

# Investigation of the Influence of Skewed Slots and Degmagnetization Effects to Line Start Permanent Magnet Assistance Synchronous Reluctance Motors

**Tien Ho Manh**

University of Transport and Communications, Vietnam  
hotien.ktd@utc.edu.vn

**Dinh Bui Minh**

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam  
dinh.buiminh@hust.edu.vn

**Tu Pham Minh**

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam  
tu.phamminh@hust.edu.vn-mail address

**Vuong Dang Quoc**

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam  
vuong.dangquoc@hust.edu.vn  
(corresponding author)

*Received: 2 September 2022 | Revised: 27 October 2022 | Accepted: 27 October 2022*

## ABSTRACT

A permanent magnet assistance synchronous reluctance motor can start directly with a net voltage or a power converter via a torque control method. However, this motor has usually a higher irreversible demagnetization level in comparison with interior permanent magnet motors, due to the fewer permanent magnets in rotor slots. In order to cope with this disadvantage, different arrangements of permanent magnets in the rotor of the line-start permanent magnet assistance synchronous reluctance motor are proposed in this paper. The V magnet shape taking skewed slots and demagnetization effect into account with the short circuit current are investigated by the finite element approach. The efficiency, torque, and output power of the proposed model have been also improved. Finally, the rotor with 3V layered magnets is prototyped to verify the efficiency of the proposed motor.

**Keywords**-line start permanent magnet assistance-synchronous reluctance; magnet shape; skewed slots; demagnetization effect; finite element method

## I. INTRODUCTION

The Line Start Permanent Magnet Assistance Synchronous Reluctance Motors (LS-PMA-SynRMs) have recently been presented [1]. Authors in [2] proposed the design and optimization of low torque ripple and high torque by using ferrite magnets. However, the influences of skewed slots and the demagnetization effects on the LS-PMA-SynRMs have not been mentioned so far. In this paper, different arrangements of permanent magnets in the rotor taking skewed slots and the demagnetization effects taken into account are developed to improve the efficiency of the LS-PMA-SynRMs. Normally,

when the LS-PMA-SynRMs are operated at synchronous speed, the secondary copper loss is eliminated, leading to efficiency improvement. Less use of rare-earth materials, low cost, comparable constant-power speed range, maximum torque, and the efficiency of the LS-PMA-SynRMs will be considered. In particular, this is an interesting choice for the application of electric traction because it has a special flux barrier and suitable magnet arrangement to improve the constant torque in a wide speed range. In addition, it has a lesser risk of irreversible demagnetization in the short circuit and overheat temperature [3].

## II. MODEL OF LS-PMA-SYNRMS

A 7.5kW LS-PMA-SynRM model for 6 poles is shown in Figure 1. The stator has 36 slots and the 3-layer rotor has 36 round bars. The magnetic material and silicon steels are made in N38UH and 35A350. The response of magnets to the demagnetizing influence is considered in this research. Depending on the type of magnets, the motor can either need to withstand irreversible demagnetization during a transient short circuit fault condition or only require that the demagnetization is avoided during the steady-state operating conditions. For the sintered NdFeB and SmCo magnets, an irreversible demagnetization will occur at the very high temperature of 120°C. On the other hand, the irreversible demagnetization will occur at low temperatures (i.e. below 0°C) for the sintered ferrite magnets. Initially, a temperature value can be fixed to investigate the possible demagnetization. For example, it will be assumed that the magnet temperature is 160°C. Hence, for a selected N38UH magnet, the knee point of the demagnetization curve at the specified temperature is indicated in Figure 2 [4]. It can be seen that the flux density of the B-H curve becomes non-linear. The irreversible demagnetization in the magnet will appear if the flux density is depressed below that point. It should be noted the curve becomes non-linear at about 0.2T – at 160°C (the magnet temperature of our model). When the knee point is higher, then the magnet is more easily irreversibly demagnetized at high temperatures.

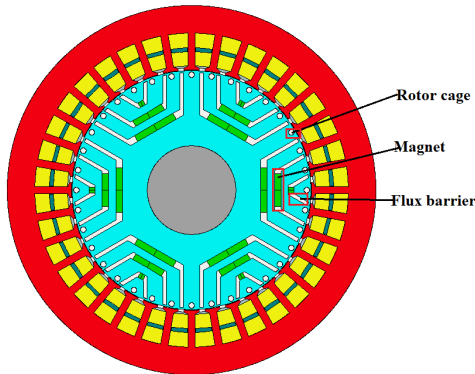


Fig. 1. LS-PMA-SynRM with 3V layer.

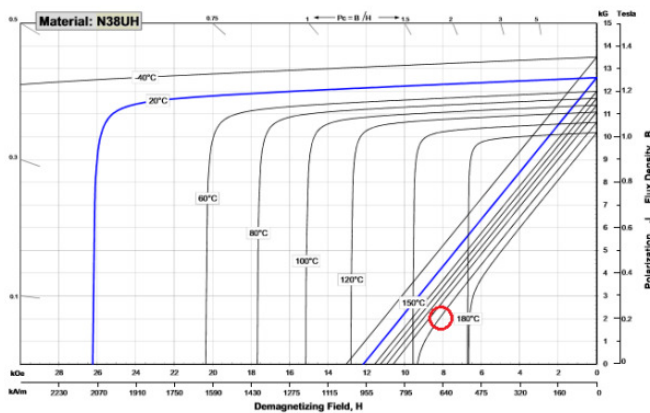


Fig. 2. Magnetization curve of B-H.

The magnetization curve with a higher temperature is pointed out in Figure 3 [4]. The curve of irreversible demagnetization is described by [5-7]:

$$B = B_r + \mu_0 \mu_r H_c - E \cdot e^{-k_1(k_2 + H_c)} \quad (1)$$

where  $B$  is the flux density,  $B_r$  is the remanent flux density,  $H_c$  is the coercivity,  $E$  is the electrical field, and  $k_1$  and  $k_2$  are respectively the factors of the magnets, where the factor  $k_2$  can be defined as [8, 9]:

$$k_2 = \frac{\ln[(B_r + (\mu_r - 1)\mu_0 H_c) \frac{1}{E}]}{k_1} - H_c \quad (2)$$

The working point of the magnet is defined by the intersection between the flux density ( $B_m$ ) and magnetic field strength ( $H_m$ ). A point on the demagnetization characteristic of the magnet and load line is also defined. The slope of the load line is called the permeance coefficient. It is determined principally by the ratio of the magnet length to the air gap ( $g$ ). When the machine is under load, the whole load line shifts, usually to the left, so the working point is shifted further down the curve. Once it overcomes the knee of the curve, an irreversible loss of magnetization will appear. Via the computation of demagnetization, it is certain that the operating point of the magnet stays above the knee. It also needs to be ensured that the magnet is working with the worst-case parameters. In particular, temperature is critical. For high-energy magnets, when the temperature increases, both  $B_r$  and  $H_c$  decrease. However, the knee moves to a lower value of  $H_m$ , which means that less current is required to demagnetize the magnet. The maximum current that can flow and its phase angle must also be considered. Normally, the current is regulated by the inverter unit, thus the maximum current is tightly controlled. But there may be fault conditions where the current reaches a value much higher than normal.

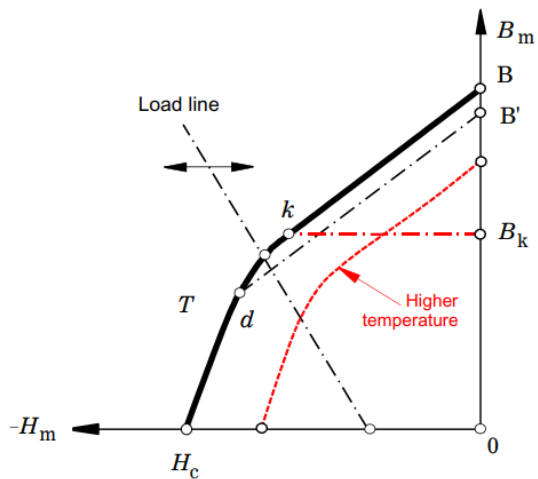


Fig. 3. B-H magnetization curve with higher temperature.

In the worst conditions, for a three-phase short-circuit, the temperature of the magnet is 25°C. For that, the current can be presented as [4]:



$$I_{d[sc]} = \frac{E_{q1}}{X_d} \quad (3)$$

where  $E_{q1}$  is the Root Mean Square (RMS) of the Electromotive Force (EMF) per phase and  $X_d$  is the  $d$ -axis synchronous reactance. Normally, this current is 2 to 5 times the maximum working current. The design parameters of LS-PMA-SynRM are given in Table I.

The schematic diagram of the proposed PMA-SynRM machine is presented in Figure 4. In the proposed motor, the magnet shape is arranged with the U-shape. This arrangement is regarded as requisite for the efficient operation in the type of I-W-U shape. The total weight of the magnet segment and copper winding is presented in Table II.

TABLE I. GEOMETRY PARAMETERS OF LS- PMA-SYNRM

Parameter	Values	Unit
Slot number	36	slot
Outer stator	210	mm
Inner stator	132	mm
Tooth width	5	mm
Slot depth	19	mm
Motor length	180	mm
Stator Lam length	180	mm
Magnet length	150	mm
Magnet size	5×2	mm
Air gap	0.4	mm
Turn per coil	20	turn

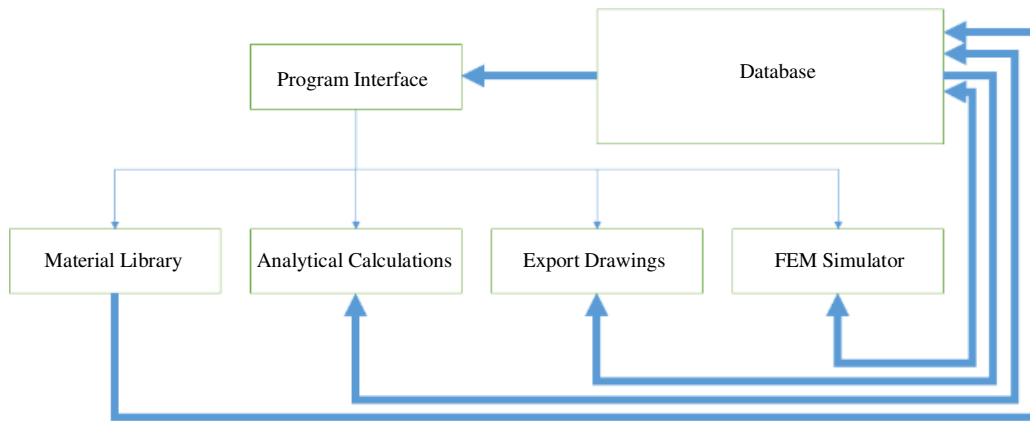


Fig. 4. Schematic diagram.

TABLE II. WEIGHT OF LS- PMA-SYNRM

Parameter	Kg
Stator lam (back iron)	8.26
Stator Lam (yooth)	5.245
Stator lamination (total)	13.5
Armature winding (active)	4.138
Armature EWdg (front)	1.027
Armature EWdg (rear)	1.027
Armature winding (total)	6.193
Rotor lam (back iron)	5.37
IPM magnet pole	4.833
Rotor lamination (total)	10.74
Magnet	0.5
Total	33.98

$$T_{EM} = \frac{3p}{2} [\lambda_{pm} \cdot i_d + (L_d - L_q) \cdot i_d \cdot i_q] \quad (4)$$

$$L_d = \frac{\lambda_d - \lambda_{pm}}{i_d} \Big|_{i_d = 0} \quad (5)$$

$$L_q = \frac{\lambda_q}{i_q} \Big|_{i_q = 0} \quad (6)$$

where  $\lambda_{pm}$  is the flux linkage generated by the PM field,  $i_d$  and  $i_q$  are respectively the direct and quadrature axis currents, and  $L_d$  and  $L_q$  are the direct and quadrature axis inductances. The term  $\lambda_{pm}$  depends on magnet sizes. The terms ( $L_d$  and  $L_q$ ) are computed following to the rotor magnet barrier and magnet pole U shape.

### III. LS-PMA-SYNRM PERFORMANCE IN DIRECT START AND TRACTION

The proposed PMA-SynRM machine was designed and analyzed in traction. The electromagnetic torque is formed from two components of magnetic reluctance torques. The Permanent Magnet (PM) component is produced via the interaction between the air-gap magnetic field and the armature reaction magnetic field. The reluctance component is instead based on the asymmetry between the magnetic circuit of  $d$ -axis and  $q$ -axis. The electromagnetic torque can then be defined as [5-9]:

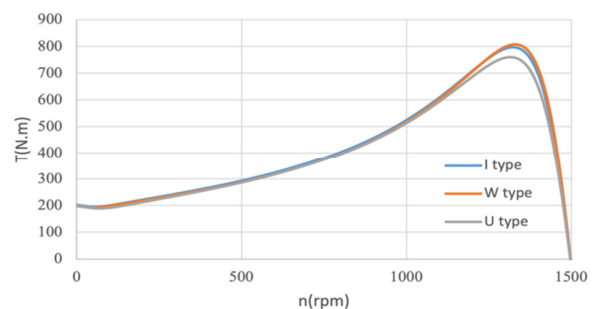


Fig. 5. Comparison of dynamic starting torque and speed with different types of I-W-U.

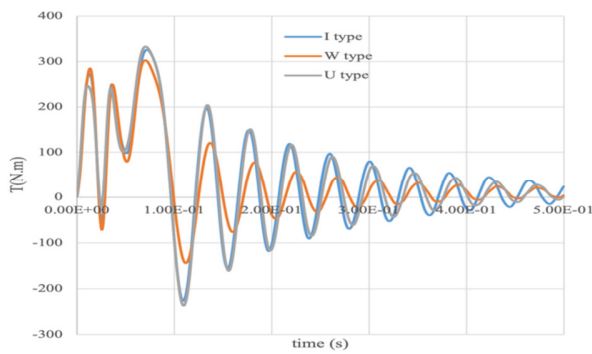


Fig. 6. Comparison of dynamic starting torque and time with different types of I-W-U.

The dynamic starting torque and speed with different types of I-W-U are shown in Figures 5 and 6. For the U-shape, the peak torque has been verified at 1500rpm. When increasing the phase current density up to  $10\text{A/mm}^2$ , the peak torque is 340Nm. It should be noted that the LS-PMA-SynRM can be pre-determined by 11kW under the direct start. If this motor is operated with the power inverter, the maximum power is doubled, and the peak torque is about 800Nm (Figure 5). The parameter comparison of the LS-PMA-SynRM with different types of I-W-U is given in Table III.

TABLE III. PARAMETER COMPARISON OF LS-PMA-SYNRM

Parameters	I shape	W shape	U shape	Unit
Shaft torque	53.67	54.267	55.667	Nm
Input power	9123.3	9213.4	9273.3	W
Output power	8644.1	8474.3	8744.1	W
Total losses	519.5	522.1	529.15	W
System efficiency	93.24	94.3	94.68	%
Armature DC copper loss	340	340	340	W
Magnet loss	178.8	168.3	168	W
Stator iron loss	10.33	10.33	10.33	W
Phase terminal voltage	289.1	289.1	289.1	V
Harmonic distortion line-line terminal voltage	4.89	4.289	3.089	V
Harmonic distortion phase terminal voltage	13.7	12.2	11.27	%
Back EMF line-line voltage	111	112	114	%

TABLE IV. TEMPERATURE OF LS-PMA-SYNRM

No	Component	T (°C)
1	T (ambient)	40
2	T (housing - active)	78.717
3	T [stator lam (back iron)]	84.833
4	T (stator surface)	88.403
5	T (rotor surface)	88.65
6	T (airgap banding)	88.651
7	T (magnet)	88.197
8	T (airgap banding)	88.651
9	T (rotor lamination)	88.674
10	T (shaft - center)	87.375
13	T (active winding minimum)	87.352

It can be seen that the maximum efficiency is 94%. The total loss errors between the three cases are smaller than 5%. For the errors on the harmonic distortion line, they are lower than 14%. The temperatures of the LS-PMA-SynRM are presented in Table IV. The maximum temperature of the

winding is  $90.7^\circ\text{C}$ , which is much lower than isolation class H ( $180^\circ$ ). The temperature of the magnet is  $88.17^\circ\text{C}$ . The dynamic starting torque and speed of the U shape is shown in Figure 7. The output power of the U shape is indicated in Figure 8. The map distribution of the rotor temperature with the U shape is shown in Figure 9. The maximum temperature is  $88.7^\circ\text{C}$ .

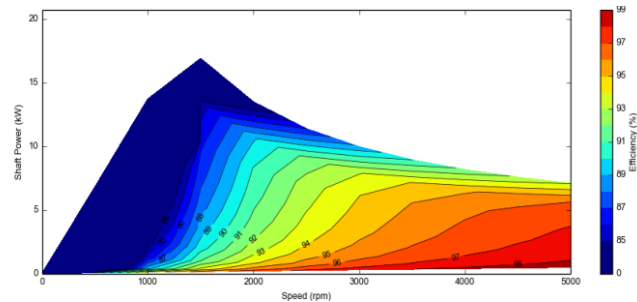


Fig. 7. Dynamic starting torque and speed of U shape design.

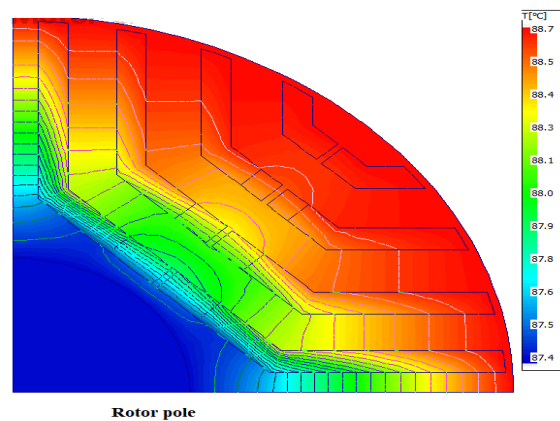


Fig. 8. Output power and speed map of the U shape.

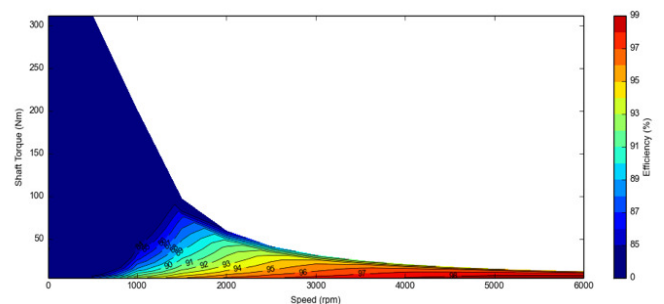


Fig. 9. Rotor temperature with the U shape.

#### IV. CONCLUSION

In this paper, the influence of skewed slots and demagnetization effects on the LS PMA-SynRM has been successfully investigated and analyzed. The values of dynamic starting torque and speed with different types of I-W-U have been also compared. The obtained results have shown that for the U shape, the torque and power density are noticeable improved, and the volume of magnets is the lowest. For a short starting time, the speed is considered as a constant in

comparison with the I and W shapes. The validation of the selected machine in the traction application and the full map of dynamic starting torque and power have been presented. In particular, the comparison of efficiency performances have been investigated with three different shapes, i.e. I, W, and U. Thermal simulation is also implemented to validate overheat capacity.

#### ACKNOWLEDGEMENT

This study was funded by the University of Transport and Communications, Hanoi, Vietnam under grant number T2022-DT-005.

#### REFERENCES

- [1] H. Kim *et al.*, "Study on Analysis Method of Asymmetric Permanent Magnet Assistance Synchronous Reluctance Motor Considering Magnetic Neutral Plane Shift," *IEEE Transactions on Applied Superconductivity*, vol. 30, no. 4, pp. 1–4, Jun. 2020, <https://doi.org/10.1109/TASC.2020.2968012>.
- [2] C. Gong and F. Deng, "Design and Optimization of a Low-Torque-Ripple High-Torque-Density Vernier Machine Using Ferrite Magnets for Low-Speed Direct-Drive Applications," in *2021 IEEE International Electric Machines & Drives Conference (IEMDC)*, Feb. 2021, pp. 1–8, <https://doi.org/10.1109/IEMDC47953.2021.9449586>.
- [3] D. B. Minh, N. H. Phuong, V. D. Quoc, and H. B. Duc, "Electromagnetic and Thermal Analysis of Interior Permanent Magnet Motors Using Filled Slots and Hairpin Windings," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8164–8167, Feb. 2022, <https://doi.org/10.48084/etasr.4683>.
- [4] M. Olszewski, "Oak Ridge National Laboratory Annual Progress Report for the Power Electronics and Electric Machinery Program," Oak Ridge National Laboratory, Oak Ridge, TN, USA, ORNL/TM-2011/263, Oct. 2011.
- [5] X. Chen, J. Wang, B. Sen, P. Lazari, and T. Sun, "A High-Fidelity and Computationally Efficient Model for Interior Permanent-Magnet Machines Considering the Magnetic Saturation, Spatial Harmonics, and Iron Loss Effect," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, pp. 4044–4055, Jul. 2015, <https://doi.org/10.1109/TIE.2014.2388200>.
- [6] 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>.
- [7] D. B. Minh, V. D. Quoc, and P. N. Huy, "Efficiency Improvement of Permanent Magnet BLDC Motors for Electric Vehicles," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7615–7618, Oct. 2021, <https://doi.org/10.48084/etasr.4367>.
- [8] M. Taniguchi *et al.*, "Development of New Hybrid Transaxle for Compact-Class Vehicles," SAE International, Warrendale, PA, USA, SAE Technical Paper 2016-01-1163, Apr. 2016, <https://doi.org/10.4271/2016-01-1163>.
- [9] Y. Wang, N. Bianchi, and R. Qu, "Comparative Study of Non-Rare-Earth and Rare-Earth PM Motors for EV Applications," *Energies*, vol. 15, no. 8, Jan. 2022, Art. no. 2711, <https://doi.org/10.3390/en15082711>.

# Efficiency Analysis and Design Considerations of a Hysteretic Current Controlled Parallel Hybrid Envelope Tracking Power Supply

**Ambily Babu**

ECE Department-VTU, AMC Engineering College, India  
ambily\_babu@outlook.com  
(corresponding author)

**B. G. Shivaleelavathi**

ECE Department- VTU, JSS Academy of Technical Education, India  
shivaleelavathi1963@gmail.com

**Veeramma Yatnalli**

ECE Department- VTU, JSS Academy of Technical Education, India  
veeramma71@gmail.com

Received: 15 October 2022 | Revised: 1 November 2022 | Accepted: 2 November 2022

## ABSTRACT

This paper presents the realization of an Envelope Tracking Power Supply (ETPS), for a sinusoidal envelope input signal. A parallel hybrid topology is chosen for the implementation. In this topology, a voltage-controlled Class-AB linear amplifier stage and a current-controlled switching converter stage operate in parallel. A hysteretic current control scheme is employed to control the operation of the switching converter. Block-level implementation of the ETPS is done using Simulink 2017b, continuous mode simulation. Simulations are performed using a sinewave input envelope signal =  $1.94 + 1.2 \sin(\omega \cdot t)$ . The input frequency varied from 1MHz to 60MHz. As the input frequency is increased, the ETPS moves from the linear to the non-linear region of operation. During the transition, the slew rates of the load current and the switching current match at a particular input frequency of 2MHz while the efficiency peaks. The maximum obtained efficiency while tracking the sinewave input signal is 82.3%. The way the efficiency can be optimized by focusing on the matching of the slew rates of load and switching currents is explained. Also, an insight into the study of various circuit parameters and the trade-offs that the designer needs to consider while designing an ETPS, is provided.

**Keywords-**envelope tracking; RF power amplifiers; supply modulators; ETPS; ETPA; mobile communication; design considerations

## I. INTRODUCTION

To cater to the ever-increasing demands of customers, RF power amplifiers in mobile communications need to deal with signals that have a very high Peak to Average Power Ratio (PAPR). Operating such an RF power amplifier with a fixed supply is no longer a viable option, as it reduces the overall efficiency as well as battery lifetime between charges. Envelope tracking is an assuring supply modulation technique [1-4] that helps improving the efficiency of such an RF power amplifier. Envelope Tracking Power Supply (ETPS) is a vital block of an Envelope Tracking Power Amplifier (ETPA), since the overall efficiency of the ETPA [5] depends on the efficiency of the ETPS as given by:

$$\eta_{ETPA} = \eta_{ETPS} * \eta_{RFP A} \quad (1)$$

The primary goal of an ETPS designer is to develop an efficient power supply, operating at higher bandwidths. Various ETPS topologies exist [6-9], aiming to provide high efficiency while maintaining signal fidelity. A hysteretic current controlled parallel hybrid ETPS, as shown in Figure 1, is employed in this paper, as it operates efficiently at high frequencies.

In this topology, a switching converter will track most of the low frequency signals, at high efficiency. The remaining high frequency signals will be tracked by the linear amplifier at a lower efficiency. The linear amplifier stage also needs to compensate for the ripples created by the switching converter, which is controlled by a hysteretic current control mechanism. The hysteretic current control mechanism offers a fast response

since it employs a window comparator and hence is ideal for high bandwidth mobile applications [10-11].

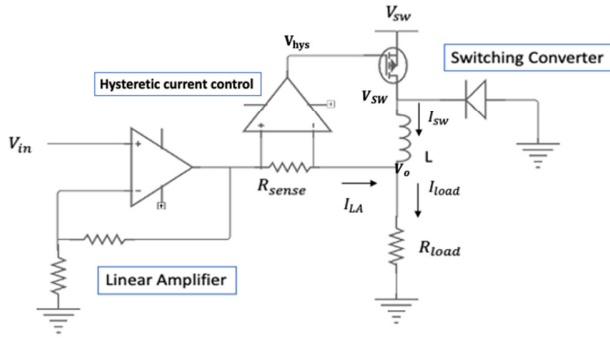


Fig. 1. Hysteretic current controlled ETSP [12].

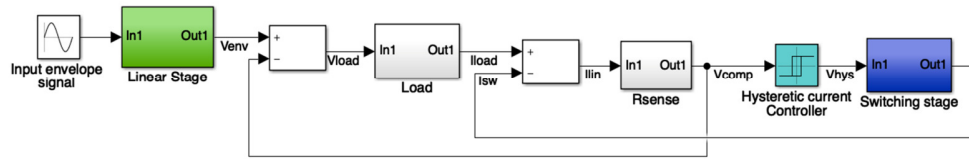


Fig. 2. Block level implementation.

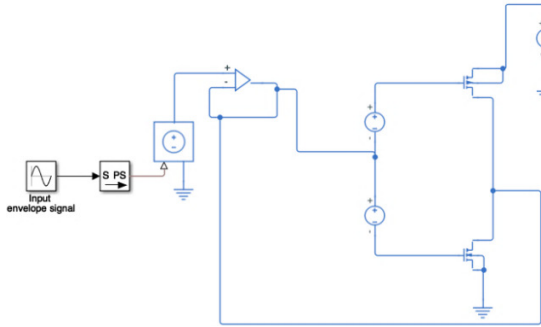


Fig. 3. Linear stage.

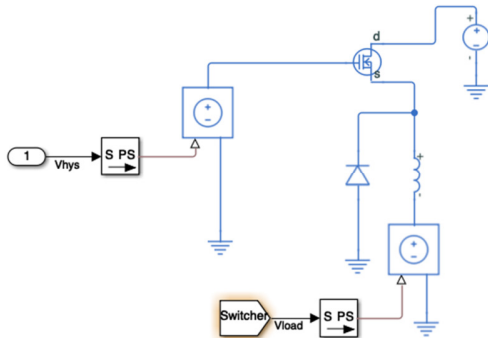


Fig. 4. Switching stage.

Linear stage current,  $I_{LA}$  is sensed by the resistor  $R_{sense}$  and the direction of this current controls the operation of the Hysteretic Controller. The value of  $R_{sense}$  is kept very small in order to sense the current accurately. A relay block in Simulink is used to function as a hysteretic controller. Hysteresis  $h$  is

## II. BLOCK LEVEL IMPLEMENTATION OF HYSTERETIC CURRENT CONTROLLED PARALLEL HYBRID ETSP

A block level implementation [13] of hysteretic current controlled parallel hybrid ETSP is depicted in Figure 2. A sinusoidal signal denoted by (2) is provided as the input envelope signal:

$$V_{in}(t) = V_{in\_dc} + V_{in\_ac} \sin(\omega t) \quad (2)$$

where  $V_{in}(t)$  is the instantaneous input signal,  $V_{in\_dc}$  the DC component of the input signal,  $V_{in\_ac} \sin(\omega t)$  the AC component of the input signal, and  $\omega = 2\pi f_{i/p}$ , where  $f_{i/p}$  is the input signal frequency.

A Class AB amplifier with an operational trans conductance amplifier input stage is used as the linear stage, as shown in Figure 3. A single stage buck converter is used as the switching stage, as shown in Figure 4.

chosen to be equal to  $\pm 7\text{mV}$ . Initially, the linear stage starts [14-18], providing current and when the voltage drop across the  $R_{sense}$  exceeds  $+7\text{mV}$ , the hysteretic controller will turn ON the switching converter. The switching stage starts providing the majority of the load current, as given by (3). The linear stage needs to source only the difference current during this time.

$$I_{load} = I_{SW} + I_{LA} \quad (3)$$

As the current provided by the switching stage increases and when the  $I_{SW}$  exceeds the  $I_{load}$ , the current through  $R_{sense}$  starts flowing in the opposite direction. When the voltage drop across  $R_{sense}$  falls below  $-7\text{mV}$  the hysteretic controller turns the switching converter OFF. The linear stage now needs to sink current.

## III. REGIONS OF OPERATION OF HYSTERETIC CURRENT CONTROLLED PARALLEL HYBRID ETSP

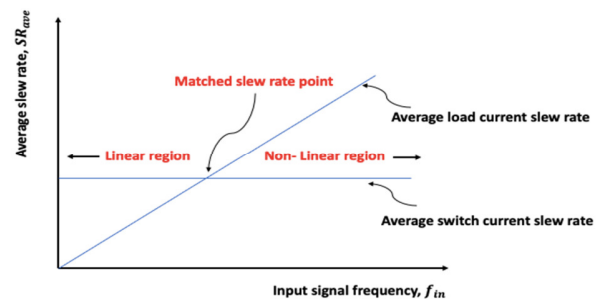


Fig. 5. Regions of operation.

The working of the hysteretic current controlled [19-24] Parallel Hybrid ETSP can be explained in 3 regions, as shown

in Figure 5. The circuit parameters considered for performing simulations in the 3 operating regions are mentioned in Table I.

TABLE I. PARAMETERS CONSIDERED FOR SIMULATION

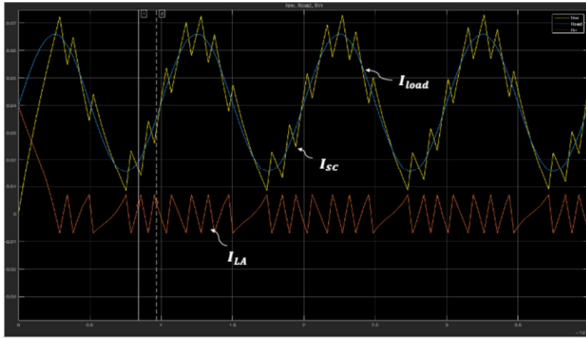
Parameter	Value
Input envelope signal	$V_s = 1.94 + 1.2 \sin(2\pi f_{ip} t)$
$R_{sense}$	$1\Omega$
$R_{load}$	$47\Omega$
$L$	$11\mu H$
$h$	$\pm 7mV$
$V_{dd}$	$5.5V$

#### A. Linear Region of Operation

In this region of operation, the slew rate of switching current is greater than the slew rate of load current as represented by (4).

$$SR_{i_{sw}} > SR_{i_{load}} \quad (4)$$

Simulations were conducted at 1MHz input frequency and the obtained results are shown in Figure 6. It can be observed that the slew rate of the switching current is greater than that of the load current. As a result, the switching frequency  $f_s$  of the switching stage increases and the calculated average switching frequency of the buck converter is 6.6MHz. The linear stage current is within the hysteresis values of  $\pm 7mA$ .

Fig. 6.  $I_{LA}$ ,  $I_{SW}$ ,  $I_{load}$  at 1MHz  $f_{in}$ .

The load current is the sum of linear stage current and the switching stage current, as is clearly visible from Figure 6. When the inductor current  $I_{SW}$  is rising, the current provided by the linear amplifier decreases and vice-versa. Hence, the three currents obey (5):

$$I_{SW}(t) + I_{LA}(t) = I_{load}(t) \quad (5)$$

$$I_{SW}(t) + \frac{V_{in}(t) - V_o(t)}{R_{sense}} = \frac{V_o(t)}{R_{load}} \quad (6)$$

Now, the switching current is:

$$\begin{aligned} I_{SW}(t) &= I_{SW\_noise}(t) + I_{SW\_input\ signal}(t) \\ &= I_{SW\_noise}(t) + \frac{V_{in}(t)}{R_{load}} \end{aligned} \quad (7)$$

Substituting the (7) in (6), we get:

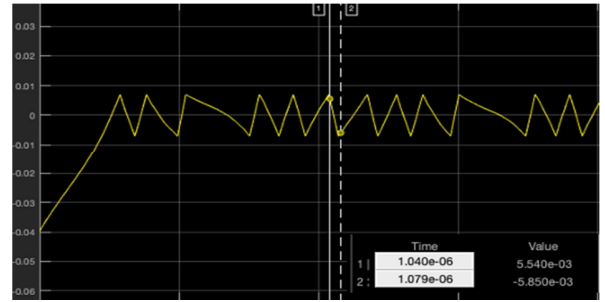
$$V_o(t) - V_{in}(t) = R_{sense} \cdot I_{SW\_noise}(t)$$

$$I_{SW\_noise}(t) = \frac{V_o(t) - V_{in}(t)}{R_{sense}} \quad (8)$$

This ripple current needs to be absorbed by the linear stage and, hence, the ripple voltage due to the switching current is given by:

$$I_{SW\_noise}(t) \cdot R_{sense} = V_o(t) - V_{in}(t) \quad (9)$$

The corresponding ripple voltage,  $V_o(t) - V_{in}(t)$  obtained during the simulation is given in Figure 7. It can be observed that the ripple voltage falls within the hysteresis value of  $\pm 7mV$ .

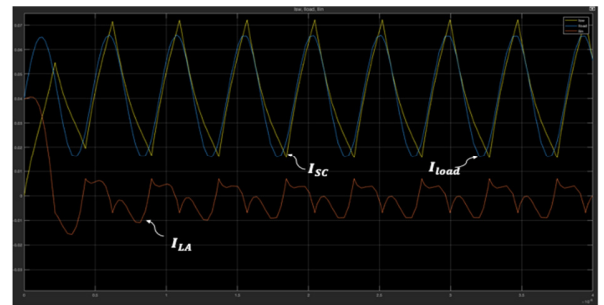
Fig. 7. Ripple voltage at 1MHz  $f_{in}$ .

#### B. Matching Slew Rate Point

The slew rate of the switching current is equal to the slew rate of the load current in this operating region, as given by (10):

$$SR_{i_{sw}} = SR_{i_{load}} \quad (10)$$

The simulation results at 2MHz input frequency are given in Figure 8.

Fig. 8.  $I_{LA}$ ,  $I_{SW}$ ,  $I_{load}$  at 2MHz  $f_{in}$ .

The average switching frequency of the converter is calculated to be 2MHz, which is equal to the input frequency. The linear stage current is slightly higher than the hysteresis values. The switching frequency is observed to be minimum at this point. The obtained corresponding ripple voltage is given in Figure 9. It can be observed that error is slightly greater than the hysteresis value of  $\pm 7mV$ .

#### C. Non-Linear Region of Operation

In this operating region, the slew rate of the switching current is lower than the slew rate of the load current as represented by (11):



$$SR i_{sw} < SR i_{load} \quad (11)$$

Simulations were conducted at 20MHz input frequency and the resulting current waveforms are given in Figure 10. As can be observed, the switching stage is able to provide only the DC components of the load current. The linear stage needs to provide the AC components of the load current. The average switching frequency of the converter is calculated to be 20MHz.

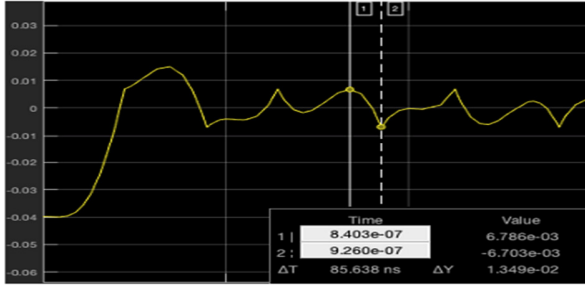


Fig. 9. Ripple voltage at 20MHz  $f_{in}$ .

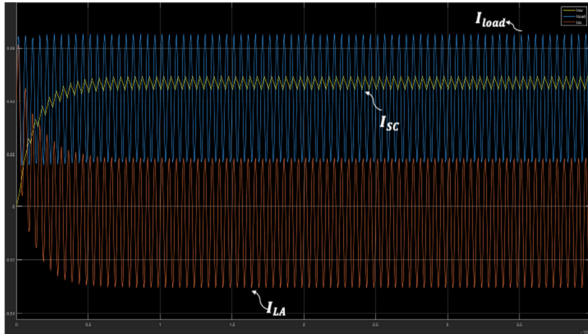


Fig. 10.  $I_{LA}$ ,  $I_{SW}$ ,  $I_{load}$  at 20MHz  $f_{in}$ .

The error ripple voltage obtained during the simulation is given in Figure 11. It can be observed that the error has significantly increased.

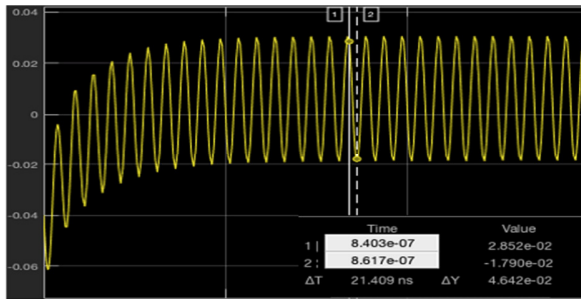


Fig. 11. Ripple voltage at 20MHz  $f_{in}$ .

#### IV. EFFICIENCY AND LOSS ANALYSIS OF ETPS

The efficiency of an ETPS is given by (12):

$$\eta_{ETPS} = \frac{P_{out(av)}}{P_{out(av)} + LA_{losses} + SC_{losses}} \quad (12)$$

where  $P_{out(av)}$  is the average output power of the ETPS,  $LA_{losses}$  represents the losses occurring in the linear amplifier stage, and  $SC_{losses}$  the losses occurring in the switching converter stage.

##### A. Linear Stage Losses

Class AB linear stage comprises of NMOS and PMOS devices, referring to Figure 3 and the occurring total losses are the sum of the losses in NMOS and PMOS devices, as stated in (13):

$$LA_{losses} = NMOS_{losses} + PMOS_{losses} \quad (13)$$

where:

$$NMOS_{losses} = (I_{SW} - I_{load}) V_{out}$$

$$PMOS_{losses} = (I_{load} - I_{SW}) (V_{dd} - V_{out})$$

##### B. Switching Stage Losses

Switching stage losses comprise of conduction losses in the diode and MOSFET, switching losses in the MOSFET, and driver losses:

$$SA_{losses} = Conduction_{losses} + Switching_{losses} + Driver_{losses} \quad (14)$$

where:

$$Conduction_{losses} = ON_{diode_{losses}} + ON_{MOS_{losses}} \quad (15)$$

$$ON_{diode_{losses}} = (1 - D) \cdot I_{SW_{dc}} \cdot V_{on_{diode}}$$

where  $D = V_{out_{dc}} / V_{dd}$ .

$$ON_{MOS_{losses}} = D \cdot I_{SW_{dc}}^2 \cdot R_{on} \quad (16)$$

where  $R_{on}$  is the ON resistance of the MOSFET.

$$Switching_{losses} = I_{SW} \cdot V_{off} \cdot (t_{on} + t_{off}) \cdot f_{sw_{ave}} / 2$$

where  $t_{on} = \frac{C_G V_G}{I_{driver(L \rightarrow H)}}$  and  $t_{off} = \frac{C_G V_G}{I_{driver(H \rightarrow L)}}$ .

$$Driver_{losses} = Q_g V_{gs} f_{sw_{ave}} \quad (17)$$

Linear region of operation:

$$f_{sw_{ave}} = \frac{R_{sense}}{L} \frac{V_{dd}}{2 \cdot h} D (1 - D \cdot \frac{V_{s_{rms}}^2}{V_{s_{dc}}^2})$$

Non-linear region of operation:

$$f_{sw_{ave}} = f_{i/p}$$

#### V. ETPS SIMULATION MEASUREMENTS FOR VARYING INPUT FREQUENCY

$V_{in} = 1.94 + 1.2 \sin(2 \cdot \pi \cdot f_{i/p} \cdot t)$  is the input signal considered for simulation. The input frequency  $f_{i/p}$  is varied over a wide range from 1MHz to 60MHz. Correspondingly, the average switching frequency of the switching converter,  $f_{sw}$  is noted down. Linear amplifier and switching converter losses are also obtained (Table II). The efficiency is calculated by (12).

It can be observed that the efficiency peaks at an input frequency of 2MHz. At this frequency,  $SR i_{sw} = SR i_{load}$  and  $f_{sw}$  is minimum. As a result, switching losses are minimum, as highlighted in Table II and the efficiency peaks, as shown in

Figure 12. For  $f_{i/p}$  lower than 2MHz,  $SR\ i_{sw} > SR\ i_{load}$  and hence the increased switching losses result in lowering the efficiency. For  $f_{i/p}$  higher than 2MHz,  $SR\ i_{sw} < SR\ i_{load}$ . In this condition, the switching converter will be able to provide only the DC components of the load current. As a result, the linear amplifier needs to provide the AC components, which results in increased linear stage losses and consequently, a drop in efficiency. Hence, the efficiency of a hysteretic current controlled parallel hybrid ETPS drops as the input signal frequency increases.

TABLE II. EFFICIENCY ANALYSIS FOR VARYING INPUT FREQUENCY

$f_{in}$ (MHz)	$f_{sw}$ (MHz)	$P_{out}$ (W)	$LA_{losses}$ (W)	$SC_{losses}$ (W)	$\eta$
1	6.6	0.094	0.00077	0.02151	80.8%
2	2	0.093	0.00217	0.0179	82.3%
5	5	0.094	0.03706	0.02333	60.9%
10	10	0.093	0.03915	0.02825	58.2%
20	20	0.093	0.03920	0.03821	54.6%
40	40	0.092	0.03906	0.05799	48.8%
60	60	0.093	0.03981	0.07759	44.2%

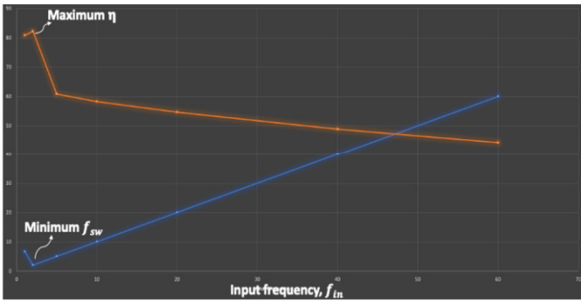


Fig. 12. Input frequency versus switching frequency and efficiency.

## VI. DESIGN CONSIDERATIONS

Various parameters like  $R_{sense}$ ,  $R_{load}$ ,  $L$ ,  $h$ ,  $V_{dd}$  need to be analyzed and considered while designing an ETPS. The main goal of an ETPS designer is to accurately reproduce the input envelope signal at maximum circuit efficiency.  $R_{sense}$  is chosen to be very much smaller than  $R_{load}$  to reduce the losses occurring in the current sense resistor. Inductor  $L$ , hysteresis  $h$ , and supply voltage  $V_{dd}'$  are parameters determined by the designer. The simulations conducted in Section V are taken as the base to study the variations in efficiency with respect to parameter variations. The Input signal considered is:

$$V_{in} = 1.94 + 1.2 \sin(2\pi f_{i/p} t)$$

### A. Inductor, $L$ Variations

The inductor value selected for simulation in Section V was 11μH. The average slew rate of the switching current is given by (18):

$$SR_{i_{sw\_ave}} = \frac{2}{L} \left(1 - \frac{V_{in\_dc}}{V_{dd}}\right) V_{in\_dc} \quad (18)$$

Reducing the inductor value to 5μH increases the slew rate of the switching current.  $SR\ i_{sw}$  exceeds  $SR\ i_{load}$  and hence  $SA_{losses}$  increases as shown in Table III. This brings down the

efficiency. Increasing the inductor value to 30μH reduces the slew rate of the switching current.  $SR\ i_{sw}$  goes below  $SR\ i_{load}$  as shown in Figure 13. As a result, the switching stage will be able to provide only the DC components of the load current and the linear stage will be forced to provide the AC components. This increases the losses occurring in the linear stage and consequently lowers the efficiency.

TABLE III. EFFICIENCY ANALYSIS FOR VARYING INDUCTOR VALUES

$f_{in} = 2\text{ MHz}$				
$L$ (μH)	$f_{sw}$ (MHz)	$LA_{losses}$ (W)	$SC_{losses}$ (W)	$\eta$
5	6.69	0.0007802	0.02147	80.1%
12	2	0.002172	0.01798296	82.3%
30	6.7	0.03757	0.02386	60.05%

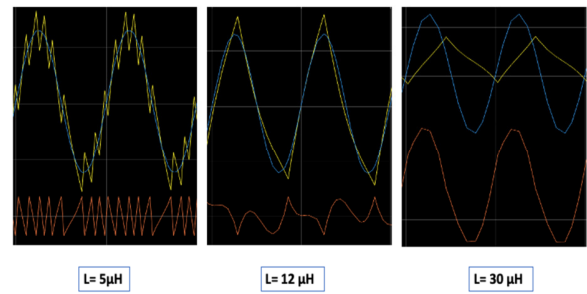


Fig. 13. Changes in slew rate as  $L$  changes.

### B. Hysteresis Variations

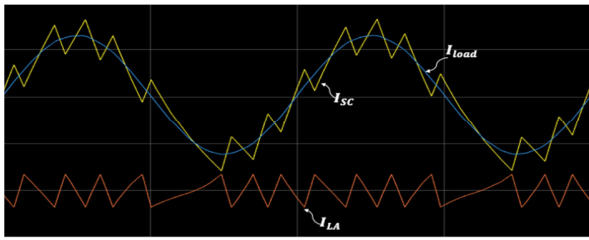
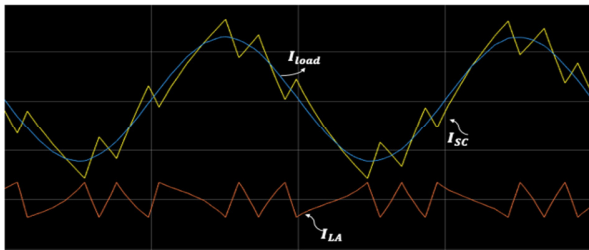
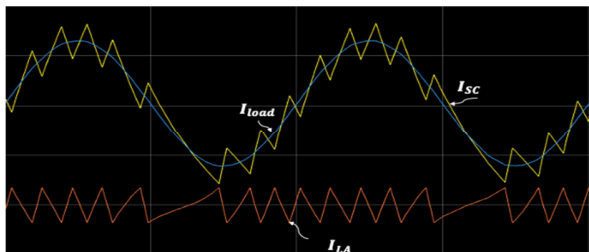
The hysteresis value selected for the simulation was  $\pm 7\text{mV}$ . When it reduces to  $\pm 2\text{mV}$ , the switching frequency increases tremendously and hence the switching losses increase as shown in Table IV. As a result, the efficiency drops. When the hysteresis value increases to  $\pm 20\text{mV}$ , the switching frequency of the converter reduces drastically and the switching converter now fails to provide the entire load current. As mentioned above, this results in increased losses occurring in linear stage. Efficiency again drops.

TABLE IV. EFFICIENCY ANALYSIS FOR VARYING HYSTERESIS VALUES

$h$ (mV)	$LA_{losses}$ (W)	$SC_{losses}$ (W)	$\eta$
2	0.006895	0.02237	76.3%
7	0.002207	0.01798	82.3%
20	0.03045	0.02001	65.1%

### C. $V_{dd}$ Variations

The selected  $V_{dd}$  for simulations in Section V is 5.5V. The average slew rate of the switching current is given by (16). Simulations are performed at 1MHz input frequency and the corresponding slew rates can be observed in Figure 14. The corresponding average switching frequency obtained from simulations is 8MHz. As the  $V_{dd}$  drops to 4.5V, the  $SR_{i_{sw\_ave}}$  drops as shown in Figure 15. The corresponding average switching frequency obtained from simulations is reduced to 6MHz.

