10 illustrate the average queue length for single path CBRP and MP CBRP. Two paths have been considered for CBRP multipath, and two different scenarios have been considered.

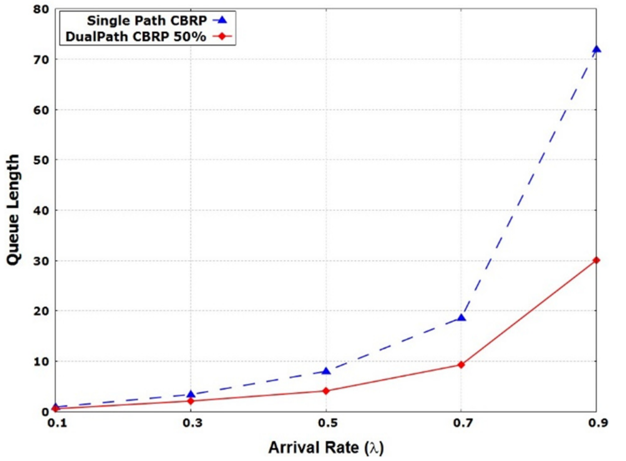


Fig. 9. Average queue length for single path and MP CBRP for scenario 1.

1) *Average Queue Length Scenario 1*

The traffic is split equally between the two paths in MP CBRP. Figure 9 shows that the average queue length for MP CBRP is better than the average queue length in single path CBRP that uses different arrival rates.

2) *Average Queue Length Scenario 2*

Traffic is not split equally between the two paths in MP CBRP model. The traffic for path 1 is 30% and for path 2 is 70%. Figure 10 shows that the average queue length for MP CBRP is better than average queue length in single path CBRP that uses different arrival rates.

Figure 11 shows the results of average queue length for scenario 1 and scenario 2. The queue length of MP-CBRP in scenario 1 is better the average queue length of MP-CBRP in scenario 2 due to the high network congestion in scenario 2.

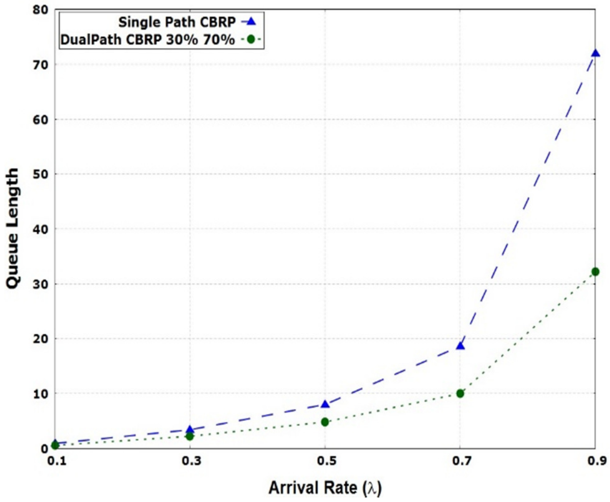


Fig. 10. Average queue length for single path and MP CBRP for scenario 2.

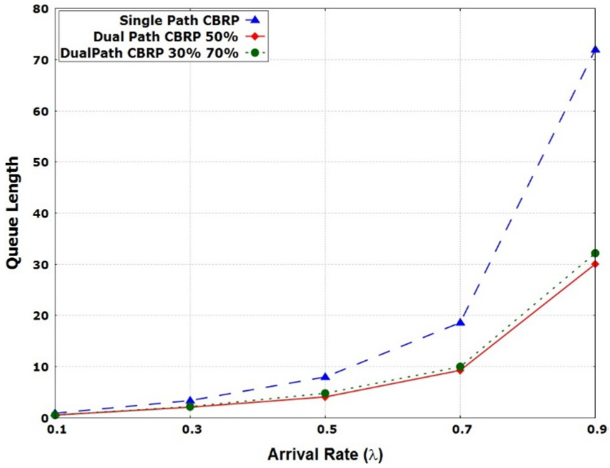


Fig. 11. Average queue length for single-path CBRP and MP-CBRP (scenarios 1 and 2).

V. CONCLUSION

This paper proposes the enhancement of the single path CBRP MANET protocol to work as an MP routing protocol. The mathematical model using the *M/M/1* queueing model was used to evaluate the performance of MP-CBRP and single path CBRP routing protocols. The results show that the average delay and average queue length in MP-CBRP are better than the average delay and average queue length in single path CBRP. In the future, the Network Simulator (NS) will be used to evaluate the MP-CBRP.

REFERENCES

[1] M. Al Mojamed, "Integrating Mobile Ad Hoc Networks with the Internet Based on OLSR," *Wireless Communications and Mobile Computing*, vol. 2020, Oct. 2020, Art. no. e8810761, https://doi.org/10.1155/2020/8810761.

[2] V. K. Verma, A. Yadav, and T. Jain, *Applications of Mobile Ad Hoc Network: Usage of Mobile Ad Hoc Network*. LAP LAMBERT Academic Publishing, 2021.

[3] R. Thiagarajan, M. R. Babu, and M. Moorthi, "Quality of Service based Ad hoc On-demand Multipath Distance Vector Routing protocol in mobile ad hoc network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 5, pp. 4957–4965, May 2021, https://doi.org/10.1007/s12652-020-01935-x.

[4] S. Soni and J. S. Shah, "QoS frameworks for Multimedia Traffic in Mobile Adhoc Networks: A Comparative Review," *Engineering, Technology & Applied Science Research*, vol. 7, no. 3, pp. 1708–1712, Jun. 2017, https://doi.org/10.48084/etasr.1131.

[5] T.-C. Wan, "Performance Evaluation of Single-Path And Multipath MANETs Routing Protocols for Dense and Sparse Topology," *International Journal of Software Engineering and Computer Systems*, vol. 3, no. 1, pp. 31–42, Feb. 2017, https://doi.org/10.15282/ijsecs.3.2017.3.0025.

[6] Z. Hui, Z. Lingli, Y. Yonghang, and C. Linlin, "A Survey of Multipath Load Balancing Based on Network Stochastic Model in MANET," in *2021 23rd International Conference on Advanced Communication Technology (ICACT)*, PyeongChang, South Korea, Feb. 2021, pp. 336–341, https://doi.org/10.23919/ICACT51234.2021.9370843.

[7] M. A. Mahdi, A. Mahdi, M. A. G. Hazber, and M. Kachout, "Performance Evaluation of Single-Path and Multipath MANETs Routing Protocols using Random Mobility Model," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 10, no. 4, Apr. 2020, Art. no. p10040, https://doi.org/10.29322/IJSRP.10.04.2020.p10040.

[8] A. Bhardwaj and H. El-Ocla, "Multipath Routing Protocol Using Genetic Algorithm in Mobile Ad Hoc Networks," *IEEE Access*, vol. 8, pp. 177534–177548, 2020, https://doi.org/10.1109/ACCESS.2020.3027043.

[9] H.-H. Choi and J.-R. Lee, "Local Flooding-Based on-Demand Routing Protocol for Mobile Ad Hoc Networks," *IEEE Access*, vol. 7, pp. 85937–85948, 2019, https://doi.org/10.1109/ACCESS.2019.2923837.

[10] Z. Chen, W. Zhou, S. Wu, and L. Cheng, "An Adaptive on-Demand Multipath Routing Protocol With QoS Support for High-Speed MANET," *IEEE Access*, vol. 8, pp. 44760–44773, 2020, https://doi.org/10.1109/ACCESS.2020.2978582.

[11] D.-G. Zhang *et al.*, "A Multi-Path Routing Protocol Based on Link Lifetime and Energy Consumption Prediction for Mobile Edge Computing," *IEEE Access*, vol. 8, pp. 69058–69071, 2020, https://doi.org/10.1109/ACCESS.2020.2986078.

[12] J. Jin, S. Ahn, and H. Oh, "A multipath routing protocol based on bloom filter for multi-hop wireless networks," in *2015 International Conference on Information Networking (ICOIN)*, Cambodia, Jan. 2015, pp. 521–522, https://doi.org/10.1109/ICOIN.2015.7057960.

[13] S. A. Mostafa, A. Mustapha, A. A. Ramli, M. A. Jubair, M. H. Hassan, and A. H. Abbas, "Comparative Analysis to the Performance of Three Mobile Ad-Hoc Network Routing Protocols in Time-Critical Events of Search and Rescue Missions," in *Advances in Simulation and Digital Human Modeling*, Cham, 2021, pp. 117–123, https://doi.org/10.1007/978-3-030-51064-0_16.

[14] F. T. AL-Dhief, N. Sabri, M. S. Salim, S. Fouad, and S. A. Aljunid, "MANET Routing Protocols Evaluation: AODV, DSR and DSDV Perspective," *MATEC Web of Conferences*, vol. 150, 2018, Art. no. 06024, https://doi.org/10.1051/matecconf/201815006024.

[15] Y.-C. Hu, D. A. Maltz, and D. B. Johnson, "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4," Internet Engineering Task Force, Request for Comments RFC 4728, Feb. 2007. https://doi.org/10.17487/RFC4728.

[16] S. A. Almazok and B. Bilgehan, "A novel dynamic source routing (DSR) protocol based on minimum execution time scheduling and moth flame optimization (MET-MFO)," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, Oct. 2020, Art. no. 219, https://doi.org/10.1186/s13638-020-01802-5.

[17] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Proceedings WMCSA'99. Second IEEE Workshop on Mobile Computing Systems and Applications*, New Orleans, LA, USA, Feb. 1999, pp. 90–100, https://doi.org/10.1109/MCSA.1999.749281.

[18] K. Karthick and R. Asokan, "Mobility Aware Quality Enhanced Cluster Based Routing Protocol for Mobile Ad-Hoc Networks Using Hybrid Optimization Algorithm," *Wireless Personal Communications*, vol. 119, no. 4, pp. 3063–3087, Aug. 2021, https://doi.org/10.1007/s11277-021-08387-2.

[19] M. Zhang and P. H. J. Chong, "Performance Comparison of Flat and Cluster-Based Hierarchical Ad Hoc Routing with Entity and Group Mobility," in *2009 IEEE Wireless Communications and Networking Conference*, Budapest, Hungary, Apr. 2009, pp. 1–6, https://doi.org/10.1109/WCNC.2009.4917894.

[20] M. Jiang, J. Li, and Y. C. Tay, "Cluster based routing protocol (CBRP) functional specification (Internet-Draft)." IETF, 1998.

[21] C. H. Ng and S. Boon-Hee, *Queueing Modelling Fundamentals: With Applications in Communication Networks*, vol. 2. Wiley, 2007.

Quality of Recycled Aggregates and Compressive Strength of No-Fines Recycled-Aggregate Concrete

Shakeel Ahmed Dahri
Department of Civil Engineering
Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
engr.shakeelahmed9@gmail.com

Bashir Ahmed Memon
Department of Civil Engineering
Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
bashir_m@hotmail.com

Mahboob Oad
Department of Civil Engineering
Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
engrmahboob04@gmail.com

Riaz Bhanbhro
Department of Civil Engineering
Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
riaz@quest.edu.pk

Ihsan Ali Rahu
Department of Civil Engineering
Quaid-e-Awam University of Engineering, Science & Technology
Nawabshah, Pakistan
Ihsanali0999@gmail.com

Abstract- This research paper presents the laboratory investigations of the compressive strength of no-fines concrete made with demolished waste as coarse aggregates used in percentages from 20% to 100%. The basic properties of aggregates were determined. Sieve analysis of both conventional and recycled aggregates was conducted to ensure the existence of well-graded aggregates in concrete. Nine concrete mixes were designed with an aggregate-cement ratio of 4. Additionally, three batches were prepared (conventional, recycled, conventional no-fines concrete) and the results were compared. For all mixes, the water-cement ratio was equal to 0.5. In each batch, 5 cylinders of standard size (total 60 samples) were prepared and cured for 28 days. The weight of the specimens was determined and compressive strength was checked in a Universal Testing Machine under gradually increasing load. A decrease in weight and compressive strength was recorded for the batches of the proposed concrete. Results show that at 40% replacement level the loss of compressive strength is 19% and the weight reduction of the samples was equal to 9%.

Keywords-recycled aggregate concrete; no-fines concrete; compressive strength; demolishing waste; recycled coarse aggregates

I. INTRODUCTION

The demand to reduce the utilization of natural resources and energy and the generation of harmful gases, applies to almost all areas of life. It is reported that every year the concrete industry alone uses 20 billion tons of aggregates, 1.5 billion tons of cement and 800 million tons of water [1]. It also produces 5% - 7% of total global annual CO_2 emissions [1]. On the other hand, the generation of construction waste has increased drastically, making waste management a major issue.

Construction/demolishing waste is usually utilized in floors and plinth protection, but it generally goes to landfills. Unavailability of such spaces forces the industry to transport waste too far leading to increase the cost of the project, while dumbing it in nearby unsuitable areas may cause environmental issues. The conventional ingredients of concrete are believed to be no eco-friendly materials due to the processes involved in their production [2]. Therefore, Recycled Aggregate Concrete (RAC) is one of the solutions in dealing with waste on site and helping to minimize the associated issues to some extent. The use of demolishing waste in new concrete has been addressed by several scholars. Authors in [3, 4] reviewed the recent development on the use of recycled aggregates. Rules and regulations and their implementation regarding the process and use of the material will help to ease the problem. Relevant regulations have been made in many countries, but their implementation has several issues [5]. The use of waste as coarse aggregates up to 100% has been addressed in literature. However, the dialog on total replacement of conventional coarse aggregates is still open. Partial replacement is preferred, particularly to evaluate the optimum recycled waste dosage with respect to strength and durability [6]. The improvement of these and other properties of concrete have also been addressed by making use of different additives, i.e. waste fly ash, waste

---

Corresponding author: Mahboob Oad

perlite powder, waste cellular concrete [7], silica fume [8], ground granulated blast furnace slag [9], etc.

There are different properties of concrete at fresh and hardened state with varying dosages of recycled aggregates utilized in various ways [8, 10-13]. However, it is observed that the optimum reported dosages by several scholars differ in general and are situation-based in particular. Also, scattering in the results of the same parameters of aggregates or concrete is noticed. This requires more work on the use of recycled aggregates to reach a more unanimous conclusion. Different types of concretes used in the industry have been studied with regard to the use of indigenous materials or waste products. No-fines concrete is one of them. No-fines concrete, also known as porous, pervious, permeable, and cellular concrete, is a porous concrete obtained when fine aggregates are omitted. The most widespread applications of no-fines concrete include road pavements and surface treatments to permit water drainage. In relation to this application, the knowledge of the durability of no-fines concrete has been limited to the study of freezing and thawing, shrinkage, thermal expansion, capacity to percolate water through the cement paste, and abrasion resistance. In addition, no-fines concrete is lightweight and can be utilized in civil engineering applications to reduce the self-weight of the members. Its application in building structures is limited due to the requirement of reinforcement which disturbs the voids present in the concrete. Also, due to the voids in concrete the bonding strength is assumed to be affected. However, it may be used in non-load bearing members to reduce their self-weight. An overview in [14] summarizes the work devoted to the topic along with developing the correlation between various properties of concrete. While producing no fines concrete, gradation of aggregates, aggregate-cement (a/c) and water-cement (w/c) ratios may be chosen appropriately, particularly when the compressive strength is the major parameter of consideration. However, to a limited extent, unit weight and apparent texture also depend on these factors. The strength of concrete is a key property with respect to durability. Due to its porosity, the strength of no-fines concrete is less than conventional concrete's [15, 16]. Therefore, its use is limited to pathways and parking [16], road pavement with limited traffic [17, 18], and blocks instead of bricks [19, 20]. Several attempts have been conducted to improve the strength of no-fines concrete by using partial replacement (up to 20%) of coarse aggregates with fine aggregates [18, 21, 22], reduced w/c and a/c ratios [23-26], fibers [27-29], admixtures including superplasticizers [26, 30], and un-crushed aggregates [31], recycled coarse aggregates [32], or mixes of un-crushed and recycled aggregates [33]. Despite these efforts, still the strength of no-fines concrete is questionable.

Although much work is devoted to various properties of no-fines concrete, only a few studies address the use of demolishing waste in the development of no-fines concrete. The use of waste in no-fines concrete will help the sustainable improvement and lessen the waste management burden on a project. This study aims to investigate the effect of demolishing waste as coarse aggregates on the compressive strength of no-fines concrete. Demolished waste was used in no-fines concrete in percentages from 20% to 100% with increment of 10% and the produced concrete was studied and evaluated.

II. MATERIALS AND TESTING

In the current work, the demolished waste of a roof projection of single-story house was used (Figure 1). Recycled Coarse Aggregates (RCA) were produced by hammering the waste down to a maximum size of 25mm. After sorting the impurities the aggregates were washed and allowed to dry (Figure 2). The conventional coarse aggregates were purchased from the local market. The source was approved quarries of Jamshoro hills.

Fig. 1. Demolishing concrete waste.

Fig. 2. Recycled and conventional aggregates.

Ordinary Portland Cement (OPC) under the brand name Pak Land of grade ps-42.5N was purchased from the local market. The source of water used for washing and mixing was the local water supply. The water pH was equal to 6.9.

A. Basic Properties

Basic properties, i.e. water absorption, specific gravity, abrasion, and soundness of both types of coarse aggregates was determined following the standard procedures. The obtained results are listed in Table I.

B. Gradation of Coarse Aggregates

To achieve proper strength of the hardened concrete, good gradation of coarse aggregates is a major aspect. Therefore, the gradation of the aggregates was done following the standard sieve analysis procedure. The passing percentage of the aggregates on each sieve is compared and plotted in Figure 3. It can be observed that both types of aggregates, conventional and recycled, confirm the permissible limits.

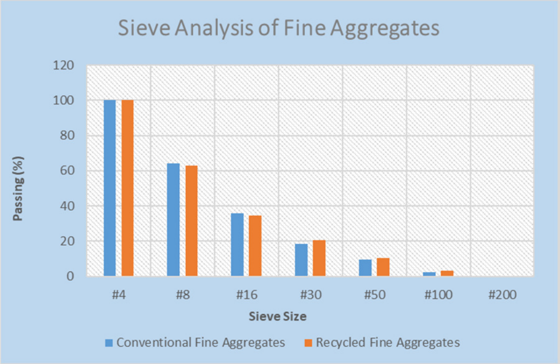


Fig. 3. Sieve analysis of coarse aggregates.

C. Specimen Preparation

To achieve the proposed aim of the research, 8 batches (B4 – B12) of the proposed no-fines concrete were prepared. The Conventional Aggregates (CA) were replaced with RCA in varying percentage from 20% to 100% with increment of 10%. In each batch, 5 standard size cylindrical specimens were cast with 1:4 cement/aggregate ratio. The w/c ratio used was equal to 0.5. A higher value of the parameter was adopted considering the higher water absorption of Recycled Aggregates (RA). Additionally, 3 batches (B1 – B3) were also cast to compare the results of the proposed concrete. The first batch (B1) was cast using conventional concrete (1:2:4 mix). In the second batch, 1:2:4 mix was used, but with all RCA. The third batch (B3) was No-fines conventional concrete with Conventional coarse Aggregates (NCA). The details of all the batches used for this research work along with the quantities of the materials used to cast the specimens in each batch are listed in Table II.

Batching of the ingredients was done using weight batching, followed by mixing the ingredients in a concrete mixer. The moulds were prepared in standard fashion and were filled in 3 layers. Compaction of the specimens of batch B1 and B2 was accomplished using a table vibrator whereas the compaction of no-fines concrete was done by rodding. After 24-hours the specimens were de-moulded and left to air dry for one day (Figure 4). The curing of the specimens was done by fully immersing them in potable water (Figure 5).

Fig. 4. Specimens.

Fig. 5. Curing of specimens.

TABLE I. BASIC PROPERTIES OF COARSE AGGREGATES

Property	CA	RCA	Difference (%)
Water absorption	1.18	4.92	316.95
Specific gravity	2.61	2.31	-11.49
Abrasion	23.7	30.2	27.4
Soundness	4.10	5.86	42.93

TABLE II. DETAILS OF CONCRETE MIXES

Batch	Description	Cement (kg)	FA (kg)	CA (kg)	RCA (kg)	Water (kg)	28 days curing
B1	Conventional concrete	8.807	19.571	36.7	--	3.96	5
B2	RCA	8.807	19.571	--	36.69	3.96	5
B3	NCA	12.33	--	51.4	--	5.54	5
B4	20% RCA and 80% CA	12.5	--	31.7	7.94	5.62	5
B5	30% RCA and 70% CA	12.5	--	27.8	11.9	5.62	5
B6	40% RCA and 60% CA	12.5	--	23.8	15.88	5.62	5
B7	50% RCA and 50% CA	12.5	--	19.9	19.91	5.62	5
B8	60% RCA and 40% CA	12.5	--	15.8	23.83	5.62	5
B9	70% RCA and 30% CA	12.5	--	11.9	27.80	5.62	5
B10	80% RCA and 20% CA	12.5	--	7.9	31.77	5.62	5
B11	90% RCA and 10% CA	12.5	--	3.9	35.74	5.62	5
B12	100% RCA	12.33	--	--	39.71	5.62	5

D. Specimen Testing

After the completion of curing time, the specimens were allowed to dry for 24 hours at room temperature. The weight of all the specimens was determined. The average weight of the specimens in each batch is listed in Table III. Each specimen was tested in a Universal Testing Machine (UTM) under gradually increasing crushing load (Figure 6). The load was recorded and used to compute the compressive strength. The average compressive strength of 5 samples of each batch was evaluated with regard to their weight (Table III).

III. RESULTS AND DISCUSSION

The basic properties of the aggregates presented in Table I show deviations of properties of the RCA from CA. The water absorption is about 3 times the water absorption of the CA. There was a 11% reduction in specific gravity, and 48% and

42% increase in the abrasion and soundness of the aggregates. The results of the parameters of RCA are better than the findings reported in [33-35]. However, the deviation from CA is mainly attributed to the mortar attached with the aggregates, the age and mix ratio of the old concrete, and the conditions to which it is exposed during service life. These factors affect the properties of the aggregates, thus need careful consideration or modified processing techniques to improve the performance of the aggregates.

Fig. 6. Specimen testing.

TABLE III. AVERAGE WEIGHT AND COMPRESSIVE STRENGTH

Batch	RCA (%)	CA (%)	Ratio	Weight (kg)	Compressive strength (MPa)
B1	0	100	1:2:4	13.62	30.86
B2	100	0	1:2:4	12.27	26.41
B3	0	100	1:4	11.61	20.43
B4	20	80	1:4	11.15	18.43
B5	30	70	1:4	10.54	17.40
B6	40	60	1:4	10.62	16.51
B7	50	50	1:4	10.25	15.42
B8	60	40	1:4	10.18	14.23
B9	70	30	1:4	10.11	12.46
B10	80	20	1:4	10.18	11.74
B11	90	10	1:4	10.00	10.35
B12	100	0	1:4	9.90	9.19

The weight of the samples of the proposed concrete are compared in Figures 7 to 9. Figure 7 shows the comparison with the weight of specimens of conventional (1:2:4) concrete. Decrease in the weight of specimens is observed with increase in the replacement level of aggregates. 18% to 27% reduction in the weight of the proposed specimens may be observed in comparison to the conventional (1:2:4) concrete, whereas, the comparison result to Recycled Aggregate (1:2:4) Concrete (RAC) ranged between 9% and 19%. When compared with the weight of conventional no-fines concrete the reduction is in the range of 4% to 15%. This shows that the absence of fine aggregates in the concrete mix results in reduced weight of the product leading to light-weight concrete. However, the strength parameter should also be checked before deciding the optimum percentage of replacement. Otherwise, a higher percentage of RA shows higher reduction in the weight of the product. The percentage difference of the weight of the proposed samples in comparison to B1, B2, and B3 samples is listed in Table IV.

The average compressive strength of the proposed concrete samples is compared with the compressive strength of all three control mixes in Figures 10 to 12. From these Figures it may be noted that the proposed concrete exhibited a reduction in comparison to all the control mixes. Lowest residual strength was recorded at the highest replacement level, and was 30%, 35%, and 45% in comparison to B1, B2, and B3 respectively. This shows that increasing dosage of the RA in no-fines concrete adversely affects the compressive strength of the product. Thus, when producing no fines concrete, the gradation of aggregates and the c/a and w/c ratios may also be chosen appropriately particularly since compressive strength is a major parameter of consideration [37]. The details of the difference of compressive strength are listed in Table V.

TABLE IV. WEIGHT DIFFERENCE PERCENTAGE

Batch	With B1	With B2	With B3
B1	--	9.91	14.77
B2	-9.91	--	5.39
B3	-14.77	-5.39	--
B4	-18.10	-9.09	-3.92
B5	-22.61	-14.10	-9.21
B6	-22.01	-13.43	-8.50
B7	-24.74	-16.47	-11.71
B8	-25.27	-17.05	-12.33
B9	-25.74	-17.57	-12.87
B10	-25.29	-17.07	-12.34
B11	-26.61	-18.54	-13.90
B12	-27.29	-19.30	-14.70

TABLE V. COMPRESSIVE STRENGTH DIFFERENCE PERCENTAGE

Batch	With B1	With B2	With B3
B1	--	16.87	51.03
B2	-14.44	--	29.22
B3	-33.79	-22.61	--
B4	-40.30	-30.22	-9.83
B5	-43.62	-34.11	-14.85
B6	-46.52	-37.50	-19.23
B7	-50.03	-41.60	-24.53
B8	-53.89	-46.11	-30.36
B9	-59.62	-52.81	-39.02
B10	-61.97	-55.55	-42.56
B11	-66.46	-60.80	-49.35
B12	-70.22	-65.19	-55.02

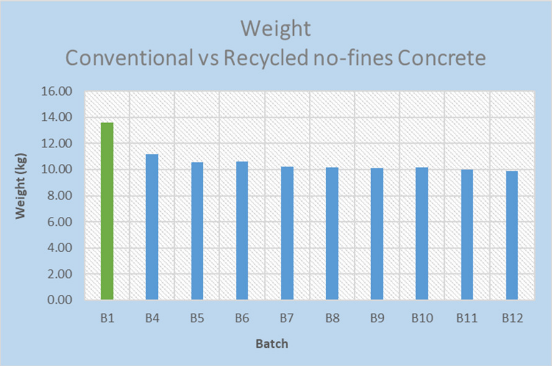


Fig. 7. Weight of B1 vs no-fines RAC.

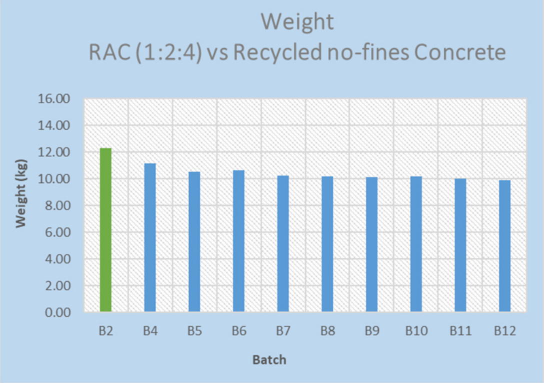


Fig. 8. Weight of B1 vs no-fines RAC.

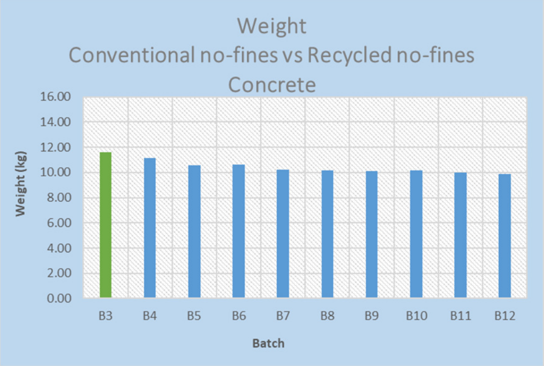


Fig. 9. Weight of B3 vs no-fines RAC.

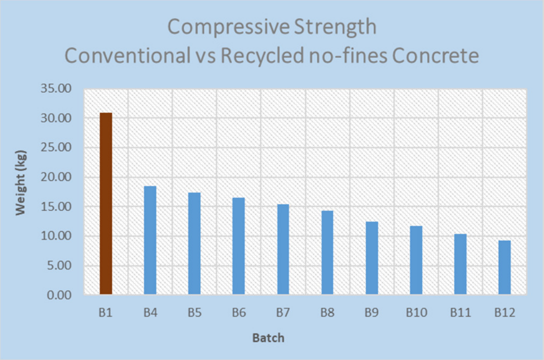


Fig. 10. Compressive strength vs B1.

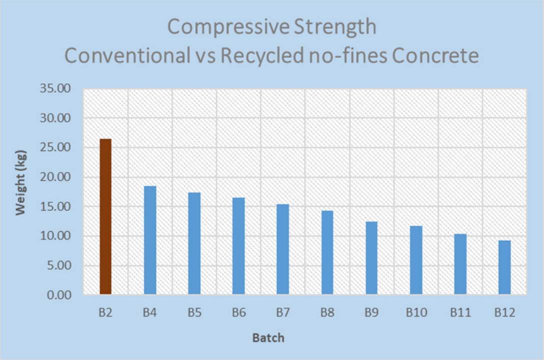


Fig. 11. Compressive strength vs B2

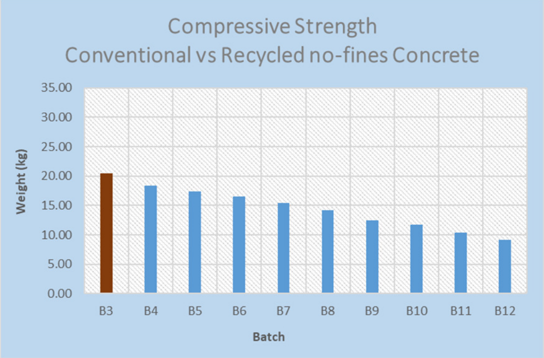


Fig. 12. Compressive strength vs B3.

IV. CONCLUSIONS

From the obtained results of this experimental study, the following conclusions can be drawn:

- Water absorption of recycled aggregates is 316% higher than that of conventional aggregates.
- Specific gravity of recycled aggregates is 11% lower than that of conventional aggregates.
- The abrasion loss of the aggregates is about 27%.
- Soundness of the aggregates is 43% less than the soundness of conventional aggregates.
- The dosage of recycled aggregates is directly proportional to the decrease in weight and compressive strength.
- In comparison to conventional concrete, 27% decrease in weight and 70% loss of compressive strength are recorded at 100% replacement level of conventional coarse aggregates.

Therefore, it may be concluded that recycled coarse aggregates from demolishing waste may be used in the production of the no-fines recycled aggregate concrete. However, considering the loss of weight and compressive strength, 40% replacement level is considered optimum as at this replacement level the decrease in weight and compressive strength is 9% and 19% respectively.

REFERENCES

[1] S. E. Shaji and J. S. Jithin, "Mechanical Properties of Concrete with Admixed Recycled Aggregate," *International Journal of Engineering Research & Technology*, vol. 6, no. 6, pp. 1–6, Apr. 2018.

[2] M. Oad, A. H. Buller, B. A. Memon, N. A. Memon, Z. A. Tunio, and M. A. Memon, "Effect of Water-Cement Ratio on Flexural Strength of RC Beams Made with Partial Replacement of Coarse Aggregates with Coarse Aggregates from Old Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3826–3831, Feb. 2019, https://doi.org/10.48084/etasr.2499.

[3] B. A. Memon and A. H. Buller, "Recent Development on Use of Demolished Concrete as Coarse Aggregates," *International Journal of Emerging Technology and Innovative Engineering*, vol. 2, no. 1, pp. 1–11, Jan. 2016, https://doi.org/10.13140/RG.2.2.34109.92648.

[4] K. P. Verian, W. Ashraf, and Y. Cao, "Properties of recycled concrete aggregate and their influence in new concrete production," *Resources, Conservation and Recycling*, vol. 133, pp. 30–49, Jun. 2018, https://doi.org/10.1016/j.resconrec.2018.02.005.

[5] W. Kubissa, R. Jaskulski, A. Koper, and J. Szpetulski, "Properties of Concretes with Natural Aggregate Improved by RCA Addition," *Procedia Engineering*, vol. 108, pp. 30–38, Jan. 2015, https://doi.org/10.1016/j.proeng.2015.06.116.

[6] C. Zheng, C. Lou, G. Du, X. Li, Z. Liu, and L. Li, "Mechanical properties of recycled concrete with demolished waste concrete aggregate and clay brick aggregate," *Results in Physics*, vol. 9, pp. 1317–1322, Jun. 2018, https://doi.org/10.1016/j.rinp.2018.04.061.

[7] M. Abed and R. Nemes, "Long-term durability of self-compacting high-performance concrete produced with waste materials," *Construction and Building Materials*, vol. 212, pp. 350–361, Jul. 2019, https://doi.org/10.1016/j.conbuildmat.2019.04.004.

[8] H. Sasanipour and F. Aslani, "Effect of specimen shape, silica fume, and curing age on durability properties of self-compacting concrete incorporating coarse recycled concrete aggregates," *Construction and Building Materials*, vol. 228, Dec. 2019, Art. no. 117054, https://doi.org/10.1016/j.conbuildmat.2019.117054.

[9] R. K. Majhi, A. N. Nayak, and B. B. Mukharjee, "Development of sustainable concrete using recycled coarse aggregate and ground granulated blast furnace slag," *Construction and Building Materials*, vol. 159, pp. 417–430, Jan. 2018, https://doi.org/10.1016/j.conbuildmat.2017.10.118.

[10] O. Damdelen, "Investigation of 30% recycled coarse aggregate content in sustainable concrete mixes," *Construction and Building Materials*, vol. 184, pp. 408–418, Sep. 2018, https://doi.org/10.1016/j.conbuildmat.2018.06.149.

[11] M. Hayles, L. F. M. Sanchez, and M. Noel, "Eco-efficient low cement recycled concrete aggregate mixtures for structural applications," *Construction and Building Materials*, vol. 169, pp. 724–732, Apr. 2018, https://doi.org/10.1016/j.conbuildmat.2018.02.127.

[12] F. Kazemian, H. Rooholamini, and A. Hassani, "Mechanical and fracture properties of concrete containing treated and untreated recycled concrete aggregates," *Construction and Building Materials*, vol. 209, pp. 690–700, Jun. 2019, https://doi.org/10.1016/j.conbuildmat.2019.03.179.

[13] S. Ismail, W. H. Kwan, and M. Ramli, "Mechanical strength and durability properties of concrete containing treated recycled concrete aggregates under different curing conditions," *Construction and Building Materials*, vol. 155, pp. 296–306, Nov. 2017, https://doi.org/10.1016/j.conbuildmat.2017.08.076.

[14] S. Ali and S. Kacha, "Correlation among properties of no fines concrete– A review," presented at the National Conference on Applications of Nano Technology in Civil Engineering, Vadodara, India, Feb. 2017.

[15] A. A. Yasin, "Investigation of Mechanical and Physical Properties of No-Fines Concrete," *International Journal of Engineering Innovation & Research*, vol. 4, no. 6, pp. 799–802, 2015.

[16] P. Mounika and K. Srinivas, "Mechanical Properties of no Fines Concrete for Pathways," *International Journal of Engineering and Techniques*, vol. 4, no. 2, pp. 68–81, 2018.

[17] A. R. Kalyane, A. Gujar, A. Jadhav, G. Agrawal, G. Mandora, and R. Mahajan, "Compressive Strength of No-Fines Concrete Compared to Conventional Concrete for Road Pavements," *International Journal of Advance Research in Science and Engineering*, vol. 6, no. 4, pp. 492–494, Apr. 2017.

[18] G. Ravish and V. K. Ahuja, "No-Fine Concrete as a Road Pavement," *International Journal of Engineering Trends and Technology*, vol. 24, no. 4, pp. 187–190, 2015.

[19] A. Malik, "An Experimental Study on Properties of No-Fines Concrete," *Imperial Journal of Interdisciplinary Research*, vol. 2, no. 10, pp. 2075–2079, 2016.

[20] K. S. Ushane, K. J. P. Kumar, and C. Kavitha, "Investigation of No-Fines Concrete in Building Blocks," *International Journal of Structural and Civil Engineering Research*, vol. 3, no. 4, pp. 170–177, 2014.

[21] M. A. Alam and S. Naz, "Experimental Study on Properties of No-fine Concrete," *International Journal of Informative and Futuristic Research*, vol. 2, no. 10, pp. 3687–3694, Jun. 2015.

[22] M. A. B. Kabir and Kh. Md. S. Islam, "Performance Study on No-Fines Concrete," in *3rd International Conference on Civil Engineering for Sustainable Development*, Khulna, Bangladesh, Feb. 2016, pp. 848–854.

[23] F. Tittarelli, A. Mobili, C. Giosue, and M. L. Ruello, "Sustainable and Durable No-fines Concrete for Vertical Applications," *International Journal of Chemical, Environmental & Biological Sciences*, vol. 1, no. 5, pp. 784–788, 2013.

[24] K. B. Thombre, A. B. More, S. R. Bhagat, and pvpit bhavdhan pune, "Investigation of Strength and Workability in No-Fines Concrete," *International Journal of Engineering Research & Technology*, vol. 5, no. 9, pp. 390–393, Sep. 2016, https://doi.org/10.17577/IJERTV5IS090206.

[25] L. JiaHao, F. C. Lian, F. Hejazi, and N. Azline, "Study of properties and strength of no-fines concrete," *IOP Conference Series: Earth and Environmental Science*, vol. 357, Nov. 2019, Art. no. 012009, https://doi.org/10.1088/1755-1315/357/1/012009.

[26] I. Gorgis, W. F. Abd, and S. Al-Mishhadani, "Acidic solution effects on no-fines concrete produced by using recycled concrete as coarse aggregate," *MATEC Web of Conferences*, vol. 162, 2018, Art. no. 02002, https://doi.org/10.1051/matecconf/201816202002.

[27] P. S. Patil, I. Sonar, and S. Shinde, "No Fine Concrete," *International Journal of Concrete Technology*, vol. 3, no. 2, pp. 18–30, Sep. 2017, https://doi.org/10.37628/ijct.v3i2.180.

[28] M. Mageswari, M. P. Karthikeyan, S. Pavithran, M. Rajkumar, and R. Govindarajan, "High Strength Permeable Pavement using no Fines Concrete," *International Journal of Civil Engineering*, vol. 3, no. 3, pp. 62–66, Mar. 2016, https://doi.org/10.14445/23488352/IJCE-V3I3P111.

[29] A. D. Ingle, A. R. Padmawar, T. R. Hirave, and D. Kumbhar, "No-Fines Concrete by using Cellulose Fiber," *International Journal of Innovative Research in Science and Engineering*, vol. 3, no. 4, pp. 253–259, 2017.

[30] F. Monika, H. Prayuda, B. Zega, and M. D. Cahyati, "Flexural and Compressive Strength on No-fines Concrete Slab using Variations of Fly Ash and Superplasticizer," *International Journal of Integrated Engineering*, vol. 11, no. 9, pp. 275–284, Dec. 2019, https://doi.org/10.30880/ijscet.00.00.0000.00.0000.

[31] M. A. Memon, M. A. Bhutto, N. A. Lakho, I. A. Halepoto, and A. N. Memon, "Effects of Uncrushed Aggregate on the Mechanical Properties of No-Fines Concrete," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2882–2886, Jun. 2018, https://doi.org/10.48084/etasr.1976.

[32] A. M. Bravo-German, I. D. Bravo-Gomez, J. A. Mesa, and A. Maury-Ramirez, "Mechanical Properties of Concrete Using Recycled Aggregates Obtained from Old Paving Stones," *Sustainability*, vol. 13, no. 6, Jan. 2021, Art. no. 3044, https://doi.org/10.3390/su13063044.

[33] S. A. Salih, I. N. Gorgis, and W. F. Abd, "Some Properties of No-Fines Concrete Produced by Using Demolished Concrete as Recycled Coarse Aggregate," *Engineering and Technology Journal*, vol. 35, no. 7, pp. 741–748, 2017.

[34] G.-F. Peng, Y.-Z. Huang, H.-S. Wang, J.-F. Zhang, and Q.-B. Liu, "Mechanical Properties of Recycled Aggregate Concrete at Low and High Water/Binder Ratios," *Advances in Materials Science and Engineering*, vol. 2013, Dec. 2013, Art. no. e842929, https://doi.org/10.1155/2013/842929.

[35] Z. A. Tunio, B. A. Memon, N. A. Memon, N. A. Lakho, M. Oad, and A. H. Buller, "Effect of Coarse Aggregate Gradation and Water-Cement Ratio on Unit Weight and Compressive Strength of No-fines Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3786–3789, Feb. 2019, https://doi.org/10.48084/etasr.2509.

Production of Light Weight Foam Concrete with Sustainable Materials

Anowar Wakaa Ali
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
a.ali1901m@coeng.uobghdad.edu.iq

Nada Mahdi Fawzi
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
nada.aljalawi@coeng.uobaghdad.edu.iq

Abstract- Most of the recent works related to the construction industry in Iraq are focused on investigating the validity of local raw materials as alternatives to the imported materials necessary for some practical applications, especially in thermal and sound insulation. This investigation includes the use of limestone dust as partial substitution of cement in combination with foam agent and silica fume to produce sustainable Lightweight Foam Concrete (LWFC). This study consists of two stages. In the first stage, trial mixes were performed to find the optimum dosage of foam agent. Limestone dust was used as a partial replacement for cement. Chemical analysis and fineness showed great similarity with cement. Many concrete mixes were prepared with the content of lime dust powder being 10%, 14%, and 18% as partial replacement of cement by weight. The results indicate that the compressive strength at 7, 28, and 90 days of age was increased for specimens with 14% limestone dust. The best results in compressive strength show an increase at 7 days and a decrease at 28 and 90 days for concrete specimens with 14% limestone dust. In addition, the results show a decrease in dry density for concrete containing 14% lime dust. In the second stage, different percentages of Polypropylene Fibers were added to the concrete, all mixes, containing a constant content of limestone dust of 14% by weight of cement, were modified using different percentages of Polypropylene Fibers (1%, 1.5 %, and 2% by volume) and the best percentage was found to be 1%. The addition of Polypropylene Fibers enhances splitting tensile and flexural strength at 28 days by 14.55% and 55% respectively.

Keywords-lightweight concrete; foam concrete; sustainable materials; polypropylene fibers ;limestone dust

I. INTRODUCTION

One of the disadvantages of conventional concrete is its high self- weight. The density of normal concrete is in the order of 2200 to 2600kg/m^3. This heavy self-weight makes it, to some extent, an uneconomical structural material. Attempts have been made in the past to reduce the self-weight of concrete in order to increase its efficiency. Light-Weight Concrete (LWC) is a concrete whose density varies from 300 to 1850kg/m^3 [1]. There are many advantages of having low density. It helps in the reduction of dead load, increases the progress of building, and lowers haulage and handling costs. The weight of a building on the foundation is an important factor in design, particularly in the case of weak soil and tall structures. In framed structures, the beams and columns have to carry the load of floors and walls. If floors and walls are made of LWC, this will result in considerable cost decrease. Another important characteristic of LWC is its relatively low thermal conductivity, a property which improves with decreasing density. In extreme climatic conditions and in the case of buildings where air-conditioning is to be installed the use of LWC with low thermal conductivity will produce considerable advantages from the points of thermal comfort and lower power consumption [2]. On the other hand, the cost of natural resources is increasing constantly, leading to the search for alternatives, such as recycled materials like rice husk ash, sawdust ash [3], silica fume, fly ash, coal bottom ash, limestone dust, marble powder, tile powder, millet husk ash, crumb rubber, etc. Also, ordinary Portland Cement (OPC), has been related to several diseases [4-5].

Basically, there are three methods for making LWC. The first way is by using porous lightweight aggregates of low apparent specific gravity. This type of concrete is known as lightweight aggregate concrete. The second way is by introducing large voids within the concrete or mortar mass. These voids should be clearly distinguished from the extremely fine voids produced by air entrainment. This type of concrete is variously knows as aerated, cellular, foamed, or gas concrete. Thirdly, by omitting the fine aggregates from the mix in such a way that a large number of interstitial voids are present. Normal-weight coarse aggregates are generally used. This concrete is called no-fines concrete [2]. Based on the purpose of use, there are three types of concrete. The first type is known as Non-Structural LWC. This is employed mainly for insulation purposes. With low unit weight, seldom exceeding 800kg/m^3, its heat insulation value is high and compressive strength is low, from 0.69 to 6.89N/mm^2 [6]. The second type is Moderate Strength LWC. The use of this concrete requires a fair degree of compressive strength, and thus they fall about midway between the structural and low density concrete. This is sometimes designed as "fill" concrete. Compressive strength varies approximately from 6.89 to 17.24N/mm^2, the insulation values are intermediate, and its density is between 800 to 1300kg/m^3 [7]. The third type is known as Structural LWC. This type can be defined as a concrete with full structural efficiency containing aggregates which fall on the other end of

---

Corresponding author: Anowar Wakaa Ali

the scale. Its minimum compressive strength is 17.24N/mm² [8]. Most of structural LWCs are capable to produce concrete with excessive compressive strength. However, thermal insulation values for structural LWC are substantially better than NWC's. Since the unit weight of structural LWC is considerably greater than that of low density concrete, the insulation efficiency is lower [9]. The compressive strength of insulating LWC increases with increasing density and water-cement ratio [10]. As for polypropylene fibers, research has shown that they increase concrete's resistance to impact loads [11].

II. MATERIALS AND MIX PROPORTIONS

A. Cement

OPC type I, manufactured by Al Kubaisa Cement Factory, was used in 3.1.1 cement in all the mixes throughout this investigation. The chemical and physical properties of this cement are presented in Tables I and II respectively. The test results indicate that the adopted cement, type N 42.5MPa conforms to the Iraqi specification No. 5/2019 [12].

TABLE I. CHEMICAL COMPOSITION OF CEMENT

Oxide composition	Result	Limits of [12]
Lime (CaO)	62	-
Silica (SiO_2)	20.1	-
Alumina (Al_2O_3)	4.24	-
Iron oxide (Fe_2O_3)	4.16	-
Sulfate (SO_3)	2.15	≤ 2.8% if $C_3A>3.5$ ≤ 2.5% if $C_3A<3.5$
Magnesia (MgO)	3.65	≤ 5%
Loss on ignition (LOI.)	3.42	≤ 4%
Insoluble residue (IR)	0.71	≤ 1.5%
Main compounds (Bogues eq.) of cement		
Tricalcium silicate (C_3S)		59.02
Dicalcium silicate (C_2S)		29.65
Tricalcium aluminate (C_3A)		4.21
Tetracalcium aluminoferrite (C_4AF)		12.65

TABLE II. PHYSICAL PROPERTIES OF CEMENT

Physical properties	Test result	Limits of [12]
Specific surface area, Blaine method, (m²/kg).	295	≥ 250 m^2/kg
Setting time -Initial setting (min) -Final setting (min)	 1:38 3:45	 ≥ 45 min ≤ 600 min
Compressive strength of mortar (MPa) 2-days 28-days	 20.4 27.5	 ≥ 10 N/m^2 ≥ 42.5 N/m^2
Soundness % (autoclave)	0.35	≤ 0.8

TABLE III. GRADING OF FINE AGGREGATES

Sieve size (mm)	Passing % of sand	Limits of [13]
10	100	100
4.75	100	90-100
2.36	100	85-100
1.18	100	75-100
0.6	67.57	60-79
0.3	36.73	12-40
0.15	4.33	0-10

B. Fine Aggregates

Natural sand supplied from Al-Ekhadir quarry was used in the concrete mixes. The physical and chemical properties of fine aggregates are shown in Tables III and IV. The test results indicate that the sand grading is within the limits specified by the Iraqi Standard No. 45/1984 [13] and lies in Zone 4.

TABLE IV. PHYSICAL AND CHEMICAL PROPERTIES OF FINE AGGREGATES

Properties of sand	Test results	Limits of [13]
Fineness modulus	2.16	-
Specific gravity	2.61	-
Absorption	0.8%	-
SO_3	0.23 %	≤ 0.5%
Dry rodded density	163.7 kg/m³	-
Materials finer than 75μm %	2.6	≤ 5.0

C. Foaming Agent

Foaming agent for cellular concrete ASTM C796-97 [14] from Dr. Foamcrete was used to produce LWC by entraining a controlled amount of air bubbles to concrete mix. Table V indicates the technical description of the foaming agent used throughout this investigation.

TABLE V. TECHNICAL DESCRIPTION OF FOAMING AGENT (DATA TAKEN FROM MANUFACTURE PRODUCTION REPORT)

Appearance	Liquid
Color	Yellow
Specific gravity	1.1
Nitrate content	Nil
Compatibility with cement	All types of Portland cement
Point	Below 0°C
Flash point	Water base
LOI	3.5

D. Mixing Water

Water was used for mixing and curing of all concrete mixes conforming to standard Iraqi specification 1703 [15].

E. Fibers

Polypropylene fibers bought from Belgian Fibers Manufacturing were used for reinforcement. The specifications and properties of the fibers used throughout the experimental work are illustrated in Table VI according to the supplier.

TABLE VI. TECHNICAL PROPERTIES OF POLYPROPYLENE FIBERS

Length of Fiber	6 mm
Diameter	34μm
Density	0.91gr/cm^3
Young modulus	3750MPa
Melting point	>164C°
Elongation	200%
Spinning oil	1%
Moisture	Max 3%
Tenacity	30Cn/tex
Alkali resistance	Excellent
Conductivity	Very low

F. Silica Fume

Silica fume commercially known as Mega Add MS(D) from the chemical company Conmix was used as partial replacement of cement throughout this investigation. Tables VII and VIII illustrate its physical and chemical properties. The results show that the used silica fume satisfies the requirements of [16].

TABLE VII. PHYSICAL ANALYSIS OF SILICA FUME

Physical Properties	**Results**	**Requirements of [16]**
Color	Grey	-
Percent retained 45µm (No.325)	10	≤ 10
Bulk density (kg/m3)	500 to 700 kg/m3	
Specific surface are (m²/gm)	>15	> 15
Specific gravity	2.2	> 105

TABLE VIII. CHEMICAL COMPOSITION OF SILICA FUME

Oxides	**Oxide content**	**Requirements of [16]**
Sio_2	89	Min. 85%
Al_2o_3	0.4	-
Fe_2o_3	1.2	-
Mgo	2.5	-
Cao	1.4	-
So_3	1.0	-
Na_2o	1.2	-

G. Limestone Dust

The limestone dust (Figure 1) was obtained from a limestone quarry in Karbal. It is finely ground in the form of a dust, most of it passing sieve No.100. The chemical composition of the lime dust conforming to [17] is given in Table IX.

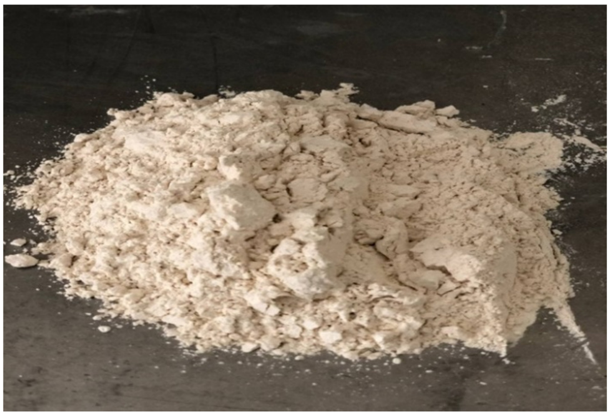


Fig. 1. The limestone dust used in the current research.

TABLE IX. CHEMICAL ANALYSIS OF LIMESTONE DUST

Chemical composition	**Constituent(%)**	**IQS 807 (%)**
Silicon dioxide (SiO_2)	5.0	5 (max)
Calcium oxide (CAO)	52.2	80 (max)
Aluminium oxide (Al_2O_3)	0.66	5 (max)
Ferric oxide (Fe_2O_3)	0.08	5 (max)
Magnesium oxide (MgO)	0.51	5 (max)
Sulphur oxide (SO_3)	0.0	-
LOI	41.49	-

H. High-Range Water Reducing Admixture(HRWRA)

The HRWRA Master Glenium 51 produced by BASF Company, satisfying the ASTM C 494/C 494M-17 [18] was used. The type F was used as concrete superplasticizer. The technical data of this type of admixture are shown in Table X.

TABLE X. TECHNICAL DESCRIPTION OF THE SUPERPLASTICIZER

Properties	**Technical description**
Appearance	Viscous liquid.
Color	Light brown
Density	1.1kg/lt.
pH	6.6

III. MIXING, CASTING, AND CURING

The mixing procedure is important in order to obtain the required workability and homogeneity. Mixing was performed with an electric mixer. The fine aggregates used were in Saturated Surface Dry (SSD) condition. The mixing sequence was the following: the required quantity of dry cement was added to the fine aggregates, the mixing continued for one minute, then the two thirds of the required quantity of water were added to the dry materials. The remaining water and the required quantity of foaming agent were mixed. Drill was used and in the head of the drill it was put the so-called mesh or paddle which rotates at high speed, 14000rpm, with a power supply of 220-240V/50Hz. Then the foam agent was added to mix. The mixer was stopped when a good homogeneous mix was produced. Before casting, the mould sides and the base were oiled slightly to prevent mortar sticking to the surfaces. The casting was carried out to two layers. Each layer was compacted by using a large vibrating table. The specimens were left under polyethylene sheets in the laboratory for one day after casting, then they were removed from the molds and kept in closed polyethylene bags for 7, 28, and 90 days until testing time.

TABLE XI. MIX SYMBOLS

Symbols	**Type**	**Content**
MR	Cube	Mortar + foam agent+ silica fume+ HRWRA
ML1	Cube	Mortar + foam agent+ silica fume+ HRWRA + lime dust 10%
ML2	Cube	Mortar + foam agent+ silica fume+ HRWRA + lime dust 14%
ML3	Cube	Mortar + foam agent+ silica fume+ HRWRA + lime dust 18%
MF1	Prism	Mortar + foam agent+ silica fume+ HRWRA + lime dust 14% + polypropylene fiber 1%
MF2	Prism	Mortar + foam agent+ silica fume+ HRWRA + lime dust 14% + polypropylene fibe1.5%
MF3	Prism	Mortar + foam agent+ silica fume+ HRWRA + lime dust 14% + polypropylene fiber 2%
MF4	Cylinder	Mortar + foam agent+ silica fume+ HRWRA + lime dust 14% + polypropylene fiber 1%

A. Mix Proportions

Several mixes were tried in order to find the best mixing of cement to sand, and several ratios were used (1:2.25,1:2,1:3,1:1.5,1:2.5). The optimal ratio of 1:2 was used. Experimental mixtures were then made by adding different amounts of foam agent to the water allocated to them (1, 3, and

5ml) where it was found that adding 3ml to the water gives the required consistency of the mortar. A reference mixture was made as described in Table XII, three limestone dust ratios (10%, 14%, and 18%) were used. The best replacement rate of the weight of cement was found to be 14%. This percentage was used in a mixture in which three different ratios of polypropylene fiber (1%, 1.5%, 2%) were used. It was found that best fiber percentage was (1%). 10% silica fume was used as a partial replacement of cement and 1% of HRWRA.

TABLE XII. DETAILS OF CONCRETE MIXES

Mixes	Cement (Kg/m^3)	Sand (Kg/m^3)	w/b	Silica fume (gm)	Foam agent (ml)	HRWRA (ml)	Fiber (gm)	Lime dust (gm)
MR	300	600	0.4	30	3	1.5	-	-
ML1	300	600	0.4	30	3	1.5	-	30
ML2	300	600	0.4	30	3	1.5	-	42
ML3	300	600	0.4	30	3	1.5	-	54
MF1	300	600	0.4	30	3	1.5	17	42
MF2	300	600	0.4	30	3	1.5	26	42
MF3	300	600	0.4	30	3	1.5	34	42
MF4	300	600	0.4	30	3	1.5	17	42

IV. RESULTS AND DISCUSSION

A. *Determination of the Consistency of Mortar (Flow Test)*

The amount of water required for standard consistency of plain mortar or foamed mortar was determined by the flow test. The flow table was prepared according to [19]. The amount of mixing water was sufficient to produce a flow of 110±5mm expressed as a percentage of the original base diameter of the flow mould. The required amounts of water for standard consistency of mortars, expressed in terms of w/c ratio, are shown in Table XIII.

TABLE XIII. DETAILS OF THE MORTAR MIXES WHEN USING 1:2 MIX

Foaming Agent (ml)	w/c	Flow 110 ± 5mm
1	0.4	106
3	0.4	112
5	0.4	117

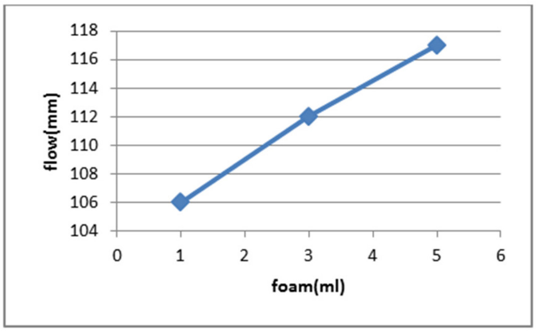


Fig. 2. The relationship between foam and flow.

B. *Compressive Strength Test*

This test was conducted on 50×50×50mm cubes using an electrical testing machine with a capacity of 2000KN at fixed load according to [20] .The average of three cubes was recorded for each test conducted at ages of 7, 28, and 90 days as shown in Table XIV**.**

TABLE XIV. COMPRESSIVE STRENGTH RESULTS

Mixes	Compressive Strength (MPa)		
	Age (days)		
	7	28	90
MR	10.45	15.77	20.32
ML1	8.95	12.32	19.93
ML2	10.73	20.86	25.58
ML3	7.41	12.71	17.14

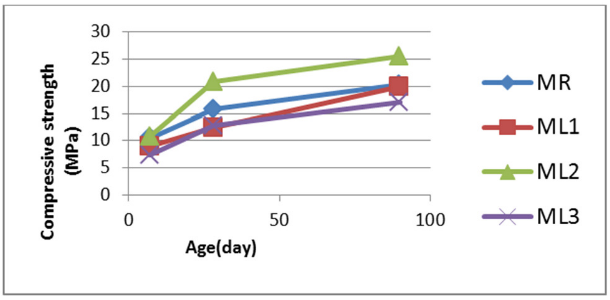


Fig. 3. Compressive strength at 7, 28, and 90 days.

C. *Dry Density Test*

Dry density was determined from the dried weight (105°C for 24 hours) and the measured volume by a ruler. Three cubes were measured in each tested sample. The density was found by weighting the specimens and dividing the weight by their measured volume. The dry density was tested at 7, 28, and 90 days as shown in Table XV.

TABLE XV. DRY DENSITY RESULTS

Mixes	Dry density (kg/m^3)		
	Age (days)		
	7	28	90
MR	1484	1623	1730
ML1	1478	1601	1679
ML2	1460	1572	1722
ML3	1354	1429	1620

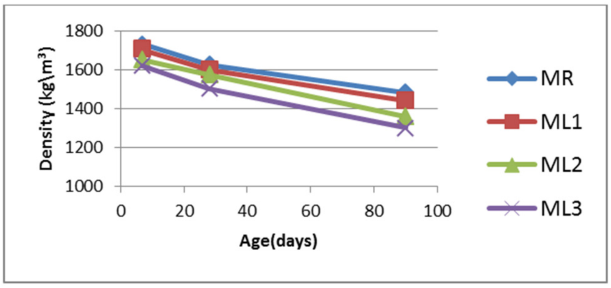


Fig. 4. Density-age relationship.

D. *Splitting Tensile Strength Test*

Splitting tensile strength test was performed according to [21] in which 100×200mm cylindrical specimens were used. They were tested using an electrical testing machine with a capacity of 2000KN. This test was conducted at 7, 28, and 90 days as shown in Table XVI.

TABLE XVI. SPLITTING TENSILE STRENGTH RESULTS

Mix	Splitting tensile strength (MPa)		
	Age (days)		
	7	28	90
ML2	0.82	1.65	2.04
ML4	1.01	1.89	2.21

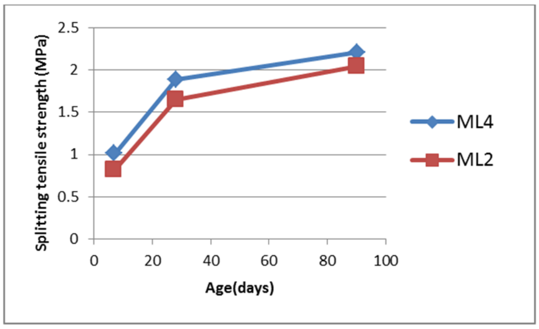


Fig. 5. Splitting tensile strength at 7, 28, and 90 days.

E. *Flexural Strength Test*

The flexural strength was determined according to [22] by using the center point method. The prism-shaped specimens with dimensions of 70×70×380mm were simply supported with 300mm span and were tested at the age of 7, 28, and 90 days as shown in Table XVII.

TABLE XVII. FLEXURAL STRENGTH RESULTS

Mix	Flexural Strength(MPa)		
	Age (days)		
	7	28	90
ML2	1.6	2.45	2.94
MF1	3.1	3.8	4.77
MF2	2.33	2.7	3.69
MF3	1.75	2.34	2.83

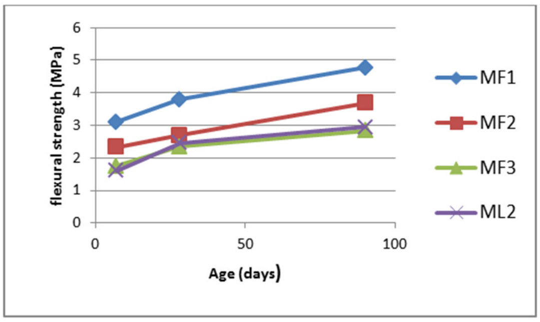


Fig. 6. Flexural strength at 7, 28, and 90 days.

V. V. CONCLUSION

Based on the results obtained from the experimental investigation, the following conclusions can be drawn:

- The water requirement for LWC is decreased as the percentage of foaming agent increases and the maximum decrease in water for all concrete mixes was 0.4 as a percentage of cement weight.
- LWC with 3ml foam agent gave better results of density and compressive strength.
- Compressive strength, flexural strength, and splitting tensile strength of LWC with limestone dust is slightly lower than those of normal concrete at all ages.
- The added foam leads to an increase in the flow. Three amounts of foam were added (1, 3, 5ml) and the flow rate was 106, 112, 117mm respectively. Three percentages of foam (10%, 14%, and 16%) were added to concrete as partial replacement of cement to produce sustainable LWC.
- Flexural strength and splitting tensile strength of LWC with 1% polypropylene fiber are higher than those of LWC without fiber at all ages.
- The best results were for specimens of LWC having a combination of polypropylene fibers (1%) and limestone dust (14%) due to the further physical effect of micro steel fibers in addition to the effect of limestone dust.

The results of this research confirm that we can use limestone dust to produce green LWC and enhance its properties with polypropylene fibers.

REFERENCES

[1] F. S. Fulton, *Concrete Technology*, 7th ed. The Portland Cement Institute, 1994.

[2] M. S. Shetty and A. K. jain, "Chapter 12," in *Concrete Technology: Theory and Practice*, 8th ed., S. Chand Publishing, 2019.

[3] N. Bheel, A. W. Abro, I. A. Shar, A. A. Dayo, S. Shaikh, and Z. H. Shaikh, "Use of Rice Husk Ash as Cementitious Material in Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4209–4212, Jun. 2019, https://doi.org/10.48084/etasr.2746.

[4] N. Bheel, R. A. Abbasi, S. Sohu, S. A. Abbasi, A. W. Abro, and Z. H. Shaikh, "Effect of Tile Powder Used as a Cementitious Material on the Mechanical Properties of Concrete," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4596–4599, Oct. 2019, https://doi.org/10.48084/etasr.2994.

[5] B. H. A. Bakar, A. T. Noaman, and H. M. Akil, "Cumulative Effect of Crumb Rubber and Steel Fiber on the Flexural Toughness of Concrete," *Engineering, Technology & Applied Science Research*, vol. 7, no. 1, pp. 1345–1352, Feb. 2017, https://doi.org/10.48084/etasr.854.

[6] C09 Committee, "ASTM C321-17: Specification for Lightweight Aggregates for Insulating Concrete," ASTM International. https://doi.org/10.1520/C0332-17.

[7] C09 Committee, "ASTM C331-17: Specification for Lightweight Aggregates for Concrete Masonry Units," ASTM International. https://doi.org/10.1520/C0331_C0331M-17.

[8] C09 Committee, "ASTM C330/C330M - 17a: Specification for Lightweight Aggregates for Structural Concrete," ASTM International. https://doi.org/10.1520/C0330_C0330M-17A.

[9] H. J. Goodman, "Low-density concrete," in *Fulton's Concrete Technology*, 7th ed., B. J. Addis, Ed. Portland Cement Institute, 1994, pp. 281–285.

[10] N. M. Fawzi, "Properties of lightweight Concrete with a view to thermal insulation and acoustic impedance," M.S. thesis, University of Baghdad, Baghdad, Iraq, 1997.

[11] N. Aljalawi, "Behavior of polypropylene fiber reinforced concrete under impact loads and explosive," *Journal of the College of Engineering of Baghdad University*, vol. 11, no. 3, pp. 59–68, Sep. 2005.

[12] *Iraqi Standard No. 5: Portland Cement*. Baghdad, Iraq: Central Organizationfor Standardization and Quality Control, 2019.

[13] *Iraqi Specification, No. 45 Aggregate from Natural Sources for Concrete and Construction.* Baghdad, Iraq, 1984.

[14] C09 Committee, "ASTM C796 - 97: Standard Test Method for Foaming Agents for Use in Producing Cellular Concrete Using Preformed Foam," ASTM International. https://doi.org/10.1520/C0796-97.

[15] *Iraqi Specification, No .1703: Water used in concrete*. Baghdad, Iraq: Central Organization for Standardization and Quality Control, 1992.

[16] C09 Committee, "ASTM C1240 - 03a: Specification for Silica Fume Used in Cementitious Mixtures," ASTM International. https://doi.org/10.1520/C1240-03A.

[17] *Iraqi Standard No. 807: Lime used in construction and in the production of building materials*. Baghdad, Iraq: Central Organization for Standardization and Quality Control, 1982.

[18] C09 Committee, "ASTM C494 / C494M - 17: Standard Specification for Chemical Admixtures for Concrete," ASTM International. https://doi.org/10.1520/C0494_C0494M-17.

[19] C01 Committee, "ASTM C230 / C230M - 21: Standard Specification for Flow Table for Use in Tests of Hydraulic Cement," ASTM International. https://doi.org/10.1520/C0230_C0230M-21.

[20] C01 Committee, "ASTM C109 / C109M - 16a: Standard Test Method for Compressive Strength of Hydraulic Cement Mortars (Using 2-in. or [50-mm] Cube Specimens)," ASTM International. https://doi.org/10.1520/C0109_C0109M-16A.

[21] C09 Committee, "ASTM C496 / C496M - 11: Standard Test Method for Splitting Tensile Strength of Cylindrical Concrete Specimens," ASTM International. https://doi.org/10.1520/C0496_C0496M-11.

[22] C09 Committee, "ASTM C293 / C293M - 16: Standard Test Method for Flexural Strength of Concrete (Using Simple Beam With Center-Point Loading)," ASTM International. https://doi.org/10.1520/C0293_C0293M-16.

Effects of Shaft Grouting on the Bearing Behavior of Barrette Piles: A Case Study in Ho Chi Minh City

Phu-Huan Vo Nguyen
Infrastructural Technique Department
Faculty of Civil Engineering
Ho Chi Minh City Open University
Ho Chi Minh City, Vietnam
huan.vnp@ou.edu.vn

Phu-Cuong Nguyen
Advanced Structural Engineering Laboratory
Structural Engineering Department
Faculty of Civil Engineering
Ho Chi Minh City Open University
Ho Chi Minh City, Vietnam
cuong.pn@ou.edu.vn

Abstract—The shaft-grouted method has been applied on high-rise buildings in Ho Chi Minh City for the purpose of increasing the bearing capacity of barrette piles. The Exim Bank Building foundation, using two kinds of shaft-grouted barrette piles, was 65m (TP1) and 85m (TP2) in depth. To assess the bearing capacity, this project assembly used the O-cell tools installed at 49m depth below the pile head level. Shaft grouting was performed from -25m to the TP1 pile toe level and -65m to the TP2 pile toe level. This work is based on the data from the O-cell experiments at the construction site and the results of finite element simulation in Plaxis software. The effectiveness of shaft grouting was analyzed and the length and position of the ejector were evaluated and compared in order to find the best solution for applying shaft grouting with the aim to ensure safety and mitigate economic problems.

Keywords-bearing capacity; barrette pile; shaft grouting; O-cell test; Plaxis

I. INTRODUCTION

High-rise building foundations in Ho Chi Minh City are constructed by using bored piles and barrette piles to ensure the total loads of structures. The shaft resistance using the grouting method has been reported in [1-3, 7, 8, 11]. Authors in [7] carried out laboratory tests in sand to find the behavior of density, soil gradation, and stress on the shaft resistance in grouting methods. The increase of shaft resistance was indicated by the low mobility grout. Authors in [2, 3] studied the bored piles in Ho Chi Minh City. In the results, the shaft resistance development in bored piles gains maximum resistance after a movement of only 3mm to 4mm. The Osterberg test (O-cell test) in principle is exactly the same as the static compression test, and it is dedicated to bored piles and barrettes. The testing principle is to apply the load directly on the pile tip or the pile body using a device called Osterberg cell. It is quite possible to use the pile load, the lateral soil friction, and the tip resistance as counterweights to increase the load. The Osterberg test method will give a result in which the two curves of load and displacement at the pile tip and pile head are built independently. The purpose of this paper is to evaluate the effectiveness of shaft grouting method on the barrette piles in Ho Chi Minh City and compare it with the no shaft-grouted method in order to find the best position to install the ejector using the Finite Element Method (FEM).

II. PROJECT INFORMATION

The Eximbank building is located at District 1, Ho Chi Minh City. It consists of 5 basements and 40 storys with a construction area of 3518m^2. Its design adopts 2 barrette test piles of 2800×800mm size and 15MN load [3]. The TP1 pile is 65.3m in depth near the borehole no. 2. The TP2 pile is 85.3m in depth near the borehole no. 8. Their designed to be grouted lengths are 40m for TP1 and 21m for TP2 respectively from their upward tips. The geological survey report states the soil condition at the test piles as follows: soft clay stratum at a depth of 7m, sandy stratum at the tips at 41m, mixed clay at 52m and finally a dense and very dense sandy stratum. O-Cell hydraulic jack is installed 16m from the pile tip and strainometers are installed along its body as shown in Figure 4 of [3]. Figure 1 shows how to install the equipment on the cross-sectional area of the two piles in which the main 32-dia steel bar of each pile is joined with the loop reinforcement to its frame. The total main steel bar area of the pile is 289cm^2 and the pile cross-sectional area is 2.24m^2. The grout pipe with a diameter of 60mm is mounted around the perimeter of the steel frame throughout the length of the pile and 39.7m and 21.2m lower than the lengths of TP1 and TP2 respectively. The grout pipe must be perforated to allow the grout to escape and it can be capped to ensure a secure grouting.

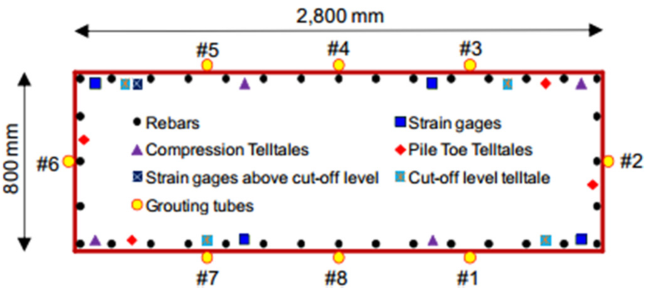


Fig. 1. Cross section and details of installation equipment for piles.

Corresponding author: Phu-Huan Vo Nguyen

III. PILE LOAD-BEARING ANALYSIS

A. *Test Results in Site*

The results of the force-displacement relationship curve of the O-cell test performed on the two test piles TP1 and TP2 can be seen in Figure 7 of [3]. The force acting on the O-cell is evaluated and adjusted subjecting to the pile self-load and the water pressure at the depth of the O-cell placement. After the duration of holding the maximum load is over, the displacement values of the lower O-cell are 9.0mm and 5.9mm, and the pile tip displacement values are 4.6mm and 2.2mm respectively. The displacement values of the upper O-cell are 6.7mm and 6.9mm and the pile end displacement values are 1.2mm and 0.8mm.

B. *FEM Analysis*

The determination of the input data for the model is made on the basis of the correlation equations and empirical researches [9, 12-14] as follows:

- The Young's modulus, E_{50ref}, is determined from the drained 3-dimensional (3-D) consolidation test with the selected chamber pressure level σ_3 in accordance with the actual state of the soil profile.
- Drained shear strength parameters are taken from the drained 3-dimensional consolidation test or as the effective shear strength values in the undrained 3-D consolidation test. For the soil profile without the 3-D CU test, CD can be obtained from the direct shear strength test, but its reliability is not high. The undrained shear strength parameter also does not take the friction angle of the ground soil, $\varphi_u = 0$, into account but only the adhesion force of the soil C_u. C_u values are obtained through the undrained 3-D test, the in-situ shear vane test, or the 1-dimensional unconfined compressive test.
- For dense layers of sand or over-consolidated layers of clay, the expansion angle ψ exists. Normally, according to the instructions of Plaxis software, $\psi = \varphi - 30^0$. In other cases, the expansion angle is equal to 0.

The parameters of the Hardening Soil model are determined directly from the geological survey report test results or from the correlation equations of the previous researches to obtain the values summarized in Table I.

TABLE I. SOIL PARAMETERS IN FEM

No.	Backfill	Layer 1	Layer 2	Layer 3	Layer 3b	Layer 4	Layer 6
Type	HS	HS	HS	HS	HS	HS	HS
γ_{unsat} (kN/m^3)	18.56	14.61	18.42	18.35	19.23	20.38	19.93
γ_{sat} (kN/m^3)	18.80	15.59	19.31	19.21	19.94	20.66	20.62
E^{ref}_{50} (kN/m^2)	11210	4000	18000	60000	80000	95000	90000
E^{ref}_{oed} (kN/m^2)	11210	4000	18000	60000	80000	95000	90000
E^{ref}_{ur} (kN/m^2)	56030	20000	90000	250000	370000	350000	330000
m	0.75	1	0.85	0.5	0.5	0.75	0.6
ν_{ur}	0.2	0.2	0.2	0.2	0.2	0.2	0.2
c (kN/m^2)	22.41	13.77	36	4.1	4.1	218.4	1.7
$\varphi^{(o)}$	0	0	4.8	32.49	32.49	0	37.00
$\psi^{(o)}$	0	0	0	2.49	2.49	0	7.00
R_{inter}	0.85	0.7	0.7	0.85	0.85	0.85	0.85

To make load capacity comparison easier, the 4 used models in Plaxis 2D software were:

- Model 1: TP1 pile with grouting
- Model 2: TP1 pile without grouting.
- Model 3: TP1 pile with grouting only from -26.3m to -48.5m.
- Model 4: TP1 pile with grouting only from -48.5m to -65.5m.

Firstly, the TP1 pile with grouting will be simulated, because it is most similar to the actual observation. After obtaining its results (note that the load level is gradually adjusted until failure) the parameters change. Pile TP1 becomes a non-grouted pile, and also from that model, the set of parameters change in order to simulate the models 3 and 4 (the grouting length from -26.3m to -48.5m and -48.5m to -65.5m respectively). The obtained results are used to draw the relationship graph of the curve and the load of the four investigated models. We convert the test load of the Osterberg method into the conventional static test load using an equivalent method, from which we have enough basis to evaluate the increase in the load capacity of the TP1 pile according to the four different models.

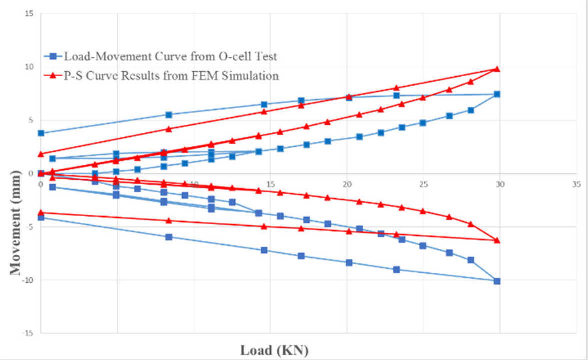


Fig. 2. Movement and load curves obtained from actual and simulated observations.

The O-cell upper displacement is 7.44mm/9.82mm on site. The O-cell lower displacement is -10.08mm/-6.27mm on site. The total strain of the O-cell is 17.52/16.09mm. This result from Figure 2 shows that the measured results and the finite element modeling have high accuracy. However, in terms of displacement between the upper and lower O-cells, there is an inversion. The field displacement value is smaller than that of the Plaxis model, i.e. 7.44/9.82mm. It should be noted that the lateral friction is enhanced due to the better grouting ability (regarding the safety issue). For the lower O-cell, in fact, the tip resistance has not been applied to the load-carrying capacity of the pile, so the actual displacement is larger than that of the Plaxis model -10.08/-6.27mm).

IV. DISSCUSION

This analytical model is rather similar to the investigated field, so the authors take this set of soil parameters to continue to simulate the remaining models with the condition of increasing load until the failure of the pile. The results of models 1-4 are displayed in Figures 3-6.

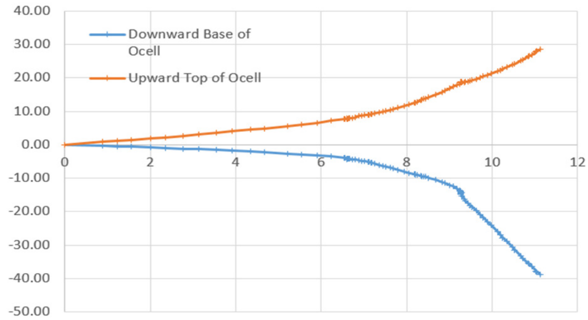


Fig. 3. Model 1. The TP1 pile with grout reaches 305% of maximum load.

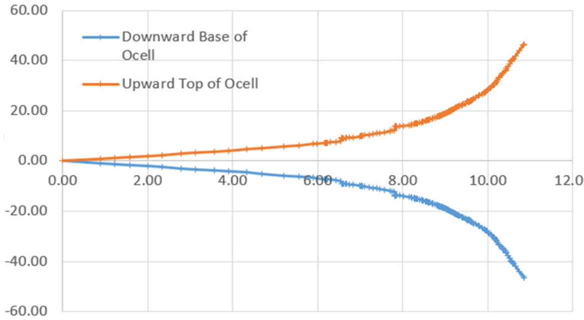


Fig. 4. Model 2. The TP1 pile without grout reaches 255% of maximum load.

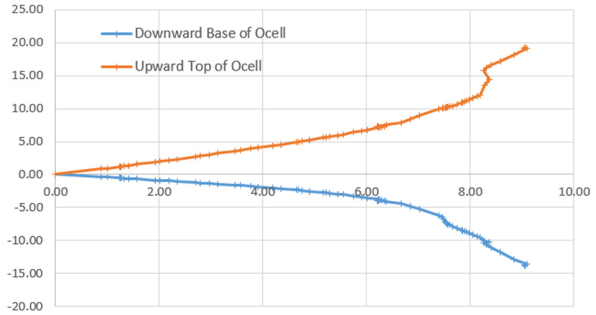


Fig. 5. Model 3. The TP1 pile with grout from -26.3m to -48.5m reaches 280% of maximum load.

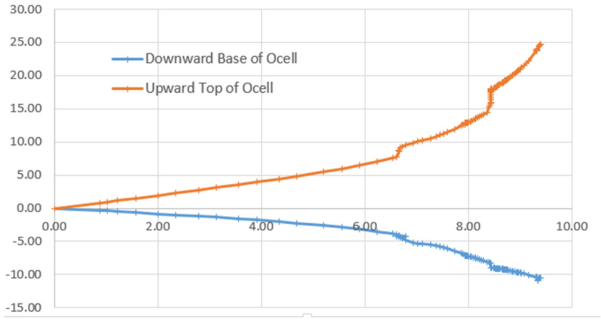


Fig. 6. Model 4. The TP1 with grout from -48.5m to -65.5m reaches 280% of maximum load

The ultimate load capacity is determined based on the shape of the load-displacement relationship curve:

$$S = f(P), \log S = f(\log P) \quad (1)$$

In many cases, it is necessary to combine with other curves like:

$$S = f(\log t), P = f(S/\log t) \quad (2)$$

With the load-displacement relationship curve, the ultimate load-carrying capacity is determined as follows:

- Case 1: the P-S relationship curve has a clear bending point. The ultimate load-carrying capacity is determined directly on the curve, the load corresponding to the point where the curve begins to change suddenly or the curve is closely parallel to the displacement axis.
- Case 2: the P-S relationship curve changes slowly. It is quite difficult or impossible to accurately determine the bending point. The ultimate load-carrying capacity is determined by different graphical methods.

The analysis and evaluation of the test results by the Osterberg method are still based on the principles of conventional static testing, in which the Q-S relationship curve is an important tool of the quantitative analysis of the load-carrying capacity of the piles that a mechanical problem usually covers. The goal of load testing in most cases is not only to determine the bearing capacity of the pile foundation, but also to get a more specific view than the one from the general theory and the calculated results about the interaction between the pile and the composing ground so that we can offer more appropriate behaviors in the design of the pile foundation and in the future construction processing. We can build a load-settlement curve (similar to a conventional static compression test curve) in the following way:

At any displacement S, the downward force affecting the soil is:

$$Q'{\downarrow} = Q{\downarrow} + w_2' \quad (3)$$

where $Q{\downarrow}$ is the downward force by the O-cell at the displacement S and w_2 is the self-load of the pile from the O-cell downwards (buoyancy must be taken into account if we are below the groundwater level).

At the same displacement S, the upward force carried by the soil is:

$$Q'{\downarrow} = (Q{\downarrow} - w_1') \times F \quad (4)$$

where w_1' is the self-load of the pile from the O-cell upwards (buoyancy must be taken into account if we are below the groundwater level) and F is the coefficient that takes into account the difference in the lateral friction between the normal static compression and the Osterberg compression. With F = 1.00 for the pile in the rock and F = 0.95 for the pile in the loose soil.

At the displacement S, the sum of the upward and downward forces will be subjected by the soil:

$$P = Q'{\downarrow} + Q'{\downarrow} = Q{\downarrow} + w2' + (Q{\downarrow} - w1') \times F \quad (5)$$

However, in the case of downward compression, the soil is already partially imposed by the self-load of the pile, so the applied load at the top of the pile is:

$$P_e = P - w_1' - w_2' = Q\downarrow + (Q\downarrow - w_1') \times F - w_1' \quad (6)$$

At a load of P_e (loading downwards), the settlement of the pile is $S + \Delta\delta$ where the increment $\Delta\delta$ refers to the elastic compressive strain of the pile.

Repeating the above steps with different positions S, we can draw a static compression curve corresponding to the data pair (P_e, $S + \Delta\delta$). If at a certain position that goes beyond the range of a curve (for example, the lateral resistance curve of the pile segment on the jack), we choose one of the following two solutions:

- Take the upward force as constant (regarded as the extreme value) and equal to the tested maximum lateral resistance.
- Extrapolating the next segment of the lateral resistance curve. However, this extrapolation has errors, because if the pile body is very rough (convex), then the lateral resistance reaches its ultimate value more slowly. Conversely, the less convex and concave the pile body is, the faster the lateral resistance reaches its ultimate value.

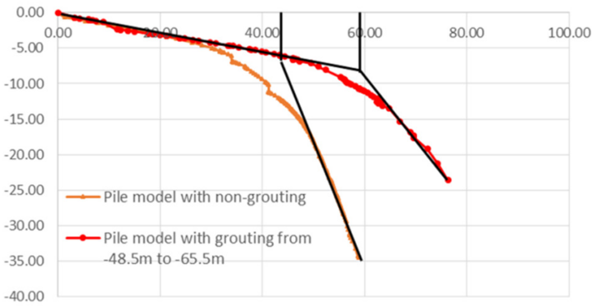


Fig. 7. Load capacity of the pile in the models 2 and 4.

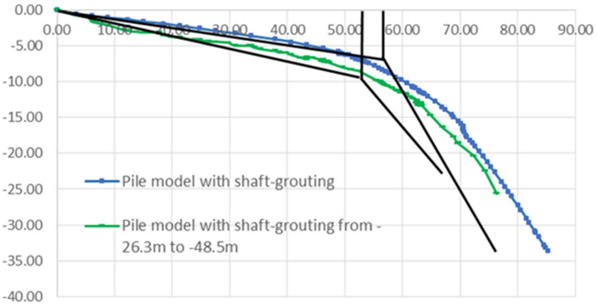


Fig. 8. Load capacity of the pile in the models 1 and 3.

From the above analysis, we will convert equivalently the test results of the Osterberg method into those of the conventional test method to plot the P-S relationship curve from the data pair (P_e, $S + \Delta\delta$) and then use the graphical method to determine the load capacity. The results of the shape of the load-displacement relationship curve are displayed in Figures 7-9.

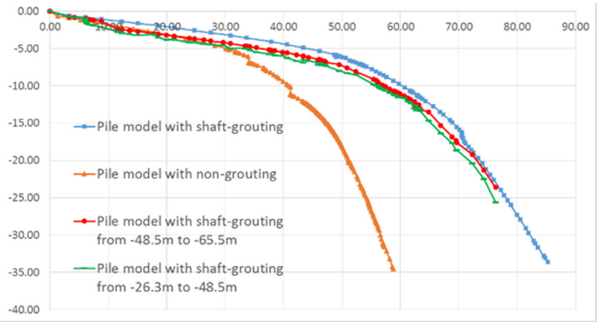


Fig. 9. Summary of the pile load-carrying capacity of the 4 models.

V. CONCLUSION

Through the experimental study from the O-cells of the TP1 pile in the Eximbank building on the ability to improve the load-carrying capacity of the lateral walls, the following conclusions can be drawn:

- Improvement of the bearing capacity of the pile by the lateral grouting method is highly effective with results of 61.7MN and 41.8MN with and without grout respectively. The friction performance of the pile body with grouting increases by 1.48 times. It is different from the previous studies that we compared the analytical equations and the field results.
- Grouting of each segment at depths of 22.2m (from -26.3m to -48.5m) and 16.8m (from -48.5m to -65.3m) corresponding to the maximum load of 59.2MN and 59.4MN is evaluated as being similar to the load-carrying capacity. The FEM simulation results show that the bigger grouting length and the bigger grouting depth will give higher values of bearing capacity. Simulating grouting in each segment is very beneficial for the design of allowable loads, both in terms of safety and cost aspects.
- The FEM (Plaxis) can simulate and evaluate the O-cell and the lateral grouting tests of piles. This method can be used for computational correction after the field results are available for detailed design.

ACKNOWLEDGMENT

The current research is supported by the Ho Chi Minh City Open University (Project No. E2019.08.3).

REFERENCES

[1] B. H. Fellenius, "Tangent Modulus of Piles Determined from Strain Data," presented at the Foundation Engineering: Current Principles and Practices, 1989, pp. 500–510.

[2] B. H. Fellenius, "Analysis of results of an instrumented bidirectional-cell test," *Geotechnical Engineering Journal of the SEAGS & AGSSEA*, vol. 46, no. 2, pp. 64–67, 2015.

[3] H. M. Nguyen, B. H. Fellenius, A. J. Puppala, P. Aravind, and Q. T. Tran, "Bidirectional Tests on Two Shaft-Grouted Barrette Piles in Mekong Delta, Vietnam," *Geotechnical Engineering Journal of the SEAGS & AGSSEA*, vol. 47, no. 1, pp. 15–25, Mar. 2016.

[4] L. M. Zhang and L. F. Chu, "Calibration of Methods For Designing Large-Diameter Bored Piles: Ultimate Limit State," *Soils and Foundations*, vol. 49, no. 6, pp. 883–895, Dec. 2009, https://doi.org/10.3208/sandf.49.883.

[5] "Report on Barrtte load testing TP2- Eximbank Tower, Ho Chi Minh City, Viet Nam," 138131–2, Sep. 2013.

[6] "Pile Foundation - Design Standard," National Standards, VN 10304: 2014.

[7] S. Pooranampillai, S. Elfass, W. Vanderpool, and G. Norris, "Large Scale Laboratory Testing of Low Mobility Compaction Grouts for Drilled Shaft Tips," *Geotechnical Testing Journal*, vol. 33, no. 5, pp. 397–409, Sep. 2010, https://doi.org/10.1520/GTJ102658.

[8] G. D. Plumbridge, B. D. Littlechild, S. J. Hill, and M. Pratt, "Full-scale shaft grouted piles and barrettes in Hong Kong–A First," in *Proceedings of the Nineteen Annual Seminar of the Geotechnical Division of the Hong Kong Institution of Engineers*, Hong Kong, China, 2000, pp. 159–166.

[9] T. Nguyen, V. Q. Lai, D. L. Phung, and T. P. Phan, "Shaft Resistance of Shaft-grouted Bored Piles and Barrettes Recently Constructed in Ho Chi Minh City," *Geotechnical Engineering*, vol. Vol. 50, no. 3, pp. 155–162, Sep. 2019.

[10] J. Zhou, X. Gong, and R. Zhang, "Model tests comparing the behavior of pre-bored grouted planted piles and a wished-in-place concrete pile in dense sand," *Soils and Foundations*, vol. 59, no. 1, pp. 84–96, Feb. 2019, https://doi.org/10.1016/j.sandf.2018.09.003.

[11] X. Zhao, G. Zhou, G. Zhao, L. Kuang, and X. Hu, "Fracture controlling of vertical shaft lining using grouting into neighboring soil deposits: A case study," *Soils and Foundations*, vol. 57, no. 5, pp. 882–891, Oct. 2017, https://doi.org/10.1016/j.sandf.2017.08.018.

[12] N. Mangi, D. K. Bangwar, H. Karira, S. Kalhoro, and G. R. Siddiqui, "Parametric Study of Pile Response to Side-by-Side Twin Tunneling in Stiff Clay," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5361–5366, Apr. 2020, https://doi.org/10.48084/etasr.3290.

[13] T. A. Rind, H. Karira, A. A. Jhatial, S. Sohu, and A. R. Sandhu, "Particle Crushing Effect on The Geotechnical Properties of Soil," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4131–4135, Jun. 2019, https://doi.org/10.48084/etasr.2730.

[14] 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.

Variation Evaluation of Path Characteristic and Site Amplification Factor of Earthquake Ground Motion at Four Sites in Central Japan

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

Abstract-The considered parameters in seismic design vary, with the Earthquake Ground Motion (EGM) having the largest variation. Since source characteristic, path characteristic, and Site Amplification Factor (SAF) influence the EGM, it is crucial to appropriately consider their variations. Source characteristic variations are mainly considered in a seismic hazard analysis, which is commonly used to evaluate variations in EGM. However, it is also important to evaluate variations in path characteristic and SAF with only a few studies having individually and quantitatively examined the variations of these two characteristics. In this study, based on strong-motion observation records obtained from four sites in central Japan, the three characteristics were extracted from seismograms using the concept of spectral inversion. After removing the source characteristic, the path characteristic and SAF were separated, and the variations in these two characteristics were quantified. To separate and obtain each characteristic from the observed record, one constraint condition must be imposed, whereas the variations in the constraint condition must be ignored. In that case, the variations in the constraint condition are included in the variations of the separated characteristics. In this study, this problem was solved by evaluating the variation in the constraint condition, which is the SAF at a hard rock site, by the use of the vertical array observation record at the site.

Keywords-earthquake ground motion; path characteristic; site amplification factor; variation

I. INTRODUCTION

In seismic design, the cross-section of a structure is determined in such a way that the cross-sectional resistance exceeds the section force caused by an earthquake. Since the parameters to be considered in seismic design vary, the cross-section of a structure should be set considering the possibility that the seismic load acting on the structure exceeds the design value. The Earthquake Ground Motion (EGM) has the largest variation among the considered parameters in seismic design. Therefore, it is critical to accurately evaluate the variation in EGM. Seismic hazard analysis [1] is a method for considering variations in EGM. In seismic hazard analysis, probabilistic EGM is evaluated by considering variations in parameters such as seismic magnitude and hypocenter location.

EGM is determined by three characteristics: source characteristic, path characteristic, and Site Amplification Factor (SAF). These three characteristics must be considered carefully because the EGM varies significantly depending on the evaluation method of these characteristics [2]. In seismic hazard analysis, source characteristic is the major variation evaluation object, and path characteristic and SAF are treated deterministically. However, since these two characteristics vary in reality, it is necessary to evaluate the variations in path characteristic and SAF separately in order to conduct rational seismic design. Although there are many research cases on the variation in amplification characteristics, the definitions of amplification characteristics vary, such as the ratio of spectral acceleration to the predicted value by ground motion prediction equations [3], the amplification at the object location with respect to the reference location [4], the ratio of the vertical array strong-motion observation record on the ground surface to the borehole [5, 6], and the residual of the observed EGM to the probabilistically evaluated EGM [7]. The SAF to be considered when the EGM is separated into the three characteristics must include not only the amplification by the shallow subsurface but also that by the deep subsurface, because the amplification factor by a shallow subsurface is significantly smaller than the actual SAF [8, 9]. Therefore, the variation in SAF for each frequency, including the effect of deep subsurface must be evaluated.

There are cases in which the variations of SAF, including the effect of deep subsurface, were studied using the spectral ratio of seismograms observed at two sites [10, 11]. However, these studies did not separate path characteristic and SAF, and because path characteristic was treated deterministically, their results included variations of path characteristic and SAF. There are no case studies in which the variations in path characteristics and SAF were quantified after separation. Spectral inversion [12, 13] is a technique for separating each characteristic from the observed records. The current study separates path characteristic and SAF based on the concept of spectral inversion and conducts quantitative evaluation of each variation.

Corresponding author: Takashi Nagao

II. METHOD

A. Target Sites and Target Earthquakes

This study focuses on four sites (GIFH20, GIFH24, GIFH28, and NGNH20) in central Japan that are part of the KiK-net [14] strong-motion seismograph network as shown in Figure 1. KiK-net provides EGMs recorded at the ground surface and in the borehole simultaneously, and P-S logging reveals the S-wave velocity profile leading to the seismograph installation position in the borehole. Strong-motion records and ground data were made available to the public by [15]. However, seismographs in boreholes are rarely installed at seismic bedrock in which S-wave velocity exceeds 3km/s. Therefore, the S-wave velocity profile from the ground surface to the seismic bedrock was obtained by combining the information from KiK-net and J-SHIS [16] as described in [8]. The obtained S-wave velocity profiles are shown in Figure 2. Among the sites, GIFH28 has the thickest and NGNH20 the thinnest sediment site. In GIFH24, the depth at which the S-wave velocity exceeds 1500m/s is shallower than that in NGNH20. The multiple reflection theory calculates transfer functions assuming horizontally stratified ground conditions (Figure 3). In the frequency range of 0.1–10Hz, which is important in seismic engineering, the amplification factor of GIFH24 is the smallest with values less than 2 in the frequency range lower than 7Hz. Because the soil layer thickness with an S-wave velocity of 720m/s is relatively thick (= 25m), the amplification factor is as high as 5 at 6Hz in NGNH20. First-order peaks occur at 2Hz in GIFH20 and 0.6Hz in GIFH28, in which the sedimentary layer is the thickest.

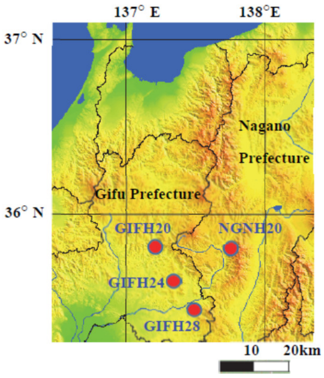


Fig. 1. Target sites.

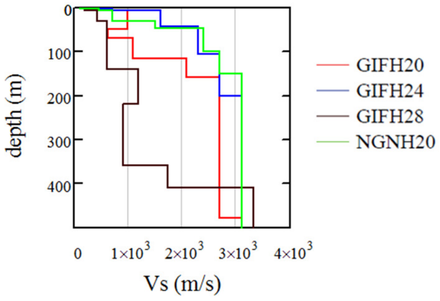


Fig. 2. S-Wave velocity profiles.

Table I shows the list of the considered events. The epicenters of the target events, as well as the locations of the observation sites, are shown in Figure 4. The red circles indicate seismograph installation sites, and the blue cross marks indicate the epicenters. The epicenters do not exist on the west side of the observation sites, but they are distributed evenly in the other directions.

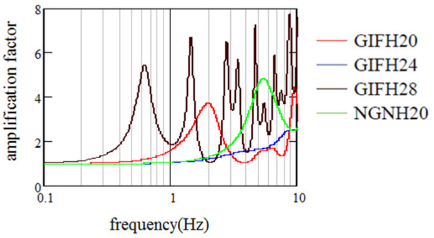


Fig. 3. Transfer functions.

TABLE I. TARGET EARTHQUAKES

No	Date	Time	Lon (deg)	Lat (deg)	Depth (km)	*M*	Site*
1	Sep/27/2020	13:13	137.80	35.10	45	5.3	b, c, d
2	Apr/23/2020	13:44	137.32	35.64	3	5.5	b, d
3	Feb/13/2018	14:39	137.59	35.87	3	4.1	b, c, d
4	Dec/6/2017	00:13	137.97	36.38	10	5.3	b, d
5	Jun/25/2017	07:02	137.59	35.87	7	5.6	b, c, d
6	Dec/6/2016	09:05	137.34	36.01	5	4.5	b, c, d
7	May/25/2015	14:28	139.64	36.05	56	5.5	a, b, d
8	Mar/4/2015	00:04	136.80	35.34	40	4.6	a, b, c
9	Dec/3/2014	23:19	137.12	35.29	45	4.2	a, b, c, d
10	Sep/16/2014	12:28	139.86	36.09	47	5.6	a, b, d
11	Jun/1/2012	17:48	139.88	36.03	44	5.1	a, b, c, d
12	Apr/25/2012	05:22	140.68	35.72	43	5.5	a, b, d
13	Jan/28/2012	07:43	138.98	35.49	18	5.4	a, b, c, d
14	Jun/30/2011	08:21	137.95	36.19	4	5.1	a, b, d
15	Apr/16/2011	11:19	139.94	36.34	79	5.9	a, b, c, d
16	Mar/16/2011	03:33	137.30	36.00	0	4.0	a, b, c, d

* a: GIFH20, b: GIFH24, c: GIFH28, d: NGNH20

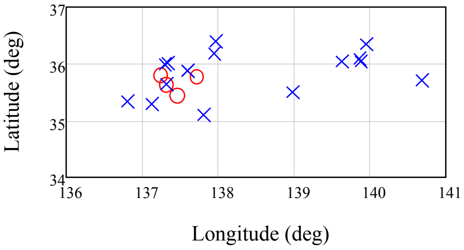


Fig. 4. Target Sites and epicenters.

Because EGM records of small seismic magnitude (*M*) have a low signal-to-noise (S/N) ratio, events of seismic magnitude $M = 4.0$ or larger are considered. Additionally, EGM records of $M = 6.0$ or larger are excluded to avoid the effect of the rapture process at the fault plane. With the ω^{-2} model [17], only EGM records with good S/N ratios in the range of 0.3Hz or higher were used. To discuss the amplification characteristics of S-waves, 10.24s records were extracted for the S-wave arrival. Before calculating the Fourier spectra, baseline correction was conducted, and the extracted records were tapered. Based on the root mean square of the Fourier spectra of seismograms recorded horizontally in two directions, a Parzen window with a bandwidth of 0.3Hz was used to smooth the spectra.

B. Separation Method of the Three Characteristics

The three characteristics are separated from the observed records according to the concept of spectral inversion. This study focuses on seismograms of the same event recorded at multiple stations. The Fourier spectra of the observed records can be expressed as the product of the source characteristic, path characteristic, and SAF as shown in (1):

$$O_{ij}(f) = S_i(f)P_{ij}(f)G_j(f) \quad (1)$$

where O is the observed EGM, S is the source characteristic, P is the path characteristic, G is the SAF, i is the event number, j is the station number, and f is the frequency. The values of each characteristic for each frequency can be obtained by solving the equations simultaneously using (2) in which the equation is expressed as a common logarithm.

$$\log O_{ij} = \log S_i + \log P_{ij} + \log G_j \quad (2)$$

The path characteristic can be expressed by (3):

$$P_{ij}(f) = \frac{1}{r_{ij}^{\,n}} \exp\left(-\pi f r_{ij} / QVs\right) \quad (3)$$

where r is the hypocentral distance and n is a value representing the geometric attenuation of the body wave, with its typical value being 1.0, but it is pointed out that $n = 2$ when the hypocentral distance is significantly long [18]. In this study, $n = 2$ was assumed for a dataset with hypocentral distances of 120km or more. Vs is the average S-wave velocity along the propagation path, and $Vs = 3.8$km/s in the target area. Q is the Q-value, which represents the inelastic attenuation of EGM.

There is a trade-off relationship among the characteristics: for example, a product of large path characteristics and small SAF can result in the same value as a product of small path characteristics and large SAF. Therefore, one constraint condition must be imposed to solve the equations simultaneously. The constraint condition is set as the SAF at one station and calculated as a theoretical SAF assuming one-dimensional (1D) ground structure [19]. The SAF of a significantly small rock site is chosen for the constraint condition. However, even at rock sites, EGM is amplified in many cases because of weathering [20]. Furthermore, because of the heterogeneity of a three-dimensional (3D) ground structure, the SAF assuming 1D ground structures often underestimates the actual SAF [2, 21]. Therefore, it is desirable that the SAF used as a constraint condition to be as small as possible and the S-wave velocity profile down to seismic bedrock is disclosed. As shown in Figure 3, the amplification factor at GIFH24 is significantly small in the frequency band below 10Hz, and we can assume that the difference in SAF between the actual and the one calculated by assuming a 1D structure is negligible. As a constraint condition, this study uses the 1D amplification factor (1DA) at GIFH24.

III. RESULTS AND DISCUSSION

A. SAF Variation in Low-Frequency Bands at Hard Rock Sites

Path characteristic and SAF can be separated in the spectral inversion and variations of the two characteristics can be discussed. However, the variation cannot be evaluated for the constraint condition. The constraint condition, which is the SAF at the reference hard rock site, also varies. Therefore, variations in the obtained characteristics include variations in the constraint conditions. Because seismographs are installed at the ground surface (S) and in boreholes (B) at KiK-net sites, the variation can be evaluated using the spectral ratio (S/B). However, even if the seismograph in the borehole is installed on the seismic bedrock, S/B cannot be considered as the SAF because borehole records include the upward EGM (E) and downward EGM (F), which are reflected at the layer boundary ($E + F$). Because E and F are identical on the ground surface, EGM is $2E$, and the S/B becomes $2E/(E + F)$. However, the SAF is defined as the spectral ratio of the EGM at the ground surface to that obtained by doubling the incident EGM (E) at the bedrock ($2E/2E$). Figure 5 compares the S/B ($= 2E/(E + F)$) with the SAF ($= 2E/2E$) under horizontally stratified ground conditions, using GIFH24 and NGNH20 as examples. The red line represents the SAF, and the blue line represents the S/B. Although the difference between the two is large in the frequency band higher than 1Hz, it is tiny in the frequency band lower than 1Hz. Therefore, the variation in S/B in the frequency band lower than 1Hz of these two sites is the variation in the SAF. Figure 6 shows the variation in S/B calculated with the observation record. The gray line represents the individual S/B and the red line represents the average value. Each characteristic is separated on a logarithmic scale in the spectral inversion, as shown in (2). Therefore, unless otherwise specified, these results are on a logarithmic scale hereafter. The value of S/B agrees with the theoretical value on average in the frequency band of 1Hz or lower, although there is some dispersion.

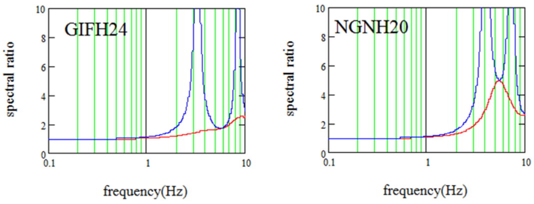


Fig. 5. Comparison of $2E/(E + F)$ and $(2E/2E)$.

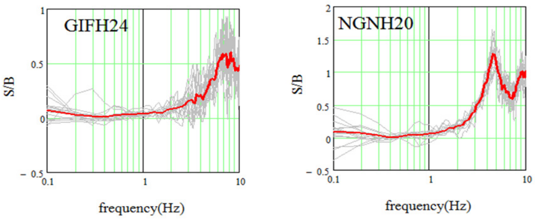


Fig. 6. Variation in S/B.

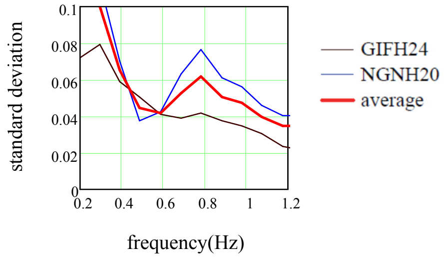


Fig. 7. SD of S/B.

Figure 7 shows the standard deviation (SD) of *S*/*B* in the frequency band lower than 1Hz. The SD is large in frequencies lower than 0.4Hz because of the poor S/N ratio in the seismograms of small amplitude. The SD of the SAF of hard rock sites in the frequency band lower than 1Hz was determined to be 0.048 using the average SD in the range of 0.488 to 1.074Hz. Figure 8 shows the histogram of the *S*/*B*, referring to the results at frequencies 0.586 and 0.781Hz. The probability distribution can be considered normal.

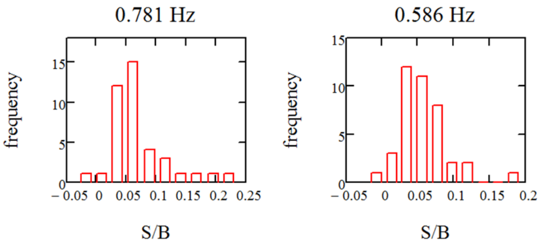


Fig. 8. Histogram of *S*/*B*.

B. Evaluation of Q-Value and Path Characteristic Variation

Because the variation in the SAF at hard rock sites was obtained when considering the frequency band of 1Hz or lower, the variation in the path characteristics can be evaluated based on the results. The *Q*-value used to evaluate path characteristics varies by region, and its frequency dependence has been mentioned. There have been no previous studies on the *Q*-value for the regions covered in this study, but the *Q*-values for other Japanese regions are $Q = 83 f^{0.73}$ [22] and $Q = 33 f^{1.0}$ [23]. In this study, we first evaluate the *Q*-value for the target area. Based on the simultaneous records from two locations (i.e. GIFH24 and NGNH20), we can evaluate the *Q*-value using (4) and combining (2) and (3).

$$Q = \frac{\pi f \left(r_j - r_k\right) \log_{10} e \Big/ Vs}{\log O_k - \log O_j - \log G_k + \log G_j - n \log r_j + n \log r_k} \qquad (4)$$

where *j* and *k* are the observation stations.

SAF in the frequency band higher than 1Hz cannot be determined at this point, but for frequencies lower than 1Hz the SAF of the two stations can be considered as 1.0. Therefore, the *Q*-value can be determined by (4). The *Q*-value was obtained as in (5) regarding the average *Q*-value for each frequency, and the minimum spectral residual, which is to be described later.

$$Q = 70 f^{0.90} \qquad (5)$$

Figure 9 compares the average *Q*-value with the value obtained by (5). The blue line represents the average *Q*-value, and the red line represents the value according to (5).

The variation in path characteristic is evaluated based on the obtained *Q*-value. By applying (2) to the records of two sites *j* and *k* (i.e. GIFH24 and NGNH20), the spectral residual (*SR*) is obtained as follows:

$$SR = \log O_j - \log O_k - \log P_j + \log P_k - \log G_j + \log G_k \qquad (6)$$

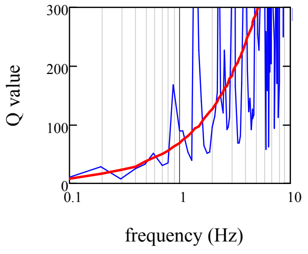


Fig. 9. *Q*-value.

The *SR* distribution is displayed in the left panel of Figure 10. The gray lines represent individual results, and the red line represents the average value. If path characteristic and SAF do not vary, *SR* becomes zero. Although the average *SR* is generally zero for frequencies lower than 1Hz, the variation in *SR* is large. The SD of *SR* is shown in the right panel of Figure 10. The average value of SD in the frequency band 0.293-1.074Hz was determined to be 0.184. The histograms of *SR* at frequencies 0.586 and 0.781Hz are shown in Figure 11. Although the result at 0.586Hz deviates from the normal distribution, the frequency distribution at 0.781Hz can be considered normal.

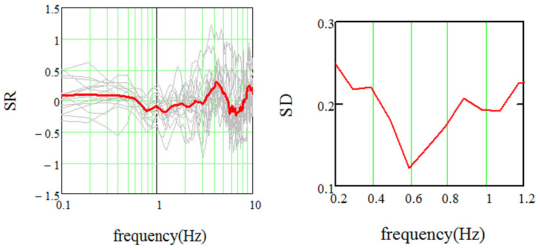


Fig. 10. Distribution and SD of *SR*.

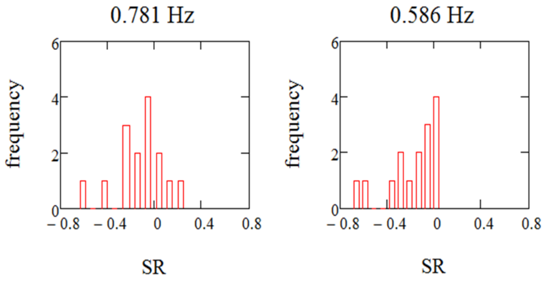


Fig. 11. Histogram of *SR*.

This *SR* variation includes the variations in path characteristics and SAFs of the two sites. Assuming that the path characteristic and SAF at each site follow independent normal distributions, the following equation holds:

$$\sigma_{SR}^2 = 2\left(\sigma_P^2 + \sigma_G^2\right) \qquad (7)$$

where σ is the SD.

By using the SD of SAF obtained in the previous section, the SD of the path characteristic was determined to be 0.121.

C. SAF and its Variation

Based on the above results, the variation in SAF at the site where the effect of the sedimentary layer cannot be ignored is evaluated. Because the path characteristics vary slightly with frequency, the SD obtained in the previous section is assumed to be applicable at all frequencies. Furthermore, because GIFH24 is a site with a significantly small amplification factor in the frequency band lower than 10Hz, it is assumed that the average value of SAF at GIFH24 agrees with the amplification factor assuming 1D ground structure (1DA) considering ground structure leading to the seismic bedrock. Regarding the variation in SAF at GIFH24, it is assumed that the SD value obtained in the previous section for the low-frequency band can be applied to all frequencies. Based on the simultaneous records at GIFH24 and NGNH20, SAF and its variations of NGNH20 are evaluated by solving (6) for SAF at NGNH20 as shown in the left panel of Figure 12. The gray lines represent the individual results, and the red line represents the average value. The right panel of Figure 12 shows the average value and the SD of SAF. The red line represents the average value and the blue line represents the SD. The SD is less dependent on the frequency. The SD has an average value of 0.323 in the 0.195–10Hz frequency band. Because the SD here includes variation in path characteristics at two sites and variation in SAF at GIFH24, these variations are removed using the same approach as in (7). The average values of the SD of SAF at NGNH20 in the frequency bands of 1–10Hz and 0.195–10Hz were determined to be 0.282 and 0.270 respectively. Furthermore, the SAF at GIFH28 is discussed. The procedures are the same as in NGNH20. The left panel of Figure 13 shows the SAF, with the gray lines representing the individual results and the red line representing the average value.

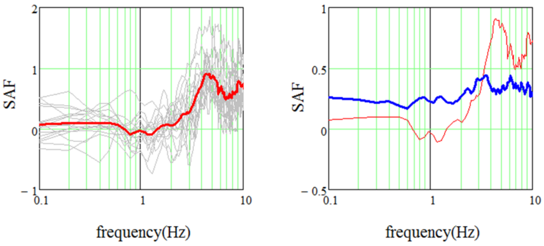


Fig. 12. Average and SD of SAF (NGNH20).

The right panel of Figure 13 shows the average SAF (red) and the SD (blue) lines. The SD in the frequency band higher than 2Hz is slightly larger than that in the lower-frequency band. The average value of SD in the frequency band of 0.195–10Hz is 0.316. When the variation in the SAF of GIFH24 and path characteristics of the two sites were removed from the SD, the average value of SD of SAF at GIFH28 was 0.061 in the 0.195–1.1Hz frequency band, 0.277 in the 1–10Hz frequency band, and 0.262 in the 0.195–10Hz frequency band.

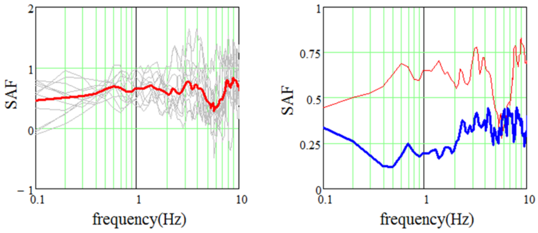


Fig. 13. Average and SD of SAF (GIFH28).

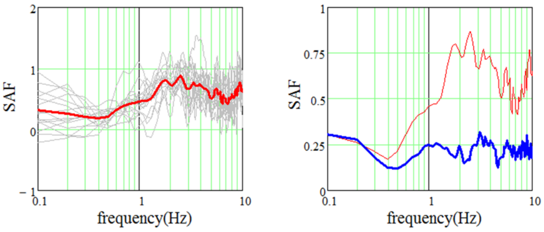


Fig. 14. Average and SD of SAF (GIFH20).

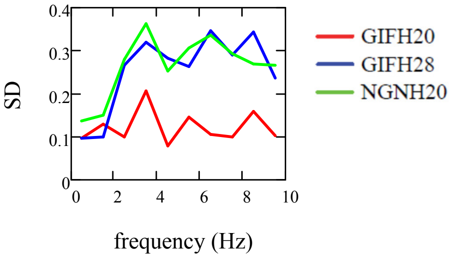


Fig. 15. Average SD of SAF.

The left panel of Figure 14 shows the SAF for GIFH20, with gray lines representing the individual results and the red line representing the average value. The right panel of Figure 14 shows the average value and SD of SAF, with the red line representing the average value and the blue line the SD. The average SD in the frequency band from 0.195 to 10Hz is 0.217. When the variation of the SAF of GIFH24 and the path characteristics of the two sites were removed from this SD, the average value of the SD of SAF at GIFH20 was 0.080 in the 0.195–1.1Hz frequency band, 0.130 in 1–10Hz, and 0.125 in 0.195–10Hz.

Because the SD at each site varies with frequency, Figure 15 shows the average SD in case of SAF with frequency. The SD is small for frequencies lower than 2Hz, in the range of 0.10-0.15, but a large value of approximately 0.30 can be obtained in the case of GIFH28 and NGNH20 when considering frequencies higher than 3Hz. However, at GIFH20, the SD is almost constant in all frequency bands and is smaller than those at the other two sites. The average value of SAF varies significantly with frequency at the three sites, but the SD does not vary significantly with frequency and is not dependent on the average value of SAF. Because the SAF in the low-frequency band is mainly determined by the deep subsurface structure, a low SD of the SAF in the low-frequency band indicates a minor local change in the deep subsurface structure near the observation sites. However, the SAF in the high-frequency band mainly depends on the shallow subsurface structure, implying that the shallow subsurface structure is changing in a complex manner around the site where the SD of the SAF is large in the high-frequency band. However, the sites where the SDs of SAFs do not increase even in the case of high-frequency bands, such as in GIFH20, are considered to be sites at which the surrounding shallow subsurface structure does not change significantly.

D. Comparison of 1DA and SAF

Figure 16 compares the SAF in this study to 1DA used in design practice. The red line represents the average SAF, the blue line represents the 1DA considering shallow and deep subsurface, and the black line represents the 1DA considering shallow subsurface only. Note that those values are expressed on an arithmetic scale. Compared to the 1DA considering deep subsurface, the shallow subsurface 1DA significantly underestimates the amplification factors. This result is in accordance with the findings in [8]. Except for the second-order peak at 1.5Hz of GIFH28, the 1DA considering deep subsurface underestimates the amplification factor compared to the SAFs in this study. Previous studies [2, 20] have also highlighted this trend. Except for the very thin sediment site, NGNH20, the SAFs at the other two thick sediment sites have high amplification factors over a wide-frequency band, whereas 1DA has large amplifications only at certain frequencies. Therefore, when considering the EGM amplification in seismic design, it is necessary to use a method capable of evaluating the actual amplification factors, such as spectral inversion, rather than 1DA.

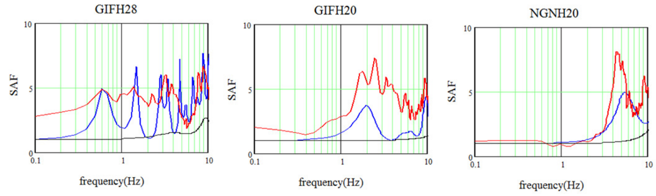


Fig. 16. SAF and 1DA.

IV. CONCLUSION

In this study, the source characteristic, path characteristic, and SAF were separated from observation records using the spectral inversion approach by utilizing the seismic records of four strong-motion observation sites in central Japan. The variation in path characteristics and SAF were quantified. The conclusions of this study are:

- SAF and path characteristic can be separated by spectral inversion by imposing a single constraint condition, namely, SAF at a hard rock site. However, the variation cannot be evaluated for the constraint condition. Therefore, the variations obtained for each characteristic contain the variation in the constraint conditions. In this study, the SD of the constraint condition was evaluated using the variation in the spectral ratio of seismograms at the ground surface to those in the borehole. The obtained SD of the constraint condition was 0.048.
- The SDs of the *SR* of two hard rock sites' records were evaluated. The variation in path characteristic was evaluated by assuming that path characteristic and SAF follow independent normal distributions. The SD of the path characteristic was estimated to be 0.121. Finally, the variation in the SAF was evaluated. The SD of the SAF was independent of the average SAF and ranged from 0.10 to 0.15 in the frequency band of 2Hz or lower. In the high-frequency band, the SD of SAF was approximately 0.30 at GIFH28 and NGNH20 and approximately 0.10 at GIFH20. The SD in the high-frequency region varies significantly from point to point. This difference in SD is due to the difference in the uniformity of the shallow subsurface around the site.
- The 1DA underestimates the SAF even when the ground structure down to the seismic bedrock is considered. At thick sediment sites, large SAF is observed in a wide-frequency band, and the envelope of the SAF is completely different from that of the 1DA. Therefore, precise evaluation of the SAF is critical for rational seismic design.

ACKNOWLEDGMENT

Strong-motion records and ground data were provided by the National Research Institute for Earth Science and Disaster Resilience. Data-sorting was aided by Yusuke Yamane. Yasuhiro Fukushima helped us evaluate the hypocentral distances.

REFERENCES

[1] N. A. Abrahamson, "State Of The Practice Of Seismic Hazard Evaluation," presented at the ISRM International Symposium, Nov. 2000, Art. no. ISRM-IS-2000-014.

[2] T. Nagao and Y. Fukushima, "Source- and Site-Specific Earthquake Ground Motions: Application of a State-of-the-Art Evaluation Method," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5882–5888, Aug. 2020, https://doi.org/10.48084/etasr.3612.

[3] N. Morikawa *et al.*, "Strong motion uncertainty determined from observed records by dense network in Japan," *Journal of Seismology*, vol. 12, no. 4, pp. 529–546, Oct. 2008, https://doi.org/10.1007/s10950-008-9106-2.

[4] K. Kato, K. Aki, and M. Takemura, "Site amplification from coda waves: Validation and application to S-wave site response," *Bulletin of the Seismological Society of America*, vol. 85, no. 2, pp. 467–477, Apr. 1995.

[5] T. Satoh, H. Kawase, and T. Sato, "Evaluation of local site effects and their removal from borehole records observed in the Sendai region, Japan," *Bulletin of the Seismological Society of America*, vol. 85, no. 6, pp. 1770–1789, Dec. 1995.

[6] C.-H. Kuo, K.-L. Wen, C.-M. Lin, N.-C. Hsiao, and D.-Y. Chen, "Site amplifications and the effect on local magnitude determination at stations of the surface–downhole network in Taiwan," *Soil Dynamics and Earthquake Engineering*, vol. 104, pp. 106–116, Jan. 2018, https://doi.org/10.1016/j.soildyn.2017.10.003.

[7] S. S. Bora, F. Cotton, F. Scherbaum, B. Edwards, and P. Traversa, "Stochastic source, path and site attenuation parameters and associated variabilities for shallow crustal European earthquakes," *Bulletin of*

Earthquake Engineering, vol. 15, no. 11, pp. 4531–4561, Nov. 2017, https://doi.org/10.1007/s10518-017-0167-x.

[8] T. Nagao, "Seismic Amplification by Deep Subsurface and Proposal of a New Proxy," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5157–5163, Feb. 2020, https://doi.org/10.48084/etasr.3276.

[9] T. Nagao, "Maximum Credible Earthquake Ground Motions with Focus on Site Amplification due to Deep Subsurface," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6873–6881, Apr. 2021, https://doi.org/10.48084/etasr.3991.

[10] Y. Fukushima and T. Nagao, "Variation of Earthquake Ground Motions with Focus on Site Amplification Factors: A Case Study," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4355–4360, Aug. 2019, https://doi.org/10.48084/etasr.2856.

[11] T. Nagao and P. Lu, "A simplified reliability estimation method for pile-supported wharf on the residual displacement by earthquake," *Soil Dynamics and Earthquake Engineering*, vol. 129, Feb. 2020, Art. no. 105904, https://doi.org/10.1016/j.soildyn.2019.105904.

[12] D. J. Andrews, "Objective Determination of Source Parameters and Similarity of Earthquakes of Different Size," in *Earthquake Source Mechanics*, Washington DC, USA: American Geophysical Union (AGU), 1986, pp. 259–267, https://doi.org/ 10.1029/GM037p0259.

[13] T. Iwata and K. Irikura, "Source Parameters of the 1983 Japan Sea Earthquake Sequence," *Journal of Physics of the Earth*, vol. 36, no. 4, pp. 155–184, 1988, https://doi.org/10.4294/jpe1952.36.155.

[14] S. Aoi, K. Obara, S. Hori, K. Kasahara, and Y. Okada, "New strong-motion observation network : KiK-net," *Eos, Transactions, American Geophysical Union*, vol. 81, 2000.

[15] *Strong-motion seismograph networks (K-NET, KiK-net).* https://www.kyoshin.bosai.go.jp/ (accessed Sep. 10, 2021).

[16] H. Fujiwara, S. Kawai, K. X. Hao, N. Morikawa, and H. Azuma, "J-SHIS - An integrated system for sharing information on national seismic hazard maps for Japan," in *16th World Conference on Earthquake,* Santiago Chile, Jan. 2017, Art. no. 0372.

[17] K. Aki, "Scaling law of seismic spectrum," *Journal of Geophysical Research (1896-1977)*, vol. 72, no. 4, pp. 1217–1231, 1967, https://doi.org/10.1029/JZ072i004p01217.

[18] K. Aki and P. G. Richards, *Quantitative Seismology*, 2nd ed. Mill Valley, CA, USA: University Science Books, 2009.

[19] S. Drouet, P. Triantafyllidis, A. Savvaidis, and N. Theodulidis, "Comparison of Site-Effects Estimation Methods Using the Lefkas, Greece, 2003 Earthquake Aftershocks," *Bulletin of the Seismological Society of America*, vol. 98, no. 5, pp. 2349–2363, Oct. 2008, https://doi.org/10.1785/0120080004.

[20] J. H. Steidl, A. G. Tumarkin, and R. J. Archuleta, "What is a reference site?," *Bulletin of the Seismological Society of America*, vol. 86, no. 6, pp. 1733–1748, Dec. 1996.

[21] Y. Fukushima, T. Nagao, J. Oshige, and I. Suetomi, "Ground motion evaluation for intra-plate earthquake by different site amplification factors and source models," in *7th International Conference on Earthquake Geotechnical Engineering*, Roma, Italy, Jun. 2019, pp. 2484–2492.

[22] K. Yoshimoto, H. Sato, and M. Ohtake, "Frequency-dependent attenuation of P and S waves in the Kanto area, Japan, based on the coda-normalization method," *Geophysical Journal International*, vol. 114, no. 1, pp. 165–174, Jul. 1993, https://doi.org/10.1111/j.1365-246X.1993.tb01476.x.

[23] A. Moya, J. Aguirre, and K. Irikura, "Inversion of Source Parameters and Site Effects from Strong Ground Motion Records using Genetic Algorithms," *Bulletin of the Seismological Society of America*, vol. 90, no. 4, pp. 977–992, Aug. 2000, https://doi.org/10.1785/0119990007.

Drag Reduction Using Biomimetic Sharkskin Denticles

Dinesh Bhatia
School of Aerospace
University of Nottingham Ningbo China
Ningbo, China
dinesh.Bhatia@nottingham.edu.cn

Yingxue Zhao
School of Aerospace
University of Nottingham Ningbo China
Ningbo, China
zhaoyingxue1998@163.com

Devinder Yadav
School of Aerospace
University of Nottingham Ningbo China
Ningbo, China
Devinder.Yadav@nottingham.edu.cn

Jian Wang
Faculty of Science, Engineering and Computing
Kingston University
London, U.K.
J.Wang@kingston.ac.uk.

Abstract-This paper explores the use of sharkskin in improving the aerodynamic performance of aerofoils. A biomimetic analysis of the sharkskin denticles was conducted and the denticles were incorporated on the surface of a 2-Dimensional (2D) NACA0012 aerofoil. The aerodynamic performance including the drag reduction rate, lift enhancement rate, and Lift to Drag (L/D) enhancement rate for sharkskin denticles were calculated at different locations along the chord line of the aerofoil and at different Angles of Attack (AOAs) through Computational Fluid Dynamics (CFD). Two different denticle orientations were tested. Conditional results indicate that the denticle reduces drag by 4.3% and attains an L/D enhancement ratio of 3.6%.

Keywords-biomimetics; sharkskin denticles; flow control; drag reduction; CFD

I. INTRODUCTION

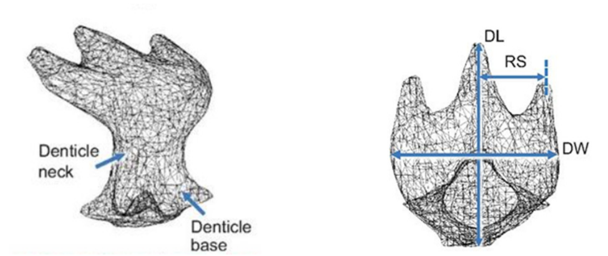
With the rapid growth of aviation in developing countries and the continuing growth of the existing aviation markets in developed countries, more people can enjoy air travels. The growing need for passengers has caused further environmental concerns since aircraft emissions can alter the atmospheric concentration of greenhouse gases that are related to climate change [1]. Moreover, economic concerns have been raised by airline companies. The supply of fossil fuels is limited and is depleting [2], therefore, much higher fuel prices and volatility occur [3]. The increase in jet fuel prices would increase air traffic and airport operation costs. Therefore, the aviation industry is focusing on environmental-friendly and green solutions to offset increasing fuel costs and higher carbon emissions. Drag reduction and improvement in aerodynamic efficiency is one way of reducing carbon emissions and fuel consumption [4].

Through the evolution of living organisms, there have been many examples in nature with traits which offer some insights into the design and optimization of engineering systems. The sharskin is one of these optimized structures. It has evolved to ensure that sharks can swim efficiently. Biomimetic analysis of the sharskin denticle structure has become a popular research topic. Biomimicry of the sharskin denticles has yielded drag reduction up to 10% compared with the corresponding smooth surface [5]. Applications inspired by the sharskin denticles have already been used in the Olympics. It was reported that the sharskin inspired speedo-swimsuit worn by Michael Phelps, the famous 8 times gold medals winner at the 2008 Olympics, helped in achieving 3-4% of drag reduction for swimmers [6]. This paper aims to investigate the impact of incorporating sharskin on 2D NACA0012 aerofoil by placing the denticles at strategic locations on the aerofoil. A detailed biomimetic study of the sharskin is conducted, following which the denticle sizing and modeling process are determined. Simulations are conducted and the results of sharskin on lift coefficient (cl), drag coefficient (cd), and L/D ratio at different positions along chord line at different AOAs are presented. Furthermore, the sharskin denticle working principle towards drag reduction and lift enhancement was analyzed to highlight the flow mechanism.

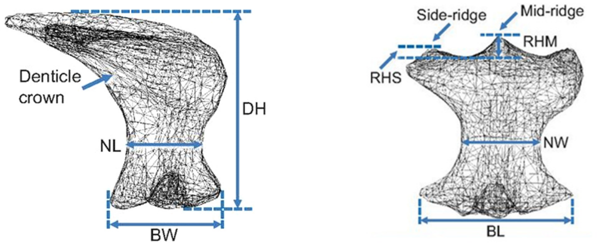
II. BIOMIMETIC STUDY OF SHARSKIN DENTICLES

Sharskin is one of the most popular bionic study topics for flow control and drag reduction and will be introduced in greater detail in this section. In order to observe the flow control mechanisms of sharskin denticles, the denticle geometry was first investigated. Sharskin is covered with rigid bony denticles. According to Figure 1(c), a sharskin denticle has a plate-like upper section with several ridges and narrows to a thin neck that roots into the skin [7]. The sharskin selected in this project belongs to a species of shortfin mako sharks called Isurus Oxyrinchus. This kind of shark is one of the fastest marine fish reaching a maximum speed of 70km·h^{-1} [8]. Some of the most detailed descriptions are shown in Figure 1 [7] and Table I [8].

Corresponding author: Dinesh Bhatia



(a) Sharskin denticle layout (b) Top view of a shark denticle



(c) Side view of a shark denticle (d) Front view of a shark denticle

Fig. 1. Micro-CT scanned sharskin denticle geometry. Reproduced with permission from the Journal of Experimental Biology 217, 10 (2014). Copyright 2014 Company of Biologist Ltd. [7].

TABLE I. SHARSKIN DENTICLE PROTOTYPE DIMENSIONS [8]

Components	Dimensions (μm)
DL (denticle length)	151
DW (denticle width)	125
RS (spacing between adjacent ridges)	51
DH (denticle height)	113
BW (denticle base width)	119
NL (denticle neck length)	45.1
NW (denticle neck width)	50.9
BL (denticle base length)	83.8
RHM (height of middle ridge)	21
RHS (height of side ridge)	11

III. METHODOLOGY

The methodology adopted in this paper was divided into three aspects. Initially, the sharskin was modeled to ensure dynamic similarity and effective translation of the denticles from the hydrodynamic to an aerodynamic environment. Computational Aided Design (CAD) modeling of sharskin denticles was conducted and incorporated on the NACA0012 aerofoil. The CFD simulations of sharskin denticles and aerofoil were performed with ANSYS FLUENT.