Fig. 14.  $SR_{isw}$  and  $SR_{iload}$  at  $V_{dd} = 5.5V$ .Fig. 15.  $SR_{isw}$  and  $SR_{iload}$  at  $V_{dd} = 4.5V$ .Fig. 16.  $SR_{isw}$  and  $SR_{iload}$  at  $V_{dd} = 6.5V$ .

As the  $V_{dd}$  increases to 6.5V, the  $SR_{isw\_ave}$  increases as shown in Figure 16. The obtained corresponding average switching frequency has increased to 9MHz. Hence, by varying the supply voltage  $V_{dd}$  of the switching converter, the slew rates of the switching current can be varied to match with the slew rates of the load current and thereby to obtain maximum efficiency.

## VII. CONCLUSION

The current paper presents a novel idea to the hysteretic current controlled parallel hybrid ETPS designers, that they should focus on matching the slew rates of switching and load currents for all frequencies of operation, in order to obtain maximum efficiency at all operating frequencies. In this paper, the hysteretic current controlled parallel hybrid ETPS is implemented in Simulink and simulations were performed using a sinusoidal envelope input signal. Efficiency analysis is performed by varying the input frequency from 1MHz to 60MHz. It was observed that the efficiency peaks for an input frequency of 2MHz, when the slew rates of the load current and the switching current matches. Ripple voltage was found to be increasing from 5.5mV to 28mV as the region of operation moved from linear to non-linear. Various circuit parameters that the designer needs to consider while designing an ETPS and its trade-offs are also discussed in length. The simulated ETPS could work for a peak efficiency of 82.3% for a switching frequency of 2MHz, while tracking the sinusoidal

envelope input signal at 2MHz input frequency. One major advantage of the current work is that it is using a single level switching converter whereas most of the existing works are using multi-level switching converters, which increases complexity.

## REFERENCES

- [1] Z. Wang, "Demystifying Envelope Tracking: Use for High-Efficiency Power Amplifiers for 4G and Beyond," *IEEE Microwave Magazine*, vol. 16, no. 3, pp. 106–129, Apr. 2015, <https://doi.org/10.1109/MMM.2014.2385351>.
- [2] B. Kim *et al.*, "Push the Envelope: Design Concepts for Envelope-Tracking Power Amplifiers," *IEEE Microwave Magazine*, vol. 14, no. 3, pp. 68–81, Feb. 2013, <https://doi.org/10.1109/MMM.2013.2240851>.
- [3] A. K. Kwan, M. Younes, R. Darraji, and F. M. Ghannouchi, "On Track for Efficiency: Concurrent Multiband Envelope-Tracking Power Amplifiers," *IEEE Microwave Magazine*, vol. 17, no. 5, pp. 46–59, Feb. 2016, <https://doi.org/10.1109/MMM.2016.2525118>.
- [4] P. Asbeck and Z. Popovic, "ET Comes of Age: Envelope Tracking for Higher-Efficiency Power Amplifiers," *IEEE Microwave Magazine*, vol. 17, no. 3, pp. 16–25, Mar. 2016, <https://doi.org/10.1109/MMM.2015.2505699>.
- [5] A. Babu, B. G. Shivaaleelavathi, and V. Yatnalli, "Enhanced Power Added Efficiency for RF Power Amplifiers in Mobile Communication," *International Journal of Grid and Distributed Computing*, vol. 13, no. 2, pp. 1543–1553, Dec. 2020.
- [6] H. He, T. Ge, and J. Chang, "A review on supply modulators for Envelope-Tracking Power Amplifiers," in *2016 International Symposium on Integrated Circuits (ISIC)*, Sep. 2016, pp. 1–4, <https://doi.org/10.1109/ISICIR.2016.7829694>.
- [7] X. Ruan, Y. Wang, and Q. Jin, "A review of envelope tracking power supply for mobile communication systems," *CPSS Transactions on Power Electronics and Applications*, vol. 2, no. 4, pp. 277–291, Dec. 2017, <https://doi.org/10.24295/CPSS/TPPEA.2017.00026>.
- [8] M. Vasić, O. Garcia, J. A. Oliver, P. Alou, and J. A. Cobos, "Survey of architectures and optimizations for wide bandwidth envelope amplifier," in *2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC)*, Novi Sad, Serbia, Sep. 2012, pp. LS8d.1-1-LS8d.1-7, <https://doi.org/10.1109/EPEPEMC.2012.6397529>.
- [9] A. Babu, B. G. Shivaaleelavathi, and V. Yatnalli, "A Comprehensive Review on Supply Modulators and Control Strategies for Envelope Tracking RF Power Amplifiers in Mobile Communication," *IETE Journal of Research*, Oct. 2021, <https://doi.org/10.1080/03772063.2021.1987998>.
- [10] M. Azhar and A. Shabbir, "5G Networks: Challenges and Techniques for Energy Efficiency," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2864–2868, Apr. 2018, <https://doi.org/10.48084/etasr.1623>.
- [11] I. Androulidakis and G. Kandus, "Mobile Phone Brand Categorization vs. Users' Security Practices," *Engineering, Technology & Applied Science Research*, vol. 1, no. 2, pp. 30–35, Apr. 2011, <https://doi.org/10.48084/etasr.19>.
- [12] H. Meng, "Analysis and design consideration of hybrid supply modulator for envelope tracking power amplifier," in *2014 9th IEEE Conference on Industrial Electronics and Applications*, Hangzhou, China, Jun. 2014, pp. 151–154, <https://doi.org/10.1109/ICIEA.2014.6931149>.
- [13] F. Wang *et al.*, "An Improved Power-Added Efficiency 19-dBm Hybrid Envelope Elimination and Restoration Power Amplifier for 802.11g WLAN Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 12, pp. 4086–4099, Sep. 2006, <https://doi.org/10.1109/TMTT.2006.885575>.
- [14] J. Choi, D. Kim, D. Kang, and B. Kim, "A Polar Transmitter With CMOS Programmable Hysteretic-Controlled Hybrid Switching Supply Modulator for Multistandard Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 7, pp. 1675–1686, Jul. 2009, <https://doi.org/10.1109/TMTT.2009.2021880>.

- [15] J. Ham *et al.*, "CMOS Power Amplifier Integrated Circuit With Dual-Mode Supply Modulator for Mobile Terminals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 1, pp. 157–167, Jan. 2016, <https://doi.org/10.1109/TCSI.2015.2512703>.
- [16] J.-S. Paek *et al.*, "A – 137 dBm/Hz Noise, 82% Efficiency AC-Coupled Hybrid Supply Modulator With Integrated Buck-Boost Converter," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 11, pp. 2757–2768, Aug. 2016, <https://doi.org/10.1109/JSSC.2016.2604296>.
- [17] Y. Jing and B. Bakkaloglu, "A High Slew-Rate Adaptive Biasing Hybrid Envelope Tracking Supply Modulator for LTE Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 9, pp. 3245–3256, Sep. 2017, <https://doi.org/10.1109/TMTT.2017.2678476>.
- [18] Y. Wang, X. Ruan, Y. Leng, and Y. Li, "Hysteresis Current Control for Multilevel Converter in Parallel-Form Switch-Linear Hybrid Envelope Tracking Power Supply," *IEEE Transactions on Power Electronics*, vol. 34, no. 2, pp. 1950–1959, Oct. 2019, <https://doi.org/10.1109/TPEL.2018.2835640>.
- [19] D. Kimball, T. Nakatani, J. Yan, P. T. Theilmann, and I. Telleiz, "High efficiency envelope tracking power amplifiers for wide modulation bandwidth signals (invited)," in *2014 Asia-Pacific Microwave Conference*, Sendai, Japan, Aug. 2014, pp. 103–106.
- [20] J. Kim *et al.*, "Wideband envelope amplifier for envelope-tracking operation of handset power amplifier," in *2014 44th European Microwave Conference*, Rome, Italy, Jul. 2014, pp. 1352–1355, <https://doi.org/10.1109/EuMC.2014.6986695>.
- [21] J. Kim, D. Kim, Y. Cho, D. Kang, and B. Kim, "Envelope-tracking power amplifier with enhanced back-off efficiency using average switch current control of supply modulator," in *2013 Asia-Pacific Microwave Conference Proceedings (APMC)*, Seoul, Korea, Aug. 2013, pp. 435–437, <https://doi.org/10.1109/APMC.2013.6694822>.
- [22] H. He, Y. Kang, T. Ge, L. Guo, and J. S. Chang, "A 2.5-W 40-MHz-Bandwidth Hybrid Supply Modulator With 91% Peak Efficiency, 3-V Output Swing, and 4-mV Output Ripple at 3.6-V Supply," *IEEE Transactions on Power Electronics*, vol. 34, no. 1, pp. 712–723, Jan. 2019, <https://doi.org/10.1109/TPEL.2018.2827396>.
- [23] J.-S. Paek, D. Kim, Y. Choo, Y.-S. Youn, J. Lee, and T. B.-H. Cho, "Design of Boosted Supply Modulator With Reverse Current Protection for Wide Battery Range in Envelope Tracking Operation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 1, pp. 183–194, Jan. 2019, <https://doi.org/10.1109/TMTT.2018.2879323>.
- [24] A. Shabbir, H. R. Khan, S. A. Ali, and S. Rizvi, "Design and Performance Analysis of Multi-tier Heterogeneous Network through Coverage, Throughput and Energy Efficiency," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2345–2350, Dec. 2017, <https://doi.org/10.48084/etasr.1256>.

# Seismic Fragility of a Single Pillar-Column Under Near and Far Fault Soil Motion with Consideration of Soil-Pile Interaction

**Foudhil Lemsara**

LGC-ROI, Department of Civil Engineering, Faculty of Technology, University of Batna 2, Algeria  
f.lemsara@univ-batna2.dz  
(corresponding author)

**Tayeb Bouzid**

LGC-ROI, Department of Civil Engineering, Faculty of Technology, University of Batna 2, Algeria  
tayeb.bouzid@univ-batna2.dz

**Djarir Yahiaoui**

LGC-ROI, Department of Civil Engineering, Faculty of Technology, University of Batna 2, Algeria  
d.yahiaoui@univ-batna2.dz

**Belgacem Mamen**

Department of Civil Engineering, Faculty of Science and Technology, University of Abbès Laghrou  
Khenchela, Algeria  
belgacem.mamen@univ-khenchela.dz

**Mohamed Saadi**

LGC-ROI, Department of Civil Engineering, Faculty of Technology, University of Batna 2, Algeria  
m.saadi@univ-batna2.dz

*Received: 10 October 2022 | Revised: 30 October 2022 | Accepted: 31 October 2022*

## ABSTRACT

The soil-structure interaction is a significant challenge faced by civil engineers due to the complexity potential in terms of seismic fragility evaluation. This paper presents a seismic fragility estimation of a single pier considering seismic ground motion types. Furthermore, sand type, pile diameter, pier height, and mass variation were considered to estimate their effect on the seismic fragility of the concrete pier. Incremental dynamic analysis was performed using a beam on a nonlinear Winkler foundation model. The analysis model condition compared near- and far-ground motion effects. Dynamic analysis and fragility assessment of the single-pier structure showed that low mass center produced less vulnerability of the concrete pier in the two cases of the sand type under near- and far-ground motions. The near and far earthquake simulations at complete failure probability had a difference of less than 5% when  $0.65s < T_1 < 1s$  and  $2.4 < T_1/T_2$ , but the opposite was shown when  $T_1 < 0.5s$  and  $3 < T_1/T_2$  were present together.

**Keywords-**soil-pile-structure interaction; incremental dynamic analysis; BNWFmodel; ground motion types; seismic fragility

## I. INTRODUCTION

Under seismic loading, liquefiable soil increases the seismic fragility of expansion bearing, piles, and embankment soil. The fragility of common components, such as columns, depends on the overlying liquefiable sand. The effect of soil strength for clay and sandy sites on the seismic performance of

skewed bridge components and its relation with the skew angle was studied in [1, 2]. The seismic fragility of a pile was studied with different seismic demands in [3]. The nonlinear Winkler foundation model is widely used to study soil-pile and soil-pile-structure interactions using nonlinear dynamic analysis [4]. In [5], a comparison of soil-structure interaction and fix-base model effects on a structure's seismic fragility was presented.



A near-fault earthquake severely damages structures more than a far-fault one. This comparison was investigated in [6] through a numerical simulation of the seismic damage of arch bridges with three earthquake indicators: area intensity, energy rate, and cumulative absolute velocity. The demand/capacity ratio of an arch bridge pier is higher under near- than far-fault earthquakes [7]. In addition to [4], pushover analysis was used to study soil-pile and soil-pile-structure interactions in [8]. In a single pile, columns fail before pile foundations [9]. Soil permeability can be considered a complex property due to its influence on seismic fragility with different soil types [10]. In [11], the impact of the soil-pile interaction on the seismic fragility of wharf structures was investigated by comparing two systems: with and without soil-pile interaction. Most recent studies were based on the Winkler foundation model to study the seismic fragility of the Soil-Pile-Structure interaction (S-P-S). This model is a good instrument for investigating the seismic fragility of concrete bridge-soil systems using stepwise and LASSO regression [12]. In [13], the mitigation of pile group spacing on scour and liquefaction effects was studied.

The same results with [5] were shown in [14] considering a skewed bridge. In [15], the effect of irregular configuration and the stiffness of the substructure on the vulnerability of the bridge was shown. The soil-pile and soil-pile-structure interactions were investigated more comprehensively in [16]. The effects of the soil-structure interaction on the seismic fragility of a bridge were investigated in [17], considering the effects of wave passage. The selection of foundation type is an essential key for the construction of a superstructure and estimating its seismic fragility. The pile foundation is an improvement in superstructure performance compared to the shallow foundation [18]. The effects of soil stiffness and pile flexibility on the seismic fragility of the pile were assessed in [19] using finite element analysis. In addition to the effect of infrastructure parameters on the seismic fragility of the structure, there is the influence of the pier parameters, such as the higher-order mode of the pier bridge column that produces overestimated seismic demands of the pile foundation in terms of curvature and displacement [20]. The lateral performance of an eccentrically braced frame was investigated in [21]. The effectiveness of retrofitting a reinforced concrete structure was shown in [22]. In [23], the seismic performance of a base-isolated nuclear power plant structure was investigated taking into account near and far-fault earthquake influence.

This study investigated the effects of near- and far-fault earthquakes on the seismic fragility of a single-column pile, taking into account the soil-pile spring system's  $p$ - $y$  curve using the Seismostruct software and incremental dynamic analysis.

## II. NEAR AND FAR GROUND MOTION SELECTIONS

The main benchmark for selecting near- and far-fault earthquakes is the closest distance from the fault (Rjb) where the near-fault has an Rjb of less than 15Km [24]. This study used the records of earthquakes from [25]. The ground motions were characterized by a median of PGV/PGA equal to 113 for near- and 119 for far-fault. In Figure 1(a, b), the thin lines define the individual spectra, and the thick line defines the mean spectrum of near (NR) and far (FR) ground motions.

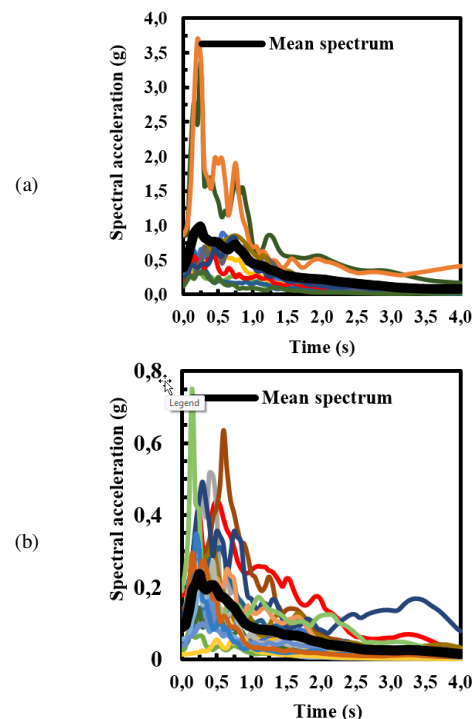


Fig. 1. Acceleration response spectra of (a) near-fault earthquakes, (b) far-fault earthquakes.

## III. NUMERICAL MODELING OF SOIL-PILE INTERACTION

The Seismostruct software offers different concrete materials and steel reinforcing bars. The nonlinear Mander [26] and the Menegetto-Pinto models of steel reinforcing are shown in Figures 2 and 3. The modeling of the soil-pile interaction was based on the design of the  $p$ - $y$  curve relationship:  $p$  is the soil reaction and  $y$  is the lateral deflection. The nonlinear Winkler foundation model (BNWF) [4] was employed, defined as multilinear curves. The structure was modeled using a 3D formulation, whereas the soil was modeled using a 1D formulation based on the reference  $p$ - $y$  model as [4]:

$$p = 0.9p_u \cdot \tanh\left(\frac{k+z}{0.9p_u}y\right) \quad (1)$$

where  $k$  is the initial modulus of the subgrade reaction in soil and  $p_u$  is the ultimate bearing capacity.

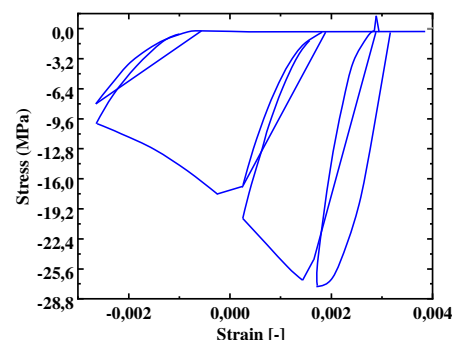


Fig. 2. Stress-strain curve of concrete mander material.



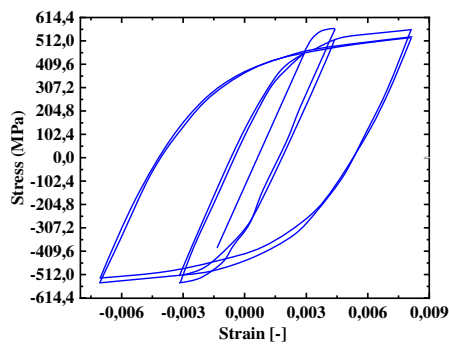


Fig. 3. Stress-strain curve of steel Menegetto-Pinto model.

The properties of each  $p$ - $y$  curve derive from the parameters of homogeneous soil layers considering spring spacing of 0.5. Two different pile diameter deposits, each with two sand types, loose and dense, and three different masses were considered: 1500, 3000, and 4500kN. For simplification, an abbreviation is proposed concerning different models: LS for the loose sand model, DS for the dense sand model, and  $d$  for pile diameter, where  $d=d_1=1.5$ m and  $d=d_2=2$ m. Two different heights ( $H$ ) of piers, equal to 5 and 10m, were used. The pile length ( $L$ ) was fixed at 30m as in [27]. The loose and dense sand had a friction angle of  $30^\circ$  and  $45^\circ$ , respectively. Figure 4 shows the soil-pile interaction model designed in Seismostruct software.

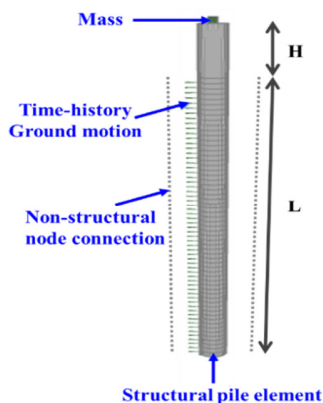


Fig. 4. Soil-pile spring interaction.

The calculation of the probability exceedance of a structure at a limit state is defined in (2), which is given by [15, 28-29]:

$$p\left(\frac{D}{PGA}\right) = \Phi\left(\frac{\ln(PGA) - \mu}{\sigma}\right) \quad (2)$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviation of the logarithmic of the peak ground acceleration when the pier reaches the threshold of performance level D at each limit state, and  $\Phi$  is the standard normal cumulative distribution function. To estimate the probability of pier damage, the analysis procedure was (Figure 5):

1. Modeling of soil-pile interaction and vertical load value using Seismostruct software.
2. Define the drift value at each limit state of different models using nonlinear static analysis.

3. Select ground motion records (near- and far-fault earthquakes).
4. Generate the incremental dynamic analysis and find the median and standard deviation of the incremental dynamic analysis.
5. Drawing of seismic fragility curve.

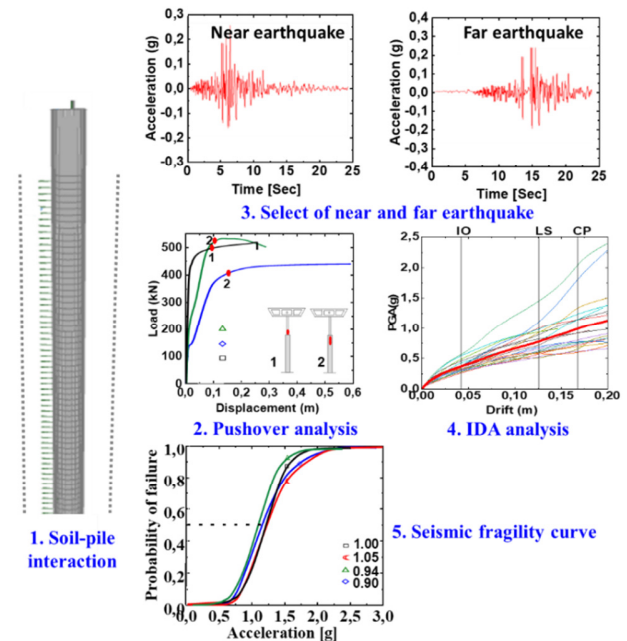


Fig. 5. Analysis procedure.

## IV. RESULTS AND DISCUSSION

### A. Limit States and Incremental Dynamic Analysis

Figure 6(a) shows the form of the incremental dynamic analysis example. The 20 thin lines define the individual incremental dynamic analysis and the thick red line defines its mean response. Figure 6(b) shows the comparison of the mean response. The mean response of the incremental dynamic analysis increases with mass center under the two ground motion types. After realizing the incremental dynamic analysis series for each model, their mean and standard deviation were calculated. The mean and standard deviation of incremental dynamic analysis in the case of pile diameter of 1.5 and 2.0 m, and  $H=5$ m are shown in Tables I and II, respectively, and the mean value decreases with mass increase.

TABLE I. MEDIAN AND STANDARD DEVIATION VALUES OF IDA WITH PILE DIAMETER EQUAL TO 1.5 M

Mass (KN)		1500		3000		4500	
Earthquake type		Near	Far	Near	Far	Near	Far
Mean ( $\mu$ )	Loose sand	0.94	1.03	0.38	0.37	0.23	0.22
	Dense sand	0.90	1.00	0.40	0.38	0.25	0.23
Standard deviation ( $\sigma$ )	Loose sand	0.36	0.37	0.32	0.35	0.28	0.31
	Dense sand	0.28	0.36	0.26	0.26	0.23	0.22

TABLE II. MEDIAN AND STANDARD DEVIATION VALUES OF IDA WITH PILE DIAMETER EQUAL TO 2M.

Mass (KN)		1500		3000		4500	
Earthquake type		Near	Far	Near	Far	Near	Far
Mean ( $\mu$ )	Loose sand	1.73	1.70	0.82	0.76	0.50	0.47
	Dense sand	1.98	2.01	0.80	0.55	0.42	0.48
Standard deviation ( $\sigma$ )	Loose sand	0.36	0.32	0.32	0.32	0.30	0.32
	Dense sand	0.24	0.32	0.23	0.49	0.17	0.24

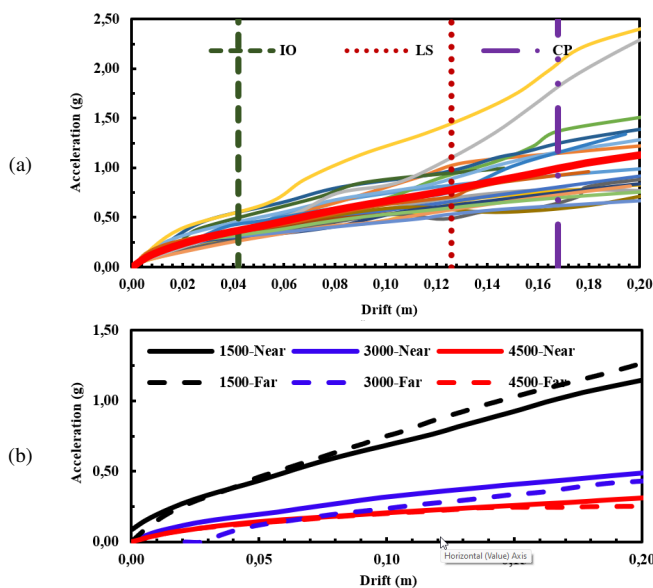


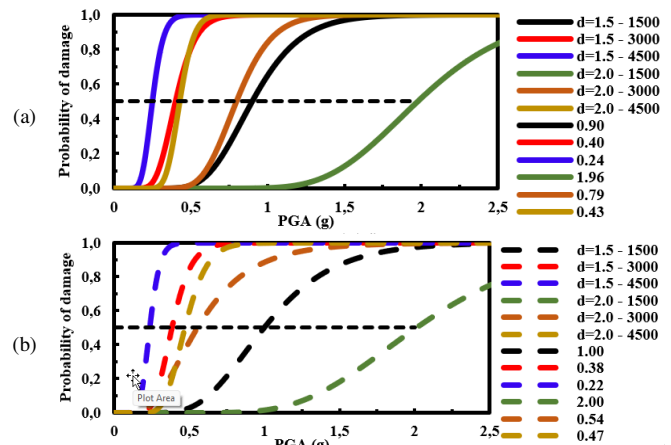
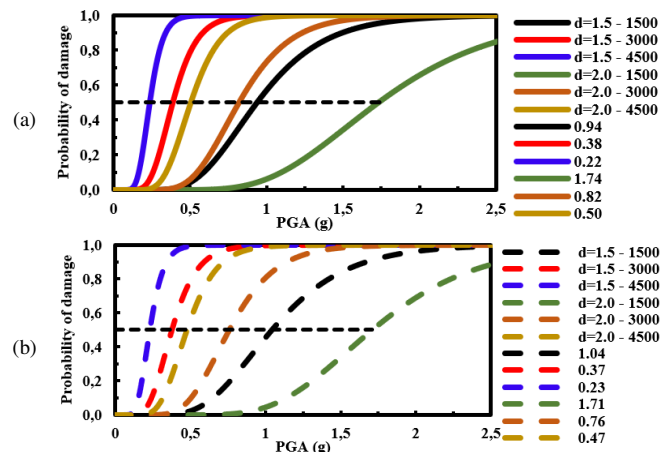
Fig. 6. (a) Incremental dynamic analysis under NR ground motion, (b) Example of mean response comparison of IDA.

### B. Fragility Curves

Utilizing parameters from the mean and standard deviation of incremental dynamic analysis, and the standard normal cumulative distribution function (2), the fragility curves are expected and generated in the design figures for the collapse prevention state.

### C. Effect of Mass

Depending on the variation of pier height ( $H$ ), the results of the numerical simulation are classified into two parts: Figures 7 and 8 for  $H=5\text{m}$  and Figures 9 and 10 for  $H=10\text{m}$ . Figure 7(a) presents the seismic fragility of the concrete pier in case the pile is embedded in dense sand under a near-fault ground motion. This figure shows that when placing a different mass of 1500, 3000, and 4500kN, the pier structure needs a peak ground acceleration equal to 0.90, 0.40, and 0.25g, respectively, to exceed 50% of failure probability. These results show that the high mass needs a low Peak Ground Acceleration (PGA) to exceed 50% of failure probability under near-fault ground motion. The seismic fragility curves under far-fault ground motion show the same trend. All figures show similar trending results with a mass variation.

Fig. 7. Seismic fragility of pier with dense sand condition and  $H=5\text{m}$ : (a) Near (NR) ground motions, (b) Far (FR) ground motions.Fig. 8. Seismic fragility of a pier with loose sand conditions and  $H=5\text{m}$ : (a) Near (NR) ground motions, (b) Far (FR) ground motions.

### D. Near and Far Fault Ground Motion Comparison

These results show that high mass needs a low PGA to exceed 50% of failure probability under near-fault ground motion. A comparison of seismic fragility curves under far-fault ground motion shows the same trend. Additionally, all Figures show similar trending results with mass variation. Figures 7 and 8 showed that the lower difference between PGA of near- and far-ground motion at 50% of failure probability in the case of  $H=5\text{m}$  are: LS-d1-3000, DS-d1-3000, LS-d1-4500, LS-d1-4500, LS-d2-4500. These models keep this most negligible difference value at complete failure probability (100%). The fundamental periods of the different models are 0.81, 0.69, 0.98, 0.85, and 0.56s, and their  $T_1/T_2$  ratios were 2.89, 2.46, 2.88, 2.5, and 2.94, respectively.

On the other hand, some models have a slight difference at only 50% failure probability. The last figures show that in models presenting  $R < 5\%$ , the far-fault produced a higher failure probability than the near-fault ground motion due to its higher value of PGV/PGA. These models were LS-d2-1500 and LS-d2-3000 in the case of  $H=5\text{m}$  and DS-d2-1500 in the case of  $H=10\text{m}$ . The fundamental periods of these models were

0.41, 0.56, and 0.64s, and their  $T_1/T_2$  ratios were 3.72, 3.5, and 1.82, respectively. These Figures showed a higher probability of failure of the pier structure in far-fault earthquakes than in near-ground motions, as observed in [31], while a higher probability of failure was presented in [23] in near-fault earthquakes in all events.

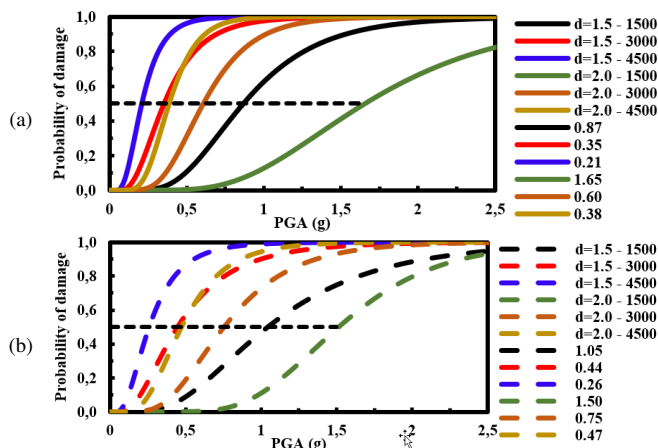


Fig. 9. Seismic fragility of pier with dense sand condition and  $H=10\text{m}$ : (a) Near (NR) ground motions, (b) Far (FR) ground motions.

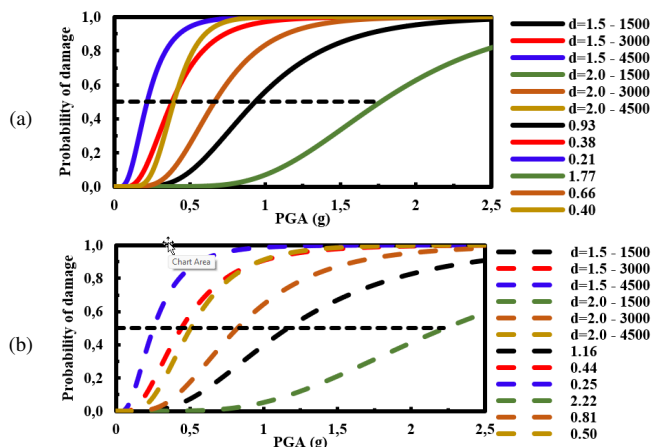


Fig. 10. Seismic fragility of pier with loose sand condition and  $H=10\text{m}$ : (a) Near (NR) ground motions, (b) Far (FR) ground motions.

## V. CONCLUSIONS

This study compared the seismic fragility of a single concrete pier under near- and far-fault earthquakes. The effects of mass variation, closest ground motion type, and pile diameter were considered with two sand models, loose and dense. The seismic fragility of a concrete pier exceeded 50% and 100% of failure probability when placing a higher mass value at the pier end in all cases of pile section. The difference between the peak ground acceleration of near- and far-fault earthquakes at complete failure probability was less than 5% when the analysis modal conditions ( $0.65\text{s} < T_1 < 1\text{s}$  and  $2.4 < T_1/T_2$ ) were verified. The pier structure was highly affected by a higher value of PGV/PGA regardless of the near and far ground motions if one of two conditions was given:

$1 < T_1 < 0.65\text{s}$  or  $T_1/T_2 < 2$ . The higher difference between the peak ground acceleration of near and far earthquakes at complete failure probability was concluded if the two following conditions are presented together:  $T_1 < 0.5\text{s}$  and  $3 < T_1/T_2$ .

## REFERENCES

- [1] B. Aygün, L. Dueñas-Osorio, J. E. Padgett, and R. DesRoches, "Efficient Longitudinal Seismic Fragility Assessment of a Multispan Continuous Steel Bridge on Liquefiable Soils," *Journal of Bridge Engineering*, vol. 16, no. 1, pp. 93–107, Jan. 2011, [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0000131](https://doi.org/10.1061/(ASCE)BE.1943-5592.0000131).
- [2] A. R. Ghotbi, "Performance-based seismic assessment of skewed bridges with and without considering soil-foundation interaction effects for various site classes," *Earthquake Engineering and Engineering Vibration*, vol. 13, no. 3, pp. 357–373, Sep. 2014, <https://doi.org/10.1007/s11803-014-0248-7>.
- [3] A. R. Ghotbi, "Performance-based seismic assessment of a large diameter extended pile shaft in a cohesionless soil," *Earthquake Engineering and Engineering Vibration*, vol. 14, no. 1, pp. 177–188, Mar. 2015, <https://doi.org/10.1007/s11803-015-0015-4>.
- [4] A. Tombari, M. H. El Naggar, and F. Dezi, "Impact of ground motion duration and soil non-linearity on the seismic performance of single piles," *Soil Dynamics and Earthquake Engineering*, vol. 100, pp. 72–87, Sep. 2017, <https://doi.org/10.1016/j.soildyn.2017.05.022>.
- [5] S. P. Stefanidou, A. G. Sextos, A. N. Kotsoglou, N. Lesgidis, and A. J. Kappos, "Soil-structure interaction effects in analysis of seismic fragility of bridges using an intensity-based ground motion selection procedure," *Engineering Structures*, vol. 151, pp. 366–380, Nov. 2017, <https://doi.org/10.1016/j.engstruct.2017.08.033>.
- [6] N. Simos, G. C. Manos, and E. Kozikopoulos, "Near- and far-field earthquake damage study of the Konitsa stone arch bridge," *Engineering Structures*, vol. 177, pp. 256–267, Dec. 2018, <https://doi.org/10.1016/j.engstruct.2018.09.072>.
- [7] I. Mohseni, H. A. Lashkariani, J. Kang, and T. H. K. Kang, "Dynamic Response Evaluation of Long-Span Reinforced Arch Bridges Subjected to Near- and Far-Field Ground Motions," *Applied Sciences*, vol. 8, no. 8, Aug. 2018, Art. no. 1243, <https://doi.org/10.3390/app8081243>.
- [8] G. Houda, B. Tayeb, and D. Yahiaoui, "Key parameters influencing performance and failure modes for interaction soil–pile–structure system under lateral loading," *Asian Journal of Civil Engineering*, vol. 19, no. 3, pp. 355–373, Apr. 2018, <https://doi.org/10.1007/s42107-018-0033-4>.
- [9] M. Čosić, R. Folić, and B. Folić, "Fragility and reliability analyses of soil - pile - bridge pier interaction," *Facta universitatis - series: Architecture and Civil Engineering*, vol. 16, no. 1, pp. 93–111, 2018, <https://doi.org/10.2298/FUACE170420008C>.
- [10] L. Su *et al.*, "Seismic fragility analysis of pile-supported wharves with the influence of soil permeability," *Soil Dynamics and Earthquake Engineering*, vol. 122, pp. 211–227, Jul. 2019, <https://doi.org/10.1016/j.soildyn.2019.04.003>.
- [11] L. Su, H. P. Wan, Y. Dong, D. M. Frangopol, and X.-Z. Ling, "Seismic fragility assessment of large-scale pile-supported wharf structures considering soil-pile interaction," *Engineering Structures*, vol. 186, pp. 270–281, May 2019, <https://doi.org/10.1016/j.engstruct.2019.02.022>.
- [12] Y. Xie and R. DesRoches, "Sensitivity of seismic demands and fragility estimates of a typical California highway bridge to uncertainties in its soil-structure interaction modeling," *Engineering Structures*, vol. 189, pp. 605–617, Jun. 2019, <https://doi.org/10.1016/j.engstruct.2019.03.115>.
- [13] X. Wang, A. Ye, and B. Ji, "Fragility-based sensitivity analysis on the seismic performance of pile-group-supported bridges in liquefiable ground undergoing scour potentials," *Engineering Structures*, vol. 198, Nov. 2019, Art. no. 109427, <https://doi.org/10.1016/j.engstruct.2019.109427>.
- [14] H. R. Noori, M. M. Memarpour, M. Yakhchalian, and S. Soltanieh, "Effects of ground motion directionality on seismic behavior of skewed bridges considering SSI," *Soil Dynamics and Earthquake Engineering*, vol. 127, Dec. 2019, Art. no. 105820, <https://doi.org/10.1016/j.soildyn.2019.105820>.

- [15] S. Soltanieh, M. M. Memarpour, and F. Kilanehei, "Performance assessment of bridge-soil-foundation system with irregular configuration considering ground motion directionality effects," *Soil Dynamics and Earthquake Engineering*, vol. 118, pp. 19–34, Mar. 2019, <https://doi.org/10.1016/j.soildyn.2018.11.006>.
- [16] K. Sekhri, D. Yahiaoui, and K. Abbache, "Inelastic Response of Soil-Pile-Structure Interaction System under Lateral Loading: A Parametric Study," *Jordan Journal of Civil Engineering*, vol. 14, no. 2, pp. 250–266, 2020.
- [17] O. M. O. Ramadan, S. S. F. Mehanny, and A. A.-M. Kotb, "Assessment of seismic vulnerability of continuous bridges considering soil-structure interaction and wave passage effects," *Engineering Structures*, vol. 206, Mar. 2020, Art. no. 110161, <https://doi.org/10.1016/j.engstruct.2019.110161>.
- [18] M. Ansari, M. Nazari, and A. K. Panah, "Influence of foundation flexibility on seismic fragility of reinforced concrete high-rise buildings," *Soil Dynamics and Earthquake Engineering*, vol. 142, Mar. 2021, Art. no. 106521, <https://doi.org/10.1016/j.soildyn.2020.106521>.
- [19] D. Forcellini, "Analytical Fragility Curves of Pile Foundations with Soil-Structure Interaction (SSI)," *Geosciences*, vol. 11, no. 2, Feb. 2021, Art. no. 66, <https://doi.org/10.3390/geosciences11020066>.
- [20] X. Chen, N. Xiang, and C. Li, "Influence of higher-order modes of slender tall pier bridge columns on the seismic performance of pile foundations," *Soil Dynamics and Earthquake Engineering*, vol. 142, Mar. 2021, Art. no. 106543, <https://doi.org/10.1016/j.soildyn.2020.106543>.
- [21] F. Abdelhamid, D. Yahiaoui, M. Saadi, and N. Lahbari, "Lateral Reliability Assessment of Eccentrically Braced Frames Including Horizontal and Vertical Links Under Seismic Loading," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8278–8283, Apr. 2022, <https://doi.org/10.48084/etasr.4749>.
- [22] M. Saadi and D. Yahiaoui, "The Effectiveness of Retrofitting RC Frames with a Combination of Different Techniques," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8723–8727, Jun. 2022, <https://doi.org/10.48084/etasr.4979>.
- [23] V. B. Tran, S. M. Nguyen, T. H. Nguyen, V. H. Nguyen, T. T. H. Doan, and D. D. Nguyen, "The Influence of Near- and Far-field Earthquakes on the Seismic Performance of Base-Isolated Nuclear Power Plant Structures," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9092–9096, Oct. 2022, <https://doi.org/10.48084/etasr.5156>.
- [24] M. Bhandari, S. D. Bharti, M. K. Shrimali, and T. K. Datta, "Seismic Fragility Analysis of Base-Isolated Building Frames Excited by Near- and Far-Field Earthquakes," *Journal of Performance of Constructed Facilities*, vol. 33, no. 3, Jun. 2019, Art. no. 04019029, [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0001298](https://doi.org/10.1061/(ASCE)CF.1943-5509.0001298).
- [25] "PEER Ground Motion Database - PEER Center." <https://ngawest2.berkeley.edu/> (accessed Nov. 06, 2022).
- [26] J. B. Mander, M. J. N. Priestley, and R. Park, "Theoretical Stress-Strain Model for Confined Concrete," *Journal of Structural Engineering*, vol. 114, no. 8, pp. 1804–1826, Aug. 1988, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1988\)114:8\(1804\)](https://doi.org/10.1061/(ASCE)0733-9445(1988)114:8(1804)).
- [27] N. Gerolymos, V. Drosos, and G. Gazetas, "Seismic response of single-column bent on pile: evidence of beneficial role of pile and soil inelasticity," *Bulletin of Earthquake Engineering*, vol. 7, no. 2, Apr. 2009, Art. no. 547, <https://doi.org/10.1007/s10518-009-9111-z>.
- [28] M. Miari and R. Jankowski, "Incremental dynamic analysis and fragility assessment of buildings founded on different soil types experiencing structural pounding during earthquakes," *Engineering Structures*, vol. 252, Feb. 2022, Art. no. 113118, <https://doi.org/10.1016/j.engstruct.2021.113118>.
- [29] X. Guo, Y. Wu, and Y. Guo, "Time-dependent seismic fragility analysis of bridge systems under scour hazard and earthquake loads," *Engineering Structures*, vol. 121, pp. 52–60, Aug. 2016, <https://doi.org/10.1016/j.engstruct.2016.04.038>.
- [30] D. Vamvatsikos and C. A. Cornell, "Incremental dynamic analysis," *Earthquake Engineering & Structural Dynamics*, vol. 31, no. 3, pp. 491–514, 2002, <https://doi.org/10.1002/eqe.141>.
- [31] A. Mosleh, M. S. Razzaghi, J. Jara, and H. Varum, "Seismic fragility analysis of typical pre-1990 bridges due to near- and far-field ground motions," *International Journal of Advanced Structural Engineering (IJASE)*, vol. 8, no. 1, pp. 1–9, Mar. 2016, <https://doi.org/10.1007/s40091-016-0108-y>.

# Spatial Distribution and Health Risk Assessment based on Groundwater Fluoride Enrichment in Gaya, Bihar, India

**Shaz Ahmad**

Department of Civil Engineering, National Institute of Technology Patna, India  
shaza.phd19.ce@nitp.ac.in

**Reena Singh**

Department of Civil Engineering, National Institute of Technology Patna, India  
reena@nitp.ac.in  
(corresponding author)

Received: 6 October 2022 | Revised: 27 October 2022 | Accepted: 4 November 2022

## ABSTRACT

Fluoride can be hazardous to the body if consumed in excess. Fluoride exposure in humans can occur through the consumption of contaminated groundwater. The purpose of this study is to determine the amount of fluoride present in the area under investigation and to determine the health risks associated with its ingestion by adults and children. In accordance with the method recommended by the Environmental Protection Agency, the assessment of the health risk was done using the Chronic Daily Intake (CDI) and Hazard Quotient ( $HQ_{\text{fluoride}}$ ). The fluoride in groundwater varies from 0.37 - 2.70mg/L in the study area. According to the sampling results of all locations, the  $HQ_{\text{fluoride}}$  value in adult males ranges from 0.38 to 2.77, 0.34 to 2.45 in adult females, and 0.41–3.0 in children, which is above the permissible limits for most of the sampling locations, indicating significantly higher health risks. The outcomes of this study could be beneficial for organizations in charge of promoting human health.

**Keywords-**Gaya; fluoride; spatial distribution; health risk assessment

## I. INTRODUCTION

Globally, there is a substantial increase in anthropogenic, natural, and man-made contaminants in the water supply in rural regions. Many sources of effluents contribute to freshwater reservoir contamination, including household, agricultural, and industrial effluents, causing global freshwater scarcity [1]. Groundwater fluoride levels are also increasing due to several geological processes, such as geothermal springs, volcanic activity, tectonic processes, weathering, and other geological processes resulting from the contact of rock and water [2]. The mobilization of fluoride through the leaching of rocks and the over-utilization of groundwater is the root cause of fluoride contamination [3]. Moreover, several anthropogenic activities also contribute to fluoride production, namely coal combustion, the use of fertilizers containing phosphates, and cement manufacturing [4]. Industries like semiconductor production, coal-fired power plants, aluminum smelters, etc. emit wastewater enriched with fluoride [5]. The chemical weathering of underground minerals is more likely to cause fluoride contamination in arid and semi-arid regions. Weathering rocks and releasing volcanic particles in the atmosphere also contribute to groundwater fluoride levels [6], which in turn may contaminate agricultural fields [7].

Fluoride-contaminated groundwater is causing adverse effects on human health [8]. A significant amount of fluoride is accumulated in different parts of the body due to the intake of contaminated water [9]. Many areas suffer from fluoride consumption related crippling skeletal fluorosis. There is a possibility that fluoride can adversely affect the immune system of the human body [11]. Chronic kidney diseases have been shown to be accelerated by the continuous ingestion of fluoride [12]. The population eating fish and tea is more likely to be exposed to fluoride because these foods contain higher concentrations [13]. Since excess fluoride causes deter mental effects on human health, various defluoridation techniques have been developed. Out of various techniques, adsorption is cost effective and efficient. Various materials, such as nanomaterials [14], clay [15], chitosan [16], industrial waste [17], carbonaceous [18], alumina [19], calcium [20], and metal oxides [21] have been used as defluoridation adsorbents.

## II. MATERIALS AND METHODS

### A. Study Area

Gaya district covers a total area of 4976km<sup>2</sup> and lies between latitudes 24°30' and 25°06', and longitudes 84°24' and 85°30'. As a result of the continental monsoon, a hostile



environment prevails throughout the studied region. Hills surround Gaya on three sides and a river on the other side, which causes the region to experience seasonal temperatures. In the summer months (May-July), the temperature ranges from 20.5°C to 44.5°C. The city receives about 214cm of rain

annually between July and October. A temperature range of -4°C to 28°C is expected during winter. Sherghati is located in the Gaya district of Bihar, and the study was conducted in the surrounding area of Sherghati as shown in Figure 1.

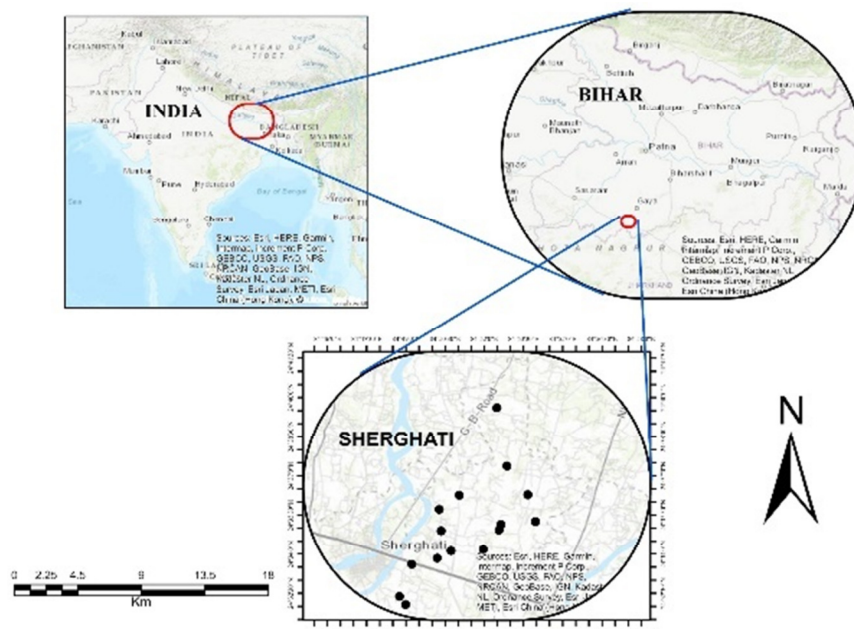


Fig. 1. Map of the study area.