A. Sizing of Sharskin Denticles

In order to quantify and analyze the sharskin aerodynamic properties, aerodynamic scaling was conducted to ensure that appropriate sizing of the shark-skin denticles could be incorporated on the aerofoil. Key dynamic similarity parameters such as the AOA, Mach Number (Ma), and Reynolds Number (Re) for the sharskin denticle model and real sharskin denticles should be computed to ensure their similarity [9]. Sizing was conducted with the aim to perform experimental testing of the sharskin denticle post numerical simulations. As a result, dynamic similarity exercise was conducted, keeping in mind the wind tunnel specifications, at the University of Nottingham Ningbo China (UNNC).

TABLE II. WIND TUNNEL SPECIFICATIONS AND SHARK HABITAT CONDITIONS

AF1300 wind tunnel UNNC	Isurus Oxyrinchus habitat and speed limit [10]
Temperature range: 5°C to 40°C	Prefers to inhabit in offshore waters with temperatures from 17 to 20°C
Nominal air velocity: 0 to 36m/s.	Maximum speed of 70km·h⁻¹.

According to the given specifications from the wind tunnel manual and shark habitat, it is noted that the maximum Ma in wind tunnel is less than 1, which means the effects of compressibility are very small, thereby Ma computation is no longer required. Therefore, by deriving the Re equations of the sharskin denticle model and the Re for real sharskin denticle prototype, given the known data of air and sea water density, shark velocity in seawater, air velocity designed in simulation, and the dynamic viscosity of sea water and air, the scale factor of sharskin denticle model dimensions versus the real sharskin denticle prototype dimensions were calculated. The derivation of equations is given below.

$$Re_{\infty} = \frac{\rho_{\infty} V_{\infty} l_{\infty}}{\mu_{\infty}} = Re_T = \frac{\rho_T V_T l_T}{\mu_T} \quad (1)$$

where Re_{∞} is the Reynolds number for real shark denticles in sea water; Re_T is the Reynolds number for the shark denticle model in the air, ρ_{∞} is the mean density of seawater at 17-20°C which is 1025.19kg/m^3 [10], ρ_T is the mean density of air at 20°C (1.2kg/m^3), V_{∞} is the fastest velocity of Isurus Oxyrinchus (19.4m/s), V_T is the simulated air velocity (10m/s), μ_{∞} is the mean viscosity of seawater at 17-20°C (0.001Pa/s), μ_T is the viscosity of air at 20°C (1.81×10^{-5}Pa/s), l_{∞} is the height of the sharskin denticle prototype (DH=113×10^{-6}m) [7], and l_T is the height of the sharskin denticle model (m).

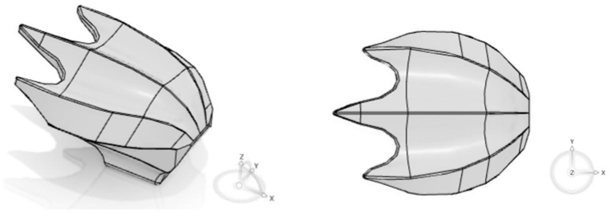
$$\text{scale factor} = \frac{l_T}{l_{\infty}} = \frac{\rho_{\infty} V_{\infty} \mu_T}{\rho_T V_T \mu_{\infty}} = 30 \quad (2)$$

Thus, when the wind speed was chosen to be 10m/s, the scale factor for shark denticle model dimension versus the real shark denticle dimension was calculated as 30 in (2). The detailed sharskin denticle model dimensions are listed in Table II.

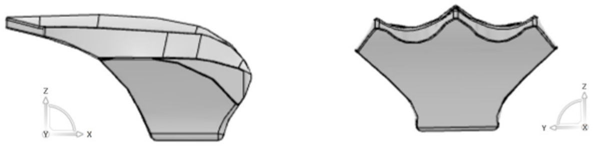
TABLE III. SHARSKIN DENTICLE MODEL DIMENSIONS

Components	Dimensions (μm)
DL	4530
DW	3750
RS	1530
DH	3390
BW	3570
NL	1353
NW	1527
BL	2514
RHM	630
RHS	330

After calculating the sharskin denticle model dimensions, the modeling process using 3DExperience part design and generative wireframe & surface was performed. The sharskin denticle is shown in Figure 2 below.



(a) Isometric view of a sharskin denticle (b) Top view of a shark denticle



(c) Side view a of sharskin denticle (d) Rear view of a shark denticle

Fig. 2. Sharskin denticle model geometry.

B. 2D simulation Setup and Validation

Once the sharskin and NACA0012 aerofoil were modeled, a comparative study was carried out between a standard NACA0012 aerofoil and the NACA0012 aerofoil with the sharskin denticles incorporated on it. The simulation case studies included 2D simulations of the NACA0012 aerofoil with the sharskin denticles put in the normal direction (as observed on the sharks) and NACA0012 aerofoil with the denticles put in the reverse direction as shown in Figure 3(c)-(d). The purpose of this simulation was to find the drag reduction rate, the L/D enhancement rate for the experiment group when sharkskin was put at different positions along the aerofoil chord line, and when the aerofoil was rotated at different AOAs. To eliminate the other factors that may interfere with the simulation results, all simulation settings were set to be the same for both the conventional aerofoil and the simulated denticle cases. An unstructured mesh of triangular shape throughout the control domain with 160,000 elements was chosen based on the mesh independent study. The mesh possesses high density elements clustered around the aerofoil surface as shown in Figure 3(a). The dimensions of the mesh and the boundary conditions were chosen from [11]. A C-shaped domain was used with the velocity inlet being placed at 10 times chord length (10c) from the aerofoil leading edge and 16c from the trailing edge. The height of the domain is 10c. The pressure outlet was placed 16c from the aerofoil trailing edge. A no-slip boundary condition was used for the aerofoil wall. Mesh elements were clustered around the aerofoil and denticle surface to ensure accurate resolution of flow features (Figure 3(b)). An inflation layer was adopted with a minimum first layer height of 5×10^{-7}m, 10 maximum layers, and a growth rate of 1.4 maintained from the aerofoil and denticle wall. A y+ value of 0.1 was maintained. All simulations were performed in ANSYS FLUENT. Solver settings were chosen based on [11]. The SST k-omega model was selected. It has shown great efficacy and accuracy in simulating micro-features at a scale similar to the sharskin denticles' and has accurately depicted the flow structure around these microfeatures [4]. As a result, the authors believe that the computational setup is suitable, reliable, and can accurately depict the flow features for the sharskin denticles and aerofoil presented in this paper.

TABLE IV. DETAILED COMPUTATIONAL SETUP

Grid	**Unstructured triangular 160,000 elements**
CFD Model	SST k-omega 2 Eqns
Solver	Steady state pressure based
Scheme	Second order coupled for pressure, momentum, and turbulence. Least square cell-based discretization for the gradient
Input parameters	Velocity: 10 ms−1, Turbulence intensity (Tu): 5%

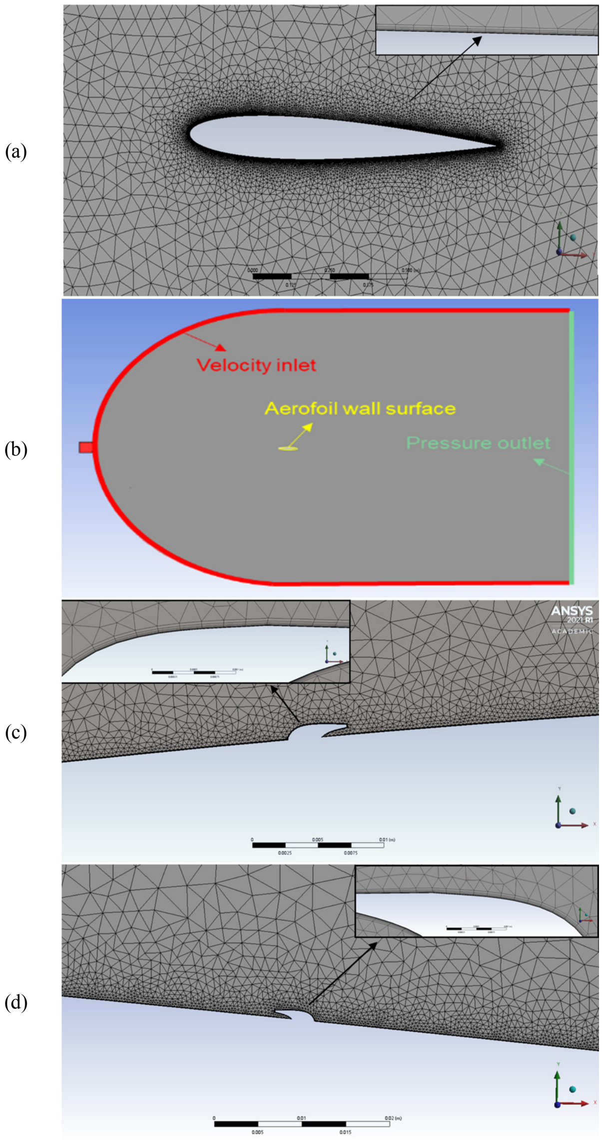


Fig. 3. Sharskin denticle model geometry. (a) NACA0012 control group, (b) simulation boundary conditions, (c) normal direction arrangement, (d) reverse direction arrangement

C. Validation and Reliability of Simulation Results

In order to ensure accurate simulation results, the result validation process was necessary, which included mesh independence check, benchmarking, and experimental result verification. It was assured that the sharskin denticles were modeled, meshed, and simulated according to the standard procedure, and that the simulation results were reliable. The cd value was monitored as the number of mesh elements increased. It can be observed from Figure 4 that the number of elements increased from 0.2×10^5 to 1.6×10^5 before the condition of steady cd value was met. Thus, the corresponding element size of 0.003c was selected for edge mesh sizing, and 0.05c was chosen for body mesh sizing.

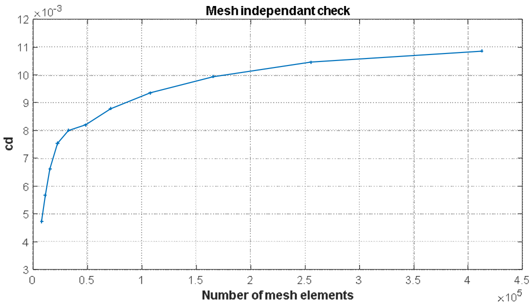


Fig. 4. Mesh independent check.

Moreover, the simulation data of cl and cd for smooth surface NACA0012 aerofoil was benchmarked with the existing data [11]. Figure 5 shows that the simulated results of cl and cd show good correlation with the existing simulation data.

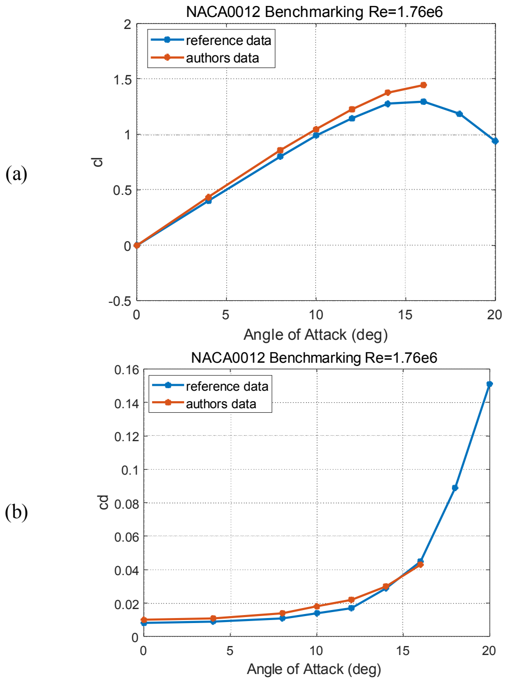


Fig. 5. Benchmarking results. (a) cl vs. AOA (°), (b) cd vs. AOA (°).

Lastly, the simulation results of NACA0012 aerofoil at *Re*=6×10^5 were compared with the X-Foil data at *Re*=6×10^5 [12]. Figure 6 shows that the cl and cd at different AOAs for the current study's simulation results are closely matched to the X-Foil data and show a very similar trend. Therefore, it can be concluded that the results presented for the NACA0012 aerofoil indicate reliability, repeatability, and accurate capture of the flow phenomenon by the applied simulation methods.

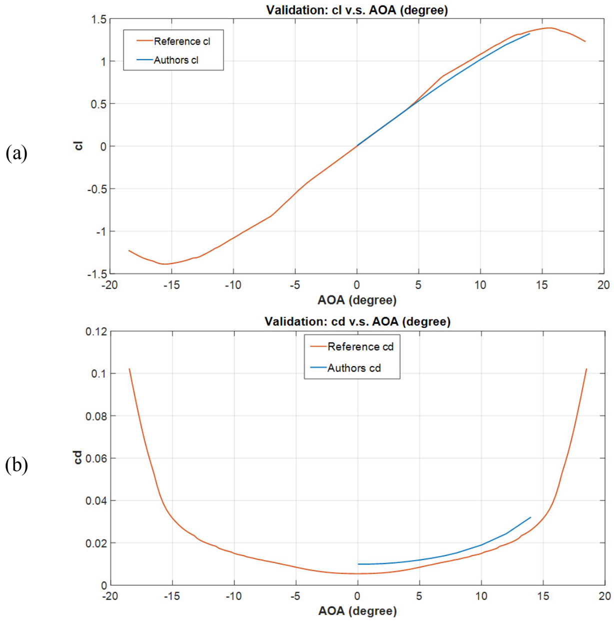


Fig. 6. Simulation verification with experimental data. (a) cl vs. AOA. (°), (b) cd vs. AOA (°).

IV. RESULTS AND DISCUSSION

A. Optimal Location of Sharskin Denticles

In this section, the simulation results of cl, cd, and L/D of sharskin denticles being put at normal and reverse directions are compared with the simulation results of a control group, i.e. NACA0012 aerofoil smooth surface at *Re*=6×10^5. The location of the denticles on chord line that has shown the highest amount of drag reduction and the highest amount of L/D enhancement is considered to be the optimal location of sharskin denticles.

1) Normal Direction Simulation Results

The denticles at normal direction were put on different locations, from 0.1c to 0.7c with a step change of 0.1c on NACA0012 aerofoil. Drag reduction rate, lift enhancement rate, and L/D enhancement rate were calculated and shown in Figures 7-9. Figures 7 and 9 show the drag reduction rate and L/D enhancement rate versus AOA (degrees). It is observed from Figures 7 and 9 that the optimal position of sharskin denticles at the normal direction on NACA0012 aerofoil for drag reduction and L/D enhancement is 0.16c, achieving maximum drag reduction and L/D enhancement of 3% and 1.5% achieved at 0° AOA and 4° AOA respectively. Moreover, it is observed from Figure 8 that the Lift enhancement rates for normal direction simulation at 1° to 2° AOA are all negative. At an AOA above 2°, the lift coefficient does not improve much, ranging between -1% to 1%. Furthermore, it was found that for normal direction simulations, promising effects of drag reduction and L/D enhancement tend to occur when the denticles were put close to the leading edge and at low AOAs.

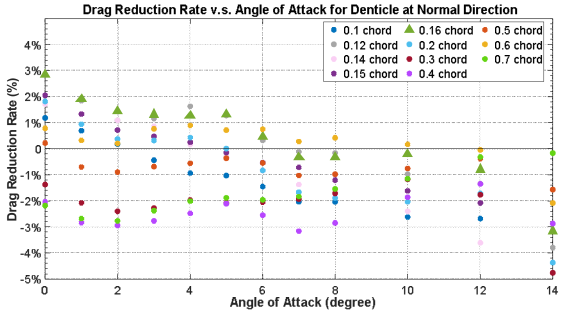


Fig. 7. Drag reduction rate vs. AOA for denticle at normal direction.

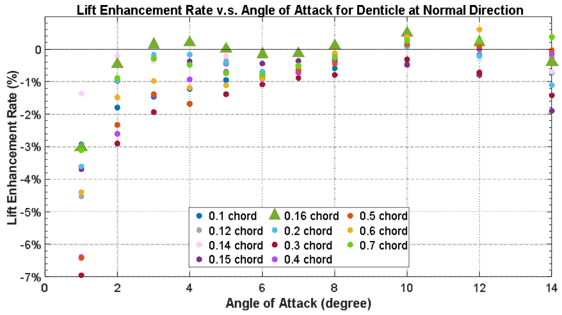


Fig. 8. Lift enhancement rate vs. AOA for denticle at normal direction.

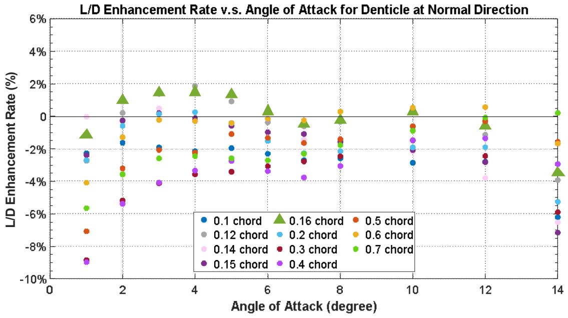


Fig. 9. L/D enhancement rate vs. AOA for denticle at normal direction.

2) *Reverse Direction Simulation Results*

Since the denticle placing in the normal direction in experiment group 1 is not very promising at higher AOAs, in order to further explore the performance of this denticle model in drag reduction and L/D enhancement, the sharskin denticles were put in the reverse direction to discover their effect. The sarkskin were put on different locations, from 0.1c to 0.7c on the NACA0012 aerofoil. The drag reduction coefficient, lift enhancement coefficient, and L/D enhancement coefficient of denticles placed in reverse direction were calculated and shown Figures 10-12.

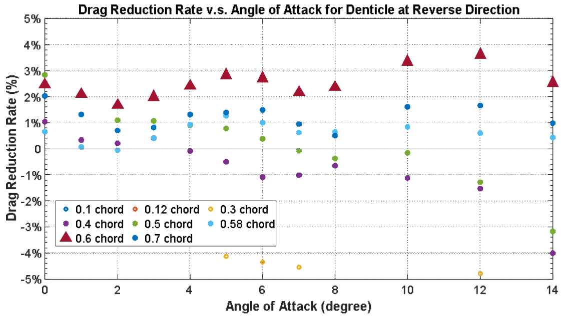


Fig. 10. Drag reduction rate vs. AOA at reverse direction.

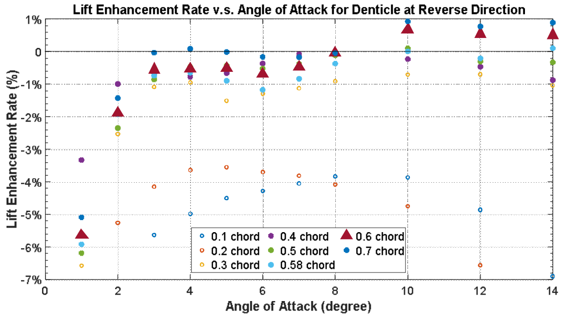


Fig. 11. Lift enhancement rate vs. AOA at reverse direction.

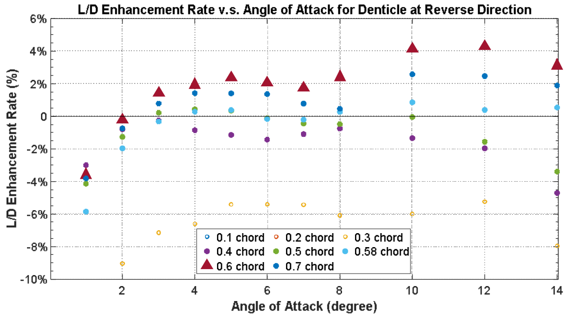


Fig. 12. L/D enhancement rate vs. AOA for denticle at reverse direction.

From Figures 10 and 12, the optimal position of sharskin at the reverse direction on NACA0012 aerofoil for drag reduction and L/D enhancement is found to be 0.6c. The maximum drag reduction and L/D enhancement for denticles placed in the reverse direction at 0.6c are 4.3% and 3.6% achieved both at 12° AOA. Moreover, the lift enhancement rate for reverse direction simulations varies from -1% to 1% after 2-degree AOA, as seen in Figure 11. Additionally, it was discovered that for the reverse direction arrangement, promising effects of drag reduction and L/D enhancement occur mostly when the denticles were placed close to the trailing edge at high AOAs.

B. *Sharskin Denticle Flow Mechanism*

Moreover, to find out the sharskin denticle drag reduction mechanism, the optimal solution of sharskin placed at the normal direction at 0.16c at 0° AOA and the optimal solution of sharskin in reverse direction at 0.6c at 12° AOA were compared with the control group of NACA0012 aerofoil at 0° and 12° AOA. The reason for this comparison was to make certain that the best scenario was analyzed.

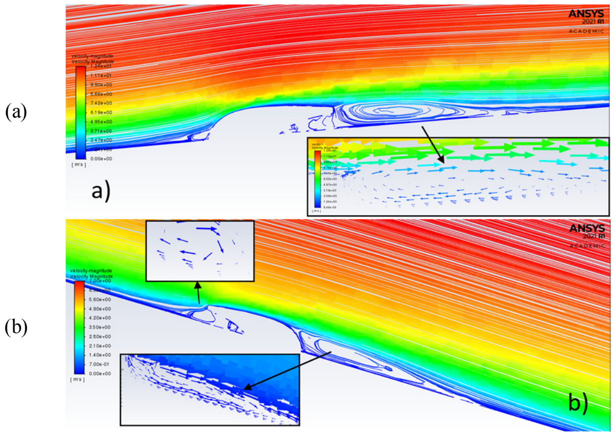


Fig. 13. Velocity magnitude pathline plot. (a) 0.16c normal direction at 0° AOA, (b) 0.6c reverse direction at 12° AOA.

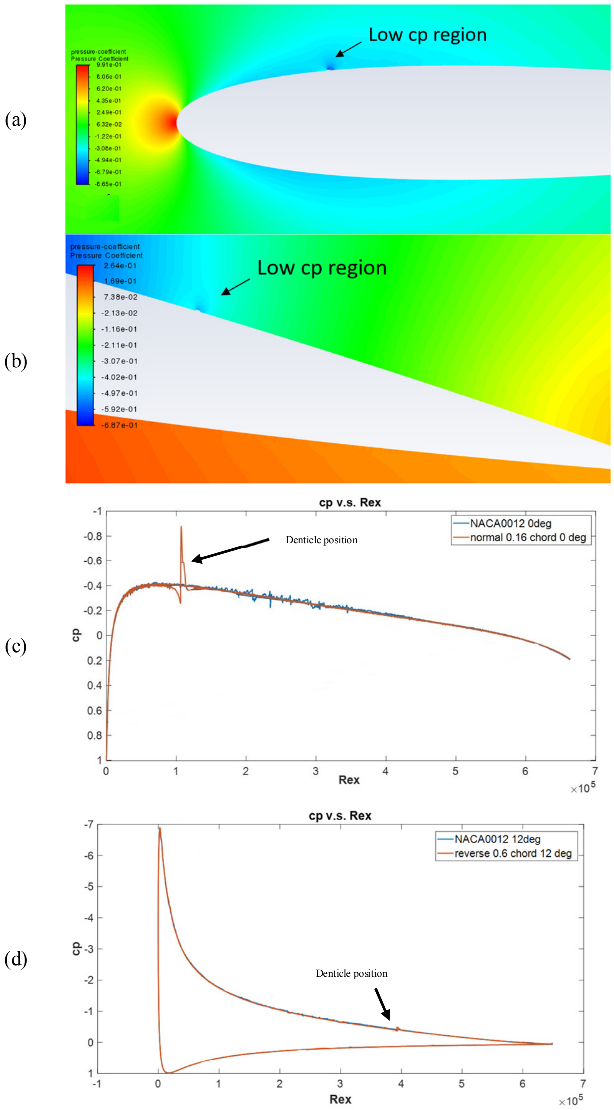


Fig. 14. Pressure coefficient contour plot and cp v.s. Rex plots. (a) 0.16c normal direction 0° AOA, (b) 0.6c reverse direction 12° AOA, (c) 0.16c normal direction 0° AOA, (d) 0.6c reverse direction 12° AOA.

The separation bubble at both the front and rear of the sharskin denticles are discovered from Figure 13. The separation bubble generated extra suction of airflow along the streamwise direction. Through the separation and reattachment of separation bubbles, the airflow accelerates and tends to stick closer to the aerofoil profile and thus it exhibits better aerodynamic performance. Moreover, from Figure 14(a)-(b), the denticles also generated a low-pressure zone at the top and downstream of the sharskin. The low-pressure zone has caused a greater pressure difference between the upper and lower surfaces of the aerofoil, which has a positive effect on lift enhancement. Meanwhile, since the pressure coefficient further downstream of the sharskin is bigger at the sharskin low-pressure zone, a favorable pressure difference along the streamwise direction is formed. This favorable pressure difference would eliminate some of the drag forces. This favorable pressure gradient can be numerically observed in Figure 14(c)-(d) where the cp for optimal solution of sharskin denticles is compared with the cp from the control group of NACA0012 smooth surface.

Additionally, the skin friction coefficient also suggested the drag reduction effect of sharskin denticles. It is observed from Figure 15(a)-(b) that the skin friction coefficient highly fluctuated in the NACA0012 control group. The skin friction coefficient after applying the sharskin is smoother than that from the control group at the same AOA, which suggests that the airflow was attached closer to the aerofoil surface after flowing through the sharskin denticles, causing less skin friction coefficient and less turbulent separation bubbles, therefore the aerodynamic efficiency of the aerofoil was improved.

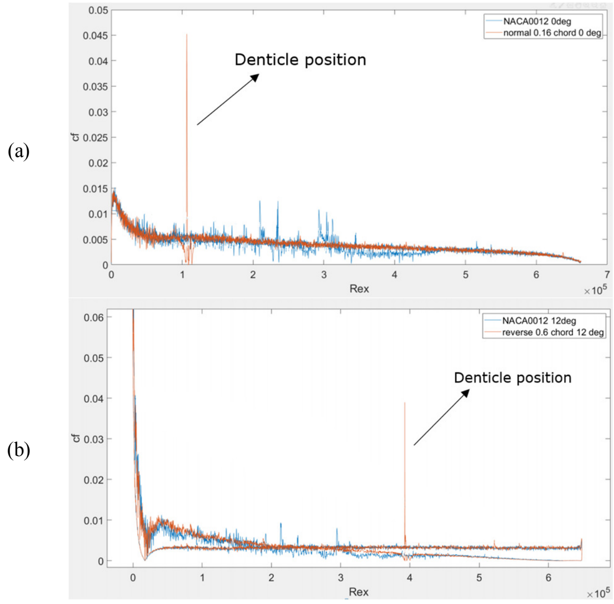


Fig. 15. Skin friction coefficient comparison with NACA0012 aerofoil. (a) 0.16c normal direction 0° AOA, (b) 0.6c reverse direction 12° AOA.

D. Undesired Pressure Gradient

Except from the sharskin denticle mechanism in drag reduction, an important feature of high-pressure region right upstream of the sharskin position was also noticed.

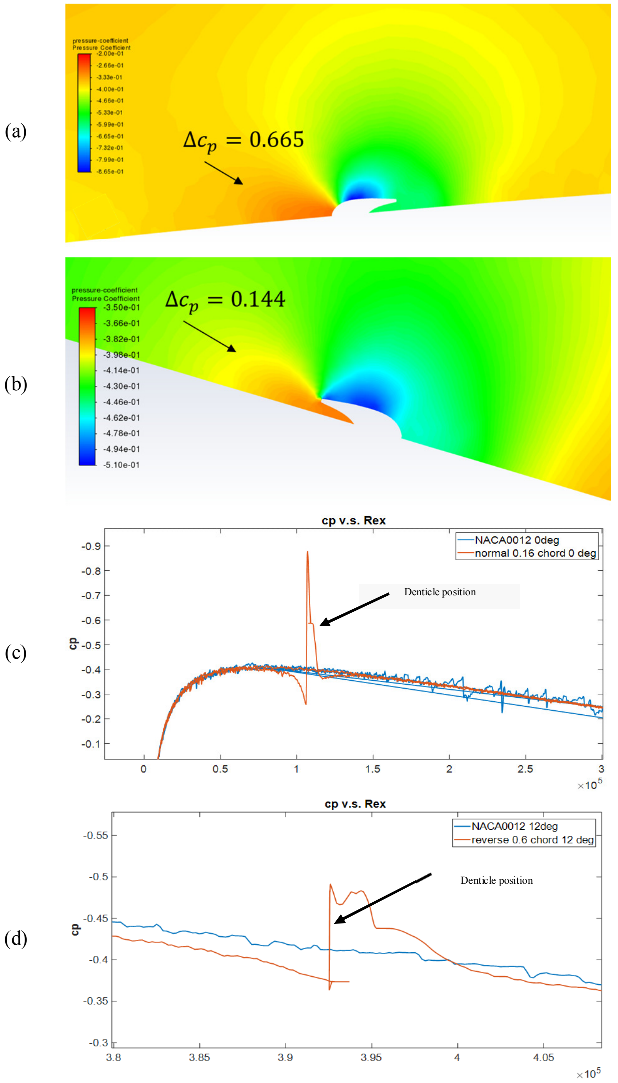


Fig. 16. Comparison between the optimal denticle simulation locations and NACA0012 control group. (a) Pressure coefficient contour, 0.16c normal direction 0° AOA, (b) pressure coefficient contour 0.6c reverse direction 12° AOA, (c) cp vs. Rex for normal direction 0.16c and NACA0012 at 0° AOA, (d) cp vs. Rex for reverse direction 0.6c and NACA0012 at 12° AOA.

According to Figure 16(a)-(b), the sharskin denticles at both normal and reverse directions have a high-pressure zone at the upstream of the sharskin which may be caused by the blockage phenomena of denticle frontal area when air flows through the denticles. This high-pressure zone would create an unfavorable pressure difference along the streamwise direction that tends to generate more drag, thus weakens the sharskin drag reduction effect. For the optimized case of normal direction in experimental group 1, the pressure difference is roughly 0.665 according to Figure 16(c), while for the optimized case of reverse direction in experimental group 2, the pressure difference is around 0.144 according to Figure 16(d). The higher the pressure difference between the upstream and downstream of sharskin is, the less drag reduction effect it would have. The reason why the reverse direction has a better drag reduction performance than the normal direction may be that the sharskin at reverse direction have a smaller unfavorable pressure difference compared to the normal direction case.

V. CONCLUSIONS

In this paper, simulations and wind tunnel experiments were conducted to attain drag reduction and L/D enhancement using strategically placed sharkskin within the laminar boundary layer. Based on the results, the following conclusions can be drawn:

- During the simulations, the maximum drag reduction and L/D enhancement, for normal direction and 0.16c sharskin position, are 3% and 1.5% achieved at 0^o AOA and 4^o AOA respectively. While the maximum drag reduction and L/D enhancement, for the reverse direction at 0.6c sharskin position, are 4.3% and 3.6% achieved both at 12^o AOA.
- The presence of the sharskin denticles results in a favorable pressure gradient aft of the sharskin which leads to a damping of disturbances within the boundary layer. This leads to transition delay and drag reduction.
- The presence of the sharskin denticles also results in an undesired pressure gradient and weakens the drag reduction effect due to the blockage phenomena of sharskin frontal area.
- As the shape of the sharskin denticles has not been optimized, a study to identify an optimized the sharskin frontal area design closely would result in a smaller unfavorable pressure gradient and greater drag reduction.

ACKNOWLEDGEMENTS

The authors wish to thank the Ningbo Science and Technology Bureau for their funding under the Ningbo Natural Science Foundation program with project code: 2019A610116. The authors would also like to thank the University of Nottingham Ningbo China and Kingston University London.

REFERENCES

[1] D. A. Pamplona and C. J. P. Alves, "Civil Aircraft Emissions Study and Pollutant Forecasting at a Brazilian Airport," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5217–5220, Feb. 2020, https://doi.org/10.48084/etasr.3227.

[2] A. A. Khaskheli, G. D. Walasai, A. S. Jamali, Q. B. Jamali, Z. A. Siyal, and A. Mengal, "Performance Evaluation of Locally-Produced Waste Cooking Oil Biodiesel with Conventional Diesel Fuel," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3521–3524, Dec. 2018, https://doi.org/10.48084/etasr.2333.

[3] "U.S. Product Supplied for Crude Oil and Petroleum Products," *US Energy Information Administration*. https://www.eia.gov/dnav/pet/pet_cons_psup_dc_nus_mbbl_m.htm (accessed Sep. 10, 2021).

[4] D. Bhatia *et al.*, "Transition Delay and Drag Reduction using Biomimetically Inspired Surface Waves," *Journal of Applied Fluid Mechanics*, vol. 13, no. 4, pp. 1207–1222, Dec. 2019, https://doi.org/10.36884/jafm.13.04.30316.

[5] B. Dean and B. Bhushan, "Shark-skin surfaces for fluid-drag reduction in turbulent flow: a review," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1929, pp. 4775–4806, Oct. 2010, https://doi.org/10.1098/rsta.2010.0201.

[6] G. D. Bixler and B. Bhushan, "Fluid Drag Reduction with Shark-Skin Riblet Inspired Microstructured Surfaces," *Advanced Functional Materials*, vol. 23, no. 36, pp. 4507–4528, 2013, https://doi.org/10.1002/adfm.201203683.

[7] L. Wen, J. C. Weaver, and G. V. Lauder, "Biomimetic shark skin: design, fabrication and hydrodynamic function," *Journal of Experimental Biology*, vol. 217, no. 10, pp. 1656–1666, May 2014, https://doi.org/10.1242/jeb.097097.

[8] F.-W. Patricia, D. Guzman, B. Iñigo, I. Urtzi, B. J. Maria, and S. Manu, "Morphological Characterization and Hydrodynamic Behavior of Shortfin Mako Shark (Isurus oxyrinchus) Dorsal Fin Denticles," *Journal of Bionic Engineering*, vol. 16, no. 4, pp. 730–741, Jul. 2019, https://doi.org/10.1007/s42235-019-0059-7.

[9] D. Bolster, R. E. Hershberger, and R. J. Donnelly, "Dynamic similarity, the dimensionless science," *Physics Today*, vol. 64, no. 9, pp. 42–47, Sep. 2011, https://doi.org/10.1063/PT.3.1258.

[10] F. L. Bachleda, *Dangerous Wildlife in California & Nevada: A Guide to Safe Encounters At Home and in the Wild*, 1st ed. Birmingham, AL, USA: Menasha Ridge Press, 2002.

[11] *Simulation CFD External Flow Validation: NACA 0012 Airfoil.* Autodesk, 2015.

[12] "Polars for NACA 0012 AIRFOILS (n0012-il)," *Airfoil Tools.* http://airfoiltools.com/airfoil/details?airfoil=n0012-il#polars (accessed Sep. 10, 2021).

Application of An Adaptive Network-based Fuzzy Inference System to Control a Hybrid Solar and Wind Grid-Tie Inverter

Dinh-Nhon Truong
Faculty of Electrical and Electronics Engineering
Ho Chi Minh City University of Technology and Education Ho Chi Minh City, Vietnam
nhontd@hcmute.edu.vn

Van-Thuyen Ngo
Faculty of Electrical and Electronics Engineering
Ho Chi Minh City University of Technology and Education Ho Chi Minh City, Vietnam
thuyen.ngo@hcmute.edu.vn

Mi-Sa Nguyen Thi
Faculty of Electrical and Electronics Engineering
Ho Chi Minh City University of Technology and Education Ho Chi Minh City, Vietnam
misa@hcmute.edu.vn

An-Quoc Hoang
Faculty of Vehicle and Energy Engineering
Ho Chi Minh City University of Technology and Education Ho Chi Minh City, Vietnam
hanquoc@hcmute.edu.vn

Abstract-In this paper, the application of an Adaptive Network-based Fuzzy Inference System (ANFIS) to control a hybrid solar and wind grid-tie inverter in order to reduce power oscillations and enhance power quality is presented. To extract the maximum power from the PV system, a Perturb and Observe (P&O) algorithm is presented that tracks the Maximum Power Point (MPP). Time-domain simulation results of the studied system are performed in MATLAB/SIMULINK under different operating conditions such as changing irradiation and short-circuit faults in the power grid. From the simulation results, it can be concluded that the designed ANFIS controller and the proposed P&O algorithm perform better than the traditional PI controller and improve transient responses under severe operating conditions.

Keywords-hybrid; grid-tie inverter; perturb and observe; ANFIS

I. INTRODUCTION

Sunlight and wind are the furthermost promising renewable energy sources. Due to the randomness of the solar irradiance and the accessibility of the wind, a combination of solar and wind energy production configuration can be a highly reliable source of electrical energy. A general planning framework for integrating solar and wind energies in a Hybrid Power System (HPS) was proposed to exploit the solar-wind complementarity and to stabilize the combined power output on a case study in Pakistan in [1]. HPS techno-economics for a single-family residence in the region of Yambol, Bulgaria was analyzed in [2] to determine optimum system configuration to minimize the excess energy produced. A hybrid optimization model for electric renewable software and its techno-economic feasibility to develop a hybrid wind-solar model was analyzed in [3]. The benefits of a water pumping power system using HPS in order to supply water to remote areas and rural zones have been studied in [4].

To generate electricity for the power grid, an inverter is required. In [5], a 3-level inverter-based grid-tie hybrid solar and wind energy system was presented with the mitigation of power quality problems. For enhancing the power quality problems caused by the switching of the small-scale grid-tie inverter under the operation states of solar systems, some solutions have been mentioned in [6]. In terms of developing a grid-tie inverter, a Shunt Series Switched Grid Tied Inverter (SSS-GTI) was proposed in [7]. This inverter structure can operate in shunt or cascade-connected mode. Another structure to improve efficiency by reducing the switching losses was introduced in [8]. Research and application of inverter controllers make an important contribution to the quality of the power supplied to the power system. Besides the traditional PI controllers [9, 10], a double closed-loop controller for voltage and current was adopted in [11], using sliding mode control [12]. Also, new and modern control techniques have been applied, such as the fuzzy logic controller [13].

The purpose of this paper is to present the applicability of the Adaptive Network-based Fuzzy Inference System (ANFIS) controller to improve the stability of the hybrid grid-tie inverter.

II. STUDIED SYSTEM CONFIGURATION

The studied system configuration is introduced in Figure 1. It includes a 5kWp rooftop solar PV system and a 3kW wind turbine system connected to the power grid through a hybrid grid-tie inverter [9, 14]. The mathematical models of the proposed system are described below.

Corresponding author: Dinh-Nhon Truong

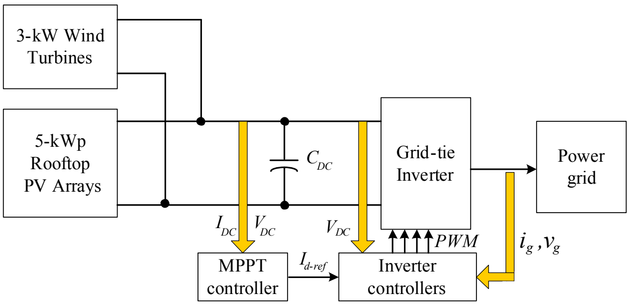


Fig. 1. Block diagram of the studied system.

A. PV Array Model

An equivalent circuit diagram of the 5kWp PV array is referred to [13, 15]. In this paper, a Sunpower SPR-415E-WHT-D PV panel is selected. The I-V curve for a rated 5kWp PV system affected by irradiance is shown in Figure 2. The rated parameters of this PV panel are:

- Maximum Power: 414.8W
- Cells per module: 128
- Voltage at Maximum Power Point (MPP): 85.3V
- Short-circuit current: 6.09A
- Current at MPP: 5.69A

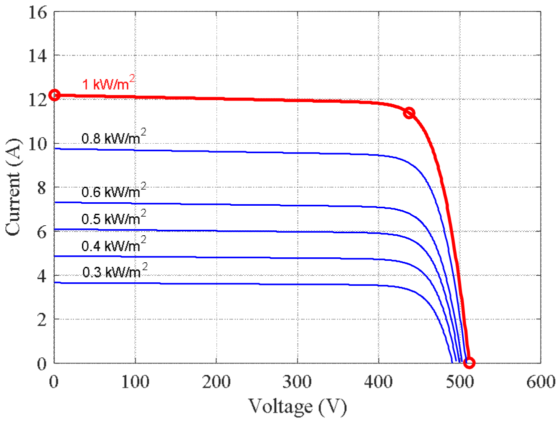


Fig. 2. Irradiance effect on the I-V curve.

B. The PMSG-based Wind Turbine Model

A small-scale Permanent magnet synchronous generator (PMSG)-based vertical wind turbine was used in this paper. The equivalent voltage equation of the studied wind PMSG in per unit (p.u.) projected to the *dq*-axis was expressed in [16, 17]. In this research, the MSFD5-3.000W wind turbine is selected with the following parameters:

- Rated power: 3.000W
- Max power: 3.500W
- Rated wind speed: 8m/s
- Generator type: Three-phase PMSG
- Working voltage: 192V DC
- Power supply method: Three-phase whole-wave bridge rectifier constant DC charger.

III. DESIGNED CONTROLLERS

A. P&O Algorithm for MPP Tracking

For optimizing the efficiency of the PV system, an MPP tracker algorithm using the Perturb and Observe (P&O) controller is applied [18, 19]. Its flowchart can be seen in Figure 3.

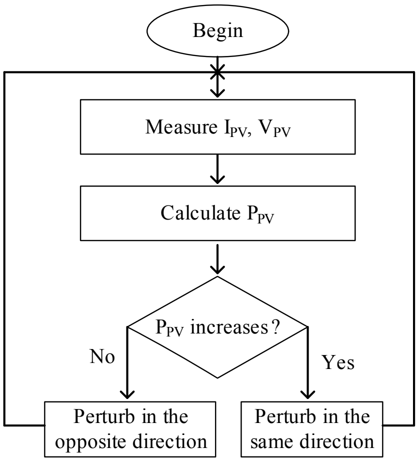


Fig. 3. Flowchart of the P&O algorithm for MPP tracking.

B. ANFIS Controller for DC Voltage Control

In this section, an ANFIS controller is proposed to replace the PI controller in the DC voltage controller block of the hybrid grid-tie inverter as presented in Figure 4 [20, 21].

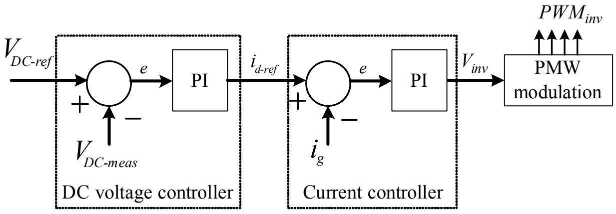


Fig. 4. Controllers for the hybrid grid-tie inverter.

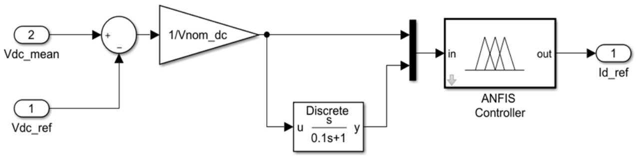


Fig. 5. The designed ANFIS controller for DC voltage regulator.

The structure of ANFIS can be seen in [22-24]. However, for enhancing the accuracy of this controller, besides the error of DC voltage (*e*), additional information from its derivative (d*e*/d*t*) is added to train the ANFIS controller (see Figure 5). As a result, after the training process, the minimal training root mean square error is RMSE = 0.000019. The surface rule is plotted in Figure 6. The selected ANFIS information are:

- Number of nodes: 35
- Number of linear parameters: 27
- Number of nonlinear parameters: 18
- Total number of parameters: 45
- Number of training data pairs: 302.401
- Number of fuzzy rules: 9

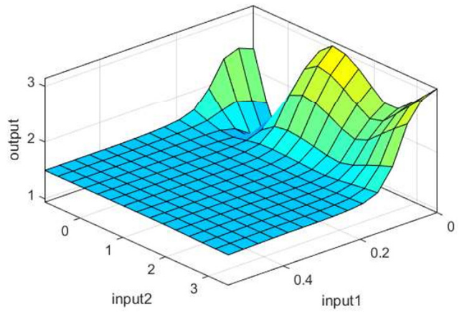


Fig. 6. Surface rule of the designed ANFIS.

IV. SIMULATION RESULTS

To evaluate the effectiveness of the proposed controller simulations were performed in the time-domain of the studied hybrid solar and wind system under different operating conditions to compare the contributions of the designed P&O and ANFIS controllers to control the hybrid grid-tie inverter and thus enhance dynamic stability. The simulations were carried out in MATLAB/SIMULINK. The mathematical model is shown in Figure 7 and the parameters of the studied system using PI controllers are illustrated in Figure 8.

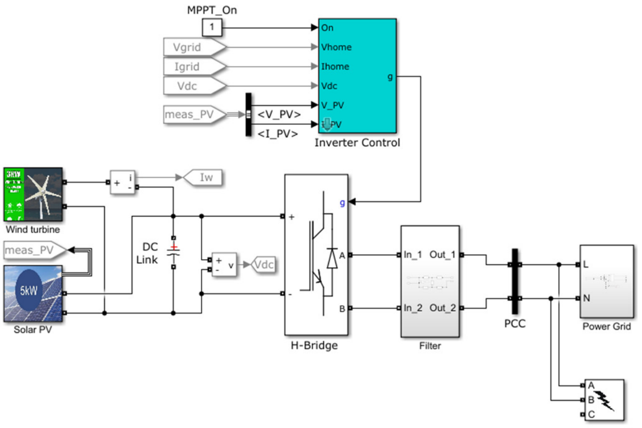
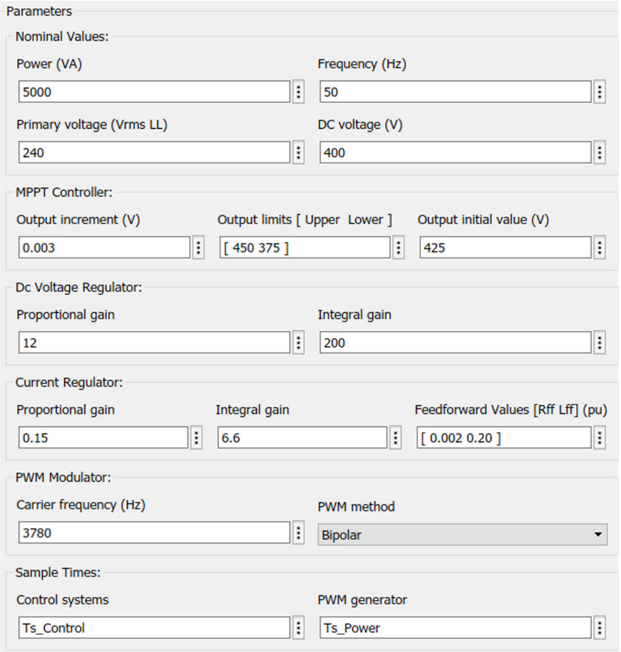


Fig. 7. The mathematical model in SIMULINK.

Figure 9(a) presents the response of the solar PV system when the solar irradiance changes and Figure 9(b) shows the output power of the solar PV system. It is easy to observe that the P&O controller helps keeping the output power of the solar PV system close to the peak value of the maximum power.



Parameters

Nominal Values:

Power (VA): 5000

Frequency (Hz): 50

Primary voltage (Vrms LL): 240

DC voltage (V): 400

MPPT Controller:

Output increment (V): 0.003

Output limits [Upper Lower]: [450 375]

Output initial value (V): 425

Dc Voltage Regulator:

Proportional gain: 12

Integral gain: 200

Current Regulator:

Proportional gain: 0.15

Integral gain: 6.6

Feedforward Values [Rff Lff] (pu): [0.002 0.20]

PWM Modulator:

Carrier frequency (Hz): 3780

PWM method: Bipolar

Sample Times:

Control systems: Ts_Control

PWM generator: Ts_Power

Fig. 8. Parameters of the inverter control block using the PI controller.

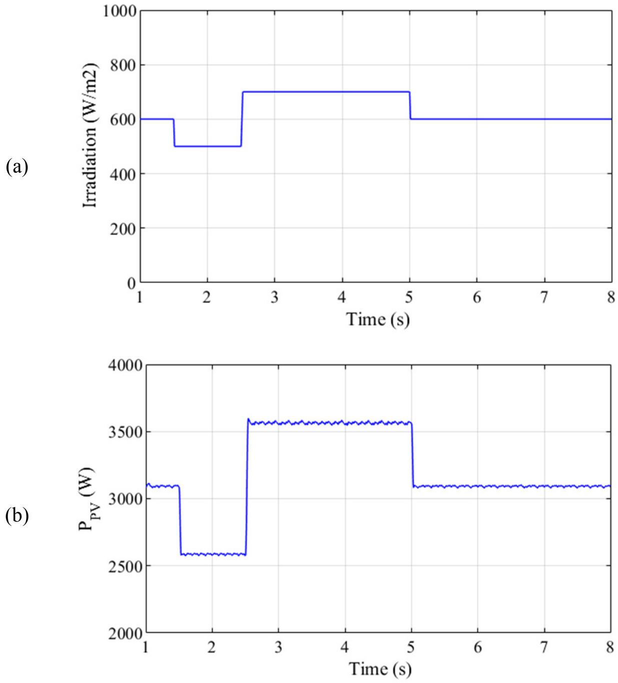


Fig. 9. Responses of the studied system with changing irradiance. (a) Irradiation, (b) active power of the solar PV system.

The fuzzy rules designed in Figures 10-12 show the comparative transient responses in the time-domain of the studied system with the PI controller (red lines) and with the designed ANFIS controller (blue lines) when a single-phase short-circuit fault occurs at the output of the inverter at 3.5s and at the power grid at 6.0s. The wind speed is 6m/s and the irradiation varies from 700W/m^2 to 600W/m^2.

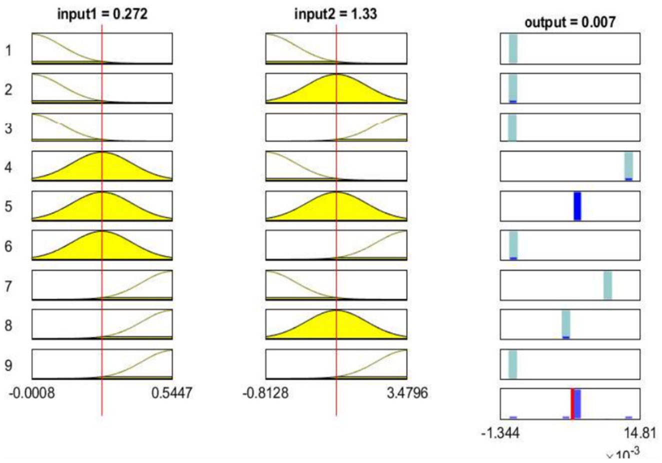


Fig. 10. Relationship between inputs and output of the ANFIS controller in the fuzzy rules.

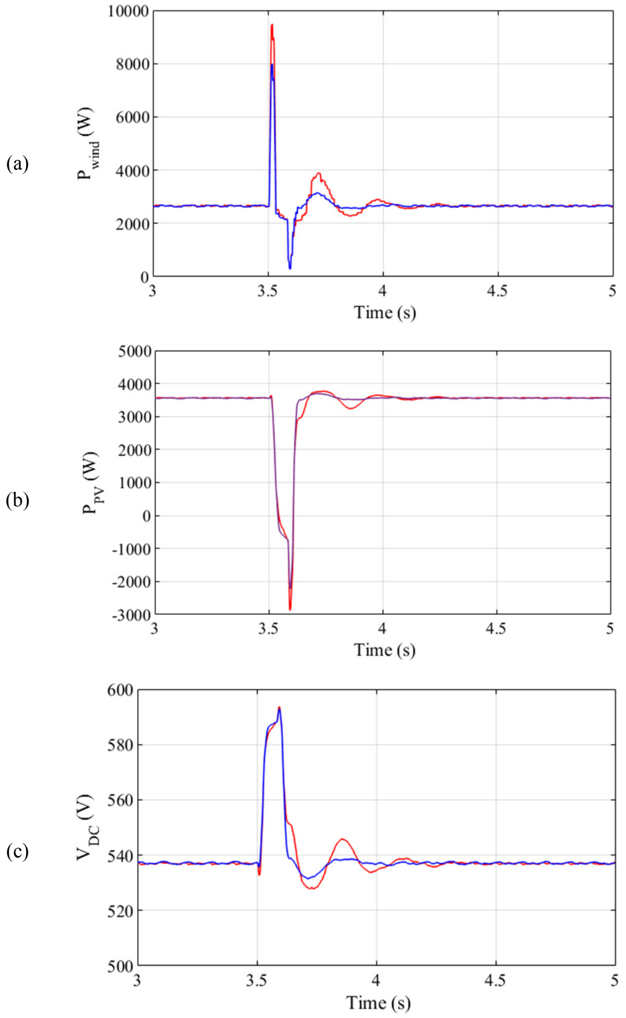


Fig. 11. Response of the studied system when a single-phase short-circuit occurs. (a) Active power of the wind system, (b) active power of the solar system, (c) voltage at the DC-link of the inverter.

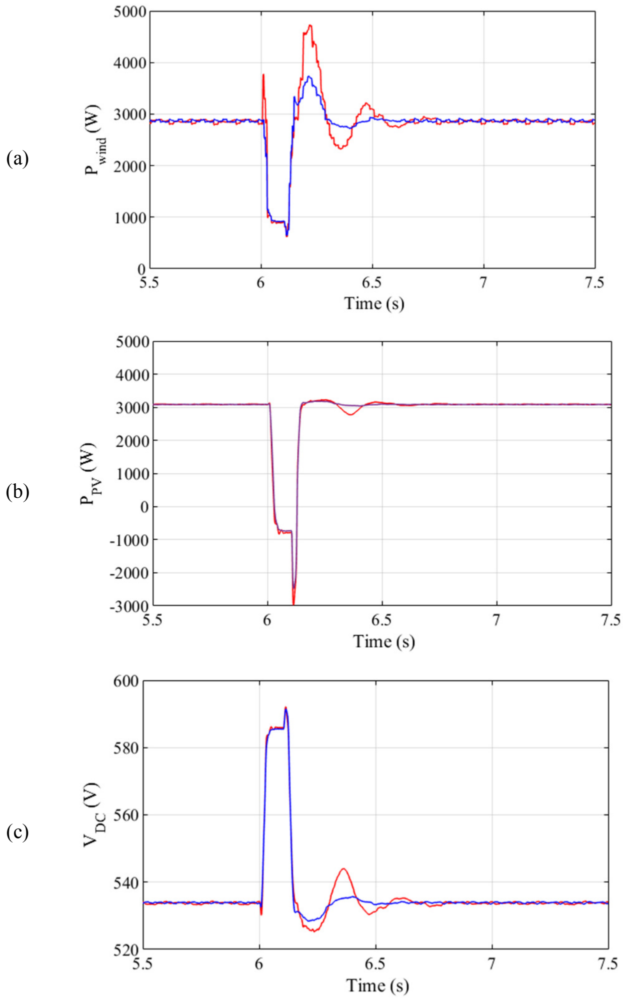


Fig. 12. Response of the studied system when a short-circuit fault occurs at the power grid. (a) Active power of the wind system, (b) active power of the solar system, (c) voltage at the DC-link of the inverter.

It is shown from the comparative results that when the faults occurred, the responses of all parameters oscillated. However, transient fluctuations were better controlled with the proposed ANFIS controller than with the PI controller with regard to overshoot and settling time.

V. CONCLUSIONS

In this paper, the comparative stability improvement of a hybrid grid-connected solar and wind system has been presented. To extract the maximum power from the solar PV system, a P&O controller is proposed for MPP tracking. Furthermore, the ANFIS controller was designed to keep stable the DC voltage of the hybrid inverter. Comparative time-domain simulation results of the studied system under severe faults were performed. It can be concluded from the results that the proposed P&O and ANFIS controller enhances the transient performance of the system with regard to overshoot and settling time reduction of the hybrid grid-tie inverter.

ACKNOWLEDGMENT

This paper was supported by the Ministry-level science and technology program with the title "Application of IoT technology in monitoring, control, and management of integrated electrical systems", grant number CT.2019.04.

REFERENCES

[1] A. Naeem, N. Ul Hassan, and N. Arshad, "Design of Solar-Wind Hybrid Power System by using Solar-Wind Complementarity," in *2020 4th International Conference on Green Energy and Applications (ICGEA)*, Singapore, Mar. 2020, pp. 100–105, https://doi.org/10.1109/ICGEA49367.2020.239713.

[2] T. Pehlivanova and A. Atanasov, "Selection of Optimal Configuration for Autonomous Solar-Wind Hybrid System for Household Consumption," in *2020 7th International Conference on Energy Efficiency and Agricultural Engineering (EE AE)*, Ruse, Bulgaria, Nov. 2020, pp. 1–4, https://doi.org/10.1109/EEAE49144.2020.9279017.

[3] N. Diaw, L. Thiaw, O. Ba, T. T. Soe, A. Swathi, and G. Giridhar, "Techno-Economic Feasibility of Wind-Solar Hybrid Systems For Rural Electrification of Sioure Village in Sahel," in *2019 IEEE 2nd International Conference on Renewable Energy and Power Engineering (REPE)*, Toronto, Canada, Nov. 2019, pp. 196–201, https://doi.org/10.1109/REPE48501.2019.9025141.

[4] W. Obaid, A.-K. Hamid, and C. Ghenai, "Hybrid MPPT Controlled Solar/Wind Power System for Pumping System," in *2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, Ras Al Khaimah, United Arab Emirates, Nov. 2019, https://doi.org/10.1109/ICECTA48151.2019.8959772.

[5] M. Malik and P. R. Sharma, "Power Quality Improvement in Hybrid Photovoltaic and Wind Power System Using 3 Levels Inverter," in *2020 International Conference on Advances in Computing, Communication Materials (ICACCM)*, Dehradun, India, Aug. 2020, pp. 253–260, https://doi.org/10.1109/ICACCM50413.2020.9212821.

[6] M. Farhoodnea, A. Mohamed, H. Shareef, and H. Zayandehroodi, "Power quality impact of grid-connected photovoltaic generation system in distribution networks," in *2012 IEEE Student Conference on Research and Development (SCOReD)*, Pulau Pinang, Malaysia, Dec. 2012, https://doi.org/10.1109/SCOReD.2012.6518600.

[7] S. Sebastian, E. A. S. Varhhese, and E. J. Varhhese, "Mitigation And Improvement Of Power Quality Using Shunt Series Switched Grid Tied Inverter (SSS-GTI)," in *2021 7th International Conference on Electrical Energy Systems (ICEES)*, Chennai, India, Feb. 2021, pp. 5–8, https://doi.org/10.1109/ICEES51510.2021.9383726.

[8] K. Suresh and S. Balagopalan, "A three-stage buck in buck, boost in boost grid-tied inverter," in *2018 International Conference on Power, Signals, Control and Computation (EPSCICON)*, Thrissur, India, Jan. 2018, https://doi.org/10.1109/EPSCICON.2018.8379590.

[9] V. Chaudhary, A. Bhargava, and S. Bhasin, "Modeling and Simulation of gird connected hybrid power system integrated with solar PV/Wind and controlled by Voltage Regulator," in *2019 International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, Jul. 2019, pp. 1316–1320, https://doi.org/10.1109/ICCES45898.2019.9002036.

[10] L. A. Alwal, P. K. Kihato, and S. I. Kamau, "A Review of Control Strategies for Microgrid with PV-Wind Hybrid Generation Systems," in *Proceedings of the Sustainable Research and Innovation Conference,* Kenya, May 2018, pp. 243–250.

[11] F. F. Mammadov, "Fuzzy Logic Controller Application in HybriD Solar and Wind Energy System," in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkey, Sep. 2019, https://doi.org/10.1109/IDAP.2019.8875960.

[12] S. Younsi and N. Hamrouni, "Control of Grid Connected Three-Phase Inverter for Hybrid Renewable Systems using Sliding Mode Controller," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 9, no. 11, 47/30 2018, https://doi.org/10.14569/IJACSA.2018.091146.

[13] J. Ma, K. L. Man, T. O. Ting, N. Zhang, S.-U. Guan, and P. W. H. Wong, "Approximate Single-Diode Photovoltaic Model for Efficient I-V Characteristics Estimation," *The Scientific World Journal*, vol. 2013, Nov. 2013, Art. no. e230471, https://doi.org/10.1155/2013/230471.

[14] G. Mustafa, M. H. Baloch, S. H. Qazi, S. Tahir, N. Khan, and B. A. Mirjat, "Experimental Investigation and Control of a Hybrid (PV-Wind) Energy Power System," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6781–6786, Feb. 2021, https://doi.org/10.48084/etasr.3964.

[15] H. Camur, Y. Kassem, and E. Alessi, "A Techno-Economic Comparative Study of a Grid-Connected Residential Rooftop PV Panel: The Case Study of Nahr El-Bared, Lebanon," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6956–6964, Apr. 2021, https://doi.org/10.48084/etasr.4078.

[16] D. N. Truong and V. T. Ngo, "Study of the Dynamic Stability Enhancement of PMG-Based Wind Turbine Generator Fed to Power Grid Using an SVC," *Journal of Science & Technology*, vol. 101, pp. 82–86, 2014.

[17] R. I. Putri, S. Adhisuwignjo, and M. Rifa'i, "Design of Simple Power Converter For Small Scale Wind Turbine System For Battery Charger," in *2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE)*, Yogyakarta, Indonesia, Nov. 2018, pp. 169–173, https://doi.org/10.1109/ICITISEE.2018.8721018.

[18] B. W. Bekit, L. D. Seneviratne, and J. F. Whidborne, "Robust adaptive control for robot manipulators," in *Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Tokyo, Japan, Jun. 1997, https://doi.org/10.1109/AIM.1997.652970.

[19] C. Wang, M. Chen, X. Zhang, and M. Gao, "An analog MPPT controller without multiplier for PV applications based on simplified P amp;O method," in *2017 IEEE Energy Conversion Congress and Exposition (ECCE)*, Cincinnati, OH, USA, Oct. 2017, pp. 2296–2300, https://doi.org/10.1109/ECCE.2017.8096446.

[20] M. Y. Allani, D. Mezghani, F. Tadeo, and A. Mami, "FPGA Implementation of a Robust MPPT of a Photovoltaic System Using a Fuzzy Logic Controller Based on Incremental and Conductance Algorithm," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4322–4328, Aug. 2019, https://doi.org/10.48084/etasr.2771.

[21] I. Vairavasundaram, V. Varadarajan, P. J. Pavankumar, R. K. Kanagavel, L. Ravi, and S. Vairavasundaram, "A Review on Small Power Rating PV Inverter Topologies and Smart PV Inverters," *Electronics*, vol. 10, no. 11, Jan. 2021, Art. no. 1296, https://doi.org/10.3390/electronics10111296.

[22] V.-T. Bui and D.-N. Truong, "Improvement of ANFIS Controller for SVC Using PSO," in *Computational Intelligence Methods for Green Technology and Sustainable Development*, 2021, pp. 293–302, https://doi.org/10.1007/978-3-030-62324-1_25.

[23] L. Wang and D.-N. Truong, "Stability Enhancement of a Power System With a PMSG-Based and a DFIG-Based Offshore Wind Farm Using a SVC With an Adaptive-Network-Based Fuzzy Inference System," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 7, pp. 2799–2807, Jul. 2013, https://doi.org/10.1109/TIE.2012.2218557.

[24] V. H. Nguyen, H. Nguyen, M. T. Cao, and K. H. Le, "Performance Comparison between PSO and GA in Improving Dynamic Voltage Stability in ANFIS Controllers for STATCOM," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4863–4869, Dec. 2019, https://doi.org/10.48084/etasr.3032.

A Hybrid Deep Learning and Optimized Machine Learning Approach for Rose Leaf Disease Classification

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

Abstract-Analysis of the symptoms of rose leaves can identify up to 15 different diseases. This research aims to develop Convolutional Neural Network models for classifying the diseases on rose leaves using hybrid deep learning techniques with Support Vector Machine (SVM). The developed models were based on the VGG16 architecture and early or late fusion techniques were applied to concatenate the output from a fully connected layer. The results showed that the developed models based on early fusion performed better than the developed models on either late fusion or VGG16 alone. In addition, it was found that the models using the SVM classifier had better efficiency in classifying the diseases appearing on rose leaves than the models using the softmax function classifier. In particular, a hybrid deep learning model based on early fusion and SVM, which applied the categorical hinge loss function, yielded a validation accuracy of 88.33% and a validation loss of 0.0679, which were higher than the ones of the other models. Moreover, this model was evaluated by 10-fold cross-validation with 90.26% accuracy, 90.59% precision, 92.44% recall, and 91.50% F1-score for disease classification on rose leaves.

Keywords-hybrid deep learning; neural networks; rose leaf diseases; support vector machine

I. INTRODUCTION

Roses are widely produced and exported globally. In 2019, the export value of roses was more than 175 million US dollars. The top five countries with the highest export rankings are Netherlands, Denmark, Uganda, Germany, and Canada [1]. In cultivating roses, pest problems such as insect infestations are often encountered along with pathogens caused by fungi, viruses, and bacteria [2]. There are also diseases caused by nutritional deficiencies such as in nitrogen, iron, zinc, and magnesium. Disease symptoms can be detected in roots, stems, branches, leaves, and buds or flowers. Especially the leaves are a source of various infectious disease symptoms. However, classifying an infected disease requires skill and experience. For example, rose mosaic disease is a common disease worldwide and can sometimes be caused by more than one pathogen [3]. Image processing methods for plant disease classification are currently being studied [4] combined with machine learning such as Support Vector Machine (SVM) [5] or K-Nearest Neighbors (KNN) [6]. For example, the authors in [7] classified 4 rose leaf diseases using machine learning with with at least 94% accuracy. In addition to machine learning, other methods such as deep learning and neural networks are applied to recognize and classify plant diseases. The author in [8] developed a Convolutional Neural Network (CNN) model, which applied MobileNet and transfer learning for rose disease classification. Over 30 and at least 15 rose diseases can be observed on the leaves [2].

Most of the research used a single perspective or a single set of images as the dataset for model training. However, in the deep learning model training, it is necessary to use images with multiple perspectives, such as image augmentation, and include image segmentation to highlight features that appear on the image. This approach usually increases the accuracy of the model. Moreover, there are currently very few studies that have applied hybrid deep learning models to classify plant diseases. Therefore, this research aims to develop rose leaf disease classification models using hybrid deep learning. Besides, this work also compares the performance of models using different classifiers, namely softmax function and SVM.

II. RESEARCH METHODOLOGY

A. Image Data Collection

This study classifies rose diseases by identifying the symptoms on leaves based on image processing and CNN. A program was developed using the Google search engine and ChromeDriver utility to search and download rose leaf images with dimensions of at least 224 pixels. All downloaded images were rechecked and labeled. Moreover, the author took photos of rose leaves with and without diseases with an Android mobile phone. Therefore, a dataset of 4,032 downloaded and taken pictures was formed. The imagres were categorized to 16 different classes with regard to the shown rose disease (15 diseases + 1 normal/control) as shown in Table I. Finally, all images were resized and cropped to the dimension of 224×224 pixels.

Corresponding author: Sumitra Nuanmeesri

TABLE I. ROSE LEAF IMAGES CLASSIFICATION

No.	Label	Disease	Caused by	Amount (images)
1	BS	Black spot	Fungi	295
2	DM	Downy mildew	Fungi	272
3	PM	Powdery mildew	Fungi	281
4	BM	Black mold	Fungi	245
5	BB	Botrytis blight	Fungi	234
6	VW	Verticillium wilt	Fungi	220
7	CLS	Cercospora leaf spot	Fungi	279
8	ATN	Anthracnose	Fungi	276
9	RT	Rust	Fungi	280
10	RM	Rose mosaic	Virus	270
11	RR	Rose Rosette	Virus	236
12	RSD	Rose spring dwarf	Virus	281
13	SLR	Strawberry latent ringspot	Virus	183
14	IB	Insect bites	Insect	192
15	ND	Nutrient deficiencies	-	238
16	NM	Normal (disease-free)	-	250

B. *Image Augmentation*

During the deep learning training mstage many images are needed to increase the performance of the model. If a particular class has a small number of images, it can affect classification accuracy. Thus, the author increased the number of training images by using the image augmentation technique, including vertical and horizontal flips, rotation (45, -45, 90, -90 degrees), shearing (45 and -45 degrees), and random zoom-in up to 200%. Thus, the number of the images increased from 4,032 to 40,320 images. The example of image augmentation is shown in Figure 1.

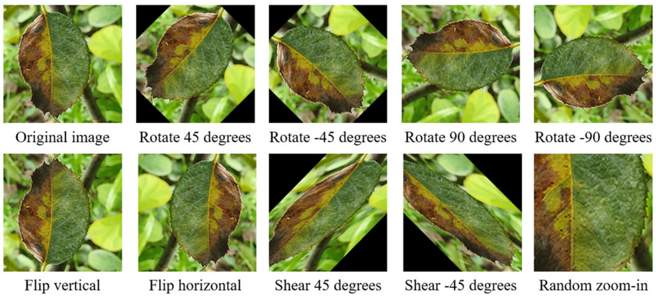
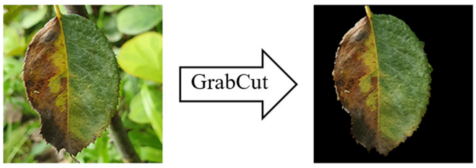

Original image, Rotate 45 degrees, Rotate -45 degrees, Rotate 90 degrees, Rotate -90 degrees
Flip vertical, Flip horizontal, Shear 45 degrees, Shear -45 degrees, Random zoom-in

Fig. 1. Example of image augmentation.

C. *Image Processing*

Image processing was applied to emphasize the physical appearance of the rose leaves. The preliminary step is that the rose leaf was separated from the background pixels with the GrabCut [9] method based on the graph-cut technique. Such an output image is illustrated in Figure 2.



Original image Image without background pixels

Fig. 2. The background pixels are removed by the GrabCut method.

The images obtained after the removal of the background pixels were subjected to image color-spacing and image thresholding processing.

1) *Hue, Saturation, Value Color Space*

Hue, Saturation, Value (HSV) is a color space model which includes ranges of color type between 0 to 360 degrees, vibrancy, and color brightness. This work focused on color ranges around 120 degrees, which are related to the green color of a rose leaf. All green areas were desaturated with the lower saturation as grayness. As a result, any color not related to green was accelerated to become more emphatic.

2) *Truncated Adaptive Gaussian Thresholding*

Truncated Adaptive Gaussian Thresholding (TAGT) is a combination technique between truncate and adaptive Gaussian thresholding. For truncate thresholding, an image without background pixels is processed. Pixels greater than the threshold value $th_{val} = 127$, were assigned that value [10]:

$$out(x,y) = \begin{cases} \text{threshold}, & \text{if } in(x,y) > th_{val} \\ in(x,y), & \text{otherwise} \end{cases} \quad (1)$$

where $in(x,y)$ refers to the input pixel coordination, and $out(x,y)$ refers to the output pixel coordination.

Next, the threshold was adjusted from the weighted sum of the block size of the pixel neighborhood at 7×7 using adaptive Gaussian and binary thresholding, as in (2) [10].

$$out(x,y) = \begin{cases} max_{val}, & \text{if } in(x,y) > th(x,y) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where max_{val} refers to the maximum value assigned to the pixels and $th(x,y)$ refers to the individual threshold calculation of each pixel.

3) *Double Inverse-Binary Thresholding*

Double Inverse-Binary Thresholding (DIBT) is a thresholding method with twice applied inverse-binary thresholding. First, the image without background pixels is taken through a thresholding process between inverse-binary and binary, where the threshold values are set to 100 and 0 respectively. The resulting image from the first step was adjusted to the threshold value of 127 by inverse-binary thresholding, calculated in (3) [10].

$$out(x,y) = \begin{cases} 0, & \text{if } in(x,y) > th(x,y) \\ max_{val}, & \text{otherwise} \end{cases} \quad (3)$$

This results in an image emphasizing the expected coordinates of the suspected position of disease infection or wilt on the rose leaf. Each original image will result to 3 more images, namely HSV, TAGT, and DIBT, as shown in Figure 3.

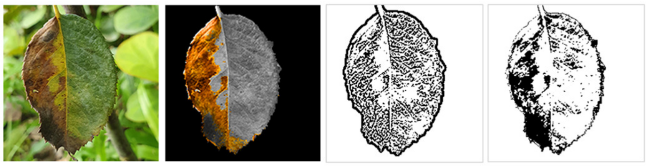

Original HSV TAGT DIBT

Fig. 3. Example of each image processing.

D. Hybrid Deep Learning Modeling

In this step the CNN models for the classification of diseases on a rose leaf were developed. Twelve models were developed as follows.

1) Visual Geometry Group16-Based CNN Model

Visual Geometry Group16 (VGG16) model is a CNN architecture presented in [11]. The input images of the VGG16-based CNN model were set to 224×224 pixels for processing through 16 weight layers, including 13 convolution layers and 3 fully connected layers. All convolution layers have a 3×3 kernel size, 1 pixel of padding size, and the Rectified Linear Unit (ReLU) activation function. Spatial pooling followed with 5 max-pooling layers with a 2×2 pixel filter and stride 2. Further, 1 flatten layer was included before feeding the output to the fully connected layers. Furthermore, the softmax activation function was applied with 1,000 output classes in the last fully connected layer. Thus, the total trainable parameters of this model were 138,357,544 as shown in Figure 4.

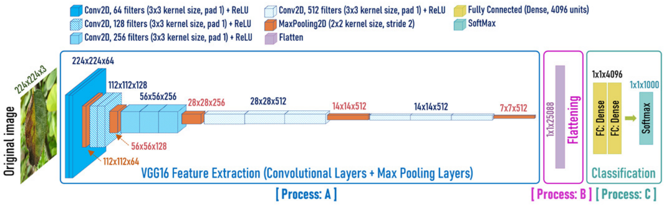


Fig. 4. VGG16-based CNN model architecture.

According to Figure 4, the "Process: A" refers to feature extraction layers, "Process: B" refers to the flatten layer, and "Process: C" refers to the fully connected layers. This research used the original image dataset with 16 labeled output classes (see Table I) to develop the VGG16-based CNN model. Therefore, the softmax layer (in Process: C) was set to 16 instead of 1000 classes. The overall trainable parameters of this model were 134,326,096.

2) Early Fusion Model

The early fusion (EF) model was developed based on the VGG16-based CNN model. It starts with the images that have undergone each image thresholding processing separately (original, HSV, TAGT, and DIBT images) to each channel of CNN for extracting features (see Process: A in Figure 4) with the VGG16 architecture. The outputs obtained for each dataset were fused and flattened before being classified with the fully connected layers, as shown in Figure 5. The sum of the trainable parameters was 484,041,973.

3) Late Fusion Model

The VGG16-based CNN model was extended to the Late Fusion (LF) model in this work. The LF model starts with each processed image as input (as the EF model) for each CNN channel, then fused each output obtained after classification by the softmax activation function. After the fusion of the results obtained from the 4 image types, they were classified by 2 dense layers of size 4,096 and were finalized with the softmax function to 16 output classes (Figure 6). The total trainable parameters were 586,665,136.

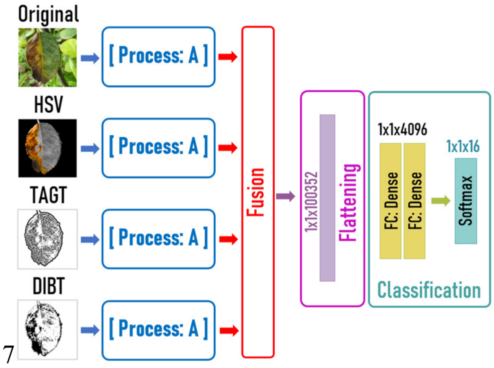


Fig. 5. VGG16-based early fusion model.

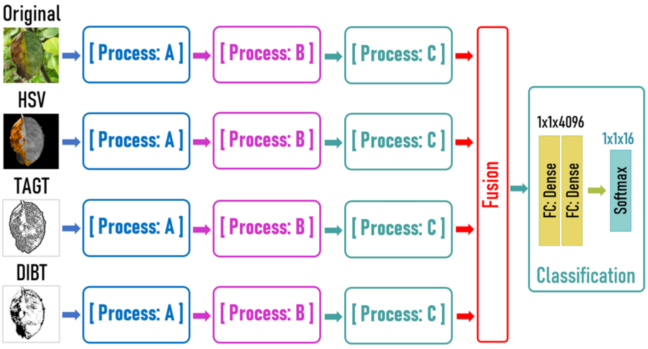


Fig. 6. VGG16-based late fusion model.

4) VGG16-Based SVM Models

According to the VGG16-based CNN model, the softmax activation function was used to classify the final output at the last layer. In contrast, this VGG16-based SVM model applied the SVM classifier instead of the softmax activation function, as shown in Figure 7.

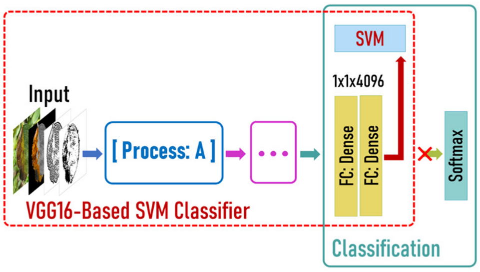


Fig. 7. VGG16-based SVM model.

SVM is a popular classifier for supervised learning algorithms, especially for binary classification. In this work, multi-class SVM is required to classify 16 classes of rose leaf disease images. There are several classifiers for multi-class SVM. In this work various multi-class SVM classifiers were applied including L2-SVM, Categorical Hinge Loss SVM (CHL-SVM), and Weston-Watkins SVM (WW-SVM), to the VGG16-based CNN models as follows.

The L2-SVM is based on the optimization of L2 norm and Squared Hinge Loss (SHL) which is calculated in (4) and (5).

$$y = w^T x_i + b \quad (4)$$

$$L = \sum_{i=1}^{N} \max(0, 1 - \hat{y}_i y)^2 \quad (5)$$

where w^T refers to the weight of dataset T, x_i refers to the augmentation of sample data vectors, b refers to the bias, N refers to the number of samples in a dataset, $\hat{y}$ refers to the actual class, and y refers to the predicted class.

Then, the SHL was optimized with a minimum of Euclidean norm and a large error penalty. Thus, the L2-SVM was formulated in (6).

$$\min \frac{1}{2}\|w\|_2^2 + C\sum_{i=1}^{N} \max(0, 1 - \hat{y}_i(w^T x_i + b))^2 \quad (6)$$

where $\|w\|_2$ refers to the Euclidean norm (L2 norm regularization), and C refers to the large error penalty for misclassification in which $C > 0$.

The CHL-SVM or multi-class hinge loss function was implemented with TensorFlow 2 based Keras and is calculated in (7) [12].

$$L = \max\big(0, 1 + \max\big((1 - \hat{y})y\big) - \sum_{i=1}^{N} \hat{y}_i y_i\big) \quad (7)$$

Regarding the WW-SVM or Weston-Watkins hinge loss [13], the linear classifier was calculated in (8), and the optimization was formulated in (9).

$$L = \sum_{i=1}^{l} \sum_{m \neq y_i} \max(0, 1 + y_i - y_m) \quad (8)$$

$$\min \frac{1}{2}\sum_{m=1}^{k}\|w_m\|^2 + C\sum_{i=1}^{l} \sum_{m \neq y_i} \max(0, 1 + y_i - y_m) \quad (9)$$

where w refers to the weight, k refers to the number of classes, l refers to the number of samples in the dataset, y refers to the predicted class, and C refers to the large error penalty for misclassification in which $C > 0$.

Finally, the last fully connected layer with the softmax function was replaced with the multi-class SVM to classify the final output. Therefore, 3 VGG16-based SVM models, namely VGG16 & L2-SVM, VGG16 & CHL-SVM, and VGG16 & WW-SVM, were developed with different multi-class SVM classifiers.

5) Early Fusion-Based SVM Models

The early fusion-based SVM model applied the 3 different multi-class SVM classifiers to the last fully connected layer with softmax of the EF model. Thus, 3 EF models were developed, namely EF & L2-SVM, EF & CHL-SVM, and EF & WW-SVM.

6) Late Fusion-Based SVM Models

The LF model included two softmax functions at two layers: before fusion and at the last layer of the model. Thus, the softmax classifier of the last layer was bypassed and replaced with the multi-class SVM classifiers. Therefore, 3 late fusion-based SVM models were developed: LF & L2-SVM, LF & CHL-SVM, and LF & WW-SVM.

By default, all models are based on VGG16 and CNN architecture. The image dataset was randomly split into 70%, 15%, and 15% for training, validating, and testing respectively. The hyperparameters were set as follows: the batch size was 64, the learning rate was 0.001, and training took 200 epochs. The models were compiled with the Adam optimizer.

E. Model Evaluation

All models were evaluated and validated by the accuracy and loss value during the training processing. In addition, the VGG16-based CNN, EF, and LF models that applied the softmax function at the fully connected layer were evaluated using the categorical cross-entropy loss [14]:

$$\text{Cross-entropy} = -\sum_{i=1}^{m}\sum_{j=1}^{n} y_{i,j}\log\big(p_{i,j}\big) \quad (10)$$

where m refers to the total of input, n refers to the number of classes, $y_{i,j}$ refers to the input i of class j, $p_{i,j}$ refers to the probability of the predicted class j by input i.

Further, the k-fold cross-validation was used to estimate the learning skill of the model based on an unseen dataset. The k-fold cross-validation is mainly used to measure performance for machine learning models but can also be applied to deep learning models. Thus, all 12 models were evaluated by 10-fold cross-validation in this work. The performance of the models was validated on accuracy (ACC) [15], precision (PREC) [16], recall (REC) [17], and F1-Score.

III. RESULTS

All 12 models were trained, validated, and tested and their performances were compared.

A. Model Training and Validation Performance

The results showed that the models developed with the early fusion technique performed better than late fusion and VGG16 models. Especially, the model developed with the early fusion method and categorical hinge loss for the SVM (EF & CHL-SVM) gave the best accuracy among the models as shown in Table II.

TABLE II. ACCURACY AND LOSS VALUES RESULTS BETWEEN TRAINING AND VALIDATION OF THE MODELS

Models	**Training**		**Validation**	
	ACC	**Loss**	**ACC**	**Loss**
VGG16 (Softmax)	94.71	0.1847	83.28	0.4063
VGG16 & L2-SVM	95.09	0.1656	83.79	0.3643
VGG16 & CHL-SVM	96.03	0.1185	84.32	0.2607
VGG16 & WW-SVM	95.59	0.1404	84.25	0.3090
EF (Softmax)	97.09	0.0654	87.31	0.1438
EF & L2-SVM	97.37	0.0514	87.62	0.1130
EF & CHL-SVM	98.28	0.0309	88.33	0.0679
EF & WW-SVM	97.91	0.0443	88.07	0.0974
LF (Softmax)	96.44	0.0981	84.19	0.2159
LF & L2-SVM	96.86	0.0770	84.87	0.1695
LF & CHL-SVM	97.21	0.0597	85.44	0.1313
LF & WW-SVM	96.97	0.0714	85.02	0.1570

According to Table II, the EF & CHL-SVM model yielded a training accuracy of 98.28% and a training loss of 0.0309. In addition, it gave a validation accuracy of 88.33% and a validation loss of 0.0679. The accuracy and loss error between training and validation of the EF & CHL-SVM are shown in Figure 8. The validation accuracy of the EF & CHL-SVM model for rose leaf disease classification can be normalized and displayed as a confusion matrix, as shown in Figure 9. The EF

& CHI-SVM model was able to classify the disease-free (NM) rose leaves with 98.95% accuracy. The most accurate classifications of rose leaf diseases were VW, IB, and DM, with 98.48%, 94.08%, and 92.68% accuracy respectively. For most of the other diseases the accuracy was higher than 87%, except for BB, ATN, CLS, and SLR which had less than 83%. Especially, the SLR disease had the lowest accuracy of 74.44%.

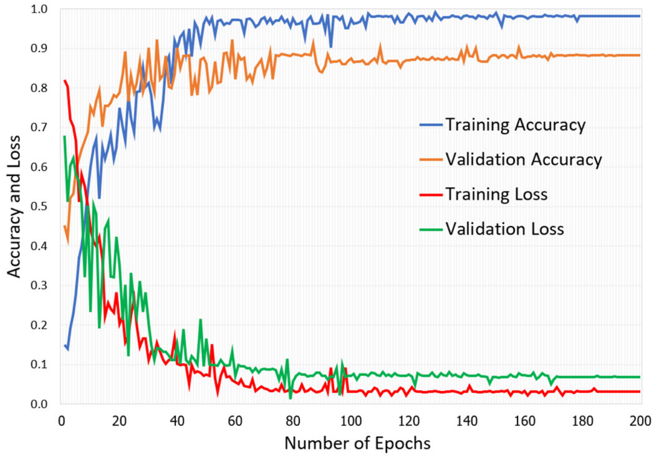
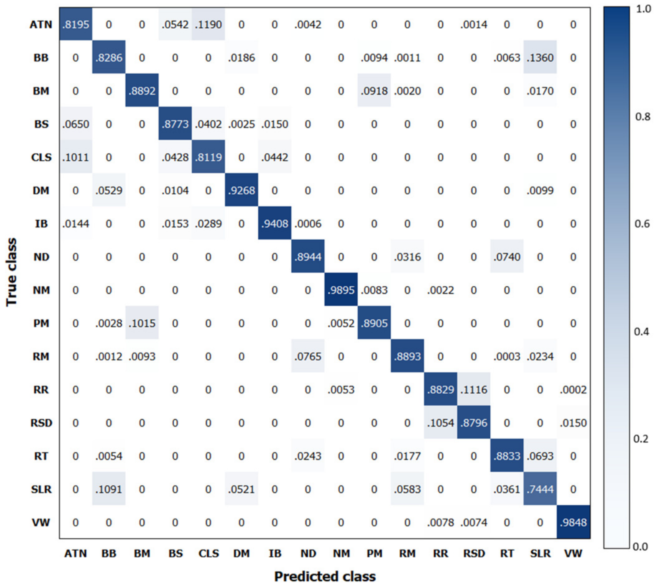


Fig. 8. The performance of training and validation of the EF & CHL-SVM model.



	ATN	BB	BM	BS	CLS	DM	IB	ND	NM	PM	RM	RR	RSD	RT	SLR	VW
ATN	.8195	0	0	.0542	.1190	0	0	.0042	0	0	0	0	.0014	0	0	0
BB	0	.8286	0	0	0	.0186	0	0	0	.0094	.0011	0	0	.0063	.1360	0
BM	0	0	.8892	0	0	0	0	0	0	.0918	.0020	0	0	0	.0170	0
BS	.0650	0	0	.8773	.0402	.0025	.0150	0	0	0	0	0	0	0	0	0
CLS	.1011	0	0	.0428	.8119	0	.0442	0	0	0	0	0	0	0	0	0
DM	0	.0529	0	.0104	0	.9268	0	0	0	0	0	0	0	0	.0099	0
IB	.0144	0	0	.0153	.0289	0	.9408	.0006	0	0	0	0	0	0	0	0
ND	0	0	0	0	0	0	0	.8944	0	0	.0316	0	0	.0740	0	0
NM	0	0	0	0	0	0	0	0	.9895	.0083	0	.0022	0	0	0	0
PM	0	.0028	.1015	0	0	0	0	0	.0052	.8905	0	0	0	0	0	0
RM	0	.0012	.0093	0	0	0	0	.0765	0	0	.8893	0	0	.0003	.0234	0
RR	0	0	0	0	0	0	0	0	.0053	0	0	.8829	.1116	0	0	.0002
RSD	0	0	0	0	0	0	0	0	0	0	0	.1054	.8796	0	0	.0150
RT	0	.0054	0	0	0	0	0	.0243	0	0	.0177	0	0	.8833	.0693	0
SLR	0	.1091	0	0	0	.0521	0	0	0	0	.0583	0	0	.0361	.7444	0
VW	0	0	0	0	0	0	0	0	0	0	0	.0078	.0074	0	0	.9848

Fig. 9. The validation accuracy of the EF & CHL-SVM model in a confusion matrix.

B. Model Evaluation

All developed models were tested and evaluated by 10-fold cross-validation with a test dataset. The result showed that the performances of the EF-based models were higher than the LF-based and VGG16-based models'. Regarding the EF-based models, the EF & CHL-SVM had the highest performance with 90.26% accuracy, 90.59% precision, 92.44% recall, and 91.50% F1-score as shown in Table III.

TABLE III. CROSS-VALIDATION PERFORMANCE

Models	ACC	PREC	REC	F1-Score
VGG16 (Softmax)	86.16	87.03	89.05	88.03
VGG16 & L2-SVM	86.37	87.27	89.16	88.20
VGG16 & CHL-SVM	86.90	87.81	89.47	88.63
VGG16 & WW-SVM	86.62	87.53	89.30	88.41
EF (Softmax)	89.54	90.20	91.90	91.04
EF & L2-SVM	89.85	90.51	92.14	91.32
EF & CHL-SVM	90.26	90.59	92.44	91.50
EF & WW-SVM	90.08	90.63	92.42	91.51
LF (Softmax)	87.62	88.56	90.04	89.30
LF & L2-SVM	87.83	88.78	90.19	89.48
LF & CHL-SVM	88.31	89.13	90.68	89.90
LF & WW-SVM	88.05	88.88	90.43	89.65

IV. CONCLUSION

This research developed 12 models for classifying rose diseases from the symptoms that appear on the rose leaves using a CNN model based on VGG16 architecture and image processing. The classification of rose diseases consists of 16 classes (9 classes for diseases caused by fungi, 4 for virus diseases, 1 for insect bit, 1 for nutrient deficiencies, and 1 disease-free class). The 12 developed CNN models were divided into three groups: VGG16, EF, and LF. In addition, each group was divided into two classifier types: softmax and SVM. The softmax function was used in 3 models, namely the VGG16-based CNN, EF, and LF models. The utilized multi-class SVM classifiers were L2-SVM, CHL-SVM, and WW-SVM. There were 4,032 rose leaf images for model training. The images were resized to 224×224 pixels and underwent image augmentation, resulting in a dataset of 40,320 images. These images were subjected to image processing, including removal of background pixels, HSV color space, TAGT, and DIBT, to emphasize their features. Both TAGT and DIBT are based on image thresholding processing. Ultimately, the dataset was split to 70% for training, 15% for validation, and 15% for model testing by 10-fold cross-validation.

The results showed that the EF-based models gave the highest training, validation, and testing performance values, followed by the LF-based and the VGG16-based models. In addition, the models developed with the SVM classifier performed higher than the models using the softmax function. The model using CHL-SVM showed the highest performance, followed by the models using WW-SVM, L2-SVM, and softmax function. Thus, the EF & CHL-SVM, a developed model based on the early fusion method and employing the SVM categorical hinge loss function was the most suitable model for classifying diseases on rose leaves with an accuracy of at least 88.33%. The models developed in [7, 8] had accuracy not less than 94%, which is higher than the accuracy of the CHL-SVM model in this work. However, these two studies only classified 4 rose leaf diseases, unlike this study which classified 15.

Moreover, it is evident that image processing can improve rose leaf disease classification accuracy, especially when the features are fused. Besides, it was found that SVM gave better results as a classifier than the softmax activation function, which is consistent with the findings in [18].

Regarding further work, the author plans to develop a model based on U-net deep learning and a transfer learning approach to detect and classify diseases on plants and then integrate it to the Internet of Things.

ACKNOWLEDGMENT

The author wishes to thank the Institute for Research and Development, Suan Sunandha Rajabhat University, for supporting this research.

REFERENCES

[1] "Top countries for Export of Roses," *NationMaster*. https://www.nationmaster.com/nmx/ranking/export-of-roses (accessed Sep. 21, 2021).

[2] R. K. Horst and R. A. Cloyd, *Compendium of Rose Diseases and Pests. Second Edition*, 2nd ed. St. Paul, MI, USA: The American Phytopathological Society, 2007.

[3] I. Vazquez-Iglesias *et al.*, "High throughput sequencing and RT-qPCR assay reveal the presence of rose cryptic virus-1 in the United Kingdom," *Journal of Plant Pathology*, vol. 101, no. 4, pp. 1171–1175, Nov. 2019, https://doi.org/10.1007/s42161-019-00307-5.

[4] S. Minaee, M. Jafari, and N. Safaie, "Design and development of a rose plant disease-detection and site-specific spraying system based on a combination of infrared and visible images," *Journal of Agricultural Science and Technology*, vol. 20, no. 1, pp. 23–36, Jan. 2018.

[5] D. Das, M. Singh, S. S. Mohanty, and S. Chakravarty, "Leaf Disease Detection using Support Vector Machine," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, Jul. 2020, pp. 1036–1040, https://doi.org/10.1109/ICCSP48568.2020.9182128.

[6] A. A. Bharate and M. S. Shirdhonkar, "A review on plant disease detection using image processing," in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, Palladam, India, Dec. 2017, pp. 103–109, https://doi.org/10.1109/ISS1.2017.8389326.

[7] K. Swetharani and V. Prasad, "Design and Implementation of an Efficient Rose Leaf Disease Detection using K-Nearest Neighbours," *International Journal of Recent Technology and Engineering*, vol. 9, no. 3, pp. 21–27, Sep. 2020, https://doi.org/10.35940/ijrte.C4213.099320.

[8] A. Rajbongshi, T. Sarker, Md. M. Ahamad, and Md. M. Rahman, "Rose Diseases Recognition using MobileNet," in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Istanbul, Turkey, Oct. 2020, https://doi.org/10.1109/ISMSIT50672.2020.9254420.

[9] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut': interactive foreground extraction using iterated graph cuts," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 309–314, Aug. 2004, https://doi.org/10.1145/1015706.1015720.

[10] "OpenCV: Miscellaneous Image Transformations," *OpenCV*. https://docs.opencv.org/master/d7/d1b/group__imgproc__misc.html (accessed Sep. 21, 2021).

[11] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, Apr. 2015, Accessed: Sep. 21, 2021. [Online]. Available: http://arxiv.org/abs/1409.1556.

[12] "tf.keras.losses.CategoricalHinge," *TensorFlow*. https://www.tensorflow.org/api_docs/python/tf/keras/losses/CategoricalHinge (accessed Sep. 21, 2021).

[13] J. Weston and C. Watkins, "Support Vector Machines for Multi-Class Pattern Recognition," presented at the 7th European Symposium on Artificial Neural Networks, Bruges, Belgium, Jan. 1999.

[14] L. Poomhiran, P. Meesad, and S. Nuanmeesri, "Improving the Recognition Performance of Lip Reading Using the Concatenated Three Sequence Keyframe Image Technique," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6986–6992, Apr. 2021, https://doi.org/10.48084/etasr.4102.

[15] S. Nuanmeesri, "Development of community tourism enhancement in emerging cities using gamification and adaptive tourism recommendation," *Journal of King Saud University - Computer and Information Sciences*, Apr. 2021, https://doi.org/10.1016/j.jksuci.2021.04.007.

[16] S. Nuanmeesri and W. Sriurai, "Thai Water Buffalo Disease Analysis with the Application of Feature Selection Technique and Multi-Layer Perceptron Neural Network," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6907–6911, Apr. 2021, https://doi.org/10.48084/etasr.4049.

[17] A. N. Saeed, "A Machine Learning based Approach for Segmenting Retinal Nerve Images using Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5986–5991, Aug. 2020, https://doi.org/10.48084/etasr.3666.

[18] H. Hasan, H. Z. M. Shafri, and M. Habshi, "A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models For Hyperspectral Image Classification," *IOP Conference Series: Earth and Environmental Science*, vol. 357, Nov. 2019, Art. no. 012035, https://doi.org/10.1088/1755-1315/357/1/012035.

AUTHOR'S PROFILE

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

Innovative Design and Analysis of an Electrically Small Reconfigurable Antenna for GPS and Blue Tooth Applications

Harshitha Reddy Katireddy
Department of ECE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur, Andhra Pradesh, India
govardhanee_ec@kluniversity.in

M. Venkata Narayana
Department of ECE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur, Andhra Pradesh, India
mvn@kluniversity.in

Govardhani Immadi
Department of ECE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur, Andhra Pradesh, India
govardhanee_ec@kluniversity.in

Abstract-In this paper, an electrically small, planar antenna with broad side radiation pattern is presented. The design contains a dipole and a segmented circular loop which works equivalent to that of a magnetic dipole. A circular patch with slots is used to provide impedance matching. In general, electrically small antennas suffer from narrow bandwidth. In this paper, the reconfigurability of the small antenna for two different applications, 1.5GHz and 2.4GHz, is discussed. This reconfigurability was achieved by using a BAR 64-03W pin diode to adjust the resonant frequency. Two reconfigurable frequency bands were achieved at 1.5GHz and 2.4GHz with broad side radiation patterns.

Keywords-electrically small antenna; PIN diode; end fire radiation pattern; NFRP elements

I. INTRODUCTION

The components used in wireless communication systems such as antennas, filters, multiplexers, modulators, etc. are continuously miniaturized and intelligentized [1-3]. Frequency reconfigurable electrically small antennas are often suitable and are used for compact, wearable devices [4, 5]. An antenna with a broad front-to-back ratio and an equivalent pattern of E and H plane radiation is shown in [6]. The research groups in [7-10] and [11-15] have extensively researched antennas with the combination of electric and magnetic dipoles. The quarter wavelength patch with vertical orientation is commonly used in [7] as a magnetic dipole, which is coupled with the electric dipole to design complementary antennas with a wavelength of $0.25\lambda_0$ [8-12]. NFRP elements were used in [13] to design electric and magnetic dipoles, contributing to the design of dual band electrically small antennas. Most of the tiny electrically small antennas are not reconfigurable except from [4] and [5]. Three magnetic NFRP elements were paired with two electric NFRP elements in [5] to achieve the reconfigurability of their radiation pattern. A planar, reconfigurable complementary antenna with a single layer pattern was indicated in [16]. In addition to the reconfigurability of polarization and radiation reconfigurability patterns, due to its powerful spectrum use, frequency reconfigurability has gained a lot of attention [17-20]. Utilizing appropriate antenna size, different frequency reconfigurable monopole antennas [17-19], patch antennas [20-22] and quasi-Yagi Uda antennas [23-24] have been created. In order to obtain a unidirectional radiation pattern which increases the antenna size, a broad ground plane is required. By eliminating the ground plane, the complementary Huygens source antenna is used to achieve high front-to-back ratios with less distance. This paper introduces an electrically small reconfigurable antenna. To produce the magnetic dipole that is combined with the rectangular bar electric dipole to obtain the end fire radiation pattern, a segmented circular loop is used. This antenna is 35mm×35mm×1.6mm in size and has a single coating. In addition, the design is expanded by inserting a pin diode in the electric dipole to achieve frequency reconfigurability. Two resonant bands were obtained at 1.5 and 2.4GHz and the end fire radiation pattern was maintained. This antenna was fabricated in order to obtain a practical realization of the results. A good agreement between the simulated and the experimental results was obtained.

II. ANTENNA MECHANISM

The magnetic dipole is considered to be a loop with a uniform current distribution. Due to its low radiation resistance and high reactance, its impedance matching is poor. The diameter of the outer loop is 16.8mm at the resonant frequency and the current distribution is uniform and flows along the Y-axis (Figure 1). The electrical and magnetic dipoles are excited in a perpendicular manner in order to design this planar, electrically small antenna. The strip line and the

---

Corresponding author: Govardhani. Immadi circular loop on the top and bottom layers of the substrate are mounted at this end. The length and width of the patch are calculated by (1)-(3).

$$W = \frac{V_o}{2F_r}\sqrt{\frac{2}{\varepsilon_r+1}} \quad (1)$$

where v_0 is the velocity of light in free space.

$$L = \frac{c}{2F_r\sqrt{\varepsilon_{reff}}} - 2\Delta l \quad (2)$$

$$\Delta l = 0.412h\frac{(\varepsilon_{reff}+0.03)(w+0.26h)}{(\varepsilon_{reff}-0.258)(w+0.8h)} \quad (3)$$

where Δl is the extension in length due to fringing effects.

The effective dielectric constant is given by:

$$\varepsilon_{reff} = \frac{\varepsilon_r+1}{2} + \frac{\varepsilon_r-1}{2}\left[1 + \left[\frac{12h}{w}\right]\right]^{-1/2} \quad (4)$$

The actual length and width of the ground plane can be calculated by (5) and (6):

$$L_g = 6h + L \quad (5)$$

$$W_g = 6h + W \quad (6)$$

where L is the length of the patch, W its width, and h is the thickness of the substrate.

Length and width of the feed line are calculated by:

$$B = \frac{60\Pi^2}{\sqrt{\varepsilon_r Z_C}} \quad (7)$$

where Z_c is the characteristic impedance and ε_r the relative permittivity of the substrate.

$$L_f = \frac{\lambda_0}{4\sqrt{\varepsilon_{eff}}} \quad (8)$$

The radius of the patch is calculated by:

$$a = \frac{F}{\sqrt{1+\frac{2h}{\Pi\varepsilon_r}[\ln\left(\frac{\Pi f}{2h}\right)+1.7726]}} \quad (9)$$

Small antennas are those in which the largest dimension is less than the one-tenth of the wave length ($\lambda/10$). Small antennas are those having $K.a<1$, where K=wave number=$\frac{2\pi}{\lambda}$=0.0314. For the proposed antenna the obtained $K.a$ value is: $K.a=0.0314\times16.8=0.527<1$.

III. ANTENNA DESIGN CONSIDERATION

Figure 1 represents the configuration of the electrically small planar antenna. It mainly consists of a segmented loop and strip line fabricated on a FR4 substrate with 4.4 relative permittivity, 1.6mm thickness, 35mm length, and 35mm width. The top layer of the antenna is fabricated with a strip line and it is fed by an edge feed acting as an electric dipole. The bottom layer of the antenna, i.e. in the ground plane, contains an outer circle with radius R_4 and an inner circle with radius R_3 (see Table I). The proposed antenna was designed and analyzed experimentally. The prototype of the antenna was fabricated and the result was verified by a vector network analyzer (Figure 2). Figure 3 shows the fabricated antenna.

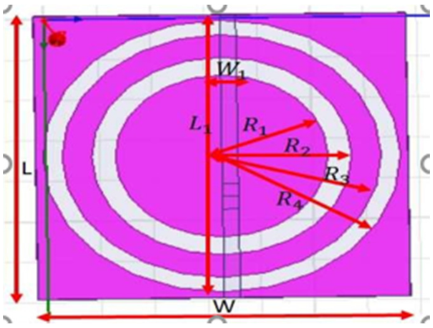


Fig. 1. The designed planar electrically small antenna.

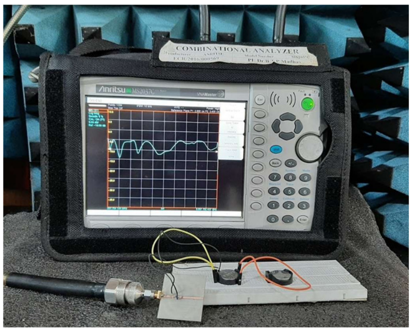
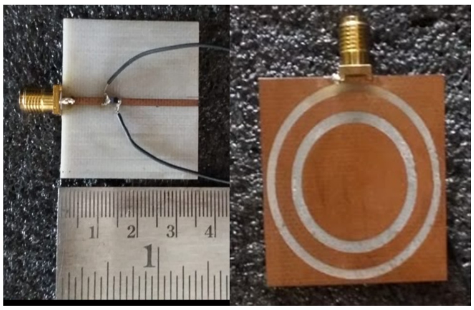


Fig. 2. Antenna parameter testing using VNA.

TABLE I. DIMENSIONS OF THE DESIGNED PLANAR ELECTRICALLY SMALL ANTENNA

Parameter	L	W	L_1	W_1	R_1	R_2	R_3	R_4
Value (mm)	35	35	35	1.5	10	12.2	15	16.8



(a) (b)

Fig. 3. Images showing the prototype of the proposed planar electrically small antenna: (a) top view, (b) bottom view.

IV. RESULTS AND DISCUSSION

Return loss is the most preferred parameter to test an antenna. Instead of using 2 different antennas at different frequencies, a single antenna can switch between 2 frequencies. In the proposed model, the antenna can switch between 1.5GHz and 2.4GHz. Figure 4 shows the return loss of the antenna when the diode is in the OFF state. During the OFF state, the antenna resonates at 2.4GHz, with a return loss value of -15.8dbm for the simulated and -13.01dbm for the actual antenna. Figure 5 represents the equivalent circuit diagram of a pin diode in ON and OFF state. During the ON state, the antenna resonates at 1.5GHz and only the resistor and inductor values are used. During the OFF state, the antenna resonates at

2.4GHz and the resistor, capacitor, and inductor values are used. Figure 6 shows the return loss of the antenna when the diode is at the ON state. The proposed antenna resonates at 1.5GHz with a return loss value of -15.07dB for the simulated model and -13.08dB for the fabricated model.

Voltage Standing Wave Ratio (VSWR) is a significant parameter for estimating the behavior of the RF components at high frequencies. The proposed antenna resonates at 2.4GHz with circular polarization when the diode is in the OFF state. Figure 8 shows the 3D radiation pattern of the antenna with broad side radiation. Maximum radiation exists in the XY plane. The maximum possible gain is about 1dB at θ=0°. Figures 9-10 represent the patterns of the antenna at various azimuthal and elevation planes.

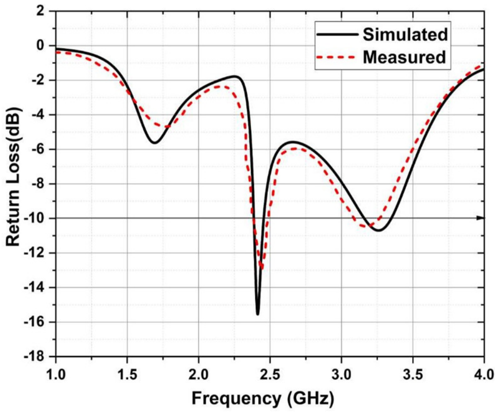


Fig. 4. Return loss of the antenna when the diode is in the OFF state.

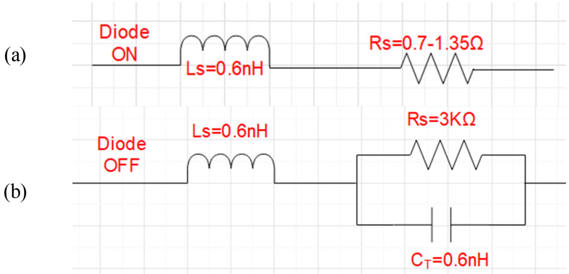


Fig. 5. Equivalent circuit diagram of a pin diode. (a) ON, (b) OFF.

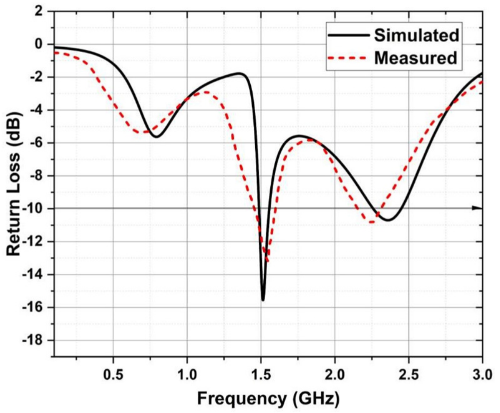


Fig. 6. Return loss when the diode is in the ON state.

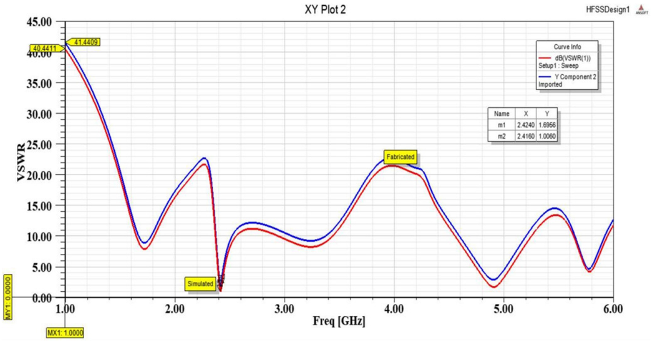


Fig. 7. VSWR of the antenna.

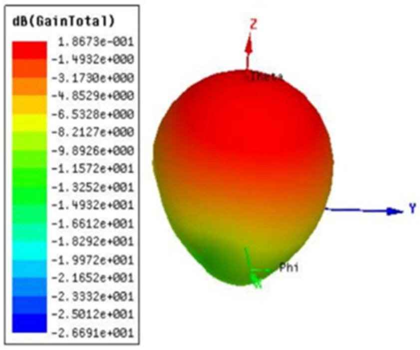


Fig. 8. 3D radiation pattern of the antenna.

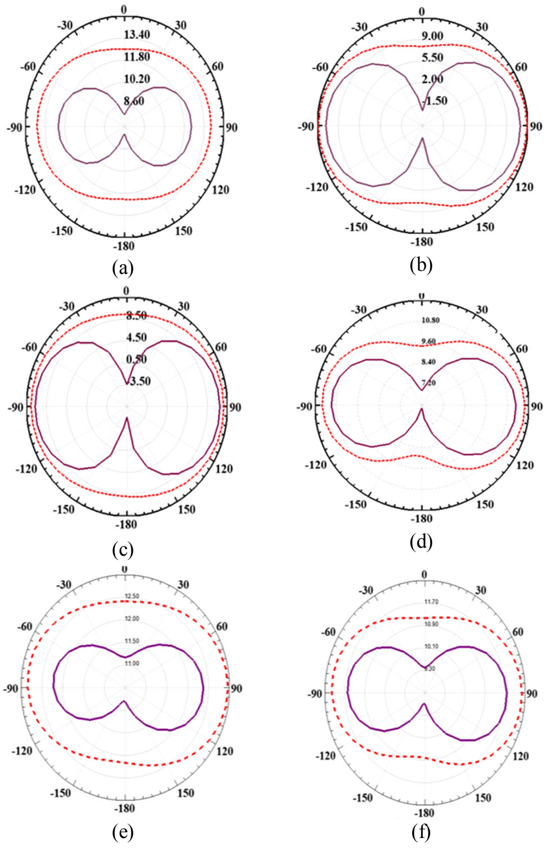
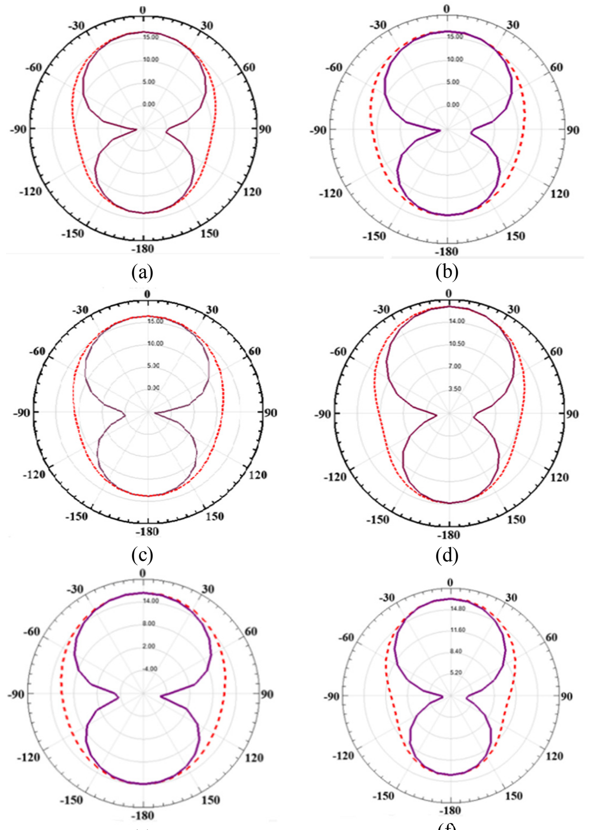


Fig. 9. Azimuth pattern at *θ* = (a) 0°, (b) 30°, (c) 60°, (d) 90°, (e) 120°, (f) 180°. Purple line: simulated results, red line: measured results.



(a) (b)

(c) (d)

(e) (f)

Fig. 10. Elevation pattern at φ= (a) 0°, (b) 30°, (c) 60°, (d) 90°, (e) 120°, (f) 180°. Purple line: simulated results, red line: measured results.

The comparison with other proposed designs in Table II shows that that the proposed antenna offers better results in terms of return loss and gain.

TABLE II. PERFORMANCE COMPARISON OF THE PROPOSED ANTENNA WITH OTHER, KNOWN ANTENNAS

Ref	**Substrate**	**Reconfigurable**	**Area (mm^3)**	**Center Frequency**	**Gain (db)**
[1]	FR-4 Epoxy	Frequency	45×40×1.6	1.28GHz 1.62GHz	-------
[6]	RT Duroid 5880	Frequency	45×38×0.254	2.4GHz	-------
[7]	Polyimide	Frequency	45×35×3.2	1.6GHz	0.9dB
[8]	FR4	Frequency	95×100×14	1.71GHz	--------
Proposed	FR4	Frequency	35×35×1.6	1.5GHz 2.4GHz	1.86dB

V. CONCLUSION

Based on the complementary concept, an electrically small, coplanar antenna was realized by combing the equivalent magnetic dipole of a segmented loop and the electric dipole of a strip line. The antenna is designed on a FR4 epoxy substrate, having a small area of 35×35×1.6mm^3. The antenna showed satisfying broadside radiation performance, exhibiting a gain of 1.86dB. It has been found that the resonance frequencies of both the electric dipole and the magnetic dipole (segmented loop) can be easily tuned with PIN diodes. Therefore, a frequency-reconfigurable complementary antenna was further developed. Two reconfigurable broadside radiation bands have been achieved at 1.5 and 2.4GHz.

REFERENCES

[1] W. B. Hong, K.-H. Back, and S. Ko, "Millimeter- wave 5G antennas for smartphones: overview and experimental demonstration," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6250-6261, December 2017, https://doi.org/10.110 9/TAP.2017.2740963.

[2] H. Maema and T. Fukusako, "Radiation efficiency improvement for electrically small and low-profile antenna by stacked elements," *IEEE Antennas and Wireless Propagation Lett*ers, vol. 13, pp. 305-308, January 2014, https://doi.org/10.1109/LAWP.2014.2305304.

[3] M.-C. Tang, B. Zhou, and R. W. Ziolkowski, "Low- profile, electrically small, Huygens source antenna with pattern-reconfigurability that covers the entire azimuthal plane," *IEEE Transactions on Antennas and Propagation*, vol.65, no. 3, pp. 1063-1072, March 2017, https://doi.org/10.1109/TAP.2016.2647712.

[4] A. Chlavin, "A new antenna feed having equal E -and H-plane patterns," *Transactions of the IRE Professional Group on Antennas and Propagation*, vol. 2, no. 3, pp. 113–119, Jul. 1954, https://doi.org/ 10.1109/T-AP.1954.27983.

[5] L. Siu, H. Wong, and K. M. Luk, "A dual-polarized magneto-electric dipole antenna with dielectric loading," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 3, pp. 616–623, March 2009, https://doi.org/10.1109/TAP.2009.2013430.

[6] H. Wong, K. M. Mak, and K. M. Luk, "Wideband shorted bowtie patch antenna with electric dipole," *IEEE Transactions on Antennas and Propagation*, vol. 56, no.7, pp. 2098–2101, July 2008, https://doi.org/ 10.1109/TAP.2008.924765.

[7] L. Ge and K. M. Luk, "A low-profile magneto- electric dipole antenna," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 4, pp. 1684–1689, January 2012, https://doi.org/10.1109/LAWP.2016. 2519942.

[8] L. Ge and K. M. Luk, "Linearly polarized and dual- polarized magneto-electric dipole antennas with reconfigurable beamwidth in the H-plane," *IEEE Transactions on Antennas and Propagation,* vol. 64, no. 2, pp. 423-431, February 2016, https://doi.org/10.1109/TAP.2015.2505000.

[9] L. Ge, M. J. Li, Y. J. Li, H. Wong, and K. M. Luk, "Linearly polarized and circularly polarized wideband dipole antennas with reconfigurable beam direction," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 4, pp. 1747-1755, April 2018, https://doi.org/10.1109/TAP. 2018.2797520.

[10] F. Wu, K. M. Luk, "Wideband tri-polarization reconfigurable magneto-electric dipole antenna," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 4, pp. 1633-1641, April 2017, https://doi.org/ 10.1109/TAP.2017.2670521.

[11] M.-C. Tang, Z. T. Wu, T. Shi, H. Zeng, W. Lin, and R. W. Ziolkowski, "Dual-linearly-polarized, electrically small, low-profile, broadside radiating, Huygens dipole antenna," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 8, pp. 3877-3885, August 2018, https://doi.org/ 10.1109/TAP.2018.2840835.

[12] W. Lin, R. W. Ziolkowski, "Electrically small, low- profile, Huygens circularly polarized antenna," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 2, pp. 636-643, February 2018, https://doi.org/ 10.1109/TAP.2017.2784432.

[13] J. O.-Y, Y. M. Pan, and S. Y. Zheng, "Center-fed unilateral and pattern reconfigurable planar antennas with slotted ground plane," *IEEE Transactions on Antennas and Propagation,* vol. 66, no. 10, pp. 5139-5149, October 2018, https://doi.org/10.1109/TAP.2018.2860 046.

[14] N. N. Trong, A. Piotrowski, and C. Fumeaux, "A frequency-reconfigurable dual-band low-profile monopolar antenna," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 7, pp. 3336-3343, July 2017, https://doi.org/10.1109/TAP.2017.2702664.

[15] L. Ge and K. M. Luk, "Frequency-reconfigurable low- profile circular monopolar patch antenna," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 7, pp. 3443-3449, July 2014, https://doi.org/10.1109/TAP.2014.2318077.

[16] A. Mansoul, F. Ghanem, M. R. Hamid, and M. Trabelsi, "A Selective Frequency-Reconfigurable Antenna for Cognitive Radio Applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 515–518, 2014, https://doi.org/10.1109/LAWP.2014.2311114.

[17] M. W. Young, S. Yong and J. T. Bernhard, "A miniaturized frequency reconfigurable antenna with single bias, dual varactor tuning," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 3, pp. 946-951, March 2015, https://doi.org/10.1109/TAP.2015.238 8776.

[18] N. Nguyen-Trong and C. Fumeaux, "Tuning Range and Efficiency Optimization of a Frequency-Reconfigurable Patch Antenna," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 1, pp. 150–154, Jan. 2018, https://doi.org/10.1109/LAWP.2017.2778272.

[19] A. Petosa, "An Overview of Tuning Techniques for Frequency-Agile Antennas," *IEEE Antennas and Propagation Magazine*, vol. 54, no. 5, pp. 271–296, Oct. 2012, https://doi.org/10.1109/MAP.2012.6348178.

[20] I. Govardhani *et al.*, "Design of high directional crossed dipole antenna with metallic sheets for UHF and VHF applications," *International Journal of Engineering & Technology*, vol. 7, no. 1.5, pp. 42–50, Dec. 2017, https://doi.org/10.14419/ijet.v7i1.5.9120.

[21] D. V. Doan, K. Nguyen, and Q. V. Thai, "A Novel Fuzzy Logic Based Load Frequency Control for Multi-Area Interconnected Power Systems," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7522–7529, Aug. 2021, https://doi.org/10.48084/etasr.4320.

[22] N. A. Zainurin, S. a. B. Anas, and R. S. S. Singh, "A Review of Battery Charging - Discharging Management Controller: A Proposed Conceptual Battery Storage Charging – Discharging Centralized Controller," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7515–7521, Aug. 2021, https://doi.org/10.48084/etasr.4217.

[23] Z. A. Shamsan, "Statistical Analysis of 5G Channel Propagation using MIMO and Massive MIMO Technologies," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7417–7423, Aug. 2021, https://doi.org/10.48084/etasr.4264.

[24] G. Imamdi, M. V. Narayan, A. Navya, and A. Roja, "Reflector array antenna design at millimetric (mm) band for on the move applications," vol. 13, no. 1, pp. 352–359, Jan. 2018.

The Effect of Low Velocity Impact Loading on Self-Compacting Concrete Reinforced with Carbon Fiber Reinforced Polymers

Julnar Muhsin Abd
Civil Engineering Department
College of Engineering
University of Baghdad
j.abd1901m@coeng.uobaghdad.edu.iq

Ikram Faraoun Ahmed
Civil Engineering Department
College of Engineering
University of Baghdad
Ikram.faroun@coeng.uobaghdad.edu.iq

Abstract-Self-Compacting Concrete (SCC) reduces environmental noise and has more workability. This research presents an investigation of the behavior of SCC under mechanical loading (impact loading). Two types of cement have been used to produce SCC mixtures, Ordinary Portland Cement (OPC) and Portland Limestone Cement (PLC), which reduces the emission of carbon dioxide during the manufacturing process. The mixes were reinforced with Carbon Fiber Reinforced Polymer (CFRP) which is usually used to improve the seismic performance of masonry walls, to replace lost steel reinforcements, or to increase column strength and ductility. Workability tests were carried out for fresh SCC. Prepared concrete slabs of 500×500×50mm were tested for low-velocity impact loading at ages of 28, 56, and 90 days after water curing. The results were compared with the ones of non-reinforced SCC mixes and show a significant effect on the impact resistance after the SCC was reinforced with CFRP. The strongest impact resistance was recorded for reinforcing slabs made from OPC SCC, while for the reinforced concrete slabs produced from PLC the results were less, but at a close rate.

Keywords-self-compacted concrete;carbon fiber reinforced polymer; impact loading test; Portland limestone cement

I. INTRODUCTION

Self-Compacted Concrete (SCC) is a kind of concrete that is cast without being compacted and fills the formwork with its own weight. SCC provides some advantages in comparison with traditional concrete, including enhanced building quality, quicker construction activity, lower costs, increased in-situ concrete quality under challenging casting loads, decreased accidents, noise and vibrations, while it assists in the attainment of higher surface quality [1]. One of the drawbacks of SCC is its large cost, which is due to the usage of large amounts of Portland Cement (PC) and chemical additives. The SCC cost will be lowered if the mineral additions replace a portion of the PC, particularly if the mineral additives are industrial waste or by-products. Furthermore, the use of mineral additives in the manufacturing of SCC has not only economical but also environmental benefits [2]. Cement manufacturing emits a large quantity of carbon dioxide and necessitates a large amount of energy [3]. Blending limestone with conventional OPC or intergrading limestone and cement clinker produces Portland Limestone Cement (PLC). The increased use of PLC in the concrete industry is mostly owed to its lower raw material usage and energy savings in clinker production [4]. Components made of Fiber Reinforced Polymer (FRP) have been utilized extensively and effectively in seismic enhancement projects [5].

Traditional metal and ceramic materials are being displaced in a wide range of high-performance implant applications due to their high specific strength, superior corrosion resistance, low weight, and design flexibility. The fiber-matrix interface is essential to the performance of fiber-reinforced composites. A visible interface can guarantee that stress is effectively transmitted from the weaker matrix to the stronger fiber, reducing stress concentrations and improving the overall mechanical characteristics [6]. Throughout their functioning lives, many structures or structural parts are likely to be subjected to impact forces or even repeated impacts. Impact forces can occur by a variety of causes. One example is the impact of moving automobiles on columns and walls of structures in accidents. Other instances include the effect of dropped items from high elevations, the effects of projectiles in wars or terror acts and the effects of water on hydraulic infrastructure, offshore structures, or bridge piers. Impact forces apply short-term dynamic loads to the material, subjecting it to unusual and unwanted stresses, particularly in brittle materials such as concrete. This type of loading necessitates improved microstructural performance in order to be efficiently absorbed. Additive fibers can improve concrete responsiveness under static and dynamic loads. Several testing techniques have been developed in order to get a satisfactory measurement of FRC impact resistance. The drop-weight, projectile impact Charpy pendulum, and the explosive and Repeated Blows Drop-Weight Impact (RBDWI) tests are among them [7].

II. MATERIAL CHARACTERIZATION

A. Cement

In this study, two types of cement were employed, PLC and OPC. The PLC was manufactured in Karbala city in Iraq by the

---

Corresponding author: Julnar Muhsin Abd

Lafarge Company, it is commercially known as Karasta, and it was purchased at a local market. To minimize exposure, it was stored in a dry area. Tables I and II show the chemical composition and the physical properties of the Karasta cement. This cement complied with the EN 197-1:2011 [8]. OPC commercially known as Mass was also utilized. Its physical and chemical properties conform to the Iraqi Standard No. 5 of 2019 [9] and are given in Tables III and IV.

TABLE I. CHEMICAL COMPOSITION OF KARASTA CEMENT

Compound	Percentage by weight	Limit of EN 197-1:2011
Magnesia (MgO)	1.83	≤ 5%
Sulfate (SO_3)	2.35	≤ 2.5
CaO	62.11	
L.O.I	7.44	≤ 4
SiO_2	18.39	
Al_2O_3	4.63	
Fe_2O_3	2.77	
Total	99.52	
C_3A	7.58	
Chloride content	0.01	≤ 0.10 %
I.R	1.38	≤ 1.5

TABLE II. PHYSICAL PROPERTIES OF KARASTA CEMENT

Physical properties	Test result	Limits of EN 197-1:2011
Specific surface area (Blaine method) cm^2/g	4875	≥ 2500
Initial setting (min)	90	≥ 45
Compressive strength (MPa) at:		
2 days	8	≥ 10 - 20
28 days	34	≥ 42.5

TABLE III. CHEMICAL COMPOSITION OF MASS CEMENT

Oxides	Test results	Limits of the Iraqi standard No.5/2019 for OPC
CaO	62.32	-
SiO_2	21.34	-
Al_2O_3	4.93	-
Fe_2O_3	5.43	-
MgO	2.12	Max 5%
SO_3	2.35	Max2.8%
L.O.I.	1.72	Max 4%
LSF	0.86	0.66-1.02
Saturation coefficient	0.83	Max 1.5%
Constituent% according to Bogue equations		
C_3S	76.28	-
C_2S	3.73	-
C_3A	6.72	-
C4AF	16.5	-

TABLE IV. PHYSICAL PROPERTIES OF MASS CEMENT

Physical properties	Test results	Limits of the Iraqi standard No.5/2019 for OPC
Specific surface area (Blaine method) m^2/kg	376	min 250
Setting time (Vicat's apparatus)		
Initial setting (min)	2:5	≥ 45
Final setting (hr)	4:00	≤ 10
Soundness using autoclave method (mm)	0.12	≤ 0.8
Compressive strength (MPa) at:		
2 days	26	≥ 20
28 days	47	≥ 42.5

B. Fine Aggregates

As fine aggregates, natural sand from the Al-Ukhaider area was used. Table V shows the physical and chemical properties and Table VI shows the sieve analysis of the sand used, which conforms to the limits of Iraq specification No. 45/1984 Zone 2 [10].

TABLE V. CHEMICAL AND PHYSICAL PROPERTIES OF FINE AGGREGATES

Property	Test result	I.Q.S.45: 1984 limits
Specific gravity	2.6	-
Absorption (%)	0.72	-
Density (kg/m^3)	1580	-
Sulphate content (SO_3)	0.2	0.50% (max)

TABLE VI. SIEVE ANALYSIS

Sieve no.	Passing %	Limits of Iraqi specification no.45/1984
10 mm	100	100
4.75 mm	93.3	100-90
2.36 mm	77.7	100-75
1.18 mm	66.6	90-55
600 μm	54.4	59-35
300 μm	26.3	30-8
150 μm	3.1	100

C. Coarse Aggregates (Gravel)

The maximum size of the crushed gravel used was 10mm. The gravel was obtained from the Al-Nabaie area in Iraq. Its physical properties were examined in the laboratory of the Ministry of Construction and Housing, Building Research Department, and it satisfies the Iraqi standard IQS No. 45/1984 as shown in Table VII.

TABLE VII. GRADING OF NATURAL COARSE AGGREGATES

Sieve size (mm)	Cumulative passing (%)	Limits of IS NO.45/1984 for 14-15mm
20	100	100
14	98	100-90
10	75	85-50
5	4	10-0
Sulphate cotent (SO_3)	0.6	Maximum 1%

D. Limestone Powder

Limestone powder was used as an inert mineral filler in order to maintain and improve the segregation resistance, to increase the amount of powder (cement and filler), and to increase the workability and density of concrete. Its chemical composition is shown in Table VIII.

TABLE VIII. CHEMICAL COMPOSITION OF LIMESTONE POWDER

Oxides	%
CaO	48.28
MgO	3.97
L.O.I	42.48
IR	2.00

E. High Range Water Reducing Admixture (HRWRA)

Betonac-1030 is a highly efficient polycarboxylate polymer based superplasticizer that is intended to preserve workability and improve slump-flow life and thus it provides concrete of extremely high workability without segregation, achieving high initial compressive strength. It permits the constant production of self-compacting concrete and complies with the criteria of ASTM C494 Type F and DIN EN 934-2 [11]. Typically, 0.6% to 1.5% of cement weight is used, although this can be modified to fulfill a broader range of concrete specifications. The technical description of Betonac-1030 is presented in Table IX.

TABLE IX. TECHNICAL DESCRIPTION OF BETONAC-1030

Calcium chloride	nail
Solid contents	40%
PH	7.5 at 20°C
Packaging	Available in 220kg drums or 1100kg IBCs
Storage/shelf life	At least 1 year if stored in originally sealed packaging and protected from direct sunbeam, extreme heat, and frost.

*Manufacturer properties

F. Water

In concrete mixes and specimen treatment, tap water from the university campus was employed. The temperature of the mixing water was kept constant at 25.2°C

G. Carbon Fiber Reinforced Polymer

The carbon fiber used is a mid-strength unidirectional woven carbon fiber fabric commercially known as SikaWrap-300 C, intended for either dry or wet use. It is a fabric that may be used in a variety of strengthening purposes. It is flexible and adaptable to various surface planes and geometries (columns, beams, piles, walls, silos, etc). Also, it has a low density and contributes as little weight as possible. In comparison with conventional strengthening, it is highly cost effective. Table X indicates the technical properties of the carbon fibers used according to the manufacturer company.