### B. Sample Collection and Analysis

An analysis of groundwater samples was conducted in 15 villages/towns in the Sherghati region of Gaya (Bihar). The fluoride concentration at each sampling site was determined by the average values of 5 samples taken from each site. Fresh groundwater samples were collected from shallow bore wells and private tube wells between March and May, 2022. The samples were collected using pre-cleaned high-density polythene 500mL bottles. A pair of gloves was worn when handling the sampling bottles to avoid contamination. Water was collected from the bore wells after 5 minutes of draining. The sample bottles were tightly capped to prevent contaminants from entering the bottle and were appropriately labelled. Samples were stored in iceboxes until they were transported to the laboratory, where they were preserved at a temperature of 4°C for further chemical testing. Each step of the sample collection, transportation, and analysis was carefully monitored to ensure that the samples were not contaminated or agitated. Fluoride concentration was determined using the SPADNS method described in APHA-AWWAPFC (1994).

### C. Statistical Analysis

The Inverse Distance Weighting (IDW) model was used to interpolate the variation of fluoride concentration in groundwater for the surrounding region of the study area using ArcGIS 10.6 application. The IDW model is a deterministic method to interpolate data, where the weight assigned to perceptual points is inversely proportional to their distances

from each other. A decrease in distance will result in a greater weightage in the calculations and vice versa [22].

### D. Fluoride Health Risks

Hazard Index (HI) values measure the non-carcinogenic risks associated with fluoridated groundwater for children and adults. These values are based on the results of studies conducted on these subjects. As outlined in the USEPA guidelines for assessing health outcomes, any hazard index up to 1 is considered to be within the allowable limit for health risk assessment [23]. Regarding non-carcinogenic risk, HI > 1 indicates a very high fluorosis risk to exposed people, whereas HI < 1 indicates a reasonable fluorosis risk. HI < 1 is considered acceptable for fluoride [24]. As a part of this analysis, the potency of fluoride in an individual is calculated by determining the Chronic Daily Intake (CDI) [25].

$$CDI = \frac{IC \times IR \times EF \times ED}{BW \times AT} \quad (1)$$

where CDI is the chronic daily intake (mg/kg body weight/day), IC is the concentration of fluoride as determined by water samples taken from the ground (mg/L), IR is the daily drinking water intake rate (L/day), its value is presently taken as 4L/day for an adult male, 3L/day for an adult female, and 1L/day for a child [25]. EF refers to the total drinking water exposure frequency (365 days/year). ED refers to the duration of exposure which is taken as 64 years for adult males, 67 years for adult females, and 12 years for children [26]. BW refers to the average body weight (65kg for adult males, 55kg for adult females, and 15kg for children) [27]. AT corresponds to the



average exposure time (i.e. 23,360 days for an adult male, 24,455 for an adult female, and 4380 days for children).

The present study aims to identify non-carcinogenic health risks resulting from drinking high fluoride contaminated groundwater, that can be estimated using the HQ estimates. The following equation was used:

$$HQ_{\text{fluoride}} = \frac{CDI}{RfD} \quad (2)$$

In the case of fluoride, a recommendation is 0.06mg/kg.BW/day as a human's oral reference dose (RfD) [25]. When the  $HQ_{\text{fluoride}}$  value is above 1, it is considered potentially dangerous. In contrast, values of  $HQ_{\text{fluoride}}$  less than 1 are regarded as acceptable considering non-carcinogenic risk to individuals.

### III. RESULTS AND DISCUSSION

#### A. Fluoride Spatial Distribution

Among all sources of fluoride ingestion in the human body, drinking water is the most significant contributor. Based on the analysis of all samples, the results show that most of the study area is at high risk of fluoride contamination due to groundwater contamination. The value of fluoride in the study area lies between 0.37 and 2.70 mg/L, exhibiting elevated fluoride content, above the WHO-permissible limit (1.0mg/L). Among the sampled sites, Site-7 had the maximum fluoride level (2.70mg/L) and Site-9 had the minimum fluoride level (0.37mg/L). It is clear from the results of this study that the local groundwater of Sherghati region in the Gaya district is severely contaminated by fluoride due to the geology of the aquifers in this region, since fluoride is endemic in the Indian subcontinent and varies greatly from region to region. Based on these findings, the study categorized fluoride concentrations and their effects into 5 categories. Accordingly, the samples are listed in decreasing order of concentrations/classes: most samples (46.7%) were collected in Class-2 (0.5–1.5mg/L), followed by Class-3 (1.5–4.0 mg/L) with 40%, and Class-1 (0.5mg/L<sup>-1</sup>) with 13.3%. No samples were found in Class-4 (4.0 -10.0mg/L) and Class-5 (> 10 mg/L). The spatial variation of fluoride is shown in Figure 2.

#### B. Evaluation of Health Effects

CDI for adult males ranged from 0.02 to 0.17mg/kg of body weight/day. The highest CDI value is recorded at Site-7 (0.17) and the minimum at Site-9 (0.02). CDI varied from 0.02 to 0.15 for adult females in this study area, with maximum at Site-7 (0.15) and minimum at Site-9 and Site-14 (0.02). A child's exposure to fluoride usually results in a greater risk than to an adult. Children have a smaller body size, which allows them to accumulate more contaminants per body weight. Children in this region were found to have a CDI value between 0.02 and 0.18. The highest CDI is observed at Site-7 and lowest Site-9.  $HQ_{\text{fluoride}}$  was calculated for each sampling site based on the sampling results. It was found that the  $HQ_{\text{fluoride}}$  ranged between 0.38 and 2.77 for adult males. The highest  $HQ_{\text{fluoride}}$  for an adult male was found at Site-7 (2.77) and the minimum at Site-9 (0.38). The  $HQ_{\text{fluoride}}$  level for adult females varied from 0.34 to 2.45, with the highest found at Site-7 and the minimum at Site-9. For children, the  $HQ_{\text{fluoride}}$  varied from 0.41 to 3.0, with the

highest value observed at Site-7 and the lowest at Site-9. The CDI and HQ values are shown in Table I. A comparison of the HI values for fluoride in 3 different age groups is shown in Figure 3. The result shows that the mean HI values for children, teenagers, and adults in the present study were greater than the safe limit (HI <1). In accordance with previous studies, it is concluded that children whose HI values exceeded the acceptable threshold were at the most significant risk of fluoride toxicity.

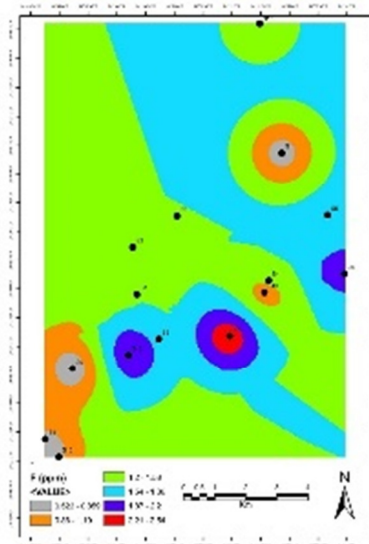


Fig. 2. Spatial variation of fluoride in the study area.

TABLE I. CDI AND HQ VALUES

Site label	F <sup>-</sup> (ppm)	CDI <sub>male</sub> (mg/kg body weight/day)	CDI <sub>female</sub> (mg/kg body weight/day)	CDI <sub>child</sub> (mg/kg body weight/day)	HQ <sub>male</sub>	HQ <sub>female</sub>	HQ <sub>child</sub>
Site-1	1.48	0.09	0.08	0.1	1.52	1.35	1.64
Site-2	1.41	0.09	0.08	0.09	1.45	1.28	1.57
Site-3	1.53	0.09	0.08	0.1	1.57	1.39	1.7
Site-4	1.4	0.09	0.08	0.09	1.44	1.27	1.56
Site-5	1.68	0.1	0.09	0.11	1.72	1.53	1.87
Site-6	2.17	0.13	0.12	0.14	2.23	1.97	2.41
Site-7	2.7	0.17	0.15	0.18	2.77	2.45	3
Site-8	1.75	0.11	0.1	0.12	1.79	1.59	1.94
Site-9	0.37	0.02	0.02	0.02	0.38	0.34	0.41
Site-10	0.72	0.04	0.04	0.05	0.74	0.65	0.8
Site-11	0.64	0.04	0.03	0.04	0.66	0.58	0.71
Site-12	2.34	0.14	0.13	0.16	2.4	2.13	2.6
Site-13	0.83	0.05	0.05	0.06	0.85	0.75	0.92
Site-14	0.42	0.03	0.02	0.03	0.43	0.38	0.47
Site-15	1.24	0.08	0.07	0.08	1.27	1.13	1.38

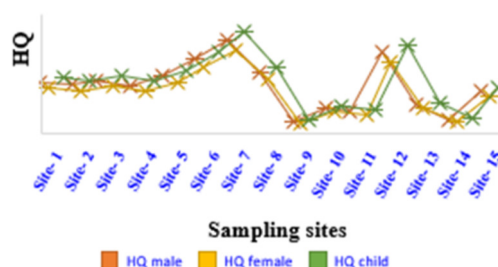


Fig. 3. HQ comparison in the study area.

### C. Health Effects of Fluoride

Fluoride is an essential micronutrient that is present in drinking water as well as in various food sources. Although fluoride benefits the body, it can also have detrimental effects. Fluoride concentration in drinking water ranging from 0.6 to 1.2mg/L benefits bone and tooth development [29]. Fluoride can cause serious health problems in infants, children, and adults if the amount accumulated in the body exceeds the permissible limit. A person will not be affected instantly by fluoride, but it will accumulate in the brain and slowly deteriorate the body's ability to function. Fluoride exposure over a long time is associated with widespread health issues [30]. Fluoride plays a vital role in protecting the newly formed teeth of children, and it plays an equally important role in preventing tooth decay in adults. It has been shown that more than 1.5mg/L of fluoride can adversely affect bone homeostasis (skeletal fluorosis) and enamel growth (dental/enamel fluorosis). Nearly 65% of the total cases of endemic fluorosis are caused by water containing excess fluoride [31]. Initially, dental fluorosis can be described by enamel discoloration, which may turn into discrete or aggregate pitting [32]. Fluoride concentrations greater than 1mg/L may cause neurotoxicity, which can interfere with memory and learning. Developing brains are more susceptible to toxicants than mature brains. The presence of fluoride in groundwater above 10mg/L accelerates the occurrence of neurological problems, hypertension, and cancer. Children exposed to fluoride may also suffer from obesity [33]. Other researchers have reported that there is either no effect or a negative correlation between fluoride and obesity [34]. However, the results of these studies are still inconclusive.

It is suggested that 80% of all diseases worldwide are caused by poor water quality. In arid and semi-arid regions, 65% of endemic fluorosis is caused by drinking water with high fluoride levels [35, 36]. Excessive fluoride consumption contributes to tooth decay and bone deterioration [37]. Fluoride intake has also been shown to be positively correlated with children's intelligence levels. Children consuming less fluoride had lower intelligence test scores than those consuming water with greater fluoride [6]. Fluoride has been identified as a carcinogen in some epidemiological studies. There is, however, some evidence that fluoride can cause osteosarcoma (bone cancer) [38, 39]. In children aged between 6 and 8 years, fluoride-caused cancer is more prevalent than in adults [39]. Neurotoxicity can be caused by fluoride concentrations between 2 and 4mg/L [40]. Many researchers have reported that the fluoride concentration adversely affects children and their thinking ability, more than it does to adults. A similar test was conducted to check the IQ of children in the region affected by fluoride and found lower Intelligence Quotient (IQ) scores in individuals who consumed more fluoride. Moreover, fluoride negatively impacts reaction times and visuospatial skills [41]. Those suffering from kidney diseases have a greater risk of developing fluorosis, even if their drinking water contains a permissible amount of fluoride, because their kidneys do not excrete fluoride as effectively. Additionally, it has been shown to affect the reproductive system as well. Fluoride in groundwater/drinking water appears to have a

tremendous impact on human health, and its occurrence and distribution have received worldwide attention.

### IV. CONCLUSION

There is a chronic fluoride contamination of groundwater in Sherghati, Gaya district. The purpose of this study is to investigate the clinical effects of the chronic exposure to fluoride. Clinical effects consisted of abnormal tooth enamel in children, joint pain and deformities of the limbs and spine in adults. Fluoride slowly accumulates in the body and may damage important organs in a longer stretch of time. A high level of chronic fluoride toxicity has been reported in Sherghati. Fluoride-safe water, education, and awareness of the dangers of fluoride toxicity are all needed in order to tackle the current situation. According to the study results, the area is considered high-risk for fluoride-related problems and requires considerable attention.

### REFERENCES

- [1] C. Singaraja, S. Chidambaram, P. Anandhan, M. V. Prasanna, C. Thivya, and R. Thilagavathi, "A study on the status of fluoride ion in groundwater of coastal hard rock aquifers of south India," *Arabian Journal of Geosciences*, vol. 6, no. 11, pp. 4167–4177, Nov. 2013, <https://doi.org/10.1007/s12517-012-0675-6>.
- [2] M. Vithanage and P. Bhattacharya, "Fluoride in the environment: sources, distribution and defluoridation," *Environmental Chemistry Letters*, vol. 13, no. 2, pp. 131–147, Jun. 2015, <https://doi.org/10.1007/s10311-015-0496-4>.
- [3] S. Ali, S. K. Thakur, A. Sarkar, and S. Shekhar, "Worldwide contamination of water by fluoride," *Environmental Chemistry Letters*, vol. 14, no. 3, pp. 291–315, Sep. 2016, <https://doi.org/10.1007/s10311-016-0563-5>.
- [4] D. Ortiz-Perez *et al.*, "Fluoride-induced disruption of reproductive hormones in men," *Environmental Research*, vol. 93, no. 1, pp. 20–30, Sep. 2003, [https://doi.org/10.1016/S0013-9351\(03\)00059-8](https://doi.org/10.1016/S0013-9351(03)00059-8).
- [5] F. Shen, X. Chen, P. Gao, and G. Chen, "Electrochemical removal of fluoride ions from industrial wastewater," *Chemical Engineering Science*, vol. 58, no. 3, pp. 987–993, Feb. 2003, [https://doi.org/10.1016/S0009-2509\(02\)00639-5](https://doi.org/10.1016/S0009-2509(02)00639-5).
- [6] W. Genxu and C. Guodong, "Fluoride distribution in water and the governing factors of environment in arid north-west China," *Journal of Arid Environments*, vol. 49, no. 3, pp. 601–614, Nov. 2001, <https://doi.org/10.1006/jare.2001.0810>.
- [7] S. Waghmare, T. Arfin, N. Manwar, S. Rayalu, N. Labhsetwar, and D. Lataye, "Adsorption behavior of eggshell modified polyalthia longifolia leaf based alumina as a novel adsorbents for fluoride removal from drinking water," *International Journal of Advance Research and Innovative Ideas in Education*, vol. 1, no. 5, pp. 904–925, 2015.
- [8] S. S. Waghmare and T. Arfin, "Fluoride induced water pollution issue and its health efficacy in India- A review," *International Journal of Engineering Research and General Science*, vol. 3, no. 5, pp. 345–358, 2015.
- [9] N. Kumar, A. A. Mahessar, S. A. Memon, K. Ansari, and A. L. Qureshi, "Impact Assessment of Groundwater Quality using WQI and Geospatial tools: A Case Study of Islamkot, Tharparkar, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5288–5294, Feb. 2020, <https://doi.org/10.48084/etasr.3289>.
- [10] T. Arfin, S. Waghmare, S. Rayalu, D. Lataye, S. Dubey, and S. Tiwari, "Adsorption Behaviour of Modified Zeolite as Novel Adsorbents for Fluoride Removal from Drinking Water: Surface Phenomena, Kinetics and Thermodynamics Studies," *International Journal of Science, Engineering and Technology Research*, vol. 4, no. 12, pp. 4114–4124, Dec. 2015.
- [11] S. S. Waghmare and T. Arfin, "Fluoride Removal from Water by various techniques: Review," *International Journal of Innovative Science, Engineering & Technology*, vol. 2, no. 9, pp. 560–571, 2015.

- [12] S. S. Waghmare, T. Arfin, N. Manwar, D. H. Lataye, N. Labhsetwar, and S. Rayalu, "Preparation and Characterization of Polyalthia longifolia Based Alumina as a Novel Adsorbent for Removing Fluoride from Drinking Water," *Asian Journal of Advances Basic Sciences*, vol. 4, no. 1, pp. 12–24, Aug. 2015.
- [13] A. N. Laghari, Z. A. Siyal, D. K. Bangwar, M. A. Soomro, G. D. Walasai, and F. A. Shaikh, "Groundwater Quality Analysis for Human Consumption: A Case Study of Sukkur City, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2616–2620, Feb. 2018, <https://doi.org/10.48084/etasr.1768>.
- [14] S. Waghmare, D. Lataye, T. Arfin, and S. Rayalu, "Defluoridation by Nano-Materials, Building Materials and Other Miscellaneous Materials: A Systematic Review," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 4, no. 12, pp. 11998–12010, Dec. 2015, <https://doi.org/10.15680/IJRSET.2015.0412046>.
- [15] T. Arfin and S. Waghmare, "Fluoride Removal by Clays, Geomaterials, Minerals, Low Cost Materials and Zeolites by Adsorption: A Review," *International Journal of Science, Engineering and Technology Research*, vol. 4, no. 11, pp. 3663–3676, Nov. 2015.
- [16] S. Waghmare and T. Arfin, "Defluoridation By Adsorption With Chitin - Chitosan-Alginate - Polymers - Cellulose - Resins - Algae And Fungi - A Review," *International Research Journal of Engineering and Technology*, vol. 2, no. 6, pp. 1179–1197, Sep. 2015.
- [17] S. Waghmare and T. Arfin, "Fluoride removal by industrial, agricultural and biomass wastes as adsorbents: review," *International Journal of Advance Research and Innovative Ideas in Education*, vol. 1, no. 4, pp. 628–653, Oct. 2015.
- [18] S. S. Waghmare and T. Arfin, "Fluoride Removal from Water By Carbonaceous Materials: Review," *International Journal of Pavement Engineering*, vol. 2, no. 9, pp. 355–361, Jul. 2015.
- [19] S. S. Waghmare and T. Arfin, "Fluoride Removal from Water by Aluminium Based Adsorption: A Review," *Journal of Biological and Chemical Chronicles*, vol. 2, no. 1, pp. 1–11, 2015.
- [20] T. Arfin and S. S. Waghmare, "Fluoride Removal from Water By Calcium Materials: A State-Of-The-Art Review," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 4, no. 9, pp. 8090–8102, Sep. 2015, <https://doi.org/10.15680/IJRSET.2015.0409013>.
- [21] T. Arfin and S. Waghmare, "Fluoride removal from water by mixed metal oxide adsorbent materials: A state-of-the-art review," *International Journal of Engineering Sciences & Research Technology*, vol. 4, pp. 519–536, Jul. 2015.
- [22] A. Raja and T. Gopikrishnan, "Drought Analysis Using the Standardized Precipitation Evapotranspiration Index (SPEI) at Different Time Scales in an Arid Region," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9034–9037, Aug. 2022, <https://doi.org/10.48084/etasr.5141>.
- [23] "OSWER 9355.4-24, Supplemental guidance for developing soil screening levels for superfund sites. Washington, DC, USA: U.S. Environmental Protection Agency, 2002.
- [24] Health Canada, *Canada Health Act Annual Report, 2004-2005*. Canada: Health Canada, 2005.
- [25] OSWER 9355.4-24, Supplemental guidance for developing soil screening levels for superfund sites. Washington, DC, USA: U.S. Environmental Protection Agency, 2001.
- [26] "World Health Statistics 2013 - World | ReliefWeb." <https://reliefweb.int/report/world/world-health-statistics-2013>.
- [27] *Nutrient requirements and recommended dietary allowances for Indians: A Report of the Expert Group of the Indian Council of Medical Research*. Indian Council of Medical Research, 2010.
- [28] *Guidance for Performing Aggregate Exposure and Risk Assessments*. Washington DC, USA: Office of Pesticide Programs, 1999.
- [29] B. Karthikeyan and E. Lakshmanan, "Fluoride in Groundwater: Causes, Implications and Mitigation Measures," in *Fluoride Properties, Applications and Environmental Management*, S. D. Monroy, Ed. 2011, pp. 111–136.
- [30] *Risk Assessment Guidance for Superfund (RAGS): Part A*. Washington DC, USA: EPA, 1989.
- [31] S. Ahmad, R. Singh, T. Arfin, and K. Neeti, "Fluoride contamination, consequences and removal techniques in water: a review," *Environmental Science: Advances*, 2022, <https://doi.org/10.1039/D1VA00039J>.
- [32] P. Mondal, D. Mehta, and S. George, "Defluoridation studies with synthesized magnesium-incorporated hydroxyapatite and parameter optimization using response surface methodology," *Desalination and Water Treatment*, vol. 57, no. 56, pp. 27294–27313, Dec. 2016, <https://doi.org/10.1080/19443994.2016.1167628>.
- [33] S. Ali *et al.*, "Concentration of fluoride in groundwater of India: A systematic review, meta-analysis and risk assessment," *Groundwater for Sustainable Development*, vol. 9, Oct. 2019, Art. no. 100224, <https://doi.org/10.1016/j.gsd.2019.100224>.
- [34] N. K. Mondal, R. Bhaumik, T. Baur, B. Das, J. K. Datta, and P. Roy, "Studies on Defluoridation of Water by Tea Ash: An Unconventional Biosorbent," *Chemical Science Transactions*, vol. 1, no. 2, pp. 239–256, 2012, <https://doi.org/10.7598/cst2012.134>.
- [35] A. Narsimha and V. Sudarshan, "Assessment of fluoride contamination in groundwater from Basara, Adilabad District, Telangana State, India," *Applied Water Science*, vol. 7, no. 6, pp. 2717–2725, Oct. 2017, <https://doi.org/10.1007/s13201-016-0489-x>.
- [36] A. J. Felsenfeld and M. A. Roberts, "A Report of Fluorosis in the United States Secondary to Drinking Well Water," *The Journal of the American Medical Association*, vol. 265, no. 4, pp. 486–488, Jan. 1991, <https://doi.org/10.1001/jama.1991.03460040062030>.
- [37] S. K. Jha, V. K. Mishra, D. K. Sharma, and T. Damodaran, "Fluoride in the Environment and Its Metabolism in Humans," in *Reviews of Environmental Contamination and Toxicology*, D. M. Whitacre, Ed. New York, NY, USA: Springer, 2011, pp. 121–142.
- [38] C.-Y. Yang, M.-F. Cheng, S.-S. Tsai, and C.-F. Hung, "Fluoride in Drinking Water and Cancer Mortality in Taiwan," *Environmental Research*, vol. 82, no. 3, pp. 189–193, Mar. 2000, <https://doi.org/10.1006/enrs.1999.4018>.
- [39] E. B. Bassin, "Association Between Fluoride in Drinking Water During Growth and Development and the Incidence of Osteosarcoma for Children and Adolescents," Ph.D. dissertation, Harvard School of Dental Medicine, Boston, MA, USA, 2001.
- [40] S. Dobaradaran, A. H. Mahvi, S. Dehdashti, S. Dobaradaran, and R. Shoara, "Correlation of fluoride with some inorganic constituents in groundwater of Dashtestan, Iran.," *Fluoride*, vol. 42, no. 1, pp. 50–53, 2009.
- [41] W. Guissouma, O. Hakami, A. J. Al-Rajab, and J. Tarhouni, "Risk assessment of fluoride exposure in drinking water of Tunisia," *Chemosphere*, vol. 177, pp. 102–108, Jun. 2017, <https://doi.org/10.1016/j.chemosphere.2017.03.011>.

# GIS-based Flood Risk Mapping: The Case Study of Kosi River Basin, Bihar, India

**Niraj Kumar**

Department of Civil Engineering, National Institute of Technology Patna, India  
nirajdsi10@gmail.com  
(corresponding author)

**Ramakar Jha**

Department of Civil Engineering, National Institute of Technology Patna, India  
rjha43@gmail.com

*Received: 27 September 2022 | Revised: 1 November 2022 and 6 November 2022 | Accepted: 7 November 2022*

## ABSTRACT

Flood risk mapping aims to create an easily read and rapidly accessible map to prioritize the mitigation effects. This study presents an empirical approach to flood risk mapping through the integration of Analytical Hierarchy Process (AHP) and Geographic Information System (GIS) techniques. SRTM 30m DEM is processed using ArcGIS 10.3 software. The study methodology includes the selection of the study area, the identification of the factors responsible for flood and collection of the required data, the generation of the desired thematic layers, and their integration to produce the flood risk map. Geomorphic, hydrologic, and socio-economic analyses are carried out to generate the thematic layers, namely slope, district's distance to active stream, highest elevation, drainage density, rainfall, population density, and land use-land cover. AHP is used to determine the relative impact weight of the thematic layers. The influence of each thematic layer and the scale values provided based on the weights and score calculated by the AHP are used to integrate the layers in GIS environment to prepare the flood risk map. Consistency ratios are determined from the judgment process to validate the reliability of the proposed approach and results. The study classified the area falling in the basin under different risk zones with Purnia and Madhepura having large areas under high risk. This study may aid decision and policymakers in the evaluation and rapid assessment of flooding phenomena in the region.

**Keywords-***flood; risk; geomorphic; hydrologic; socio-economic; AHP; SRTM; GIS*

## I. INTRODUCTION

Floods are one of the most devastating and recurring natural disasters, having huge impact on human lives and property [1, 2]. Floods are caused mainly due to heavy rainfalls, whereas landslides, dam breaches, and earthquakes are some other causes. In addition to these, human activities such as overpopulation, settlement increase in the way of river courses, drainage alterations, deforestation, and climate change are also responsible [3, 4]. Water is the most important factor among the basic needs of life [5]. Rainwater can be captured as a means of preventing flood [6]. Floods are difficult to contain and control, thus effective mitigation/management of flood risk is essential [7, 8]. Flood risk management needs to recognize the interconnections between infrastructures, economic systems and the role of human factors in assessing and managing the risk [8]. Many researchers have used various methodologies to generate flood risk maps [7, 9-19]. Earlier conventional methods were utilized for comparative risk analysis of hazards [9]. To model and predict the magnitude of flood risk areas, integrated Analytical Hierarchy Process (AHP) and Geographic Information System (GIS) analysis techniques are used [10-

12]. GIS has been used for the evaluation of satellite-based rainfall estimates [20] and to assess long-term water quality changes [21]. AHP is a theory of relative measurement of intangible criteria [22]. It was proven to be an effective tool for solid waste management in addition to risk analysis [23, 24]. Toposheet maps, satellite photos, Landsat-8, and Digital Elevation Model (DEM) are utilized to create thematic layers in a GIS environment to delineate the flood hazard potential zone in [7]. The integration of GIS and Hierarchical Multi-criteria Analysis (HMA) offer a low-cost methodology to produce vulnerability maps [13]. Landsat images were utilized to identify the land-use for vulnerability analysis in [25]. Flood hazard zone maps are a valuable tool to identify the nature of floods and their effects on human life [14]. A flood hazard zone map can be an effective tool in the planning of the future development of a city [15]. Flood hazard and vulnerability maps are integrated to generate the flood risk map.

The present study aims to map the flood risk of Kosi river basin in Bihar using remote sensing and GIS which can help evaluating the efficiency of drainage network infrastructure and the development efforts needed to reduce flood risk [14].

## II. STUDY AREA

In the present study, flood risk mapping of Kosi river basin has been conducted. Kosi river is responsible for many flood incidents in the region during the last 30 years. The river has shifted nearly 150Km westward with extensive flooding in the basin region during the last 200 years [26]. The main reason of this westward movement and frequent flooding in the region is the sedimentation of the basin [27]. The cause of shifting of the river and the sedimentation in the region is the geomorphic and hydrological behavior of the basin [28]. The Kosi river is a major tributary of the Ganga river. The Kosi river originates in the Himalayan region in Nepal. After draining a large area in Tibet and Nepal it enters into Bihar (India) region near Bhimnagar and joins the Ganga river near Kursheela, Katihar district, Bihar. The drainage basin is located between latitudes  $25.35^{\circ}$  –  $26.36^{\circ}$  N and longitudes  $86.53^{\circ}$  –  $87.59^{\circ}$  E. The Kosi river basin is surrounded by the Great Himalayas in the north, by Mahananda river basin in the east, by the Burhi Gandak basin in the west, and by the Ganga River in the south. The Kosi river basin (main stem) extends over 10 districts: Araria, Bhagalpur, Khagaria, Katihar, Madhubani, Madhepura, Munger, Purnea, Saharsa, and Supaul of Bihar in the Indian region.

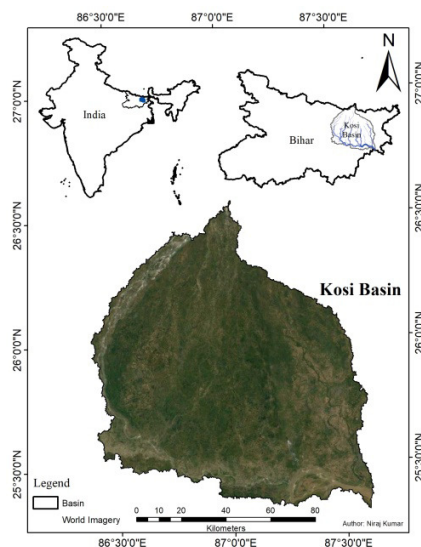


Fig. 1. Kosi river basin (generated by the ArcGIS 10.3).

## III. DATA SOURCE

The current study comprises a collection of data to generate the desired thematic layers. Shuttle Radar Topography Mission 1 Arc-Second Global elevation data and Landsat 8 images were extracted from the USGS Earthexplorer platform for geomorphic analysis and socio-economic analysis. Hydrologic analysis was carried out using monthly rainfall data collected from the Indian Meteorological Department, New Delhi. Population data were collected from the Census of India, Ministry of Home affairs, Govt. of India (<http://www.censusindia.gov.in/>). ArcGIS online services and iBHUGOAL-BiHar infrastrUcture mapping- Geomatics Oriented Application Model Geographical Information System

(GIS) (<http://gis.bih.nic.in/>) were used to create the Indian administrative boundaries and the district maps of Bihar.

## IV. METHODOLOGY

The flood risk mapping consisted of several steps which include the selection of the study area, the identification of the factors related to the flood events, the collection of the required data, the generation of the desired thematic layers based on the factors responsible for hazard and vulnerability with the help of the available data, the integration of all the thematic layers to produce flood hazard and vulnerability maps, the generation of the flood risk map, and the validation of the result [7, 9-19]. In this study ArcGIS 10.3 software (ArcMap 10.3) is used. The ArcSWAT tool is used as an add-in in ArcGIS10.3 software to delineate the watershed and to generate stream networks. Spatial Analyst tools and Data Management tools are used in ArcMap 10.3 for generation of the final flood risk map. The various thematic layers based on the data availability are generated. The generation of thematic layers has been divided into three parts, namely geomorphic, hydrologic, and socio-economic.

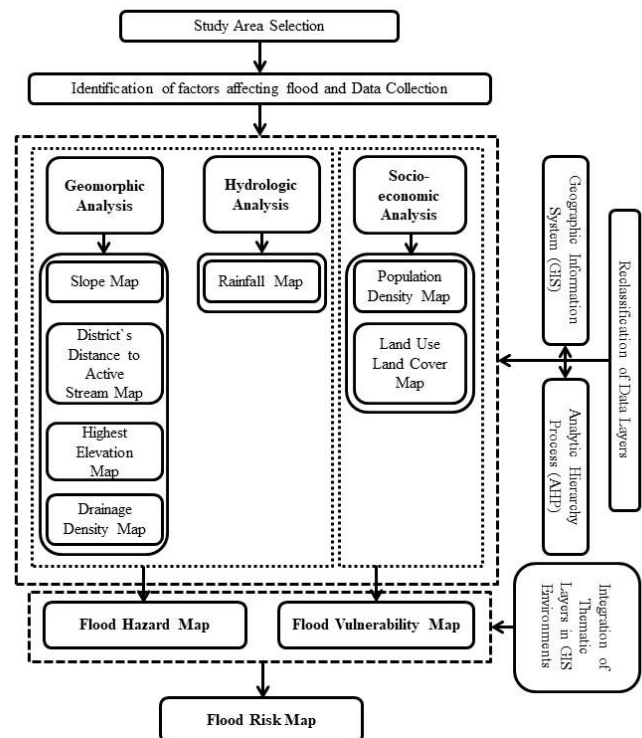


Fig. 2. Methodology.

The geomorphic analysis consists of the generation of the slope map, the district's distance to the active stream map, the highest elevation map, and the drainage density map. The SRTM 1 Arc-Second Global elevation DEM has been used to extract these maps [29]. The slope map of the study area has been extracted using the Surface tool under Spatial Analyst tool extension in ArcMap 10.3. The drainage density map was generated from the stream network of the basin using the Density tool under the Spatial Analyst tool extension. The



district's highest elevation map and the distance to active streams map were generated using the Zonal tool under the Spatial Analyst tool extension. The final highest elevation and distance to active stream maps were generated using the Interpolation tool under the Spatial Analyst tool extension to create continuous maps within the basin boundary. The Interpolation tool interpolates a raster surface from points using the Inverse Distance Weighted (IDW) technique.

Hydrologic analysis consists of the generation of the rainfall map of the study area. The 10-year mean annual rainfall data collected for the 10 districts of the basin region were imported in ArcGIS. These values are joined in the point feature of each district. Then, the final rainfall map is generated using the Interpolation tool to create a continuous raster rainfall map within in the basin boundary. The socio-economic analysis consists of the generation of the population density map and the Land Use-Land Cover (LULC) map of the study area. The final population density map using the population data for the 10 districts is generated in the similar way as the rainfall map is generated using the IDW tool. The LULC map is generated by processing the Landsat 8 images using the Classification tool extension in ArcMap 10.3. The final generated maps are classified into 5 classes using natural breaks, which are further reclassified using the Reclassify tool under the Spatial Analyst tool extension. The reclassified maps are given values from 1 for very low risk to 5 for very high risk. The influence of each thematic layer and the scale values provided based on the weights and score calculated by AHP is used to integrate the layers in GIS environment to prepare the flood hazard and vulnerability maps using the Weighted Overlay tool under the Spatial Analyst tool extension. The fundamental 9-point intensity scale [22] helped in multiple pairwise criteria comparisons in AHP. AHP has been used by various researchers for risk mapping as a vital tool [7, 9-19].

TABLE I. THE FUNDAMENTAL 9-POINT INTENSITY IMPORTANCE SCALE [22]

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgment slightly favor one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgment strongly favor one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favored very strongly over another. Its dominance is demonstrated in practice
8	Very very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation

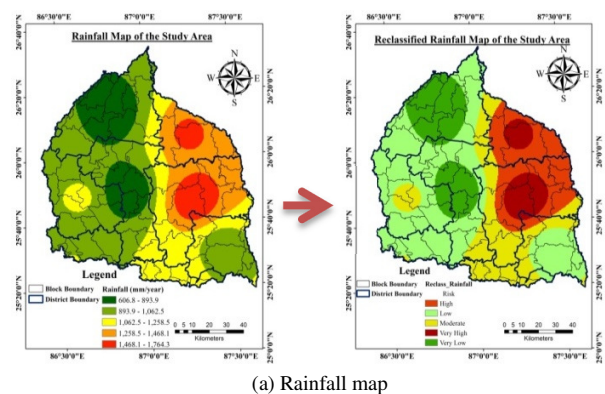
The flood hazard and vulnerability maps are multiplied using the Raster Calculator tool (Map Algebra tool) under the Spatial Analyst tool extension to generate the final flood risk map.

## V. RESULTS AND DISCUSSION

### A. Thematic Layer Generation

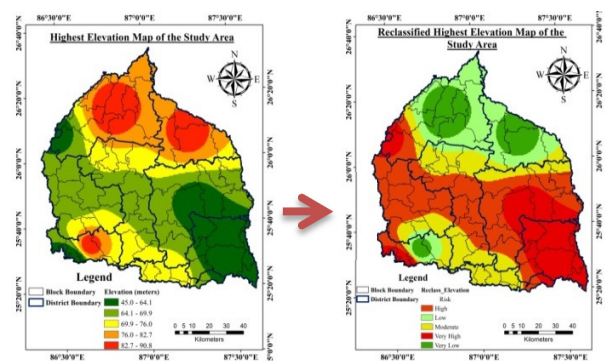
In the present study, 7 thematic layers have been created using the collected data sets. The rainfall map is generated for hydrological analysis. Similarly, slope, district's distance to active stream, highest elevation, and drainage density maps are generated for geomorphic analysis and population density and LULC maps for socio-economic analysis. These thematic layers are classified into 5 classes using natural breaks, which are further reclassified.

#### 1. Hydrological factors

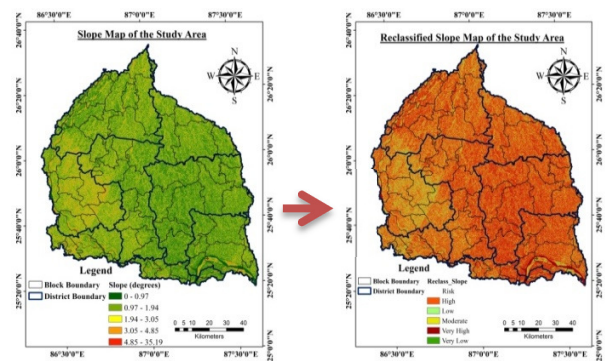


(a) Rainfall map

#### 2. Geomorphic factors

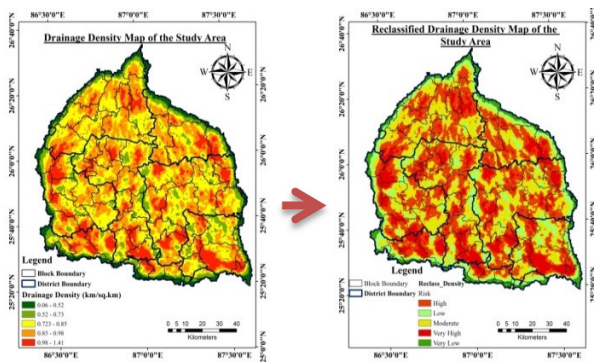


(b) Highest elevation map

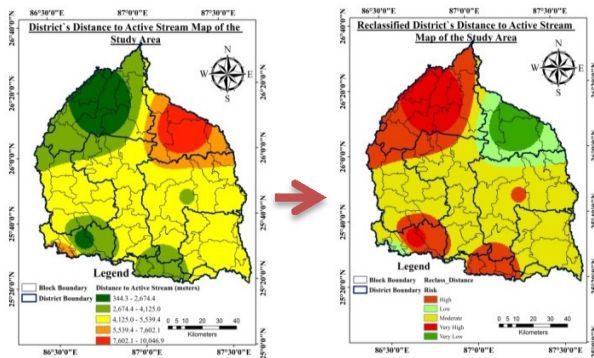


(c) Slope map



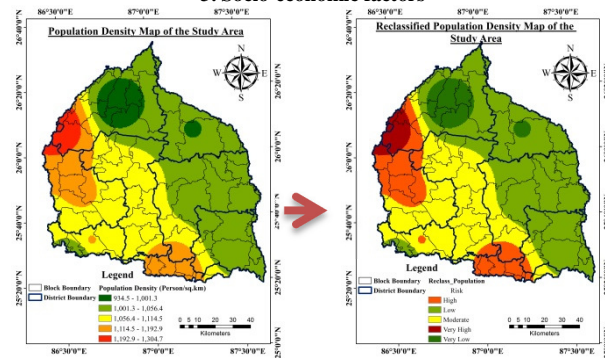


(d) Drainage density map

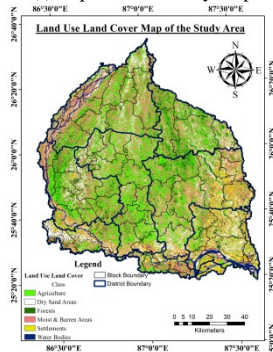


(e) Districts' distance to active stream map

### 3. Socio-economic factors



(f) Population density map



(g) Land use land cover map

Fig. 3. Thematic layers of the study area.

### B. AHP for Flood Risk Assessment

The flood risk assessment is the combination of flood hazard and vulnerability analyses. The analysis of a region using AHP starts with building a hierarchy. The understanding of various factors is used for developing the hierarchy to generate the final flood hazard and vulnerability maps separately. The hierarchy generation comprises of 2 steps, namely the calculation of the reciprocal matrix and the calculation of the normalized matrix. The generated matrices are followed by the generation of matrices of various sub-factors. The calculation of the reciprocal matrix followed by the calculation of the normalized matrix is carried out for each criterion to assign the weight and influence percentage. Further, the score is assigned to different cell attributes of each criterion. After assigning the weight and influence percentage to the criteria and the score to the sub-categories of the criteria the flood hazard and vulnerability maps are generated using the Weighted Overlay tool.

TABLE II. WEIGHTED FLOOD HAZARD RANKING

Criterion	Priority/weight	Categories (cell attributes)	Re-classed attributes	Score	Scale value
Rainfall (mm/year)	0.435	606.8-893.9	1	Very low	1
		893.9-1062.5	2	Low	1
		1062.5-1258.5	3	Moderate	2
		1258.5-1468.1	4	High	5
		1468.1-1764.3	5	Very high	9
Highest elevation (m)	0.265	45.0-64.1	5	Very high	9
		64.1-69.9	4	High	5
		69.9-76.0	3	Moderate	2
		76.0-82.7	2	Low	1
		82.7-90.8	1	Very low	1
Slope (deg)	0.154	0.0-0.97	5	Very high	9
		0.97-1.94	4	High	5
		1.94-3.05	3	Moderate	2
		3.05-4.85	2	Low	1
		4.85-35.19	1	Very low	1
Drainage density (km/km²)	0.09	0.064-0.524	1	Very low	1
		0.524-0.725	2	Low	1
		0.725-0.852	3	Moderate	2
		0.852-0.979	4	High	5
		0.979-1.413	5	Very high	9
District's distance to active stream (m)	0.055	344.3-2674.4	5	Very high	9
		2674.4-4126.0	4	High	5
		4126.0-5539.4	3	Moderate	2
		5539.4-7602.1	2	Low	1
		7602.1-10046.9	1	Very low	1

### 1) Flood Hazard Mapping

The 5 thematic layers generated in geomorphic and hydrologic analysis are used to generate the flood hazard map. High rainfall is responsible for frequent flooding in the region and the silting of the river streams, thus it was assigned the higher weight. The next higher weight was assigned to the elevation criterion, followed by criteria like slope, drainage density, and distance to active streams.

### 2) Flood Vulnerability Mapping

The 2 thematic layers generated in socio-economic analysis are used to generate the flood vulnerability map. The population is much more responsible for the flooding vulnerability of the region than LULC.

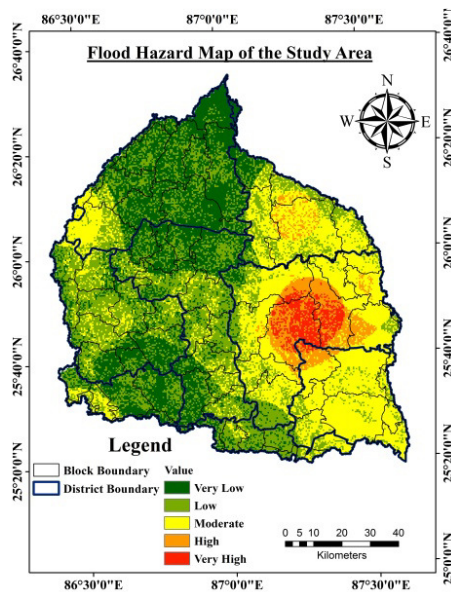


Fig. 4. Flood hazard map.

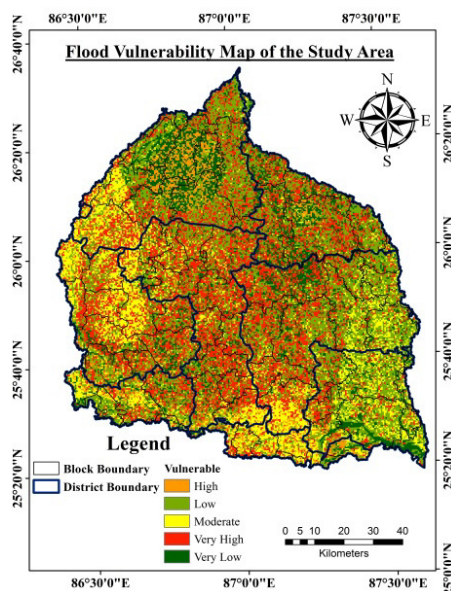


Fig. 5. Flood vulnerability map.

TABLE III. WEIGHTED FLOOD VULNERABILITY RANKING

Criterion	Priority/weight	Categories (cell attributes)	Re-classed attributes	Score	Scale value
Population density (persons/km <sup>2</sup> )	0.75	934.5-1001.3	1	0.623	1
		1001.3-1056.5	2	1.213	1
		1056.5-1114.5	3	2.405	2
		1114.5-1192.9	4	4.658	5
		1192.9-1304.7	5	9.000	9
LULC class	0.25	Agriculture		9.000	9
		Settlement		6.193	6
		Dry sand		4.437	4
		Open land		2.750	3
		Water body		1.402	1
		Forest		0.750	1

### 3) Flood Risk Mapping

In the present study, the final risk map is generated by simply multiplying the classified flood hazard and vulnerability maps using the Raster Calculator tool. The generated map is then classified into 5 main classes, namely Very Low, Low, Moderate, High, and Very High using the Reclassify tool. The tabulate area tool (Zonal tool) is used to calculate the district area and block area under different risk classes.

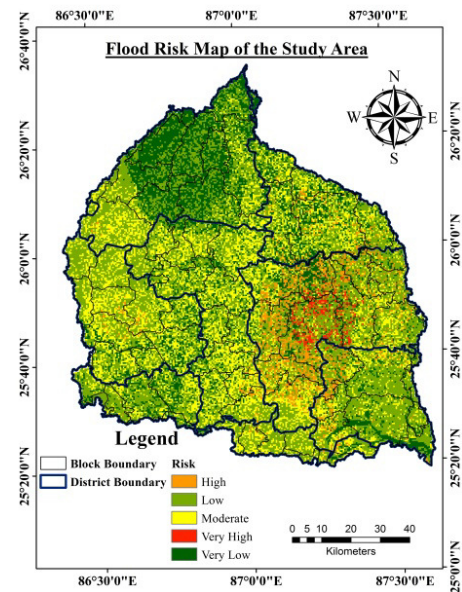


Fig. 6. Flood risk map.

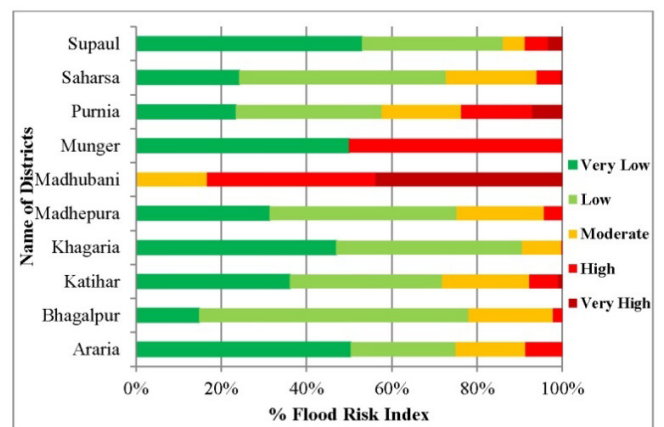


Fig. 7. Bar chart showing the district area percentage under different risk classes.

The present study shows that all 10 districts of the region have more or less some area under High Risk with Madhubani having the largest percentage of area under Very High Risk and Purnia and Madhepura districts having large areas under High Risk. The regions of Supaul have large areas under Very Low Risk and thus are considered as safer. Araria, Munger and Supaul have large areas under Very Low Risk. Almost all the districts have approximately equal percentage of area under Moderate Risk. Raniganj of Araria district, Krityanand Nagar,

Damdaha, Barhara, Bhawanipur and Banmankhi of Purnia district, and Supaul of Supaul district have the largest areas under High Risk.

## VI. CONCLUSION

The current study presents an empirical approach for flood risk mapping in the Kosi river basin region through the integration of AHP and GIS techniques. This approach can aid decision and policymakers in the evaluation and assessment of flooding phenomena in the region. It will help evaluating the efficiency of drainage network infrastructure and development efforts needed to reduce flood risk. Consistency ratios are determined from the judgment process to validate the reliability of the proposed approach and results. This approach can be applied to many flood prone regions, where data availability is poor and mapping resources are limited.

The present study classified 10 districts and 76 blocks of the basin region into risk areas. The block level analysis of risk zones will help the decision makers in taking steps at root level. The final flood risk map obtained after combining the different layers in GIS environment has been validated with previous research works. The generation of the maps is not based on some previous inundation data, but on the integrated approach of different factors responsible. Thus, it is not only a hydrological phenomenon, but an integrated response of the river basin. Some of the regions in the map have high hazard risk, even the regions that are not inundated due to the importance of other factors. The use of high resolution images and DEMs increases the mapping efficiency. The results obtained in this study justify the integration of AHP and GIS techniques which gives a powerful tool for decision making procedures in flood risk mapping because these techniques allow a coherent and efficient use of the spatial data.

## REFERENCES

- [1] S. N. Jonkman, "Global Perspectives on Loss of Human Life Caused by Floods," *Natural Hazards*, vol. 34, no. 2, pp. 151–175, Feb. 2005, <https://doi.org/10.1007/s11069-004-8891-3>.
- [2] S. I. Khan *et al.*, "Satellite Remote Sensing and Hydrologic Modeling for Flood Inundation Mapping in Lake Victoria Basin: Implications for Hydrologic Prediction in Ungauged Basins," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 85–95, Jan. 2011, <https://doi.org/10.1109/TGRS.2010.2057513>.
- [3] C. Clark, "Deforestation and Floods," *Environmental Conservation*, vol. 14, no. 1, pp. 67–69, 1987.
- [4] Z. W. Kundzewicz *et al.*, "Flood risk and climate change: global and regional perspectives," *Hydrological Sciences Journal*, vol. 59, no. 1, pp. 1–28, Jan. 2014, <https://doi.org/10.1080/02626667.2013.857411>.
- [5] L. Kumar *et al.*, "Analysis of Groundwater Quality Parameter of Hand-Dug Well and Their Health Risk Assessment of Study Area Tharpakar," *International Journal of Modern Research in Engineering & Management*, vol. 2, no. 5, pp. 41–48, May 2019.
- [6] V. Kumar *et al.*, "GIS-Based Analysis of a Rainwater Harvesting System in the Multipurpose Hall of Quaid-e-Awam University of Engineering, Science, and Technology," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8837–8842, Aug. 2022, <https://doi.org/10.48084/etasr.4995>.
- [7] Malsawmtluanga and C. Vabeihmo, "Assessment of Flood Hazard Zonation Using Geographic Information System and Analytical Hierarchy Process: A Case Study of Tlawng River Watershed in Sairang, Mizoram, India," *Nature Environment and Pollution Technology*, vol. 21, no. 1, pp. 101–109, Mar. 2022, <https://doi.org/10.46488/NEPT.2022.v21i01.011>.
- [8] S. Jonkman and R. Dawson, "Issues and Challenges in Flood Risk Management—Editorial for the Special Issue on Flood Risk Management," *Water*, vol. 4, no. 4, pp. 785–792, Oct. 2012, <https://doi.org/10.3390/w4040785>.
- [9] R. W. Kates and J. X. Kaspersen, "Comparative risk analysis of technological hazards (a review)," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 22, pp. 7027–7038, Nov. 1983.
- [10] Y. O. Ouma and R. Tateishi, "Urban Flood Vulnerability and Risk Mapping Using Integrated Multi-Parametric AHP and GIS: Methodological Overview and Case Study Assessment," *Water*, vol. 6, no. 6, pp. 1515–1545, Jun. 2014, <https://doi.org/10.3390/w6061515>.
- [11] A. Ghosh and S. Kar, "Application of analytical hierarchy process (AHP) for flood risk assessment: a case study in Malda district of West Bengal, India," *Natural Hazards*, vol. 94, pp. 349–368, Oct. 2018, <https://doi.org/10.1007/s11069-018-3392-y>.
- [12] N. Kazakis, I. Kougias, and T. Patsialis, "Assessment of flood hazard areas at a regional scale using an index-based approach and Analytical Hierarchy Process: Application in Rhodope–Evros region, Greece," *Science of The Total Environment*, vol. 538, pp. 555–563, Dec. 2015, <https://doi.org/10.1016/j.scitotenv.2015.08.055>.
- [13] K. Loumi and A. Redjem, "Integration of GIS and Hierarchical Multi-Criteria Analysis for Mapping Flood Vulnerability: The Case Study of M'sila, Algeria," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7381–7385, Aug. 2021, <https://doi.org/10.48084/etasr.4266>.
- [14] A. A. Ghezelsoufloo and M. Hajibigloo, "Application of Flood Hazard Potential Zoning by using AHP Algorithm," *Civil Engineering Research Journal*, vol. 9, no. 5, pp. 150–159, Feb. 2020, <https://doi.org/10.19080/CERJ.2019.09.555775>.
- [15] B. Büchele *et al.*, "Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks," *Natural Hazards and Earth System Sciences*, vol. 6, no. 4, pp. 485–503, Jun. 2006, <https://doi.org/10.5194/nhess-6-485-2006>.
- [16] B. Das, S. C. Pal, and S. Malik, "Assessment of flood hazard in a riverine tract between Damodar and Dwarakeswar River, Hugli District, West Bengal, India," *Spatial Information Research*, vol. 26, no. 1, pp. 91–101, 2018, <https://doi.org/10.1007/s41324-017-0157-8>.
- [17] P. Dash and J. Sar, "Identification and validation of potential flood hazard area using GIS-based multi-criteria analysis and satellite data-derived water index," *Journal of Flood Risk Management*, vol. 13, no. 3, 2020, Art. no. e12620, <https://doi.org/10.1111/jfr3.12620>.
- [18] R. S. Ajin, R. R. Krishnamurthy, M. Jayaprakash, and P. G. Vinod, "Flood hazard assessment of Vamanapuram River Basin, Kerala, India: An approach using Remote Sensing & GIS techniques," *Advances in Applied Science Research*, vol. 4, no. 3, pp. 263–274, May 2013.
- [19] K. S. Vignesh, I. Anandakumar, R. Ranjan, and D. Borah, "Flood vulnerability assessment using an integrated approach of multi-criteria decision-making model and geospatial techniques," *Modeling Earth Systems and Environment*, vol. 7, no. 2, pp. 767–781, Jun. 2021, <https://doi.org/10.1007/s40808-020-00997-2>.
- [20] T. G. Gebremicael *et al.*, "Satellite-based rainfall estimates evaluation using a parsimonious hydrological model in the complex climate and topography of the Nile River Catchments," *Atmospheric Research*, vol. 266, Mar. 2022, Art. no. 105939, <https://doi.org/10.1016/j.atmosres.2021.105939>.
- [21] L. Kumar *et al.*, "Assessment of physicochemical parameters in groundwater quality of desert area (Tharpakar) of Pakistan," *Case Studies in Chemical and Environmental Engineering*, vol. 6, Dec. 2022, Art. no. 100232, <https://doi.org/10.1016/j.csee.2022.100232>.
- [22] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York, NY, USA: McGraw-Hill, 1980.
- [23] S. Joel, M. L. Ernest, and A. Ajapnwa, "Application of Analytic Hierarchy Process Decision Model for Solid Waste Management Strategy in Yaoundé, Cameroon," *The Journal of Solid Waste Technology and Management*, vol. 45, no. 4, pp. 502–517, Nov. 2019, <https://doi.org/10.5276/JSWTM/2019.502>.

- [24] T. Hazra, B. Maitra, and S. Goel, "Development and Application of a Multi-Criteria Decision Making (MCDM) Tool for Solid Waste Management: Kolkata as a Case Study," in *Advances in Solid and Hazardous Waste Management*, S. Goel, Ed. Cham: Springer International Publishing, 2017, pp. 275–299, [https://doi.org/10.1007/978-3-319-57076-1\\_14](https://doi.org/10.1007/978-3-319-57076-1_14).
- [25] T. L. Dammalage and N. T. Jayasinghe, "Land-Use Change and Its Impact on Urban Flooding: A Case Study on Colombo District Flood on May 2016," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3887–3891, Apr. 2019, <https://doi.org/10.48084/etasr.2578>.
- [26] C. V. Gole and S. V. Chitale, "Inland Delta Building Activity of Kosi River," *Journal of the Hydraulics Division*, vol. 92, no. 2, pp. 111–126, Mar. 1966, <https://doi.org/10.1061/JYCEAJ.0001406>.
- [27] N. A. Wells and J. A. Dorr Jr., "Shifting of the Kosi River, northern India," *Geology*, vol. 15, no. 3, pp. 204–207, Mar. 1987, [https://doi.org/10.1130/0091-7613\(1987\)15<204:SOTKRN>2.0.CO;2](https://doi.org/10.1130/0091-7613(1987)15<204:SOTKRN>2.0.CO;2).
- [28] T. Chakraborty, R. Kar, P. Ghosh, and S. Basu, "Kosi megafan: Historical records, geomorphology and the recent avulsion of the Kosi River," *Quaternary International*, vol. 227, no. 2, pp. 143–160, Nov. 2010, <https://doi.org/10.1016/j.quaint.2009.12.002>.
- [29] N. Kumar and R. Jha, "Morphometric Analysis of Kosi River Basin, Bihar, India Using Remote Sensing and GIS Techniques," in *Climate Change Impacts on Water Resources: Hydraulics, Water Resources and Coastal Engineering*, R. Jha, V. P. Singh, V. Singh, L. B. Roy, and R. Thendiyath, Eds. Cham: Springer International Publishing, 2021, pp. 469–481, [https://doi.org/10.1007/978-3-030-64202-0\\_40](https://doi.org/10.1007/978-3-030-64202-0_40).

# Design of a Shielded Room against EMP Signal as per MIL-STD 461

**Venkata Sai Charishma Pathala**

Department of EECE, Gandhi Institute of Technology and Management, Deemed to be University, India  
cpathala@gitam.in  
(corresponding author)

**Pappu V. Y. Jayasree**

Department of EECE, Gandhi Institute of Technology and Management, Deemed to be University, India  
jpappu@gitam.in

Received: 30 September 2022 | Revised: 3 November 2022 | Accepted: 5 November 2022

## ABSTRACT

Electromagnetic shielding is the best technique to protect equipment from the Electromagnetic Pulse (EMP) signal. This paper explains how effectively the equipment will be protected within a shielded room against EMP signals. The shielded room is designed with different points of entry used to provide electrical connections to the Equipment Under Test (EUT) in a honeycomb structure for ventilation to protect the equipment from the EMP signal. The shielded room with four points of entry and honeycomb structures is designed, analyzed theoretically, and simulated in the CST Studio. The points of entry (PoE) and the honeycomb structure are designed based on MIL-STD-461 E/F/G (by following this standard the maximum frequency of EMP signal is 100MHz). It is observed that by increasing the size of the PoE the shielding effectiveness value decreases by 20dB for perfect electrical conductor (PEC) material of 2mm thickness. It is concluded that the equipment will be more protected when it is placed nearer to the front wall or in the middle of the shielded room. The performance of the shielded room will not be affected with honeycomb structures which will provide 220dB Shielding Effectiveness (SE).

**Keywords-**Electromagnetic Pulse (EMP); Point of Entry (PoE); honeycomb structure; MIL-STD-461 E/F/G; Shielding Effectiveness (SE); Perfect Electrical Conductor (PEC)

## I. INTRODUCTION

In the present electromagnetic environment, the protection of equipment related to medical or military applications is of utmost priority. Equipment shielding is achieved by designing a shielded room. Usually, the shielded room will be designed by placing gaskets during the design of the door of the room, a honeycomb structure for ventilation, and apertures for the routing of cables. The main motto of this shielded room is to protect the electronic equipment from the interference of the dangerous signal strength of EMPs. The most effective method to safeguard the equipment from Electromagnetic Interference (EMI) is shielding and filtering [1-4]. The performance of a shield is defined by its SE value which is obtained from taking the difference between the received power with shield material and without it [5]. The shielding technique is used in many fields like aerospace engineering or medical and military equipment. In aerospace applications, different combinations of shielding materials are used to provide shielding to the aircraft [6]. Even the human brain can be protected from mobile radiation by placing shields coated with different types of shielding materials [7].

The simulated data for electromagnetic calculations can be acquired by various numerical methods [8], but the computations usually require long time. So, the analytic method has a better performance than the numerical methods. The analytical methods are classified in two types, i.e. based on aperture coupling between the aperture [9, 10] and based on the transmission line theory. The design of metallic shielded rooms is used to reduce the leakage signals from external electromagnetic fields. According to the IEEE Standard 299, the SE is used for the calculation of the effectiveness of metallic shielded enclosures [11]. The materials used in the design of shielded rooms are considered based on their conductivity and the frequency of operation. In this paper, the material used in the design was a perfect electrical conductor, in theoretical calculations the conductivity value of the materials is not considered while calculating the SE of the shielded room. The calculations of SE can be done with many methods. One of these is to add three parameters (Reflection loss ( $R$ ), Absorption loss ( $A$ ), Multiple reflections ( $M$ )) to obtain the SE value.

$$SE_{dB} = R_{dB} + A_{dB} + M_{dB} \quad (1)$$



Another method for the calculation of electrical SE (ESE) is achieved by considering the difference between the field signal in the presence and the absence of the shield as in (2).  $E_A$  and  $E_P$  are the electric field intensities without and with an enclosure, respectively.

$$\text{ESE} = 20 \log \frac{E_A}{E_P} [\text{dB}] \quad (2)$$

The SE value depends on the size and the number of Points of Entry (PoE). The PoE is defined by the frequency of the input signal. The maximum frequency of the EMP signal is 100MHz, so that the PoE are designed with a cut of frequency greater than the frequency of the input signal. Leakage of signal occurs when the wavelength of the signal is shorter than the size of the PoE.

## II. DESIGN OF THE SHIELDED ROOM

The shielded room was designed to protect equipment against EMP signals. The EMP signal has a high amplitude and short duration. The characteristics of an EMP signal is a double exponential voltage pulse with amplitude of 55kV, rise/fall time 2.3ns, pulse width 22.73ns as per MIL-STD 461 E/F/G [12] (Figure 1). The EMP signal is given as an input to the shielded room, the equipment in the shielded room is protected by providing all power connections to the equipment. The design of PoE for a shielded room must consider the frequency component of the EMP signal. The frequency of the EMP signal is calculated by taking the inverse of its period. The maximum frequency of the EMP signal is 100MHz, considered from the military standard MIL-STD 461 E/F/G and the PoE are designed such that they should not allow the entry of the EMP signal.

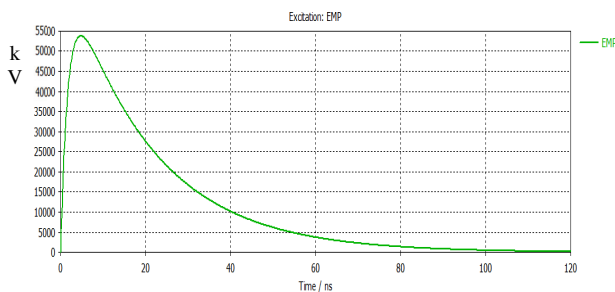


Fig. 1. EMP signal waveform.

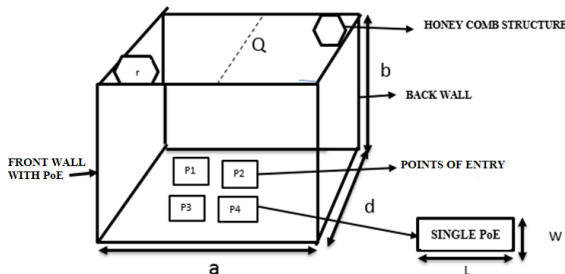


Fig. 2. A shielded room with different PoE.

The shielded room with different PoE for routing the power cables and honeycomb structure to provide ventilation to the equipment is shown in Figure 2.

The shielded room was designed with 4 PoE with different cutoff frequencies to protect. The PoE is considered as the waveguide. The maximum frequency of the EMP signal is 100MHz so that the waveguides are designed with a cutoff frequency greater than that, so that the equipment is protected from the EMP signal. The frequency of the PoE is shown in Table I. Honeycomb structures were considered for ventilation. Two honeycomb structures with cutoff frequency of 40GHz were placed in parallel at the top of the room. The design dimensions of the shielded room with honeycomb structure are also given in Table I.

TABLE I. SHIELDED ROOM DIMENSIONS

Parameters	Length	Width	Height
RF shielded room	$d = 2\text{m}$	$a = 2\text{m}$	$b = 2\text{m}$
Honeycomb panel	$r = 0.45\text{m}$		
PoE	$l$ (m)	$w$ (m)	Cutoff frequency ( $f_c$ )
P1	0.1	0.05	1.5GHz
P2	0.075	0.0375	2 GHz
P3	0.06	0.03	2.5 GHz
P4	0.05	0.025	3 GHz

## III. MATHEMATICAL ANALYSIS

### A. Single Point of Entry

The effectiveness of the metallic shielded room can be determined mathematically by considering a single PoE by using the Transmission Line Model (TLM) method [13]. The equivalent circuit of the shielded room with PoE is represented here.

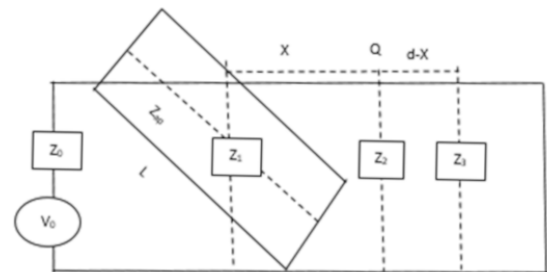


Fig. 3. Circuit model of the shielded room with PoE.

From the equivalent circuit model, Thevenin's theorem is applied to calculate the value of SE. The Q point represents the position of the equipment. The length of the PoE is  $L$ . The applied input voltage is denoted by  $V_0$ , the input impedance of the equivalent circuit is  $Z_0$ , a single PoE is considered in the circuit model, and its impedance is  $Z_{PoE}$ . The SE value is calculated at distance  $X$  from the aperture.

$$V_1 = \frac{V_0 Z_{PoE}}{Z_0 + Z_{PoE}} \quad (3)$$

$$Z_{PoE} = \frac{1}{2} j Z_0 \tan\left(\frac{P_0 l}{2}\right) \quad (4)$$

$$Z_1 = \frac{Z_0 Z_{PoE}}{Z_0 + Z_{PoE}} \quad (5)$$



The dominant mode of propagation in the rectangular waveguide is the TE<sub>10</sub> mode and its characteristic impedance is:

$$Z_g = \frac{Z_0}{\sqrt{1-(f)^2}} \quad (6)$$

where  $f = \frac{\lambda}{2a}$  and the propagation constant  $P_g$  is:

$$P_g = P_0 * \sqrt{1 - (f)^2} \quad (7)$$

where  $P_0 = \frac{2\pi}{\lambda_0}$ .

$Z_g$  and  $P_g$  are imaginary frequencies below the cutoff frequency. Then  $V_1$ ,  $Z_1$ , which are the short circuit terminals of the wave guide to Q are transformed by attributing an equivalent voltage  $V_2$ , source impedance  $Z_2$ , and load impedance  $Z_3$ :

$$V_2 = \frac{V_1}{\cos(P_g X) + j \left( \frac{Z_1}{Z_g} \right) \sin(P_g X)} \quad (8)$$

$$Z_2 = \frac{Z_1 + j l \tan(P_g X)}{1 + j \left( \frac{Z_1}{Z_g} \right) \tan(P_g X)} \quad (9)$$

$$Z_3 = j Z_g \tan(d - X) \quad (10)$$

The voltage at point Q is:

$$V_Q = \frac{V_2 Z_3}{(Z_2 + Z_3)} \quad (11)$$

The SE is:

$$SE = -20 \log_{10} |2V_Q / V_0| \quad (12)$$

### B. Multiple Points of Entry

A shielded room with multiple PoE of different sizes was designed to protect the equipment from the EMP signal. The maximum frequency of the EMP signal should be considered while designing the PoE, so that no external signal is allowed inside the room. The multiple PoE of the side wall of the shielded room can be seen in Figure 4.

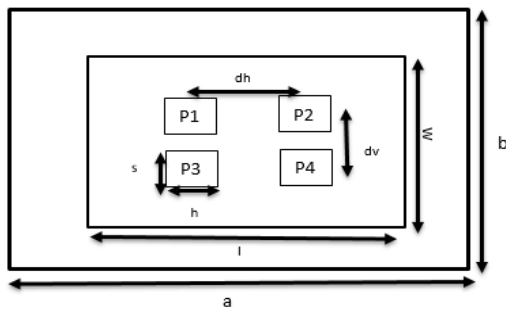


Fig. 4. Hardening of EMP signal with multiple PoE.

A shielded room with length and breadth  $a$  and  $b$ , respectively, with an array of entries with dimensions of  $l$ ,  $w$  is considered. The distance between the PoE is defined as  $dh$  and  $dv$ .  $\lambda_0$  and  $Y_0$  are the free-space wavelength and the intrinsic admittance. Each PoE is considered as a rectangular hole denoted as  $d$ , where  $h$ ,  $s$  are its length and width.

$$d = 0.636(h + s) \quad (13)$$

The SE calculation for multiple PoE when an EMP signal is applied to shielded room follows. The admittance of the equivalent circuit is given by [19]:

$$\frac{Y_{PoEs}}{Y_0} = -j \frac{3d_h d_v \lambda_0}{\pi d^3} + j \frac{288}{\pi \lambda_0 d^2} \left[ \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \left( \frac{\epsilon_m n^2}{d_v^2} + \frac{\epsilon_n m^2}{d_h^2} \right) J_1^2(X) \right] \quad (14)$$

The argument of the Bessel function is:

$$X = \frac{[\pi d(m^2/d_h^2 + n^2/d_v^2)/2]^{1/2}}{(m^2/d_h^2 + n^2/d_v^2)/2]^{1/2}} \quad (15)$$

$Z_{PoEs} = 1/Y_{PoEs}$  represents the PoE array connecting the free space [14]. Figure 4 shows multiple PoE placed on the shielded room wall. The effective wall impedance  $Z_{PoEs}$  is the inverse of  $Z_{PoEs}$ . Using an impedance ratio concept,  $Z_{PoEs}$  becomes:

$$Z_{ah}^l = Z_{ah} \left( \frac{W * L}{a * b} \right) \quad (16)$$

where the length  $l$  and width  $w$  of multiple PoE are:

$$l = \frac{dh}{2} + [(m-1) * dh] + \frac{dh}{2} \quad (17)$$

$$w = \frac{dv}{2} + [(n-1) * dv] + \frac{dv}{2} \quad (18)$$

where  $m$  and  $n$  represent the number of apertures in length and width of the array, respectively.

$$V_1 = \frac{(V_0 * Z_{ah}^l)}{(Z_0 + Z_{ah}^l)} \quad (19)$$

$$Z_1 = \frac{(Z_0 * Z_{ah}^l)}{(Z_0 + Z_{ah}^l)} \quad (20)$$

The SE is:

$$SE = -20 \log_{10} |2V_Q / V_0| \quad (21)$$

### C. Honeycomb Structure

The attenuation constant for the wave guide is given below [15]. The attenuation constant is used for the calculation of the SE value of the wave guide:

$$\alpha = w(\mu \epsilon)^{\frac{1}{2}} \sqrt{\frac{f_c}{f} - 1} \quad (22)$$

where  $f_c$  is the cutoff frequency of the waveguide. The attenuation constant of the waveguide in dB can be calculated by placing the cutoff frequency value in (22). The SE value for a single hexagon cell (honeycomb structure) is calculated by using (23):

$$SE_{dB} = 17.5 \frac{d}{g} \sqrt{1 - \left( \frac{g_f}{96654} \right)^2} \quad (23)$$

In [16], an infinite array of parallel-plate waveguides is analyzed by the Wiener-Hopf method. The resulting equation is shown below. The first part gives the SE of the unit cell of a hexagon waveguide, while the second term is the SE of an infinite array of parallel-plate waveguides [24].

$$SE_{dB} = 17.5 \frac{d}{g} \sqrt{1 - \left( \frac{g_f}{96654} \right)^2} - 20 \log_{10} \frac{2k_g}{\pi} \cos \varphi \quad (24)$$

where  $k$  is the wave number,  $g$  is a transverse dimension of the waveguide, and  $\varphi$  is the angle of an incident wave.

The SE equation for a hexagon structure has been modified by adding a third term to (24), so the performance of the hexagon structure at lower frequency is increased. The normalized frequency during the design of the honeycomb structure should be greater than 5 times the radius of the cell of the honeycomb structure. SE value does not depend on the number of hexagon cells in the honeycomb structure, but on the size and the length of a single cell.

$$SE_{dB} = 17.5 \frac{d}{g} \sqrt{1 - \left(\frac{g_f}{96654}\right)^2} - 20 \log_{10} \frac{2k_g}{\pi} \cos \varphi - 20 \log_{10} \frac{2R_g}{f} \quad (25)$$

#### IV. SIMULATIONS

##### A. Shielded Room with a Single Point of Entry

A shielded room with a size of 2m×2m×2m with a single PoE of size 0.075m×0.0375m is shown in Figure 5. The probe was placed at the center of the shielded room. The position of the probe varied inside the room to find the fields inside the room with an EMP signal transmitted continuously to the shielded room. The material used in the design of the shielded room is Perfect Electrical Conductor (PEC). The thickness of the shielded material is 2mm.

SE value varies with the size of the PoE. A PoE of size 10cm×5cm is considered and the SE value is 255dB. If the size of the aperture is decreased, the SE value is increased as shown in Table II. The size of the PoE depends on the frequency of the EMP signal. Figure 6 shows the SE value for 4 apertures.

TABLE II. SE VALUES FOR DIFFERENT POINTS OF ENTRY

PoE	a (m)	b (m)	SE (dB)
P1	0.1	0.05	290
P2	0.075	0.0375	310
P3	0.06	0.03	320
P4	0.05	0.025	330

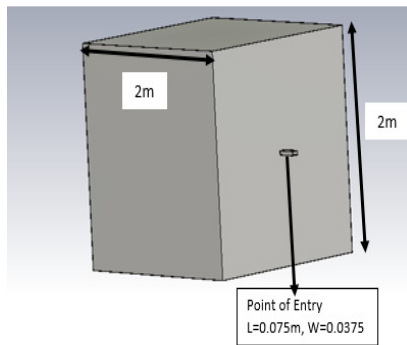


Fig. 5. Shielded room with a single PoE.

##### B. Shielded Room with Multiple Points of Entry

The shielded room with different entries was designed in CST simulation tool as shown in Figure 6. The SE value for an individual PoE is shown in Figure 7.

Let us consider shielded rooms with 4 PoE with different dimensions. The dimensions of an entry depend on its cutoff frequency. The cutoff frequency should be greater than the

frequency of the EMP signal. In this condition the SE value meets the military standards MIL-STD-188-125-1 and 2. Figure 8 shows the total SE value of 4 different dimensions. The SE value decreases with the frequency of operation, which is below the resonant frequency. The resonant frequency of the shielded room was calculated by considering the room dimensions along with mode of propagation. The electric field probe is considered at the center of the shielded room, so that maximum SE value is obtained at that point. The SE value for 4 different PoE is 250dB.

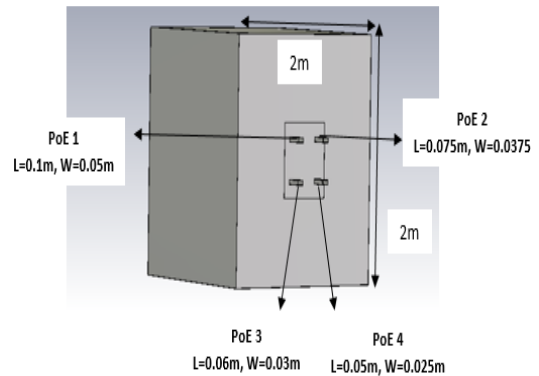


Fig. 6. Shielded room with 4 PoE with different sizes.

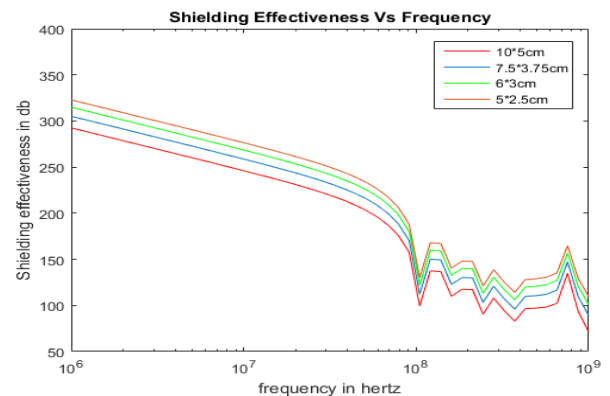


Fig. 7. SE value for different PoE sizes when an EMP signal is given as input.

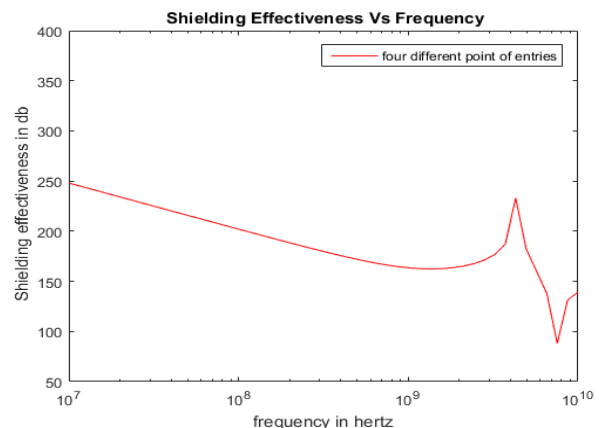


Fig. 8. Effect of SE when an EMP is applied to multiple PoE.

### C. Shielded Room with Honeycomb Structures

The shielded room was designed with honeycomb structures used for ventilation, placed at the top of the room. The honeycomb structures were designed such that no external signal would enter the shielded room. Figure 9 shows the honeycomb structure. The honeycomb structures placed on the shielded room are shown in Figure 10. The SE value is calculated when the EMP signal is applied to the room.

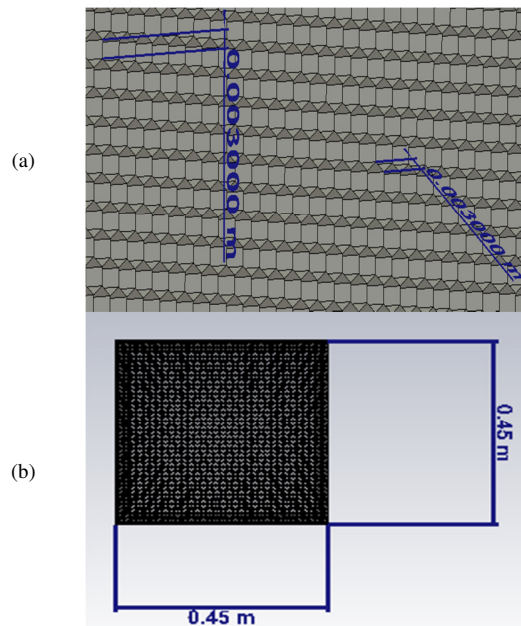


Fig. 9. Cell dimension details of the honeycomb structure, (b) total honeycomb structure.

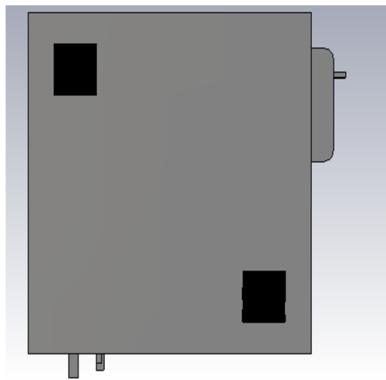


Fig. 10. Shielded room with 2 honeycomb structures.

### D. Shielded Room with Points of Entry and Honeycomb Structure

The shielded room was designed with 4 PoE and 2 honeycomb structures, as shown in Figure 11. The EMP signal is incident on the shielded room with a voltage of 55kv. The SE value is calculated by placing the field probe at the center of the room. The total shielded room SE value is given in Figure 12.

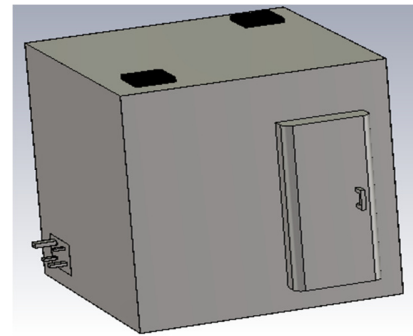


Fig. 11. Shielded room with 4 PoE and 2 honeycomb structures.

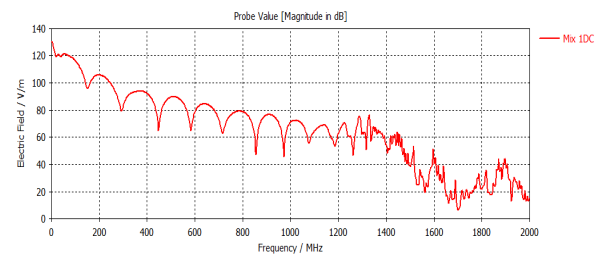


Fig. 12. The electric field value for a totally shielded room with 4 PoE and 2 honeycomb structures.

Total SE is calculated by taking the difference between the electric field with and without shield.

### E. Effect of Varying Field Probe

The EMP signal is incident directly to the shielded room, which has dimensions of  $2\text{m} \times 2\text{m} \times 2\text{m}$  with a single PoE. The EMP signal is parallel to the shielded enclosure. In this case the polarization type is fixed, so the field probe inside the enclosure is varied so that the SE value varies with changes in the position of the probe. The field probe gives the position of the equipment so that maximum SE is noted. The shielded room with a single PoE is shown in Figure 4. The change in the value of SE with the position of the field probe is shown in Figure 8. Consider a point Q at a different distances  $x_1$ ,  $x_2$ ,  $x_3$  from the front panel of the shielded room. The SE value increases when the position of the probe is near the wall of the shielded room as clearly observed in Figure 13.

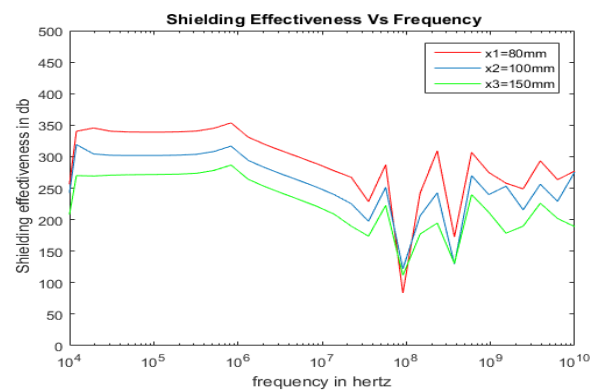


Fig. 13. Varying position of the field probe in the shielded room when the EMP is applied.

By observing the graph of Figure 13, we note that the SE value increases when the position of the probe is to the shielded room front wall and at the center of the room, and decreases when the probe is near to the back wall of the shielded room.

## V. CONCLUSION

Theoretical calculations and simulation analysis were conducted in this paper for a shielded room against EMP signals. The transmission line method was used in the theoretical calculations of the shielded room with 4 different PoE dimensions. The smallest PoE with 0.05m length and 0.025m width has better SE value of 330dB than a PoE with greater size, with 0.1m length and 0.05m width, where the SE value decreases to 250dB. Simulation analysis is done with 3DEM CST modeling tool to verify the combined effect of multiple PoE and honeycomb structures. The total SE value for a shielded room with 4 PoE and 2 honeycomb structures is 140dB.

The novelty of this paper is that it was conducted as per MIL-STD-188-125-1/2 to protect the equipment from external EMP signals. This design of shielded room is useful for RF engineers that need to estimate the interference levels in sensitive equipment due to EMP signals.

## REFERENCES

- [1] J. Chen and J. Wang, "A Three-Dimensional Semi-Implicit FDTD Scheme for Calculation of Shielding Effectiveness of Enclosure With Thin Slots," *IEEE Transactions on Electromagnetic Compatibility*, vol. 49, no. 2, pp. 354–360, Feb. 2007, <https://doi.org/10.1109/TEMC.2007.893329>.
- [2] A. Rabat, P. Bonnet, K. E. K. Drissi, and S. Girard, "Analytical Formulation for Shielding Effectiveness of a Lossy Enclosure Containing Apertures," *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 5, pp. 1384–1392, Jul. 2018, <https://doi.org/10.1109/TEMC.2017.2764327>.
- [3] F. G. K. Abdulla and R. Abdulla, "A Comparative Application for Evaluating Composite Fabrics Used in Electromagnetic Shielding," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2156–2159, Dec. 2017, <https://doi.org/10.48084/etasr.1480>.
- [4] A. Khodadadi, M. H. Nazari, and S. H. Hosseini, "Designing an Optimal Lightning Protection Scheme for Substations Using Shielding Wires," *Engineering, Technology & Applied Science Research*, vol. 7, no. 3, pp. 1595–1599, Jun. 2017, <https://doi.org/10.48084/etasr.1175>.
- [5] Q. Xu, Y. Huang, X. Zhu, L. Xing, Z. Tian, and C. Song, "Shielding Effectiveness Measurement of an Electrically Large Enclosure Using One Antenna," *IEEE Transactions on Electromagnetic Compatibility*, vol. 57, no. 6, pp. 1466–1471, Sep. 2015, <https://doi.org/10.1109/TEMC.2015.2477455>.
- [6] S. C. A. Bikkina and P. V. Y. Jayasree, "Development of a Wire Mesh Composite Material for Aerospace Applications," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9310–9315, Oct. 2022, <https://doi.org/10.48084/etasr.5201>.
- [7] S. S. Pudipeddi, P. V. Y. Jayasree, and S. G. Chintala, "Polarization Effect Assessment of Sub-6 GHz Frequencies on Adult and Child Four-Layered Head Models," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8954–8959, Aug. 2022, <https://doi.org/10.48084/etasr.5096>.
- [8] P. Dehkhoda, A. Tavakoli, and R. Moini, "An Efficient and Reliable Shielding Effectiveness Evaluation of a Rectangular Enclosure With Numerous Apertures," *IEEE Transactions on Electromagnetic Compatibility*, vol. 50, no. 1, pp. 208–212, Oct. 2008, <https://doi.org/10.1109/TEMC.2007.911922>.
- [9] T. Martin, M. Backstrom, and J. Loren, "Semi-empirical modeling of apertures for shielding effectiveness simulations," *IEEE Transactions on Electromagnetic Compatibility*, vol. 45, no. 2, pp. 229–237, Feb. 2003, <https://doi.org/10.1109/TEMC.2003.810818>.
- [10] M. Edrisi and A. Khodabakhshian, "Simple Methodology for Electric and Magnetic Shielding Effectiveness Computation of Enclosures for Electromagnetic Compatibility use," *Journal of Electromagnetic Waves and Applications*, vol. 20, no. 8, pp. 1051–1060, Jan. 2006, <https://doi.org/10.1163/156939306776930312>.
- [11] "IEEE Standard Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures," *IEEE Std 299-2006 (Revision of IEEE Std 299-1997)*, pp. 1–52, Oct. 2007, <https://doi.org/10.1109/IEEESTD.2007.323387>.
- [12] A. Wraight, W. D. Prather, and F. Sabath, "Developments in Early-Time (E1) High-Altitude Electromagnetic Pulse (HEMP) Test Methods," *IEEE Transactions on Electromagnetic Compatibility*, vol. 55, no. 3, pp. 492–499, Jun. 2013, <https://doi.org/10.1109/TEMC.2013.2241442>.
- [13] M. P. Robinson *et al.*, "Analytical formulation for the shielding effectiveness of enclosures with apertures," *IEEE Transactions on Electromagnetic Compatibility*, vol. 40, no. 3, pp. 240–248, Dec. 1998, <https://doi.org/10.1109/15.709422>.
- [14] *MIL-STD-188-125-2: High-altitude Electromagnetic Pulse (Hemp) Protection For Ground-based C4I Facilities Performing Critical, Time-urgent Missions, Part 2: Transportable Systems*. Washington DC, USA: Department of Defence, 1999.
- [15] H.-R. Im, I.-K. Jung, J.-G. Yook, and H.-R. Song, "Analysis of EMP Penetration into an Enclosure with Electromagnetic Shielding Material," in *2018 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)*, Boston, MA, USA, Jul. 2018, pp. 31–32, <https://doi.org/10.1109/USNC-URSI.2018.8602763>.
- [16] W. A. Bereuter and D. C. Chang, "Shielding Effectiveness of Metallic Honeycombs," *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-24, no. 1, pp. 58–61, Oct. 1982, <https://doi.org/10.1109/TEMC.1982.304013>.

# The Effect of Rigid Inclusions on the Dynamic Response of Highway Embankment

**Faris Bouabdallah**

Department of Civil Engineering, Laboratory LMGHU, 20 August 1955 University of Skikda, Algeria  
f.bouabdallah@univ-skikda.dz  
(corresponding author)

**Kamel Goudjil**

Department of Civil Engineering, University of Mohamed-Cherif Messaadia, Algeria  
k\_goudjil@yahoo.com

**Messast Salah**

Department of Civil Engineering, Laboratory LMGHU, 20 August 1955 University of Skikda, Algeria  
msalah2007msalah@gmail.com

*Received: 8 October 2022 | Revised: 28 October 2022 | Accepted: 6 November*

## ABSTRACT

Soft soils are widespread in many areas of the East-West highway of Algeria, covering large areas. Construction projects such as highway embankments, airfields, roads, and railways in such areas experience several problems due to their low strength and permeability. This study used a fully dynamic analysis to evaluate the dynamic response of a highway embankment. The numerical analysis was carried out using the FLAC 2D version 8.10 software. The main objective was to evaluate the effect of rigid inclusions on the dynamic response of highway embankments by monitoring settlement, shear-strain curves, and maximum soil accelerations and displacements at different depth points. The obtained results showed that the configuration of the rigid inclusions significantly influences the dynamic response of the model. For instance, the settlement at 0.9m is 16% higher than at 0.2m from rigid inclusion.

*Keywords-inclusions; dynamic response; flac; settlement; shear strain curve*

## I. INTRODUCTION

Several studies have been carried out to better understand the behavior of soft soils, due to the spread of soft soils, which cover several areas in Algeria, and develop disaster prevention measures, estimate their parameters, and study the effect of some materials such as rubber concrete and lime on the behavior of composite clayey soils [1-5]. Structures such as highway or railway embankments, tanks, walls, or slopes face several problems with such soft foundations. These problems relate to bearing capacity failures, intolerable total and differential settlements, large lateral pressures, and slope instabilities. Several studies showed that the principal cause of several landslides, which affected the East-West highway in Algeria, is the presence of clayey formations [6-7]. A variety of ground improvement techniques have been used to address the concerns on soft soil foundations, such as geosynthetic [8-11], geogrid [12-14], pile [15-16], PVD [17], soil nail [18-19], chemical injection and grouting [20], stone column [21-24], and rigid inclusions [25-27]. However, the suitable solution depends on several parameters, such as the thickness of the soft

soil, the nature of the project to build, the requirements of the client, etc.

Rigid inclusions are a ground improvement technique that has a significant advantage over deep foundations, as they can provide the needed settlement control and retain the shallow foundation support of the structure. Rigid inclusions, known by different names such as Controlled Modulus Columns (CMC), pile-supported earth platforms, jet grouting columns, or soil column reinforcement, are vertical elements across soil layers with low bearing capacity and/or high compressibility, extending down to a more resistant layer. Due to their higher stiffness compared to surrounding soil, they support a portion of the loads applied at the ground's surface. Therefore, the loads taken by the soft soil can be reduced to an acceptable level of soil-bearing capacity and settlement. Several studies investigated the effect of rigid inclusions to improve soft soils. Some studies used the Stress Reduction Ratio (SRR) to find the optimal design [28-29]. However, they did not consider several factors that can potentially influence the SRR, such as the characteristics of the subgrade soils and the influence of dynamic loadings. Other approaches adopted rigid inclusions techniques to improve soft soil intended to support a railroad

embankment. The rigid inclusions reduced the ground settlements by 15-20%, depending on the subgrade plasticity index [30]. Several numerical studies investigated the seismic behavior of highway embankments in improved soft soils. These studies used the linear elastic-perfectly plastic constitutive model with a Mohr-Coulomb failure criterion to represent the behavior of the soil and LTP [31-36]. Some studies used the constitutive models UBCYST and PM4silt to take into account the nonlinear behavior of soils during seismic motions [37-40].

This paper presents a fully dynamic analysis and evaluation of the seismic response of highway embankments over rigid inclusions to improve soft soil. The UBCHYST model [41] was adopted to consider the dynamic characteristics of the soil during seismic movements. The effect of ground motions of the CHI-CHI earthquake and two types of rigid inclusion configurations, floating and placed in hard clay, were studied.

## II. NUMERICAL MODELING

### A. The Geometry of the Numerical Model

A 3-layer soil model was considered, where 10m thick soft soil was built over a 5m thick hard soil layer. The highway was constructed on an embankment with an 11.75m width and a slope angle of approximately 18°. The toe of the embankment was at an elevation of 15m. The embankment includes a load transfer platform 0.6m thick as part of the embankment materials. To simplify the problem, the soil layers were supposed to be horizontal, as shown in Figure 1. The water table was not set up in the model and drained conditions were considered. A finite difference mesh with a size of 0.25cm for each element was chosen to avoid wave distortion during seismic wave transmission through the soil [42]. The constitutive model UBCHYST can be used with low and high-permeability soils, such as clayey, silty, and granular. In this hysteretic model, the shear modulus is a function of the stress ratio and varies throughout the loading cycle to give hysteretic stress-strain loops of varying amplitude and area damping throughout the earthquake excitation. A detailed description of all UBCHYST model parameters was given in [41].

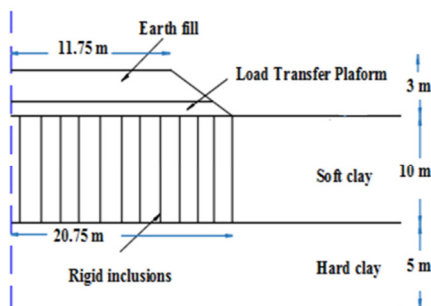


Fig. 1. Geometry of the numerical model

TABLE I. RIGID INCLUSION PARAMETERS

Density (kg/m <sup>3</sup> )	Shear modulus <i>G</i> (MPa)	Young modulus <i>E</i> (MPa)
25000	12.5	30

Tables I and II present the parameters of the soil and structural elements. Soil characteristics were taken from [31, 39, 43]. The calibration parameters of the constitutive UBCHYST model were determined based on laboratory experiments [39, 43]. Since the soft soil used in [31] was modeled as the Mohr-Coulomb constitutive model, a cyclic shear test was conducted to determine the calibration parameters of the UBCHYST model using the modulus reduction curves from [44].

TABLE II. SOIL PARAMETERS OF THE UBCHYST MODEL

Soil parameters	Stiff clay	Soft clay	LTP	Earth fill
Bulk modulus <i>K</i> (MPa)	76	8.33	41.66	76
Shear modulus <i>G</i> (MPa)	38	3.8	19	38
Density (kg/m <sup>3</sup> )	2000	1600	2000	2000
Friction angle $\Phi$ (deg)	30	25	25	35
Cohesion <i>C</i> (kPa)	15.2	5	50	0
Tensile <i>T</i>	0	0	0	0
Dilation angle <i>dil</i>	0	0	0	0
Stress rate factor <i>Rf</i>	1	0.8	0.98	1
Stress rate exponent <i>n</i>	2	2	2.5	2
First cycle factor <i>Mod1</i>	0.75	0.75	1	0.75
Large strains exponent <i>Rm</i>	0.5	0.5	0.5	0.5
Large strains factor <i>dfac</i>	0	0	0	0

### B. Rigid Vertical Elements

Numerical analysis was conducted using 12 rigid inclusions per row in the soft soil, with 10m length and 0.6m diameter. The separation between elements in both directions was equal to 1.8m. The rigid inclusion characteristics were taken from [31].