TABLE X. CARBON FIBER PROPERTIES

Laminate nominal thickness	0.167mm	
Laminate nominal cross section	167mm^2 per m width	
Dry fiber density	1.82g/cm^3	
Dry fiber thickness	0.167mm based on fibre content	
Area density	304g/m^2 ±10g/m^2 (carbon fibres only)	
Dry fiber tensile strength	4000N/mm^2	ISO 10618
Dry fiber modulus of elasticity in tension	230000N/mm^2	
Dry fiber elongation at break	1.7 %	

*Manufacturer properties

H. Epoxy Impregnation Resin

Sikadur-330 is an anisotropic epoxy-based impregnating resin and adhesive in two components. It can be utilized in dry and wet applications. The wet application technique requires a primer resin for bonding. It is produced for manual immersion procedures and is simple to mix and apply with a trowel and an insemination roller. It has excellent adhesion to vertical and overhanging surfaces and excellent adherence to a wide range of substrates. Its mechanical characteristics are exceptional and there is no need for a different primer. Table XI indicates the technical properties of Sikadur-330

TABLE XI. PROPERTIES OF SIKADUR-330

Composition	**Epoxy resin**
Color component	A: white paste B: grey paste A + B mixed: light grey paste
Density	1.30 ± 0.1kg/l (component A+B mixed) at 23°C
Modulus of elasticity in flexure	~ 3 800 N/mm^2 (7 days at 23°C) (DIN EN 1465)
Tensile strength	~ 30 N/mm^2 (7 days at +23°C) (ISO 527)
Modulus of elasticity in tension	~ 4 500 N/mm^2 (7 days at +23°C) (ISO 527)
Tensile adhesion strength	Concrete fracture (> 4N/mm^2) on sandblasted substrate (EN ISO 4624)

I. Concrete Mixes

To achieve the research objectives, two types of mixes were prepared according to ACI-211 [12]. The details of the mixes used throughout this investigation are given in Table XII.

TABLE XII. DETAILS OF THE PREPARED MIXES

Mix symbol	**Description of the mix**
P1	Reference mix with OPC
R1P	Mix with OPC reinforced with carbon fibers
P2	Mix with Karasta LPC
R2P	Mix with Karasta LPC reinforced with carbon fibers

TABLE XIII. MAIN DETAILS OF THE USED MIXES

Mix	**Cement Kg/m3**		**FA kg/m^3**	**CA kg/m^3**	**LD kg/m^3**	**W kg/m^3**	**SP kg/m^3**	**W/p**
	OPC	**PLC**						
P1	400		684	760	80	168	6	0.42
P2		400	684	760	80	168	6	0.42

J. Mixing Procedure

The mixing process was carried out in accordance with the ASTMC-192/02 standard [13] and cohesive concrete mixtures were obtained. The mixing steps are: 1. Fine Saturated Surface Dry (SSD) aggregates were mixed with the one-third of water. 2. The cement and the mineral additives were added along with another one-third of water, and then the ingredients were mixed for one minute. 3. The SSD coarse aggregates were added with the last third of the water and one third of the superior plasticizer. The mixture was mixed for a minute and a half, and then it was left for half a minute to rest. 4. Then, two thirds of the remaining superior plasticizer were added and the mixture was mixed well for a minute and a half. 5. The mixture was discharged for testing and pouring. The total mixing time is about 5 minutes.

III. RESULTS AND DISCUSSION

A. Fresh Properties

1) Slump Flow

Stagnant flow results are shown in Table XV. All results are within the acceptance criteria of EFNARC, 2005 [14] (Table XIV).

TABLE XIV. ACCEPTANCE CRITERIA [14]

Method	**Unit**	**Typical range of values**	
		Minimum	Maximum
Slump flow by Abrams cone	mm	650	800
T50 cm slump flow	sec	2	5
V-funnel	sec	6	12
Time increase, V-funnel at T=5min	sec	0	3
U-box	h1-h2 (mm)	0	30
L-box	h2/h1	0.8	1.0

Fig. 1. Slump test.

TABLE XV. SCC RHEOLOGICAL PROPERTIES RESULTS

Mix	**Slump flow (mm)**	**T500 (sec)**	**V-funnel time (sec)**	**L-box height ratio (h2/h1)**
OPC	667	5	11	0.825
PLC	7 44	3	8	0.941

2) *T500mm*

Slump flow times (T_{500}) were in the range of 3-5s. All results are within the acceptance criteria of [14].

3) *V-Funnel Time*

The results are consistent with the reported SCC accepted standards [14]. The stagnation values for binary and ternary mixtures that include OPC and PLC are shown in Table XIV. Leveling results range from 667 to 744mm. Therefore, all mixtures are assumed to have good consistency and workability from the packaging point of view.

4) *L-Box*

The L-box height ratios varied from 0.843 to 0.948 (Table XV). From the obtained results it was found that the fluidity was good for all the mixtures since it was greater than 0.8, which is considered the minimum critical value [14].

B. Mechanical Properties (Impact Resistance)

The device used to measure impact resistance contains three parts:

- The main support part: It consists of an iron bracket that is strong enough to support the examination device during the examination and withstand the impact. It also contains another support for which the examination form is placed. It is made of angle iron and is well fixed to prevent the movement of the form during the examination.
- The iron structure for dropping the block: It is a cylindrical tube with an inner diameter of 10.5cm which is fixed with supports that prevent it from moving during the examination. This tube contains openings of different heights (0.5, 1.0, 1.7m). The height of 1m was used to examine the reinforced slabs for this research.
- The falling block: It is an iron ball, weighing 3kg with a diameter of 9.5cm. This mass is thrown several times from a height of 1m on the model.

Fig. 2. L-box test.

C. Examination Method

Each model was placed on the supporting iron structure, and then the iron block was dropped from the specified height of 1m on the model. The number of blows that were required for the first crack to occur and the number of strikes that caused failure are shown in Table XVI. The average of 3 slabs was taken for each variable at ages of 28, 56, and 90 days according to [15]. The failure pattern was observed in the unreinforced slab, which is divided into several pieces. The typical failure behavior of unreinforced concrete specimens under impact resistance is shown in Figure 4. The stages of failure in the reinforced slab were categorized as: in the first stage the first crack occurs, during the second stage fragmentation occurs, and at the third stage failure occurs with cracks in the corners of the slab and the occurrence of a full hole so that the reinforcement appears. The typical failure behavior of reinforced concrete specimens under impact resistance is shown in Figure 5.

Reinforcement of SCC improves the impact resistance because the carbon fiber reinforced with polymer used in this research has good mechanical properties to curb cracking and absorb high energy when withdrawn from the concrete, which makes the reinforced slabs to have high strength and flexibility and makes them capable of withstanding external shocks better than SCC without reinforcement. The impact resistance (number of blows) increases with the curing age, as can be seen in Figure 3. This is consistent with the findings in previous studies that examined the reaction of low impact velocity Carbon Fiber Reinforced Polymer (CFRP) epoxy laminates.

TABLE XVI. SLAB UNDER THE IMPACT RESISTANCE TEST

Mix	**Average impact resistance**					
	28 days		56 days		90 days	
	Number of blows required for					
	first crack	failure	first crack	failure	first crack	failure
P1	1	2	1	3	1	3
P1R	2	40	2	54	2	70
P2	1	2	1	2	1	3
P2R	2	37	2	49	2	65

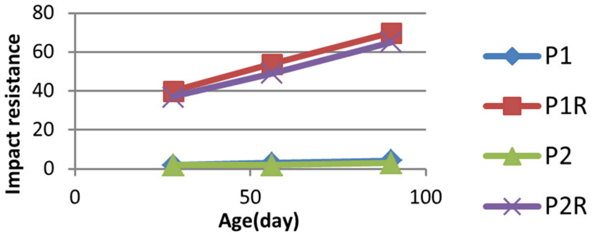


Fig. 3. Impact resistance with different ages of curing.

Fig. 4. Typical failure behavior of unreiforced concrete slabs.

Fig. 5. Typical failure behavior of reiforced concrete slabs.

It is known that damage caused by low impact velocity (where there is no hole) in CFRP laminates rises with the impact energy when submitted to impact loads. As a result, the design of fibers and the sequence of stacking play a crucial part in the compliance under these temperature and loading circumstances with CFRP composites. Thermal variation does not produce interlaminar strains in the plain-woven laminate employed, where the fibers were oriented parallel in each ply, and damage extension after impact does not rise significantly at low temperatures [16]. Based on the extensive experimental data, reinforcing of SCC with the use of recycled CFRP fibers increases the mechanical characteristics and impact resistance of specimens [17].

IV. CONCLUSION

The following conclusions can be drawn from the results of the experimental work:

- The use of limestone powder as filler for SCC and the replacement percentage do not affect the properties of solid concrete.
- Carbon fibers can be used for reinforcement and consolidation and they show high impact resistance.
- Impact resistance was more powerful and reached 70 blows in slabs containing OPC at the age of 90 days.
- The results between SCC with OPC slabs and SCC with PLC slabs are not so paradoxical.

This research contributes to the understanding of the impact behavior of concrete slabs made with carbon fibers. In practice, concrete slabs reinforced with CFRP and produced from PLC can be used in many different applications. The SCC made with locally available materials could be widely used instead of conventional concrete at complex construction sites. This reinforcing system is used to increase the strength and ductility of columns and the loading capacity of structural elements, enabling changes in use, alterations and refurbishment, and structural design and construction defect corrections. The use of lime dust and PLC to produce sustainable SCC is environmentally beneficial since the carbon dioxide emissions and energy consumption are reduced by reducing the amount of cement being manufactured and used.

REFERENCES

[1] N. A. Memon, M. A. Memon, N. A. Lakho, F. A. Memon, M. A. Keerio, and A. N. Memon, "A Review on Self Compacting Concrete with Cementitious Materials and Fibers," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2969–2974, Jun. 2018, https://doi.org/10.48084/etasr.2006.

[2] M. Uysal and M. Sumer, "Performance of self-compacting concrete containing different mineral admixtures," *Construction and Building Materials*, vol. 25, no. 11, pp. 4112–4120, Nov. 2011, https://doi.org/10.1016/j.conbuildmat.2011.04.032.

[3] N. Bheel, A. S. Memon, I. A. Khaskheli, N. M. Talpur, S. M. Talpur, and M. A. Khanzada, "Effect of Sugarcane Bagasse Ash and Lime Stone Fines on the Mechanical Properties of Concrete," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5534–5537, Apr. 2020, https://doi.org/10.48084/etasr.3434.

[4] F. Shaker, A. Rashad, and M. Allam, "Properties of concrete incorporating locally produced Portland limestone cement," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 2301–2309, Dec. 2018, https://doi.org/10.1016/j.asej.2017.04.005.

[5] N. Askarizadeh and M. R. Mohammadizadeh, "Numerical Analysis of Carbon Fiber Reinforced Plastic (CFRP) Shear Walls and Steel Strips under Cyclic Loads Using Finite Element Method," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2147–2155, Dec. 2017, https://doi.org/10.48084/etasr.1279.

[6] S. R. Abid, M. L. Abdul-Hussein, N. S. Ayoob, S. H. Ali, and A. L. Kadhum, "Repeated drop-weight impact tests on self-compacting concrete reinforced with micro-steel fiber," *Heliyon*, vol. 6, no. 1, Jan. Art. no. e03198, 2020, https://doi.org/10.1016/j.heliyon.2020.e03198.

[7] M. A. Sanjuán and C. Argiz, "The new european standard on common cements specifications EN 197-1:2011," *Materiales de Construcción*, vol. 62, no. 307, pp. 425–430, Sep. 2012, https://doi.org/10.3989/mc.2012.07711.

[8] *Iraqi Specification No. 5 : Portland Cement*. Baghdad, Iraq: Central Agency for Standardization And Quality Control, 2019.

[9] *Iraqi Specification No. 45: Aggregates*. Baghdad, Iraq: Central Agency for Standardization And Quality Control, 1984.

[10] C09 Committee, "Specification for Chemical Admixtures for Concrete," ASTM International, West Conshohocken, PA, USA. https://doi.org/10.1520/C0494_C0494M-12.

[11] ACI Committee 211, *ACI PRC-211.1-91: Standard Practice for Selecting Proportions for Normal, Heavyweight, and Mass Concrete (Reapproved 2009)*. ACI, 2002.

[12] "Standard Practice for Making and Curing Concrete Test Specimens in the Laboratory," ASTM International, West Conshohocken, PA, USA, ASTM C192/C192M-07.

[13] BIBM, CEMBUREAU, ERMCO, EFCA, and EFNARC, *The European Guidelines for Self-Compacting Concrete: Specification, Production and Use*. 2005.

[14] ACI Committee 544, "Design Considerations for Steel Fiber Reinforced Concrete," ACI, West Conshohocken, PA, USA, ACI 544.4R-88, 1988.

[15] T. Gómez-del Río, R. Zaera, E. Barbero, and C. Navarro, "Damage in CFRPs due to low velocity impact at low temperature," *Composites Part B: Engineering*, vol. 36, no. 1, pp. 41–50, Jan. 2005, https://doi.org/10.1016/j.compositesb.2004.04.003.

[16] M. Mastali and A. Dalvand, "The impact resistance and mechanical properties of self-compacting concrete reinforced with recycled CFRP pieces," *Composites Part B: Engineering*, vol. 92, pp. 360–376, May 2016, https://doi.org/10.1016/j.compositesb.2016.01.046.

Optimal Shedding Against Voltage Collapse Based on Genetic Algorithm

Mohamed Ali Zdiri
CEM Laboratory
Engineering School of Sfax
Sfax, Tunisia
mohamed-ali.zdiri@enis.tn

Ahmed S. Alshammari
Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
ahm.alshammari@uoh.edu.sa

Ahmed A. Alzamil
Department of Electrical Engineering
University of Hail
Hail, Saudi Arabia
aa.alzamil@uoh.edu.sa

Mohsen Ben Ammar
CEM Laboratory
Engineering School of Sfax
Sfax, Tunisia
mohsen.benammar@enis.tn

Hsan Hadj Abdallah
CEM Laboratory
Engineering School of Sfax
Sfax, Tunisia
hsan.hajabdallah@enis.tn

Abstract-The prevalent tendency in power transmission systems is to operate closer and closer to the energy limit, rendering system voltage instability a commonly widespread phenomenon. It is, therefore, necessary that certain remedial corrective controls need be undertaken whenever these systems tend towards failure. In this respect, load shedding stands as a major correction mechanism and such a failure can be prevented and nominal system voltage can be resumed. It is worth noting however that load shedding must be implemented very carefully to ensure the satisfaction of both the customer and the electricity-production company. In this context, our focus of interest is laid on load and machine shedding against voltage collapse as an effective corrective method. It is important to note that such a problem turns out to be commonly defined as an optimization problem under constraints. Using genetic algorithms as resolution methods, the application of the proposed methods was implemented on the 14-node IEEE test network, while considering a number of different case studies.

Keywords-power transmission; load and machine shedding; voltage collapse; genetic algorithm; 14-node IEEE test network

I. INTRODUCTION

Electric power plays a crucial role in almost every domain, particularly lighting, communication, and transportation. However, electricity production and transmission are not without serious hazards, likely to disrupt their normal operation and continuity. Actually, numerous disturbances are known to take place worldwide every year, often resulting in noticeable blackouts. Large-scale blackouts, though very rare to occur, could cause huge losses and severe damages to the society and the economy. In this respect, the electrical grid might well undergo or experience a severe voltage collapse phenomenon, bringing about serious consequences [1]. Such a phenomenon often finds its origin in unexpected cascades of events, causing sudden voltage drops, usually leaving little time, often just a few minutes, for the necessary protective measures to be implemented. Hence, the voltage stability margins are assigned even a greater effective role to meet market needs [2-5].

The inherent difficulties in analyzing the various voltage collapse associated problems result primarily from the more or less non-linear behavior of the electrical grid various components. In this regard, the incident cases reported to take place over the past few decades prove that most of the voltage collapse problems have been recorded to occur after a significant disturbance or sudden increase in the electrical grid load, wherein the latter turns out to be weakened and its reactive consumption increased. Accordingly, the phenomenon is characterized with a gradual voltage decrease in one or more consuming regions, immediately accelerating within a few minutes. Hence, the voltage degradation, at the load level, turns out to be so significant that it leads to service interruptions, culminating directly in increased reactive grid losses and voltage drops as major outcomes. Failure in voltage regulation, at the group terminal level, engenders an acceleration of the voltage plane degradation, likely to culminate in cascade tripping, of groups and lines, as well as voltage collapse of the general electricity network. Such faults could well affect the

---

Corresponding author: Mohamed Ali Zdiri

power converters' operation, bringing about a total shutdown of the entire electrical production chain [6-9].

The electricity companies usually charge the transmitted energy to the consumer per MWh, while aspiring for an optimization of the production cost. Accordingly, each of the generators' and plants' performance and contribution must be determined in such a way as to ensure the effective minimization of power producing cost throughout the power system operation process. In effect, the evolution in electrical energy consumption has resulted in significant increases in power transportation and generating costs, hence the importance of an efficient strategy aiming at minimizing the electrical energy exploitation and generation costs. The environmental impact of such processes turns out to be increasingly important and seriously considered when developing the electricity power production and transmission procedures. To achieve these goals, several optimization methods and strategies have been applied [10-11]. The classic methods would serve to solve the single-objective optimization problem (production cost optimization) whereas the most efficient generator supplied grid would not be useful. These methods were designed to obtain voltages in the necessary boundary zones reliable enough to ensure that the grid would remain in an operating point, far from a voltage collapse level, with maintained stability. In [12], for instance, the authors proposed a method in which the determination of the load shedding localization relies heavily on the two buses' phase angle sensitivity. More recently, however, several researchers have developed distinct load shedding methods, enabling to adjust the frequency and voltage within the required or demanded limits while reducing the amount of load shedding [13-18].

In the present paper, we opted for the genetic-algorithm methods which are suitable for treating a single-objective optimization problem. Noteworthy, this field of interest is very dynamic and is still exhibiting continuous development. Accordingly, the study's major objective is focused on applying these optimization methods through implementing genetic algorithms for optimal load and machine shedding against voltage collapses. Hence, to ensure the system's continuity within the available permitted limits while maintaining minimum production cost, two distinct shedding modes are envisaged, namely load shedding and machine shedding. In this context, the choice of the appropriate shedding scheme rests exclusively on the computation of the limit's required power. To this end, a test is administered on the load distribution program to set the limit's value. The process implementation procedure was conducted via the IEEE 14-node standard test network.

II. SHEDDING AGAINST VOLTAGE COLLAPSE

With regard to the permanent mode, the study of voltage collapse provides a solution to the magnitudes of an electrical grid in normal balanced operation. In this respect, the relevant quantities involve the voltages, the node injected powers as well as the line flowing power, and the relevant currents and losses are deduced therefrom. Voltage drops accompany the power transfer between the consumption and the production points, under normal operating conditions. These drops are usually of a small percentage of the normal voltage. Actually, voltage collapse is most often due to increased load, lack of reactive power and/or short circuit. In effect, the process of shedding against voltage collapse depends on the knowledge of the required power limit value. At this level, an electrical grid is considered to be stable, from a voltage point of view, once each node's respective voltages are set within the allowable limit. In the load-distribution scheme, the power required is gradually increased until exceeding the admissible voltage limits. In this case, the associated power value is dubbed as the demanded limit power. It is in terms of the PD value (the demanded power) that the fit shedding type can be actually defined. Accordingly, for a PD greater than the demanded limit power (pdlim), the load shedding process is carried out, and, inversely, however, machine shedding is affected. Figure 1 illustrates the way the appropriate shedding type can be selected against voltage collapse.

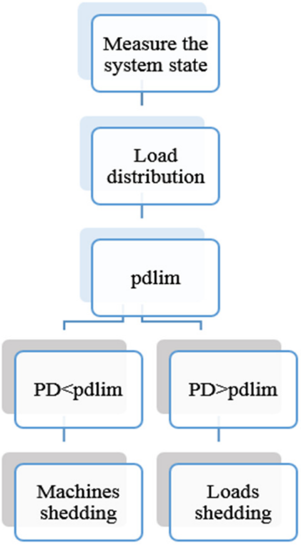


Fig. 1. Shedding type.

III. FORMULATION OF THE GENETIC ALGORITHM BASED OPTIMAL SHEDDING PROBLEM

The loads' and generators' optimal shedding with static constraints takes into account the nodes' voltage limits, the lines' powers transmitted, and the generators' outputs. In our problem formulation, two objective functions persist, one related to the loads, while the other deals with machines.

A. Loads and Machines' Shedding Objective Functions

The objective function, relevant to the load nodes to be minimized is:

$$F(x) = \sum_{i=1}^{nl} (C_{ti}\,\lambda_i) \quad (1)$$

where C_{ti} denotes the load shedding cost at node *I*, λ_{ti} stands for the load-shedding factor at node *I*, and *nl* designates the number of consuming nodes. The loads' associated cost is fixed ahead in accordance with the nodes' importance.

As for the generator node related objective cost function, it is formulated in the form of a generated power function, as illustrated through the following equation:

$$f(x) = \sum_{i=0}^{Ng}(a_i + b_i\, pg_i + c_i\, pg_i^{2}) \quad (2)$$

where Ng designates the number of generators, pg_i stands for the generated power, and a_i, b_i and c_i are the generated powers' relevant coefficients.

This problem type involves special control variables to be defined subsequently. Control variables designate the problem-associated input variables, likely to be adjusted to optimize the loads' and machines' objective functions as well as the constraints' adjustment. These variables are of the form: $x = \{v_i,\ P_i\ and\ \lambda_i\}$, where v_i represents the voltage at node *i*, P_i denotes the active power at node i, and λ_i designates the shedding factor at node *i*.

B. Constraints

1) Equality Constraints

For loads:

$$P_{Gi} - (1-\lambda_i)\, P_{Di} - \sum_{j=1}^{n}|V_i||V_j|\big(G_{ij}\cos(\delta_i - \delta_j) + B_{ij}\sin(\delta_i - \delta_j)\big) = 0 \quad (3)$$

with *i* = 1, ..., *n* is the number of loaded nodes, *n* denotes the set of nodes connected to *i*, PD_i designates the load active power at node's i level, G_{ij} represents the real part of the element in the nodal admittance matrix (*Ybus*) by corresponding row *i* and column *j*, B_{ij} is the imaginary part of the element in the nodal admittance matrix by corresponding row *i* and column *j*, and δ_i stands for the V_i voltage.

For generators:

$$P_{Gi}\,(1-\lambda_i) - P_{Di} - \sum_{j=1}^{n}|V_i||V_j|\big(G_{ij}\cos(\delta_i - \delta_j) + B_{ij}\sin(\delta_i - \delta_j)\big) = 0 \quad (4)$$

where *i* = 1, ..., *n* designates the number of machine nodes, and *n* denotes the *i* connected nodes set.

2) Inequalities Constraints

These constraints reflect the limits set on the power system physical devices, along with the limits created to ensure the system's security. Indeed, for the sake of maintaining the electrical system's safety, transmission lines bear certain limits regarding the generated power, owing to thermal losses occurring at the conductor level, such as:

$$pg_i^{min} < pg_i < pg_i^{max} \quad (5)$$

where pg_i^{min} and pg_i^{max} designate the minimum and maximum generated power and pg_i denotes the power generated at the node *i*.

Additionally, the voltage level must be maintained within the allowable range to ensure the grid's smooth running smoothly, while preserving customer satisfaction. Any voltage drops would engender severe disturbances with respect to any load type, specifically:

$$v_i^{min} < v_i < v_i^{max} \quad (6)$$

where v_i^{min} designates a minimum voltage equal to 0.9pu, v_i^{max} denotes a maximum voltage of the range of 1.1pu, and v_i represents the node *i* corresponding voltage.

As the load and generator shedding factors are limited to specific minimum and maximum values, the entirety of the λ_i relevant values must be set between λ_i^{min} and λ_i^{max} ([0, 1]), such as:

$$\lambda_i^{min} < \lambda_i < \lambda_i^{max} \quad (7)$$

where λ_i stands for the shedding factor at the level of node *i*.

C. The Genetic Algorithm

Genetic algorithms are optimization algorithms with natural evolution drawn techniques: crossing, mutation, selection etc.. They help provide solutions to unresolved problems likely to be computed in an algorithmic finite-time manner [19]. Figure 2 depicts the genetic algorithm's structure, while the flowchart in Figure 3 highlights our suggested problem-solving algorithm sample. λg and λc stand for the machine and load associated shedding factors respectively.

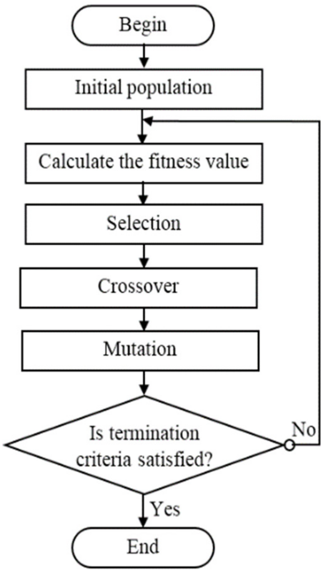


Fig. 2. Genetic algorithm structure.

The algorithmic steps shown in Figure 3 are:

Population initialization

If PD > pdlim

Calculation of the Pc and Qc load active and reactive power, respectively.

Solving these load values distribution problem: if the voltage is within its limit, we return to step 1, otherwise:

Formulate the optimization problem and determine the control parameters' limits

Solve the optimization problem using the genetic algorithm.

The best ultimately withdrawn individual represents the load's optimal shedding factor result.

Carry out the load shedding process.

Display the results.

If PD < pdlim

Same work with the generators' constraints and objective function.

End.

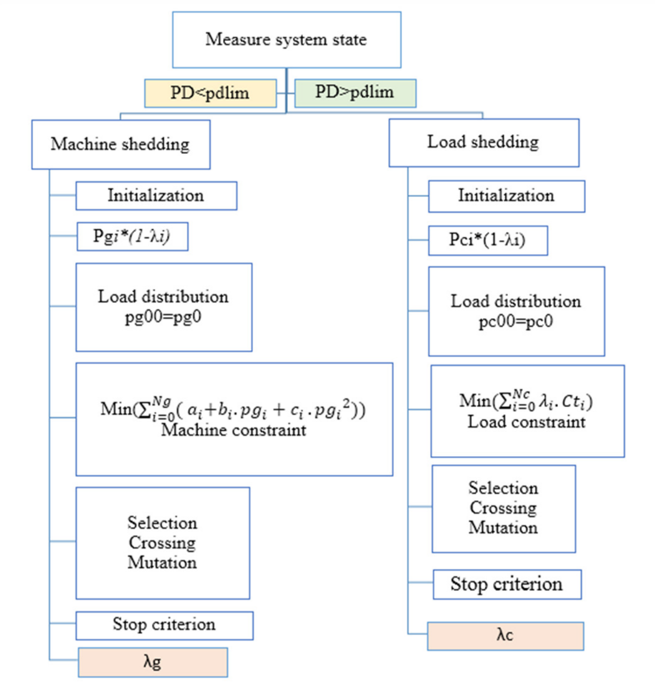


Fig. 3. Optimal load and machine shedding algorithm flowchart.

IV. APPLICATION AND SIMULATION RESULTS

A. Application on the 14-Node IEEE Test Network

For the implementation purposes of our scheme, we considered the 14-node IEEE test network, involving 5 generators, 9 loads and 20 lines, shown in Figure 4.

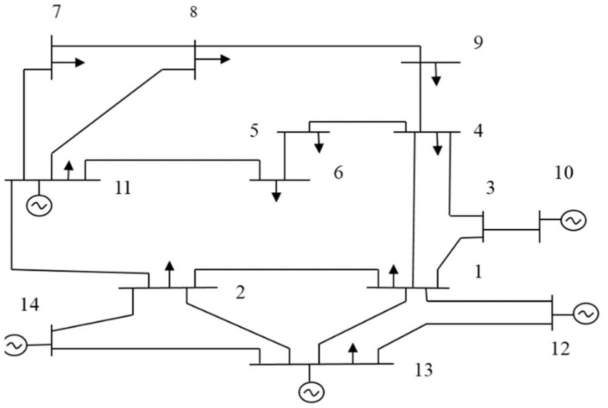


Fig. 4. Structure of the 14-node IEEE test network.

Line data, production limits, specific coefficient of each generator, voltage limits relevant data, load shedding cost functions associated coefficients, and the node data are illustrated in Tables I-IV. It is important to note that all the nodes limits' shedding factors are set between 0 and 1.

B. Simulation Results

MATLAB was utilized to solve the voltage collapse problem. The PD used is equal to 8pu, which exceeded the limit demanded power pdlim=5.4pu. Accordingly, the shedding type is loads shedding. Table V highlights the applied genetic algorithm associated parameters. The relevant convergence profile is illustrated in Figure 5. Accordingly, one could well note that the genetic algorithm appears to converge starting from generation number 55, which corresponds to the most optimal solution.

TABLE I. NODE RELEVANT DATA

Line number	Liaison	Impedance (pu)	Line number	Liaison	Impedance (pu)
1	1-2	0.01335+j0.04211	11	10-2	0.05695+j0.17388
2	1-3	j0.20912	12	10-11	0.04699+j 0.19797
3	1-4	j0.55618	13	11-1	0.06701+j0.17103
4	3-4	j0.11001	14	12-6	0.09498+j0.1989
5	10-1	0.05811+j0.17632	15	12-7	0.12291+j0.25581
6	4-5	0.03181+j0.0845	16	12-8	0.06615+j0.13027
7	4-9	0.12711+j0.27038	17	12-2	j0.25202
8	5-6	0.08205+j0.19207	18	13-3	j0.17615
9	7-8	0.22092+j0.19988	19	14-10	0.01938+j 0.05917
10	8-9	0.17093+j0.34802	20	14-2	0.05403 +j0.22304

TABLE II. PRODUCTION LIMITS AND PROPER COST COEFFICIENT RELEVANT TO EACH GENERATOR

Node N°	pgmin	pgmax	a($/h)	b($/M.W.h)	c($/M.(W^2).h)
10	0.1	0.7	1469	40	0.13
11	0.1	1.2	450	46	0.11
12	0.15	1	1050	40	0.028
13	0.2	1.4	1245	40	0.0354
14	0.1	2.5	1660	36	0.0211

TABLE III. DATA RELATING TO VOLTAGE LIMITS AND COEFFICIENTS OF LOAD SHEDDING COST FUNCTIONS

Node N°	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
V_{min} (pu)	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
V_{max} (pu)	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
Cost (Ct) ($)	100	400	200	550	900	250	450	500	50	*	*	*	*	*

* See Table II.

TABLE IV. LINE RELEVANT DATA

Node N°	**Type**	**Active power generated (pu)**	**Active power consumed (pu)**	**Reactive power consumed (pu)**	**Voltage (pu)**
1	P-Q	0	1.2550	0.1024	-
2	P-Q	0	0.1995	0.0420	-
3	P-Q	0	0	0	-
4	P-Q	0	0.7745	0.4358	-
5	P-Q	0	0.3912	0.1313	-
6	P-Q	0	0.3544	0.1523	-
7	P-Q	0	0.0919	0.0473	-
8	P-Q	0	0.2363	0.1523	-
9	P-Q	0	0.1602	0.0420	-
10	P-V	1.0502	0	0	1.0450
11	P-V	1.0502	0.2941	0.1969	1.0100
12	P-V	1.5753	2.4732	0.4988	1.0700
13	P-V	1.0502	0 .5697	0.3340	1.0900
14	Balance	2.2910	0	0	1.0600

TABLE V. GENETIC ALGORITHM PARAMETERS

Population size	800
Generation number	100
Probability of mutation	0.05
Crossover probability	0.9

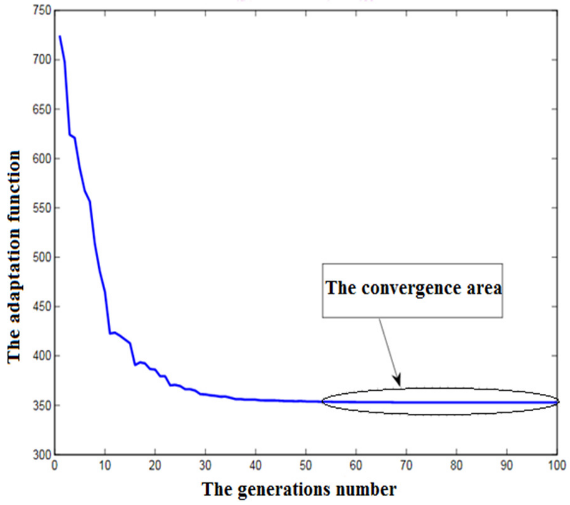


Fig. 5. The convergence curve of the genetic algorithm.

The cost and load-shedding factor profile, relevant to each load node, is depicted in Figure 6. We can see that the load shedding process proves to persist at the level of buses number 1, 3, 4, 6, and 9, displaying different importance degrees, specifically:

- "Low" with regard to nodes number 3 and 4,
- "Medium" with regard to node number 6, and,
- "Totally unloaded" regarding the nodes 1 and 9.

It should be noted that any increase in the load-shedding factor turns out to be closely associated with a load-cost reduction.

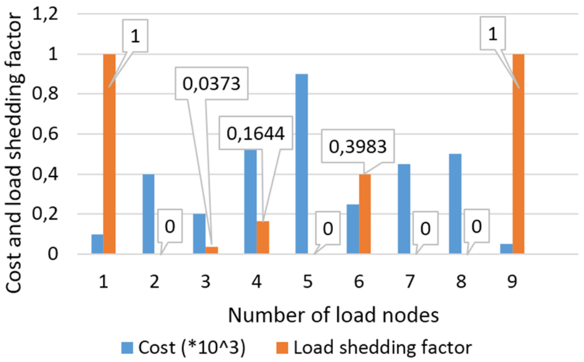


Fig. 6. The cost and load-shedding factor profiles of consumer nodes.

The pre and post load shedding power profiles of the aggregate node loads are highlighted in Figure 7. It can be seen that each load respective power turns out to decrease:

- Slightly at the level of nodes 2, 5, 7 and 8 due to the very high cost,
- totally at the level of nodes 1 and 9 due to the very low cost, and
- around 40% and 60% at the nodes 4 and 6, as cost at the level of node 4 appears to be rather high as compared to the other nodes (1, 3, 6 and 9).

The loaded nodes’ pre and post load shedding voltage profiles are illustrated in Figure 8. It can be seen that there is an increase in the voltages of nodes 4, 5, 6, and 9, which is not within the allowable limit of stability before the load shedding. After load shedding, the voltages of the various load nodes are within their allowable limits, something that maintains the stability and continuity of the electrical network.

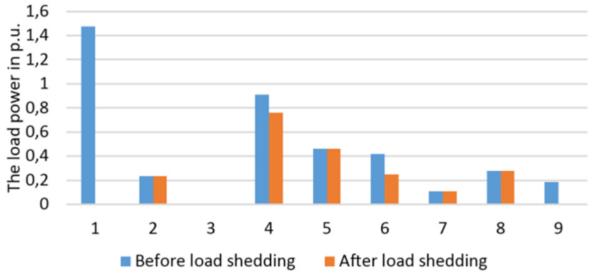


Fig. 7. The consumer nodes' pre and post load shedding power.

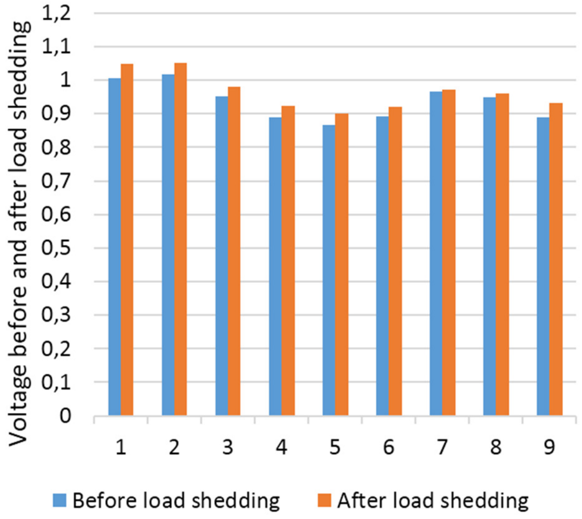


Fig. 8. The consumer nodes' pre and post load shedding voltage.

V. CONCLUSION

Several known registered voltage collapse events indicate that most of the observed power grids operate closely to their stability limits. This issue turns out to be further compounded coupled with the electricity market liberalization. Hence, for the sake of avoiding the electrical grid associated blackouts and for saving more energy, the implementation of an optimal shedding process on a 14-node IEEE test network was undertaken in this paper. In this context, two shedding types have been considered, namely, load and machine shedding. For the shedding type determination, a test has been administered on the load distribution program to set the exact shedding types. In order to avoid the voltage collapse problem, an implementation of the genetic algorithm for the shedding of a 14-node IEEE test network was realized. Based on the obtained simulation results, we were able to highlight the advanced optimization approach's remarkable performance and achieved effectiveness in terms of avoiding the blackouts while minimizing the production cost. The valuable contribution of the present research lies not only in minimizing the production cost, but also in combatting overloading lines, for an effective balance between energy generation and consumption. In view of future development, the proposed work can be expanded to other electrical grids containing renewable energy sources.

REFERENCES

[1] W. Lu, "Le délestage optimal pour la prévention des grandes pannes d'électricité," Ph.D. dissertation, Institut National Polytechnique de Grenoble, Grenoble, France, 2009.

[2] A. Kazemi and B. Badrzadeh, "Modeling and simulation of SVC and TCSC to study their limits on maximum loadability point," *International Journal of Electrical Power & Energy Systems*, vol. 26, no. 8, pp. 619–626, Oct. 2004, https://doi.org/10.1016/j.ijepes.2004.04.008.

[3] T. Van Cutsem, "Voltage instability: phenomena, countermeasures, and analysis methods," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 208–227, Feb. 2000, https://doi.org/10.1109/5.823999.

[4] P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004, https://doi.org/10.1109/TPWRS.2004.825981.

[5] P. Kundur, *Power System Stability and Control*, 1st ed. New York, NY, USA: McGraw-Hill Education, 1994.

[6] M. A. Zdiri, B. Bouzidi, and H. Hadj Abdallah, "Performance investigation of an advanced diagnostic method for SSTPI-fed IM drives under single and multiple open IGBT faults," *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering*, vol. 38, no. 2, pp. 616–641, Jan. 2019, https://doi.org/10.1108/COMPEL-04-2018-0181.

[7] M. A. Zdiri, B. Bouzidi, and H. Hadj Abdallah, "Synthesis of an improved diagnosis method for open IGBT faults dedicated to FSTPI fed IM drives," *Journal of Electrical Systems*, vol. 15, no. 2, pp. 314–330, Jun. 2019.

[8] M. A. Zdiri, B. Bouzidi, M. Ben Ammar, and H. Hadj Abdallah, "SSTPI-IM Reconfiguration and Diagnostic under OCF Appearance Used in PV System," *International Journal of Renewable Energy Research*, vol. 11, no. 1, pp. 20–30, Mar. 2021.

[9] M. A. Zdiri, M. Ben Ammar, B. Bouzidi, R. Abdelhamid, and H. H. Abdallah, "An Advanced Switch Failure Diagnosis Method and Fault Tolerant Strategy in Photovoltaic Boost Converter," *Electric Power Components and Systems*, vol. 48, no. 18, pp. 1932–1944, Nov. 2020, https://doi.org/10.1080/15325008.2021.1909182.

[10] J. J. Ford, H. Bevrani, and G. Ledwich, "Adaptive load shedding and regional protection," *International Journal of Electrical Power & Energy Systems*, vol. 31, no. 10, pp. 611–618, Nov. 2009, https://doi.org/10.1016/j.ijepes.2009.06.004.

[11] C. Moors and T. Van Cutsem, "Determination of optimal load shedding against voltage instability," in *13th Power System Computation Conference (PSCC)*, Trondheim, Norway, Jan. 1999.

[12] T. N. Le, H. A. Quyen, T. T. B. Phan, N. A. Nguyen, and T. P. Trieu, "Select Location for Load Shedding In Power System*," in *2018 4th International Conference on Green Technology and Sustainable Development (GTSD)*, Ho Chi Minh City, Vietnam, Nov. 2018, pp. 13–17, https://doi.org/10.1109/GTSD.2018.8595559.

[13] Y. Tofis, S. Timotheou, and E. Kyriakides, "Minimal Load Shedding Using the Swing Equation," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2466–2467, May 2017, https://doi.org/10.1109/TPWRS.2016.2614886.

[14] J. Wang *et al.*, "Strategy on Interruptible Load Selection for Precise Load Shedding System of Source-Grid-Load Friendly Interaction System," in *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, Changsha, China, Nov. 2019, pp. 1765–1769, https://doi.org/10.1109/EI247390.2019.9062030.

[15] T. Shekari, A. Gholami, F. Aminifar, and M. Sanaye-Pasand, "An Adaptive Wide-Area Load Shedding Scheme Incorporating Power System Real-Time Limitations," *IEEE Systems Journal*, vol. 12, no. 1, pp. 759–767, Mar. 2018, https://doi.org/10.1109/JSYST.2016.2535170.

[16] C. Wang, H. Yu, L. Chai, H. Liu, and B. Zhu, "Emergency Load Shedding Strategy for Microgrids Based on Dueling Deep Q-Learning," *IEEE Access*, vol. 9, pp. 19707–19715, 2021, https://doi.org/10.1109/ACCESS.2021.3055401.

[17] K. Soleimani and J. Mazloum, "Designing a GA-Based Robust Controller For Load Frequency Control (LFC)," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2633–2639, Apr. 2018, https://doi.org/10.48084/etasr.1592.

[18] T. Le and B. L. N. Phung, "Load Shedding in Microgrids with Consideration of Voltage Quality Improvement," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6680–6686, Feb. 2021, https://doi.org/10.48084/etasr.3931.

[19] F. A. Alshammari, G. A. Alshammari, T. Guesmi, A. A. Alzamil, B. M. Alshammari, and A. S. Alshammari, "Local Search-based Non-dominated Sorting Genetic Algorithm for Optimal Design of Multimachine Power System Stabilizers," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7283–7289, Jun. 2021, https://doi.org/10.48084/etasr.4185.

Automated Pavement Distress Detection Using Image Processing Techniques

Iman Hashim Abbas
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
i.altaie1901m@coeng.uobaghdad.edu.iq

Mohammed Qadir Ismael
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
drmohammedismael@coeng.uobaghdad.edu.iq

Abstract-Pavement crack and pothole identification are important tasks in transportation maintenance and road safety. This study offers a novel technique for automatic asphalt pavement crack and pothole detection which is based on image processing. Different types of cracks (transverse, longitudinal, alligator-type, and potholes) can be identified with such techniques. The goal of this research is to evaluate road surface damage by extracting cracks and potholes, categorizing them from images and videos, and comparing the manual and the automated methods. The proposed method was tested on 50 images. The results obtained from image processing showed that the proposed method can detect cracks and potholes and identify their severity levels with a medium validity of 76%. There are two kinds of methods, manual and automated, for distress evaluation that are used to assess pavement condition. A committee of three expert engineers in the maintenance department of the Mayoralty of Baghdad did the manual assessment of a highway in Baghdad city by using a Pavement Condition Index (PCI). The automated method was assessed by processing the videos of the road. By comparing the automated with the manual method, the accuracy percentage for this case study was 88.44%. The suggested method proved to be an encouraging solution for identifying cracks and potholes in asphalt pavements and sorting their severity. This technique can replace manual road damage assessment.

Keywords-pavement distress; AEOP; python code; image processing

I. INTRODUCTION

Road traffic accidents are globally becoming a big issue [1]. According to recent World Health Organization reports, more than 1.25 million people are killed, and 50 million are injured in road traffic accidents annually [2]. Road defects play a significant role in the occurrence of accidents [3]. Road networks are an essential part of our daily life. Pavements deteriorate over time for a variety of causes. The load of the travelling vehicles may cause a pavement to deteriorate. Transverse cracks, longitudinal cracks, block cracks, and alligator cracks are the four types of cracks found in early pavement deterioration. If these early deteriorations are not addressed, potholes emerge, making the road more unsafe. Rehabilitation procedures, such as fixing potholes, will cost 10 to 20 times as much as resealing cracks. Pavement detection and rating are important to maintain the cost of repairing road deterioration low [4]. Traditionally, the cracks in the pavement are manually inspected. Manual inspection is not only time-consuming, inefficient, and prone to errors, but it can also lead to security incidents. Most researches aim to locate and fix pavement cracks as fast as possible [5]. Several studies have been carried out to create a system for automated detection of pavement cracks that solves the restrictions of the manual technique [6].

According to threshold values obtained from histograms of an image, the authors in [7] did an entropy and image dynamic thresholding to automatically divide pavement crack pixels into crack and non-crack pixels. Every binary image achieved by using the threshold value was classified into non-overlapping blocks in which the occurrence of cracks was confirmed using the entropy function and the thresholding operation although the findings are not consistent with every examined image, because of the various lighting circumstances while utilizing a single threshold. Authors in [8] presented another use of image processing on crack identification, focusing and getting an extra specific crack image free of noise. The given approach appears hopeful, but it is time-consuming because it analyzes a one-lane 10m road for a few minutes. Authors in [9] suggested a mixed method for automatic crack identification based on 3D ultra laser-imaging pavement data. Matched filtering was applied to show the cracks, while tensor voting was used to find the main directions of the cracks, and the minimal spanning tree was utilized to identify the crack path. Meanwhile, authors in [10] developed an approach for detecting and segmenting cracks in pavement images by utilizing a steerable filter with two shifted and rotatable tails to achieve satisfactory crack segmentation accuracy. Authors in [11] recently made an automated algorithm-based image processing approach for segmenting and improving pavement crack identification utilizing 3D pavement images.

Although image processing techniques have been used effectively in crack finding, there are yet certain concerns regarding the accuracy and efficiency that must be addressed. Authors in [12] gave an image processing algorithm, special for the rapid evaluation of cracking in the pavement surface. Authors in [13] invented a photogrammetric method to

---

Corresponding author: Iman Hashim Abbas

categorize and enumerate the number of pavement cracks automatically. Authors in [14] proposed a novel automatic crack discovery approach depending on a segment extending for sophisticated pavement movie frames. Interrelated segments are connected to produce a crack by analyzing the relationship between connected domains, and the character of crack trend could be best employed in crack differentiation. Real pavement surface images were utilized to investigate the method's execution, and the results demonstrated that the surface pavement crack could be accurately and automatically classified. Authors in [15] proposed a way to detect and categorize the defects on road pavement surfaces automatically by using an anisotropy measure. Authors in [16] designed a 3-stage method for crack investigation through high-contrast images. The method is based on the detection of cracks in the pavement through the evaluation of curves and the mathematical formation in the presence of a percentage of noise in the processed image. It was depended on the mathematical morphology and evaluation of curvature that detects the crack-like patterns in a noisy environment. Authors in [17] evaluated the achievement of 6 frequently applied segmentation methods. According to their findings, the dynamic optimization-based method surpasses the other algorithms. Authors in [18] state that the goal of incorporating the Shuffled Frog-Leaping Algorithm (SFLA) into the Electromagnetism-like Mechanism (EM) is to improve pavement crack properties in lighting invariant images. Authors in [19] created a pavement distress detection method that helps reduce noise and sharpen the linear characteristics of raw images. Authors in [20] utilized wavelet transform on image processing in order to automatically locate and highlight cracks. The suggested method has restrictions because it is hard to detect cracks due to sharp surface consistencies that create noise in the last image. Because crack pixels have more darkness than the surrounding pavement pixels, some thresholding approach has been used by many studies to locate cracks on the pavement.

II. STUDY OBJECTIVES

The essential goals of this study are:

- To develop new algorithms for automatic detection and classifications of pavement cracks and potholes in images and videos.
- To compare the manual method (visual survey) and the automated method.

III. METHODOLOGY

With the high-performance optical sensor technology and the increasing number of algorithms that rely on computer vision, the development of civil engineering-related applications increased. Most of these applications are dependent on image and video processing techniques by emphasizing the specific properties of the image and increasing the probability of correct detection [21]. Two python codes were used in this study. The first one, which dealt with images, attempted to categorize the types of 2D pavement distress and assess their severity. The second, which dealt with videos, attempted to evaluate sections of road pavement.

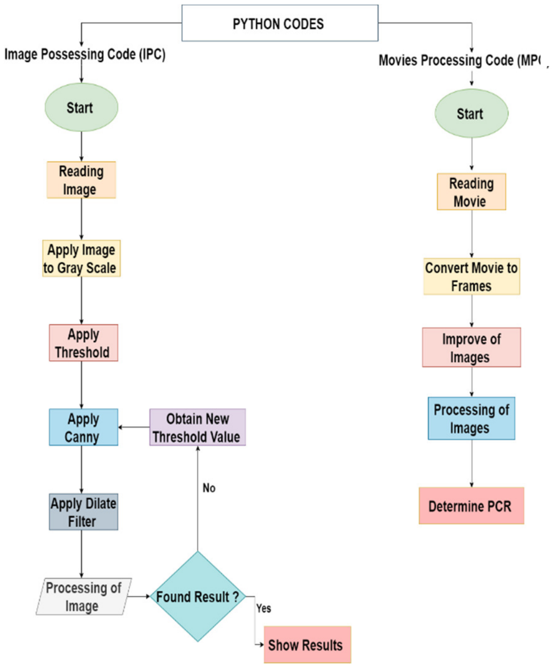


Fig. 1. Work flowchart.

A. Image Possessing Code (IPC)

1) Image Capturing and Reading

Images obtained from the road were used to detect cracks (horizontal, vertical, and alligator) and potholes and to assess their severity. The images were captured from a GoPro 8 hero camera that was positioned with an angle of 90° and a height of 1.10m from the pavement surface as shown in Figures 2 and 3. A captured image must be in jpg format for the code to understand it. The reading of the image is an essential phase of the computer vision flow chart. It enables the framework to get the information.

Fig. 2. The angle between the camera and the surface of the pavement.

2) Apply Blur

Step 2 is used to reduce noise by Gaussian blur, which is the most commonly used smoothing technique to eliminate noises in images. In this technique, an image should be convolved with a Gaussian kernel to produce the smoothed image. It can be considered as a nonuniform low-pass filter that preserves low spatial frequency [22].

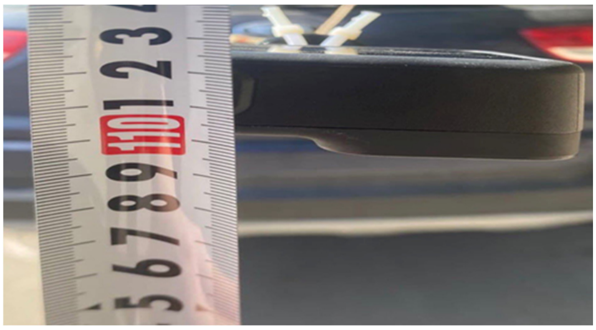


Fig. 3. The height of the camera from the surface of the pavement.

3) Covert to Grayscale and Resize

For many image processing applications, color information doesn't help in identifying important edges or other features. So, the image will be converted from RGB to grayscale to neutralize colors. Also the image is resized to 200×200 pixels.

4) Apply Thresholding

It is a sort of image segmentation in which the pixels of an image are changed to make the image easier to analyze. Thresholding is a famous technique used in image processing algorithms to extract features. Thresholding is most commonly used to identify regions of interest in an image while disregarding areas of no concern [23]. For a given image, this work involves creating a histogram of grayscale values to find the vertices in the image. A threshold is then chosen according to the restricted area confined between two peaks. The failure areas of the pavement within the captured image are often marked by abrupt changes in the grayscale level of adjacent regions.

5) Apply Canny (Edge Detection)

Canny edge detection sets the edges of each pixel in the image to determine the color value after converting the image from grayscale to binary, ensuring that there are only two colors (black and white) in the image. The Canny edge detection algorithm process consists of 4 steps:

- Apply a Gaussian filter to smooth the image and remove the noise because edge detection is prone to noise. A a 5×5 Gaussian filter is used to remove the noise from the image.
- Find the intensity gradients in the image: The image is then smoothed and filtered in both horizontal and vertical dimensions with a Sobel kernel to have the initial derivative in horizontal and vertical direction, G_x and G_y. The edge gradient and direction for each pixel are calculated from:

$$Edge\_Gradient(G) = \sqrt{G_y^2 + G_y^2} \quad (1)$$

$$Angle(\theta) = \tan^{-1}(\frac{G_y}{G_x}) \quad (2)$$

The direction of the gradient is always perpendicular to the edges. It is rounded to one of these different angles: vertical, horizontal, and the two diagonal directions [24].

- Non-maximum suppression: after the identification of the extent and the direction of the gradient, the image is completely inspected to discard any unwanted pixels. Every pixel is assessed to look if it is a regional maximum in its neighbourhood in the gradient's direction [24, 25].
- Hysteresis Thresholding: this step decides if the edges are real or not. This requires two thresholding values, $minVal$ and $maxVal$. Pixels with intensity gradient of the edges greater than $maxVal$ are guaranteed to be edges, while those with an intensity gradient less than $minVal$ are guaranteed to be non-edges and are therefore eliminated. The ones located between the two criteria are sorted as edges or non-edges according to their connection. If they are linked to "sure-edge" pixels, they are regarded to be edges. If not, they will be removed as well [24].

6) Dilate Filter Application

A dilate filter is applied to fill the empty holes. A 5×5 kernel size is used. Dilation is applied to the binary images. The primary action of dilation on a binary image is that the boundaries of regions of foreground pixels are continuously increased (for example, white pixels). As a result, foreground pixel areas grow in size, while gaps within those regions shrink. Image dilation is used over erosion because erosion contracts the object. Since the noise is gone, it won't return, but the object area increases. It is useful in connecting the broken parts of an object back together.

7) Image Processing

The processing on the images consists of the following steps:

- Detect all contours in the image (finding all objects): Contours can be defined as a curve linking all continuous points (along the border) of matching color or intensity. The contours are important tools for object recognition and shape evaluation [26]. A find-contours function retrieves all the contours in the image that it can detect. Contours can appear in an image in a variety of ways. Some might be grouped in other contours etc. This technique is used to simplify locating the contours of interest and to understand the hierarchy in which the contours are nested [26].
- Calculate contour area: Image moments assist in calculating some properties usually selected to have attractive features. It is used to define objects and to detect uncomplicated features of the image like area, intensity, centroid, orientation [26], etc. Image moment can be defined as a weighted average (moment) of the image pixels' intensities
- Filter areas larger than the selected minimum crack area: Minimum crack size can be variant using the software UI, and changes will be seen in real-time for crack detection.
- Calculate arc length: arc length is used to calculate the perimeter of the contour. This is used to know more about the crack length and later to decide its type.
- Get the arc borders edges that approach a curve or a polygon with another curve such that the distance between them is less or equal to the particular accuracy, then calculate the number of curves/polygon endpoints to know how many endpoints are in the crack. For example, longitudinal and transverse have two endpoints and the alligator type has a minimum of 5 endpoints.

- Get the arc type, based on these conditions:

  If width < 100 and h > w × 2, it is a longitudinal.

  If w > h × 2 and h < 100, it is a transverse.

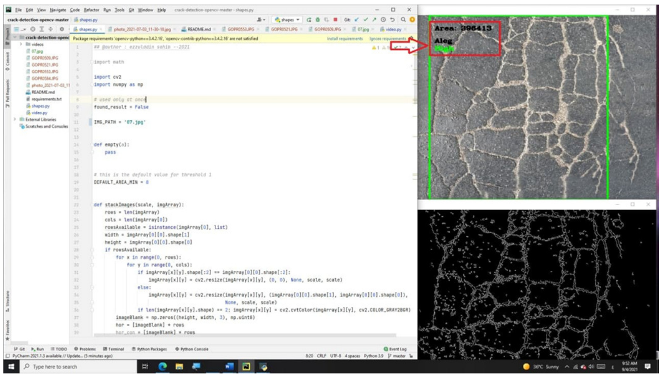
  If the endpoints of the edge are larger than 5 and the shape is adequate, it is an alligator crack.

  Else, it is a pothole.

- Draw contours on the screen.
- Draw the arc path to detect crack areas and their width.
- Get the severity based on user-entered values for low, medium, and high severity.
- Draw result texts and border around the crack.

8) *Show the Results*

An example can be seen in Figure 4.



```
## @author : ezzuldin sahib --2021

import math

import cv2
import numpy as np

# used only at once
found_result = False

IMG_PATH = '07.jpg'


def empty(a):
    pass


# this is the default value for threshold 1
DEFAULT_AREA_MIN = 8


def stackImages(scale, imgArray):
    rows = len(imgArray)
    cols = len(imgArray[0])
    rowsAvailable = isinstance(imgArray[0], list)
    width = imgArray[0][0].shape[1]
    height = imgArray[0][0].shape[0]
    if rowsAvailable:
        for x in range(0, rows):
            for y in range(0, cols):
                if imgArray[x][y].shape[:2] == imgArray[0][0].shape[:2]:
                    imgArray[x][y] = cv2.resize(imgArray[x][y], (0, 0), None, scale, scale)
                else:
                    imgArray[x][y] = cv2.resize(imgArray[x][y], (imgArray[0][0].shape[1], imgArray[0][0].shape[0]),
                                                None, scale, scale)
                if len(imgArray[x][y].shape) == 2: imgArray[x][y] = cv2.cvtColor(imgArray[x][y], cv2.COLOR_GRAY2BGR)
        imageBlank = np.zeros((height, width, 3), np.uint8)
        hor = [imageBlank] * rows
        hor_con = [imageBlank] * rows
```

Fig. 4. The obtained result.

B. *Movie Processing Code (MPC)*

1) *Import Video*

The video is imported from the computer as an MP4 type. Video reader instructions are used by the code to read the movie.

2) *Calculate the Total Number of Frames*

The total number of frames in the movie is calculated by using a high-speed (240 fps) camera.

3) *Select the Frames that will Be Processed*

For surveying speed of 80km/hr or 22m/s and w=3m, L=2.20m frame dimensions, 1 frame must be selected from every 22 frames. For example, if the period of video is 1s (240 frames), the code will select 11 frames (1, 23, 45, 67, 89, 111, 133, 155, 177, 199, 221). These numbers apply only to the type of camera Hero 8 Black because they are calculated based on the characteristics of this camera.

4) *Change RGB Frames to Greyscale and then to Binary*

This step is essential in enabling the code to work with only two colors.

5) *Regulate the Frames to Matrices*

6) *Measure the Standard Deviation of Each Frame by Utilizing STD Orders.*

7) *Regulate the Frames to to Have the Same Pixel Size*

8) *Filtering*

To decrease the frame noise, the image should be filtered. In MPC, and after experimentation with many kinds of filters, the most useful filter that can be utilized is the Gaussian filter. This kind of filter is dealing with almost all noises in the frame.

9) *Color Conversion*

In order to differentiate undamaged pavement from the impairment area in the same image, the colors are converted to conclude the involved areas in white and black for the normal pavement.

10) *Edge Detection*

Edge detection is conducted by the Canny algorithm.

11) *Total Edge Values*

Total edge values are found for every column and every row in the frame.

12) *Finding the Distress Kind*

The values of total edges were compared to find the distress kind in the frame. The damage in the pavement and its severity are found by measuring the white to black ratio in the frame.

13) *Determine the PCR of the Pavement*

IV. RESULTS AND DISCUSSION

The traditional method of assessment of pavement condition is to walk or drive down the road and collect the data manually. This way of road pavement surface detection depends on the experience level of the employees and it is time-consuming, hazardous, and subjective. Therefore, an effort has been made to fully automate the data collection process. To examine the possibility of the program to identify the types of failures and their severity from different images, 50 failure images were obtained and processed. The program correctly identified 38 of them. In the remaining images, the program either failed to detect the crack, identified it incorrectly, or gave a wrong severity of the crack. So, the examined images have an accuracy percentage of about 76%. This medium accuracy may be caused by environmental factors like shadows, dust, or weather conditions. The accuracy percentage of the distress detection may be enhanced by taking clearer images and having criteria for proper capture of the images with a steady source of light.

The proposed Automated Evaluation Of Pavement (AEOP) method detects the defects of roads automatically and classifies the most common distresses (fatigue, transversal cracks, longitudinal cracks, and potholes) by using computer vision techniques and image processing. The obtained classification results contain distress name/class and severity level as demonstrated in Figures 7-10. The manual method of assessment was conducted by the specialized department that is responsible for the evaluation and maintenance of city roads. The engineers did the manual assessment by using the

Pavement Condition Index (PCI) to quantify the condition of the road based on a scale from 0 to 100. They measured the Pavement's integrity and surface condition and rated the sections as very poor, poor, fair, satisfactory, good, or excellent according to [27]. By comparing the automated with the manual survey for the 8 sections, the percentage of accuracy results of distress detection for this case study is 88.44%.

TABLE I. PERCENTAGE DIFFERENCES BETWEEN PCI AND AEOP

No.	Sections	PCI	AEOP
1	Sec 1	68%	75.23%
2	Sec 2	69%	78.03%
3	Sec 3	71%	75.45%
4	Sec 4	72%	80.26%
5	Sec 5	75%	81.46%
6	Sec 6	60%	76.25%
7	Sec 7	65%	80.02%
8	Sec 8	70%	75.34%

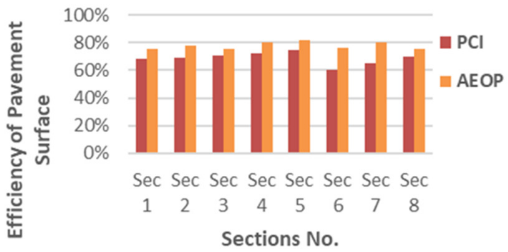


Fig. 5. Percentage differences between PCI and AEOP surveying.

In Figure 5, we can see that the largest differences between the two methods occurred in sections 7 and 6. The smallest difference between the two methods occurred in section 3. When the two methods' results were compared, as shown in Table I, the results are nearly similar with a small variation. The discrepancy in the results can be attributed to several factors, including the type of pavement evaluation method, or maybe due to other issues like environmental factors such as stains, lane marks of the road, adjacent vehicles, shadows of side road trees, or texture differences among different pavement surfaces that affect the results of the automated detection of the crack. Among the most crucial comparisons to make is the difference in the time survey between the two techniques. The survey committee estimated the time required for the field survey and provided a report on the pavement's efficiency. The differences between the manual method and the automated approach are shown in Table II.

TABLE II. TIME CONSUMPTION

No.	Time (min)		
	Section No.	PCI	AEOP
1	Section 1	210	20
2	Section 2	152	24
3	Section 3	200	25
4	Section 4	120	29
5	Section 5	245	27
6	Section 6	150	30
7	Section 7	118	25
8	Section 8	200	21

The time required to do a manual survey vs the AEOP approach is significantly different. The AEOP took less time than the manual method, as shown in Figure 6. The time needed to finish the manual assessment of the 8 sections of the highway by the engineers of the maintenance department of the mayoralty of Baghdad was about 23hr. and the time needed for finishing the automated way was about 3hr.

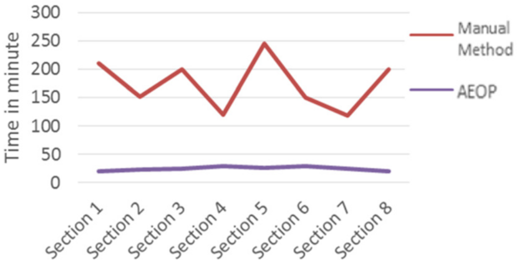
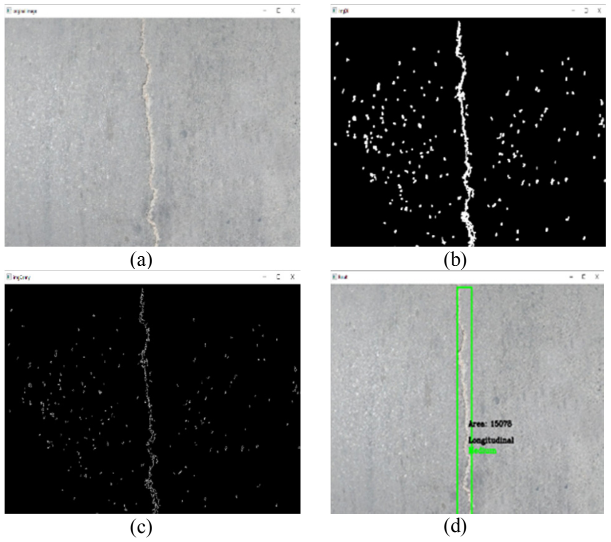
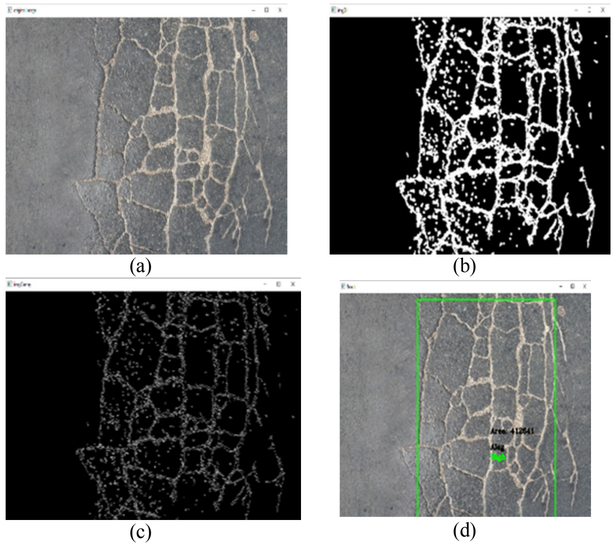


Fig. 6. Required time for completing each method in each section.