### C. Boundary Conditions and Interfaces

Artificial boundaries were used in the static analysis to depict the semi-infinite nature of the soil. The side boundaries were fixed in the horizontal direction while the bottom part was fixed in all directions. However, these boundary conditions were replaced in the dynamic analysis by quiet boundaries along the bottom of the model to minimize the effects of reflected waves at the bottom and free-field boundaries on the sides to avoid wave reflections[45]. To apply quiet boundary conditions along the same boundary as the dynamic input, the dynamic input must be applied as a shear stress boundary, because the effect of the quiet boundary will be nullified if the input is applied as an acceleration or velocity wave. To do this, the velocity record was converted into a shear stress boundary condition using

$$\sigma_s = factor \times (\rho \times C_s) \times v_s \quad (1)$$

where  $\sigma_s$  is the applied shear stress,  $\rho$  is the mass density of the material at the boundary,  $C_s$  is the speed of s-wave propagation through the medium at the boundary,  $v_s$  is the input shear particle velocity, and the *factor* is generally equal to 2.

Figures 2 and 3 show the time history acceleration and velocity of the CHI-CHI earthquake. Regarding soil-rigid inclusion interfaces, since the soil foundations of this study were clayey soils, the interfaces were assumed to have a zero-friction angle and the same cohesive strength as the surrounding soil [46]. The shear and normal stiffness values were determined using (2), based on the FLAC manual [45].



$$K_n = K_s = 10 \left( \frac{K+1.3 G}{\Delta Z_{min}} \right) \quad (2)$$

where  $K$  and  $G$  are the bulk and shear modulus of the adjacent soil, respectively, and  $\Delta Z_{min}$  is the smallest edge of the adjacent soil element.

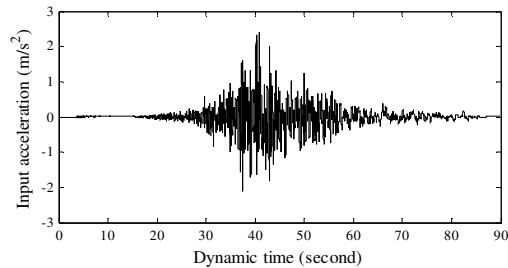


Fig. 2. Input acceleration time history of the CHI-CHI earthquake in 1999.

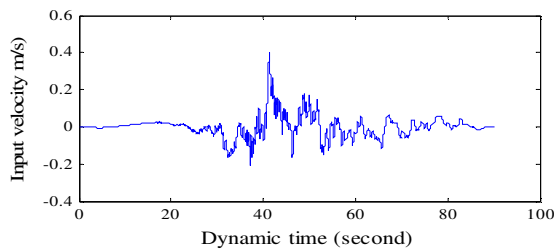


Fig. 3. Input velocity time history of the CHI-CHI earthquake in 1999.

### III. RESULTS AND DISCUSSION

The soil response in terms of settlements at the bottom of the embankment and along the depth, and shear stresses and strains were evaluated for both cases of Rigid Inclusions (RI), placed on hard clay and floating. All these results took into account only the influence of the dynamic loading on the soft soils improved with rigid inclusions. In the first step, an initial stress state was generated. Then, the vertical reinforcements were installed, and the last static calculation step considered the activation of the earth platform and the highway embankment. The required seismic boundary conditions, free field boundaries, and quiet boundaries were added for the dynamic analysis, and the dynamic analysis was executed by applying the corresponding horizontal wave using shear stress (2) at the base of the models.

#### A. Settlements at the Bottom of the Embankment

Figure 4 shows the histories of soil settlements at the base of LTP for different points relative to the RI. It is noticeable that the settlement at 0.9m from RI is greater than at 0.2m. For example, the settlement of the point at 0.9m is 16% higher than at 0.2m from RI. The settlement decreases the closer the point is to the rigid inclusion. This can be explained by the effect of shear stiffness provided via the interaction of the rigid elements with the surrounding soil, which leads to reduce the settlement of the surrounding soil.

#### B. Settlements along the Depth

Figure 5 shows the histories of soil settlements at different depths along the centreline axis. It is clear that the maximum

settlements (6.4cm) occur at the top of the LTP layer and are reduced with depth. The maximum settlement in the middle of soft clay (3.3cm) is 23.6% greater compared to those in its base (0.78cm).

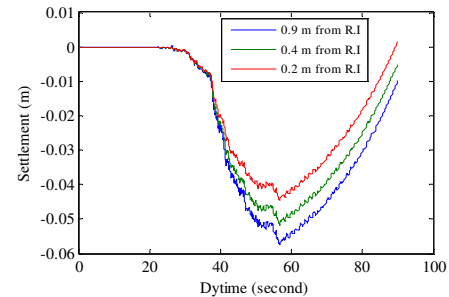


Fig. 4. Settlement at the bottom of the embankment.

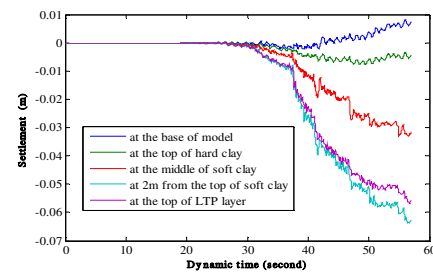


Fig. 5. Settlements along the depth.

#### C. Shear Strains and stress curves

Figures 6-8 show the hysteresis curve that relates shear stresses and shear strains obtained after dynamic loading in the soil.

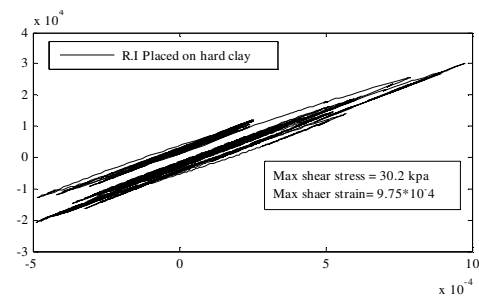


Fig. 6. Shear stress and strain curve at the middle of hard clay.

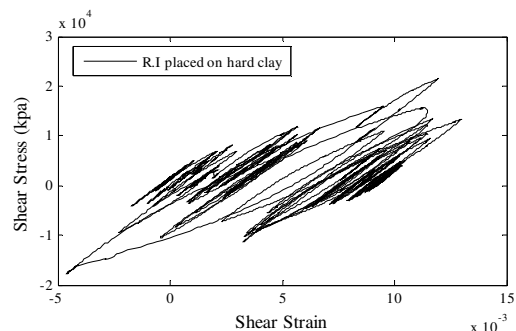


Fig. 7. Shear stress and strain curve at the middle of soft clay.

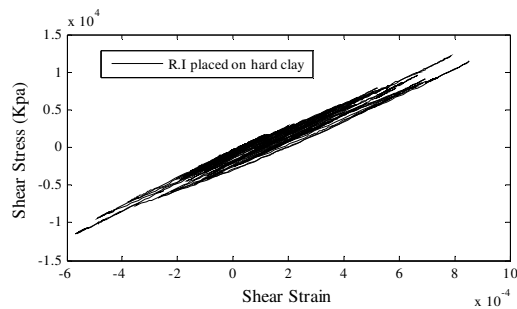


Fig. 8. Shear stress and strain curve at the middle of LTP.

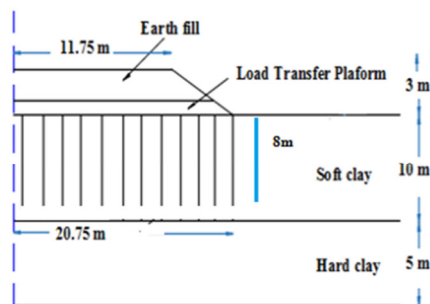


Fig. 9. Floating rigid inclusions.

The values obtained were recorded in the middle of the soft clay and the LTP layer. Figures 3-5 show maximum shear strain at  $9.75 \times 10^{-4}$ ,  $13 \times 10^{-3}$ , and  $8.55 \times 10^{-4}$ , respectively. The distribution of the maximum shear strain is higher in the middle of the soft clay. At the LTP layer and the middle hard clay, the maximum shear strains are very close, but the shear stresses differ. These results are due to the higher value of shear stiffness of the hard clay and the LTP layer compared to the soft clay. Figure 9 shows 8m long floating rigid inclusions embedded in the soft soil. Shear stresses and strains at the same points were estimated to show the efficiency of rigid inclusions. Figures 10 and 11 show the hysteresis curve that relates shear stresses and strains obtained after dynamic loading in the soil for floating rigid inclusions. The obtained values were recorded at the same points for rigid inclusions placed on hard clay. The maximum shear strain value in the middle of soft clay was  $15.2 \times 10^{-3}$ , having an increase of  $2.2 \times 10^{-3}$ . Also, the maximum value of shear strain in the LTP layer reached  $9.63 \times 10^{-4}$ , having an increase of  $1.08 \times 10^{-4}$ .

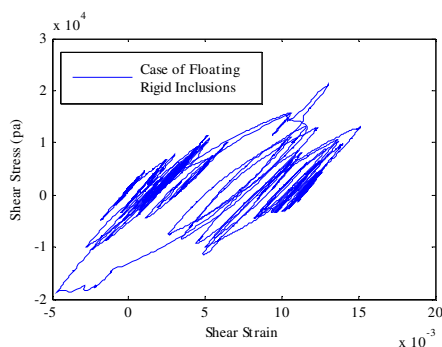


Fig. 10. Shear stress and strain curve at the middle of soft clay.

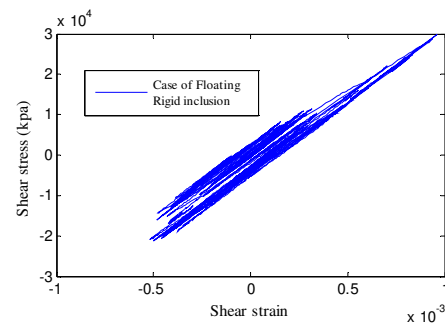


Fig. 11. Shear stress and strain curve at the middle of LTP.

#### IV. CONCLUSION

A fully dynamic numerical modeling analysis was conducted to evaluate the seismic response of a highway embankment constructed over reinforced soft soil. Two configuration types were considered to study the effect of rigid inclusions, placed on hard soil and floating. The numerical modeling results of all configurations were compared in terms of settlement, shear stress, and shear strain. Settlements at the base of the embankment were evaluated in three different positions relative to the position of the rigid inclusion. The settlement was important at 0.9m from the rigid inclusions while the minimum settlement was recorded at 0.2m. This can be explained by the shear stiffness provided via the interaction of the rigid elements with the surrounding soil, which reduces the settlement of the surrounding soil. The maximum value of shear strains was recorded in the middle of the soft clay compared to the LTP layer and the middle of the hard clay for all cases. The cohesion of the LTP layer and the hard clay increases the strength of the soil and reduces shear strain. In floating rigid inclusions, the maximum shear strain in the middle of soft clay was  $15.2 \times 10^{-3}$ , which is an increase of  $2.2 \times 10^{-3}$ . The maximum value of shear strain in the LTP layer reached  $9.63 \times 10^{-4}$ , which is an increase of  $1.08 \times 10^{-4}$ . These results are due to the shear force that can develop along the rigid inclusion/soil interface, which is a function of the cohesive strength of the interface. Several studies investigated the effect of various parameters such as rigid inclusion diameter, spacing, seismic force intensity, etc. However, few studies investigated the influence of rigid inclusions using a fully dynamic analysis. This paper presented a dynamic study using the UBCHYST model to highlight the nonlinear dynamic behavior of soft soils.

#### REFERENCES

- [1] S. M. Aissa Mamoune, "Contribution à la mesure, prévision et modélisation du comportement des sols expansifs," MSc Thesis, University of Tlemcen, Algeria, 2002.
- [2] M. Afes, "Contribution à la détermination des paramètres de gonflement des sols, et étude de l'argile de Mila (Algérie) traitée à la chaux," Ph.D. dissertation, Lyon, INSA, 1996.
- [3] D. Athmania, A. Benaissa, A. Hammadi, and M. Bouassida, "Clay and Marl Formation Susceptibility in Mila Province, Algeria," *Geotechnical and Geological Engineering*, vol. 28, no. 6, pp. 805–813, Nov. 2010, <https://doi.org/10.1007/s10706-010-9341-5>.
- [4] F. Klouche, E. Kara Mostefa, A. Sekkel, M. Maliki, and M. Bouguenaya, "Effect of waste concrete on the compressibility of clay from the Bouhanak region (Algeria)," *Euro-Mediterranean Journal for*

- Environmental Integration*, vol. 7, no. 2, pp. 191–200, Jun. 2022, <https://doi.org/10.1007/s41207-022-00308-4>.
- [5] H. Trouzine, M. Bekhiti, and A. Asroun, "Effects of scrap tyre rubber fibre on swelling behaviour of two clayey soils in Algeria," *Geosynthetics International*, vol. 19, no. 2, pp. 124–132, Apr. 2012, <https://doi.org/10.1680/gein.2012.19.2.124>.
- [6] B. E. Kechebour and A. Talah, "Durability of New Roads: Case Study of the Algerian East West Highway," *International Journal of Structural and Civil Engineering Research*, vol. 7, no. 1, pp. 87–91, Feb. 2018, <https://doi.org/10.18178/ijscer.7.1.87-91>.
- [7] A. Benaissa and M. A. Bellouche, "Propriétés géotechniques de quelques formations géologiques propices aux glissements de terrains dans l'agglomération de Constantine (Algérie)," *Bulletin of Engineering Geology and the Environment*, vol. 57, no. 3, pp. 301–310, Mar. 1999, <https://doi.org/10.1007/s100640050049>.
- [8] J. Han and M. A. Gabr, "Numerical Analysis of Geosynthetic-Reinforced and Pile-Supported Earth Platforms over Soft Soil," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 128, no. 1, pp. 44–53, Jan. 2002, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2002\)128:1\(44\)](https://doi.org/10.1061/(ASCE)1090-0241(2002)128:1(44)).
- [9] P. Ariyaratne and D. S. Liyanapathirana, "Review of existing design methods for geosynthetic-reinforced pile-supported embankments," *Soils and Foundations*, vol. 55, no. 1, pp. 17–34, Feb. 2015, <https://doi.org/10.1016/j.sandf.2014.12.002>.
- [10] T. Thyagaraj, Ed., "Ground Improvement Techniques and Geosynthetics: IGC 2016 Volume 2," in *IGC 2016 Vol. 2*, Chennai, India, 2019, vol. 14, <https://doi.org/10.1007/978-981-13-0559-7>.
- [11] M. Touahmia, "Performance of Geosynthetic-Reinforced Soils Under Static and Cyclic Loading," *Engineering, Technology & Applied Science Research*, vol. 7, no. 2, pp. 1523–1527, Apr. 2017, <https://doi.org/10.48084/etasr.1035>.
- [12] M. S. S. Almeida, M. Ehrlich, A. P. Spotti, and M. E. S. Marques, "Embankment supported on piles with biaxial geogrids," *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, vol. 160, no. 4, pp. 185–192, Oct. 2007, <https://doi.org/10.1680/jge.2007.160.4.185>.
- [13] H. L. Liu, C. W. W. Ng, and K. Fei, "Performance of a Geogrid-Reinforced and Pile-Supported Highway Embankment over Soft Clay: Case Study," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 133, no. 12, pp. 1483–1493, Dec. 2007, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2007\)133:12\(1483\)](https://doi.org/10.1061/(ASCE)1090-0241(2007)133:12(1483)).
- [14] B. Mazouz, T. Mansouri, M. Baazouzi, and K. Abbeche, "Assessing the Effect of Underground Void on Strip Footing Sitting on a Reinforced Sand Slope with Numerical Modeling," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9005–9011, Aug. 2022, <https://doi.org/10.48084/etasr.5131>.
- [15] R. P. Chen, Z. Z. Xu, Y. M. Chen, D. S. Ling, and B. Zhu, "Field Tests on Pile-Supported Embankments over Soft Ground," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 136, no. 6, pp. 777–785, Jun. 2010, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0000295](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000295).
- [16] C. Yun-min, C. Wei-ping, and C. Ren-peng, "An experimental investigation of soil arching within basal reinforced and unreinforced piled embankments," *Geotextiles and Geomembranes*, vol. 26, no. 2, pp. 164–174, Apr. 2008, <https://doi.org/10.1016/j.geotexmem.2007.05.004>.
- [17] X. Jiang, Y. Jiang, C. Y. Wu, W. Q. Wang, J. Y. Geng, and Y. J. Qiu, "Numerical analysis for widening embankments over soft soils treated by PVD and DJM columns," *International Journal of Pavement Engineering*, vol. 21, no. 3, pp. 267–279, Feb. 2020, <https://doi.org/10.1080/10298436.2018.1461869>.
- [18] D. A. Mangnejo, S. J. Oad, S. A. Kalhor, S. Ahmed, F. H. Laghari, and Z. A. Siyal, "Numerical Analysis of Soil Slope Stabilization by Soil Nailing Technique," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4469–4473, Aug. 2019, <https://doi.org/10.48084/etasr.2859>.
- [19] W. R. Azzam and A. Basha, "Utilization of soil nailing technique to increase shear strength of cohesive soil and reduce settlement," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 9, no. 6, pp. 1104–1111, Dec. 2017, <https://doi.org/10.1016/j.jrmge.2017.05.009>.
- [20] Sina Kazemain, "Review of soft soils stabilization by grouting and injection methods with different chemical binders," *Scientific Research and Essays*, vol. 7, no. 24, Jun. 2012, <https://doi.org/10.5897/SRE11.1186>.
- [21] Z. Guetif, M. Bouassida, and J. M. Debats, "Improved soft clay characteristics due to stone column installation," *Computers and Geotechnics*, vol. 34, no. 2, pp. 104–111, Mar. 2007, <https://doi.org/10.1016/j.compgeo.2006.09.008>.
- [22] J. Castro, M. Karstunen, and N. Sivasithamparam, "Influence of stone column installation on settlement reduction," *Computers and Geotechnics*, vol. 59, pp. 87–97, Jun. 2014, <https://doi.org/10.1016/j.compgeo.2014.03.003>.
- [23] S. W. Abusharar and J. Han, "Two-dimensional deep-seated slope stability analysis of embankments over stone column-improved soft clay," *Engineering Geology*, vol. 120, no. 1, pp. 103–110, Jun. 2011, <https://doi.org/10.1016/j.enggeo.2011.04.002>.
- [24] S. Saxena and L. B. Roy, "Suitability Analysis of Stone Column Materials with PLAXIS," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8421–8425, Apr. 2022, <https://doi.org/10.48084/etasr.4761>.
- [25] J. A. B. Garcia, J. F. Rodríguez Rebolledo, D. V. dos Santos Mützenber, B. Caicedo, and G. de Farias Neves Gitirana, "Experimental Investigation of a Load-Transfer Material for Foundations Reinforced by Rigid Inclusions," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 147, no. 10, Oct. 2021, Art. no. 04021110, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002649](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002649).
- [26] L. Briançon, D. Dias, and C. Simon, "Monitoring and numerical investigation of a rigid inclusions–reinforced industrial building," *Canadian Geotechnical Journal*, Mar. 2015, <https://doi.org/10.1139/cgj-2014-0262>.
- [27] V. D. Tran, J. J. Richard, and T. Hoang, "Soft Soil Improvement Using Rigid Inclusions: Toward an Application for Transport Infrastructure Construction in Vietnam," in *New Prospects in Geotechnical Engineering Aspects of Civil Infrastructures*, HangZhou, China, Jul. 2018, pp. 89–99, [https://doi.org/10.1007/978-3-319-95771-5\\_8](https://doi.org/10.1007/978-3-319-95771-5_8).
- [28] D. Mazzei, K. Kniss, F. Elsaid, and Y. Zhang, "Rigid Inclusions Ground Improvement for a New Energy Facility: Design, Construction, and Full-Scale Embankment Load Testing and Results," pp. 101–114, Mar. 2019, <https://doi.org/10.1061/9780784482117.009>.
- [29] P. Burtin and J. Racinais, "Embankment on Soft Soil Reinforced by CMC Semi-Rigid Inclusions for the High-speed Railway SEA," *Procedia Engineering*, vol. 143, pp. 355–362, Jan. 2016, <https://doi.org/10.1016/j.proeng.2016.06.045>.
- [30] D. Wang, M. Sánchez, and J.-L. Briaud, "Numerical study on the effect of rigid inclusions on existing railroads," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 43, no. 18, pp. 2772–2796, 2019, <https://doi.org/10.1002/nag.3001>.
- [31] G. A. L. Jimenez, "Static and Dynamic behaviour of pile supported structures in soft soil," Ph.D. dissertation, Université Grenoble Alpes, 2019.
- [32] J. L. Rangel-Nunez, A. Gomez-Bernal, J. Aguirre-Gonzalez, E. Sordo-Zabay, and E. Ibarra-Razo, "Dynamic Response of Soft Soil Deposits Improved with Rigid Inclusions," presented at the The 14 th World Conference on Earthquake Engineering, Beijing, China, Oct. 2008.
- [33] M. Á. Mánica Malcom, E. Ovando-Shelley, and E. Botero Jaramillo, "Numerical Study of the Seismic Behavior of Rigid Inclusions in Soft Mexico City Clay," *Journal of Earthquake Engineering*, vol. 20, no. 3, pp. 447–475, Apr. 2016, <https://doi.org/10.1080/13632469.2015.1085462>.
- [34] M. Houda, O. Jenck, and F. Emeriault, "Soft soil improvement by rigid inclusions under vertical cyclic loading: numerical back analysis," *European Journal of Environmental and Civil Engineering*, vol. 25, no. 3, pp. 409–428, Feb. 2021, <https://doi.org/10.1080/19648189.2018.1531268>.
- [35] R. M. Patel, B. R. Jayalekshmi, and R. Shivashankar, "Effect of Reinforcement Width on Dynamic response of Basal Geosynthetic-Reinforced Embankment," *Transportation Infrastructure Geotechnology*, vol. 9, no. 4, pp. 516–542, Aug. 2022, <https://doi.org/10.1007/s40515-021-00188-1>.

- [36] S. C. Edem, M. Padmavathi, and V. Padmavathi, "Dynamic Response Analysis of Highway Embankment with Different Fill Material Modifications," *International Journal of Latest Technology in Engineering, Management & Applied Science*, vol. 4, no. 10, 2015.
- [37] R. Hadidi, Y. Moriwaki, J. Barneich, R. Kirby, and M. Mooers, "Seismic Deformation Evaluation Of Lenihan Dam under 1989 Loma Prieta Earthquake," presented at the Tenth U.S. National Conference on Earthquake Engineering, Anchorage, Alaska, Jul. 2014.
- [38] F. R. Olaya and L. M. Cañabi, "Seismic Assessment of a Dam on a Clayey Foundation," in *Proceedings of the 4th International Conference on Performance Based Design in Earthquake Geotechnical Engineering* (Beijing 2022), Cham, 2022, pp. 1959–1967, [https://doi.org/10.1007/978-3-031-11898-2\\_179](https://doi.org/10.1007/978-3-031-11898-2_179).
- [39] A. Osouli and S. Zamiran, "The effect of backfill cohesion on seismic response of cantilever retaining walls using fully dynamic analysis," *Computers and Geotechnics*, vol. 89, pp. 143–152, Sep. 2017, <https://doi.org/10.1016/j.compgeo.2017.04.007>.
- [40] J. Macedo, P. Torres, L. Vergaray, S. Paihua, and C. Arnold, "Dynamic effective stress analysis of a centreline tailings dam under subduction earthquakes," *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, vol. 175, no. 2, pp. 224–246, Apr. 2022, <https://doi.org/10.1680/jgeen.21.00017a>.
- [41] E. Naesgaard, P. Byrne, and A. Amini, "Hysteretic model for non-liquefiable soils (UBCHYST5d)," UBCHYST5d Memo, 2011.
- [42] R. L. Kuhlemeyer and J. Lysmer, "Finite Element Method Accuracy for Wave Propagation Problems," *Journal of the Soil Mechanics and Foundations Division*, vol. 99, no. 5, pp. 421–427, May 1973, <https://doi.org/10.1061/JSFEAQ.0001885>.
- [43] G. Candia and N. Sitar, "Seismic Earth Pressures on Retaining Structures in Cohesive Soils," Berkeley, CA, USA, UCB GT 13-02, Aug. 2013.
- [44] M. B. Darendeli, "Development of a new family of normalized modulus reduction and material damping curves," Ph.D. dissertation, University of Texas, Austin, TX, USA, 2001.
- [45] "FLAC2D." Itasca Consulting Group, Inc., Minneapolis, MN, USA.
- [46] L. Hazzar, M. N. Hussien, and M. Karray, "Influence of vertical loads on lateral response of pile foundations in sands and clays," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 9, no. 2, pp. 291–304, Apr. 2017, <https://doi.org/10.1016/j.jrmge.2016.09.002>.

# Energy Management of an Autonomous Photovoltaic System under Climatic Variations

**Khouloud Njeh**

CEM Laboratory, National Engineering School of Sfax, Tunisia  
njehkhouloud0@gmail.com

**Mohamed Ali Zdiri**

CEM Laboratory, National Engineering School of Sfax, Tunisia  
mohamed-ali.zdiri@enis.tn  
(corresponding author)

**Mohsen Ben Ammar**

CEM Laboratory, National Engineering School of Sfax, Tunisia  
mohsen.benamar@enis.tn

**Abdelhamid Rabhi**

Modeling, Information, and Systems Laboratory, University of Picardie Jules Verne, France  
abdelhamid.rabhi@u-picardie.fr

**Fatma Ben Salem**

CEM Laboratory, National Engineering School of Sfax, Tunisia  
fatma.bensalem@isgis.usf.tn

*Received: 26 September 2022 | Revised: 1 November 2022 | Accepted: 2 November 2022*

## ABSTRACT

Predictable and unpredictable variations are two major causes of renewable energy-related intermittency. Such a drawback could be overcome by applying new energy management strategies, particularly storage systems. This paper investigates the feasibility of introducing a photovoltaic management string, enabling the maintenance of continuous energy supply storage capacity. Accordingly, the PV system is liable to perform at maximum efficiency following the implementation of an MPPT-controlled DC/DC boost converter of the Perturb and Observe (P&O) type. Most of the energy-management approaches target goals lie mainly in optimizing the energy resources' operational and storage capacities. The current MATLAB/Simulink-based simulation study achieved results that turned out to testify well to the advanced design's high performance and efficiency in maintaining the load's energy autonomy while increasing the battery's life span.

**Keywords-**PV-battery; storage system; energy management strategy; P&O; MATLAB/Simulink

## I. INTRODUCTION

Given the increased global demand in energy supply, renewable energy development continues to undergo rapidly frequent changes. In effect, the depletion of fossil fuel reserves has enticed research and development programs to offset the deficit by turning to alternative sources of energy [1-2]. Petroleum-based resources, such as natural gas, coal, and oil, have for long been the unique available energy sources. Still, their uneven and random distribution across territories, along with their limited reserves and supply have spurred international conflict and fierce competitions. Also noteworthy,

is the fact that the fossil-based conventional sources are usually available for only a few decades, while nuclear power generating plants often produce dangerous pollutants, including radioactive waste. To meet the world's growing energy demands, it is therefore necessary to find adequate solutions and diversify the sources of energy. The current research is focused on retrieving and developing new environmentally friendly, renewable, and inexhaustible power sources. Still, energy extraction techniques require further research and development strategies to improve efficiency and reliability. The renewable energy sources fall into three categories:

- Mechanical energy (wind, waterfall, and wave hydrodynamics).
- Electrical energy (photovoltaic (PV) and thermoelectric conversion).
- Heat energy (geothermal and solar thermal).

It is worth noting that most of these power sources rely predominantly on the sun as a major source. Indeed, as a viable source of renewable energy, the sun has been the focus of intensive global research for several decades [3]. It is worth mentioning, in this respect, that the PV panels' drawn energy is not only powerful, but also convertible into electricity thanks to an array of PV installations: isolated "stand-alone" installations, grid connected installations, and hybrid facilities [4]. Regarding the present paper, the modeling and simulation of an autonomous renewable-energy electrical system are treated with respect to the following main axes:

- The theoretical study is focused on the detailed analysis of the boost converter.
- Highlights of an advanced stand-alone PV system with Maximum Power Point Tracking (MPPT) control and the relevant energy management strategy are outlined.
- Analysis and comparison with various storage technologies allow us to identify the load necessary power requirements, the climatic conditions, and the most convenient storage modes fit for an autonomously sustained PV.

The energy management strategy is used to optimize the system's regulatory and supervisory tasks while controlling the charge and discharge status to extend the battery's life. Actually, such a process is deemed critical in an environment where electrical energy consumption is constantly increasing, due to the remarkable growth of industrial development, communication, and transport. In this context, we put forward special battery-stand-equipped indirect storage systems. Similarly, a backup power supply seems essential in case of energy shortage in the batteries' stores [5]. Regarding the present work, our choice is set on applying lithium-ion batteries, given the significant advantages they offer (see Table I). More particularly, they display noticeable energy density, very fast charging process, and a diversity of life cycles. The proposed PV system involves a PV generator, lithium-ion battery, and resistive load. To this end, a special energy management strategy is enrolled to ensure the system's perpetual continuity and improve the battery's life cycle.

TABLE I. BATTERY TECHNOLOGIES AND PERFORMANCE

Technology	Lead-acid (Pb-ac)	Lithium-ion (Li-ion)	Nickel cadmium (Ni-Cd)
Energy density (Wh/kg)	25-45	80-150	20-60
Power density (W/kg)	80-150	500-2000	100-800
Discharge time (h)	0.3-3	0.3-5	0.3-3
Return (%)	A few days	A few months	Less than 1 month
Efficiency (%)	60-98	90-100	60-80
Cycle life	300-1500	>1500	300-1500
Cost (\$/kWh)	100-300	500-1200	200-800

The advanced PV system exhibits not only a high power output, but also an abrupt response to climate as well as load variations. Accordingly, a relevant energy management strategy is considered, whereby the system's continuous functioning and battery's life cycle could be maintained. The strategy is so simple to implement that it requires a low processing amount, and enhances the system's overall storage efficiency and lifetime while allowing to avoid total load disconnect.

## II. SIZING AND MODELING OF PHOTOVOLTAIC CHAIN COMPONENTS

The general diagram of the PV convention chain under examination is depicted in Figure 1.

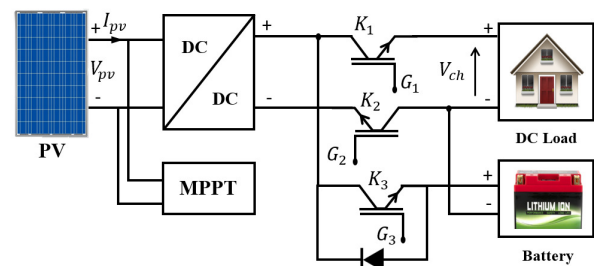


Fig. 1. Block diagram of the PV conversion chain.

### A. PV System Modeling

For an effective PV design to be achieved, accurate measurements and a reliable mathematical approach are necessary. More particularly, the design control rests on the knowledge of the subsystems' various technical characteristics. Basically, the design background requires fulfilling the following steps [6]:

- Selecting the PV modules: the voltage to be delivered, the appropriate technology, and the peak power.
- Determining the storage capacity and battery type.
- Pinpointing the power converters' fit sizes.

### B. Photovoltaic Single Model

The single-diode PV cell corresponding circuit displays two resistors: A shunt resistor  $R_{sh}$  connected in parallel, designating the leakage current in the PN junction, and a series resistor  $R_s$  connected in series, which stands for the connections' various resistances [7]. On substituting the two currents  $I_d$  (diode current) and  $I_{sh}$  with their relevant expressions, the following equation is drawn:

$$I_{PV} = I_{ph} - I_s \times \left( \exp \left[ \frac{V_{PV} + R_s I_{PV}}{V_t} \right] - 1 \right) - \frac{V_{PV} + R_s I_{PV}}{R_{sh}} \quad (1)$$

where  $V_t = k \cdot T / q$ ,  $k$  is the Boltzmann's constant,  $T$  the cell temperature, and  $q$  the electron's charge.

### C. Boost Converter Model

A step-up boost converter is frequently used to transform a low-input voltage into a high output value [8]. It incorporates a



continuous input voltage  $U$ , an inductor  $L$ , a diode  $D$ , and two capacitors  $C_e$  and  $C$ .

#### D. P&O MPPT Control Strategy

Thanks to its simplicity, exclusively requiring current and voltage measurements of the PV array, the Perturb and Observe (P&O) method is considered an effective strategy frequently applied in the MPPT research area. The idea of this approach lies in regularly changing the voltage of the PV array, by comparing the previously delivered power to that actually delivered, following a particular disturbance. Due to this technique, identifying the power peak turns out to be a feasible procedure, regardless of temperature or irradiation fluctuations. Figure 2, depicts the P&O approach corresponding algorithm [9-10]. The associated duty cycle  $d(t)$  is highlighted by the following expression:

$$d(t) = d(t-1) + q \cdot \delta_d \quad (2)$$

where  $q$  is either equal to 0 or  $\pm 1$ , depending on the PV power and voltage variation, and  $\delta_d$  denotes the disturbance value.

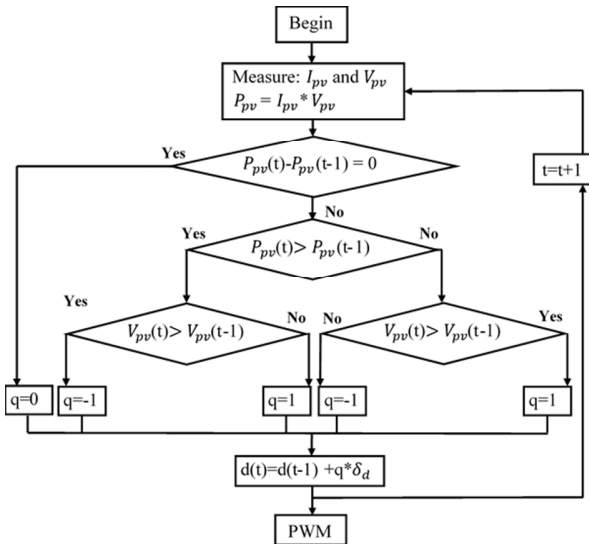


Fig. 2. Flowchart of the P&O MPPT technique.

#### E. Energy Computation Requirement of a Standalone PV System

The first step to follow when sizing a PV installation consists in estimating the energy consumption amount and determining the load-specific profile [11]. Hence, the consumed energy corresponding expression turns out to be:

$$E_{DC} = P_{DC} \times t_j \quad (3)$$

where  $E_{DC}$  denotes the energy consumed on a daily basis,  $t_j$  stands for the daily usage time, and  $P_{DC}$  designates the daily-consumed power.

#### F. Sizing a Photovoltaic Generator

The total number of modules necessary for maintaining the load provision constitutes an initial step in sizing a solar generator [12].

The PV generator's peak power is determined by:

$$P_C = \frac{1000 \cdot E_{Tot}}{G \cdot F_{CG}} \quad (4)$$

where  $E_{Tot}$  denotes the total energy consumed on a daily basis,  $G$  refers to the irradiation, and  $F_{CG}$  designates the global factor correction.

The total number  $N_{Tot}$  of PV modules is computed as:

$$N_{Tot} = \frac{E_{Tot}}{\eta_M \cdot G \cdot S_M \cdot F_{CG}} \quad (5)$$

where  $S_M$  denotes the module area and  $\eta_M$  the module efficiency.

Depending on the load and module peak power,  $N_{tot}$  is determined by:

$$P_C = G_0 \cdot S_M \cdot \eta_s \Rightarrow N_{Tot} = \frac{P_C}{P_{CM}} \quad (6)$$

where  $P_{CM}$  designates the PV module's peak power.

The number of modules in series is:

$$N_{MS} = \frac{V_{bat}}{V_{max}} \quad (7)$$

where  $N_{MS}$  represents the number of PV modules connected in series and  $V_{bat}$  and  $V_{max}$  are the battery voltage and the maximum power voltage, respectively.

The number of parallel modules is determined by:

$$N_{MP} = \frac{N_{Tot}}{N_{MS}} \quad (8)$$

The size of the photovoltaic surface is computed as:

$$Surface = N_{MP} \cdot M_{MS} \cdot S_M \quad (9)$$

The peak PV power is calculated as:

$$P_{pv} = \frac{E_{ch}}{G \cdot (1 - pertes)} \quad (10)$$

The storage system capacity is sized by:

$$C_{batt} = \frac{E_s \cdot N_j}{U_{batt} \cdot \eta_{batt} \cdot P_{db}} \quad (11)$$

where  $P_{db}$  denotes the battery discharge depth and  $N_j$  the number of autonomy days.

### III. ENERGY MANAGEMENT ALGORITHM

An adequate operation of the autonomous PV system entails optimal management of the energy flows exchanged among the system's various components. At this level, it is necessary to introduce our special management system architecture [13-16]. We anticipate that the batteries are fully charged at the start of the PV installation to monitor the entire

system [13]. The energy management strategy is mainly concerned with monitoring the charge state and protecting the battery (minimum charge state  $EDC_{min} = 30\%$  and maximum charge state  $EDC_{max} = 95\%$ ) as depicted in Figure 3. It should be also noted that the PV/battery system involves 4 operating modes, as illustrated in Figure 4, where  $P_{Avd}$  is the difference between the power of the PV generator and the load.

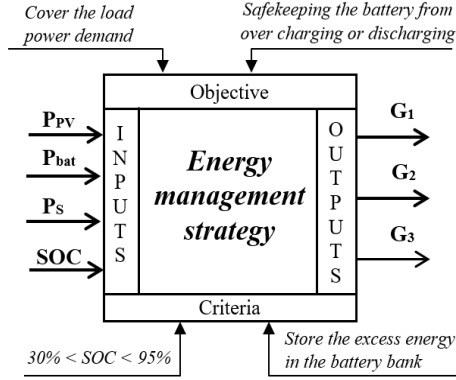


Fig. 3. ULM diagram of the energy management algorithm.

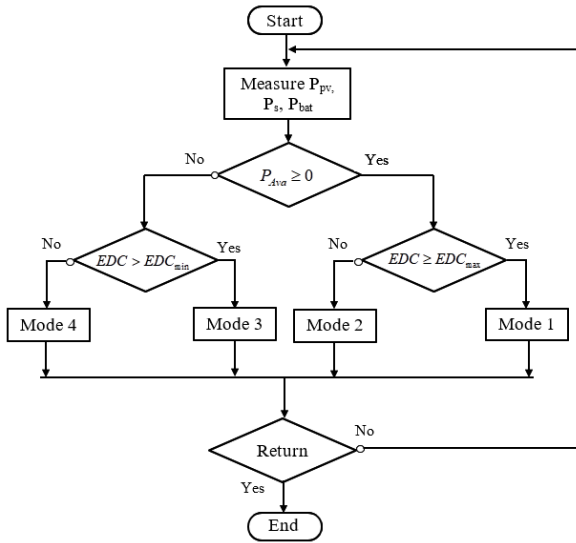


Fig. 4. Operating mode management flowchart.

In case of excessive PV energy:

- MODE 1: The batteries are charged. The PV generator is adequately ready to meet the demand.
- MODE 2: The PV generator is fit to satisfy the load. The produced energy excess is to be stored in the batteries.

In case of lack of PV energy:

- MODE 3: Only the batteries are to supply the charge.
- MODE 4: The batteries are completely discharged with no PV production.

The management algorithm helps supervise the energy exchanges between the installation elements, depending on the two scenarios highlighted in Figure 4:

- Supply of receivers and storage by PV panels if PV energy is sufficient
- Supply of receivers exclusively by the storage system if the PV energy is insufficiency.

#### IV. GLOBAL SYSTEM SIMULATION

Figure 5 shows the system corresponding diagram for a 5kWp PV field, 220V (50Ah) storage lithium-ion battery, and a boost converter, modeled and simulated in MATLAB/Simulink.

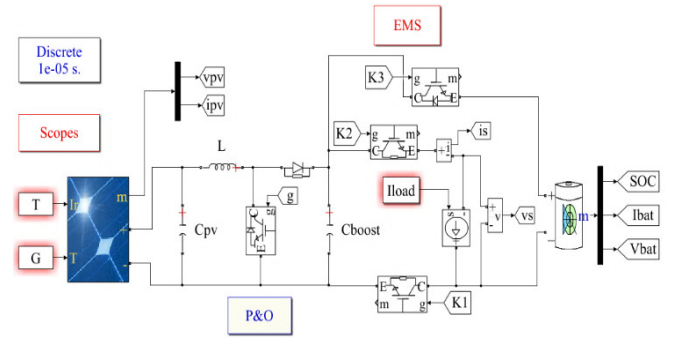


Fig. 5. The global system's diagram in MATLAB/Simulink.

##### A. Simulation Results Considering One-day Profiles

Figure 6 depicts a day-long sunshine and temperature actual representations, as applied for the subsequent simulation results. To analyze the system's operation during the interval [5h, 18.5h], a simulation of the global system was performed over a sunny day. As proof of the management algorithm's effective robustness, Figure 7 outlines two different curves:

- The first curve depicts the maximum power drawn by the  $P_{pv}$  photovoltaic panels.
- The second curve highlights the load power profile  $P_s$ .

The open state (Level 0) or closed state (Level 1) operating moments, relevant to the three power switches ( $K_1$ ,  $K_2$ , and  $K_3$ ), are illustrated in Figure 8. Based on Figure 8:

- Suppose that the PV panels' generated power is greater than that of the load, which is the case for the intervals [(6h-7h24), (7h36-9h), (9h54-10h54), (11h54-16h36), and (17h-18h30)]. In this case, an excess production mode will be adopted at these intervals considering modes 1 and 2. As shown in Figure 8, the power switches  $K_1$  and  $K_2$  are off, and  $K_3$  is on during Mode 1 operation, and  $K_1$ ,  $K_2$ , and  $K_3$  are closed during operation in Mode 2.
- Suppose that the consumption amount is lower than the production rate (i.e. when the panel's generated power proves to be lower than the load required power), as it is the case for the intervals [(7h25-7h35), (9h-9h53), (10h55-1153), (16h37-17h), modes 3 and 4 will then be applied, as

the consumption rate turns out to exceed the production rate. The curve depicted in Figure 8 shows that in mode 3 the switch  $K_1$  is off and the switch  $K_2$  is on, while in mode 4, all the switches  $K_1$ ,  $K_2$ , and  $K_3$  are on.

- Once the system's power consumption and production rates are equal to zero, the system will be disconnected from its load, as is the case during the period span [18:30-6:30].

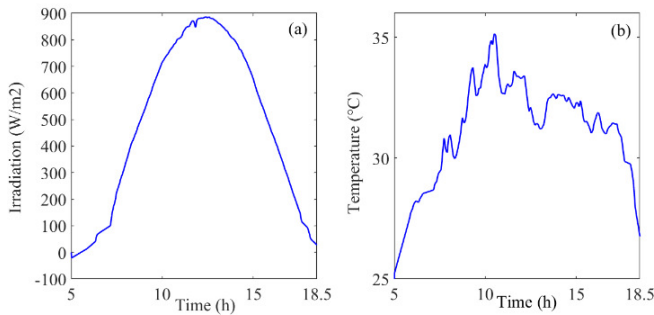


Fig. 6. Sunshine and daytime temperature.

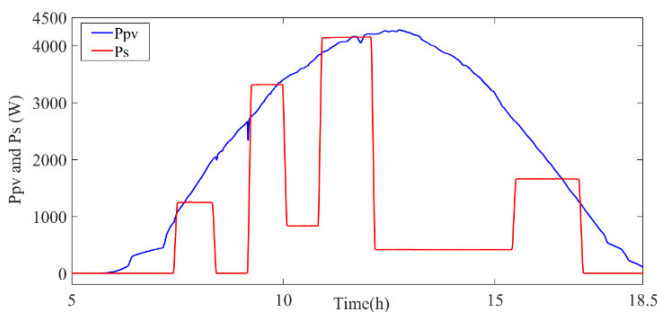


Fig. 7. PV system generated power and load power profile.

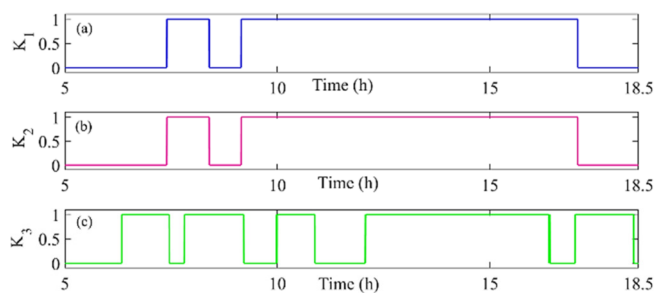


Fig. 8. Operation of the three switches.

The curve appearing on Figure 9 depicts the load as well as the batteries respective voltage  $V_{bat}$ , which is almost equal to the standard reference value of 240V. Regarding the illustrated curve in Figure 10, it highlights well the variations in the battery's charge levels over time. Note that the battery's SOC (State of Charge) appears to vary depending on the power demand divergence from the charge and the PV generator's power supply.

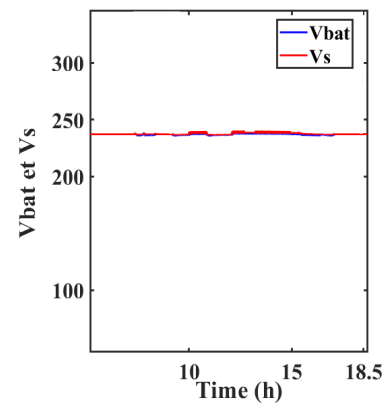


Fig. 9. Load and battery voltage.

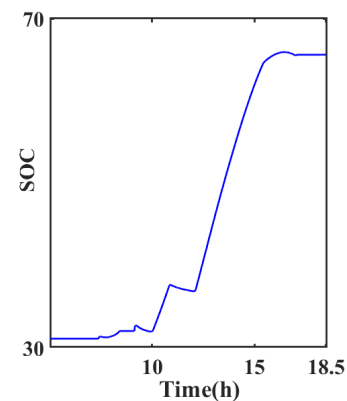


Fig. 10. Battery charge status.

Under the advanced strategy, the continuous bus-provided voltage can be maintained at the desired voltage level of 240V. Accordingly, the DC bus voltage is subject to significant voltage fluctuations on applying this method.

## V. CONCLUSION

The novelty of the current study lies in the investigation of the capacity of an autonomous PV-battery conversion system to maintain the load with continuous energy supply under distinct climatic variations. In this respect, a MATLAB/Simulink study context enabled the execution of a simple energy-management technique with various switching modes. The suggested energy management strategy has been able to provide the load with a continuous energy supply, maintained even in the conditions of solar unavailability. In addition, it helps in extending the battery's lifespan, by dealing with the frequent charging and discharging processes. The simulation study results, achieved by the proposed energy-management strategy under distinct operating modes, verified the scheme's remarkable performance and effectiveness not only in maintaining the PV-battery system's continuity, but also in increasing the battery's lifecycle and durability.

## REFERENCES

- [1] C. R. Kumar. J and M. A. Majid, "Renewable energy for sustainable development in India: current status, future prospects, challenges, employment, and investment opportunities," *Energy, Sustainability and*

- Society*, vol. 10, no. 1, Jan. 2020, Art. no. 2, <https://doi.org/10.1186/s13705-019-0232-1>.
- [2] J. A. Gow and C. D. Manning, "Development of a model for photovoltaic arrays suitable for use in simulation studies of solar energy conversion systems," in *1996 Sixth International Conference on Power Electronics and Variable Speed Drives (Conf. Publ. No. 429)*, Nottingham, UK, Sep. 1996, pp. 69–74, <https://doi.org/10.1049/cp:19960890>.
- [3] F. R. Arroyo M. and L. J. Miguel, "The Role of Renewable Energies for the Sustainable Energy Governance and Environmental Policies for the Mitigation of Climate Change in Ecuador," *Energies*, vol. 13, no. 15, Jan. 2020, Art. no. 3883, <https://doi.org/10.3390/en13153883>.
- [4] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, and R. Gorini, "The role of renewable energy in the global energy transformation," *Energy Strategy Reviews*, vol. 24, pp. 38–50, Apr. 2019, <https://doi.org/10.1016/j.esr.2019.01.006>.
- [5] D. Sera, "Real-time Modelling, Diagnostics and Optimised MPPT for Residential PV Systems," Ph.D. dissertation, Queensland University of Technology, Brisbane, Australia, 2009.
- [6] W. Bensaci, "Modélisation et simulation d'un système photovoltaïque adapté par une commande MPPT," M.S. thesis, Université Kasdi Merbah, Ouargla, Algeria, 2012.
- [7] W. Z. Fam and M. K. Balachander, "Dynamic performance of a DC shunt motor connected to a photovoltaic array," *IEEE Transactions on Energy Conversion*, vol. 3, no. 3, pp. 613–617, Sep. 1988, <https://doi.org/10.1109/60.8076>.
- [8] Z. R. Labidi, H. Schulte, and A. Mami, "A Systematic Controller Design for a Photovoltaic Generator with Boost Converter Using Integral State Feedback Control," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 4030–4036, Apr. 2019, <https://doi.org/10.48084/etasr.2687>.
- [9] S. Hota, M. K. Sahu, and J. M. R. Malla, "A Standalone PV System with a Hybrid P&O MPPT Optimization Technique," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2109–2112, Dec. 2017, <https://doi.org/10.48084/etasr.1374>.
- [10] F. Z. Kebbab, L. Sabah, and H. Nouri, "A Comparative Analysis of MPPT Techniques for Grid Connected PVs," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8228–8235, Apr. 2022, <https://doi.org/10.48084/etasr.4704>.
- [11] S. Abada, "Étude et optimisation d'un générateur photovoltaïque pour la recharge d'une batterie avec un convertisseur Sepic," M.S. thesis, Université Laval Québec, 2011.
- [12] X. Pei, S. Nie, Y. Chen, and Y. Kang, "Open-Circuit Fault Diagnosis and Fault-Tolerant Strategies for Full-Bridge DC–DC Converters," *IEEE Transactions on Power Electronics*, vol. 27, no. 5, pp. 2550–2565, Feb. 2012, <https://doi.org/10.1109/TPEL.2011.2173589>.
- [13] D. Keiner, M. Ram, L. D. S. N. S. Barbosa, D. Bogdanov, and C. Breyer, "Cost optimal self-consumption of PV prosumers with stationary batteries, heat pumps, thermal energy storage and electric vehicles across the world up to 2050," *Solar Energy*, vol. 185, pp. 406–423, Jun. 2019, <https://doi.org/10.1016/j.solener.2019.04.081>.
- [14] M. B. Ammar, M. A. Zdiri, and R. B. Ammar, "Fuzzy Logic Energy Management between Stand-alone PV Systems," *International Journal of Renewable Energy Research (IJRER)*, vol. 11, no. 3, pp. 1238–1249, Sep. 2021.
- [15] M. B. Roberts, A. Bruce, and I. MacGill, "Impact of shared battery energy storage systems on photovoltaic self-consumption and electricity bills in apartment buildings," *Applied Energy*, vol. 245, pp. 78–95, Jul. 2019, <https://doi.org/10.1016/j.apenergy.2019.04.001>.
- [16] J. Goop, E. Nyholm, M. Odenberger, and F. Johnsson, "Impact of electricity market feedback on investments in solar photovoltaic and battery systems in Swedish single-family dwellings," *Renewable Energy*, vol. 163, pp. 1078–1091, Jan. 2021, <https://doi.org/10.1016/j.renene.2020.06.153>.

# Transient Analysis of the Fuzzy Logic-based Speed Control of a Three-phase BLDC Motor

**S. Raja**

Department of Electrical & Electronics Engineering, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, India  
rajas@kanchiuniv.ac.in

**M. Rathinakumar**

Department of Electrical & Electronics Engineering, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, India  
rathinamari@rediffmail.com  
(corresponding author)

Received: 16 October 2022 | Revised: 26 October 2022 | Accepted: 2 November 2022

## ABSTRACT

The energy-efficient motor is a vital requirement for modern industrial, automobile, and home appliance drives. Among the special machines, the Brushless DC Motors (BLDCMs) are more attractive to the application engineer because they offer high power-to-weight ratio, compact construction, do not require consistent maintenance, and have an efficiency margin of 85–90%. This study uses FPGA to create fuzzy logic-based speed control for a three-phase BLDCM with transient time domain characteristics. The fuzzy logic-based controller algorithm is implemented using the FPGA Xilinx Spartan board, which receives the actual speed from a position sensor located inside the BLDCM. It calculates the ratio of the duty cycle of the Pulse Width Modulation (PWM) pulse. It is triggered by the driver system of the BLDCM drive to attain the reference speed. The performance characteristics of a BLDC motor drive such as steady-state error, peak overshoot, speed drop under loaded conditions, and restoration time after loaded conditions were analyzed using MATLAB/Simulink 2014. The hardware configuration setup is validated by the suggested system's simulation response.

**Keywords-**fuzzy logic; Brushless DC Motor (BLDCM); steady state error; peak overshoot; transients; Field Programmable Gate Array (FPGA); Pulse Width Modulation (PWM)

## I. INTRODUCTION

Energy-efficient industrial drives are becoming more and more necessary in emerging industrial and residential applications. One such advancement is in BLDCMs [1-3], which have much more efficiency than the single-phase induction motor, high torque-to-power ratio, require less maintenance, and have a fast reaction to commands. BLDC motors need an electronic commutator, which is a major shortcoming. Many researchers describe various control methods applied to BLDCM drive applications. Control of the BLDCM is commonly employed with PWM control techniques. This method is simple and efficient. Its main drawback is the slow response, which is always coupled with hysteresis current control. Hysteresis-current control yields a faster speed response than PWM control. In order to achieve well-regulated speed-to torque control, the Proportional (P) or Proportional Integral (PI) actions were introduced. A PI controller offers low deviation from the desired set point. A wide range of control actions can be accomplished [4-9].

Linear control theory is highly complex and requires extensive knowledge to develop efficient controller designs. Digital signal processors, ultra-high speed integrated chips, and FPGAs are employed to design the digital control algorithm, adding to the complexity of the overall system design [10]. When building an energy-efficient industrial drive application, several considerations must be made. Some of the key problems are the compatibility of the selected hardware technology, the efficiency of the application, power utilization, and the area that must suit the requirements of the algorithm to be implemented. FPGA technology has lower power consumption and a high level of parallelism in nature, which offers higher throughput compared to DSP [11]. Hence, the FPGA Xilinx Spartan 3E is used in this paper to implement a PWM module, a speed sensing module, and a fuzzy logic control algorithm for the speed control of a BLDCM drive. The above module program is written in VHSIC (Very High-Speed Integration Circuit) Hardware Description Language (VHDL).



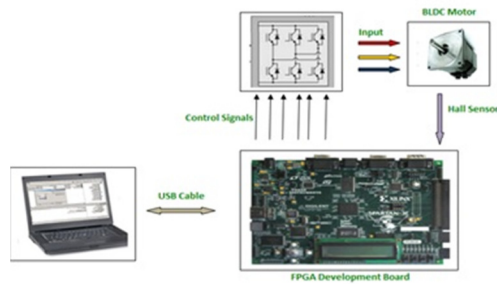


Fig. 1. Proposed block diagram of the 3-phase BLDCM control.

## II. SENSOR BASED CONTROLLER FOR BLDC MOTOR

The speed of the BLDC motor is measured by using a Hall sensor [11-12]. It is calculated from the output of the Hall sensor, which senses information about the rotor position of the BLDC motor. The error between the reference and the actual speed of the BLDC motor is given to the fuzzy logic speed controller, which generates the command or reference current. The hysteresis current regulator provides the firing currents to the inverter, whose output is generated based on the difference between the actual phase currents and the reference currents. In this paper, the three position sensors A1, A2, and A3 are installed at the non-rotating end of the motor. They measure the rotor position every 120 degrees. One of the Hall sensors' condition changes. The electrical rotation of a BLDC motor does not imply the motor's single mechanical rotation, which is decided by the rotor pairs [9]. The commutation instants with the PWM signal are driven to the 3-phase inverter of the BLDCM. This helps to create a constant torque and to reach the desired machine speed [13, 17].

## III. CONCEPT OF THE FUZZY CONTROLLER

The fuzzy control algorithm has 2 inputs and 1 output. One input is the deviation of the set point speed with the actual speed and the other input is the derivative of error. The output of the fuzzy system is the input of the hysteresis control block. The hysteresis control takes care of the limit of the load current within a limited range [7]. The internal process of the fuzzy logic system is shown in Figure 3 [18]. The membership of the error input and its change spread in  $(-13, 13)$  and  $(-5, 5)$ , respectively [19]. This is because the fuzzy logic controller depends on model behavior knowledge of the data. The membership value for the input and output function of the error values in earlier works with PI controller falls in the zone of  $(-13, 13)$ . Similarly, the change in error value falls in the zone of  $(-5, 5)$  in my earlier works regarding the PI controller. This is the main reason the membership function and the change in error values in the present fuzzy logic controller have been chosen as  $(-13, 13)$  and  $(-5, 5)$ , respectively. The considered membership function is a triangular membership function. In this paper, a triangular membership function of a fuzzy variable is used as shown in Figure 2. The membership function is divided into 7 categories [8]. The next form finds the new duty cycle values from the fuzzy rule base. The 49 rules are formed based on the previous experience and the knowledge about the behavior of the system using the "IF THEN" statements.

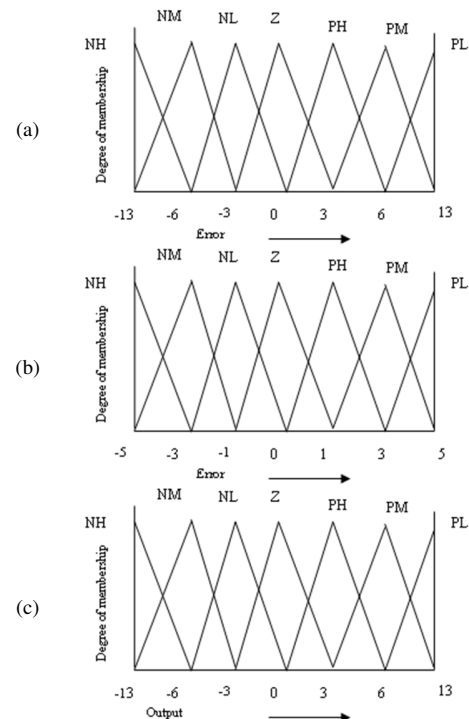


Fig. 2. Evaluated values of the membership function of the fuzzy control algorithm.

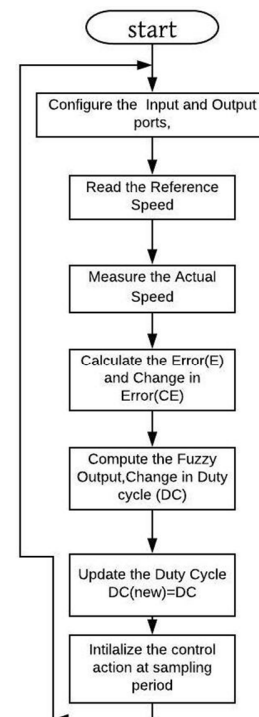


Fig. 3. Flow chart of the speed regulation control algorithm.

This new duty cycle output is in the fuzzy variable, which is to be converted into a real-time numerical value by using the defuzzification process. In this work, the centroid defuzzification method is used. The switches receive a variable

duty cycle of the PWM signal. The control regulates the chopping of the provided voltage to the motor thereby controlling the motor speed [20].

#### IV. SIMULATION RESULTS

Simulations were conducted in MATLAB/Simulink. The fuzzy-based control algorithms are tested against different speed and load conditions. The simulation results are shown in Figure 5.

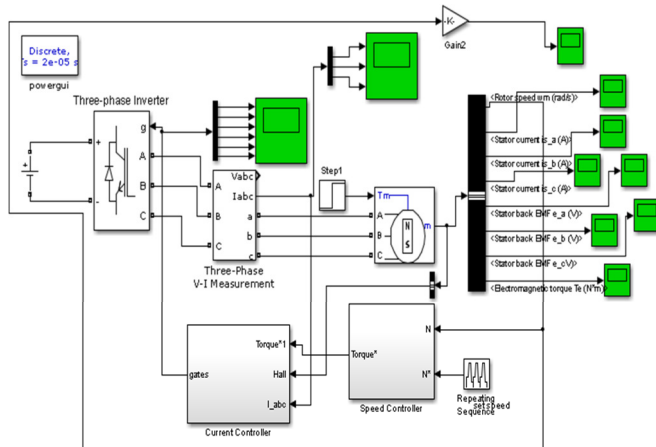


Fig. 4. Three-phase BLDCM fuzzy control algorithm modeled in Simulink 2014.

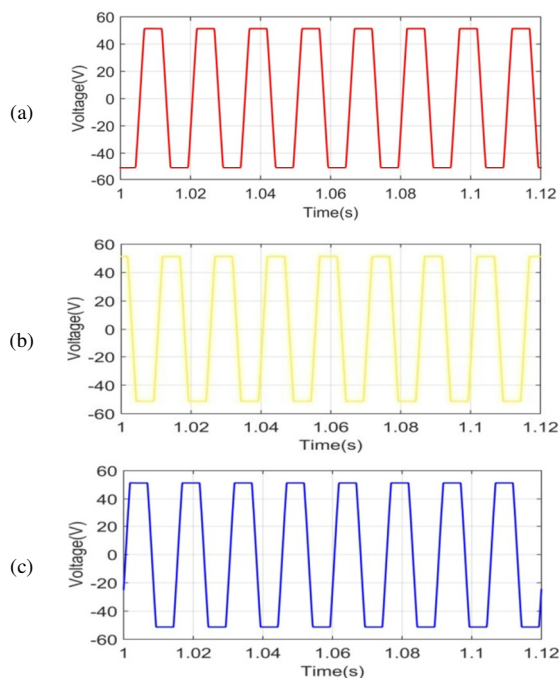


Fig. 5. (a) R phase BEMF waveform, (b) Y phase BEMF waveform, (c) B phase BEMF waveform at the desired rotor speed of 2000 rpm.

The peak overshoot in the speed response of a BLDCM drive using the fuzzy control algorithm is depicted in Figure 7. It can be understood that the motor speed settled down at the

reference value with less oscillations and lower magnitude and the speed variable is smooth, for the scheme of fuzzy logic control. From Figure 7, a steady state error with a value of 0.025 % occurs in the BLDC motor, when it is operated during the speed of 2000 rpm.

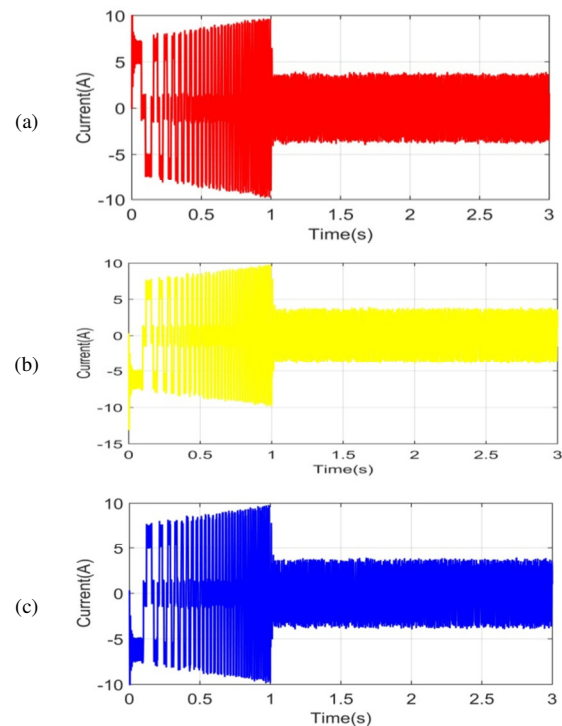


Fig. 6. (a) R phase stator current, (b) Y phase stator current waveform, (c) B phase stator current at the desired rotor speed of 2000 rpm.

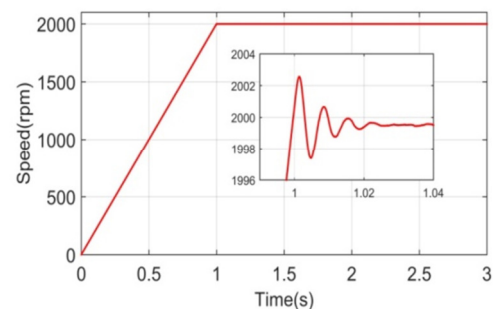


Fig. 7. Simulation rotor speed response when the BLDCM runs at the desired rotor speed of 2000 rpm.

TABLE I. BLDCM SPECIFICATIONS

Specifications	Parameters	Unit
Rated power	60	W
Rated voltage	24	V
Rated torque	0.15	Nm
Rated speed	3000	rpm

#### V. HARDWARE SETUP

The suggested hardware includes a 3-phase BLDC motor, a 3-phase inverter that uses MOSFETs, an opto isolator, driving

circuits, etc. (Figure 9). The Xilinx Spartan FPGA produced the PWM signals. These signals are driven by an IC buffer and the output from the buffer is fed to the opto isolator to isolate the PWM signals from the high voltage section.

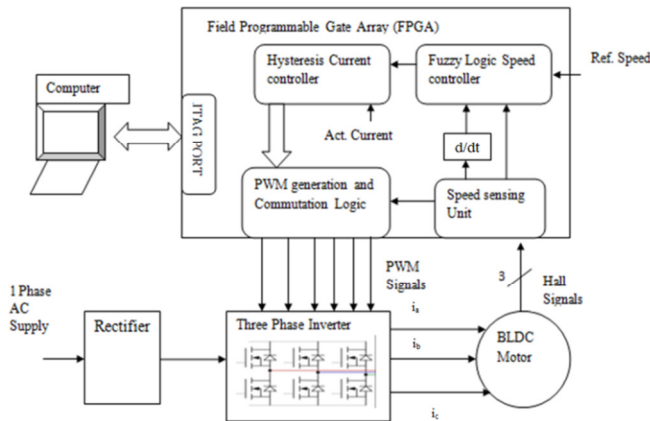


Fig. 8. Detailed block diagram of the fuzzy logic assisted drive – real time implementation by Xilinx FPGA.

The properly installed Hall effect sensor on the 3-phase BLDC motor transmits a signal to the FPGA board, which evaluates the speed and gets the user-desired speed. It compares the reference and actual speeds and calculates the difference. After receiving the speed error, the fuzzy logic controller calculates the PWM signal's duty cycle to produce the desired speed. The PWM signal is combined with the commutation logic and is fed to the driver system of the motor. Electronic commutation is needed for motor rotation with the help of the commutation cycle. The FPGA board creates the PWM and the commutation signal depicted in Figure 10.

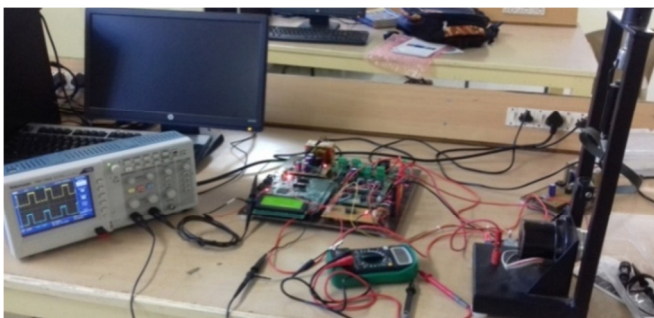


Fig. 9. Real time implementation of the suggest system by Xilinx FPGA.

The Xilinx Integrated System Environment 11.2 is used to edit, compile, synthesize, and download the code onto the Xilinx Spartan 3E board. VHDL language was used to develop the control algorithms like fuzzy logic, hysteresis logic, PWM logic, and speed estimation modules. The switching pulses for the 6 MOSFETs of the driver section and the R-Y, Y-B, and B-R phase voltages of the 3-phase of the BLDC motor were recorded with a digital scope. Channel 1 of the oscilloscope shows the switching pulse for the upper switch of the first leg of the VSI (Figure 10(a)), Channel 2 (Figure 10(a)), shows the PWM output of the fuzzy logic controller, which is fed the

switching pulse for the lower switch of the first leg of the VSI. The R-Y, Y-B, and B-R phase voltages of the 3-phase BLDCM running under 2000 rpm were recorded.

TABLE II. SUMMARY OF THE USED COMPONENTS IMPLEMENTED WITH THE XILINX FPGA

Logic components (in numbers)	Available	Utilized components	% of used components
Xilinx slices	768	297	38
Flip flops	1536	177	11
LUT	1536	472	30

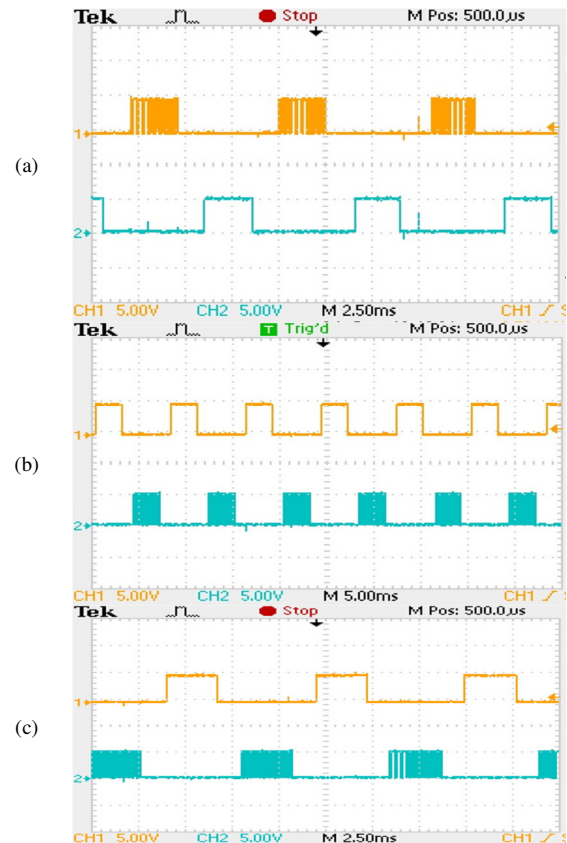


Fig. 10. Switching pulses at the (a) first leg of the inverter, (b) second leg of the inverter, (c) third leg of the inverter when BLDCM runs at 2000 rpm.

## VI. TRANSIENT RESPONSE ANALYSIS

The speed drops at a step change in load by 0.5% when 80% of the maximum load is applied to the fuzzy controller-based BLDC motor drive. The speed drops at a step change in load by 0.18% when 20% of the maximum load is applied to a fuzzy controller-based BLDCM drive. Figure 10 demonstrates that a restoration time of 0.024s is required after a sudden change in the load to reach the set speed when 80% of the rated load is applied. Figure 12 demonstrates that a restoration time of 0.02s is required after a sudden change in the load to reach the set speed when 20% of the rated load is applied. The transient performance of the BLDCM drive by a fuzzy algorithm is measured and recorded for various speeds as shown in Table III. The performance characteristic curves are plotted from the simulation results. Figure 15 shows that the

speed drops at a step change in load by 0.18 to 0.5% when the applied load varies from 20% to 80% of the maximum load for the fuzzy controller-based BLDCM drive. Hence, when the load torque increases, the speed drop also increases. The restoration time required also increases from 0.02 to 0.025s after a sudden change in the load to reach the set speed when the increased minimum load reaches the full load, as shown in Figure 16.

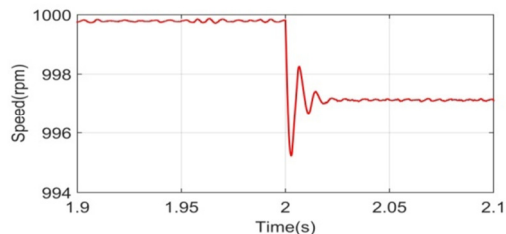


Fig. 11. Rotor speed response under 1000 rpm for the 3-phase BLDCM when 80% of the rated load is applied at 2s.

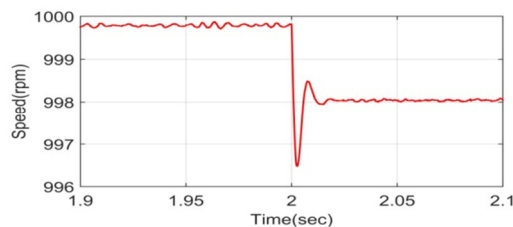


Fig. 12. Rotor speed response under 1000 rpm for the 3-phase BLDCM when 60% of the rated load is applied at 2s.

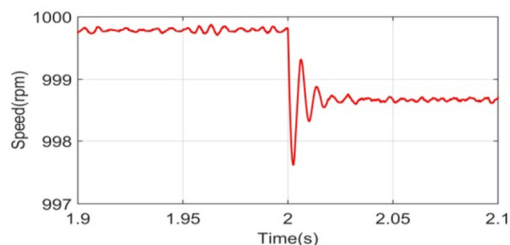


Fig. 13. Rotor speed response under 1000 rpm for the 3-phase BLDCM when 40% of the rated load is applied at 2s.

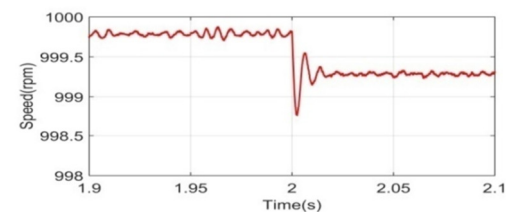


Fig. 14. Rotor speed response under 1000 rpm for the 3-phase BLDCM when 20% of the rated load is applied at 2s.

The cross comparison of Figures 11-16 clearly shows the enhanced control of the fuzzy logic controller in speed and load variations of the BLDCM. While using the fuzzy logic controller and the speed response in steady state condition, the actual speed of the BLDCM is almost the same as the set point speed. The overshoots present in the speed response are lesser and the load disturbance rejection ability is good. The fuzzy

logic controller gives a better transient response according to the applied set speed and load variation conditions. The proposed fuzzy logic-based controller of the BLDCM drive is suited for applications where the adjustable speed and load variations frequently occur.

TABLE III. TRANSIENT ANALYSIS OF FUZZY CONTROL ALGORITHMS IMPLEMENTED BY XILINX FPGA

Percentage of load applied at rated load	Speed drop (rpm)	Restore time (s)	Steady state error (%)
20	998.2	0.02	0.7
40	997.5	0.02	1.3
60	996.5	0.023	1.9
80	995	0.024	2.9

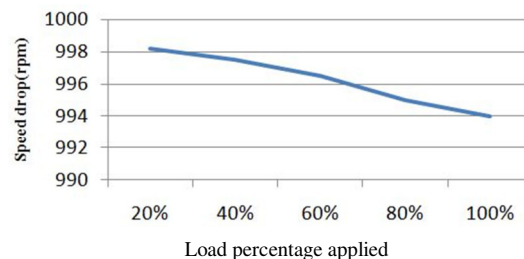


Fig. 15. Variation of load torque to the motor and the speed drop.

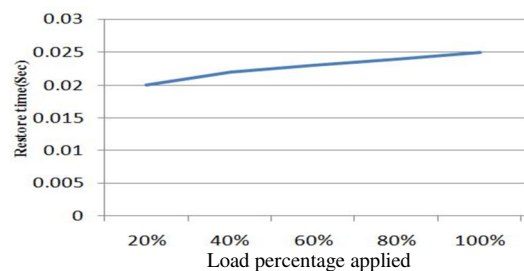


Fig. 16. Variation of load torque to the motor and restoration time to recover the desired speed.

## VII. CONCLUSION

The findings demonstrate that better characteristics of transient time domain (in comparison with the PI controllers) were obtained when a digital fuzzy control algorithm is implemented on an FPGA in BLDCM speed control [7]. By operating the BLDCM under various load situations, a load test was carried out to confirm the efficacy of the fuzzy control. The performance of BLDCM drive will be enhanced even further by the use of the adaptive fuzzy- neural mechanism.

## ACKNOWLEDGMENT

The authors would like to express their gratitude to the Management of SCSVMV for providing the lab facilities required to complete the research work successfully.

## REFERENCES

- [1] A. Emadi, *Handbook of Automotive Power Electronics and Motor Drives*. Boca Raton, FL, USA: CRC Press, 2005.
- [2] C.-L. Huang and S.-C. Yang, "Torque Ripple Reduction for BLDC Permanent Magnet Motor Drive using DC-link Voltage and Current Modulation," in *2021 IEEE International Future Energy Electronics*



- Conference (IFEEC), Taipei, Taiwan, Aug. 2021, <https://doi.org/10.1109/IFEEC53238.2021.9661795>.
- [3] X. Wang *et al.*, "Stray Flux-Based Rotation Angle Measurement for Bearing Fault Diagnosis in Variable-Speed BLDC Motors," *IEEE Transactions on Energy Conversion*, vol. 36, no. 4, pp. 3156–3166, Sep. 2021, <https://doi.org/10.1109/TEC.2021.3079444>.
  - [4] I. Elomary, A. Ahmed, and I. Lhassane, "Optimization Design of the BLDC Motor Using Backtracking Search Algorithm," *International Review of Electrical Engineering (IREE)*, vol. 16, no. 2, pp. 167–173, Apr. 2021, <https://doi.org/10.15866/iree.v16i2.17411>.
  - [5] S. B. Ozturk, W. C. Alexander, and H. A. Toliyat, "Direct Torque Control of Four-Switch Brushless DC Motor With Non-Sinusoidal Back EMF," *IEEE Transactions on Power Electronics*, vol. 25, no. 2, pp. 263–271, Oct. 2010, <https://doi.org/10.1109/TPEL.2009.2028888>.
  - [6] S. B. Ozturk, W. C. Alexander, and H. A. Toliyat, "Direct Torque Control of Four-Switch Brushless DC Motor With Non-Sinusoidal Back EMF," *IEEE Transactions on Power Electronics*, vol. 25, no. 2, pp. 263–271, Oct. 2010, <https://doi.org/10.1109/TPEL.2009.2028888>.
  - [7] M. Perotti and F. Fiori, "A Closed Loop Delay Compensation Technique to Mitigate the Common Mode Conducted Emissions of Bipolar PWM Switched Circuits," *IEEE Transactions on Power Electronics*, vol. 36, no. 5, pp. 5450–5459, Feb. 2021, <https://doi.org/10.1109/TPEL.2020.3031349>.
  - [8] M. Karthik *et al.*, "Evaluation of electromagnetic intrusion in brushless DC motor drive for electric vehicle applications with manifestation of mitigating the electromagnetic interference," *International Journal of Ambient Energy*, Oct. 2020, <https://doi.org/10.1080/01430750.2020.1839546>.
  - [9] T. C. Siong, B. Ismail, S. F. Siraj, M. F. Mohammed, and M. F. N. Tajuddin, "Implementation of Fuzzy Logic controller for permanent magnet brushless DC motor drives," in *2010 IEEE International Conference on Power and Energy*, Kuala Lumpur, Malaysia, Aug. 2010, pp. 462–467, <https://doi.org/10.1109/PECON.2010.5697627>.
  - [10] E. Monmasson and M. N. Cirstea, "FPGA Design Methodology for Industrial Control Systems—A Review," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 4, pp. 1824–1842, Dec. 2007, <https://doi.org/10.1109/TIE.2007.898281>.
  - [11] S. Raja, D. M. Rathinakumar, and A. S. Viswanathan, "Performance Analysis of Digital Controller for BLDC Motor Drive System," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 6, no. 6, pp. 10825–10831, 2007, <https://doi.org/10.15680/IJIRSET.2017.0606129>.
  - [12] V. G. K and J. D. Shree, "Implementation of a High Power Quality BLDC Motor Drive Using Bridgeless DC to DC Converter with Fuzzy Logic Controller," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9178–9185, Oct. 2022, <https://doi.org/10.48084/etasr.5213>.
  - [13] D. B. Minh, V. D. Quoc, and P. N. Huy, "Efficiency Improvement of Permanent Magnet BLDC Motors for Electric Vehicles," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7615–7618, Oct. 2021, <https://doi.org/10.48084/etasr.4367>.
  - [14] A. Prakash and C. Naveen, "Combined strategy for tuning sensor-less brushless DC motor using SEPIC converter to reduce torque ripple," *ISA Transactions*, Jul. 2022, <https://doi.org/10.1016/j.isatra.2022.06.045>.
  - [15] J.-W. Reu, J. Hur, B.-W. Kim, and G.-H. Kang, "Vibration reduction of IPM type BLDC motor using negative third harmonic elimination method of air-gap flux density," in *2010 IEEE Energy Conversion Congress and Exposition*, Atlanta, GA, USA, Sep. 2010, pp. 1745–1752, <https://doi.org/10.1109/ECCE.2010.5618125>.
  - [16] M. Hussain, A. Ulasay, H. S. Zad, A. Khattak, S. Nisar, and K. Imran, "Design and Analysis of a Dual Rotor Multiphase Brushless DC Motor for its Application in Electric Vehicles," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7846–7852, Dec. 2021, <https://doi.org/10.48084/etasr.4345>.
  - [17] P. Ubare and D. N. Sonawane, "Performance Assessment of the BLDC Motor in EV Drives using Nonlinear Model Predictive Control," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8901–8909, Aug. 2022, <https://doi.org/10.48084/etasr.4976>.
  - [18] R. O. Reddy, S. Kautish, V. P. Reddy, N. S. Yadav, M. M. Alanazi, and A. W. Mohamed, "Effects of Integrated Fuzzy Logic PID Controller on Satellite Antenna Tracking System," *Computational Intelligence and Neuroscience*, vol. 2022, Mar. 2022, Art. no. e7417298, <https://doi.org/10.1155/2022/7417298>.
  - [19] T. Muthamizhan, B. Shivaji, and M. Aijaz, "Fuzzy logic controller based Multilevel Inverters integrated Speed Control of Induction Motors," in *2022 3rd International Conference for Emerging Technology (INCET)*, Belgaum, India, Feb. 2022, <https://doi.org/10.1109/INCET54531.2022.9824251>.
  - [20] F. Belgacem, M. Mostefai, M. Yahia, and A. Belgacem, "Optimization of Photovoltaic Water Pumping System Based on BLDC Motor for Agricultural Irrigation with Different MPPT Methods," *Periodica Polytechnica Electrical Engineering and Computer Science*, vol. 66, no. 4, pp. 315–324, Oct. 2022, <https://doi.org/10.3311/PPee.20140>.



# Cutting Parameter Optimization based on Online Temperature Measurements

**Abdelillah Djamal Kara Ali**

Department of Mechanical Engineering, University of Tlemcen, Algeria  
dj\_kara\_ali@yahoo.fr  
(corresponding author)

**Nasreddine Benhadji Serradj**

Department of Mechanical Engineering, University of Tlemcen, Algeria  
kawther4499@yahoo.fr

**Mohamed El Amine Ghernaout**

Department of Mechanical Engineering, University of Tlemcen, Algeria  
e\_amine2001@yahoo.fr

Received: 20 September 2022 | Revised: 26 October 2022 | Accepted: 27 October 2022

## ABSTRACT

The deformation of metallic materials during the machining operation requires a significant amount of energy. During the chip formation process and due to the plastic deformation of the metal and the friction along the tool-part interface, the thermal loads generated are strongly impacted by the cutting factors. Thus, the choice of optimized cutting conditions is essential to control the quality of the work required. The aim of the present experimental study is to optimize the cutting parameters using temperature measurements. The average temperature of the cutting tool is studied using a FLIR A325sc type infrared camera. Optimal cutting parameters for each performance metric were obtained using the Taguchi techniques.

**Keywords-machining; cutting conditions; optimization; thermography; temperature measurement**

## I. INTRODUCTION

The choice of cutting conditions is a very important factor in machining. With the increasingly rapid development of machine tools and the use of new high-performance materials for the tool-part couple, cutting temperature is an essential concern in machining. To be highly competitive, the objectives sought by the mechanical manufacturing industries are to produce high quality products with low cost and time constraints. To achieve them, one of the solutions is to optimize machining parameters such as cutting speed, depth of cut and feed. The effects of cutting parameters (cutting speed, depth of cut, feed per revolution,) on surface roughness, tool wear, cutting forces related to turning operations have been studied in depth. Authors in [1] used the Response Surface Method (RSM) with a factorial design to predict surface roughness when machining high strength materials. Authors in [2] studied the machinability of hard steel in turning by the response surface method. Authors in [3] used a multiplicative model to predict surface roughness and tool vibration. The context of the current work is based on the use of full factorial type experimental designs based on the experimental characterization of the temperature measurement for the

optimization of the cutting conditions during the turning operation. Many techniques have been developed during the recent years that address this type of problem. Authors in [4] propose a general classification of the different techniques for modeling and optimizing input-output relationships in metal cutting processes. They are classified into two methods, classical optimization techniques and unconventional techniques. The current work is oriented on the study of the heat flow in the zone of the cut represented by the measurement of the temperature. Various techniques of temperature measurement have been studied and evaluated [5-8]. The optical infrared method has been used in [9-10].

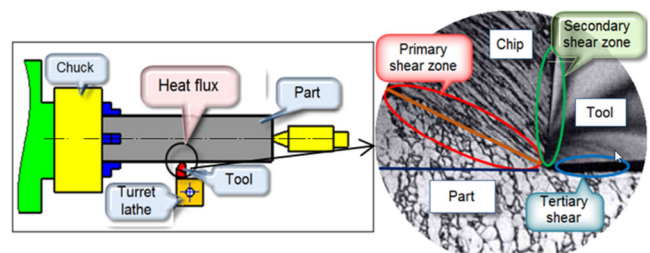


Fig. 1. Sources of heat generation in machining.

At the heart of the process, two main phenomena are in interaction:

- A very strong plastic deformation in the shear zones.
- The friction of the chip on the rake face of the tool.

One may also add the friction of the clearance face on the newly machined surface. The energy generated is then largely transformed into a significant heat flux which modifies and deteriorates the machined surface and the cutting tool shown in Figure 1.

In this paper, an experimental measurement methodology is proposed and the results are presented using an acquisition instrumentation, equipped with a single infrared camera with advantages such as contactless measurement, measurement without or with very little disturbance between the surface of the studied object and its surrounding environment, the capacity for real-time measurement, wide range of operating temperatures, the ability to adapt to any type of material, and a unit processing of the measured values. The correlation between the cutting conditions and the temperature appears on the flowchart represented in Figure 2.

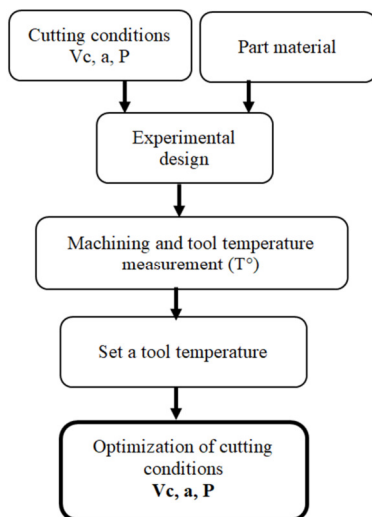


Fig. 2. Correlation between cutting conditions and temperature.

In industry and in order to make the best choices, very often one empirical method (Taylor's law, Kronenberg model, etc.) is used to calculate and optimize the cutting parameters before machining. Knowing that the low cumulative gains generate significant profits, we propose an approach that allows refining the manufacturing conditions after following the experimental tests.

## II. MATERIALS AND METHODS

The infrared camera used is the Flir A325sc, coupled with the FLIR Research software. The inserts are made of tungsten carbide with a beak radius  $r = 0.4\text{mm}$  (Figure 3). The specimens are made of S185 steel, a material widely used in industry. For the machining we used a conventional parallel lathe made by TOS TRENCIN type (Figure 4).



Fig. 3. Thermal camera FLIR A325sc and tool holder with the used insert.



Fig. 4. Machining station with tool, specimen, and camera.

The conducted experimental tests and the calibration of the camera are widely discussed in [11]. Before using the results, the data are retrieved from the camera software (FLIR Quick Plot) for viewing the machining sequences. Figure 5 represents the acquisition chain.



Fig. 5. Chain of acquisition and processing of thermal data.

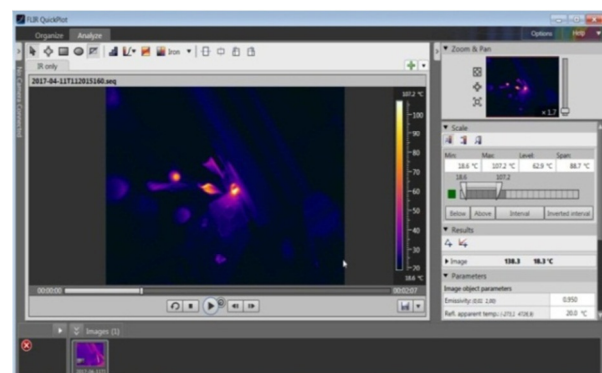


Fig. 6. Software interface.

The recovery of the temperatures is made after tracing the curve represented by Figure 7. The curve is divided into three zones:

- Zone 1 represents the machining initiation.

- Zone 2 determines the stabilization part of the section, from the latter we recover the temperatures for analysis.
- Zone 3 represents the tool release and the end of the turning operation.

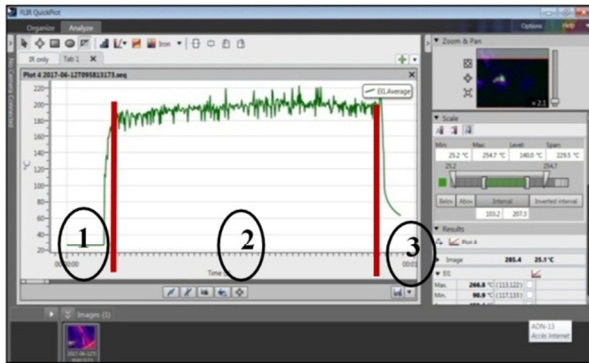


Fig. 7. Software interface and curve areas studied.

### III. TEMPERATURE PREDICTION MODEL

The relationship between the inputs which are the machining parameters (cutting speed, feed, depth of cut) as independent variables vary according to an  $N^3$  type design of experiments. In a  $4^3$  type experiment, each parameter has 4 levels, which gives us 64 experiments (full plan). We carried out 64 machining operations, each time using a new cutting edge. The expanded linear model of the type second-order equation is:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{11}X_1^2 + b_{22}X_2^2 + b_{33}X_3^2 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 \quad (1)$$

where  $Y = T$  is the estimated response,  $X_1, X_2, X_3$  respectively represent the control factors (cutting speed, feed, and depth of cut), and  $b_0, \dots, b_{23}$  are the coefficients of the related factors. Constants and exponents are determined by the least squares method. After careful reading of the sequences recorded by the camera, we determined the temperature of the tip of the tool according to the cutting parameters as shown in Table I.

TABLE I. CUTTING PARAMETERS

Parameter	Code	Levels			
		1	2	3	4
Cutting speed $V_c$ (m/min)	$X_1$	78	112	157	220
Feed rate $a$ (mm/rev)	$X_2$	0.08	0.11	0.14	0.20
Depth of cut $P$ (mm)	$X_3$	0.25	0.5	0.75	1
Turning conditions	Machining without lubrication				

The analysis of variance in Table II shows that the interaction feed and depth of cut clearly influence the temperature ( $P$ -value = 0.04). After processing, we obtain the prediction equation (2):

$$T^{\circ} \text{ tool} = 89.6 + 0.054 V_c - 638 a - 19.2 P - 0.000215 V_c^2 + 1640 a^2 + 32.3 P^2 + 0.572 V_c a + 0.036 V_c P + 368 a P \quad (2)$$

The residual distribution analysis is based on a powerful test for the detection of non-normality. The normal probability plot is shown in Figure 8. It is clear that the residuals lie close

to the reference line, which implies that the errors are normally distributed.

TABLE II. ANALYSIS OF VARIANCE

Source	DF	Adj SS	Adj MS	F-value	P-value
Regression	9	34124.1	3791.57	12.63	0.000
$V_c$	1	8.0	7.95	0.03	0.871
$a$	1	677.8	677.83	2.26	0.139
$P$	1	38.2	38.18	0.13	0.723
$V_c \times V_c$	1	14.4	14.39	0.05	0.828
$a \times a$	1	438.5	438.46	1.46	0.232
$P \times P$	1	260.5	260.46	0.87	0.356
$V_c \times a$	1	116.4	116.40	0.39	0.536
$V_c \times P$	1	17.9	17.85	0.06	0.808
$a \times P$	1	1333.7	1333.65	4.44	0.040
Error	54	16212.6	300.23		
Total	63	50336.7			

TABLE III. COEFFICIENTS OF EQUATION (2)

Term	Coef	SE Coef	T-value	P-value	VIF
Constant	89.6	43.3	2.07	0.043	
$V_c$	0.054	0.335	0.16	0.871	67.55
$a$	-638	424	-1.50	0.139	75.61
$P$	-19.2	53.8	-0.36	0.723	48.28
$V_c \times V_c$	-0.000215	0.000982	-0.22	0.828	53.63
$a \times a$	1640	1357	1.21	0.232	63.50
$P \times P$	32.3	34.7	0.93	0.356	32.25
$V_c \times a$	0.572	0.918	0.62	0.536	17.03
$V_c \times P$	0.036	0.146	0.24	0.808	13.11
$a \times P$	368	175	2.11	0.040	14.92

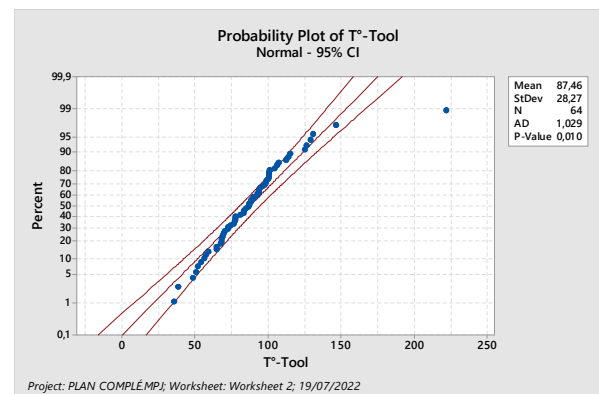


Fig. 8. Normality line of the distribution of residuals.

### IV. RESULTS AND DISCUSSION

The test results were divided according to the method of measurement, either in zone 2 or zone 3. The results obtained during the first method (zone 2) have been discussed in [12]. We present the results obtained in zone 3 where the temperature measurement is made just after the release of the tool. The comparison between the experimental values and the predicted values for the obtained temperature model, is represented graphically in Figure 9 for the complete design of the experiment. The latter shows that the experimental and the estimated values are close with an accuracy of 88.87%. These results prove the accuracy of the model. Furthermore, they confirm that the model can be effectively used to predict the machining temperature with a 95% confidence interval.

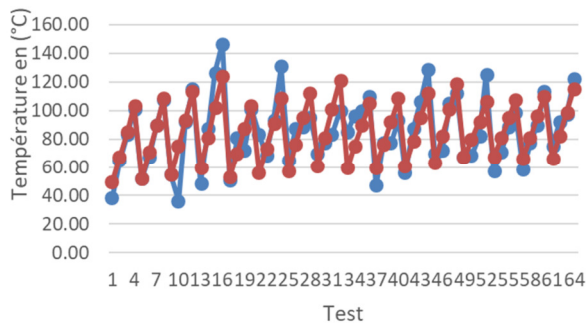


Fig. 9. Experimental and predicted curves.

Figure 10 shows the different interactions. The interaction  $V_c \times p$  (between the cutting speed and the depth of cut) is noticed to have an imposing effect. On the other hand, the interaction  $a \times p$  (between the feed and the depth of cut) and the interaction  $V_c \times a$  (between cutting speed and feed) are less significant. Most of the energy released is spent in deforming the chips both cutting and plastically. Part of this energy is transformed into heat by friction at the two interfaces (tool-chip and tool-workpiece). The generation of heat during machining causes the temperature to rise in the process area. It turns out that the cutting speed is the imposing factor on the cutting temperature. To visualize the influence of the cutting parameters on the temperature ( $T^\circ$ ), the response surfaces (2D) are presented below.

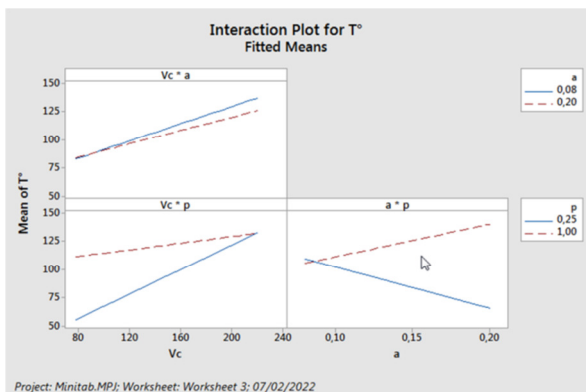


Fig. 10. Interactions of cutting conditions.