(a) (b) (c) (d)

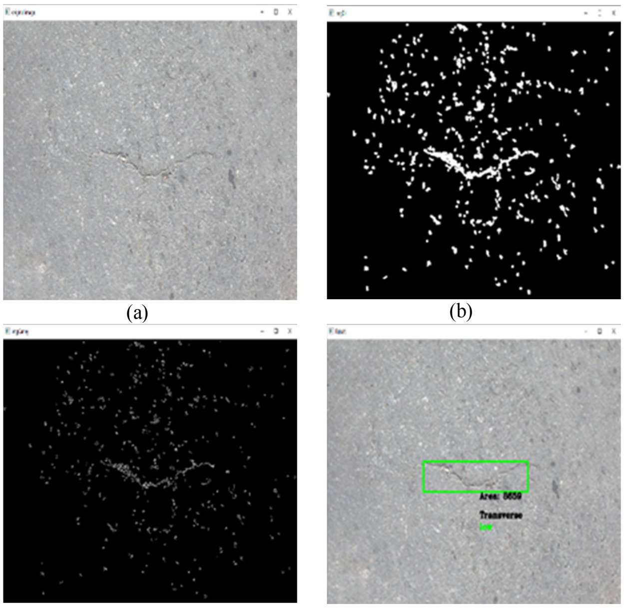
Fig. 7. Longitudinal crack: (a) Original image, (b) binary image, c) after applying Canny threshold, (d) result.



(a) (b) (c) (d)

Fig. 8. Alligator crack: (a) Original image, (b) binary image, (c) after applying Canny threshold, (d) result.

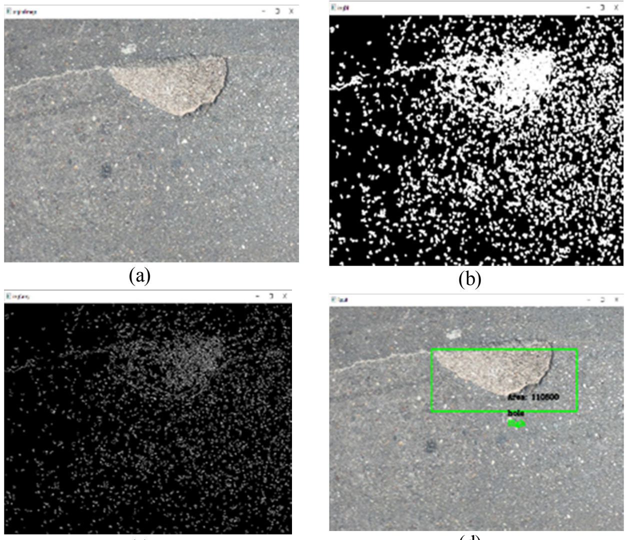
The cost of surveying is also an important issue to be evaluated. The manual method requires a committee with at least 3 experienced road engineers or technicians who work in road maintenance, and the use of these specialists means an augmented cost that will be spent on the evaluation of the road section. In addition, the evaluation of all the road sections requires a long time, and increasing time also means increased cost of the assessment of the whole road. On the other hand, the evaluation of the performance of the road's pavement using the AEOP method doesn't need committees or many workers. It requires just one person who knows how to use the code and a vehicle supplied with a camera.



(a) (b)

(c) (d)

Fig. 9. Transverse crack: (a) Original image, (b) binary image, c) after applying Canny threshold, (d) result.



(a) (b)

(c) (d)

Fig. 10. Pothole: (a) Original image, (b) binary image, c) after applying Canny threshold, (d) result.

V. CONCLUSIONS

The presented results support the following conclusions:

- It is useful to utilize the image processing approach for identifying pavement failures. The presented algorithm has a lot of promise in terms of automated crack and pothole detection, since it can provide quick, nearly precise, and cost-effective results, which is important in pavement management systems.
- The suggested study is an encouraging solution for identifying cracks and potholes in asphalt pavements and sorting their severity by automated survey. By comparing the automated with the manual survey, the percentage of accuracy results of distress detection for this case study is 88.44%.
- There is a very large difference in the time needed to conduct the two methods. The AEOP required about 87% less than the manual method.
- Also, the automated method is much safer because it is generally conducted at prevailing traffic, it is easy to operate, it does not require committees, just one person to do it, and so it costs less than the manual method.
- The automated technique can replace the traditional road measurements.

Improving the automated crack detection can be done by using a vehicle with more developed equipment like multiple sensors, or a lighted camera to overcome the shadowing problem. There is a need to develop the code through cooperation with artificial software experts who can add new algorithms and thresholds to the current code in order to increase its performance and make it more efficient in exploring more types of failures on paved roads.

ACKNOWLEDGEMENT

The authors would like to thank the engineers of the maintenance department of the Mayoralty of Baghdad for their cooperation.

REFERENCES

[1] I. J. Mrema and M. A. Dida, "A Survey of Road Accident Reporting and Driver's Behavior Awareness Systems: The Case of Tanzania," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6009–6015, Aug. 2020, https://doi.org/10.48084/etasr.3449.

[2] M. Touahmia, "Identification of Risk Factors Influencing Road Traffic Accidents," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2417–2421, Feb. 2018, https://doi.org/10.48084/etasr.1615.

[3] A. Detho, S. R. Samo, K. C. Mukwana, K. A. Samo, and A. A. Siyal, "Evaluation of Road Traffic Accidents (RTAs) on Hyderabad Karachi M-9 Motorway Section," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2875–2878, Jun. 2018, https://doi.org/10.48084/etasr.1920.

[4] D. Akhila and V. Preeja, "A Novel Technique for Automatic Road Distress Detection and Analysis," *International Journal of Computer Applications*, vol. 101, no. 10, pp. 18–23, 2014.

[5] C. Chen *et al.*, "Automatic Pavement Crack Detection Based on Image Recognition," in *International Conference on Smart Infrastructure and Construction*, Cambridge, UK, Jul. 2019, pp. 361–369, https://doi.org/10.1680/icsic.64669.361.

[6] T. S. Tran, V. P. Tran, H. J. Lee, J. M. Flores, and V. P. Le, "A two-step sequential automated crack detection and severity classification process for asphalt pavements," *International Journal of Pavement Engineering*, Oct. 2020, https://doi.org/10.1080/10298436.2020.1836561.

[7] H. Oliveira and P. L. Correia, "Automatic road crack segmentation using entropy and image dynamic thresholding," in *17th European Signal Processing Conference*, Glasgow, UK, Aug. 2009, pp. 622–626.

[8] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," *Pattern Recognition Letters*, vol. 33, no. 3, pp. 227–238, Feb. 2012, https://doi.org/10.1016/j.patrec.2011.11.004.

[9] G. Sollazzo, K. C. P. Wang, G. Bosurgi, and J. Q. Li, "Hybrid Procedure for Automated Detection of Cracking with 3D Pavement Data," *Journal of Computing in Civil Engineering*, vol. 30, no. 6, Nov. 2016, Art. no. 04016032, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000597.

[10] S. Li, Y. Cao, and H. Cai, "Automatic Pavement-Crack Detection and Segmentation Based on Steerable Matched Filtering and an Active Contour Model," *Journal of Computing in Civil Engineering*, vol. 31, no. 5, Sep. 2017, Art. no. 04017045, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000695.

[11] B. Li, K. C. P. Wang, A. Zhang, Y. Fei, and G. Sollazzo, "Automatic Segmentation and Enhancement of Pavement Cracks Based on 3D Pavement Images," *Journal of Advanced Transportation*, vol. 2019, Feb. 2019, Art. no. e1813763, https://doi.org/10.1155/2019/1813763.

[12] Y. Huang and B. Xu, "Automatic inspection of pavement cracking distress," *Journal of Electronic Imaging*, vol. 15, no. 1, Jan. 2006, Art. no. 013017, https://doi.org/10.1117/1.2177650.

[13] M. Mustaffar, T. C. Ling, and O. C. Puan, "Automated pavement imaging program (APIP) for pavement cracks classification and quantification-a photogrammetric approach," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, no. B4, pp. 367–372, 2008.

[14] F. Liu, G. Xu, Y. Yang, X. Niu, and Y. Pan, "Novel Approach to Pavement Cracking Automatic Detection Based on Segment Extending," in *International Symposium on Knowledge Acquisition and Modeling*, Wuhan, China, Dec. 2008, pp. 610–614, https://doi.org/10.1109/KAM.2008.29.

[15] T. S. Nguyen, M. Avila, and S. Begot, "Automatic detection and classification of defect on road pavement using anisotropy measure," in *17th European Signal Processing Conference*, Glasgow, UK, Aug. 2009, pp. 617–621.

[16] G. Pascale and A. Lolli, "Crack assessment in marble sculptures using ultrasonic measurements: Laboratory tests and application on the statue of David by Michelangelo," *Journal of Cultural Heritage*, vol. 16, no. 6, pp. 813–821, Nov. 2015, https://doi.org/10.1016/j.culher.2015.02.005.

[17] Y.-C. Tsai, V. Kaul, and R. M. Mersereau, "Critical Assessment of Pavement Distress Segmentation Methods," *Journal of Transportation Engineering*, vol. 136, no. 1, pp. 11–19, Jan. 2010, https://doi.org/10.1061/(ASCE)TE.1943-5436.0000051.

[18] J. Jiang, "Crack Enhancement Algorithm Based on Improved EM," *The Journal of Information and Computational Science*, vol. 12, pp. 1037–1043, 2015, https://doi.org/10.12733/JICS20105443.

[19] M. Gavilan *et al.*, "Adaptive Road Crack Detection System by Pavement Classification," *Sensors*, vol. 11, no. 10, pp. 9628–9657, Oct. 2011, https://doi.org/10.3390/s111009628.

[20] P. Subirats, J. Dumoulin, V. Legeay, and D. Barba, "Automation of Pavement Surface Crack Detection using the Continuous Wavelet Transform," in *International Conference on Image Processing*, Atlanta, GA, USA, Oct. 2006, pp. 3037–3040, https://doi.org/10.1109/ICIP.2006.313007.

[21] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, vol. 25, no. 3, pp. 507–515, Aug. 2011, https://doi.org/10.1016/j.aei.2011.01.002.

[22] R. Chandel and G. Gupta, "Image Filtering Algorithms and Techniques: A Review," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 10, pp. 198–202, 2013.

[23] N. M. Zaitoun and M. J. Aqel, "Survey on Image Segmentation Techniques," *Procedia Computer Science*, vol. 65, pp. 797–806, Jan. 2015, https://doi.org/10.1016/j.procs.2015.09.027.

[24] G. Jie and L. Ning, "An Improved Adaptive Threshold Canny Edge Detection Algorithm," in *International Conference on Computer Science and Electronics Engineering*, Hangzhou, China, Mar. 2012, vol. 1, pp. 164–168, https://doi.org/10.1109/ICCSEE.2012.154.

[25] R. R. Chavan, S. A. Chavan, G. D. Dokhe, M. B. Wagh, and A. S. Vaidya, "Quality Control of PCB using Image Processing," *International Journal of Computer Applications*, vol. 141, no. 5, pp. 28–32, 2016.

[26] R. S. Choras, "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems," *International journal of biology and biomedical engineering*, vol. 1, no. 1, pp. 6–16, 2007.

[27] *ASTM D6433-07(2007), Standard practice for roads and parking lots pavement condition index surveys*. West Conshohocken, PA, USA: ASTM International, 2007.

Impact Resistance of Bendable Concrete Reinforced with Grids and Containing PVA Solution

Shrooq Abd Al Kareem
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
s.shetan1901M@coeng.uobaghdad.edu.iq

Ikram Faraoun Ahmed
Department of Civil Engineering
College of Engineering
University of Baghdad
Baghdad, Iraq
ikram.faroun@coeng.uobaghdad.edu.iq

Abstract-The development of new building materials, able of absorbing more energy is an active research area. Engineering Cementitious Composite (ECC) is a class of super-elastic fiber-reinforced cement composites characterized by high ductility and tight crack width control. The use of bendable concrete produced from Portland Limestone Cement (PLC) may lead to an interest in new concrete mixes. Impact results of bendable concrete reinforced with steel mesh and polymer fibers will provide data for the use of this concrete in areas subject to impact loading. The experimental part consisted of compressive strength and impact resistance tests along with a result comparison with unreinforced concrete. Concrete samples, with dimensions of 100×100×100mm (cubes), and 500×500×50mm (slabs), were poured and were treated at ages of 28, 56, and 90 days. The compressive strength increased by 36.11%, 45.5%, and 52.4% respectively, whereas the impact resistance for samples reinforced with steel mesh and polypropylene fibers gave superior results to the conventional mixes.

Keywords-ECC concrete; impact resistance; polypropelene fiber

I. INTRODUCTION

In comparison with other building materials, concrete is more practical, durable, and economical [1]. Concrete is widely used as a building material due to its high durability, easy accessibility, and low cost [2] and is commonly used in almost every kind of construction [3]. Cracks are inevitable during the life of a concrete structure. Structures exposed to the external environment are more susceptible to cracking because they are affected by shrinkage or expansion in weight and drying in addition to other environmental conditions along with the overloading factor. These cracks affect the strength of the structures by weakening them and the mechanical properties and its durability are reduced as these cracks create a path for the penetration of harmful factors into the core of the structures [4]. Attempts have been made to reduce cracks and improve the tensile properties of concrete members by using traditionally reinforced steel bars and also through the application of restraining techniques. Although both methods provide more tensile strength to the concrete members, they do not increase the tensile strength of the concrete itself. In normal concrete and similar brittle materials, fine structural cracks develop even before loading due to drying shrinkage or other causes of volume changes. Upon loading, these micro cracks spread and open due to the stress concentration effect [5]. To give the concrete flexibility we have to modify the traditional material. This type of concrete is softer than conventional concrete and is known as the Engineered Cementitious Composite (ECC). Developed at the end of the twentieth century, the ECC demonstrates the unique properties of high strength concrete, including excellent stress-hardening properties, multiple cracking properties, and fiber bridging properties which in turn increase the ductility of a structure.

ECC, also called ultra-hard cementious composite, flexible concrete, or conflexpave, is a matrix-based composite material reinforced with short fibers with a maximum volume fraction of 2.0%. ECC blends are generally developed by adding Polyvinyl Alcohol (PVA) solutions, Polypropylene (PP), and Polyethylene (PE) fibers. The introduction of fibers aims to reduce the damage in the concrete structure that is exposed to seismic and impact loads [6]. ECC concrete, is an easy-to-form mortar-based composite reinforced with specially selected short polymer fibers. ECC acts more like ductile metal than brittle glass which leads to a wide variety of applications. The tensile stress capacity of ECC can reach 3-5%, which is much higher than the 0.01% of ordinary concrete. The compressive strength of ECC is similar to that of normal to high strength concrete. Ordinary concrete is brittle by nature while ECC is ductile. Due to this property, it has wide applications and future scopes in various fields [7]. The associated high fracture toughness and controlled slit width (typically less than 100μm) make ECC an ideal material for improving the serviceability and durability of infrastructures [8].

II. MATERIAL CHARACTERIZATION

A. Cement

The cement employed in the current research was Ordinary Portland Cement (OPC) and Portland Limestone Cement (IL). The IL was produced by LAFARGE and conforms to European Standard EN-197-1 & European Standard no 196-2-Type IL [9]. Tables I-IV show the properties and specifications of the two types of cement.

Corresponding author: Shrooq Abd Al Kareem

TABLE I. PHYSICAL TESTS OF IL ACCORDING TO EUROPEAN STANDARD EN-197-1 FOR GRADE (42.5)

Test	Test result	Specification limits	Compliance
Finance (blain) cm^2/gm	5105	At least	Matching
Initial setting time (min)	90 min	At least 45min	Matching
Compressive strength (MPa)			
2days curing	8	At least (10-20)	Not matching
28 days curing	34	At least 42.5	Not matching

TABLE II. CHEMICAL ANALYSIS OF IL ACCORDING TO EUROPEAN STANDARD NO 196-2

Oxides and phases	Karasta test results (%)	Specification limits	Compliance
CaO	62.22	-	-
SiO_2	17.18	-	
Al_2O_3	4.19	-	
Fe_2O_3	2.91	-	
MgO	1.95	Not more than 5%	Not matching
SO_3	2.44	Not more than 2.5%	Not matching
LOI	6.21	Not more than 4%	Not matching
Total	98.43	-	
CI-	0.011	-	
C_3A	6.18	Less than 0.1%	
I.R	0.47	Not more than 1.5%	

TABLE III. CHEMICAL ANALYSIS OF OPC

Oxide composition	Test results	Limits of IQS No.5 /2019 for OPC
CaO	62.32	-
SiO_2	21. 34	-
Al_2O_3	4.93	-
Fe_2O_3	5.43	-
MgO	2.12	< 5%
SO_3	2.35	< 2.8%
LOI	1.72	< 4%
LSF	0.86	0.66 – 1.02%
IR	0.83	< 1.5%
Main Compounds (Bogue's equation)		
C_3S	76.28	-
C_2s	3.73	-
C_3A	6.72	-
C_4AF	16.5	-

TABLE IV. PHYSICAL TESTS OF OPC

Physical properties	Test results	Requirements of IQS No.5/2019 for OPC
Specific surface area (Blaine method) m^2/Kg	376	> 250
Initial setting (h)	2:45	> 45min
Final setting (h)	4:20	≤ 10h
Soundness (autoclave method) %	0.12	< 0.8
Compressional strength (MPa) at 2 days	20	≥ 10
Compressional strength (MPa) at 28 days	37	≥ 42.5

B. Sand

Ekhaider natural sand was employed as fine aggregates in this study. Its physical and chemical properties are shown in Table V. Sieve analysis shows that the sand lies in Zone 2 from the tests that were carried according to the requirements of [10] as shown in Table VI.

TABLE V. PHYSICAL AND CHEMICAL TESTS OF FINE AGGREGATES

Property	Test result	I.Q.S.45: 1984 limits
Specific gravity	2.6	-
Absorption %	0.72	-
Density (kg/m^3)	1580	-
SO_3	0.2%	0.50% (Max)

TABLE VI. GRADING TESTS OF FINE AGGREGATES

Sieve no.	Passing %	Limits of Iraqi specification no.45/1984 zone 2
10mm	100	100
4.75mm	93.3	100-90
2.36 mm	77.7	100-75
1.18 mm	66.6	90-55
600 μm	54.4	59-35
300 μm	26.3	30-8
150 μm	3.1	10-0

C. Silica Fume

Silica Fume (SF) (condensed micro-silica) with an activity index of 121% conforming to ASTM C1240-15 [11] was used. The technical data for the SF are shown in Tables VII and VIII.

TABLE VII. PHYSICAL TESTS OF SILICA FUME

Physical properties	Silica fume	Limit of specification requirements ASTM C-1240
Percent Retained on 45μm (No. 325) sieve, Max.%	7	< 10
Accelerated pozzolanic strength activity Index with Portland cement at 7 days, Min. percent of control	121	> 105
Specific surface, Min (m^2/g)	15	> 15

TABLE VIII. CHEMICAL ANALYSIS OF SILICA FUME

	SiO_2 %	Loss on ignition %	Moisture content %
Sample	92.84	1.59	0.33
Chemical requirement according to ASTM-C1240-2012	85% min	6% max	3%,max

D. Superplasticizer

A third-generation superplasticizer (Sika Viscocrete -5930) was employed, which conforms to the specifications of ASTM-C494 Types G and F [12]. Superplasticizers are employed to attain sever decrease of water, enhanced flowability, and optimal cohesion. The decrease in water-cement (w/c) ratio means that the cement paste permeability reduces noticeably, hence superplasticizers can be effectively employed to enhance the properties of concrete and prevent specific defects such as honeycombing.

E. Water

Tap water was used conforming [13]

F. Polypropylene Fibers

PP fibers absorb water and resist alkalis, chemicals, and chloride. Table IX shows the properties of PP fibers.

TABLE IX. PROPERTIES OF PP FIBERS

Length	12mm
Diameter	0.032
Density Kg/m^3	910
Tensile properties	600 - 700

G. Polyvinyl Alcohol (PVA) Solution

PVA is a water-soluble and biodegradable artificial polymer. PVA is an excellent adhesive with superior bonding strength and film forming and emulsifying properties. It has excellent adhesion to both hydrophilic and hydrophobic materials. The solution was prepared to employ closely 80g of PVC in 2lt of boiled water until it disappeared completely.

H. Reinforcement Grids

Two kinds of grid reinforcement were employed, steel and polymer with 12.7mm square opening size and 0.2mm diameter.

III. EXPERIMENTAL WORK

A. Mixures

Two major categories of ECC mixtures have been designed and chosen. The first mixture was composed of OPC, sand, SF, PVA solution, and the optimum dosage of superplasticizer of about 1.3% which is within the acceptable range which is between 0.2 and 1.5% by weight of cement. The mix was reinforced with PP fibers. The second mixture was composed of the same components except that it was produced from IL cement. The mix designs are described in detail in Table X.

TABLE X. MIX PROPORTIONS

Type		Ref.	ECC
Cement (kg/m^3)	OPC	356	356
	IL		
Sand (kg/m^3)	Zone 2	320	320
SF (kg/m^3)	-	285	285
VF (kg/m^3)	PPM	-	18
Water (kg/m^3)	-	288	288
S.P (kg/m^3)	5930	4.6	4.6
Acetate (kg/m^3)	PVA	-	3.56

B. Molds Preparation

For compressive strength tests, 100×100×100mm cubes were used, and for the impact resistance test, 500×500×50mm molds were used.

IV. RESULTS AND DISCUSSION

A. Compressive Strength

The compressive strength test was conducted according to the British standard B.S.1881: part116 [14] by using 3 cubes specimens with dimensions of 100×100×100mm and with a testing machine of 2000KN capacity with loading rate of 2.5MPa/s, at a position perpendicular to the direction of casting (Figure 1-2). The compressive strength test was conducted at the ages of 28, 56, and 90 days.

It can be noted that the samples reinforced with PP fibers have better results than the reference mixtures. The increase in compressive strength was 36.11%, 45.5%, and 52.4% at 28, 56, and 90 days of water curing respectively. The difference between the two types of cement is the change in the chemical compounds of the cement The interpretation of the compressive strength test results is divided into three stages:

Fig. 1. The typical failure behavior of reference concrete specimens.

Fig. 2. The typical failure behavior of concrete specimens with 2% PP fibers under compression.

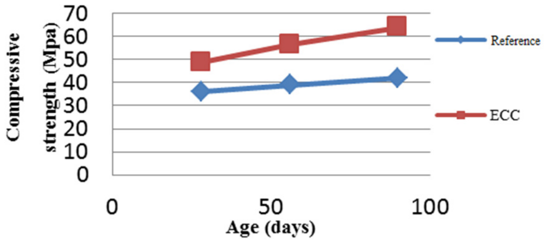


Fig. 3. Compressive strength vs curing age of mixes containing OPC.

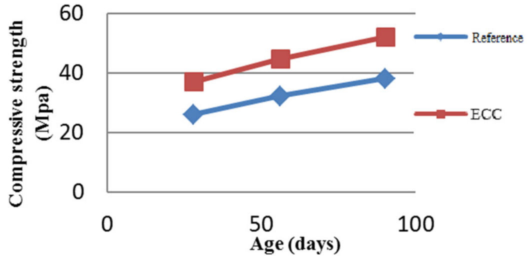


Fig. 4. Compressive strength vs curing age of mixes containing IL.

At first, there is the addition of PP fibers. These fibers were combined with the cement paste and enhanced the ability of micro cracks to spread without breaking or pulling the fibers and therefore reduced the cracks in the concrete.

The second stage is the addition of small amounts of water-soluble polymers (PVA solution) that enhance the bonding strength and durability of cementing materials.

In the third stage, and keeping in mind the recent focus on the sustainable influence on unstable constructions, the environmentally friendly SF decreases carbon dioxide emissions by decreasing the amount of cement used. SF is a material with a volume twice smaller than cement, which can easily fill the voids between cement particles, and is also a highly active pozzolanic material that reacts easily with $Ca(OH)_2$ and water to form secondary CSH through a pozzolanic reaction whereby the generated CSH fills the capillary voids and a denser microstructure and high compressive strength can be achieved (Figure 2). Time-dependent deformation behavior of bendable concrete produced by IL cement was studied in [15]. Two types of IL cement were used and PP fibers and PVA solution were utilized. The best results obtained were for mixes containing PP fibers and PVA solution [15].

B. Impact Resistance

The device used to measure impact resistance contains three parts that can be described as follows:

- The main support part: It consists of an iron bracket that is strong enough to support the examination device during the examination and withstand the impact. It also contains another support on which the examination form is placed, which is of angle iron and is well fixed to prevent the movement of the form during the examination.
- The iron structure for dropping the block: It is a cylindrical tube with an inner diameter of 10.5cm. It is fixed with supports that prevent it from moving during the examination. This tube contains an opening of 1.5m height.
- The falling block: It is an iron ball, weighing 3.4kg with a diameter of 9.5cm. This mass is thrown from the height of 1.5m on the model several times.

The model is placed on the supporting iron structure, and then the iron block is dropped from the specified height of 1.5m on the model. The number of blows at which the first crack appears and the number of blows to cause failure were recoded. The average of three 50×50×5cm specimens was considered. The test was carried out at the ages of 28, 56, and 90 days according to ACI C-544 [16]. The results are shown in Table XI.

TABLE XI. IMPACT RESISTANCE TEST RESULTS

Mix	**Reinforcement**	**Impact resistance**					
		No. of blows to cause:					
		1st crack at day			**Failure at day**		
		28	**56**	**90**	**28**	**56**	**90**
Reference (OPC)	Plain	2	4	6	3	5	6
ECC OPC	Steel	6	8	9	67	93	118
	Polymer	5	7	9	32	50	67
Reference (IL)	Plain	2	4	5	3	5	6
ECC IL	Steel	5	7	8	50	53	55
	Polymer	5	7	8	23	31	38

Dimensions: 50×50×5cm

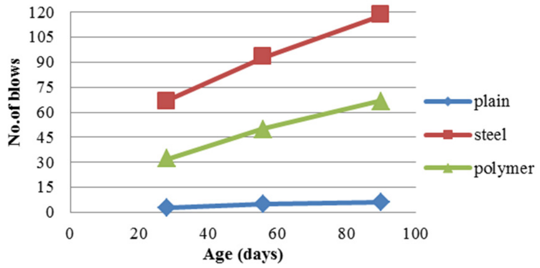


Fig. 5. Impact resistance with curing age of reference and ECC concrete mixes containing OPC.

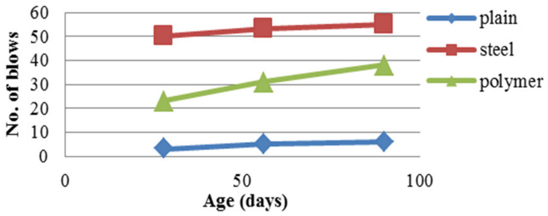


Fig. 6. Impact resistance with curing age of reference and ECC concrete mixes containing IL.

According to the test results shown in Table XI and Figures 5 and 6, the samples reinforced with steel mesh and PP fibers have the greatest results with 67, 93, and 118 blows at 28, 56, and 90 days of water curing respectively, unlike the conventional mix which showed the lowest results. The mix reinforced with a plastic mesh had mid-values of 32, 50, and 67 blows respectively. The impact resistance test results interpretation is again divided into three stages:

At first, the addition of PP fibers: these fibers incorporated with the cement paste, enhanced the ability of micro cracks to spread without breaking or pulling the fibers and therefore the cracks in concrete were reduced.

The addition of small amounts of water-soluble polymers (PVA solution): enhances the bonding strength and durability of cementing materials.

The presence of steel mesh that has higher tensile strength: The bonding between concrete and reinforcing mesh plays a major role in impact resistance, as the number of blows required to cause first crack and failure increases. This can be attributed to the effect of the reinforcement mesh which helped absorb more impact energy efficiently (Figures 7-9). The behavior of superplasticized concrete slabs reinforced with different types of polymer grids and subjected to low velocity impact loading was studied extensively in [17]. The tested 500×500×50mm and 500×500×25mm slabs had either imported or local polymer grids. The adopted falling steel mass was 3.36kg in weight and 9.5cm in diameter and the falling height was 1.2m. Three mixes were used. Many properties were tested during this research, including compressive strength and impact resistance. The test results indicated that the presence of polymer grids as concrete reinforcement enhances significantly the mechanical properties of concrete. It increases the number of blows required to cause impact failure and delays first crack and final scabbing and spalling to take place [17].

Fig. 7. Typical failure behavior of reference concrete specimens under impact.

Fig. 8. Typical failure behavior of concrete specimens including 2% PP fibers and polymer grids under impact.

Fig. 9. Typical failure behavior of concrete specimens including 2% PP fibers and steel grids under impact.

V. CONCLUSIONS

According to the obtained experimental results, the following are the most notable outcomes of the current study:

- Compressive strength increased by 36.11%, 45.5%, and 52.4% over the traditional mixture at the age of 28, 56, and 90 days of water curing respectively.
- The fiber-reinforced slab and steel mesh showed superior performance in impact resistance compared to the traditional mixture and required 67, 93, and 118 blows at 28, 56, and 90 days respectively, while the slab reinforced with fiber and polymer mesh showed satisfactory results compared to the traditional mixture with 32, 50, and 67 respectively.
- The cost of bendable concrete was reduced by using IL cement.
- Bendable concrete made with locally available materials and reinforcements can be widely used instead of conventional concrete at complex construction sites.
- It is possible to use locally manufactured IL cement in producing bendable concrete with acceptable mechanical properties reinforced with locally available grids.

REFERENCES

[1] M. T. Lakhiar, S. Sohu, I. A. Bhatti, N. Bhatti, S. A. Abbasi, and M. Tarique, "Flexural Performance of Concrete Reinforced by Plastic Fibers," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 3041–3043, Jun. 2018, https://doi.org/10.48084/etasr.2084.

[2] M. T. Lakhiar, N. Mohamad, M. a. B. Shaikh, A. A. Vighio, A. A. Jhatial, and A. A. A. Samad, "Effect of River Indus Sand on Concrete Tensile Strength," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2796–2798, Apr. 2018, https://doi.org/10.48084/etasr.1869.

[3] A. S. Buller, A. M. Buller, T. Ali, Z. A. Tunio, S. Shabbir, and M. A. Malik, "Experimental Characterization of Bacterial Concrete Against Mechanical and Durability Performance," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6703–6707, Feb. 2021, https://doi.org/10.48084/etasr.3983.

[4] V. C. Li, "On Engineered Cementitious Composites (ECC) A Review of the Material and Its Applications," *Journal of Advanced Concrete Technology*, vol. 1, no. 3, pp. 215–230, Jun. 2011.

[5] P. Archana, A. N. Nayak, S. R. Nayak, H. Vaddar, and D. S. Magnur, "Study of Strength of Polypropylene Fiber Reinforced Concrete," *International Journal of Engineering Research & Technology*, vol. 6, no. 6, May 2017.

[6] K. Ramasamy, S. Kandasamy, and K. Mani, "Influence of polymeric and non-polymeric fibers in hybrid engineered cementitious composites," *Romanian Journal of Materials*, vol. 48, no. 4, pp. 507–513, Jan. 2018.

[7] A. Chaudhary and A. K. Sharma, "Experimental Study on Flexural Behaviour of Engineering Cementitious Composite as Bendable Concrete," *International Journal for Research in Applied Science and Engineering Technology*, vol. 7, no. 6, pp. 1274–1278, Jun. 2019, https://doi.org/10.22214/ijraset.2019.6219.

[8] V. C. Li, "Strategies for High Performance Fiber Reinforced Cementitious Composites Development," in *Proceedings of International Workshop on Advances in Fiber Reinforced Concrete*, Bergamo, Italy, Sep. 2004.

[9] *Cement -Part 1: Composition, specifications and conformity criteria fo rcommon cements*. CEN, 2011.

[10] *Iraqi Specification, No. 45 Aggregate from Natural Sources for Concrete and Construction*. Baghdad, Iraq, 1984.

[11] C09 Committee, "ASTM C1240 - 03a: Specification for Silica Fume Used in Cementitious Mixtures," West Conshohocken, PA: ASTM International. https://doi.org/10.1520/C1240-03A.

[12] *ASTM C494 / C494M-15, Standard Specification for Chemical Admixtures for Concrete*. West Conshohocken, PA: ASTM International, 2015.

[13] *Iraqi Specification, No .1703: Water used in concrete*. Baghdad, Iraq: Central Organization for Standardization and Quality Control, 1992.

[14] "Testing concrete — Part 116: Method for determination of compressive strength of concrete cubes," in *British Standard*, BSI, 1983.

[15] S. S. Saeed, "Time-dependent deformation behavior of bendable concrete produced by Portland limestone cement," M.S. thesis, University of Baghdad, Baghdad, Iraq, 2016.

[16] ACI Committee 544, "Design Considerations for Steel Fiber Reinforced Concrete," ACI, West Conshohocken, PA, USA, ACI 544.4R-88, 1988.

[17] I. F. A. Almulla, "Behaviour of concrete units containing Polymer grids'," M.S. thesis, University of Baghdad, Baghdad, Iraq, 2002.

Multi-Layer Perceptron Neural Network Model Development for Chili Pepper Disease Diagnosis Using Filter and Wrapper Feature Selection Methods

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

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

Abstract-The goal of the current study is to develop a diagnosis model for chili pepper disease diagnosis by applying filter and wrapper feature selection methods as well as a Multi-Layer Perceptron Neural Network (MLPNN). The data used for developing the model include 1) types, 2) causative agents, 3) areas of infection, 4) growth stages of infection, 5) conditions, 6) symptoms, and 7) 14 types of chili pepper diseases. These datasets were applied to the 3 feature selection techniques, including information gain, gain ratio, and wrapper. After selecting the key features, the selected datasets were utilized to develop the diagnosis model towards the application of MLPNN. According to the model's effectiveness evaluation results, estimated by 10-fold cross-validation, it can be seen that the diagnosis model developed by applying the wrapper method along with MLPNN provided the highest level of effectiveness, with an accuracy of 98.91%, precision of 98.92%, and recall of 98.89%. The findings showed that the developed model is applicable.

Keywords-chili pepper diseases; feature selection; multi-layer perceptron neural network; wrapper

I. INTRODUCTION

Chili peppers are plants used and consumed by Thai people in many forms. They are fundamental spices that can enhance the flavor, odor, and color of food. Chili peppers are considered unique due to their nutritional and medical benefits, flavors, and colors, so they cannot be replaced by other plants, hence, they are highly economically important in Thailand [1]. Chili peppers must be appropriately stored and maintained before being sold. Many farmers usually face problems in chili pepper plantations. These include inevitable natural disasters, weeds, and pests. Therefore, the farmers need to be educated about chili pepper diseases in order to protect their plants appropriately. Multi-Layer Perceptron Neural Networks (MLPNNs) have been broadly applied to diagnose diseases. For example, such models were developed in [2] and in [3] for predicting lung cancers and heart diseases respectively. These two studies developed disease diagnosis models which were more than 90% effective.

The current research aims to apply MLPNNs to diagnose chili pepper diseases. It also employs the filter and wrapper feature selection methods to develop the diagnosis model. The feature selection methods will select only significant features for synthesizing the model rapidly and classifying data more effectively [4]. The most effective model in terms of data classification will be further applied to the development of a chili pepper disease diagnosis system that can assist the farmers in diagnosing the diseases and treat their plants timely and adequately.

II. LITERATURE REVIEW

A. Filter Feature Selection Methods

Filter feature selection methods are the processes that evaluate the level of effectiveness of each feature regarding its compatibility with data analysis. These processes do not rely on any learning approach. These filter methods select the data by ranking each feature in a priority queue. The priority list is based on the number specified by the user or the threshold of the selected feature. The advantages of these methods are rapid data processing and independence from learning methods [4]. This research applies two types of filter feature selection methods, namely the Information Gain (IG) and Gain Ratio (GR), as follows.

1) Information Gain (IG)

IG is a measure used for classifying data by calculating the gain value of each attribute. If any attribute has the highest gain value, it will be selected as a subset with classification power, as shown in (1) [5]. Equation (1) demonstrates the calculation of entropy, while (2) explains the calculation of the gain value [5].

$$Entropy = -\sum_{i=0}^{c-1} p_i(t)\log_2 p_i(t) \quad (1)$$

$$Gain = Entropy(\text{parent}) - \sum_{i=1}^{k} \frac{N(v_i)}{N} Entropy\left(v_i\right) \quad (2)$$

where c refers to the number of classes, $p_i(t)$ refers to the frequency value of class i for node t, $Entropy$(parent) refers to

---

Corresponding author: Sumitra Nuanmeesri the entropy of the parent node, k refers to the total number of feature values, N refers to the total amount of data of the parent node, and $N(v_i)$ refers to the total amount of instances of child node i.

2) Gain Ratio (GR)

GR is a measure that uses the gain ratio as the indicator for dividing datasets into sub-datasets based on the IG value. However, when the IG value is used for classifying datasets, there would be the occurrence of bias in favor of attributes with large numbers of values. Thus, the GR can be derived from the gain value divided by the *SplitInfo* value that is calculated in (3) [6], where k refers to the total number of split data. Therefore, the GR is calculated in (4) [5]:

$$SplitInfo = -\sum_{i=1}^{k} \frac{N(v_i)}{N} \log_2 \frac{N(v_i)}{N} \quad (3)$$

$$GR = \frac{Gain}{SplitInfo} \quad (4)$$

B. Wrapper Feature Selection

Wrapper feature selection methods are processes that select subsets from all features. They explore the feature subsets which specifically match a learning approach. Therefore, these methods can increase the effectiveness of the learning process to the greatest extent. This research applies an evolutionary wrapper for selecting features. The features, which are predictor variables, are randomized into the equation each by each. Then, their effectiveness in prediction is tested. If the level of effectiveness in prediction increases, the effective features will be maintained. In contrast, if the level of effectiveness in prediction decreases, the ineffective features will be removed [7].

C. Multi-Layer Perceptron Neural Network

MLPNN is a form of perceptron neural network with multiple layers. It is suitable for complex computational tasks. It consists of 3 layer lavels, namely the Input layer, the Hidden layer, and the Output layer. The Hidden layer could be composed of many layers, but there must be at least one layer in it [6, 8]. The data are the input of the Input layer. The output is sent out from the Output layer. The summation function of MLP is calculated in (5):

$$n = \sum_{i=1}^{k} p_i w_i + b \quad (5)$$

where n is the total sum gained from the summation function, p_i is the input of neuron i, w_i is the weight of neuron I, k is the number of Input layer neurons, and b is the bias value. The MLPNN model is illustrated in Figure 1.

D. Model's Effectiveness Evaluation

The model's effectiveness was evaluated by the confusion matrix, which can calculate the precision, recall, and accuracy of the model [9-12].

E. Related Works

Authors in [2] developed a model to predict lung cancer using the MLPNN, comparing the efficacy of disease classification with the K-Nearest Neighbor (KNN) technique. The results showed that the disease classification model using the MLPNN was more effective. Authors in [3] applied the MLPNN to predict heart diseases. The efficacy of the heart disease classification model was compared with 9 classification methods and the results showed that the MLPNN was the most effective.

Authors in [13] applied the Wrapper method to select feature data to classify cardiac arrhythmias. The model's effectiveness was tested using 10-fold cross-validation, comparing 5 modeling techniques. The results showed that the model used to classify cardiac arrhythmias using Wrapper combined with MLP had the best performance. Authors in [14] developed an MLP model by applying the Correlation-based Feature Selection (CFS) and IG to analyze Thai water buffalo diseases. The experimental results showed that the developed model by CFS and MLP was efficient with accuracy, precision, and recall greater than 99.0%.

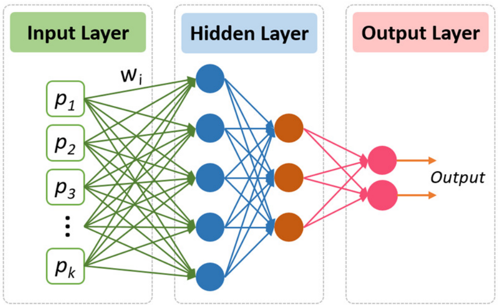


Fig. 1. Multi-Layer Perceptron Neural Network model.

III. METHODS

In this work, a MLPNN model for chili pepper disease diagnosis was developed using filter and wrapper feature selection methods. Six 6 processes were involved: data collection, data preparation, feature selection, modeling, evaluation, and deployment.

A. Data Collection

Primary data were collected from 5 agricultural professionals and 33 chili pepper farmers in Ubon Ratchathani Province of Thailand with the use of questionnaires. This is a significant chili pepper cultivation area in 3 districts, namely Muang Sam Sip District, Mueang District, and Khueang Nai. The secondary data were collected from the review in [15]. An analysis of the data revealed 14 diseases in chili peppers in those plantations, as shown in Figure 2. The details of the data are illustrated in Table I.

B. Data Preparation

At the data preparation stage, the selected data were prepared before being applied to the model. The data were prepared in the following steps: 1) data selection, 2) data cleaning, and 3) data transformation. Ultimately, 863 questionnaires were processed, and were divided into 80% for training and 20% for testing. The data used for developing the chili pepper disease diagnosis model are illustrated in Table II.

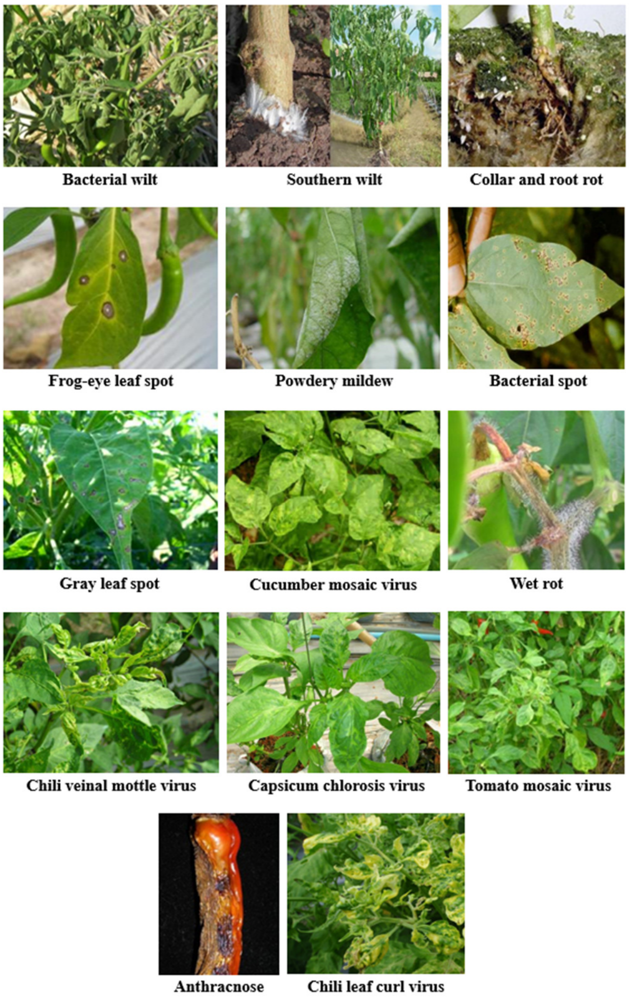


Fig. 2. Sample image of 14 chili peppers diseases.

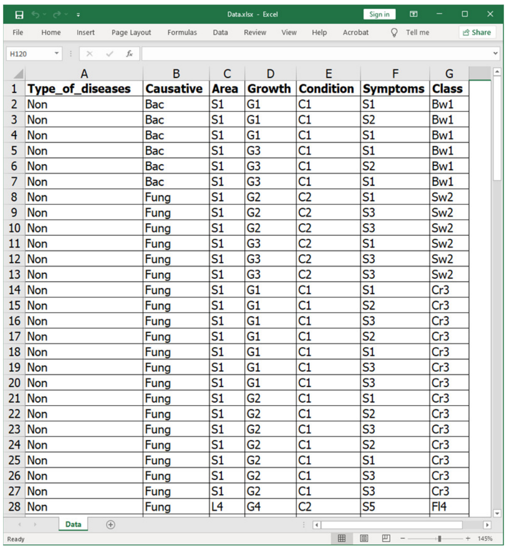
TABLE I. DATA USED IN THIS RESEARCH

No.	Data
1	Types of chili pepper diseases
2	Causative agents
3	Infected areas
4	Growth stages of infection
5	Conditions for diseases
6	Symptoms
7	Fourteen types of chili pepper diseases: 1. Bacterial wilt 2. Southern wilt 3. Collar and root rot 4. Frog-eye leaf spot 5. Powdery mildew 6. Bacterial spot 7. Wet rot 8. Gray leaf spot 9. Anthracnose 10. Cucumber mosaic virus 11. Chili veinal mottle virus 12. Chili leaf curl virus 13. Capsicum chlorosis virus 14. Tomato mosaic virus

TABLE II. FEATURES USED FOR DEVELOPING THE CHILI PEPPER DISEASE DIAGNOSIS MODEL

No.	Data	Features
1	Types of chili pepper diseases	Non: nonliving (abiotic) agents
		Liv: living (biotic) agents
2	Causative agents	Bac: bacterial agents
		Fung: fungal agents
		Vir: viral agents
3	Infected areas	S1: stems
		F2: fruits
		R3: roots
		L4: leaves
		St5: stems and roots
		Lf6: leaves and fruits
		Ls7: leaves and stems
		Fs8: fruits and stems
		Sl9: stems, leaves, and fruits
4	Growth stages of infection	G1: early stage of growth
		G2: flowering
		G3: harvest
		G4: at all stages
5	Conditions for diseases	C1: poorly drained soil
		C2: humidity and rain consecutive days
		C3: rains and wet soil
		C4: strong and dry winds
		C5: frost and strong winds
		C6: overcrowded sprouts; poor ventilation
		C7: thrips
		C8: tobacco whiteflies
6	Symptoms	S1: browned vessels
		S2: white powdery coating or brown pellets appearing on the lower surface
		S3: water soaked lesions on stems and stalks
		S4: midget growth
		S5: severely spotted leaves turn yellow and drop
		S6: black narrow elongated lesions or streaks developing throughout the stems to the top of the plant
		S7: bleached, pale veinlets
		S8: wet, macerated stalks with black-ended silvery hair covering the lesions
		S9: distorted, irregular, curly leaves
		S10: yellowing of lower leaves at the front and back
		S11: a dark ring and a yellowish halo around the ring, forming a "frog-eye" appearance on the leaves
7	Types of chili pepper diseases (classes)	Bw1: bacterial wilt
		Sw2: southern wilt
		Cr3: collar and root rot
		Fl4: frog-eye leaf spot
		Pm5: powdery mildew
		Bs6: bacterial spot
		Wr7: wet rot
		Gr8: gray leaf spot
		An9: anthracnose
		Cm10: cucumber mosaic virus
		Cv11: chili veinal mottle virus
		Cl12: chili leaf curl virus
		Cc13: capsicum chlorosis virus
		Tm14: tomato mosaic virus

Once the data in Table II had been collected, the research team verified their accuracy. Then, they were converted into a csv file to be computed in Weka 3.9, as illustrated in Figure 3. After converting the data, the authors input the data into Weka to select features by applying IG, GR, and Wrapper.



	A	B	C	D	E	F	G
1	**Type_of_diseases**	**Causative**	**Area**	**Growth**	**Condition**	**Symptoms**	**Class**
2	Non	Bac	S1	G1	C1	S1	Bw1
3	Non	Bac	S1	G1	C1	S2	Bw1
4	Non	Bac	S1	G1	C1	S1	Bw1
5	Non	Bac	S1	G3	C1	S1	Bw1
6	Non	Bac	S1	G3	C1	S2	Bw1
7	Non	Bac	S1	G3	C1	S1	Bw1
8	Non	Fung	S1	G2	C2	S1	Sw2
9	Non	Fung	S1	G2	C2	S3	Sw2
10	Non	Fung	S1	G2	C2	S3	Sw2
11	Non	Fung	S1	G3	C2	S1	Sw2
12	Non	Fung	S1	G3	C2	S3	Sw2
13	Non	Fung	S1	G3	C2	S3	Sw2
14	Non	Fung	S1	G1	C1	S1	Cr3
15	Non	Fung	S1	G1	C1	S2	Cr3
16	Non	Fung	S1	G1	C1	S3	Cr3
17	Non	Fung	S1	G1	C1	S2	Cr3
18	Non	Fung	S1	G1	C1	S1	Cr3
19	Non	Fung	S1	G1	C1	S3	Cr3
20	Non	Fung	S1	G1	C1	S3	Cr3
21	Non	Fung	S1	G2	C1	S1	Cr3
22	Non	Fung	S1	G2	C1	S2	Cr3
23	Non	Fung	S1	G2	C1	S3	Cr3
24	Non	Fung	S1	G2	C1	S2	Cr3
25	Non	Fung	S1	G2	C1	S1	Cr3
26	Non	Fung	S1	G2	C1	S3	Cr3
27	Non	Fung	S1	G2	C1	S3	Cr3
28	Non	Fung	L4	G4	C2	S5	Fl4

Fig. 3. A sample of the data used in model development.

C. Modeling

The data were divided into 4 datasets, namely 3 datasets undergoing the feature selection process and 1 original dataset used for MLPNN modeling. The research team specified the parameters for the MLPNN modeling as follows: Hidden layer = 2, Training time = 500, Learning rate = 0.3, and Momentum = 0.2. These parameters provided the best results.

D. Evaluation

The effectiveness of the chili pepper disease diagnosis model was evaluated by 10-fold cross-validation with the test dataset (20% of the total data). In addition, it was measured by examining the precision, recall, and accuracy of the developed models [6].

E. Deployment

After constructing the model, the most effective sample was applied to develop the chili pepper disease diagnosis prototype system. This system can diagnose chili pepper diseases timely and suggest proper treatments for each disease rapidly. The research model framework is illustrated in Figure 4.

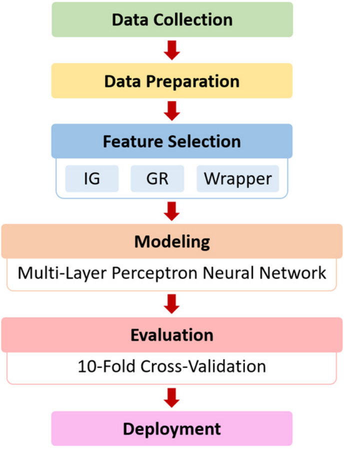


Fig. 4. The research model framework.

IV. RESULTS

The chili pepper disease diagnosis model was developed by applying filter and wrapper methods along with MLPNN. The number of features remaining after the feature selection process is shown in Table III.

TABLE III. REMAINING FEATURES AFTER FEATURE SELECTION

Feature selection method	Number of features remaining
IG	7
GR	7
Wrapper	6

The effectiveness of the models was evaluated by 10-fold cross-validation. The research results showed that the model developed by Wrapper and MLPNN (Wrapper+MLP) provided the highest results with accuracy of 98.91%, precision of 98.92%, and recall of 98.89%. Next in the rankings were IG+MLP, GR+MLP, and MLP (original data), as shown in Table IV and Figure 5. Furthermore, compared to other studies that detected or classified chili pepper diseases (Table V), the developed model by the Wrapper+MLP method gave the highest effectiveness result.

TABLE IV. EFFECTIVENESS EVALUATION RESULTS

Method	Precision (%)	Recall (%)	Accuracy (%)
MLP (original data)	92.60	92.54	92.52
IG+MLP	96.50	96.42	96.45
GR+MLP	93.34	93.20	93.26
Wrapper+MLP	98.92	98.89	98.91

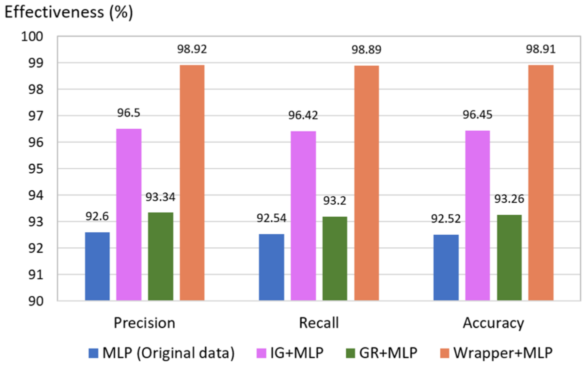


Fig. 5. Comparison between the models' effectiveness evaluation results.

TABLE V. COMPARISON WITH OTHER RESEARCH STUDIES

Method	No. of classes	Accuracy (%)
HSV model feature extraction [16]	2	80.00
Hyperspectral image + MLP [17]	2	83.26
Deep Belief Network [18]	2	91.96
Deep Learning + Support Vector Machine [19]	7	92.10
Fuzzy C-Means segmentation [20]	5	97.56
Wrapper + MLP (This study)	14	98.91

V. CONCLUSION AND DISCUSSION

This research aimed to develop a chili pepper disease diagnosis model by applying filter and wrapper feature selection methods along with the MLPNN. The data used for classifying the 14 considered chili pepper diseases were processed by 3 feature selection techniques, namely IG, GR, and Wrapper.

Once the feature selection process was completed, the selected data were applied to develop the chili pepper disease diagnosis model with the use of the MLPNN The experimental results indicated that the diagnosis model developed by the Wrapper method and the MLPNN provided the highest level of effectiveness, with an accuracy of 98.91%. This result means that the developed model can be used in a chili pepper disease diagnosis system, suggesting chili pepper disease prevention and treatment. The results of this research conform to the study conducted in [3], in which the authors applied the MLP method to develop a disease diagnosis model which proved to perform better than other known techniques. The research findings also conform to the findings in [13]. The authors employed the wrapper method and the MLP to classify the symptoms of cardiac arrhythmia. Their research results showed that the wrapper feature selection method could increase the effectiveness of the classification process.

In general, a disease on chili peppers can be identified by its symptoms. Most research is focused on image processing [16, 17, 20] and deep learning [18, 19] disease analysis from chili pepper leaves that require large image datasets to train the model. However, some symptoms that appear on chili pepper leaves may be similar, resulting in discrepancies in the efficiency and accuracy of classification. In addition, the disease identification of chili peppers from leaves is limited to some diseases with distinctive features appearing on the leaves. It is often necessary to analyze symptoms in other parts of the chili pepper plant, such as roots, stems, leaves, and fruits. Therefore, this study used the information presented on plants throughout the chili pepper life cycle and considered the environmental conditions to help identify 14 occurring diseases. Besides, the Wrapper feature selection process proved to be more efficient than the IG, and GR, and the classification without feature selection.

ACKNOWLEDGMENT

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

REFERENCES

[1] K. Lertrat, "Production, planting, processing, marketing, and chili pepper products in Thailand," *Research Community*, vol. 73, pp. 15–20, May 2007.

[2] S. Potghan, R. Rajamenakshi, and A. Bhise, "Multi-Layer Perceptron Based Lung Tumor Classification," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, Mar. 2018, pp. 499–502, https://doi.org/10.1109/ICECA.2018.8474864.

[3] K. Subhadra and B. Vikas, "Neural Network Based Intelligent System for Predicting Heart Disease," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 5, pp. 484–487, 2019.

[4] K. Sutha and J. J. Tamilselvi, "A review of feature selection algorithms for data mining techniques" *International Journal on Computer Science and Engineering*, vol. 7, no. 6, pp. 63–67, Jun. 2015.

[5] P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining*, 2nd ed. New York, USA: Pearson Education, 2019.

[6] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*, 1st ed. Boston, Massachusetts, USA: Addison-Wesley, 2005.

[7] Y. B. Wah, N. Ibrahim, H. A. Hamid, S. Abdul-Rahman, and S. Fong, "Feature selection methods: Case of filter and wrapper approaches for maximising classification accuracy," *Pertanika Journal of Science & Technology*, vol. 26, no. 1, pp. 329–340, Jan. 2018.

[8] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *International Journal of Artificial Intelligence and Expert Systems*, vol. 1, no. 1, pp. 111–122, 2011.

[9] S. Nuanmeesri, S. Chopvitayakun, P. Kadmateekarun, and L. Poomhiran, "Marigold flower disease prediction through deep neural network with multimodal image," *International Journal of Engineering Trends and Technology*, vol. 69, no. 7, pp. 174–180, Jul. 2021, https://doi.org/10.14445/22315381/IJETT-V69I7P224.

[10] S. Nuanmeesri, L. Poomhiran, and K. Ploydanai, "Improving the prediction of rotten fruit using convolutional neural network," *International Journal of Engineering Trends and Technology*, vol. 69, no. 7, pp. 51–55, Jul. 2021, https://doi.org/10.14445/22315381/IJETT-V69I7P207.

[11] A. N. Saeed, "A Machine Learning based Approach for Segmenting Retinal Nerve Images using Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5986–5991, Aug. 2020, https://doi.org/10.48084/etasr.3666.

[12] 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.

[13] A. Mustaqeem, S. M. Anwar, M. Majid, and A. R. Khan, "Wrapper method for feature selection to classify cardiac arrhythmia," in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Jeju, South Korea, 2017, pp. 3656–3659, https://doi.org/10.1109/EMBC.2017.8037650.

[14] S. Nuanmeesri and W. Sriurai, "Thai Water Buffalo Disease Analysis with the Application of Feature Selection Technique and Multi-Layer Perceptron Neural Network," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6907–6911, Apr. 2021, https://doi.org/10.48084/etasr.4049.

[15] S. Sudhi-Aromna *et al., A guide to chili pepper pests*, Nonthaburi, Thailand: The Agricultural Co-operative Federation of Thailand, Ltd., 2014.

[16] D. P. Patil, S. R. Kurkute, P. S. Sonar, and S. I. Antonov, "An advanced method for chilli plant disease detection using image processing," in *52nd International Scientific Conference On Information, Communication and Energy Systems and Technologies*, Niš, Serbia, 2017, pp. 309–313.

[17] M. Ataş, Y. Yardimci, and A. Temizel, "A new approach to aflatoxin detection in chili pepper by machine vision," *Computers and Electronics in Agriculture*, vol. 87, pp. 129–141, 2012, https://doi.org/10.1016/j.compag.2012.06.001.

[18] S. Jana, A. R. Begum, and S. Selvaganesan, "Design and analysis of pepper leaf disease detection using Deep Belief Network," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 9, pp. 1724–1731, 2020.

[19] N. N. Ahmad Loti, M. R. Mohd Noor, and S.-W. Chang, "Integrated analysis of machine learning and deep learning in chili pest and disease identification," *Journal of the Science of Food and Agriculture*, vol. 101, no. 9, pp. 3582–3594, 2021, https://doi.org/10.1002/jsfa.10987.

[20] S. Das Chagas Silva Araujo, V. S. Malemath, and K. M. Sundaram, "Symptom-Based Identification of G-4 Chili Leaf Diseases Based on Rotation Invariant," *Frontiers in Robotics and AI*, vol. 8, 2021, Art. no. 650134, https://doi.org/10.3389/frobt.2021.650134.

AUTHORS PROFILE

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

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

Magnetite Particle Presence in the Human Brain: A Computational Dosimetric Study to Emphasize the Need of a Complete Assessment of the Electromagnetic Power Deposition at 3.5 GHz

David Vatamanu
Faculty of Engineering, Lucian Blaga University
and Nicolae Balcescu Land Forces Academy
Sibiu, Romania
davidvatamanu@yahoo.com

Simona Miclaus
Nicolae Balcescu Land Forces Academy
Sibiu, Romania
simo.miclaus@gmail.com

Abstract-The growing evidence of increased magnetite nanoparticles (both endo- and exogenic) in the human brain raises the importance of assessing the entire power deposition when electromagnetic waves at GHz frequencies propagate in such tissues. This frequency range corresponds to many popular portable communication devices that emit radiation close to a human's head. At these frequencies, the current dosimetric numerical codes can not accurately compute the magnetic losses part. This is due to the lack of an implemented computational algorithm based on solving the coupled Maxwell and Landau-Lifshitz-Gilbert equations, in the case of magneto-dielectrics, considering eddy currents losses and specific properties of magnetic sub-millimetric particles. This paper focuses on analyzing the limits and the inconsistencies when using commercial dosimetric numerical software to analyze the total absorbed power in brain models having ferrimagnetic content and being exposed to 3.5GHz electromagnetic waves. Magnetic losses computed using Polder's permeability tensor as constitutive relation lead to unreliable results. However, using such software can provide a preliminary view of the electromagnetic impact of ultra- and super-high frequencies on magnetic-dielectric tissues.

Keywords-magnetic brain; magnetite particles; magnetic-dielectric; microwaves dosimetry; power loss density

I. INTRODUCTION

Magnetic nanoparticles in human tissues, and more specifically in the brain, have been scarcely investigated for their electromagnetic wave absorption at Ultra-High Frequencies (UHF) [1, 2]. The presence of biogenic magnetite (Fe_3O_4) nanocrystals in the human brain was discovered in 1992, while the concentration of 0.2-12μg magnetite in each 1g of dry cerebral tissue was identified in 2009 [3]. Besides the biogenic magnetite particles, exogenic magnetite nanocrystals, arising from atmospheric pollution, were recently discovered to be deposited in some organs. These nanocrystals were first identified in 2016 in the human brain [4] and later in other organs too (heart, blood, liver, etc.) [5, 6]. The exogenic magnetite nanoparticles present a different morphology from the endogenic [7]. Significant concerns were raised for increased health risks as these nanoparticles are regarded as highly toxic [8]. They can enter the blood flow and be transported to the brain. Moreover, they can catalyze the generation of reactive oxygen species in vivo and cause protein modification, lipid peroxidation, or DNA damage [8]. The increased magnetite content in the human brain raises another question about its contribution to the mechanism of thermal or non-thermal interactions with electromagnetic fields in high frequencies. In general, magnetite nanoparticles absorb very well electromagnetic waves at UHF frequencies [9-12]. Although their role in brain functioning is still unknown, memory mechanisms' [13] or crossing the blood-brain barrier [14-15] involvements cannot be neglected. A map of locations and mass information about magnetite nanoparticles present in the human brain was published in 2018 [16], while the concentration differences in the brains of people belonging to different geographical regions were highlighted in 2021 [17]. To answer this question and address the problem of quantification of electromagnetic power deposition in tissues having ferro- or ferrimagnetic content, a simulation was conducted on wave propagation in human brain models using commercial software. Such software suites provide radiofrequency/microwave dosimetry solutions at high frequencies but have a serious limitation regarding their capability to compute correctly the magnetic losses.

The electromagnetic waves emitted by various communication devices using UHF or Super-High Frequency (SHF) bands affect the human head in a very significant manner. Due to the eyes, ears, and mouth positions, the brain is practically always near the wave sources during all kinds of wireless communication devices usage. The brain is a target of electromagnetic radiation and the dose rate it receives has various thermal and non-thermal biological effects. The thermal effects represent the basis of the human protection standards metrics and safety limits [18, 19], while the non-thermal are still under debate and study [20-22]. On the other

---

Corresponding author: Simona Miclaus

hand, radiofrequency and microwave numerical dosimetry is continuously following the development and the popularity of emitting technologies [23-30]. Computational dosimetry using pure dielectric head models started more than 20 years ago and developed progressively using more and more accurate head models [31-40]. However, none dosimetric research provided a comparative quantification of the dissipated power when a brain contains magnetite or not. In this respect, this work underlines the limits and gaps in conducting a realistic quantification of the entire power loss in a "magnetic brain" by using available software, emphasizing its possible role to address the "magnetic brain" dosimetric problem in the future.

Micromagnetic software suites are widely used in the electronic industry, but they can't be utilized in bioelectromagnetics, as they generally consider tissues as pure dielectrics. These software suites have not implemented the proper coupled Maxwell and Landau-Lifshitz-Gilbert (LLG) equations to provide reliable results concerning the magnetic power losses when very high-frequency waves pass a tissue containing magnetic particles. Neither the eddy currents nor the specific properties (geometric and magnetic anisotropies, etc.) of the magnetic nanoparticles are considered in such approaches. This analysis introduces the subject of reconsidering the software to use in order to assess the real impact of the magnetic nanoparticles' presence in tissues.

II. LIMITATIONS OF PRESENT RADIOFREQUENCY SOFTWARE

The CST Studio Suite [41] is used to simulate and analyze antenna parameters [42-44]. CST and all similar commercial software provide also a monitor function for computing the Specific Absorption Rate (SAR) of energy deposition in the tissues. SAR is a reference metric in all standards and is defined by taking into account only the dielectric losses in the tissues. SAR is related to temperature increase (ΔT) by:

$$\mathrm{SAR} = \frac{\sigma E_{rms}^2}{\rho} = c\frac{\Delta T}{\Delta t} \quad (1)$$

where σ is the electric conductivity of the tissue, E_{rms} is the root-mean-square value of the internal/local electric (E)-field strength, ρ is the mass density, c is the specific heat of the tissue, ΔT is the temperature variation, and Δt is the time duration of the wave propagation that dissipated the heat. As SAR is expressed in W/kg and power loss density is expressed in W/m^3, they can be directly connected to the mass density.