The isotherm in Figure 11 shows that the cutting temperature increases with the feed over the interval of 0.08mm/rev and 0.14mm/rev, with a drop in temperature over the interval 0.14mm/rev and 0.2mm/rev. The thermogram in Figure 12 shows the concentration of heat on the tool tip and chip fragments. The isotherm in Figure 13 shows that the cutting temperature increases with increasing cutting speed and depth of cut. The thermogram in Figure 14 shows the heat distribution on the cut face. The isotherm in Figure 15 shows that the temperature increases for the interval of the feed rate (0.14mm/rev-0.2mm/rev) and the depth of cut (0.75mm-1mm) and it remains moderate for the lower intervals. As can be seen, for a depth of cut equal to 0.25mm, the temperature decreases. The thermogram in Figure 16 shows the concentration of heat on the nose of the tool.

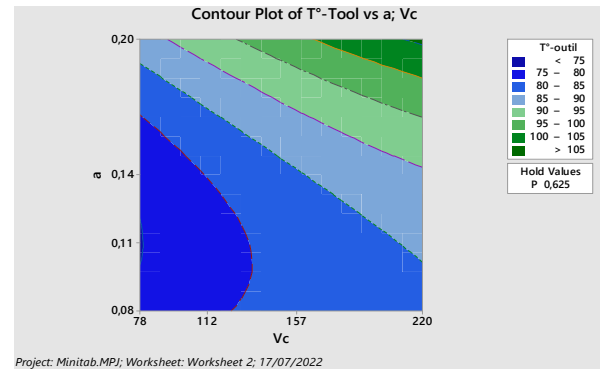


Fig. 11. Temperature isotherm as a function of cutting speed and feed rate.

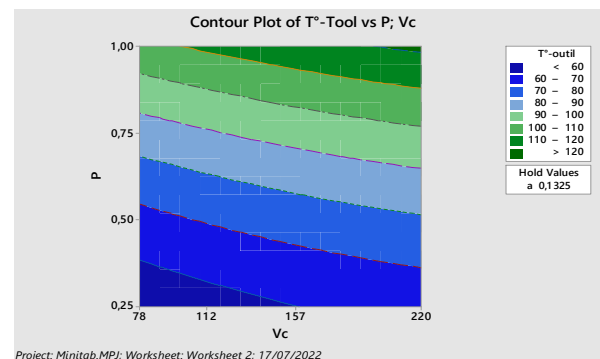
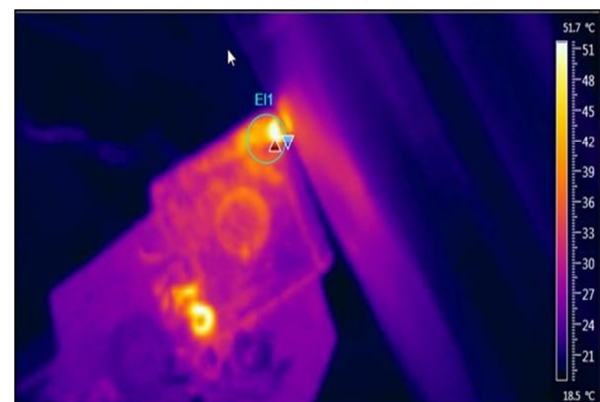
Fig. 12. Thermogram for  $V_c=78\text{m/min}$ ,  $a=0.08\text{mm/rev}$ , and  $P=0.25\text{mm}$ .

Fig. 13. Isotherm as a function of cutting speed and depth of cut.

Fig. 14. Thermogram for  $V_c=157\text{m/min}$ ,  $a=0.08\text{mm/rev}$ , and  $P=0.75\text{ mm}$ .



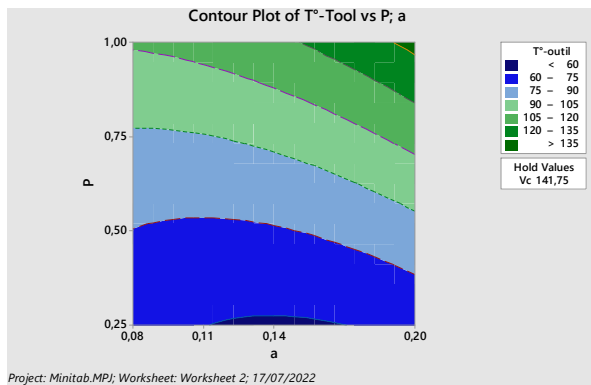
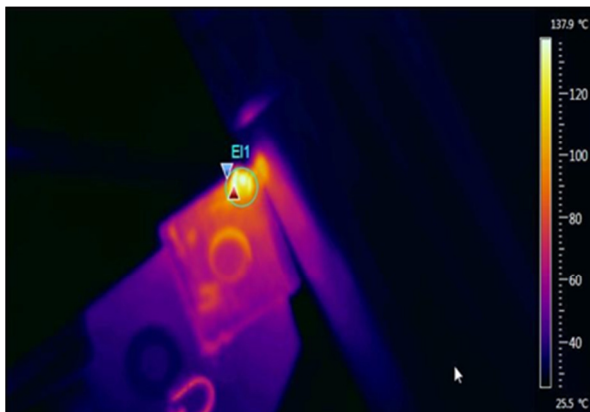


Fig. 15. Isotherm as a function of feed rate and depth of cut.

Fig. 16. Thermogram for  $V_c = 112\text{m/min}$ ,  $a = 0.14\text{ mm/rev}$ , and  $P = 1\text{ mm}$ .

## V. OPTIMIZATION OF CUTTING CONDITIONS

This goal can be achieved by intervening in the different stages of the production of the part:

- Design of the assembly to be manufactured.
- Part definition.
- Choice of manufacturing range.
- Choice of cutting conditions.

Cutting conditions can be optimized by setting an optimum cutting temperature obtained during experimentation. It can be given by the manufacturer depending on the technical characteristics of the pad. In this paper, a method is proposed to find the optimal cutting conditions. The choice of cutting conditions has a direct impact on the machining time of the part, therefore on the cost price of machining and on the productivity improvement. The optimization plot allows us to interactively modify the cut parameters of the input variables to perform sensitivity analysis and possibly improve the initial solution. We apply the optimization technique of the Minitab software used in the current work, for the measurement just after the release of the tool and under the following machining constraints:

$$78 \leq V_c \leq 220; \quad 0.08 \leq a \leq 0.2; \quad 0.25 \leq p \leq 1 \quad (3)$$

By choosing a temperature of  $100^\circ$  we obtained the results of the optimized input variables grouped in Table IV.

TABLE IV. OPTIMIZED CONDITIONS

Solution	$T^\circ$				
	$V_c$	$a$	$p$	Fit	Desirability
1	149	0.14	0.348485	99.967	0.999446
2	220	0.185560	0.256348	100.148	0.997301
3	78	0.150854	0.991703	103.348	0.939131

For desirability index equal to 0.999446, the optimal conditions are:  $V_c = 149\text{m/min}$ ,  $a = 0.14\text{mm/rev}$ ,  $p = 0.34\text{ mm}$ . The optimization diagram in Figure 17 shows the influence of the cutting conditions with the fixed cutting temperature.



Fig. 17. Impact of cutting conditions.

## VI. CONCLUSION

This work shows that a thermography technique is a useful procedure for establishing an experimental methodology for evaluating cutting temperature during the machining operations. It is therefore a method of mapping temperatures in real time, allowing the evaluation over time of surface thermal phenomena. The advances in manufacturing technology require us to estimate and predict the temperature rise at different points of the workpiece-tool interface in order to optimally adjust different parameters beforehand to improve machinability.

No previous research work has used only an infrared camera. The major difficulty in taking the machining sequence by the camera is the rolling of the chip, hence the need to choose the cutting conditions well and to have good protection and fixing. The experimental results allow us to conclude that the variations of the temperatures are very weak in permanent regimes for invariant cutting parameters, that the stresses applied by the tool on the machined material lead to a rise in temperature and that the heat flow generated during machining is largely evacuated by the chip. The experiments allowed us to determine the temperatures of the machined surface of the part as well as that of the cutting edge for different cutting parameters. The fit of the model was checked and found to be adequate at an accuracy level of 88.87% for temperature measurements just after tool release. This demonstrates that the model can be used to predict the cutting temperature. The technique could be an economical and effective method for temperature prediction and optimization of the cutting parameters.

The quality of the parts produced depends largely on the cutting conditions and this quality is an important parameter during the production of mechanical parts. The cutting fluid is



also an important parameter since it allows to limit the rise in temperature in the cutting zones and to establish the presence of a fluid film between the different surfaces. However, in order to protect the environment and the health of the personnel, the current trend is to limit and even eliminate the use of cutting fluids.

## REFERENCES

- [1] S. K. Choudhury and P. Srinivas, "Tool wear prediction in turning," *Journal of Materials Processing Technology*, vol. 153–154, pp. 276–280, Nov. 2004, <https://doi.org/10.1016/j.jmatprotec.2004.04.296>.
- [2] J.-T. Horng, N.-M. Liu, and K.-T. Chiang, "Investigating the machinability evaluation of Hadfield steel in the hard turning with Al<sub>2</sub>O<sub>3</sub>/TiC mixed ceramic tool based on the response surface methodology," *Journal of Materials Processing Technology*, vol. 208, no. 1, pp. 532–541, Nov. 2008, <https://doi.org/10.1016/j.jmatprotec.2008.01.018>.
- [3] O. B. Abouelatta and J. Mádl, "Surface roughness prediction based on cutting parameters and tool vibrations in turning operations," *Journal of Materials Processing Technology*, vol. 118, no. 1, pp. 269–277, Dec. 2001, [https://doi.org/10.1016/S0924-0136\(01\)00959-1](https://doi.org/10.1016/S0924-0136(01)00959-1).
- [4] I. Mukherjee and P. K. Ray, "A review of optimization techniques in metal cutting processes," *Computers & Industrial Engineering*, vol. 50, no. 1, pp. 15–34, May 2006, <https://doi.org/10.1016/j.cie.2005.10.001>.
- [5] P. J. T. Conradie, G. A. Oosthuizen, N. F. Treurnicht, and A. A. Shaalane, "Overview of Work Piece Temperature Measurement Techniques for Machining of Ti6Al4V#," *The South African Journal of Industrial Engineering*, vol. 23, no. 2, 2012, <https://doi.org/10.7166/23-2-335>.
- [6] N. M. M. Reddy and P. K. Chaganti, "Investigating Optimum SiO<sub>2</sub> Nanolubrication During Turning of AISI 420 SS," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3822–3825, Feb. 2019, <https://doi.org/10.48084/etasr.2537>.
- [7] A. R. Motorcu, Y. Isik, A. Kus, and M. C. Cakir, "Analysis of the cutting temperature and surface roughness during the orthogonal machining of AISI 4140 alloy steel via the Taguchi method," *Materiali in tehnologije*, vol. 50, no. 3, pp. 343–351, Jun. 2016, <https://doi.org/10.17222/mit.2015.021>.
- [8] A. Kus, Y. Isik, M. C. Cakir, S. Coşkun, and K. Özdemir, "Thermocouple and Infrared Sensor-Based Measurement of Temperature Distribution in Metal Cutting," *Sensors*, vol. 15, no. 1, pp. 1274–1291, Jan. 2015, <https://doi.org/10.3390/s150101274>.
- [9] N. A. Abukhshim, P. T. Mativenga, and M. A. Sheikh, "Heat generation and temperature prediction in metal cutting: A review and implications for high speed machining," *International Journal of Machine Tools and Manufacture*, vol. 46, no. 7, pp. 782–800, Jun. 2006, <https://doi.org/10.1016/j.ijmachtools.2005.07.024>.
- [10] I. Korkut, M. Boy, I. Karacan, and U. Seker, "Investigation of chip-back temperature during machining depending on cutting parameters," *Materials & Design*, vol. 28, no. 8, pp. 2329–2335, Jan. 2007, <https://doi.org/10.1016/j.matdes.2006.07.009>.
- [11] N. Benhadji Serradj, "Contribution to the evaluation of the thermal field at the tool-part interface for the optimization of machining conditions," Ph.D. dissertation, University of Tlemcen, Tlemcen, Algeria, 2022.
- [12] N. B. Serradj, A. D. K. Ali, and M. E. A. Ghernaout, "A Contribution to the Thermal Field Evaluation at the Tool-Part Interface for the Optimization of Machining Conditions," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7750–7756, Dec. 2021, <https://doi.org/10.48084/etasr.4235>.
- [13] M. PradeepKumar, K. Amarnath, and M. SunilKumar, "A Review on Heat Generation in Metal Cutting," *International Journal of Engineering and Management Research*, vol. 5, no. 4, pp. 193–197, Aug. 2015.
- [14] O. Riou, P.-O. Logerais, and J.-F. Durastanti, "Quantitative study of the temperature dependence of normal LWIR apparent emissivity," *Infrared Physics & Technology*, vol. 60, pp. 244–250, Sep. 2013, <https://doi.org/10.1016/j.infrared.2013.05.012>.
- [15] N. M. M. Reddy and P. K. Chaganti, "Investigating Optimum SiO<sub>2</sub> Nanolubrication During Turning of AISI 420 SS," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3822–3825, Feb. 2019, <https://doi.org/10.48084/etasr.2537>.
- [16] D. Kara Ali, N. Benhadji Serradj, and M. E. A. Ghernaout, "Qualification and Validation of an in-situ Measurement Method of the Machining Temperature," in *Mechanism, Machine, Robotics and Mechatronics Sciences*, Cham, 2019, pp. 15–27, [https://doi.org/10.1007/978-3-319-89911-4\\_2](https://doi.org/10.1007/978-3-319-89911-4_2).
- [17] *FLIR SC325*. FLIR Systems, 2010.

# Groundwater Quality Assessment Based on a Statistical Approach in Gaya District, Bihar

**Shaz Ahmad**

Department of Civil Engineering, National Institute of Technology Patna, India  
shaza.phd19.ce@nitp.ac.in  
(corresponding author)

**Reena Singh**

Department of Civil Engineering, National Institute of Technology Patna, India  
reena@nitp.ac.in

Received: 16 October 2022 | Revised: 5 November 2022 | Accepted: 6 November 2022

## ABSTRACT

India is one of the countries that face the serious problem of groundwater contamination. The current study's main objective is to evaluate the quality of the groundwater in the Serghati and its surrounding region of the Gaya district and its suitability for drinking purposes. To achieve this aim, 75 groundwater samples from the 15 sampling sites were collected during the period from March to May 2022. We measured and analyzed the major physicochemical characteristics of the water and compared them to the World Health Organization (WHO) standards. With the help of the Water Quality Index (WQI), groundwater quality was assessed. According to the study results, 3 sites have a WQI value of more than 100, which is unsuitable for drinking. Correlation matrices were used to assess groundwater quality and the extent of the interdependencies of the various parameters. Principal Component Analysis (PCA) reduces the number of significant variables. Three principal components with a total variance of 73.53% were identified and used in the analysis. Overall, the result indicates that most areas' water quality is good and safe for drinking.

**Keywords-**drinking water; physiochemical parameters; statistical analysis; Pearson's correlation; principal component analysis

## I. INTRODUCTION

Water resources play a critical role in the growth and development of civilization [1]. Nowadays, water resources have inevitably become one of the most important factors that determine economic growth. It is necessary to assess whether water resources are sufficient and sustainable, particularly in countries like India, where water often plays a multiple-faceted role in terms of economics as well as social aspects [2]. Maintaining access to safe drinking water has become an urgent requirement, even though 30% of the urban population and 90% of the rural population are still wholly dependent on contaminated surface and ground water to meet their daily needs. The availability and quality of clean drinking water are among the most important factors for promoting agricultural production and providing health care services to the general population [3]. Many states, including Andhra Pradesh, Bihar, Rajasthan, West Bengal, Jharkhand, Orissa, and Punjab, are experiencing severe water scarcity, due to factors such as the lack of environmental awareness and indiscriminate waste disposal from agriculture, mining, and anthropology, resulting in one of the most dangerous situations for development in their areas [4]. The use of fertilizers and agrochemicals and the overexploitation of groundwater, result in the contamination of

water resources [5]. Residents of arid and semi-arid regions rely heavily on groundwater [6]. Agricultural lands are also getting deteriorated due to the contamination of groundwater by excess fluoride [7, 8]. Fluoride-contaminated groundwater slowly accumulates the excess fluoride and produces adverse effects [8, 9]. Large quantities of fluoride are stored in various parts of the body due to the use of contaminated water [10]. Crippling skeletal fluorosis has become common across the globe [2]. The immune system also gets affected by the intake of excess fluoride [11]. Chronic kidney diseases may attack at a faster rate due to the intake of excess fluoride through food [12]. To remove the excess fluoride, various techniques have been developed. Among them, adsorption is found to be more effective in both removal and cost. Various materials have been developed as defluorination adsorbents, such as nanomaterials [13], clay [14], chitosan [15], industrial waste [16], carbonaceous [17], alumina [18], calcium [19], and metal oxides [20].

## II. MATERIALS AND METHODS

### A. Study Area

Gaya district stretches over 4976km<sup>2</sup> and lies between 24°30' and 25°06' latitude and 84°24' and 85°30' longitude The

entire research region has a hostile environment caused by the continental monsoon. Mountains surround Gaya on three sides, with a river on the other side, resulting in seasonal temperature variations in the region. Higher temperatures up to 45°C can be found during the summer months (May-July). The city receives about 214cm of rainfall in July and October. The winter months are known for their cold temperatures ranging from -4°C to 28°C. The study was conducted in the surrounding area of Sherghati, located in the Gaya district of Bihar.

### B. Water Sampling and Analysis

A total of 75 groundwater samples, 5 from each of the 15 sites, were taken across the research region from underground sources (maximum bore wells and hand pumps). The sampling site and geo-positions (latitude and longitude) were located with Global Positioning System (GPS). The concentrations of each individual parameter at each sampling site were determined by averaging the concentrations of each parameter. Between March and May 2022, fresh groundwater samples were collected from shallow bore wells and tube wells. The water samples were collected using pre-cleaned high-density 500mL polythene bottles. The bottles were handled with gloves to prevent sample contamination. Before sampling the water from a bore well, it was drained for 5 minutes. The samples were sealed tightly so contaminants were not able to enter the bottles. As soon as the samples were transported to the laboratory, they were encased in iceboxes and were preserved at 4°C for further chemical analysis. All steps were carefully monitored throughout the sampling process to ensure that the samples were not contaminated or agitated during the collection, transportation, and analysis. The fluoride concentration in the water was determined using the SPADNS method described in APHA1994 [21]. The chemicals and distilled water used to prepare the solution (or dilution) were analytical grades and the highest purity Millipore water, respectively. Thermo Scientific Multi-Parameter Kit measured in situ parameters like pH immediately after sampling. The anions chloride, sulphate, and phosphate were quantified using the usual  $\text{AgNO}_3$  titration with a UV spectrophotometer, turbid metric, and colorimetric methods. Authors in [3] used a technique to assess water samples' nitrate ( $\text{NO}_3$ ) content. Calcium and magnesium were determined using the conventional EDTA titration technique. All the processes used in the investigation were conventional methods for examining water and wastewater [22].

### C. Estimation of Water Quality Index (WQI)

WQI was estimated according to a three step process:

- Step-1. The weight ( $W_n$ ) characteristics of each specification are calculated to utilize the given equation:

$$W_n = K/S_n \quad (1)$$

where  $K = \frac{1}{\frac{1}{S_1} + \frac{1}{S_2} + \frac{1}{S_3} + \dots + \frac{1}{S_n}} = \frac{1}{\sum \frac{1}{S_n}}$ , where  $S_n$  is the standard desirable value of the  $n$ th specification.

- Step-2. The Sub-Index ( $Q_n$ ) is calculated to utilize the given equation:

$$Q_n = \frac{V_n - V_0}{S_n - V_0} \times 100 \quad (2)$$

where  $V_n$  is the mean concentration of the  $n$ th specification,  $S_n$  is the standard desirable value of the  $n$ th specification, and  $V_0$  represents the real values of the specifications in pure water (Generally,  $V_0 = 0$ , for most specifications except for pH, that has a  $V_0$  of 7).

- Step-3. WQI was calculated by the summation of  $W_n$  and  $Q_n$ :

$$WQI = \sum Q_n W_n / \sum W_n \quad (3)$$

and since  $\sum W_n = 1$ , we have:

$$WQI = \sum Q_n W_n \quad (4)$$

### D. Correlation

An evaluation of the correlation between two variables can be measured by the coefficient of correlation ( $r$ ), which can be defined as a measure of the degree to which the two variables are associated. As far as the range of  $r$  is concerned, it spans from -1 to +1 [23]. As a general rule, the correlation between different water quality parameters is obtained to determine their interrelationship, which helps visualizing the most effective parameters [24]. This ultimately contributes to the decision-making process related to monitoring water quality. However, a high or low value of  $r$  (for example, close to +1) is typically indicative of a strong correlation [25].

### E. Principal Component Analysis (PCA)

The PCA method of multivariate statistics is usually used to reduce the dimension of a dataset [26]. PCA produces eigenvalues and eigenvectors when applied to the covariance matrix of the initially correlated variables [27]. Several coefficients make up the eigenvector, also referred to as the loading. To obtain the Principal Components (PCs) [28], one has to multiply all the loadings by the original set of variables. PCA is a linear combination of initially correlated variables that forms a new set of orthogonal uncorrelated variables. In 1933, Hotelling developed this method and had been widely adopted in studies regarding water quality assessment [29].

## III. RESULTS AND DISCUSSION

The physicochemical parameters of the groundwater samples are presented in Table I. The pH of the drinking water at the 15 sites has been found to range between 6.78 and 7.20. Over the study area, it is evident that the water is slightly acidic to slightly basic. There is no site where the pH values exceed the permissible limit of 8.5. As far as chloride is concerned, the minimum value is 43.23mg/L and the maximum value is 203.49mg/L and all samples are within the maximum allowable limit of 1000mg/L. Sulphate concentration ranges from 15.67 to 95.18mg/L, with a mean value of 39.09mg/L and a standard deviation of 24.90mg/L. The sulphate levels do not exceed the permissible limit of 400mg/L. There is a wide range of nitrate concentrations in the samples, ranging from 3.54 to 43.23mg/L with a mean and standard deviation of 15.54 and 10.49mg/L, respectively. The nitrate levels did not exceed the permissible limit of 45mg/L. Likewise, the minimum fluoride concentration was 0.37mg/L, the maximum fluoride concentration was

2.70mg/L, with 6 locations crossing the permissible limit of 1.5mg/L. Total Hardness (TH) ranged between 92 and 332 mg/L, with an average and standard deviation of 228.93mg/L and 58.39mg/L, respectively. No sites had TH values crossing the permissible limit of 600mg/L. At some sites, the concentrations of parameters that describe the water quality were higher than those recommended by the Bureau of Standards (2012) for drinking water. Combinations of different natural factors, such as soil salinization, mineral dissolution, prolonged residence time for water-rock interactions may be responsible. However, it should be noted that anthropogenic activities also play a significant role in the degradation of groundwater quality. There is a great concern regarding the concentration of nitrate and fluoride in groundwater since their increased value adversely affects human health.

TABLE I. STATISTICAL ANALYSIS OF PHYSICOCHEMICAL PARAMETERS

	Minimum	Maximum	Mean	Std. deviation
pH	6.78	7.20	7.05	0.12
TDS	290	426.60	331.02	33.98
F <sup>-</sup>	0.37	2.70	1.378	0.7
Cl <sup>-</sup>	43.23	203.49	108.84	44.89
NO <sub>3</sub> <sup>-</sup>	3.54	43.23	15.54	10.5
SO <sub>4</sub> <sup>2-</sup>	15.67	95.18	39.09	24.91
TH	92	332	228.93	58.4
Total alkalinity	21.40	40.50	31.36	6.59

TABLE II. GROUNDWATER WQI VALUES

Sampling location	WQI
S1	81.25
S2	39.61
S3	43.6
S4	76.52
S5	90.9
S6	119.78
S7	149.15
S8	97.43
S9	20.85
S10	37.45
S11	36.26
S12	128.94
S13	47.3
S14	24.54
S15	70.32

We employed the WQI to investigate the groundwater's condition and assess whether it is suitable for drinking based on its overall quality. As WQI uses a weighted sum approach to calculate its score, it makes it easy for the data to be communicated to a broad audience because it combines all the parameters into one numerical value that can be used to convey information. Table II presents the WQI of each site. There is no doubt that 3 out of 15 sites have WQI above 100, as shown in Figure 1. The groundwater at these sites is not suitable for drinking. It was found that the groundwater had the highest WQI according to this study (149.14), which indicates that the groundwater is the most polluted. However, the lowest WQI at the least polluted site is 24.53. Generally, the groundwater quality in the study area is acceptable for most areas, but it is of very poor quality in some regions and shouldn't be used for drinking. We assessed the inter-dependencies among different

parameters by establishing the correlations between their values. The analysis outcome is shown in Table III.

### Water quality of samples

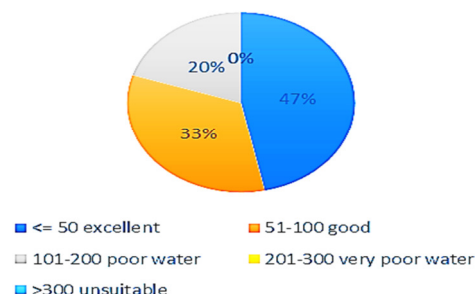


Fig. 1. Water quality based on WQI value.

TABLE III. PEARSON'S CORRELATION MATRIX

	pH	TDS	F <sup>-</sup>	Cl <sup>-</sup>	NO <sub>3</sub> <sup>-</sup>	SO <sub>4</sub> <sup>2-</sup>	TH	TA
pH	1.000							
TDS	-0.351	1.000						
F <sup>-</sup>	0.565	-0.214	1.000					
Cl <sup>-</sup>	0.315	0.340	0.242	1.000				
NO <sub>3</sub> <sup>-</sup>	-0.0022	0.361	-0.096	0.396	1.000			
SO <sub>4</sub> <sup>2-</sup>	-0.370	0.776	-0.192	0.527	0.584	1.000		
TH	0.347	-0.430	0.047	0.205	-0.094	-0.225	1.000	
TA	0.126	0.150	0.097	0.376	0.331	0.111	0.592	1.000

TABLE IV. TOTAL VARIANCE ANALYSIS

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	3.058	33.974	33.974	3.058	33.974	33.974
2	2.294	25.485	59.458	2.294	25.485	59.458
3	1.266	14.072	73.530	1.266	14.072	73.530
4	0.956	10.627	84.157			
5	0.626	6.957	91.114			
6	0.364	4.050	95.164			
7	0.227	2.525	97.689			
8	0.178	1.981	99.670			
9	0.030	0.330	100.000			

Extraction Method: PCA.

An increase in positive or negative values of  $r$  is usually indicative of a strong correlation between two variables. It can be seen that most of the parameters demonstrate a positive correlation with one another. Generally, the parameters from a common origin are strongly correlated and vice versa. Despite this, some of these parameters show a negative correlation and some a very small correlation, indicating that these components' sources may differ.

PCA was also conducted on each of the water quality parameters, and the results are shown in Table IV. The variances for each PC are individually and cumulatively based on the data collected. There is no doubt that PC1, PC2, PC3 are responsible for the majority of the Variance (73.53%), while the rest of the PCs are responsible for a much lower percentage of the Variance. Most of the time, it is recommended that the PCs capable of explaining 70% of the Variance are used to

reduce dimensionality while losing the least amount of information [30]. To determine the leading PCs from the original variables in the study, PCA was carried out using the SPSS 16.0 software. The first factor accounts for about 33.974% of the variance, corresponding to the largest eigenvalue (3.058). In the total variance calculation, the second factor, which corresponds to the second eigenvalue (2.294), is responsible for approximately 25.485% of the variance. The third factor accounts for 14.072 % of the total variance with an eigenvalue of 1.266. Among the remaining 9 factors, no one has an eigenvalue greater than 1. The significance of any factor is determined by its eigenvalue, which must be greater than 1. There appears to be a relatively higher loading for PC1 for the parameters TDS and sulphate compared to the others, as shown in Table V. Meanwhile, hardness, alkalinity, chloride, and pH loadings are found to be more in PC2, and in PC3, loading is more in fluoride.

TABLE V. COMPONENT MATRIX

	Component		
	1	2	3
pH	-0.589	0.564	0.359
TDS	0.878	0.090	0.146
F- (ppm)	-0.379	0.406	0.558
Cl- (ppm)	0.272	0.810	0.260
NO <sub>3</sub> <sup>-</sup> (mg/L)	0.595	0.453	0.006
SO <sub>4</sub> <sup>2-</sup> (mg/L)	0.849	0.283	0.149
TH (mg/l)	-0.467	0.568	-0.601
Total alkalinity (mg/l)	0.124	0.699	-0.567

Extraction Method: PCA

#### IV. CONCLUSION

Generally, there is a groundwater contamination issue in Sherghati, Gaya district. The purpose of this study is to investigate the quality of groundwater and its suitability for drinking. The available water in the region lies between slightly acidic to slightly alkaline. To analyze the status, we calculated the water quality index for each site of the sampling area. The physicochemical parameters are almost within the permissible range except for fluoride, which crosses the permissible limit at some sampling sites. WQI was accurately predicted for the groundwater samples, something that can be of great importance for analyzing and reducing environmental impacts and, ultimately, ensuring public health. We used the PCA methodology to analyze the water quality measurements, and we extracted 3 components with a total variance of 73.53%. Most of the 15 samples have WQI value below 100 and only 3 cross the value of 100. Hence, there is good quality drinking water available in the area.

#### REFERENCES

- [1] W. Guissouma, O. Hakami, A. J. Al-Rajab, and J. Tarhouni, "Risk assessment of fluoride exposure in drinking water of Tunisia," *Chemosphere*, vol. 177, pp. 102–108, Jun. 2017, <https://doi.org/10.1016/j.chemosphere.2017.03.011>.
- [2] T. Arfin and S. Waghmare, "Fluoride removal from water by various techniques: Review," *International Journal of Innovative Science, Engineering & Technology*, vol. 2, no. 9, pp. 560–571, Jul. 2015.
- [3] A. N. Laghari, Z. A. Siyal, D. K. Bangwar, M. A. Soomro, G. D. Walasai, and F. A. Shaikh, "Groundwater Quality Analysis for Human Consumption: A Case Study of Sukkur City, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2616–2620, 2018, <https://doi.org/10.48084/etasr.1768>.
- [4] S. Ahmad, R. Singh, T. Arfin, and K. Neeti, "Fluoride contamination, consequences and removal techniques in water: a review," *Environmental Science: Advances*, 2022, <https://doi.org/10.1039/D1VA00039J>.
- [5] A. Raja and T. Gopikrishnan, "Drought Analysis Using the Standardized Precipitation Evapotranspiration Index (SPEI) at Different Time Scales in an Arid Region," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9034–9037, 2022, <https://doi.org/10.48084/etasr.5141>.
- [6] K. Praveen and L. B. Roy, "Study of Groundwater Quality for Irrigation Purpose – A Case Study of Paliganj Distributary, Bihar, India," *Natural Volatiles & Essential Oils Journal*, vol. 8, no. 6, pp. 3461–3477, 2021.
- [7] S. Waghmare, D. Lataye, T. Arfin, N. Manwar, S. Rayalu, and N. Labhsetwar, "Adsorption Behavior of Eggshell Modified Polyalthia Longifolia Leaf Based Alumina as a Novel Adsorbents for Fluoride Removal from Drinking Water," *International Journal of Advance Research and Innovative Ideas in Education*, vol. 1, no. 5, pp. 904–926, 2015.
- [8] S. Ahmad and R. Singh, "Spatial Distribution and Health Risk Assessment based on Groundwater Fluoride Enrichment in Gaya, Bihar, India," *Engineering, Technology & Applied Science Research*, vol. 13, no. 1, pp. 9825–9829, Feb. 2023, <https://doi.org/10.48084/etasr.5397>.
- [9] S. S. Waghmare and T. Arfin, "Fluoride induced water pollution issue and its health efficacy in India: A review," *Citeseer*, vol. 3, no. 5, pp. 345–358, 2015.
- [10] S. Waghmare, T. Arfin, S. Rayalu, D. Lataye, S. Dubey, and S. Tiwari, "Adsorption behavior of modified zeolite as novel adsorbents for fluoride removal from drinking water: surface phenomena, kinetics and thermodynamics studies," *International Journal of Science, Engineering and Technology Research*, vol. 4, no. 12, pp. 4114–4124, 2015.
- [11] S. S. Waghmare, T. Arfin, N. Manwar, D. H. Lataye, and N. Labhsetwar, S. Rayalu, "Preparation and Characterization of Polyalthia longifolia Based Alumina as a Novel Adsorbent for Removing Fluoride from Drinking Water," *Asian Journal of Advanced Basic Sciences*, vol. 4, no. 1, pp. 12–24, Aug. 2015.
- [12] K. Praveen and L. B. Roy, "Assessment of Groundwater Quality Using Water Quality Indices: A Case Study of Paliganj Distributary, Bihar, India," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8199–8203, Feb. 2022, <https://doi.org/10.48084/etasr.4696>.
- [13] S. Waghmare, D. Lataye, T. Arfin, and S. Rayalu, "Defluoridation by Nano-Materials, Building Materials and Other Miscellaneous Materials: A Systematic Review," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 4, pp. 11998–12010, Dec. 2015, <https://doi.org/10.15680/IJIRSET.2015.0412046>.
- [14] S. S. Waghmare and T. Arfin, "Fluoride Removal by Clays, Geomaterials, Minerals, Low Cost Materials and Zeolites by Adsorption: A Review," *International Journal of Science, Engineering and Technology Research*, vol. 4, no. 11, pp. 3663–3676, 2015.
- [15] S. S. Waghmare and T. Arfin, "Defluoridation By Adsorption With Chitin - Chitosan-Alginate - Polymers - Cellulose - Resins - Algae And Fungi: A Review," *International Research Journal of Engineering and Technology*, vol. 2, no. 6, pp. 1178–1197, 2015.
- [16] T. Arfin and S. S. Waghmare, "Fluoride Removal By Industrial, Agricultural and Biomass Wastes As Adsorbents: Review," *International Journal of Advance Research and Innovative Ideas in Education*, vol. Vol-1, no. 4, pp. 2395–4396, 2015.
- [17] S. S. Waghmare and T. Arfin, "Fluoride Removal from Water By Carbonaceous Materials: Review," *International Journal of Modern Trends in Engineering and Research*, vol. 2, no. 9, pp. 355–361, 2015.
- [18] S. S. Waghmare and T. Arfin, "Fluoride Removal from Water by Aluminium Based Adsorption: A Review," *Journal of Biological and chemical Chronicles*, vol. 2, no. 1, 2015.
- [19] S. S. Waghmare, T. Arfin, "Fluoride Removal from Water By Calcium Materials: A State-Of-The-Art Review," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 4, no.

- 9, pp. 8090–8102, Sep. 2015, <https://doi.org/10.15680/IJRSET.2015.0409013>.
- [20] T. Arfin and S. S. Waghmare, "Fluoride removal from water by mixed metal oxide adsorbent materials: A state-of-the-art review," *International Journal of Engineering Sciences & Research Technology*, vol. 4, no. 9, pp. 519–536, Jul. 2015.
- [21] R. Khan and D. C. Jhariya, "Groundwater quality assessment for drinking purpose in Raipur city, Chhattisgarh using water quality index and geographic information system," *Journal of the Geological Society of India*, vol. 90, no. 1, pp. 69–76, Jul. 2017, <https://doi.org/10.1007/s12594-017-0665-0>.
- [22] X. He, P. Li, Y. Ji, Y. Wang, Z. Su, and V. Elumalai, "Groundwater Arsenic and Fluoride and Associated Arsenicosis and Fluorosis in China: Occurrence, Distribution and Management," *Exposure and Health*, vol. 12, no. 3, pp. 355–368, Sep. 2020, <https://doi.org/10.1007/s12403-020-00347-8>.
- [23] D. Ortiz-Pérez *et al.*, "Fluoride-induced disruption of reproductive hormones in men," *Environmental Research*, vol. 93, no. 1, pp. 20–30, Sep. 2003, [https://doi.org/10.1016/s0013-9351\(03\)00059-8](https://doi.org/10.1016/s0013-9351(03)00059-8).
- [24] N. Kumar, A. A. Mahessar, S. A. Memon, K. Ansari, and A. L. Qureshi, "Impact Assessment of Groundwater Quality using WQI and Geospatial tools: A Case Study of Islamkot, Tharparkar, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5288–5294, Feb. 2020, <https://doi.org/10.48084/etasr.3289>.
- [25] K. Loganathan and A. J. Ahamed, "Multivariate statistical techniques for the evaluation of groundwater quality of Amaravathi River Basin: South India," *Applied Water Science*, vol. 7, no. 8, pp. 4633–4649, Dec. 2017, <https://doi.org/10.1007/s13201-017-0627-0>.
- [26] P. K. Mohapatra, R. Vijay, P. R. Pujari, S. K. Sundaray, and B. P. Mohanty, "Determination of processes affecting groundwater quality in the coastal aquifer beneath Puri city, India: a multivariate statistical approach," *Water Science and Technology: A Journal of the International Association on Water Pollution Research*, vol. 64, no. 4, pp. 809–817, 2011, <https://doi.org/10.2166/wst.2011.605>.
- [27] P. Sahu, G. C. Kisku, P. K. Singh, V. Kumar, P. Kumar, and N. Shukla, "Multivariate statistical interpretation on seasonal variations of fluoride-contaminated groundwater quality of Lalganj Tehsil, Raebareli District (UP), India," *Environmental Earth Sciences*, vol. 77, no. 13, p. 484, Jun. 2018, <https://doi.org/10.1007/s12665-018-7658-1>.
- [28] T. A. Chandio, M. N. Khan, and A. Sarwar, "Fluoride estimation and its correlation with other physicochemical parameters in drinking water of some areas of Balochistan, Pakistan," *Environmental Monitoring and Assessment*, vol. 187, no. 8, Aug. 2015, Art. no. 531, <https://doi.org/10.1007/s10661-015-4753-6>.
- [29] D. Marghade, D. B. Malpe, N. Subba Rao, and B. Sunitha, "Geochemical assessment of fluoride enriched groundwater and health implications from a part of Yavatmal District, India," *Human and Ecological Risk Assessment*, vol. 26, no. 3, pp. 673–694, 2020, <https://doi.org/10.1080/10807039.2018.1528862>.
- [30] N. Subba Rao, C. Srihari, B. Deepthi Spandana, M. Sravanthi, T. Kamalesh, and V. Abraham Jayadeep, "Comprehensive understanding of groundwater quality and hydrogeochemistry for the sustainable development of suburban area of Visakhapatnam, Andhra Pradesh, India," *Human and Ecological Risk Assessment: An International Journal*, vol. 25, no. 1–2, pp. 52–80, Feb. 2019, <https://doi.org/10.1080/10807039.2019.1571403>.



# Performance Evaluation of Emulsion Liquid Membrane on Chlorpyrifos Pesticide Removal: Stability, Mass Transfer Coefficient, and Extraction Efficiency Studies

**Farrah Emad Al-Damluji**

Department of Environmental Engineering, College of Engineering, University of Baghdad, Iraq  
farrah.emad@coeng.uobaghdad.edu.iq  
(corresponding author)

**Ahmed A. Mohammed**

Department of Environmental Engineering, College of Engineering, University of Baghdad, Iraq  
ahmed.abedm@yahoo.com

Received: 8 October 2022 | Revised: 2 November 2022 | Accepted: 2 November 2022

## ABSTRACT

Emulsion Liquid Membrane (ELM) is an emerging technology that removes contaminants from water and industrial wastewater. This study investigated the stability and extraction efficiency of ELM for the removal of Chlorpyrifos Pesticide (CP) from wastewater. The stability was studied in terms of emulsion breakage. The proposed ELM included n-hexane as a diluent, span-80 as a surfactant, and hydrochloric acid (HCl) as a stripping agent. Parameters such as mixing speed, aqueous feed solution pH, internal-to-organic membrane volume ratio, and external-to-emulsion volume ratio were investigated. A minimum emulsion breakage of 0.66% coupled with a maximum chlorpyrifos extraction and stripping efficiency were achieved at 96.1% and 95.7% at best-operating conditions of 250/50 external-to-emulsion volume ratio, external feed solution pH 6, 250rpm mixing speed, and 1:1 internal-to-membrane volume ratio at 10min contact time without utilizing a carrier agent. A study of extraction kinetics and estimation of mass transfer coefficient was also conducted ( $3.89 \times 10^{-9}$  m/s). The results of this work can be extended to the removal of other types of pesticides from wastewater.

**Keywords-**emulsion liquid membrane; chlorpyrifos; stability; mass transfer coefficient; extraction efficiency

## I. INTRODUCTION

Emerging Contaminants (ECs) such as synthetic dyes, cosmetics, pharmaceuticals (antibiotics, hormones, and others), and pesticides are considered major issues in many parts of the world [1]. The final fate of pesticides after use has different pathways: (1) soil top layer accumulation, where pesticides may be adsorbed and deposited, (2) degradation on both the plant surface and the soil, (3) plant organs translocation, and (4) dissolution-transport of contaminants from soil and plant surfaces into water runoff causing water resource contamination [2, 3]. Chlorpyrifos (CP) is a broad-spectrum organophosphorus insecticide (OP) and has a high value of octanol-water partition coefficient  $K_{ow} = 4.70$  [4]. Exposure to chlorpyrifos can produce nausea, diarrhea, salivation, vomiting, tremor, and convulsion symptoms [5-7]. Chlorpyrifos is banned in many countries, nevertheless, monitoring still has to be carried out. The liquid membrane extraction, commonly known as Emulsion Liquid Membrane (ELM), was presented as a

substitute technique for separation by a polymeric membrane as well as liquid-liquid extraction [8, 9]. The mechanism of ELM depends on the diffusion of the solute (contaminant), passing through an organic liquid membrane [10]. The main advantages of ELM are: (1) high diffusion rate of the contaminant through the membrane, (2) high interfacial area for solute mass transfer at the internal water-oil interface due to the small-sized droplets, (3) ability to treat an assortment of compounds and elements in an industrial setting at higher speeds along with an extraordinary level of effectiveness for various solute (contaminants) volume concentrations, and (4) simultaneous performance of both extraction and stripping at the external and the internal interfaces of a system, respectively [11, 12]. Despite the promising features of ELM, its instability has impeded widespread applications on a larger scale. The stability of an emulsion is defined as liquid membrane resistance to high shear stress during solute extraction. The best stability level is essential in the ELM process [13], as it is produced using the appropriate surfactant to overcome the

difficulties of an ELM system on a practical industrialized level [14].

This study investigated chlorpyrifos extraction from contaminated water using ELMs to propose a novel method for pesticide extraction. The influences of stirring speed, internal aqueous phase-organic phase volume ratio, external feed phase pH, external phase-emulsion volume ratio, and contact time were studied to determine the best conditions to achieve minimal membrane breakage in the emulsion system and reach maximum removal efficiency.

## II. MATERIALS AND METHODS

### A. Chemicals and Equipment

This study used analytical reagent-grade chemicals. The CP used was purchased from the local market (Om Agro Chemicals, India). The chemical formula of CP ( $C_9H_{11}Cl_3NO_3PS$ ) can be seen in [19]. CP has a molar mass of 350.6g/mol, 1.4g/cm<sup>3</sup> density, 1.87×10<sup>-5</sup>mmHg vapor pressure at 25°C, 41.5-42.5°C melting point, and decomposition at 160°C [15, 16]. Hydrochloric acid (HCl) and sodium hydroxide (NaOH) were acquired from Thomas beaker in India. The organic membrane consists of a diluent and a surfactant. The nonionic surfactant was sorbitan monooleate, commonly recognized as Span80, obtained from Merck, Darmstadt, Germany (Sigma Aldrich); n-hexane was used as a diluent, obtained from Thomas beaker, India, which was selected because of the higher rate of transport of solute than heptane and kerosene [17, 18]. All laboratory tests were carried out at a room temperature of 20±2°C, using Mtops SR 30 homogenizer, Heidolph RZR 2021 compact digital mixer system, quartz cells, PG instruments T80+ UV/VIS spectrometer, Isolab centrifuge machine, ATC pH meter, and an Isolab magnetic stirrer with a temperature controller.

### B. Experimental Work

The membrane phase was produced by dissolving a specific amount of surfactant (Span 80) in n-hexane by gently stirring via a magnetic stirrer. The emulsion was established by adding the internal phase (HCl) dropwise to the organic membrane phase (Span 80 and n-hexane as a solvent) while using a homogenizer. The resulting emulsion was then added to the external phase while mixing continuously. Dispersion occurs in the shape of globule formation, where each globule is made of stripping solution droplets encased in the membrane solution. Samples were withdrawn from the mixture at certain time intervals of 1, 2, 4, 6, 8, 10, 15, and 20min. The sample containing both the emulsion and the external phase was filtered (pore size 0.22µm diameter). After extraction, a separation of the emulsion from the feed solution was carried out by gravity, and a demulsification process was carried out by applying a centrifugal force to the emulsion (4000rpm for 30min) to segregate the phases, while the contaminant would be extracted as a concentrated solution. The CP concentration in the separated external phase and the filtered samples was measured periodically using an ultraviolet spectrophotometer (UV). Figure 1 shows the ELM process.

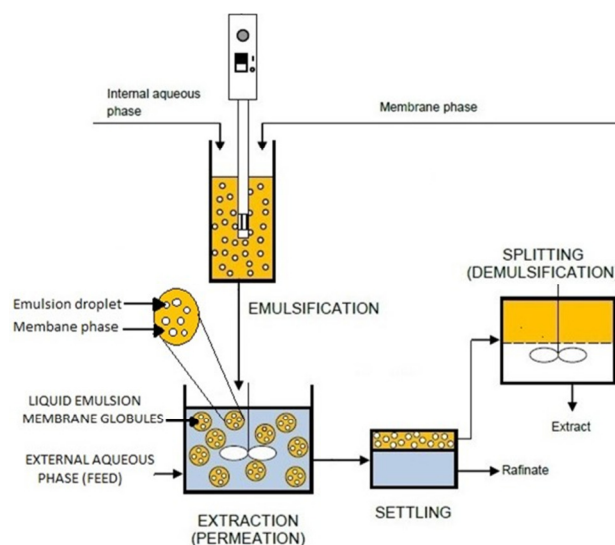


Fig. 1. Flow diagram of batch process ELM.

## III. ANALYSIS AND CALCULATIONS

### A. Extraction Process of CP

The CP concentration in the separated phase solution was measured by the ultraviolet spectrophotometer having 290nm UV wavelength. Subsequently, the percentage of CP extraction was calculated using [20]:

$$\text{Extraction Efficiency (E \%)} = \frac{C_o - C}{C_o} \times 100 \quad (1)$$

where  $C_o$  and  $C$  represent the initial CP concentration and the CP concentration after a specified time interval, respectively.

### B. Stripping

The resulting emulsion was allowed to be naturally separated from the feed solution by gravity, then a demulsification process was achieved by applying centrifugal force on the emulsion. The CP concentration extracted from the internal aqueous solution was measured, and the stripping efficiency was then estimated according to [20]:

$$\text{Stripping (S\%)} = \frac{C_{f,int}}{C_o - C_{f,ext}} \times 100 \quad (2)$$

where  $C_{f,int}$  and  $C_{f,ext}$  represent the final CP concentration in the internal and external phases, respectively.

### C. Membrane Leakage/Breakage

Stabilization is achieved by adding the surfactant, thus the extraction efficiency of any solute is mainly affected by emulsion stability [21]. The breakdown of the emulsion usually occurs after a period when the emulsion is considered unstable, causing a gradual lowering of the removal efficiency along with a significant loss of extracted species [22]. The stability of the liquid membrane is considered the most essential factor affecting the efficiency of solute removal [23]. The tracer method ( $H^+$  ion of the internal phase) was used to evaluate the stability because this tracer is easily detectable by a pH meter. Breakage is the result of tracer transfer from internal-to-external phase solutions. Any alteration in the pH of the feed

phase (i.e. leaking  $H^+$  ions) results from the emulsion breaking (rupture) because of the excretion of the HCl from the internal to the external aqueous phase. The emulsion rupture or breakage ( $\varepsilon$ ) signifies the internal aqueous phase volume leaked to the external feed aqueous phase ratio using [20]:

$$\varepsilon(\%) = \frac{V_s}{V_i} * 100 \quad (3)$$

where  $V_s$  represents the leaked volume of the stripping phase (internal phase) and  $V_i$  represents the initial volume of the internal stripping phase.  $V_s$  was computed using [24]:

$$V_s = V_b^\circ * \frac{10^{-pH_o} - 10^{-pH}}{10^{-pH} - C_{H^+}^{int}} \quad (4)$$

where  $V_b^\circ$  is the external feed phase initial volume,  $pH$  is the external feed pH having been in contact with the organic emulsion,  $pH_o$  is the initial external phase pH, and  $C_{H^+}^{int}$  is the initial acid concentration ( $H^+$ ) of the internal phase solution.