Magnetic losses are not included properly in SAR calculations of current dosimetric software packages. Magnetic losses may have 4 possible origins: a) relaxation due to the rotation of the magnetic moments (without rotation of the whole particle) – Neel relaxation, b) relaxation due to the rotation of the particle in the alternating field – Brownian relaxation, c) shifting of magnetic domain walls in multi-domain materials - hysteresis loss, and d) generation of Foucault (eddy) currents in bulk materials or at centimetric scale (resistive heating when the magnetic flux is rapidly varying) [45-48]. The first 3 mechanisms of heat dissipation are detected mainly in nanometric dimensions and depend on size, shape, crystalline anisotropy, morphology, and degree of aggregation/agglomeration of the crystals. If ferro- or ferrimagnetic (like magnetite) material is present in a dielectric tissue, the whole organ should be treated as a magneto-dielectric. Sub-millimetric magnetic particles are called micromagnetics. "Magnetic brain" falls in the magneto-dielectrics and should be treated by micromagnetics, since the dimensions of the magnetite crystals are less than 200nm if exogenic and less than 70nm if endogenic. The dynamics of the magnetization vector M in a material exposed to an alternating magnetic field H is governed by the LLG equation [49]:

$$\frac{\partial M}{\partial t} = -\frac{\gamma}{1+\alpha^2} M \times H_{eff} - \frac{\gamma\alpha}{(1+\alpha^2)M_s} M \times M \times H_{eff} \quad (2)$$

where M_s is the saturation magnetization, α is the damping constant, γ is the gyromagnetic ratio of the electron, and H_{eff} is the effective magnetic field. H_{eff} includes the actual magnetic field H obtained as a solution of the Maxwell equations and other terms which take into account the crystalline anisotropy, the Heisenberg exchange interaction (quantum mechanical effect), etc. The LLG equation states that the change rate of the magnetic moment (magnetization M) is proportional to the torque due to different interactions. The magnetic moment itself is due to the orbital rotation and spin of electrons. The Larmor precession of the spin creates a resonant frequency in the magnetic material property, defined by ferromagnetic resonance. Equation (2) shows that torques are due to the magnetic field acting on magnetic moments. So the LLG equation requires knowledge of the magnetic field, which can be obtained by solving the Maxwell equations.

The vast majority of micromagnetic solvers use the magnetostatic approximation to the Maxwell equations [41]. In that case, E- and H-fields are completely decoupled. Many micromagnetic software suites solve the LLG equation dynamically with coupling to the magnetostatic solution [49, 50]. These tools do not predict the interaction between the magnetization M and electromagnetic waves. In that case, the Maxwell equations do not include any dynamic effects, and the fields propagate through the system instantaneously. The static approximation is valid when $\tau_m / \tau \ll 1$, where τ is the characteristic time of the system dynamics, and τ_m is the magnetic diffusion time constant given by $\tau_m = \sigma\mu L^2$, where μ is the magnetic permeability and L is the system size [51].

When frequency increases, the effects of induced eddy currents and dynamic magnetic fields become significant and cannot be neglected. In such cases, it becomes necessary to renounce the static Maxwell equations [50-55]. The magneto-quasistatic approximation (eddy current approximation) of the Maxwell equations is based on neglecting the displacement current term $\partial \mathrm{D}/\partial t$. Eddy current effects are negligible only when time variations of H and M are slow enough and the system size is smaller than the skin depth. A magneto-quasistatic approximation is therefore valid whenever the system size is much smaller than the wavelength, and when the induced E-fields (due to time-varying H-fields and M) are the dominant effect over capacitive effects. This approximation neglects wave propagation effects and raises the diffusion of both H-field and induced eddy currents. Indeed, under the magneto-quasistatic approximation, H-field satisfies the following diffusion equation [49]:

$$\frac{\partial H}{\partial t} = \frac{1}{\sigma\mu_0}\nabla^2 H - \frac{\partial M}{\partial t} + \frac{1}{\sigma\mu_0}\nabla\nabla \cdot M \quad (3)$$

A similar equation is valid for the induced eddy currents, noted by *J*.

Full-wave electromagnetic simulators, such as CST Studio Suite and Ansys HFSS 3D High-Frequency Simulation Software [56], solve Maxwell's equations by defining material dispersion and anisotropy through the application of Polder's permeability tensor as the constitutive relation. This tensor is used because ferrimagnetic materials become anisotropic in the presence of a magnetizing field. From the perspective of solving the Maxwell equations properly, the LLG acts as the constitutive relation between the magnetic field H and the magnetic flux density B through $B = \mu_0(H + M)$. Given H, the magnetization M is obtained by solving the LLG equation. The LLG and Maxwell equations must be solved jointly as a coupled system of differential equations, for realistic and complete computation results. Different strategies can be adopted to solve the coupled Maxwell-LLG equations. A coupled micromagnetic-Maxwell equations solver was developed only recently, taking into account the effect of the eddy currents [49, 51]. Neglecting the effects of eddy currents in the GHz spectrum is completely inappropriate. An efficient and accurate modeling tool for multiphysics problems that encompasses electrodynamics and micromagnetics was proposed in [48]. Three major problems in modeling interactive micromagnetics and electromagnetic waves were solved using [48]: 1) coupling between Maxwell's equations and the LLG equation in every time step, 2) scale disparity of the multiphysics, 3) field discontinuity at material interfaces. However, such advanced software packages [52, 55] have not yet been implemented to treat bioelectromagnetic dosimetry issues. In the multiphysics problem, involved in ferro- and ferrimagnetic materials exposed to radiofrequency/microwaves, the dynamic Maxwell's (4) and LLG (5) equations must be considered as:

$$\begin{aligned} \nabla \times H &= \varepsilon \frac{\partial E}{\partial t} + J + \sigma E \\ \nabla \times E &= -\frac{\partial B}{\partial t} \end{aligned} \qquad (4)$$

$$\frac{\partial M}{\partial t} = \mu_0 \gamma (M \times H) - \frac{\alpha}{|M|} M \times \frac{\partial M}{\partial t} \qquad (5)$$

Consequently, using CST Studio for dielectric versus magnetic brain simulations provides an incomplete/partial solution to the role that magnetite crystals have in absorbing electromagnetic waves in the UHF and SHF spectra since it doesn't implement the coupled Maxwell-LLG equations. Further on, CST Studio was used with its underlined limitations to quantify power loss, stored electric energy, and magnetic energy densities in pure dielectric versus magneto-dielectric brain models. The absorbed power (equal to power loss P_{loss}) was calculated in each model by using the mean power loss density values and the corresponding volumes of the materials. The energy density formula is obtained as a byproduct of the energy conservation law (Poynting theorem), which is derived using Maxwell's equations with the aid of the equations of motions of the polarization P (in dielectric) and magnetization M (in magnetic) mediums:

$$-\nabla \cdot \mathrm{S} = \frac{\partial W}{\partial t} + P_{loss} \qquad (6)$$

where S, W, and P_{loss} stand for the Poynting vector (power density in $\mathrm{W/m^2}$), energy density (in $\mathrm{J/m^3}$), and power loss (W), respectively. Equation (6) can be re-written as:

$$-\nabla \cdot (\mathrm{E} \times \mathrm{H}) = \frac{\partial W_e}{\partial t} + \frac{\partial W_b}{\partial t} + P_{loss} \qquad (7)$$

$$P_{loss} = \mathrm{E.J} \qquad (8)$$

where W_e and W_b are the electric and magnetic energy densities, and J is the current density. Equation (7) can be rewritten as:

$$\begin{aligned} -\nabla \cdot (\mathrm{E} \times \mathrm{H}) &= \mathrm{E} \cdot \frac{\partial \mathrm{D}}{\partial t} + \mathrm{H} \cdot \frac{\partial \mathrm{B}}{\partial t} \\ &= \frac{\partial}{\partial t}\left(\frac{\epsilon_0 \mathrm{E}^2}{2}\right) + \mathrm{E} \cdot \frac{\partial \mathrm{P}}{\partial t} + \frac{\partial}{\partial t}\left(\frac{\mu_0 \mathrm{H}^2}{2}\right) \\ &+ \mu_0 \mathrm{H} \cdot \frac{\partial \mathrm{M}}{\partial t} \end{aligned} \qquad (9)$$

The time-averaged energy density of a harmonic electromagnetic wave, $\langle W \rangle$, can be calculated as:

$$\langle W \rangle = \frac{\epsilon_0}{4} \frac{\partial(\omega \epsilon(\omega))}{\partial \omega} |\mathrm{E}|^2 + \frac{\mu_0}{4} \frac{\partial(\omega \mu(\omega))}{\partial \omega} |\mathrm{H}|^2 \qquad (10)$$

III. COMPUTING THE ELECTROMAGNETIC LOSSES BASED ON POLDER'S PERMEABILITY TENSOR AS CONSTITUTIVE RELATION

A. Computational Models: Geometries, Electric and Magnetic Properties

Since no micromagnetic solver has yet the capacity to include eddy currents' effects, CST Studio was used for some magneto-dielectric material computations. A computational approach was applied, using magnetite particles having a radius of 1mm, to distinguish the contribution of these particles to the total absorbed power in the brain and separate the electric and magnetic energy densities. The computations were conducted in the frequency domain and used almost the same mesh dimensions in all cases (mesh properties: hexahedral, cells/wavelength = 15, spherical brain model: smallest cell = 0.2955, largest cell = 5.6774, realistic brain model: smallest cell = 0.3217, largest cell = 5.6832). A series of electromagnetic exposures were designed, computed, and analyzed to observe the contribution of the magnetite particles to the total absorbed power in the following brain models:

- S1: a simple spherical brain model (homogeneous), having the dimensions and material properties given in Table I. This model was purely dielectric.
- S2: a simple dielectric brain sphere, having a single particle of magnetite located in its center; magnetite volume was $\mathrm{V1}=682.67\mathrm{mm}^3$;
- S3: a simple dielectric brain sphere, containing N1=163 uniformly spread identical smaller particles of magnetite. Their total volume was equal to V1.
- S4: a simple dielectric brain sphere, having N2=129 uniformly spread smaller particles of magnetite (identical). Their total volume was equal to $\mathrm{V2}=540.27\mathrm{mm}^3$.

- S5: an MRI-based anatomical human head model, with non-magnetic brain tissues inserted (head of Gustav, belonging to the Bio-Models library, 2.08×2.08×2mm spatial resolution, containing 11 tissue types, 134028 voxels were used for the brain model while its mass was 1211.905g). The brain had the same volume as the simple spherical brain model (homogeneous).
- S6: an MRI-based anatomical human head model with magnetite inserted in the brain tissue (Gustav model with magnetite). A number N2=129 magnetite particles were uniformly spread in this model, with their total volume being equal to V2.

The mass concentration of magnetite in the brain was C1=1.6mg/g for S2 and S3, and C2=2.34mg/g for S4 and S6. Table I shows the dimensions of the materials used in the models. Table II presents the dielectric, magnetic, and other relevant properties of the materials at 3.5GHz. Magnetite dielectric and magnetic parameters at 3.5GHz were extracted from [49], while the brain and the rest tissues of the head were modeled based on dielectric data presented in [50].

TABLE I. BRAIN TISSUE AND MAGNETITE PARTICLES DIMENSIONS

Exposure situation	**Magnetite Volume (mm^3)**	**Brain Volume (mm^3)**	**Magnetite Mass (g)**	**Brain Mass (kg)**	**Brain sphere diameter (mm)**	**Magnetite sphere diameter (mm)**
S1	-	2144661.00	-	2.24	160	-
S2	682.67	2143978.00	3.58	2.24	160	10.92
S3	682.67	2143978.00	3.58	2.24	160	2
S4	540.28	1159170.72	2.83	1.21	130.35	2
S5	-	1159717.70	-	1.21	N/A	-
S6	540.275091	1159177.43	2.83	1.21	N/A	2

TABLE II. SIMULATION MATERIALS' PROPERTIES

	Brain	**Magnetite**	**Copper**	**FR4**
Material type	Normal	Normal	Lossy material	Normal
Real part – relative electric permittivity	45.81	4.8	N/A	4.3
Real part – relative magnetic permeability	1	0.84	1	1
Electric Conductivity (S/m)	0.767	0.233657	5.8E+07	0.00861191
Magnetic Conductivity (1/Sm)	0	7428.26	N/A	0
Thermal Conductivity (W/K/m)	0.5 (normal)	5.5 (anisotropic)	401.0	0.3
Heat capacity (kJ/K/kg)	3.7	0.6	0.39	0
Bloodflow coefficient (W/K/m^3)	40000	0	0	0
Basal metabolic rate (W/m^3)	7000	0	0	0
Mass density (kg/m^3)	1045	5240	8930	N/A

B. Characterization of Emitting Antenna and Field

Electromagnetic waves were generated by a patch antenna (made of copper and FR4 dielectric). Its geometry and dimensions are presented in Figure 1(a), and it emitted a continuous wave on 3.5GHz. This frequency is used in WiMAX communications, based on the IEEE 802.16, and in the first implementations of the 5th generation (5G) mobile broadband communications worldwide. The simulated power of the patch antenna was 0.5W, but simulations in CST showed a Voltage Standing Wave Ratio (VSWR) of 1.27 due to its mismatch loss, and the radiated power from the antenna was 0.138W due to its total efficiency of 28% (Figure 1(b)). The far-field radiation pattern of this antenna at 3.5GHz is presented in Figure 1(c). As it can be observed, the gain is low on the O-z direction (-7.8dBi), therefore the exposure of the brain model is very low.

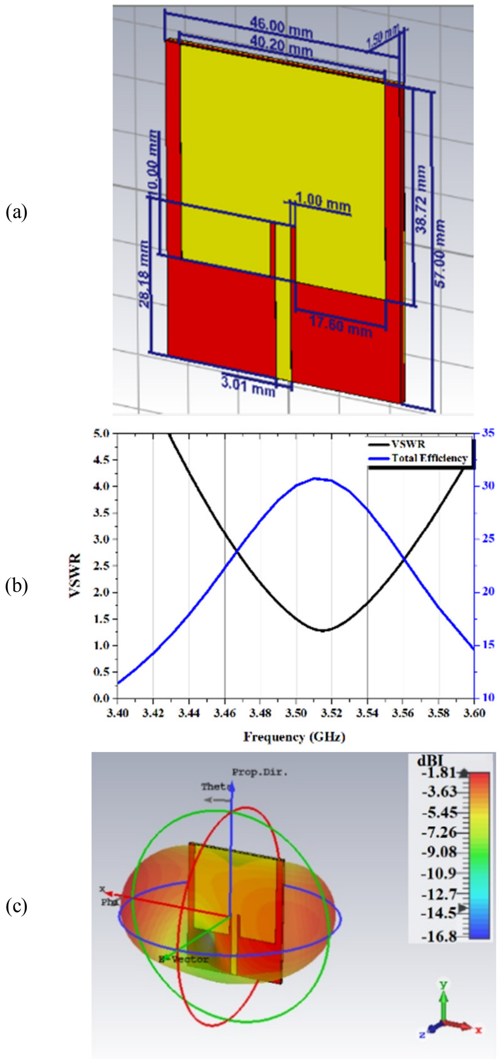


Fig. 1. The emitting patch antenna model: (a) geometric parameters, (b) stationary wave ratio and total efficiency at 3.5GHz, c) 3D radiation pattern at 3.5GHz.

C. Exposure of Non-Magnetic and Magnetic Brain Models

The geometries of the exposure models are shown in Figures 2 and 3. The distance between the antenna surface and the incidence point of the model (brain, in the case of a full head model) was always 45cm, therefore the far-field conditions of exposure were fulfilled in all cases. Figures 2(b)-(c) show the positions of the magnetite particles in the simple homogeneous spherical model of the brain (S1, S2, S3, S4). Figures 3(b) and 4 show the magnetite particles distribution in the realistic brain model (S5, S6). The results of S1-S3 cases can be compared because the mass concentration of magnetite is the same (C1). Similarly, the results of the S4-S6 cases can be compared due to the same concentration of magnetite (C2).

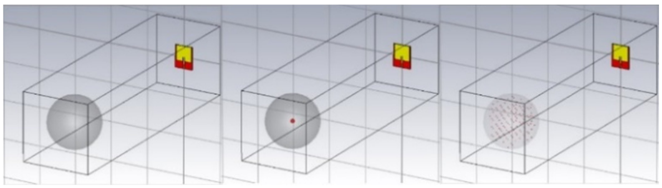

(a) (b) (c)

Fig. 2. The computational space figuring the positions of the antenna, simplified brain model (sphere) and the distribution of the magnetite particles: (a) non-magnetic brain, (b) brain containing a single magnetite particle, (c) brain containing spread magnetite particles with a total volume equal to the single-particle from situation (b).

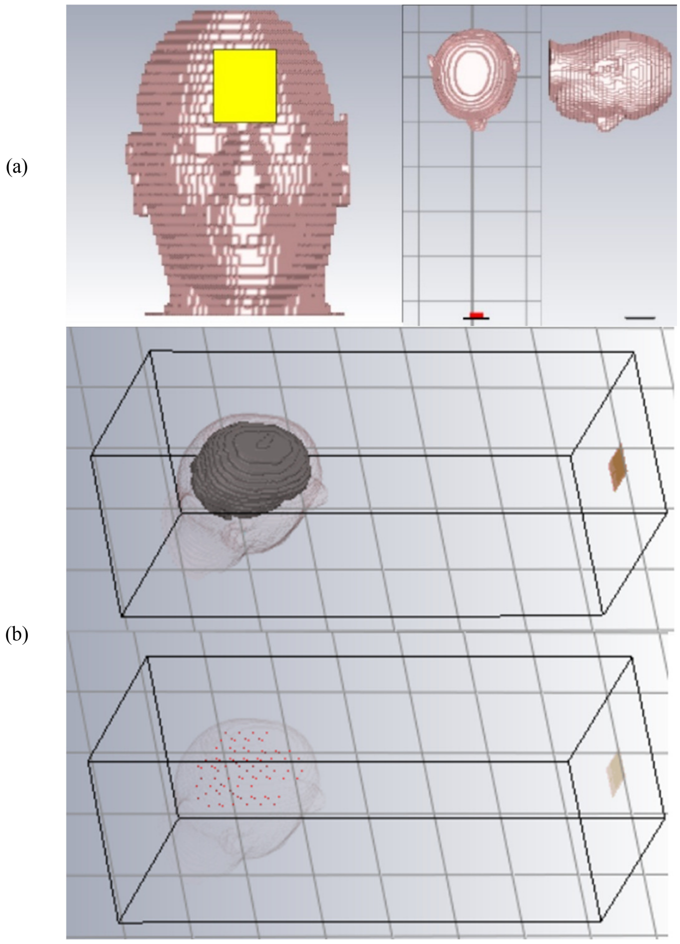


Fig. 3. (a) The human anthropomorphic head model (Gustav) and the positioning of the patch antenna, (b) figuring the non-magnetic brain in the head (up) and the distribution of the magnetic particles in the magnetic brain case.

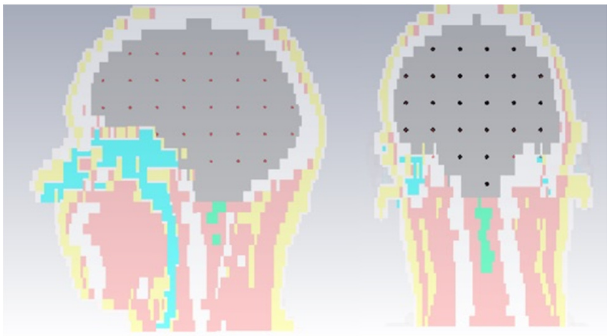


Fig. 4. Details of the distribution (homogeneous) of the magnetite particles in the realistic magnetic brain model.

IV. RESULTS

Figure 5 shows the distribution of the E-field levels in air tangential to the head model. In all S1-S6 cases, the distribution and the absolute values of the E-field strengths in the coronal plane are very similar to Figure 5. It can be observed that the incident field strengths are not exceeding 4.5V/m and they are mostly concentrated in the range between 1-3V/m in the coronal plane.

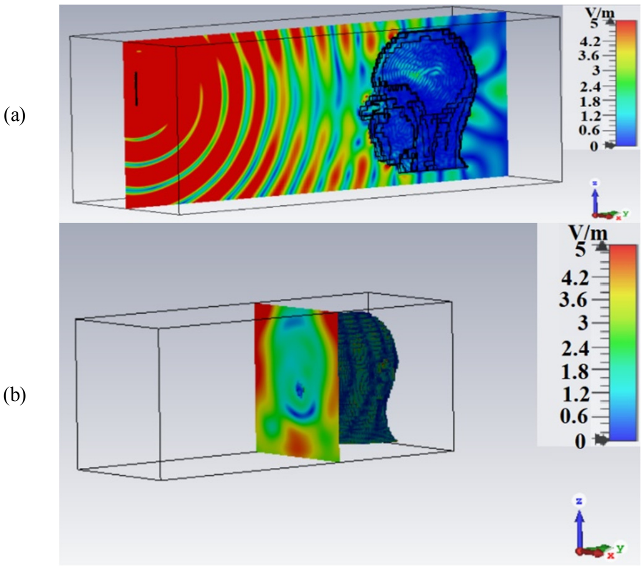


Fig. 5. Incident E-field strength in the air: (a) sagittal plane, (b) coronal plane tangential to the nose of the model.

Table III shows the following post-processed values: a) the mean and the maximum power densities in each model, b) the mean values of the electric and of the magnetic energy densities, c) the mean power loss in each model, d) the normalized power loss (the ratio between the mean power loss in the model and the mean E-field strength in the air, in the coronal plane tangent to the model surface), e) the normalized total SAR (the ratio between the total mean power loss per mass unit of the model and the mean E-field strength in the air, in the coronal plane tangent to the model surface). For example, comparing S3 with S1 suggests that power loss, its normalized value, and normalized SAR are lower when the magnetite is spread in the brain than when not present at all, which is an intriguing result. On the other hand, comparing S6 with S5, when the magnetite concentration is higher, power loss, its normalized value, and normalized SAR are higher when the magnetite is spread in the brain than when not present at all. Therefore, the presence of magnetic material seems not to enlarge the losses consistently.

TABLE III. EFFECTS OF MAGNETITE PRESENCE IN BRAIN EXPRESSED BY STORED ENERGY DENSITY AND LOSSES (INCLUDING NORMALIZED ONES)

Case	$E_{incident}$ (V/m)	Material type	Power loss density (W/m^3)		Mean energy density (J/m^3)		P_{loss} (µW)	Normalized power loss (µW/(V/m))	Average normalized total SAR ((µW/kg)/(V/m))
			Max	Mean	Electric	Magnetic			
S1	4.11	Brain	7.37	0.11	5.61E-11	5.68E-11	227.33	55.27	24.66
S2	4.12	Brain	6.77	0.11	5.61E-11	5.67E-11	227.49	55.27	24.63
		Magnetite	0.77	0.34	3.44E-11	4.01E-11			
S3	4.01	Brain	7.30	0.10	5.35E-11	5.40E-11	216.70	54.03	24.08
		Magnetite	2.18	0.23	2.92E-11	4.12E-11			
S4	3.94	Brain	8.19	0.11	5.98E-11	6.04E-11	131.17	33.27	27.41
		Magnetite	9.36	0.35	4.61E-11	6.04E-11			
S5	3.97	Brain	2.25	0.06	8.47E-11	8.47E-11	70.74	17.82	14.70
S6	3.96	Brain	2.23	0.06	8.40E-11	8.40E-11	70.89	17.89	14.76
		Magnetite	3.53	0.33	4.39E-11	7.05E-11			

Figure 6 shows the sectioning plane orientation in the models that will be used for presenting the distributions in Figures 7-9. Figure 7 shows the distribution of the electric energy density stored in the models, in the chosen section plane, for all 6 cases. In S3 and S4 images, the magnetite particles are emphasized to observe the lower levels of energy density at their positions. S5 is the dielectric brain in the head model, while S6 is the magnetic brain in the head model. The position of the magnetite particles is still observable, as the energy density is lower in their positions than in their vicinity.

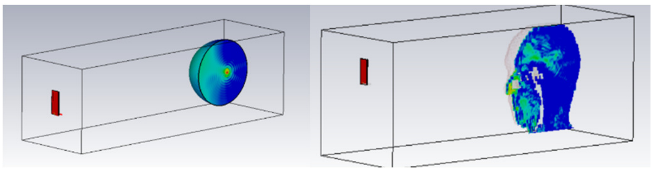


Fig. 6. Figuring the section plane applied to the models to show the spatial distributions of energy density and power loss density.

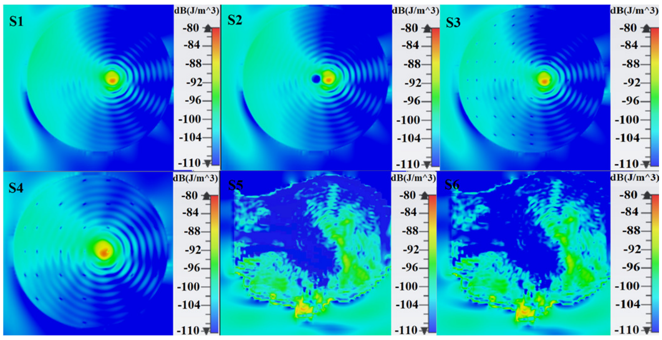


Fig. 7. Electric energy density distribution in the chosen sagittal section plane in the 6 cases.

Figure 8 shows the distribution of the magnetic energy density for all 6 cases. The position of the magnetite particles is still observable in S3 and S4, again, as a bit smaller values than in the neighboring material. No differences appear between S5 and S6 regarding the magnetic energy density accumulation in the dielectric versus the magnetic brain.

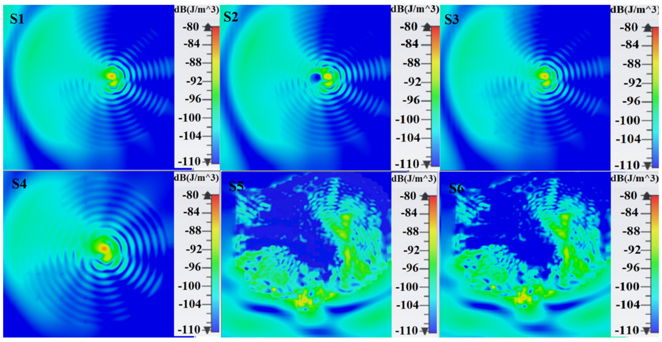


Fig. 8. Magnetic energy density distribution in the chosen sagittal section plane in the 6 situations.

Figure 9 shows the power loss density distribution map in the chosen section plane. The positions of the magnetic particles are delimited visually in all cases. Figure 10 shows an enlarged image for the Gustav head model of case S6. Even with the micromagnetic limits of CST, which is using Polder's permeability tensor to describe the magnetic properties at a larger scale and to compute the losses, obvious excess power deposition is due to the presence of magnetite when the 3.5GHz wave is propagating in the 11-tissue MRI human head model having magnetite particles.

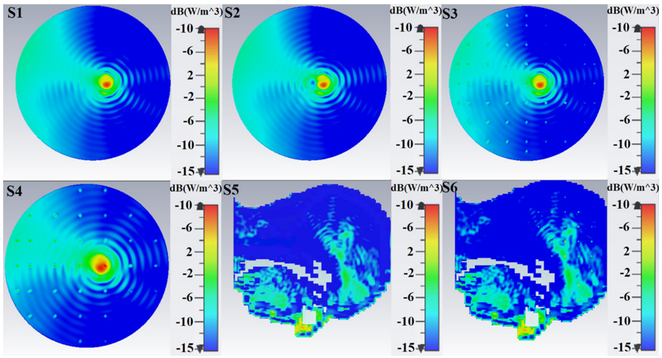


Fig. 9. Power loss density distribution in the chosen sagittal section plane in the 6 cases.

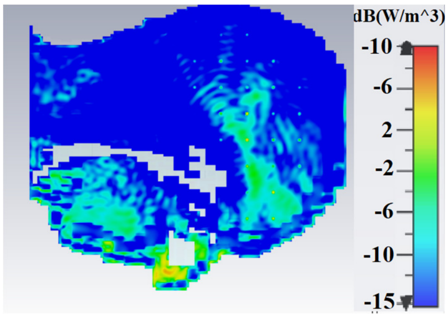


Fig. 10. Magnification of the S6 case to observe the losses due to magnetite particles present in the brain by using the computations based on Polder's permeability tensor as a constitutive relation.

Based on the results tabulated in Table III, emphasis was given to the overall differences between the models, regarding: a) the ratio between the magnetic and electric energy densities accumulated in each model, in tandem with its volume (Figure 11), b) the percentage the magnetic brain is lossier than the pure dielectric (Figure 12), and c) the power loss densities (max and average) in the magnetic and dielectric models (Figure 13). The results were grouped in two sections in each figure, corresponding to the two difference concentrations of magnetite (C1, C2). Figure 11 shows that magnetite concentration is not determining the energy density ratio between the magnetic and electric energies, as possibly higher magnetite content results in lower stored magnetic energy density. Moreover, this ratio is larger in all magnetic brain models compared to purely dielectric. The overall concentration, distribution, and dimensions of magnetite particles matter, while the problem becomes more complex when examining geometric and magnetic anisotropies, particle geometry, local computation of H_{eff}, etc.

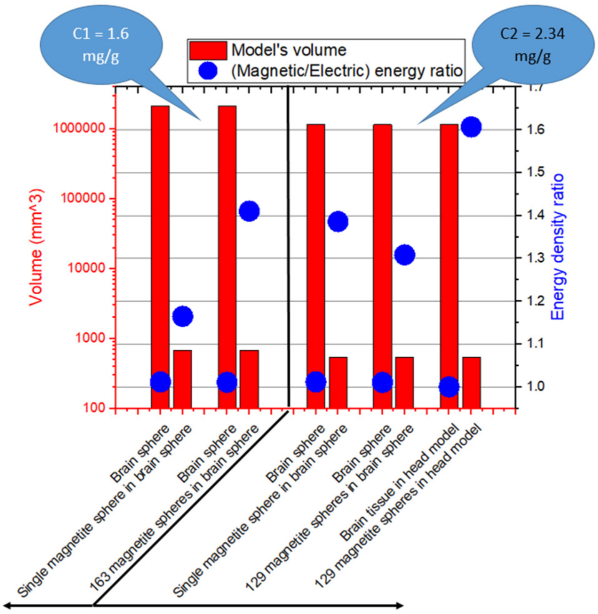


Fig. 11. The ratio of magnetic to electric energy densities stored in pure dielectric versus magnetic brain models (blue balls) in tandem with the volume of the models (red bars). Two categories are delimited due to two different concentrations of magnetite (C1, C2).

Figure 12 shows that for the same concentration of magnetite C1, the brain containing 163 smaller magnetite particles presents a total deposited power even smaller than the dielectric, while the brain containing one single larger magnetite particle may have a larger total deposited power. Practically, among all cases, the larger total power deposition in the magnetic brain happened in Gustav's head. In this case, 0.21% more total power was dissipated in the magnetic brain versus the non-magnetic one. However, the magnetic brain containing 129 spheres of magnetite showed a 5.7% lower total power loss than the similar pure dielectric brain model.

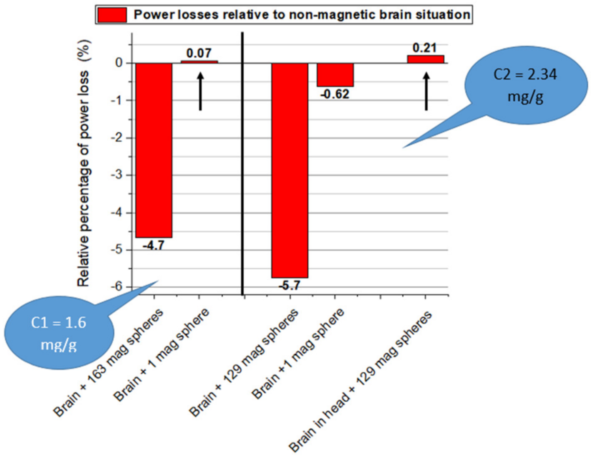


Fig. 12. Relative percentage of total loss deposition in each model: magnetic relative to non-magnetic.

Figure 13 shows comparatively that in most cases the mean power loss density was a little bit higher in the magnetic brain than in the pure dielectric. However, the existence of an exception indicates again the limitations of CST in computing realistically the magnetic losses contribution. Moreover, the larger peak values of power loss density are even more intriguing, as they seem to be present rather in the pure dielectric than in the magnetic brain model.

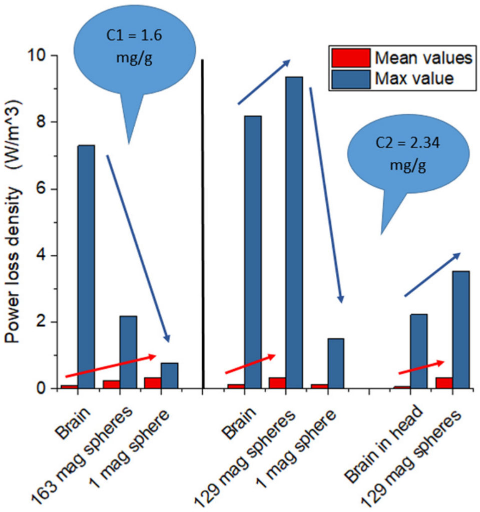


Fig. 13. Average and maximum values of each model's power loss density.

V. CONCLUSION

This paper analyzed a set of simple or respectively anthropomorphic brain models having magnetite content to emphasize the differences in the total power deposited in the tissue when it is considered purely dielectric versus when it is a magnetic-dielectric, with the limitations and the gaps of the dosimetric codes. A patch antenna emitting at 3.5GHz was modeled and positioned 45cm in front of the brain model. Two concentration amounts of magnetite were examined, while its distribution in the brain tissue was modeled either as a unique sphere or a spread of 1mm radius particles having the same total volume. Six dosimetric situations were analyzed comparatively, whereas two of them consisted in the brain belonging to the Gustav human head voxel model.

The results showed both higher and lower total power losses in the magnetic- versus non-magnetic brain models. The relative percentage of the losses on the pure dielectric brain was in the range from -5.7 to +0.21%. The larger the magnetite concentration, the larger the relative differences. If power loss density was normalized to the average E-field strength in a coronal plane tangential to the model, in air, it increased in the case of the magnetic brain of the anthropomorphic model versus the pure dielectric one. The average total SAR normalized to the E-field strength in air showed again that the contribution of the magnetic loss is not consistent when the model is changed. Practically, no clear conclusions could be extracted in connection with the magnetic loss contribution to the whole dissipated power. The results of the mean values of total power loss density showed however more reliability, as the stored magnetic energy density was in all cases dominated by the magnetic brains. Overall, the present approach aimed at underlining the difficulty in getting an accurate solution when requesting absorbed doses of electromagnetic energy in tissues that are not pure dielectrics but have magnetic features. The current computational dosimetry is limited and can only cast a shadow of knowledge about these phenomena. Since experimental dosimetry in the human brain is precluded, it would be of great importance to obtain realistic computational information about the degree of significance that magnetic particles have when they are present in dielectrics and their interaction with signals used by various wireless devices.

REFERENCES

[1] S. Miclaus, M. Racuciu, and and P. Bechet, "H -Field Contribution to the Electromagnetic Energy Deposition in Tissues Similar to the Brain But Containing Ferrimagnetic Particles, During Use of Face-Held Radio Transceivers," *Progress In Electromagnetics Research B*, vol. 73, pp. 49–60, 2017, https://doi.org/10.2528/PIERB17010101.

[2] S. Miclaus, C. Iftode, and and A. Miclaus, "Would the Human Brain Be Able to Erect Specific Effects Due to the Magnetic Field Component of an UHF Field via Magnetite Nanoparticles?," *Progress In Electromagnetics Research M*, vol. 69, pp. 23–36, 2018, https://doi.org/10.2528/PIERM18030806.

[3] J. L. Kirschvink, A. Kobayashi-Kirschvink, and B. J. Woodford, "Magnetite biomineralization in the human brain.," *Proceedings of the National Academy of Sciences*, vol. 89, no. 16, pp. 7683–7687, Aug. 1992, https://doi.org/10.1073/pnas.89.16.7683.

[4] B. A. Maher *et al.*, "Magnetite pollution nanoparticles in the human brain," *Proceedings of the National Academy of Sciences*, vol. 113, no. 39, pp. 10797–10801, Sep. 2016, https://doi.org/10.1073/pnas.1605941113.

[5] L. Calderón-Garcidueñas *et al.*, "Combustion- and friction-derived magnetic air pollution nanoparticles in human hearts," *Environmental Research*, vol. 176, Sep. 2019, Art. no. 108567, https://doi.org/10.1016/j.envres.2019.108567.

[6] D. Lu *et al.*, "Chemical multi-fingerprinting of exogenous ultrafine particles in human serum and pleural effusion," *Nature Communications*, vol. 11, no. 1, May 2020, Art. no. 2567, https://doi.org/10.1038/s41467-020-16427-x.

[7] R. Gieré, "Magnetite in the human body: Biogenic vs. anthropogenic," *Proceedings of the National Academy of Sciences*, vol. 113, no. 43, pp. 11986–11987, Oct. 2016, https://doi.org/10.1073/pnas.1613349113.

[8] Q. Zhang *et al.*, "Separation and Tracing of Anthropogenic Magnetite Nanoparticles in the Urban Atmosphere," *Environmental Science & Technology*, vol. 54, no. 15, pp. 9274–9284, Aug. 2020, https://doi.org/10.1021/acs.est.0c01841.

[9] J. L. Kirschvink, "Microwave absorption by magnetite: A possible mechanism for coupling nonthermal levels of radiation to biological systems," *Bioelectromagnetics*, vol. 17, no. 3, pp. 187–194, 1996, https://doi.org/10.1002/(SICI)1521-186X(1996)17:3<187::AID-BEM4>3.0.CO;2-#.

[10] F. C. Størmer, "Magnetite in dura and pia mater in human brain. A shield against electromagnetic radiation?," *Medical Hypotheses*, vol. 82, no. 1, Jan. 2014, Art. no. 123, https://doi.org/10.1016/j.mehy.2013.11.010.

[11] P. C. Fannin, I. Malaescu, C. N. Marin, and N. Stefu, "Microwave propagation parameters in magnetic fluids," *The European Physical Journal E*, vol. 29, no. 3, pp. 299–303, Jul. 2009, https://doi.org/10.1140/epje/i2009-10477-7.

[12] P. C. Fannin, C. N. Marin, I. Malaescu, and N. Stefu, "Microwave dielectric properties of magnetite colloidal particles in magnetic fluids," *Journal of Physics: Condensed Matter*, vol. 19, no. 3, Jan. 2007, Art. no. 036104, https://doi.org/10.1088/0953-8984/19/3/036104.

[13] E. M. Alfsen, F. C. Størmer, A. Njå, and L. Walløe, "A proposed tandem mechanism for memory storage in neurons involving magnetite and prions," *Medical Hypotheses*, vol. 119, pp. 98–101, Oct. 2018, https://doi.org/10.1016/j.mehy.2018.07.003.

[14] M. A. Busquets, A. Espargaró, R. Sabaté, and J. Estelrich, "Magnetic Nanoparticles Cross the Blood-Brain Barrier: When Physics Rises to a Challenge," *Nanomaterials*, vol. 5, no. 4, pp. 2231–2248, Dec. 2015, https://doi.org/10.3390/nano5042231.

[15] H. Nittby *et al.*, "Nonthermal GSM RF and ELF EMF effects upon rat BBB permeability," *The Environmentalist*, vol. 31, no. 2, pp. 140–148, Jun. 2011, https://doi.org/10.1007/s10669-011-9307-z.

[16] S. Khan and D. Cohen, "Using the magnetoencephalogram to noninvasively measure magnetite in the living human brain," *Human Brain Mapping*, vol. 40, no. 5, pp. 1654–1665, 2019, https://doi.org/10.1002/hbm.24477.

[17] J. Hammond, B. A. Maher, I. A. M. Ahmed, and D. Allsop, "Variation in the concentration and regional distribution of magnetic nanoparticles in human brains, with and without Alzheimer's disease, from the UK," *Scientific Reports*, vol. 11, no. 1, Apr. 2021, Art. no. 9363, https://doi.org/10.1038/s41598-021-88725-3.

[18] International Commission on Non-Ionizing Radio Protection, "Guidelines for Limiting Exposure to Time-varying Electric, Magnetic, and Electromagnetic Fields (up to 300GHz)," *Health Physics*, vol. 74, no. 4, pp. 494–522, Apr. 1998.

[19] "IEEE C95.1-2019 - IEEE Standard for Safety Levels with Respect to Human Exposure to Electric, Magnetic, and Electromagnetic Fields, 0 Hz to 300 GHz." [Online]. Available: https://standards.ieee.org/standard/C95_1-2019.html.

[20] I. Belyaev, "Duration of Exposure and Dose in Assessing Nonthermal Biological Effects of Microwaves," in *Dosimetry in Bioelectromagnetics*, Boca Raton, FL, USA: CRC Press, 2017.

[21] A. A. Pilla, "Nonthermal electromagnetic fields: From first messenger to therapeutic applications," *Electromagnetic Biology and Medicine*, vol. 32, no. 2, pp. 123–136, Jun. 2013, https://doi.org/10.3109/15368378.2013.776335.

[22] L. B. Salford, H. Nittby, and B. R. R. Persson, "Effects of Electromagnetic Fields From Wireless Communication upon the Blood-Brain Barrier," in *A Rationale for Biologically-Based Exposure Standards for Low-Intensity Electromagnetic Radiation (section 10)*, BioInitiative Working Group, 2012.

[23] P. Shrivastava and T. R. Rao, "Investigations of SAR Distributions and Temperature Elevation on Human Body at 60 GHz with Corrugated Antipodal Linear Tapered Slot Antenna," *Progress In Electromagnetics Research M*, vol. 59, pp. 111–121, 2017, https://doi.org/10.2528/PIERM17041707.

[24] T. Hamed and M. Maqsood, "SAR Calculation & Temperature Response of Human Body Exposure to Electromagnetic Radiations at 28, 40 and 60 GHz mmWave Frequencies," *Progress In Electromagnetics Research M*, vol. 73, pp. 47–59, 2018, https://doi.org/10.2528/PIERM18061102.

[25] J. Lan, T. Hong, X. Liang, and and G. Du, "Evaluation of Microwave Microdosimetry for Human Eyes with Glasses Exposed to Wireless Eyewear Devices at Phone Call State," *Progress In Electromagnetics Research M*, vol. 63, pp. 71–81, 2018, https://doi.org/10.2528/PIERM17080802.

[26] K. Y. Yazdandoost and I. Laakso, "Numerical Modeling of Electromagnetic Field Exposure from 5G Mobile Communications at 10 GHz," *Progress In Electromagnetics Research M*, vol. 72, pp. 61–67, 2018, https://doi.org/10.2528/PIERM18070503.

[27] M. Munde, A. Nandgaonkar, and and S. Deosarkar, "Low Specific Absorption Rate Antenna Using Electromagnetic Band Gap Structure for Long Term Evolution Band 3 Application," *Progress In Electromagnetics Research M*, vol. 80, pp. 23–34, 2019, https://doi.org/10.2528/PIERM18102103.

[28] M. Munde, A. Nandgaonkar, and and S. Deosarkar, "Dual Feed Wideband Annular Ring Microstrip Antenna with Circular DGS for Reduced SAR," *Progress In Electromagnetics Research B*, vol. 88, pp. 175–195, 2020, https://doi.org/10.2528/PIERB20071804.

[29] W. E. May, I. Sfar, J. M. Ribero, and and L. Osman, "Design of Low-Profile and Safe Low SAR Tri-Band Textile EBG-Based Antenna for IoT Applications," *Progress In Electromagnetics Research Letters*, vol. 98, pp. 85–94, 2021, https://doi.org/10.2528/PIERL21051107.

[30] K. Mahmoud, A. Baz, W. Alhakami, H. Alhakami, and and A. M. Montaser, "The Performance of Circularly Polarized Phased Sub-Array Antennas for 5G Laptop Devices Investigating the Radiation Effects," *Progress In Electromagnetics Research C*, vol. 110, pp. 267–283, 2021, https://doi.org/10.2528/PIERC21012005.

[31] G. M. J. V. Leeuwen, J. J. W. Lagendijk, B. J. A. M. V. Leersum, A. P. M. Zwamborn, S. N. Hornsleth, and A. N. T. J. Kotte, "Calculation of change in brain temperatures due to exposure to a mobile phone," *Physics in Medicine and Biology*, vol. 44, no. 10, pp. 2367–2379, Aug. 1999, https://doi.org/10.1088/0031-9155/44/10/301.

[32] P. Bernardi, M. Cavagnaro, S. Pisa, and E. Piuzzi, "Specific absorption rate and temperature increases in the head of a cellular-phone user," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 7, pp. 1118–1126, Jul. 2000, https://doi.org/10.1109/22.848494.

[33] A. Drossos, V. Santomaa, and N. Kuster, "The dependence of electromagnetic energy absorption upon human head tissue composition in the frequency range of 300-3000 MHz," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 11, pp. 1988–1995, Nov. 2000, https://doi.org/10.1109/22.884187.

[34] A. Hirata and T. Shiozawa, "Correlation of maximum temperature increase and peak SAR in the human head due to handset antennas," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 7, pp. 1834–1841, Jul. 2003, https://doi.org/10.1109/TMTT.2003.814314.

[35] A. Hirata and T. Shiozawa, "Correlation of maximum temperature increase and peak SAR in the human head due to handset antennas," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 7, pp. 1834–1841, Jul. 2003, https://doi.org/10.1109/TMTT.2003.814314.

[36] A. Hirata, M. Fujimoto, T. Asano, J. Wang, O. Fujiwara, and T. Shiozawa, "Correlation between maximum temperature increase and peak SAR with different average schemes and masses," *IEEE Transactions on Electromagnetic Compatibility*, vol. 48, no. 3, pp. 569–578, Aug. 2006, https://doi.org/10.1109/TEMC.2006.877784.

[37] T. Samaras, E. Kalampaliki, and J. N. Sahalos, "Influence of Thermophysiological Parameters on the Calculations of Temperature Rise in the Head of Mobile Phone Users," *IEEE Transactions on Electromagnetic Compatibility*, vol. 49, no. 4, pp. 936–939, Nov. 2007, https://doi.org/10.1109/TEMC.2007.908257.

[38] A. Hirata and O. Fujiwara, "The correlation between mass-averaged SAR and temperature elevation in the human head model exposed to RF near-fields from 1 to 6 GHz," *Physics in Medicine and Biology*, vol. 54, no. 23, pp. 7227–7238, Nov. 2009, https://doi.org/10.1088/0031-9155/54/23/013.

[39] I. Laakso, R. Morimoto, A. Hirata, and T. Onishi, "Computational Dosimetry of the Human Head Exposed to Near-Field Microwaves Using Measured Blood Flow," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 2, pp. 739–746, Apr. 2017, https://doi.org/10.1109/TEMC.2016.2633326.

[40] R. Morimoto, I. Laakso, V. D. Santis, and A. Hirata, "Relationship between peak spatial-averaged specific absorption rate and peak temperature elevation in human head in frequency range of 1–30 GHz," *Physics in Medicine and Biology*, vol. 61, no. 14, pp. 5406–5425, Jul. 2016, https://doi.org/10.1088/0031-9155/61/14/5406.

[41] *CST Studio Suite (2019)*. CST Microwave Studio.

[42] S. Ghnimi, A. Nasri, and A. Gharsallah, "Study of a New Design of the Planar Inverted-F Antenna for Mobile Phone Handset Applications," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5270–5275, Feb. 2020, https://doi.org/10.48084/etasr.3287.

[43] S. K. Bitra and S. Miriyala, "An Ultra-Wideband Band Pass Filter using Metal Insulator Metal Waveguide for Nanoscale Applications," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7247–7250, Jun. 2021, https://doi.org/10.48084/etasr.4194.

[44] K. Mekki, O. Necibi, C. Boussetta, and A. Gharsallah, "Miniaturization of Circularly Polarized Patch Antenna for RFID Reader Applications," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5655–5659, Jun. 2020, https://doi.org/10.48084/etasr.3445.

[45] M. Jazirehpour and S. A. Seyyed Ebrahimi, "Effect of aspect ratio on dielectric, magnetic, percolative and microwave absorption properties of magnetite nanoparticles," *Journal of Alloys and Compounds*, vol. 638, pp. 188–196, Jul. 2015, https://doi.org/10.1016/j.jallcom.2015.03.021.

[46] M. Jazirehpour and S. A. S. Ebrahimi, "Synthesis of magnetite nanostructures with complex morphologies and effect of these morphologies on magnetic and electromagnetic properties," *Ceramics International*, vol. 42, no. 15, pp. 16512–16520, Nov. 2016, https://doi.org/10.1016/j.ceramint.2016.07.067.

[47] H. Khurshid *et al.*, "Anisotropy effects in magnetic hyperthermia: A comparison between spherical and cubic exchange-coupled FeO/Fe3O4 nanoparticles," *Journal of Applied Physics*, vol. 117, no. 17, May 2015, Art. no. 17A337, https://doi.org/10.1063/1.4919250.

[48] X. Liu *et al.*, "Shape-dependent magnetic and microwave absorption properties of iron oxide nanocrystals," *Materials Chemistry and Physics*, vol. 192, pp. 339–348, May 2017, https://doi.org/10.1016/j.matchemphys.2017.02.012.

[49] S. Couture, X. Wang, A. Goncharov, and V. Lomakin, "A coupled micromagnetic-Maxwell equations solver based on the finite element method," *Journal of Magnetism and Magnetic Materials*, vol. 493, Jan. 2020, Art. no. 165672, https://doi.org/10.1016/j.jmmm.2019.165672.

[50] J. Leliaert and J. Mulkers, "Tomorrow's micromagnetic simulations," *Journal of Applied Physics*, vol. 125, no. 18, May 2019, Art. no. 180901, https://doi.org/10.1063/1.5093730.

[51] S. Couture and V. Lomakin, "Electromagnetic-micromagnetic simulator for magnetization-eddy current dynamics in magnetic materials and devices," in *2017 IEEE International Symposium on Antennas and Propagation USNC/URSI National Radio Science Meeting*, Jul. 2017, pp. 1117–1118, https://doi.org/10.1109/APUSNCURSINRSM.2017.8072601.

[52] Z. Yao, R. U. Tok, T. Itoh, and Y. E. Wang, "A Multiscale Unconditionally Stable Time-Domain (MUST) Solver Unifying Electrodynamics and Micromagnetics," *IEEE Transactions on*

Microwave Theory and Techniques, vol. 66, no. 6, pp. 2683–2696, Jun. 2018, https://doi.org/10.1109/TMTT.2018.2825373.

[53] H. Mosallaei and K. Sarabandi, "Magneto-dielectrics in electromagnetics: concept and applications," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 6, pp. 1558–1567, Jun. 2004, https://doi.org/10.1109/TAP.2004.829413.

[54] Q. Nguyen and A. I. Zaghloul, "Susceptibility of Nanoparticles Studied by Landau-Lifshitz-Gilbert and Snoek's Equations," in *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, Jul. 2019, pp. 1299–1300, https://doi.org/10.1109/APUSNCURSINRSM.2019.8888631.

[55] J. Leliaert, A. Vansteenkiste, A. Coene, L. Dupré, and B. Van Waeyenberge, "Vinamax: a macrospin simulation tool for magnetic nanoparticles," *Medical & Biological Engineering & Computing*, vol. 53, no. 4, pp. 309–317, Apr. 2015, https://doi.org/10.1007/s11517-014-1239-6.

[56] Ansys HFSS - 3D High Frequency Electromagnetic Simulation Software. Ansys.

[57] I. Kong, S. Hj Ahmad, M. Hj Abdullah, D. Hui, A. Nazlim Yusoff, and D. Puryanti, "Magnetic and microwave absorbing properties of magnetite–thermoplastic natural rubber nanocomposites," *Journal of Magnetism and Magnetic Materials*, vol. 322, no. 21, pp. 3401–3409, Nov. 2010, https://doi.org/10.1016/j.jmmm.2010.06.036.

[58] "Tissue dielectric properties," *IT'IS Foundation.* https://itis.swiss/virtual-population/tissue-properties/database/dielectric-properties (accessed Sep. 28, 2021).

A Deep Learning-based Mobile Application for Segmenting Tuta Absoluta's Damage on Tomato Plants

Loyani Loyani
School of Computational and Communication Science and Engineering
The Nelson Mandela African Institution of Science and Technology
Arusha, Tanzania
loyanil@nm-aist.ac.tz

Dina Machuve
School of Computational and Communication Science and Engineering
The Nelson Mandela African Institution of Science and Technology
Arusha, Tanzania
dina.machuve@nm-aist.ac.tz

Abstract-With the advances in technology, computer vision applications using deep learning methods like Convolutional Neural Networks (CNNs) have been extensively applied in agriculture. Deploying these CNN models on mobile phones is beneficial in making them accessible to everyone, especially farmers and agricultural extension officers. This paper aims to automate the detection of damages caused by a devastating tomato pest known as Tuta Absoluta. To accomplish this objective, a CNN segmentation model trained on a tomato leaf image dataset is deployed on a smartphone application for early and real-time diagnosis of the pest and effective management at early tomato growth stages. The application can precisely detect and segment the shapes of Tuta Absoluta-infected areas on tomato leaves with a minimum confidence of 70% in 5 seconds only.

Keywords-mobile applications for agriculture; Tuta Absoluta; deep learning; convolutional neural networks; segmentation

I. INTRODUCTION

Agriculture is an important economic sector of the Tanzanian economy, contributing about 29.1% of Gross Domestic Product (GDP) and 67% of total employment [1]. Tanzanian farmers grow a variety of crops for food and economic purposes. As one of the most widely grown crops in the world [2], the tomato plant is also grown in different parts of Tanzania. In 2017, 247,135 tons of tomatoes were harvested in 54,520 hectares, which is equivalent to 64% of all fruits and vegetables in the country [1]. Tomato (Solanum Lycopersicum L.) is considered a high-value crop and income resource for smallholder farmers in Sub-Saharan Africa [3]. Nevertheless, the invasive pest known as Tuta Absoluta is a major threat to tomato production [4]. It can greatly damage tomato yield to the extent that growers may give up production due to the costs and losses it causes. Yield losses can reach percentages as high as 80-100% if no control measures are taken [5]. The pest is native to South America but has spread quickly, not only across the Mediterranean basin but also across Europe, Middle East, Asia, and Africa, where it was first recorded in Algeria in 2008 [6-8]. The first case of Tuta Absoluta in Tanzania was recorded in August 2014 in Arumeru District, Arusha, and it has since spread to the other regions of the country [9]. The number of extension officers who are key facilitators in providing farmers with proper information on plant diseases and pest control is very limited to meet the farmers' demands in Tanzania [10]. Although farmers and extension officers struggle with different methods to control the pest, there has not yet been an effective mechanism to exploit the Tuta Absoluta's infestation on tomato leaves at early stages before causing great yield losses to farmers.

The application of Artificial Intelligence (AI) plays an important role in precision agriculture due to its flexibility, high performance, accuracy, and cost efficiency [11-13]. Computer vision techniques, such as deep Convolutional Neural Networks (CNNs) have shown promise to transform the agricultural field in plant disease diagnosis. Several researchers developed CNN models using image datasets for pest and disease diagnosis in plants like tomato, banana, apple, cassava, cherry, alfalfa, wheat, and grapevine [14-21]. Deploying these CNN models on mobile phones would be beneficial in making them accessible to everyone, especially farmers and agricultural extension officers. Despite the widespread use of smartphones, CNN models are currently being deployed only on a few smartphone applications. The use of mobile applications to diagnose plant diseases has received little attention in the literature. For instance, Petrellis [22] developed a windows smartphone application using image processing techniques for identifying vineyard diseases by extracting features from grape leaf images. It can be used as a standalone application or combined with a remote server, and it is also extensible to different plant diseases. Accuracy higher than 90% was achieved in recognizing the vineyard diseases. As an extension of their previous works [22-24], the authors in [25] proposed a low complexity image processing and a classification method implemented on smartphones for disease diagnosis in grape, peach, and citrus plants with an extensible set of diseases. Classification algorithms such as decision tree, random forest, Naïve Bayes, and neural networks were used in

---

Corresponding author: Loyani Loyani

their experiments. The experimental results show that the developed mobile application achieved accuracy between 80 to 98% for disease recognition. Moreover, authors in [26] developed an image processing algorithm deployed in smartphones for detecting diseases in the wheat plant using a dataset of 3637 leaf images. The smartphone application was used to capture wheat leaf images that were then diagnosed online by comparison with stored previous algorithm results with accuracy higher than 80%.

Similarly, the authors in [27] developed a CNN model based on Single Shot Multibox (SSD) architecture then deployed it on a mobile application for real-time detection of diseases in cassava plants. The model performed better with F1-scores of 79% and 54% on test data and real-world images respectively. Authors in [28] proposed two deep CNNs (Resnet152 and Inceptionv3) using a dataset of 3000 leaf images to detect Fusarium wilt race 1 and black Sigatoka banana diseases. The Inceptionv3 model was deployed on an Android mobile phone and achieved an accuracy of 99% in detecting the two diseases in real environment. Authors in [29] proposed an automated irrigation and plant-leaf disease detection system using an Android mobile application. The mobile application was used to take photos of the suspected plant leaves, which were then sent to a cloud server, where the image was processed by matching it to the plant leaf images with diseases stored in the cloud database. In this work, the classification was based on Artificial Neural Networks (ANNs). In a recent study, researchers reported the performance of a deep learning object detection model for diagnosing plant diseases and pests [30]. The model was trained on 2756 images of cassava leaves exhibiting pest and disease symptoms. It was subsequently deployed as a mobile application in Android smartphones, which was then tested in the field and proved to be 74-88% accurate in diagnosing cassava disease and pest symptoms. This diagnosis accuracy was higher than that of agricultural extension officers (40-58%) and farmers (18-31%). Authors in [31] trained 5 CNN models using the PlantVillage dataset [32] with 18,160 tomato leaf images to classify 10 labels. The CNN model ResNet50, which had better prediction accuracy, was subsequently deployed in a mobile application to classify and identify tomato plant diseases successfully.

Early and real-time detection of Tuta Absoluta's damages on tomato plants can play a vital role in managing the pest and enhancing farmers' decisions. This study presents a mobile application called TutaSegmenter [33] deployed with a deep CNN model to detect and segment the effects of a tomato leaf miner (Tuta Absoluta) on tomato plants. This intervention approach for early pest detection and effective management at primary tomato growth stages will help farmers avoid massive economic losses.

II. METHODS

We used TensorFlow framework to deploy the CNN semantic segmentation model in a smartphone to detect Tuta Absoluta's damage on tomato plants. We employed transfer learning based on the U-Net architecture which has shown best performance on the International Symposium on Biomedical Imaging (ISBI) dataset [34] for semantic segmentation. Authors in [35] introduced this U-shaped CNN architecture, which has performed exceedingly well in the biomedical image segmentation and later in many other fields, outperforming earlier segmentation methods even after being trained with only a few images [36]. We fine-tuned the model parameters to our dataset, which comprised of 1212 tomato leaf images damaged with Tuta Absoluta at early growth stages. The tomato leaf image dataset was built with images collected from experimental fields in Arusha and Morogoro regions in Tanzania. Examples of leaf images in this dataset are shown in Figure 1.

Fig. 1. Sample images from the tomato leaf image dataset.

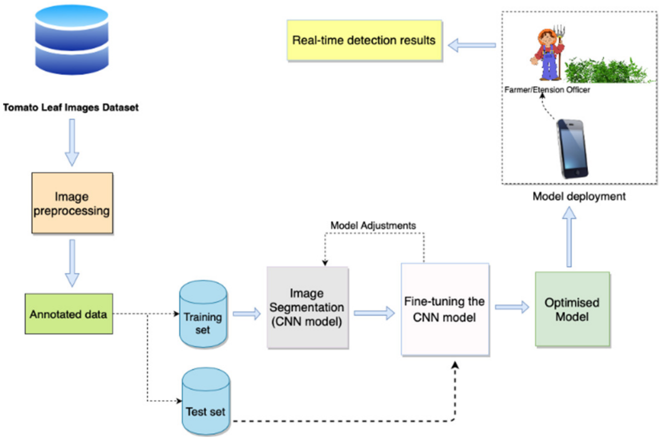


Fig. 2. Conceptual framework.

The experiments were conducted on a computer preinstalled with Windows 10 equipped with one Intel® Core™ i7-8550U 3.6GHz CPU, Intel® Iris® Plus Graphics, 512GB SSD storage, and 16GB memory. Google Collaboratory with Tesla P100-PCIE GPU and 27GB memory was utilized. We implemented our proposed network using Python v3.7.10 and the Keras [37] module with TensorFlow [38] as backend. The model's performance was then evaluated using different evaluation metrics, and the model's parameters were tuned to get an optimized model. The optimized CNN model was deployed on a Samsung Galaxy A02s Android phone to enable farmers to automatically detect and segment the affected areas on tomato plants. Figure 2 shows the conceptual framework of this work.

A. Data Preprocessing

The dataset of healthy and Tuta Absoluta-infested tomato leaf images was collected from the fields in Tanzania using Canon EOS Kiss X7 and Samsung SM-G570F cameras. The complete details of this dataset have been reported in [39] and are freely available to the research community at the open access repository in [40]. For this study, 1212 images of infested tomato plants were extracted from the dataset to develop the semantic segmentation model. Each sample plant contained an average of 6 leaves which is equivalent to 7272 leaflets for the selected images in our dataset. LabelMe [41], an open-source graphical annotation tool (see Figure 3), was then used to produce ground truth masks by manually drawing and labeling irregular polygons following the shape of the infested area. Annotations were then saved in VOC [42] format with their corresponding images. Then, the dataset was split into training and test sets in 80:20 ratio as shown in Table I.

Fig. 3. LabelMe annotation tool.

TABLE I. DATA DISTRIBUTION

Image set	Number of images
Training set	969
Test set	243
Total	1212

B. The CNN Model

The approach presented in this paper is an extension of the method described in [39] where a Tuta Absoluta semantic segmentation model based on U-Net architecture was proposed. In this work, a U-Net architecture was trained for 200 epochs using an annotated dataset of 1212 images resized to 512×512 pixels. The learning rate was set to 0.01 using the Adam [42] optimization function and the sigmoid activation function. The cross-entropy loss, Intersection over Union (IoU), and dice coefficient metrics were used to evaluate the model's performance. Then, an optimized Tuta Absoluta semantic segmentation model was deployed in an Android smartphone.

C. Model Deployment

This refers to integrating a machine learning model into an existing production environment to make practical decisions. We deployed the proposed model into a mobile phone so that farmers and extension officers can use their smartphones to automatically detect affected areas on tomato plants. To deliver reliable, user-friendly, and efficient software in a short time, agile software development methodology was used in developing the Tuta Absoluta segmentation mobile application based on users' feedback. We used the TensorFlow Lite (TFLite) to convert our suggested model into a lighter format because CNN models are complex and heavy, requiring a lot of memory and storage space to run. Converting the models into TFLite format reduces their file size, increases execution speed, and introduces optimizations that do not affect accuracy. This enables us to execute CNN models efficiently and run inference on mobile devices with limited computing and memory resources. The CNN model was converted to a mobile-compatible format using the python programming language. The application's interface and functionality were then defined using Extensible Markup Language (XML) and Kotlin programming language respectively. The mobile application software was implemented in Android Studio Integrated Development Environment (IDE) and tested in a Samsung SM-A025F Galaxy A02s smartphone. Smartphone applications are becoming increasingly popular and necessary [44]. We developed an Android application since Android holds the bigger market share in Africa of over 84% [45], where this research focuses.

Moreover, to achieve its purpose, the Tuta Absoluta segmentation mobile application was designed to capture and upload images, provide general information about tomatoes and the Tuta Absoluta pest, run inference using the CNN segmentation model, and display the segmentation results. Since the application operates completely offline, it only takes 5s to run the inference on a captured tomato leaf image and accurately segment the Tuta mines. The Tuta Absoluta segmentation mobile application is simple and easy to use, even without any guidance.

D. Software Design

In this paper, the design of the mobile application software is visualized using Unified Modelling Language (UML) diagrams. This includes the use case, activity and sequence diagrams. A use case model outlines how different categories of users interact with the system, their expectations, and the actions that the system must take to meet these objectives. Figure 4 depicts the use case diagram for the Tuta Absoluta segmentation mobile application. Each of the use cases shown is described in Table II.

Moreover, the activity diagram describes the dynamic aspect of the application software. It graphically represents a series of actions or flow of control in a system with support for iteration and concurrency. The activity diagram for Tuta Absoluta segmentation mobile application is shown in Figure 5. Also, the sequence diagram depicts how objects interact with each other for a particular scenario. It details the way operations in the application software are carried out. Figure 6 shows the sequence diagram for the Tuta Absoluta segmentation mobile application.

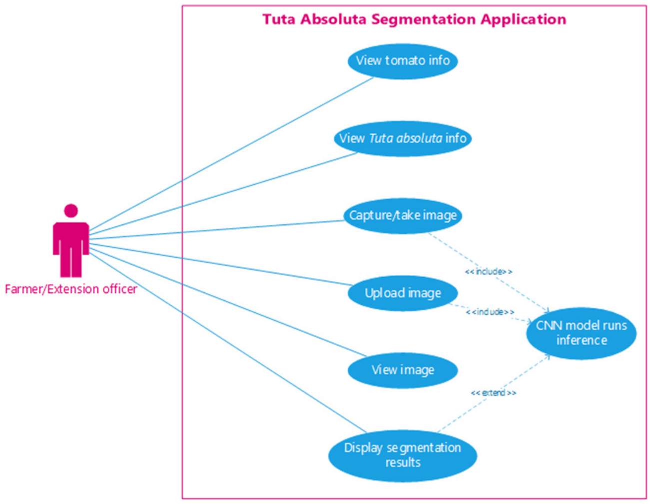


Fig. 4. The use case diagram for Tuta Absoluta segmentation application.

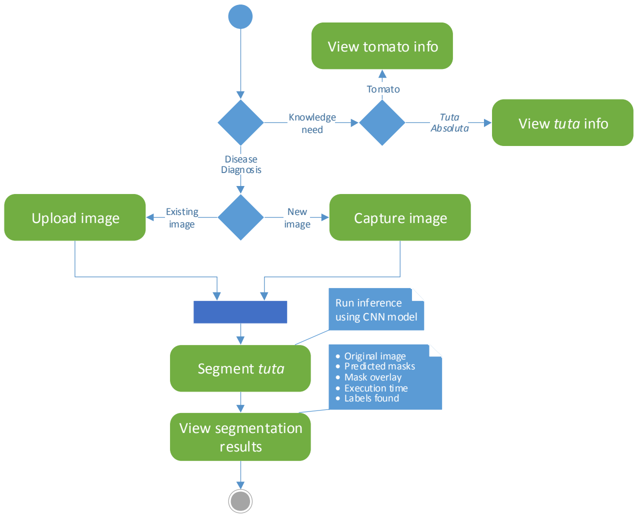


Fig. 5. The activity diagram for Tuta Absoluta segmentation application.

TABLE II. DESCIPTION OF THE USE CASES

Use case	Description	Actor (s)
View tomato info	The user(s) can display general information about tomatoes such as the scientific name, production statistics, and planting information.	Farmer or extension officer
View Tuta absoluta info	The user(s) can display general information about Tuta Absoluta such as their common and scientific names, physiology, and life cycle.	Farmer or extension officer
Capture/take photo	The user(s) can access their mobile phone’s camera to take a photo of a tomato plant. Then the system will automatically run inference on the photo using the CNN model in the background to segment Tuta mines.	Farmer or extension officer
Upload an image	The user(s) can upload a tomato plant image from their mobile phone’s gallery. Then the system will automatically run inference on the uploaded photo using the CNN model in the background to segment tuta mines.	Farmer or extension officer
View image	The user(s) can view the captured or uploaded image in the mobile application.	Farmer or extension officer
Display segmentation results	The user(s) can display the original image, segmentation results, and overlay in the mobile application.	Farmer or extension officer

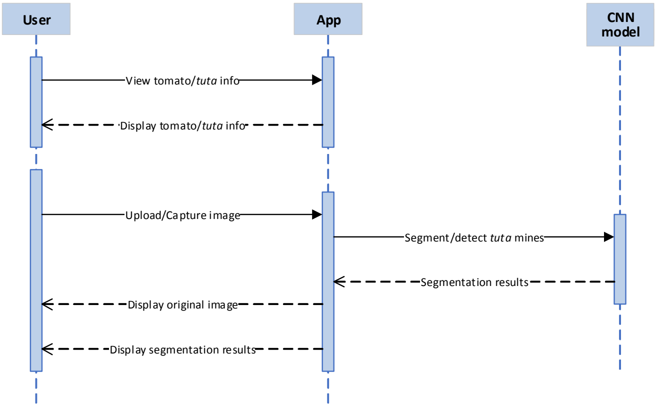


Fig. 6. The sequence diagram for Tuta Absoluta segmentation application.

III. RESULTS AND DISCUSSION

The model’s performance was evaluated using the cross-entropy loss, IoU and dice coefficient evaluation metrics. The experimental results show that the U-Net segmentation model achieved 78.60% and 82.86% of IoU and dice coefficient, respectively. Figure 7 shows the U-Net training loss curve over 200 epochs. The losses dropped rapidly during early training iterations before stabilizing around 60 epochs, indicating that the model fits well on the characteristics of our dataset at both the early and late stages of the training process.

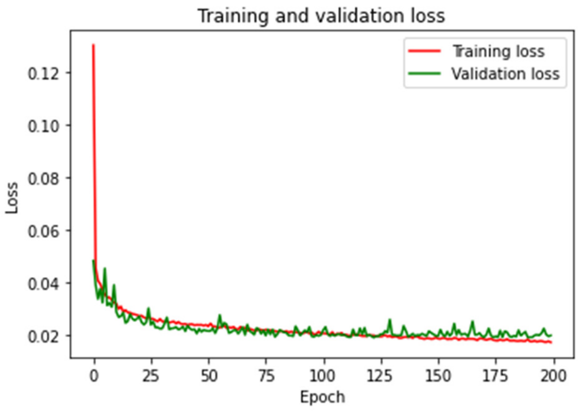


Fig. 7. Training and validation loss for U-Net.

Figure 8 illustrates the confusion matrix for our segmenter. Since our task is binary segmentation, we map 0 for the background class and 1 for Tuta, the object of interest. As mentioned above, the U-Net model was converted to TFLite format and then embedded into a mobile application in Android Studio. Deploying on mobile devices would help democratize model access while also protecting user privacy that runs inference offline.

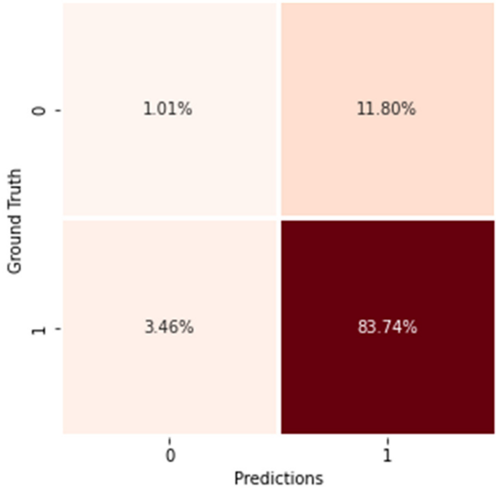
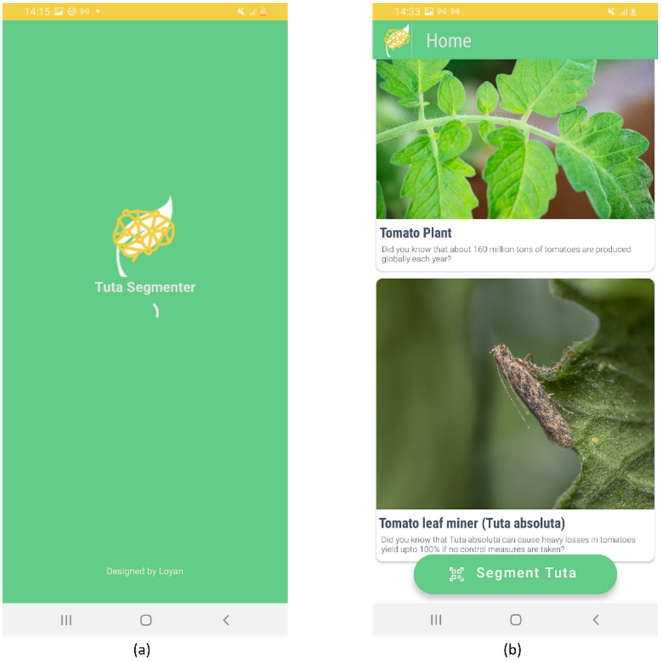


Fig. 8. Confusion matrix for the Tuta Absoluta segmentation model.



Tomato Plant

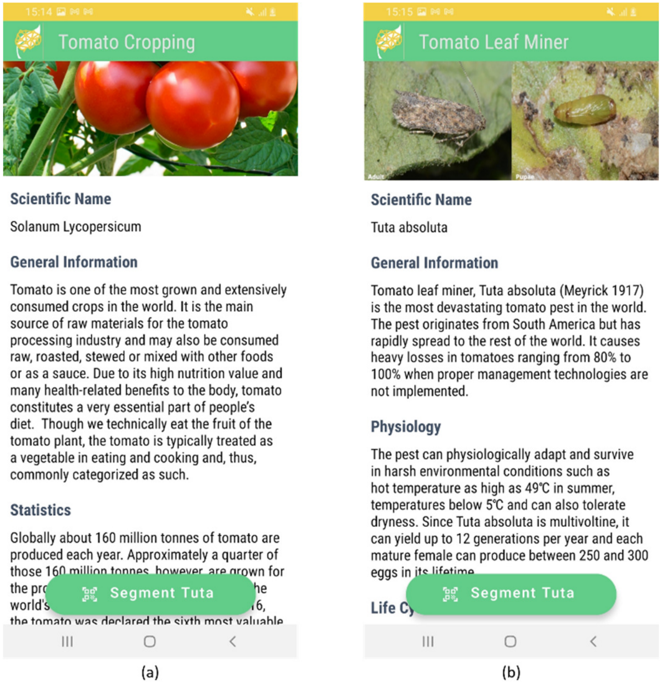
Did you know that about 160 million tons of tomatoes are produced globally each year?

Tomato leaf miner (Tuta absoluta)

Did you know that Tuta absoluta can cause heavy losses in tomatoes yield upto 100% if no control measures are taken?

Fig. 9. Tuta absoluta segmentation mobile application: (a) splash screen, (b) landing/home page.

We developed a simple and user-friendly mobile application named TutaSegmenter [33] to allow smooth interaction between the farmers/extension officers and the application. The application can be downloaded, free of charge, from the Google Playstore. The farmer/extension officer clicks on the application icon which welcomes them with a colorful splash screen that lasts only for 2s before landing on the scrollable home page. The home screen has two clickable cards with quick facts about the tomato plant and Tuta Absoluta, and a clickable floating button for Tuta segmentation, as illustrated in Figure 9. The user can read general information about tomatoes, such as their scientific name, production statistics, and cropping information such as the amount of water, soil, and fertilizer they require by clicking on the tomato plant card, as shown in Figure 10(a). By selecting the Tuta Absoluta card from the main page, the user can view general information about the pest, including common and scientific names, physiology, and lifecycle (see Figure 10(b)). This will assist the farmer in learning and comprehending the pest to adopt effective pest management and tomato plant care procedures.



Scientific Name

Solanum Lycopersicum

General Information

Tomato is one of the most grown and extensively consumed crops in the world. It is the main source of raw materials for the tomato processing industry and may also be consumed raw, roasted, stewed or mixed with other foods or as a sauce. Due to its high nutrition value and many health-related benefits to the body, tomato constitutes a very essential part of people's diet. Though we technically eat the fruit of the tomato plant, the tomato is typically treated as a vegetable in eating and cooking and, thus, commonly categorized as such.

Statistics

Globally about 160 million tonnes of tomato are produced each year. Approximately a quarter of those 160 million tonnes, however, are grown for the pr he world's 6, the tomato was declared the sixth most valuable

Scientific Name

Tuta absoluta

General Information

Tomato leaf miner, Tuta absoluta (Meyrick 1917) is the most devastating tomato pest in the world. The pest originates from South America but has rapidly spread to the rest of the world. It causes heavy losses in tomatoes ranging from 80% to 100% when proper management technologies are not implemented.

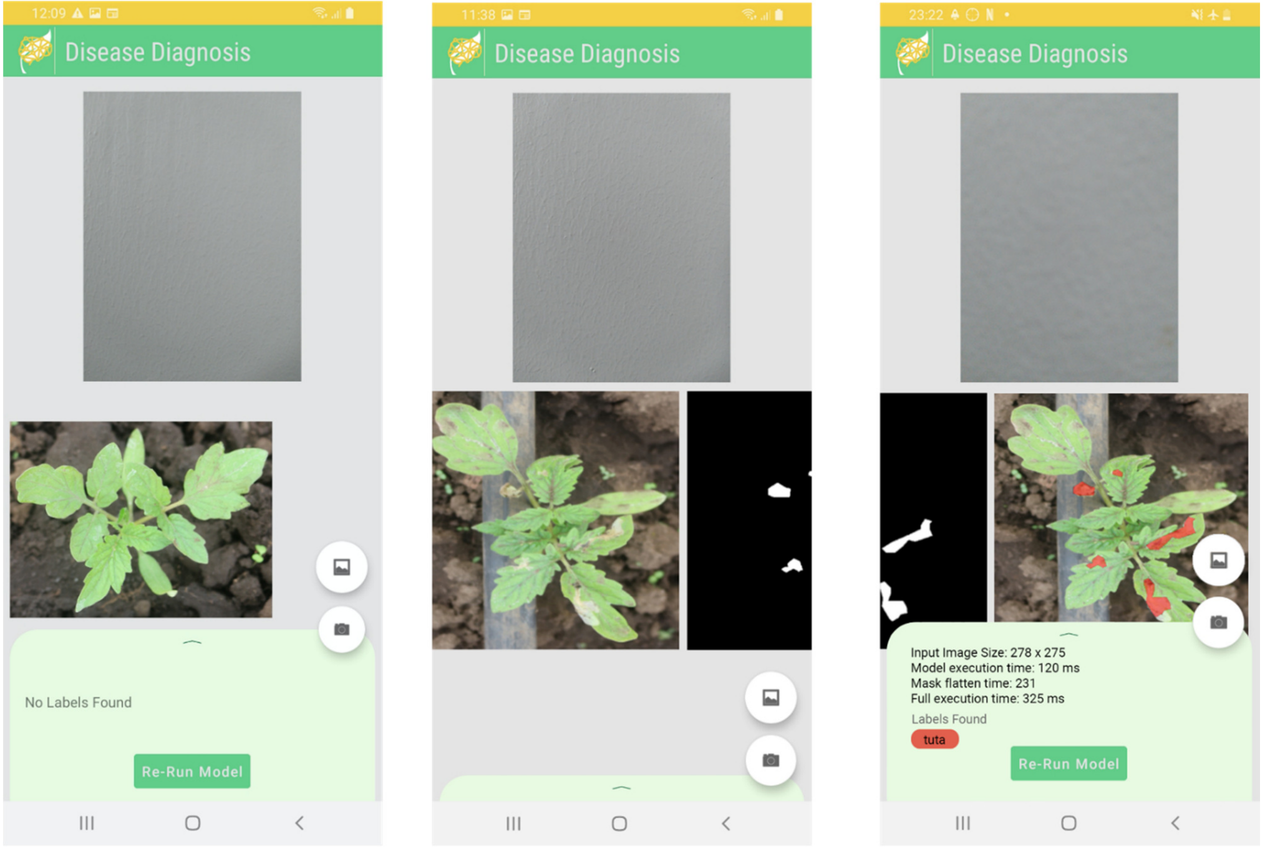
Physiology

The pest can physiologically adapt and survive in harsh environmental conditions such as hot temperature as high as 49℃ in summer, temperatures below 5℃ and can also tolerate dryness. Since Tuta absoluta is multivoltine, it can yield up to 12 generations per year and each mature female can produce between 250 and 300 eggs in its lifetime.

Life Cy

Fig. 10. Tuta Absoluta segmentation mobile application: (a) description page for tomato cropping, (b) description page for Tuta Absoluta.

Additionally, on every page of the mobile application, we placed a clickable floating button for Tuta segmentation so that the farmer can quickly navigate to the disease diagnostics page to detect and segment Tuta mines in tomato leaf images. The disease diagnostics page allows the user to capture or upload images, then shows the results (predicted masks and masks overlay) displayed in a horizontally scrollable section along with the original image. The user can also view the input image size and detection details (labels found and the execution time) in the bottom sheet layout (see Figure 11(c)). If the image is devoid of Tuta mines, the application will display a text "No labels found" as shown in Figure 11(a).



No Labels Found

Input Image Size: 278 x 275
Model execution time: 120 ms
Mask flatten time: 231
Full execution time: 325 ms
Labels Found

(a) (b) (c)

Fig. 11. Tuta Absoluta segmentation mobile application. (a) A healthy plant. (b) Original image and mask prediction. (c) Segmentation results.

IV. CONCLUSION AND FUTURE WORK

This paper presents a cost-effective mobile application for real-time detection and segmentation of Tuta Absoluta's damage on tomato plants at their early growth stages. We utilized a CNN model trained on a 1212 tomato leaf images dataset and optimized it for mobile deployment. We evaluated the performance of the deployed model on a mobile application that detects Tuta mines in tomato leaf images with minimum detection confidence of 70%. The application was able to precisely detect and segment the shapes of Tuta Absoluta-infected areas on tomato leaves in just 5s. Since the deployed model operates offline, there are no running costs once installed. The application provides early and real-time diagnosis of the devastating pest, and effective management at early tomato growth stages to avoid huge economic losses. Providing these services in the absence of a mobile network, makes the application ideal for use in remote areas, which is beneficial to smallholder farmers especially in developing countries. This study proves that the use of computer vision techniques is vital in revolutionizing disease and pest management in agriculture.

In the future, we expect to continuously improve the diagnostic model in Tuta Absoluta segmentation application and add more knowledge as it becomes available to provide farmers with an ever-improving link to expert knowledge. Also, we aim to make the application capable of determining the extent of damage and suggest appropriate actions to be taken to control the pest based on the severity status. The application will be translated to Swahili, a widely spoken language in the Eastern, Central and Southern parts of Africa.

ACKNOWLEDGEMENT

The authors would like to thank the African Development Bank (AfDB) for funding this work through Project No. P-Z1-IA0-016 under Grant No. 2100155032816.

REFERENCES

[1] *Annual Agriculture Sample Survey Crop and Livestock Report*. Tanzania: Ministry of Agriculture, 2017.

[2] S. R. Rupanagudi, B. S. Ranjani, P. Nagaraj, V. G. Bhat, and T. G, "A novel cloud computing based smart farming system for early detection of borer insects in tomatoes," in *International Conference on Communication, Information & Computing Technology*, Mumbai, India, Jan. 2015, pp. 1–6, https://doi.org/10.1109/ICCICT.2015.7045722.

[3] V. Mutayoba and D. Ngaruko, "Assessing Tomato Farming and Marketing Among Smallholders in High Potential Agricultural Areas of Tanzania," *International Journal of Economics, Commerce and Management*, vol. 6, no. 8, pp. 577–590, 2018.

[4] A. H. R. Gonring, A. H. Walerius, M. M. Picanco, L. Bacci, J. C. Martins, and M. C. Picanco, "Feasible sampling plan for Tuta absoluta egg densities evaluation in commercial field tomato," *Crop Protection*, vol. 136, Oct. 2020, Art. no. 105239, https://doi.org/10.1016/j.cropro.2020.105239.

[5] N. Desneux *et al.*, "Biological invasion of European tomato crops by Tuta absoluta: ecology, geographic expansion and prospects for biological control," *Journal of Pest Science*, vol. 83, no. 3, pp. 197–215, Aug. 2010, https://doi.org/10.1007/s10340-010-0321-6.

[6] "Phthorimaea absoluta (tomato leafminer)," *CABI*. https://www.cabi.org/isc/datasheet/49260#tosummaryOfInvasiveness (accessed Sep. 30, 2021).

[7] H. E. Z. Tonnang, S. F. Mohamed, F. Khamis, and S. Ekesi, "Identification and Risk Assessment for Worldwide Invasion and Spread of Tuta absoluta with a Focus on Sub-Saharan Africa: Implications for Phytosanitary Measures and Management," *PLOS ONE*, vol. 10, no. 8, 2015, Art. no. e0135283, https://doi.org/10.1371/journal.pone.0135283.

[8] V. Van Damme *et al.*, "Overwintering potential of the invasive leafminer Tuta absoluta (Meyrick) (Lepidoptera: Gelechiidae) as a pest in greenhouse tomato production in Western Europe," *Journal of Pest Science*, vol. 88, no. 3, pp. 533–541, Sep. 2015, https://doi.org/10.1007/s10340-014-0636-9.

[9] M. Chidege, S. Al-zaidi, N. Hassan, A. Julie, E. Kaaya, and S. Mrogoro, "First record of tomato leaf miner Tuta absoluta (Meyrick) (Lepidoptera: Gelechiidae) in Tanzania," *Agriculture & Food Security*, vol. 5, no. 1, Aug. 2016, Art. no. 17, https://doi.org/10.1186/s40066-016-0066-4.

[10] T. J. Maginga, T. Nordey, and M. Ally, "Extension System for Improving the Management of Vegetable Cropping Systems," *Journal of Information Systems Engineering and Management*, vol. 3, no. 4, Nov. 2018, Art. no. 29, https://doi.org/10.20897/jisem/3940.

[11] D. I. Patricio and R. Rieder, "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review," *Computers and Electronics in Agriculture*, vol. 153, pp. 69–81, Oct. 2018, https://doi.org/10.1016/j.compag.2018.08.001.

[12] N. C. Eli-Chukwu, "Applications of Artificial Intelligence in Agriculture: A Review," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4377–4383, Aug. 2019, https://doi.org/10.48084/etasr.2756.

[13] M. T. Linaza *et al.*, "Data-Driven Artificial Intelligence Applications for Sustainable Precision Agriculture," *Agronomy*, vol. 11, no. 6, Jun. 2021, Art. no. 1227, https://doi.org/10.3390/agronomy11061227.

[14] J. Amara, B. Bouaziz, and A. Algergawy, "A Deep Learning-based Approach for Banana Leaf Diseases Classification," in *Datenbanksysteme fur Business, Technologie und Web*, Gesellschaft für Informatik e.V., 2017, pp. 79–88.

[15] J. G. A. Barbedo, "An Automatic Method to Detect and Measure Leaf Disease Symptoms Using Digital Image Processing," *Plant Disease*, vol. 98, no. 12, pp. 1709–1716, Dec. 2014, https://doi.org/10.1094/PDIS-03-14-0290-RE.

[16] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization," *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, Apr. 2017, https://doi.org/10.1080/08839514.2017.1315516.

[17] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition," *Sensors*, vol. 17, no. 9, Sep. 2017, Art. no. 2022, https://doi.org/10.3390/s17092022.

[18] L. Mkonyi *et al.*, "Early identification of Tuta absoluta in tomato plants using deep learning," *Scientific African*, vol. 10, Nov. 2020, Art. no. e00590, https://doi.org/10.1016/j.sciaf.2020.e00590.

[19] S. P. Mohanty, D. P. Hughes, and M. Salathe, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, 2016, Art. no. 1419, https://doi.org/10.3389/fpls.2016.01419.

[20] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep Learning for Image-Based Cassava Disease Detection," *Frontiers in Plant Science*, vol. 8, 2017, Art. no. 1852, https://doi.org/10.3389/fpls.2017.01852.

[21] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Computational Intelligence and Neuroscience*, vol. 2016, Jun. 2016, Art. no. e3289801, https://doi.org/10.1155/2016/3289801.

[22] N. Petrellis, "A smart phone image processing application for plant disease diagnosis," in *6th International Conference on Modern Circuits and Systems Technologies*, Thessaloniki, Greece, May 2017, pp. 1–4, https://doi.org/10.1109/MOCAST.2017.7937683.

[23] N. Petrellis, "Plant lesion characterization for disease recognition A Windows Phone application," in *2nd International Conference on Frontiers of Signal Processing*, Warsaw, Poland, Oct. 2016, pp. 10–14, https://doi.org/10.1109/ICFSP.2016.7802948.

[24] N. Petrellis, "Mobile Application for Plant Disease Classification Based on Symptom Signatures," in *21st Pan-Hellenic Conference on Informatics*, Larissa, Greece, Sep. 2017, pp. 1–6, https://doi.org/10.1145/3139367.3139368.

[25] N. Petrellis, "Plant Disease Diagnosis for Smart Phone Applications with Extensible Set of Diseases," *Applied Sciences*, vol. 9, no. 9, Jan. 2019, Art. no. 1952, https://doi.org/10.3390/app9091952.

[26] A. Johannes *et al.*, "Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case," *Computers and Electronics in Agriculture*, vol. 138, pp. 200–209, Jun. 2017, https://doi.org/10.1016/j.compag.2017.04.013.

[27] A. Ramcharan *et al.*, "A Mobile-Based Deep Learning Model for Cassava Disease Diagnosis," *Frontiers in Plant Science*, vol. 10, 2019, Art. no. 272, https://doi.org/10.3389/fpls.2019.00272.

[28] S. L. Sanga, D. Machuve, and K. Jomanga, "Mobile-based Deep Learning Models for Banana Disease Detection," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5674–5677, Jun. 2020, https://doi.org/10.48084/etasr.3452.

[29] Ranjith, S. Anas, I. Badhusha, O. T. Zaheema, K. Faseela, and M. Shelly, "Cloud based automated irrigation and plant leaf disease detection system using an android application," in *International conference of Electronics, Communication and Aerospace Technology*, Coimbatore, India, Apr. 2017, vol. 2, pp. 211–214, https://doi.org/10.1109/ICECA.2017.8212798.

[30] L. M. Mrisho *et al.*, "Accuracy of a Smartphone-Based Object Detection Model, PlantVillage Nuru, in Identifying the Foliar Symptoms of the Viral Diseases of Cassava–CMD and CBSD," *Frontiers in Plant Science*, vol. 11, 2020, Art. no. 1964, https://doi.org/10.3389/fpls.2020.590889.

[31] S. Verma, A. Chug, A. P. Singh, S. Sharma, and P. Rajvanshi, "Deep Learning-Based Mobile Application for Plant Disease Diagnosis," in *Applications of Image Processing and Soft Computing Systems in Agriculture*, Hershey, PA, USA: IGI Global, 2019, pp. 242–271.

[32] D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," *arXiv:1511.08060 [cs]*, Apr. 2016, Accessed: Sep. 30, 2021. [Online]. Available: http://arxiv.org/abs/1511.08060.

[33] L. Loyani, "TutaSegmenter," *Google Play Store*. https://play.google.com/store/apps/details?id=org.tensorflow.lite.examples.tutasegmentation (accessed Sep. 30, 2021).

[34] I. Arganda-Carreras *et al.*, "Crowdsourcing the creation of image segmentation algorithms for connectomics," *Frontiers in Neuroanatomy*, vol. 9, 2015, Art. no. 142, https://doi.org/10.3389/fnana.2015.00142.

[35] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Munich, Germany, Oct. 2015, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.

[36] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *25th International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, Dec. 2012, pp. 2843–2851.

[37] F. Chollet, "Introduction to Keras," in *Deep Learning with Python*, 1st ed., New York, NY, USA: Manning Publications Co., 2017, pp. 60–62.

[38] M. Abadi, "TensorFlow: learning functions at scale," in *21st ACM SIGPLAN International Conference on Functional Programming*, New York, NY, USA, Sep. 2016, Art. no. 1, https://doi.org/10.1145/2951913.2976746.

[39] L. K. Loyani, K. Bradshaw, and D. Machuve, "Segmentation of Tuta Absoluta's Damage on Tomato Plants: A Computer Vision Approach,"

Applied Artificial Intelligence, Sep. 2021, https://doi.org/10.1080/08839514.2021.1972254.

[40] D. P. Rubanga et al., A Deep Learning Dataset for Tomato Pest Leafminer TUTA ABSOLUTA. 2020, https://doi.org/10.5281/zenodo.4305416.

[41] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "LabelMe: A Database and Web-Based Tool for Image Annotation," *International Journal of Computer Vision*, vol. 77, no. 1, pp. 157–173, May 2008, https://doi.org/10.1007/s11263-007-0090-8.

[42] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, Jun. 2010, https://doi.org/10.1007/s11263-009-0275-4.

[43] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980 [cs]*, Jan. 2017, Accessed: Sep. 30, 2021. [Online]. Available: http://arxiv.org/abs/1412.6980.

[44] M. N. A. Khan, A. M. Mirza, R. A. Wagan, M. Shahid, and I. Saleem, "A Literature Review on Software Testing Techniques for Smartphone Applications," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6578–6583, Dec. 2020, https://doi.org/10.48084/etasr.3844.

[45] "Mobile Operating System Market Share Africa," *StatCounter Global Stats.* https://gs.statcounter.com/os-market-share/mobile/africa (accessed Sep. 30, 2021).

Industrial Activity Land Suitability Assessment Using Delphi and AHP to Control Land Consumption

The Case Study of Bordj Bouarreridj, Algeria

Fateh Toufik Benaissa
Department of Urban Engineering
Institute of Urban Techniques Management
University of Msila
Msila, Algeria
fatehtoufik.benaissa@univ-msila.dz

Boudjemaa Khalfallah
Department of Urban Engineering
Institute of Urban Techniques Management
University of Msila
Msila, Algeria
boudjemaa.khalfalllah@univ-msila.dz

Abstract-Studies and research on land suitability for the localization of economic and industrial activities in cities are attracting academics and local policies concerned with urban planning and organization in order to make urban areas more accommodating. The current paper aims to demonstrate the ability of visualization and simulation techniques to assist planners in preserving urban real estate by combining the Delphi method with one of the multi-criteria hierarchical analysis decision-making techniques. The city of Bordj Bouarreridj in Algeria was chosen as the case study and, with the participation of experts, a set of criteria were adopted to determine land suitability. Three criteria were selected in the first phase with the use of the Delphi method. In the second phase, we applied the Analytical Hierarchy Process (AH, a hierarchical analysis method. During the third stage, the obtained results were translated into the GIS program to derive the spatial suitability map. The following are among the most important achieved results: 13.38% (521.793 km^2) of the area was regarded to be of high suitability, 67.54% (2,634.758 km^2), achieved moderate suitability, 12.94% (504,794 km^2) attained low suitability, and 6.14% (239,685 km^2) were found to be not suitable at all.

Keywords-industrial activity; Bordj Bouarerridj; land suitability; control consumption; Delphi method; AHP

I. INTRODUCTION

Both private and public sectors contribute to the increment of the area of cities. In Algeria, we find that most of the construction of industrial zones is usually at the expense of agricultural lands, which affects the land uses of cities by covering the increasing housing demand [1]. The economic transformations of cities have created factors, represented in factories and workshops, that contribute to the emergence of management and organizational problems of the territory, especially about the allocation of lands and their suitability. The planners confront a dual problem in this situation. They must create initiatives and programs that maintain ecological balance while promoting economic growth, while they must act as mediators, attempting to prevent conflict and reduce objections. Land use planning is a means of optimal allocation of land resources [2]. It is frequently utilized in the spatial allocation and characteristics of land size to scheme and arrange land usage of the studied areas. Studying and identifying land-use needs for each land-use type is critical and fundamental for evaluating land use suitability for each land mapping unit [3]. Land use suitability analyses are of considerable use in the planning of mega-cities [4]. As a result, it is critical to use a technique that allows weight estimation. One such strategy is the Analytical Hierarchy Process (AHP), which is a multi-criteria decision-making method.

The author in [5] used several variables that were related to the environmental aspect. He focused his studies on environmental indicators. In this study, the answers were obtained by experts and institutions, and the obtained data were analyzed with the AHP method, after which a map of spatial suitability was derived by the Geographic Information Systems (GIS). Authors in [6] used AHP to determine the GIS weights based on 12 variables. Their study focused on social, economic, and environmental indicators. Such methods have the largest number of variables but the participation of experts and sometimes residents is often necessary. Authors in [7] used GIS and AHP to select the best sites for urban growth in Seremban, Malaysia. Many social, economic, and environmental factors were considered to create a final land suitability map, but the notable thing in this study is that the researchers relied on the selection of factors and variables on local literature and general observations only without relying on experts or residents.

This paper presents the results obtained by integrating the Delphi approach, AHP, and GIS in analyzing the relevance of land use to economic activity to control the land consumption in urban areas. The case study area was the city of Bordj Bouarerridj.

II. MATERIALS AND METHODS

A. The Study Area

Bordj Bouarerridj is an Algerian city in the high eastern highlands. It is located in the Algerian east in the middle of two Atlas Mountains series, the Beban mountain range and the Al-Hodna mountain range with a population density of 186.4 inhabitants/hectare in 2014, according to the NSB (National Statistics Bureau). Its total area is estimated at 3901.02km^2, i.e.

Corresponding author: Fateh Toufik Benaissa the 1/600 of the total area of Algeria. The forms of land usage vary, the land appropriateness is broad, and the degree of use is high (Figure 1).

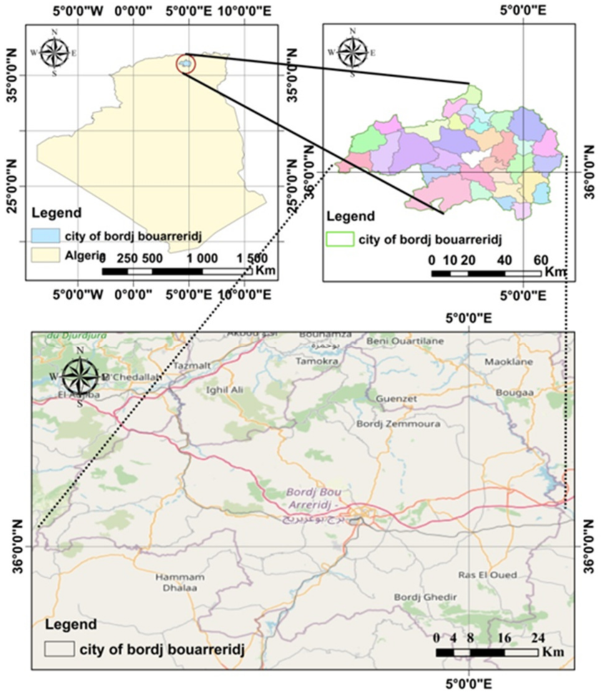


Fig. 1. The study area. Prepared by the researchers on the QGIS program, 2021.

B. Data Collection

1) Suitability Factors for Land Use for Industrial Activities

The criteria defined by the Delphi method to study the suitability of lands for economic and industrial activities, and their weights were appreciated by using the expert choice v11 program, which is an application for multi-criteria analysis through AHP. After collecting the geographical data, layers of determinant criteria, and sub-criteria were set in the open-source QGIS program. The results are displayed graphically in Figures 2-9.

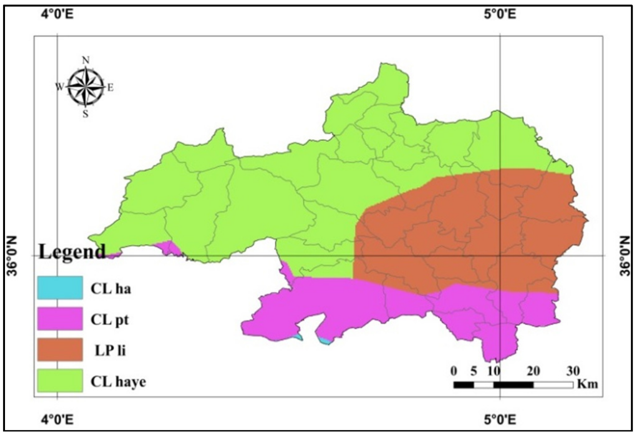


Fig. 2. Geological nature of the study area. Prepared by the researchers on the QGIS program.

2) Nature and Environment

For this factor, forest areas, hydrographical networks, and watershed basins have been identified, through the Bordj Bouarerridj city Reconstruction Plan. The suitability sub-factors were as follows (Figure 2): Forest areas: (0-5 Km), (5-9 Km), (9-13 Km), and (> 13 Km). Hydrographical network: (0-5 Km), (5-10 Km), (10-15 Km), and (> 15 Km). Watershed basins: (0-5 Km), (5-10 Km), (10-15 Km), and (>15 Km).

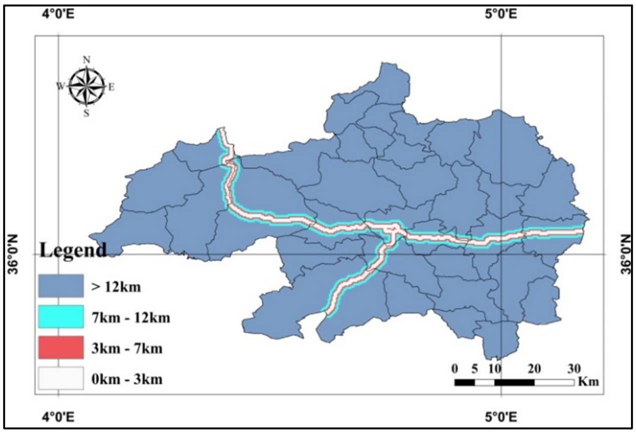


Fig. 3. Railways of the study area. Prepared by the researchers on QGIS.

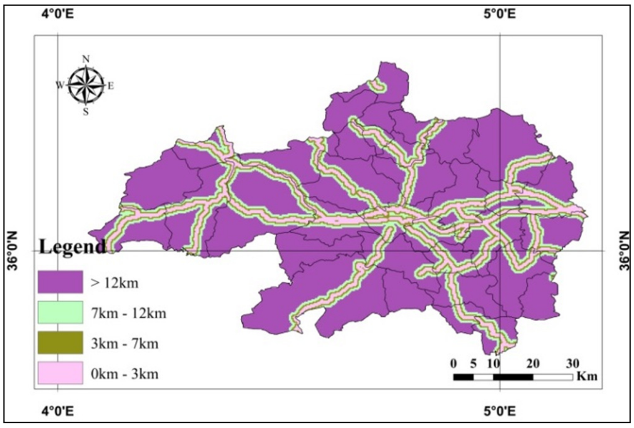


Fig. 4. Principal roads of the study area. Prepared by the researchers on QGIS.

3) Security and Protection

The relevant information was collected from the planning and reconstruction plans of the Bordj Bouarerridj state. Layers were determined in the ARC GIS program in urban tissues, archaeological areas, slopes, and geological sites according to the opinion of experts and technicians. The range of suitability for the distance from residential areas was as follows: (< 4 Km), (4-10 Km), (10-15 Km), and (> 15 Km). For slope, the range of suitability was as follows: (< 4%), (4%-8%), (8%-12%), and (> 12%). The geological nature was determined according to the permeability and type of soil.

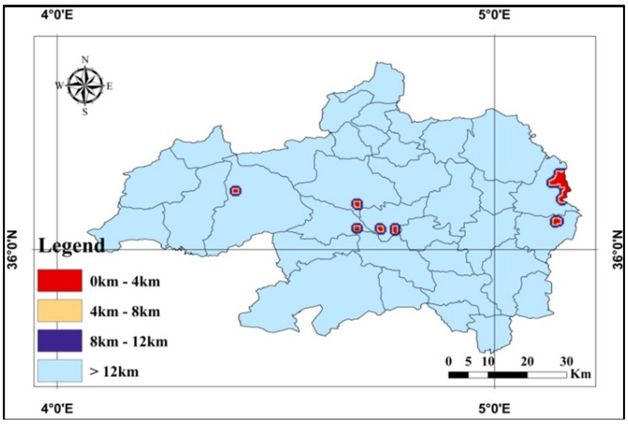


Fig. 5. Watershed basins of the study area. Prepared by the researchers on QGIS.

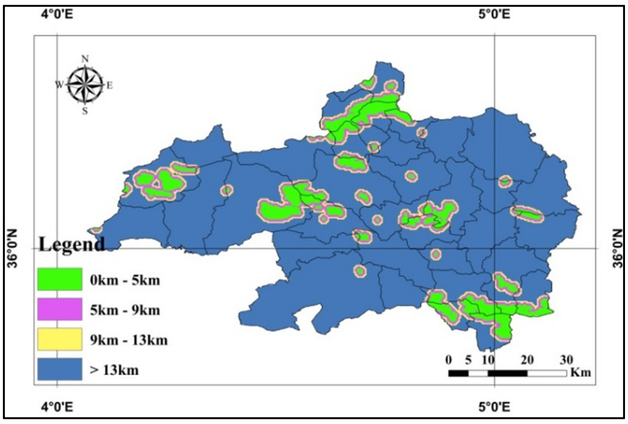


Fig. 6. Forests on the study area. Prepared by the researchers on QGIS.

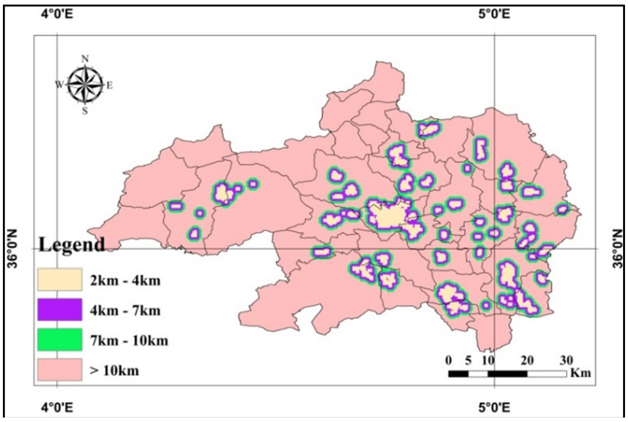


Fig. 7. Urban areas of the study area. Prepared by the researchers on QGIS.

4) Economic Determinant

For this factor, we considered all the main roads and structured railway lines of the case study. The data were collected from the planning and reconstruction plans of Bordj Bouarerridj in 2018. The range of suitability was determined by the experts as follows: (0.5-4 Km), (4-10 Km), and (> 10 Km) (Figures 3, 4).

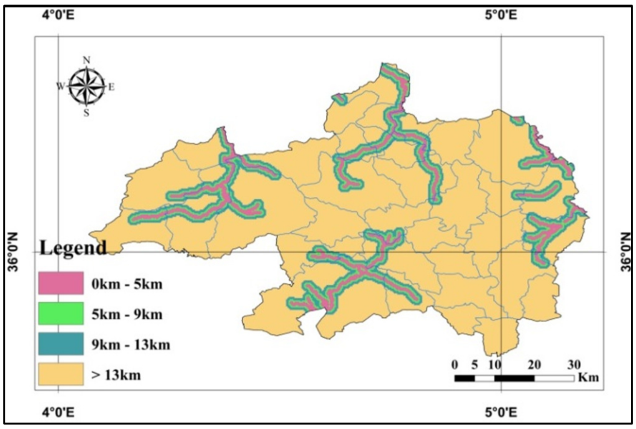


Fig. 8. Water ways of the study area. Prepared by the researchers on QGIS.

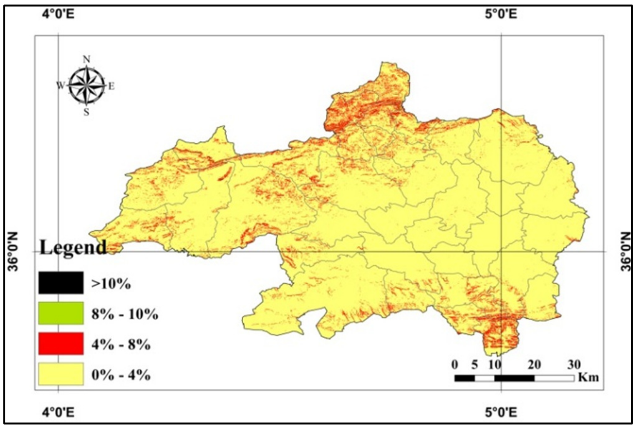


Fig. 9. Slopes of the study area. Prepared by the researchers on QGIS.

C. Methods

Industrial site selection is a spatial decision problem. Such problems typically involve a large set of feasible alternatives [8]. Land suitability for industrial activities is a somewhat complicated process considering the data and characteristics that distinguish them. Therefore many possible alternatives must take as well as appropriate multi-faceted and often conflicting standards, such as economic, environmental, climate, social, etc. The Delphi method is often applied to define the evaluation criteria as a way of structuring the method of collective communication to affect the complex problem of extracting and studying the influential factors to seek out the most appropriate sites. The suitability of lands for industrial activities in urban areas is determined by calculating the degree of importance of the percentage of significant value [9]. After, we used the AHP method which is used to solve complex decision problems [10]. That decision-making tool depends on reliable measurement and thorough evaluation to present suitable lands for economic activities [11]. It also serves as a helpful tool for capturing both subjective and objective assessment measures for checking the consistency of the evaluations and thus reducing bias in decision making [12]. This study is based on the AHP and through interactive computer programs (expert choice v11) [13]. The computer uses a hierarchical framework of criteria and comparisons between them to establish the criterion weight. The additive transformation function of the weighted summation technique is applied to get a final score for each choice. Expert Choice Network accepts no restriction on the size of the decision issued in this scenario. The DM's preferences on criteria are represented by cardinal weights [14].

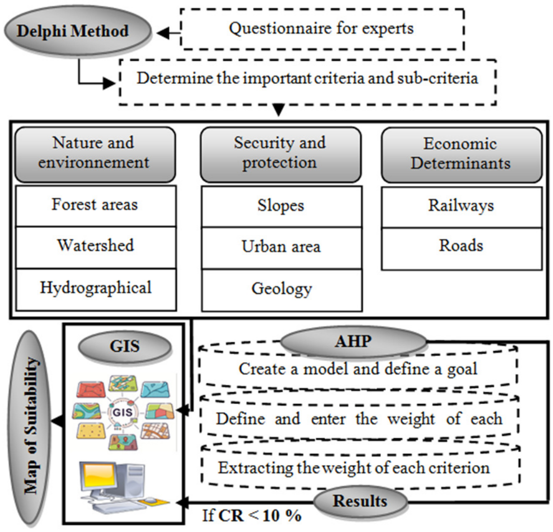


Fig. 10. The analysis used for this research.

D. Delphi Method

The Delphi method is a consensus method that aims to find general agreement among an expert panel on a specific research topic [15]. It is considered a way for organizing the group communication process to facilitate group issue solving and model building. Its principal area of application has remained that of technical forecasting [16]. In this research paper, we relied on 15 experts of various disciplines, namely geology, urbanization, economics, and law, who are familiar with the research topic and close to the decision-making authorities. Our questions were presented to them in four stages. After the first stage, we sent an anonymous summary of the resume of the experts' outlook of each previous stage and its reasoning [17]. This gradually led to the convergence of viewpoints and after the questionnaire's results were obtained. Through the questions posed to the experts and based on their opinions, criteria of approval for each standard set of sub-criteria were set. After this, the AHP series pyramid was analyzed to determine the weights of the main indicators and the sub-criterions in Expert Choice v11 software.

E. Hierarchical Analysis Process

AHP is a Multiple-Criteria Decision-Analysis (MCDA) technique used in solving various decision-making problems [18-20]. It is suitable for the evaluation of complex multi-attribute alternatives [21], because it is considered one of the most effective and easy procedures [22]. The AHP model has been applied to determine weights and ranks of all the subcategories [23]. The AHP methodology may tolerate inconsistency according to pair-wise comparison by given a measurement of evaluation inconsistency. This measurement is among the most essential parts of the priority determination process. The higher the consistency ratio, the assessment result becomes more inconsistent [24]. After the hierarchy creation, the relative importance of all resolution elements is captured and detected through binary comparisons. The Expert Choice is a decision-oriented program that provides support to decision-makers by organizing information with regarding to the complexity of a multi-solution problem into a hierarchical model consisting of objective, potential scenarios, agents, and alternatives. The priorities are determined along with the comparison matrix. We begin with a matrix that is consistent and has known priority p_i. p_i/p_j is used to compare alternatives i and j which is multiplied with the priority vector p results [25]:

$$\begin{bmatrix} p_1/p_1 & p_1/p_2 & \cdots & p_1/p_n \\ p_2/p_1 & p_2/p_2 & \cdots & p_2/p_n \\ \vdots & \vdots & \ddots & \vdots \\ p_n/p_1 & p_n/p_2 & \cdots & p_n/p_n \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = n \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} \quad (1)$$

Or grouped:

$$A\vec{p} = n\vec{p} \quad (2)$$

where $\vec{p}$ is the prioritization vector, n the matrix dimension, and A the matrix of comparison.

The author in [13] proposed a Consistency Index (CI) because priorities are only relevant if they are generated from coherent or close-to-consistent matrices, which is related to the eigenvalue method [26]:

$$\mathrm{CI} = \frac{Maks\ eigenvalue - n}{n-1} \quad (3)$$

$$Maks\ eigenvalue = \sum wi.ci \quad (4)$$

The Consistency Ratio (CR) is calculated to measure the consistency of the solution [27], and it is obtained by (5). The Random Consistency Index (RI) is shown in Table I [28].

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \quad (5)$$

TABLE I. RANDOM CONSISTENCY INDEX.

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

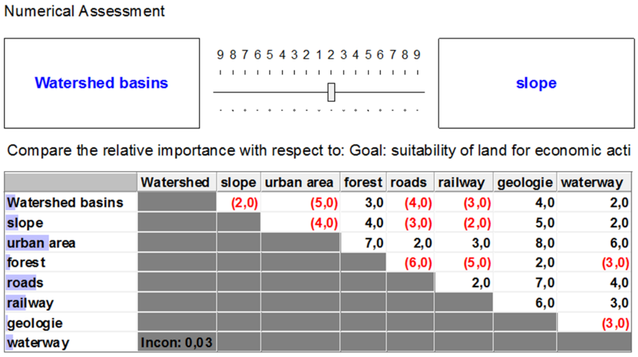
If CR ≥ 10% the data are considered inconsistent. If CR < 10% we have data consistency and the matrix is considered as having an acceptable consistency [25].

F. Geographic Information System (GIS)

GIS is a process of mapping and integrating computer-founded information with the ability of data layer management to make appropriate decisions by combining geographical information layers [29]. Data combined with an application of GIS provide alternative, flexible and scalable means [30]. The availability of digital data archives since 1987 enables the quantitative and qualitative assessment of land use patterns [31]. Accordingly, several studies have been conducted emphasizing the adoption of Delphi, AHP, or a combination of these methods, with GIS for modeling and analyzing spatial suitability [32]. It assists decision-makers to assign priority weights to decision criteria, evaluate the suitable alternatives, and visualize the results [33].

III. RESULTS AND DISCUSSION

This research relied on a MCDA using the AHP analysis to study the suitability of land use for economic activity in order to control the consumption of urban land involving the opinions of experts in the field. Eight criteria were used to quantify the suitability for economic and industrial land use, namely urban areas, slopes, forest areas, hydrographical network, watershed basins, geology, roads, and railways. A comparison matrix was implemented between the main factors and subfactors and paired-wise comparison was conducted between the sub-criteria of the same level [34]. Afterwards, the weight of each determinant was extracted (Figure 11).



Numerical Assessment

Compare the relative importance with respect to: Goal: suitability of land for economic acti

	Watershed	slope	urban area	forest	roads	railway	geologie	waterway
Watershed basins		(2,0)	(5,0)	3,0	(4,0)	(3,0)	4,0	2,0
slope			(4,0)	4,0	(3,0)	(2,0)	5,0	2,0
urban area				7,0	2,0	3,0	8,0	6,0
forest					(6,0)	(5,0)	2,0	(3,0)
roads						2,0	7,0	4,0
railway							6,0	3,0
geologie								(3,0)
waterway	Incon: 0,03							

Fig. 11. Results of AHP among main and sub criteria.

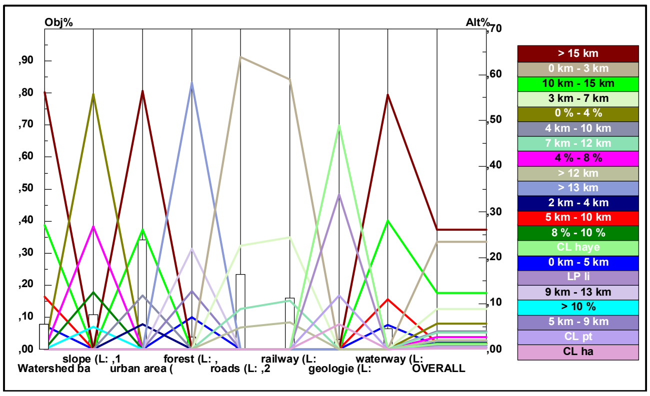
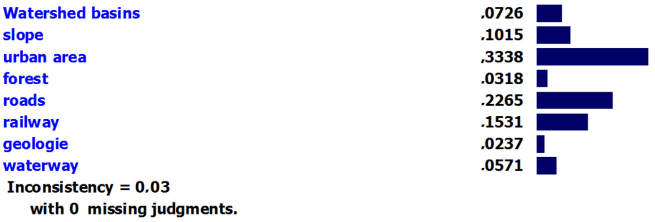


Fig. 12. Performance sensitivity for nodes.

The proportion of consistency was less than 10%, which means that the distribution between the factors was acceptable.

TABLE II. AHP KEY CRITERIA RATING

criteria	weight	rating
Road	0.2265	2
Railway	0.1531	3
Geology	0.237	8
Hydrographical network	0.571	6
Watershed basins	0.0726	5
Slope	0.1015	4
Forest	0.0318	7
Urban area	0.3338	1



Watershed basins	.0726
slope	.1015
urban area	.3338
forest	.0318
roads	.2265
railway	.1531
geologie	.0237
waterway	.0571

Inconsistency = 0.03
with 0 missing judgments.

Fig. 13. Priorities with respect to goal.

Based on the previous hierarchical analysis of the suitability of land use for economic and industrial activity, the role of GIS comes in the form of a model that summarizes all stages of work in the ARC GIS. We compiled all the pre-defined standard maps in GIS via Raster calculators with each standard multiplied by the weight generated by the AHP and Expert Choice to produce a suitable land use map. (Figure 14).

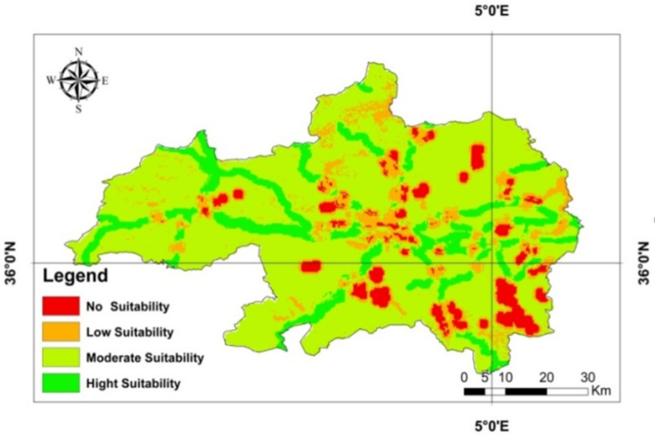


Fig. 14. The produced land suitability map for for economic and industrial activity regarding the Bordj Bouarerridj area. Prepared by the researchers on the QGIS program.

The final map gross weight was calculated by raster calculation using (6):

$$SL = (ROw * ROr) + (RWw * RWr) + (Gw * Gr) + (WWw * WWr) + (WAw * WAr) + (Sw * Sr) + (Fw * Fr) + (Uw * Ur) \quad (6)$$

where SL is the suitability of land use for economic and industrial activity, w is the weight of each criterion, r is the classification of each criterion, RO stands for road, RW for railway, G for geology, WW for waterways, WA for water, S for slope, F for forest, and U for urban areas.

Through hierarchical analysis the proportion of land-use suitability for economic and industrial activity in the city of Bordj Bouarerridj was:

- 13.38% of the total area with a total surface area of 521.793km^2 achieved high suitability.
- 67.54% of the total area with a total surface area of 504.794km^2 achieved moderate suitability.
- 12.94% of the total area with a total surface area of 2634.758km^2 achieved low suitability.

- 6.14% of the total area with a total surface area of 239.685km^2 was not suitable.

The results are summarized in Table III. In the final map (Figure 14) of the city of Bordj Bouarerridj' suitability of land usage for industrial activities, we note that the areas that achieved high suitability according to the AHP are the areas that are away from the urban tissues and near the road.

TABLE III. PERCENTAGE OF AHP STANDARD.

Classes	Area (km^2)	%
High suitability	521.793	13.38 %
Moderate suitability	2634.758	67.54 %
Low suitability	504.794	12.94 %
No suitability	239.685	6.14 %
Total	3901.023	100 %

IV. CONCLUSIONS

In this study, we utilized Delphi methodology and AHP and GIS technologies and a strategy was developed to assess the land use for economic activities and thus be able to manage the land consumption of metropolitan areas. The case study of this research was the state of Bordj Bouarerridj. A group of experts defined a set of criteria that helped determine the suitability of regions for economic activity. The AHP hierarchical analysis method was used to extract land-appropriate measures through the Expert Choice program. The consistency ratio was 0.03, which is less than 0.1, and therefore is considered favorable.

After using the AHP method, spatial analysis of the study area was carried out, with the creation of a geographical database of the criteria specified by the experts in GIS environment by creating a model using all accounted data to extract an appropriate map of land use for economic activity.

We can conclude by saying that the combination of the Delphi method and the hierarchical analysis using Expert Choice software in GIS environment is a practical scientific method that gives accurate results in determining the land use suitability for economic activity, according to the defined criteria. However, these criteria remain variable factors because they depend on the specificities of each region. The produced map may be adapted to identify regions for future housing development, evaluate the worth of a property, and assist decision-makers in achieving balanced urban development.

REFERENCES

[1] Mohamed Gherbi, "Instruments of Urban Planning in Algerian City: Reality and Challenges," *Journal of Civil Engineering and Architecture*, vol. 9, no. 7, pp. 807–812, Jul. 2015, https://doi.org/10.17265/1934-7359/2015.07.007.

[2] L. Jie, Y. Jing, Y. Wang, and Y. Shu-xia, "Environmental Impact Assessment of Land Use Planning in Wuhan City Based on Ecological Suitability Analysis," *Procedia Environmental Sciences*, vol. 2, pp. 185–191, Jan. 2010, https://doi.org/10.1016/j.proenv.2010.10.022.

[3] T. D. T., "Using GIS and AHP technique for land-use suitability analysis," in *GeoInformatics for Spatial-Infrastructure Development in Earth and Allied Sciences (GIS-IDEAS) 2006*, Nov. 2006.

[4] R. Liu, K. Zhang, Z. Zhang, and A. G. L. Borthwick, "Land-use suitability analysis for urban development in Beijing," *Journal of Environmental Management*, vol. 145, pp. 170–179, Dec. 2014, https://doi.org/10.1016/j.jenvman.2014.06.020.

[5] T. Everest, A. Sungur, and H. Özcan, "Determination of agricultural land suitability with a multiple-criteria decision-making method in Northwestern Turkey," *International Journal of Environmental Science and Technology*, vol. 18, no. 5, pp. 1073–1088, May 2021, https://doi.org/10.1007/s13762-020-02869-9.

[6] E. Memarbashi, H. Azadi, A. A. Barati, F. Mohajeri, S. V. Passel, and F. Witlox, "Land-Use Suitability in Northeast Iran: Application of AHP-GIS Hybrid Model," *ISPRS International Journal of Geo-Information*, vol. 6, no. 12, Dec. 2017, Art. no. 396, https://doi.org/10.3390/ijgi6120396.

[7] M. M. Aburas, S. H. O. Abdullah, M. F. Ramli, and Z. H. Asha'ari, "Land Suitability Analysis of Urban Growth in Seremban Malaysia, Using GIS Based Analytical Hierarchy Process," *Procedia Engineering*, vol. 198, pp. 1128–1136, Jan. 2017, https://doi.org/10.1016/j.proeng.2017.07.155.

[8] A. Rikalovic, I. Cosic, and D. Lazarevic, "GIS Based Multi-criteria Analysis for Industrial Site Selection," *Procedia Engineering*, vol. 69, pp. 1054–1063, Jan. 2014, https://doi.org/10.1016/j.proeng.2014.03.090.

[9] S. A. Delbari, S. I. Ng, Y. A. Aziz, and J. A. Ho, "An investigation of key competitiveness indicators and drivers of full-service airlines using Delphi and AHP techniques," *Journal of Air Transport Management*, vol. 52, pp. 23–34, Apr. 2016, https://doi.org/10.1016/j.jairtraman.2015.12.004.

[10] M. R. Abbaszadeh, M. Moradi, and S. T. Mehrabankhou, "Application of analytic hierarchy process in analyzing and ranking of non-financial measures that affect investor decisions," *International Journal of Business Management and Administration*, vol. 2, no. 2, pp. 31-42, Feb. 2013.

[11] Y. Wind and T. L. Saaty, "Marketing Applications of the Analytic Hierarchy Process," *Management Science*, vol. 26, no. 7, pp. 641–658, Jul. 1980, https://doi.org/10.1287/mnsc.26.7.641.

[12] P. Fernandez, S. Mourato, and M. Moreira, "Social vulnerability assessment of flood risk using GIS-based multicriteria decision analysis. A case study of Vila Nova de Gaia (Portugal)," *Geomatics, Natural Hazards and Risk*, vol. 7, no. 4, pp. 1367–1389, Jul. 2016, https://doi.org/10.1080/19475705.2015.1052021.

[13] T. L. Saaty, *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*, 3rd ed. Pittsburgh, PA, USA: RWS Publications, 2013.

[14] P. Jankowski, "Integrating geographical information systems and multiple criteria decision-making methods," *International Journal of Geographical Information Systems*, vol. 9, no. 3, pp. 251–273, May 1995, https://doi.org/10.1080/02693799508902036.

[15] P. Galanis, "The Delphi Method," *Archives of Hellenic Medicine*, vol. 35, no. 4, pp. 564–570, Jul. 2018.

[16] H. A. Linstone and M. Turoff, *The Delphi Method: Techniques and Applications*, 1st ed. Reading, MA, USA: Addison Wesley Publishing Company, 1975.

[17] S. Dehimi and M. Hadjab, "Evaluating the quality of life in urban area by using the Delphi method. A case study of M'sila city/Algeria," *Romanian Journal of Geography*, vol. 2, no. 63, pp. 193–202, 2019.

[18] H. I. Mohammed, Z. Majid, Y. B. Yamusa, M. F. M. Ariff, K. M. Idris, and N. Darwin, "Sanitary Landfill Siting Using GIS and AHP: A Case Study in Johor Bahru, Malaysia," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4100–4104, Jun. 2019, https://doi.org/10.48084/etasr.2633.

[19] E. Triantaphyllou and S. Mann, "Using the analytic hierarchy process for decision making in engineering applications: Some challenges," *The International Journal of Industrial Engineering: Theory, Applications and Practice*, vol. 2, no. 1, pp. 35–44, Feb. 1995.

[20] A. Ishizaka and M. Lusti, "An Intelligent Tutoring System for AHP," in *9th International Conference on Operational Research*, Jan. 2003, pp. 215–223.

[21] A. Emrouznejad and M. Marra, "The state of the art development of AHP (1979–2017): a literature review with a social network analysis," *International Journal of Production Research*, vol. 55, no. 22, pp. 6653–6675, Nov. 2017, https://doi.org/10.1080/00207543.2017.1334976.

[22] B. D. Nguyen, D. T. Minh, A. Ahmad, and Q. L. Nguyen, "The Role Of Relative Slope Length In Flood Hazard Mapping Using Ahp And Gis (Case Study: Lam River Basin, Vietnam)," *Geography, Environment, Sustainability*, vol. 13, no. 2, pp. 115–123, Jun. 2020, https://doi.org/10.24057/2071-9388-2020-48.

[23] F. Elmahmoudi, O. E. K. Abra, A. Raihani, O. Serrar, and L. Bahatti, "Elaboration of a Wind Energy Potential Map in Morocco using GIS and Analytic Hierarchy Process," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6068–6075, Aug. 2020, https://doi.org/10.48084/etasr.3692.

[24] M. Nilashi, K. Bagherifard, O. Ibrahim, N. Janahmadi, and M. Barisami, "An Application Expert System for Evaluating Effective Factors on Trust in B2C WebsitesTrust, Security, ANFIS, Fuzzy Logic, Rule Based Systems, Electronic Commerce," *Engineering*, vol. 3, no. 11, pp. 1063–1071, Nov. 2011, https://doi.org/10.4236/eng.2011.311132.

[25] A. Ishizaka and A. Labib, "Analytic Hierarchy Process and Expert Choice: Benefits and limitations," *OR Insight*, vol. 22, no. 4, pp. 201–220, Dec. 2009, https://doi.org/10.1057/ori.2009.10.

[26] D. Yudi and H. H. Azwir, "Reducing Defects Number of Ampoule by Considering Expected Failure Cost At Quality Control Department of PT. X," *Journal of Industrial Engineering*, vol. 2, no. 2, pp. 65–74, Nov. 2018, https://doi.org/10.33021/jie.v2i2.486.

[27] S. Panchal and A. Debbarma, "Rail-Route Planning Using a Geographical Information System (GIS)," *Engineering, Technology & Applied Science Research*, vol. 7, no. 5, pp. 2010–2013, Oct. 2017, https://doi.org/10.48084/etasr.1329.

[28] E. O. Oyatoye, B. B. Amole, and S. O. Adebiyi, "Patients' perception of quality service delivery of public hospitals in Nigeria using analytical hierarchy process," *Journal of Health Management & Information Science*, vol. 3, no. 3, pp. 66–73, Jul. 2016.

[29] H. Kazemi, S. Sadeghi, and H. Akinci, "Developing a land evaluation model for faba bean cultivation using geographic information system and multi-criteria analysis (A case study: Gonbad-Kavous region, Iran)," *Ecological Indicators*, vol. 63, pp. 37–47, Apr. 2016, https://doi.org/10.1016/j.ecolind.2015.11.021.

[30] R. Saouli, N. Benhassine, and A. Oularbi, "A Spatio-temporal retrospective of the uban sprawl of Annaba (Algeria)," *Journal of Fundamental and Applied Sciences*, vol. 12, no. 2, pp. 825–844, Nov. 2020.

[31] A. Ayache, M. Addoun, B. Hellal, and N. Ayad, "Diachronic analysis of soil occupancy using remote detection tools in tlemcen province southern (western Algeria)," *Journal of Fundamental and Applied Sciences*, vol. 13, no. 1, pp. 75–88, 2021, https://doi.org/10.4314/jfas.v13i1.5.

[32] H. Kazemi, S. Sadeghi, and H. Akinci, "Developing a land evaluation model for faba bean cultivation using geographic information system and multi-criteria analysis (A case study: Gonbad-Kavous region, Iran)," *Ecological Indicators*, vol. 63, pp. 37–47, Apr. 2016, https://doi.org/10.1016/j.ecolind.2015.11.021.

[33] P. Jankowski, "Integrating geographical information systems and multiple criteria decision-making methods," *International Journal of Geographical Information Systems*, vol. 9, no. 3, pp. 251–273, May 1995, https://doi.org/10.1080/02693799508902036.

[34] A. R. Al-shabeeb, "The Use of AHP within GIS in Selecting Potential Sites for Water Harvesting Sites in the Azraq Basin—Jordan," *Journal of Geographic Information System*, vol. 8, no. 1, pp. 73–88, Feb. 2016, https://doi.org/10.4236/jgis.2016.81008.