#### IV. RESULTS AND DISCUSSION

##### A. Effect of Internal to Membrane Phase Volume Ratio

The volume ratio of the internal to organic membrane phase must be satisfied to achieve a better transfer rate of the solute within the organic membrane and reduce emulsion interface thickness [25]. The appropriate amount of volume ratio improves solute transport through the organic membrane. Different internal-membrane phase ratios (1:3, 1:2, 1:1, 2:1, 3:1) were selected to study their influence on emulsion stability, in addition to the CP removal efficiency. Figure 2 shows the profiles of emulsion breakage percentage and extraction and stripping efficiency. Increasing the ratio from 1:3 to 1:1, i.e. equalizing membrane phase with internal phase volume, caused a decrease in breakage from 2.44% to 0.82% and increased the extraction and stripping efficiencies of CP to a maximum of 90.3% and 86.2%, respectively. This could be because at a lower volume ratio (1:3), an abundance of oil phase volume led to the creation of a thicker interface and more viscous emulsion that hindered internal phase diffusion into the membrane [20].

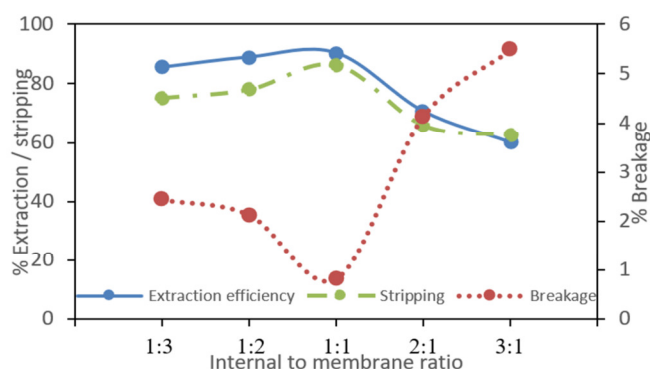


Fig.2. Effect of internal-to-membrane ratio on membrane breakage percentage, extraction, and stripping efficiencies (stirring speed: 300rpm for 5min, external pH = 7, external to emulsion phase volume ratio: 10/1).

Additionally, at low volume ratios, the lowering in extraction efficiency due to a smaller amount of the internal

phase (stripping agent) extracted the solute from the organic membrane phase. A further increase in ratio from 1:1 to 3:1, i.e. increasing the volume of the internal phase related to the organic membrane phase, caused an evident waning in stability. Breakage elevated to 5.48%, while extraction and stripping efficiencies reduced to 61.1% and 62.7%, respectively. This decrease in efficiency occurred because of the insufficient membrane volume that caused a partial entrapment of the internal droplet, hence, droplets tended to leak out of the bubble towards the aqueous feed phase [26]. Similar results were recorded in [27]. Therefore, a 1:1 volume ratio of internal to organic membrane phase was designated and considered the best ratio to achieve the highest extraction efficiency.

##### B. Effect of External to Emulsion Volume Ratio

Experiments were carried out for the following ratios: 250/10, 250/25, 250/50, 250/75, and 250/100, by altering the emulsion volume only while maintaining a fixed volume of external feed aqueous phase. Figure 3 displays the extraction, stripping efficiency, and emulsion breakage percent for different treatment ratios.

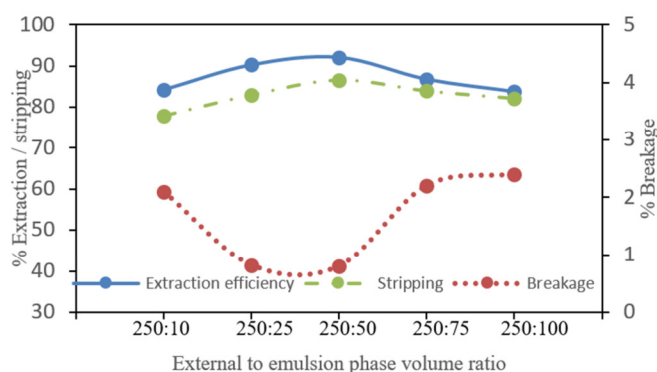


Fig. 3. Effect of external-to-emulsion phase ratio on membrane breakage percentage, extraction, and stripping efficiencies (stirring speed: 300rpm for 5min, internal to membrane ratio: 1/1, external phase pH = 7).

The results show that a decrease in the treatment ratio to 250/50 improved extraction efficiency and decreased breakage from 84.3% and 2.1% to 92.1% and 0.8%, respectively. With decreasing the ratio, the swelling phenomena develop and become fast, and also an occurrence of substantial internal droplets coalescence that will eventually grow in size [20]. This could be explained that due to the volume increase of emulsion, more quantity would be available to add to the progress of CP transport. Furthermore, a big volume of the emulsion favors an enormous number of globules formation in providing a vast surface area for exchange, thus an enhanced rate of solute transfer. Another possible reason could be because of the increase in emulsion holdup that further increases the interfacial area for a better rate of mass transfer [28]. Further decreasing the volume ratio to 250/100 causes a substantial reduction in extraction efficiency to 83.9%. This could be due to the limited interfacial area available for the solute mass transfer rate, owing to the challenges in dispersing the emulsion into the aqueous feed solution due to the emulsion's large

volume that increases its total viscosity causing an adverse effect on the interfacial area [29]. As a result, the 250/50 treatment ratio was chosen as the best ratio.

### C. Effect of Mixing Time

Mixing time can be specified as the time required to achieve maximum extraction and is a necessary factor in the ELM process [30]. Mixing time, also known as contact time, is defined as the period in which the external feed phase stays in direct contact with the emulsion while continuously stirring [31, 32]. Figure 4 shows the effect of mixing time on emulsion breakage and the percentages of extraction and stripping for times ranging from 2 to 20 minutes. When increasing mixing time, CP extraction and stripping efficiency keep increasing, reaching an optimal level at 10min of 93.8% and 94.7%, respectively, while the breakage percentage is reduced to a minimum of 0.73%. Longer contact time would cause the internal phase solution to move toward the feed solution due to the increased emulsion breakage [33, 34]. Therefore, the extracted molecules revert backward to the feed phase, decreasing extraction and stripping efficiencies. Therefore, a mixing time of 10 minutes was chosen as the optimal period to extract CP.

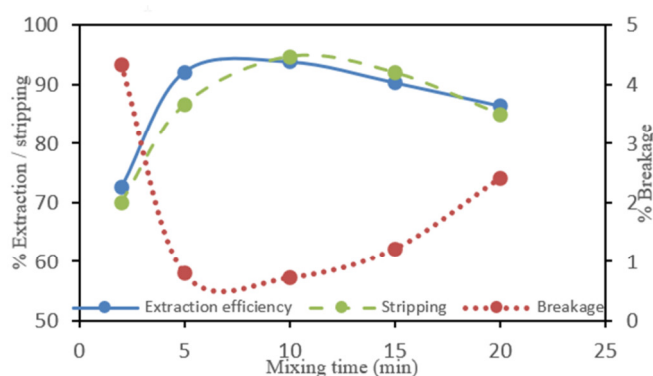


Fig. 4. Effect of mixing time on membrane breakage, and extraction and stripping efficiencies (stirring speed: 300rpm, external phase pH = 7, feed phase to emulsion ratio: 5/1, internal to organic membrane ratio: 1/1).

### D. Effect of External Phase pH

The pH of the external phase is a significant factor for extraction purposes [35]. Additionally, pH also influences membrane stability, considering that it can accelerate the emulsion demulsification process. The tests were carried out at different pH values in basic, neutral, and acidic mediums ranging from pH 3 to 8. Figure 5 presents the effects of external feed pH on extraction percentage, stripping efficiency, and organic emulsion breakage. CP extraction and stripping efficiency are exceedingly influenced by the external phase pH. For a highly acidic solution of pH=3, the extraction and stripping efficiencies were 62.72% and 85%, respectively, while the emulsion breakage was found to be at its peak at 7.41%. This could be attributed to the higher H<sup>+</sup> ion concentrations which reduce the surfactant properties [12], causing a de-stabilization process that leads to a drop in CP extraction efficiency. The higher the initial concentration of H<sup>+</sup>, the lower the extraction efficiency is [36]. When increasing pH

towards a neutral medium (pH=6), increase in both extraction and stripping efficiency occurs, reaching maximum values of 95.3% and 97.2%, respectively, while the breakage percentage is reduced to 0.69% within 10min of stirring time. However, the CP extraction efficiency decreases continuously when further increasing pH, while the breakage increases gradually at 82.8% and 1.89%, respectively, for pH=8. This could be explained by the released protons as the result of an anion exchange reaction [37]. Also, due to the increase in pH, a formation of other species occurs [36]. Therefore, it was appropriate to maintain the pH of the external feed phase at 6.

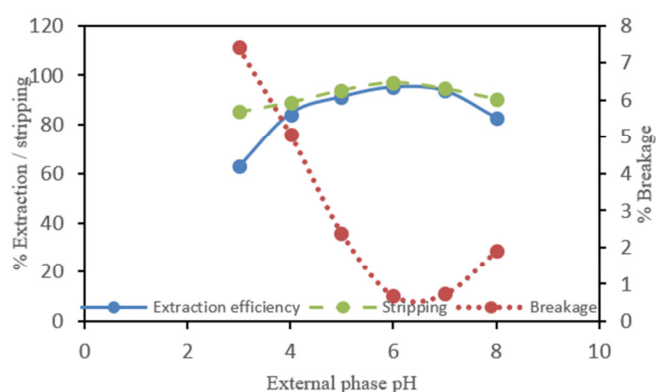


Fig. 5. Effect of external phase pH on membrane breakage, extraction, and stripping efficiency (stirring speed: 300rpm for 10min, feed phase to emulsion ratio: 5/1, internal to organic membrane ratio: 1/1).

### E. Effect of Stirring Speed

A set of experiments was conducted to study the effect of different mixing speeds (200, 250, 300, and 350rpm) on extraction efficiency, stripping efficiency, and membrane breakage, while the other parameters were: mixing time: 10 min, external phase pH=6, external phase to emulsion volume ratio: 5/1, and internal to organic membrane phase ratio: 1/1. Figure 6 displays the experimental results. At lower mixing speeds (200rpm), a minimum membrane breakage value was recorded (0.81%) due to the insufficient shear that resulted from the impeller dispersing emulsion. This could be because lower stirring speeds cause the formation of large globules, increasing the membrane thickness. Although low values of membrane breakage were noted, slow stirring speeds do not usually improve the extraction and stripping efficiencies (89.39% and 93.6%, respectively, for 200rpm stirring speed), since large globules of emulsion minimize the available area for mass transfer [38]. However, using high mixing speeds (over 250rpm) was not beneficial neither to emulsion stability nor to extraction efficiency (2.63% and 84.84%, respectively, for 350rpm). Figure 6 proves that increasing speed above 250rpm leads to the deterioration of the organic membrane. Elevated levels of membrane leakage are predicted at higher mixing speeds since the emulsion can not bear the excessive shear delivered by impellers in addition to the contactor wall [39], causing an easy rupture of the membrane wall, discharging the already extracted solute into the feed phase. As the emulsion was produced from a viscous component, the stability of the emulsion was sustained at 250rpm stirring speed



because of the rigid membrane wall. At 250rpm stirring speed, 0.66% membrane breakage, 96.1% extraction percentage, and 95.7% stripping efficiency were registered. Thus, 250rpm was considered the critical speed of stirring.

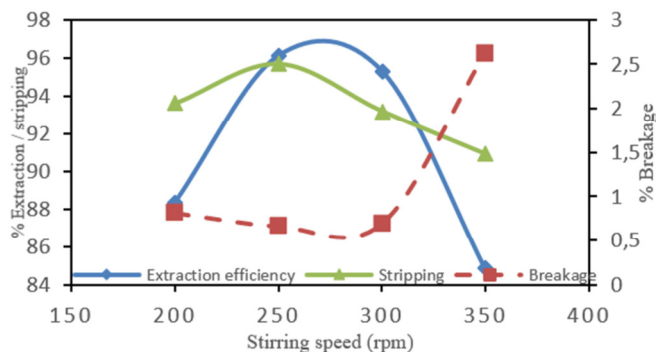


Fig. 6. Effect of stirring speed on membrane breakage, extraction and stripping efficiencies for 10min (external phase pH = 6, external to organic emulsion volume ratio: 5/1, and internal to membrane volume ratio: 1/1).

#### F. F. Evaluation of the Solute (CP) Extraction Kinetics and Estimation of Mass Transfer Coefficient

The extraction kinetics of CP using the ELM method were investigated following the approach in [40, 41], using the first-order equation:

$$\ln\left(\frac{C}{C_0}\right) = -K_{obs} \cdot t \quad (5)$$

where  $t$  represents time in minutes, and  $K_{obs}$  is the rate constant of extraction ( $\text{min}^{-1}$ ), which can be evaluated from the slope of the resulting line from the previous equation representing the  $K_{obs}$  value. Because the obtained value was positive, the extraction process follows the first-order kinetics. The obtained  $K_{obs}$  value was  $0.2173\text{min}^{-1}$ . The total mass transfer coefficient for the ELM system was found by [42]:

$$\frac{1}{K_T} = \frac{1}{K_F} + \frac{1}{K_M} \quad (6)$$

where  $K_T$  is the total mass transfer coefficient (m/s),  $K_F$  is the interfacial reaction rate constant (m/s), and  $K_M$  represents the mass transfer coefficient of the external phase (m/s), estimated by the Skelland-Lee correlation [40] given by:

$$\frac{K_M}{\sqrt{ND}} = 2.932 \cdot 10^{-7} \left( \frac{V_i + V_m}{V_i + V_m + V_e} \right) Re^{1.371} \left( \frac{d}{T} \right)^{0.548} \quad (7)$$

where  $N$  is mixing speed (rpm),  $T$  and  $d$  are the mixing tank and impeller diameter (m), respectively,  $V_i$ ,  $V_e$ , and  $V_m$  denote the internal, external, and membrane phase volumes, and  $D$  represents the species diffusivity (CP) in the organic membrane phase, determined using Wilke-Chang correlation [43] shown in (9).

$$Re = \frac{N d^2 \rho_{ext.}}{\mu_{ext.}} \quad (8)$$

where  $\rho_{ext.}$  is the density ( $\text{kg/m}^3$ ) and  $\mu_{ext.}$  is the viscosity ( $\text{kg/m}\cdot\text{s}$ ). The calculated  $Re$  value was equal to 348577.

$$D = \frac{117.3 \cdot 10^{-8} \cdot (M \cdot \phi)^{0.5} \cdot T_{emp.}}{\mu_m \cdot \phi_c^{0.6}} \quad (9)$$

where  $M$  is the solvent molecular weight (n-hexane=86.18 kg/Kmol),  $T_{emp.}$  is the temperature in kelvin,  $\phi$  is the solvent association factor (n-hexane=1),  $\mu_m$  is the membrane viscosity ( $12.224 \times 10^{-3} \text{Kg/m}\cdot\text{s}$ ), and  $\phi_c$  is the molar volume of the CP evaluated using Schroeder method [44]. The  $\phi_c$  was  $0.399 \text{m}^3/\text{kmol}$ ,  $D$  was  $6.79 \times 10^{-10} \text{m}^2/\text{s}$ , and  $K_M$  was  $2.6 \times 10^{-4} \text{m/s}$ .  $K_F$  was calculated through:

$$\ln\left(\frac{C}{C_0}\right) = -A \cdot K_F \cdot t \quad (10)$$

$K_F$  could be calculated by comparing (10) and (5):

$$K_F = \frac{K_{obs}}{A} \quad (11)$$

where  $A$  represents the emulsion-specific interfacial area, given by [45]:

$$A = \frac{A_i}{V} = \frac{6 \alpha}{d_{32}} \quad (12)$$

where  $A_i$  is the interfacial area of the emulsion droplet,  $V$  is the unit volume of the emulsion,  $\alpha$  represents the water volume fraction, and  $d_{32}$  represents the diameter of the emulsion droplet. Finally, the calculated mass transfer coefficients were:  $K_M = 2.6 \times 10^{-4} \text{m/s}$ ,  $K_F = 3.89 \times 10^{-9} \text{m/s}$ , and  $K_T = 3.89 \times 10^{-9} \text{m/s}$ .

#### V. CONCLUSIONS

This study investigated the stability and extraction efficiency of ELM for removing chlorpyrifos pesticide from wastewater. A minimum emulsion breakage of 0.66% along with the highest removal percentage of 96.1% and 95.7% stripping efficiencies were achieved at the best operating experimental conditions, which were: internal-to-membrane volume ratio: 1:1, external feed solution pH: 6, stirring speed: 250rpm, external phase-to-emulsion ratio: 250:50 at 10min contact time, no requirement of an additional carrier, and the mass transfer coefficients  $K_M$ ,  $K_F$ , and  $K_T$  had values of  $2.6 \times 10^{-4} \text{m/s}$ ,  $3.89 \times 10^{-9} \text{m/s}$ , and  $3.89 \times 10^{-9} \text{m/s}$ , respectively. In general, it can be concluded that ELM is a productive, efficient, and suitable advanced separation method for the treatment of wastewater contaminated with pesticides.

#### REFERENCES

- [1] I. A. Saleh, N. Zouari, and M. A. Al-Ghouti, "Removal of pesticides from water and wastewater: Chemical, physical and biological treatment approaches," *Environmental Technology & Innovation*, vol. 19, Aug. 2020, Art. no. 101026, <https://doi.org/10.1016/j.eti.2020.101026>.
- [2] M. Qurie *et al.*, "Removal of chlorpyrifos using micelle-clay complex and advanced treatment technology," *Desalination and Water Treatment*, vol. 57, no. 33, pp. 15687–15696, Jul. 2016, <https://doi.org/10.1080/19443994.2015.1096836>.
- [3] A. R. Jatoti, A. Q. Jakhani, K. C. Mukwana, A. N. Laghari, and M. M. Tunio, "Study of Physicochemical Properties of Commercial Drinking Bottled Water Brands," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3576–3579, Dec. 2018, <https://doi.org/10.48084/etasr.2173>.
- [4] K. A. Osman, A. I. Al-Humaid, K. N. Al-Redhaiman, and R. A. El-Mergawi, "Safety methods for chlorpyrifos removal from date fruits and its relation with sugars, phenolics and antioxidant capacity of fruits," *Journal of Food Science and Technology*, vol. 51, no. 9, pp. 1762–1772, Sep. 2014, <https://doi.org/10.1007/s13197-012-0693-0>.
- [5] N. Akhtar, M. K. Srivastava, and R. B. Raizada, "Assessment of chlorpyrifos toxicity on certain organs in rat, *Rattus norvegicus*," *Journal of Environmental Biology*, vol. 30, no. 6, pp. 1047–1053, 2009.

- [6] J. V. Peter, T. I. Sudarsan, and J. L. Moran, "Clinical features of organophosphate poisoning: A review of different classification systems and approaches," *Indian Journal of Critical Care Medicine*, vol. 18, no. 11, pp. 735–745, Nov. 2014, <https://doi.org/10.4103/0972-5229.144017>.
- [7] J. F. Risher and H. A. Navarro, "Toxicological Profile for Chlorpyrifos," US Department of Health and Human Services, Atlanta, GA, USA, Sep. 1997.
- [8] Norela Jusoh, Norasikin Othman, Norasikin Othman, and Nur Alina Nasruddin, "Emulsion Liquid Membrane Technology in Organic Acid Purification," *Malaysian Journal of Analytical Sciences*, vol. 20, no. 2, pp. 436–443, 2016.
- [9] S. Muthusaravanan *et al.*, "Optimization and extraction of pharmaceutical micro-pollutant - norfloxacin using green emulsion liquid membranes," *Desalination and Water Treatment*, vol. 156, pp. 238–244, 2019.
- [10] N. Othman *et al.*, "Easy removing of phenol from wastewater using vegetable oil-based organic solvent in emulsion liquid membrane process," *Chinese Journal of Chemical Engineering*, vol. 25, no. 1, pp. 45–52, Jan. 2017, <https://doi.org/10.1016/j.cjche.2016.06.002>.
- [11] F. Al-Ani, Q. Alsahy, and M. Al-Dahhan, "Enhancing Emulsion Liquid Membrane System (ELM) Stability and Performance for the Extraction of Phenol from Wastewater using Various Nanoparticles," *Desalination and Water Treatment*, vol. 210, pp. 180–191, Jan. 2021, <https://doi.org/10.5004/dwt.2021.26547>.
- [12] I. Benabela, A. Benderrag, B. Haddou, J. P. Canselier, and C. Gourdon, "Dye removal with emulsion liquid membrane: experimental design and response surface methodology," *Environmental Technology*, Jun. 2022, <https://doi.org/10.1080/09593330.2022.2091480>.
- [13] S. Zereshki, A. Shokri, and A. Karimi, "Application of a green emulsion liquid membrane for removing copper from contaminated aqueous solution: Extraction, stability, and breakage study using response surface methodology," *Journal of Molecular Liquids*, vol. 325, Mar. 2021, Art. no. 115251, <https://doi.org/10.1016/j.molliq.2020.115251>.
- [14] A. L. Ahmad, A. Kusumastuti, C. J. C. Derek, and B. S. Ooi, "Emulsion liquid membrane for cadmium removal: Studies on emulsion diameter and stability," *Desalination*, vol. 287, pp. 30–34, Feb. 2012, <https://doi.org/10.1016/j.desal.2011.11.002>.
- [15] S. Sheikhi, R. Dehghanzadeh, A. Maryamabadi, and H. Aslani, "Chlorpyrifos removal from aqueous solution through sequential use of coagulation and advanced oxidation processes: By-products, degradation pathways, and toxicity assessment," *Environmental Technology & Innovation*, vol. 23, Aug. 2021, Art. no. 101564, <https://doi.org/10.1016/j.eti.2021.101564>.
- [16] H. Ubaid ur Rahman, W. Asghar, W. Nazir, M. A. Sandhu, A. Ahmed, and N. Khalid, "A comprehensive review on chlorpyrifos toxicity with special reference to endocrine disruption: Evidence of mechanisms, exposures and mitigation strategies," *Science of The Total Environment*, vol. 755, Feb. 2021, Art. no. 142649, <https://doi.org/10.1016/j.scitotenv.2020.142649>.
- [17] Anil Kumar, A. Thakur, and P. S. Panesar, "Role of Operating Process Parameters on Stability Performance of Green Emulsion Liquid Membrane Based on Rice Bran Oil," *Theoretical Foundations of Chemical Engineering*, vol. 55, no. 3, pp. 534–544, May 2021, <https://doi.org/10.1134/S0040579521030118>.
- [18] Mohd Hazare Zairy Mohd Harun, A. L. Ahmad, and L. Rajandram, "Emulsion Liquid Membrane Screening for Ibuprofen Removal from Aqueous Solution," *Journal of Physical Science*, vol. 33, no. 1, pp. 109–122, Apr. 2022, <https://doi.org/10.21315/jps2022.33.1.8>.
- [19] S. S. Díaz, H. Al-Zubaidi, A. C. Ross-Obare, and S. O. Obare, "Chemical reduction of chlorpyrifos driven by flavin mononucleotide functionalized titanium (IV) dioxide," *Physical Sciences Reviews*, vol. 5, no. 11, Nov. 2020, <https://doi.org/10.1515/psr-2020-0007>.
- [20] A. A. Mohammed and N. Q. Jaber, "Stability and performance studies of emulsion liquid membrane on pesticides removal using mixture of Fe<sub>3</sub>O<sub>4</sub> nanoparticles and span80," *Environmental Advances*, vol. 9, Oct. 2022, Art. no. 100294, <https://doi.org/10.1016/j.envadv.2022.100294>.
- [21] A. H. Sulaymon and A. A. Mohammed, "Separation and Hydrodynamic Performance of Air-Kerosene-Water System by Bubble Column," *International Journal of Chemical Reactor Engineering*, vol. 8, no. 1, Mar. 2010, <https://doi.org/10.2202/1542-6580.2211>.
- [22] S. Laguel and M. H. Samar, "Removal of Europium(III) from water by emulsion liquid membrane using Cyanex 302 as a carrier," *Desalination and Water Treatment*, vol. 165, pp. 269–280, 2019.
- [23] N. D. Zaulkiflee, A. L. Ahmad, N. F. Che Lah, and M. M. H. Shah Buddin, "Removal of emerging contaminants by emulsion liquid membrane: perspective and challenges," *Environmental Science and Pollution Research*, vol. 29, no. 9, pp. 12997–13023, Feb. 2022, <https://doi.org/10.1007/s11356-021-16658-5>.
- [24] A. A. Mohammed, H. M. Selman, and G. Abukhanafer, "Liquid surfactant membrane for lead separation from aqueous solution: Studies on emulsion stability and extraction efficiency," *Journal of Environmental Chemical Engineering*, vol. 6, no. 6, pp. 6923–6930, Dec. 2018, <https://doi.org/10.1016/j.jece.2018.10.021>.
- [25] A. A. Mohammed, M. A. Atiya, and M. A. Hussein, "Simultaneous studies of emulsion stability and extraction capacity for the removal of tetracycline from aqueous solution by liquid surfactant membrane," *Chemical Engineering Research and Design*, vol. 159, pp. 225–235, Jul. 2020, <https://doi.org/10.1016/j.cherd.2020.04.023>.
- [26] M. A. Mohammed, W. O. Noori, and H. A. Sabbar, "Application of Emulsion Liquid Membrane Process for Cationic Dye Extraction," *Iraqi Journal of Chemical and Petroleum Engineering*, vol. 21, no. 3, pp. 39–44, Sep. 2020, <https://doi.org/10.31699/IJCPE.2020.3.5>.
- [27] P. Ghorbanpour and M. Jahanshahi, "Removal of zinc by emulsion liquid membrane using lecithin as biosurfactant," *Journal of Dispersion Science and Technology*, vol. 43, no. 14, pp. 2218–2226, Nov. 2022, <https://doi.org/10.1080/01932691.2021.1929287>.
- [28] A. Kumar, A. Thakur, and P. S. Panesar, "A review on emulsion liquid membrane (ELM) for the treatment of various industrial effluent streams," *Reviews in Environmental Science and Bio/Technology*, vol. 18, no. 1, pp. 153–182, Mar. 2019, <https://doi.org/10.1007/s11157-019-09492-2>.
- [29] P. S. Kulkarni and V. V. Mahajani, "Application of liquid emulsion membrane (LEM) process for enrichment of molybdenum from aqueous solutions," *Journal of Membrane Science*, vol. 201, no. 1, pp. 123–135, May 2002, [https://doi.org/10.1016/S0376-7388\(01\)00720-7](https://doi.org/10.1016/S0376-7388(01)00720-7).
- [30] H. Bouzidi, L. Otmani, R. Doufnoune, L. Zerroual, and D. Benachour, "Influence of Membrane Type on Some Electrical Properties of a Single Microbial Fuel Cell," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8492–8499, Jun. 2022, <https://doi.org/10.48084/etasr.4813>.
- [31] A. Benderrag, M. Djellali, B. Haddou, M. Daaou, and B. Bounaceur, "Experimental design and RSM on the recovery of Ni (II) ions by ELM using TX-100 as a biodegradable surfactant," *Environmental Technology*, vol. 43, no. 3, pp. 386–401, Jan. 2022, <https://doi.org/10.1080/09593330.2020.1791967>.
- [32] A. L. Ahmad, M. H. Z. Mohd Harun, M. K. Akmal Jasni, and N. D. Zaulkiflee, "Removal of Ibuprofen at Low Concentration Using a Newly Formulated Emulsion Liquid Membrane," *Membranes*, vol. 11, no. 10, Oct. 2021, Art. no. 740, <https://doi.org/10.3390/membranes11100740>.
- [33] H. M. Salman and A. A. Mohammed, "Removal of-Copper Ions-from Aqueous Solution Using Liquid-Surfactant-Membrane Technique," *Iraqi Journal of Chemical and Petroleum Engineering*, vol. 20, no. 3, 2019.
- [34] A. A. Mohammed, "Removal of Emulsified Paraffine from Water: Effect of Bubble Size and Particle Size on Kinetic of Flotation," *Iraqi Journal of Chemical and Petroleum Engineering*, vol. 8, no. 3, pp. 1–5, Sep. 2007.
- [35] E. Fouad, F. Ahmad, and K. Abdelrahman, "Optimization of Emulsion Liquid Membrane for Lead Separation from Aqueous Solutions," *Engineering, Technology & Applied Science Research*, vol. 7, no. 5, pp. 2068–2072, Oct. 2017, <https://doi.org/10.48084/etasr.1390>.
- [36] A. A. Mohammed, M. A. Atiya, and M. A. Hussein, "Studies on membrane stability and extraction of ciprofloxacin from aqueous solution using pickering emulsion liquid membrane stabilized by magnetic nano-Fe<sub>2</sub>O<sub>3</sub>," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 585, Jan. 2020, Art. no. 124044, <https://doi.org/10.1016/j.colsurfa.2019.124044>.



- [37] J. Liang, H. Li, J. Yan, and W. Hou, "Demulsification of Oleic-Acid-Coated Magnetite Nanoparticles for Cyclohexane-in-Water Nanoemulsions," *Energy & Fuels*, vol. 28, no. 9, pp. 6172–6178, Sep. 2014, <https://doi.org/10.1021/ef501169m>.
- [38] R. A. Kumbasar, "Selective separation of chromium (VI) from acidic solutions containing various metal ions through emulsion liquid membrane using trioctylamine as extractant," *Separation and Purification Technology*, vol. 64, no. 1, pp. 56–62, Nov. 2008, <https://doi.org/10.1016/j.seppur.2008.08.005>.
- [39] A. L. Ahmad, A. Kusumastuti, C. J. C. Derek, and B. S. Ooi, "Emulsion liquid membrane for heavy metal removal: An overview on emulsion stabilization and destabilization," *Chemical Engineering Journal*, vol. 171, no. 3, pp. 870–882, Jul. 2011, <https://doi.org/10.1016/j.cej.2011.05.102>.
- [40] H. P. Kohli, S. Gupta, and M. Chakraborty, "Stability and performance study of emulsion nanofluid membrane: A combined approach of adsorption and extraction of Ethylparaben," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 579, Oct. 2019, Art. no. 123675, <https://doi.org/10.1016/j.colsurfa.2019.123675>.
- [41] M. Raji, H. Abolghasemi, J. Safdari, and A. Kargari, "Response Surface Optimization of Dysprosium Extraction Using an Emulsion Liquid Membrane Integrated with Multi-Walled Carbon Nanotubes," *Chemical Engineering & Technology*, vol. 41, no. 9, pp. 1857–1870, 2018, <https://doi.org/10.1002/ceat.201700351>.
- [42] B. Reynolds, J. B. Richards, K. Horn, and K. Karraker, "Delay discounting and probability discounting as related to cigarette smoking status in adults," *Behavioural Processes*, vol. 65, no. 1, pp. 35–42, Jan. 2004, [https://doi.org/10.1016/S0376-6357\(03\)00109-8](https://doi.org/10.1016/S0376-6357(03)00109-8).
- [43] R. E. Treybal, *Mass-transfer operations*, 3rd ed. Malaysia: McGraw-Hill, 1980.
- [44] B. E. Poling, J. M. Prausnitz, and J. P. O'Connell, *Properties of Gases and Liquids*. McGraw-Hill Education, 2001.
- [45] V. Karcher, F. A. Perrechil, and A. C. Bannwart, "Interfacial Energy during the Emulsification of Water-in-Heavy Crude Oil Emulsions," *Brazilian Journal of Chemical Engineering*, vol. 32, pp. 127–137, Mar. 2015, <https://doi.org/10.1590/0104-6632.20150321s00002696>.

# Pantograph Catenary Contact Force Regulation Based on Modified Takagi-Sugeno Fuzzy Models

**Nguyen Van Hai**

Department of Electrical Engineering, University of Transport and Communications, Vietnam  
haiktd@utc.edu.vn  
(corresponding author)

**Nguyen Van Tiem**

Department of Cybernetics, University of Transport and Communications, Vietnam  
nguyenvantiem@utc.edu.vn

**Le Hung Lan**

Department of Cybernetics, University of Transport and Communications, Vietnam  
lehunglan@utc.edu.vn

**Vo Thanh Ha**

Department of Cybernetics, University of Transport and Communications, Vietnam  
vothanhha.ktd@utc.edu.vn

*Received: 27 October 2022 | Revised: 9 November 2022 | Accepted: 10 November 2022*

## ABSTRACT

In this paper, a new contact force control technique, based on the modified Takagi-Sugeno model and the parallel distributed compensation concept is developed to suppress vibrations between the pantograph and the catenary by regulating the contact force to a reference value, thereby achieving stable current collection. The proposed method uses simple and standard PID and modified Takagi-Sugeno fuzzy controllers. The two controllers guarantee the designed system's robust stability. Furthermore, based on a simplified pantograph-catenary system model, the comparative simulation results show that variations of the contact force can be almost attenuated. As a result, the system stability is guaranteed, and the performance robustness is verified.

**Keywords-***Takagi-Sugeno; pantograph; catenary; stability; modified T-S-Fuzzy*

## I. INTRODUCTION

Improving the current collection quality from the overhead line is one of the most challenging difficulties in high-speed rails. The pantograph, an articulated suspension device, ensures the overhead line's current collection, but its interaction with the contact wire results in oscillations. This system must guarantee a good quality of the existing collection, it must thus be compensated. This issue mainly affects expensive railroads because poor current collection results in performance constraints, high maintenance cost, and inconsistent service. The pantograph-catenary system, a dynamic couple system formed by the interactions of the pantograph and catenary, is often responsible for directly affecting the quality of electric transmission (PCS). Through contact forces, the pantograph and the catenary are impacted by one another. The overhead

wire's varying stiffness along the span [1-4] is a significant cause of vibration. The pantograph will vibrate and the contact force will fluctuate as it moves along the overhead wire due to the stiffness change those results in a periodic excitation. Additionally, while a moving panhead travels along the overhead wire, a flexural wave motion is created in the wire, affecting the contact force and motion pantograph. An actuator, an essential component of the PCS, applies the dynamic uplift force to the pantograph frame within the context of the active vibration control of a railway pantograph to significantly reduce the contact force's unpredictable fluctuation. Many control algorithms have been developed (simple PI controllers [5], robust controllers [6], adaptive controllers [7], backstepping controllers [8], predictive model controllers [9], intelligent controllers [10-12], and fuzzy controllers [13]) for

producing adequate uplift force to regulate the fluctuation of the contact force. To manage the oscillations brought on by the unpredictable stiffness and to regulate the contact force taken into account as a pre-specified reference value, an intelligent contact force regulator, including MZL algorithms [14-15], is given in [10]. The vibration of the contact force caused by the catenary's time-varying stiffness can be successfully suppressed, and the comparative simulation results demonstrate that the suggested method is more effective than the previous fuzzy algorithms. However, although intelligent or fuzzy control are promising techniques, there are still issues with constructing ambiguous regulations and ensuring system stability. The Parallel Distributed Compensation (PDC) idea and the modified Takagi-Sugeno (T-S) model are the foundations of the contact force control approach proposed in this work. Furthermore, the devised technique enables the execution of straightforward and joint PID controllers to ensure the robust stability of the proposed system.

## II. PROBLEM FORMULATION

The model of the pantograph – catenary system is exhibited in Figure 1. It is represented by a 2-degree of freedom mechanical system as shown in Figure 2. The motion equation of the PCS model can be written as [12]:

$$\begin{aligned} m_h \ddot{x}_h + c_h(\dot{x}_h - \dot{x}_f) + k_h(x_h - x_f) + k_{pan}(x_h - x_{cat}) &= 0 \\ k_{pan}(x_{cat} - x_h) + k_{cat}x_{cat} &= 0 \end{aligned} \quad (1)$$

$$m_f \ddot{x}_f - c_h(\dot{x}_h - \dot{x}_f) + c_f \dot{x}_f - k_h(x_h - x_f) = u$$

or:

$$m_h \ddot{x}_h + c_h(\dot{x}_h - \dot{x}_f) + k_h(x_h - x_f) + k_0 x_h = 0 \quad (2)$$

$$m_f \ddot{x}_f - c_h(\dot{x}_h - \dot{x}_f) + c_f \dot{x}_f - k_h(x_h - x_f) = u$$

where:

$$k_0 = \frac{k_{pan} k_{cat}}{k_{pan} + k_{cat}} \quad (3)$$

By taking the Laplace transformation of (2), one gets:

$$\begin{aligned} [m_h s^2 + c_h s + (k_h + k_0)] X_h(s) - (c_h s + k_h) X_f(s) &= 0 \\ -(c_h s + k_h) X_h(s) + [m_f s^2 + (c_f + c_h) s + k_h] X_f(s) &= U(s) \end{aligned} \quad (4)$$

where the term  $s$  is defined as:  $X_h(s) = \frac{F(s)}{k}$ . For that, the plant transfer function  $P(s)$  can be computed as follows:

$$P(s) = \frac{F(s)}{U(s)} = \frac{k(c_h s + k_h)}{M(s)} \quad (5)$$

where:

$$\begin{aligned} M(s) &= [m_f s^2 + (c_f + c_h) s + k_h] [m_h s^2 + c_h s + (k_h + k_0)] \\ &- (c_h s + k_h)^2 = m_f m_h s^4 + [m_f c_h + (c_f + c_h) m_h] s^3 \\ &+ [m_f (k_h + k_0) + m_h k_h + c_h c_f] s^2 \\ &+ [k_0 (c_f + c_h) + k_h c_f] s + k_h k_0 \end{aligned} \quad (6)$$

It should be noted that the dynamic uplift force  $u(t)$  is designed to regulate the contact force  $F(t)$  for the prescribed constant value  $F_r(t)=100N$ . If the time-varying stiffness of the catenary influenced by the stiffness variation coefficient  $\alpha$ , the operational speed of the train  $V$ , and the length in a span  $L$  are considered, the contact force  $f(t) = kx_h(t)$  will be rewritten as:

$$f(t) = k_0 \left( 1 + \alpha \cos\left(\frac{2\pi V}{L} t\right) \right) x_h(t) \quad (7)$$

where:

$$k = k_0 \left( 1 + \alpha \cos\left(\frac{2\pi V}{L} t\right) \right) \quad (8)$$

Clearly, the term of  $k$  can be considered as an uncertain interval parameter  $k \in [k_{min}, k_{max}]$ , which can be expressed as:

$$k_{min} = k_0(1 - \alpha), \quad k_{max} = k_0(1 + \alpha) \quad (9)$$

where  $m_h$  is the head mass,  $m_f$  the frame mass,  $x_h$  and  $x_f$  the head displacement,  $k_h$  the head stiffness of suspension,  $c_h$  the head viscous damping coefficient,  $c_f$  the frame viscous damping coefficient,  $k_{pan}$  the pantograph shoe stiffness of suspension,  $k_{cat}$  the catenary stiffness of suspension, and  $u$  the uplift force.

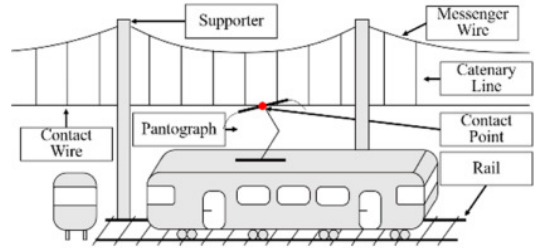


Fig. 1. Pantograph-catenary system components.

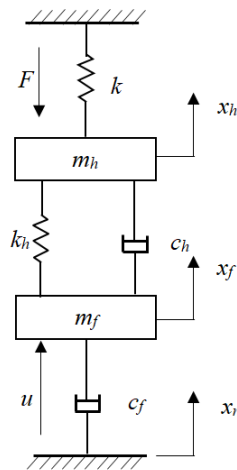


Fig. 2. The PCS 4<sup>th</sup> order model.

In this paper, stable robust controllers are developed to suppress the vibration resulting from parameters (the stiffness-

variation coefficient  $a$ , the operational speed of train  $V$  and the length in a span  $L$ ). The first solution is the robust PI controller, and the second one is the T-S fuzzy controller. Both solutions are based on the most common industrial controller, the PID controller. The parameters of the PCS used in the simulation [12] are:  $m_h = 9.1\text{kg}$ ,  $m_f = 17.2\text{kg}$ ,  $k_h = 7 \times 10^3 \text{ N/m}$ ,  $c_h = 130\text{Ns/m}$ ,  $c_f = 30\text{Ns/m}$ ,  $k_{cat} = 1.535 \times 10^6 \text{ N/m}$ ,  $\alpha = 0.3$ ,  $k_{pan} = 8.23 \times 10^4 \text{ N/m}$ ,  $V = 70\text{km/h}$ ,  $L = 80\text{m}$ , and for the given parameters, the following values are respectively computed:  $k_{min} = 54678$ ,  $k_{max} = 101550$ .

### III. PARTITION METHOD

The PI controller is given as:

$$C(s) = K_p + \frac{K_I}{s} \quad (10)$$

Then the closed loop transfer functions of (5) and (10) can be written as:

$$W(s) = \frac{C(s)P(s)}{1 + C(s)P(s)} = \frac{k(K_p s + K_I)(c_h s + k_h)}{sM(s) + k(K_p s + K_I)(c_h s + k_h)} \quad (11)$$

where  $k \in [k_{min}, k_{max}]$ .

The closed loop system characteristic equation is defined as:

$$H(s) = sM(s) + k(K_p s + K_I)(c_h s + k_h) \quad (12)$$

where:

$$M(j\omega) = R_M(\omega) + jI_M(\omega) \quad (13)$$

Based on the robust D-partition technique [15], it can be seen that the stability region boundary in the  $K_p - K_I$  space is the solution of the following equation:

$$H(j\omega) = [k(k_h K_I - c_h \omega^2 K_p) - \omega I_M(\omega)] + j[k\omega(K_p k_h + K_I c_h) + \omega R_M(\omega)] \quad (14)$$

Equation (14) is equivalent to:

$$-kc_h \omega^2 K_p + k k_h K_I = \omega I_M(\omega) \quad (15)$$

where the functions are respectively expressed as:

$$K_p = -\frac{\omega c_h I_M(\omega) + k_h R_M(\omega)}{k(c_h^2 \omega^2 + k_h^2)} \quad (16)$$

$$K_I = -\frac{\omega[c_h \omega R_M(\omega) - k_h I_M(\omega)]}{k\omega(c_h^2 \omega^2 + k_h^2)}$$

In (16), when  $\omega$  varies from 0 to  $\infty$ , the obtained values of  $K_p$  and  $K_I$  are considered as the stability region boundary in the form of the single curve for a specific value of  $k$ . This boundary will be the group of curves when  $k$  varies within the interval  $k \in [k_{min}, k_{max}]$ . The robust stability region for the given PCS is shown in Figure 3.

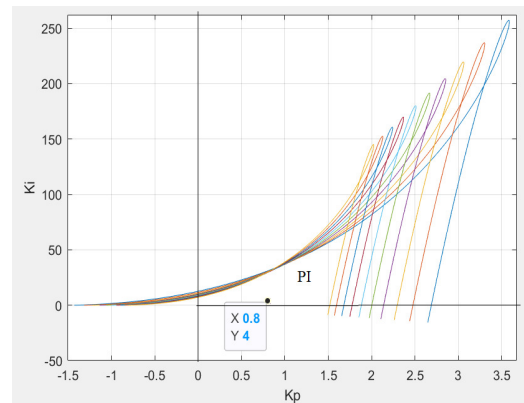


Fig. 3. Robust stability region of the PI controller.

### IV. MODIFIED T-S-FUZZY CONTROLLER

The T-S model is one of the most well-liked modeling frameworks among the different fuzzy modeling topics [16-17]. Since it can approximate any smooth nonlinear control system, the T-S fuzzy model is considered a universal approximator. Additionally, several feedback control strategies may be used with T-S fuzzy models. The PDC idea is the foundation for the control law that is most frequently employed, and for this concept, the fuzzy controller and T-S fuzzy model share the same fuzzy rules and sets [18]. A linear controller is created for each local linear plant by the PDC principle to provide stability and the desired performance of the local linear closed loop system, compensating a corresponding conclusion in the rules of the T-S plant model. The final nonlinear control is an amorphous amalgamation of the many regulations and control operations. Finding a combined Lyapunov function that can satisfy all fuzzy subsystems results in a sufficient condition that guarantees the entire system's stability. To determine the well-known Lyapunov role, Linear Matrix Inequalities (LMIs) must be solved numerically. The primary downsides of the PDC design technique are the complexity of the computations without a solution guarantee and the difficulty in locating the combined Lyapunov function for the numerous fuzzy subsystems. To address these problems, a novel fuzzy logic controller with two consequents in each rule—a numerator component and a denominator part—is presented in [19].

Additionally, the numerator and denominator coefficients are calculated so that the total closed-loop system acts as a linear system. Further, the notion above is expanded for continuous systems in the suggested method in [14]. The closed-loop T-S fuzzy control system acts as a polytope of linear systems based on the recently developed technique. Instead of employing a problematic method to find a joint Lyapunov function as in the previous approaches, the system stability may be readily tested using simple, graphical solid stability criteria.

First of all, the interval  $[k_{min}, k_{max}]$  is divided into  $r$  overlapping subintervals. The linear subintervals are defined as the fuzzy sets and form the universe of discourse. The  $r+2$  triangular membership functions corresponding to the subintervals of the following form are presented in Figure 4.

For each  $i$ -th interval, a represented value  $k_i$  is chosen. For that, a fuzzy variable  $M_i$  with the related membership functions  $\mu_i$  is defined. Furthermore, the two interval boundaries of  $k_{min}$  and  $k_{max}$  are also added. For the correspondence of the  $i$ -th interval, the local plant transfer function can be obtained.

$$P_i(s) = \frac{F(s)}{U(s)} = \frac{k_i(c_h s + k_h)}{M(s)}, \quad i = 1, \dots, r+2 \quad (17)$$

where the function  $M(s)$  is defined as:

$$M(s) = m_f m_h s^4 + [m_f c_h + (c_f + c_h) m_h] s^3 + [m_f (k_h + k_0) + m_h k_h + c_h c_f] s^2 + [k_0 (c_f + c_h) + k_h c_f] s + k_h k_0 \quad (18)$$

We also have:

$$c_1^i = \frac{k_i c_h}{m_h m_f}, \quad c_0^i = \frac{k_i k_h}{m_h m_f}, \quad d_3^i = \frac{m_f c_h + (c_f + c_h) m_h}{m_h m_f}, \quad d_2^i = \frac{[m_f (k_h + k_0) + m_h k_h + c_h c_f]}{m_h m_f}, \quad d_1^i = \frac{k_0 (c_f + c_h) + k_h c_f}{m_h m_f}, \quad d_0^i = \frac{k_h k_0}{m_h m_f} \quad (19)$$

The local transfer function in (17) will be then expressed as:

$$P_i(s) = \frac{F(s)}{U(s)} = \frac{c_1^i s + c_0^i}{s^4 + d_3^i s^3 + d_2^i s^2 + d_1^i s + d_0^i} \quad (20)$$

or in the form of the following differential equation:

$$f^{(5)}(t) = -d_3^i f^{(4)}(t) - d_2^i f^{(3)}(t) - d_1^i f''(t) - d_0^i f'(t) + c_1^i u''(t) + c_0^i u'(t) \quad (21)$$

For each above local plant, the local PI controller is used:

$$C_i(s) = \frac{U(s)}{E(s)} = K_p^i + \frac{K_I^i}{s} = \frac{K_p^i s + K_I^i}{s}, \quad i = 1, \dots, r+2 \quad (22)$$

or, in the form of the following differential equation:

$$c_1^i u''(t) + c_0^i u'(t) = -K_p^i c_1^i f''(t) - (K_p^i c_0^i + K_I^i c_1^i) f'(t) - K_I^i c_0^i f(t) + K_I^i c_1^i f_r \quad (23)$$

The local PI controller parameters can be finally chosen by using the proposed robust D-partition and some optimal performance criterions. By using the modified fuzzy T-S model [14-15], the  $i$ -th fuzzy IF-THEN rule for the describing plant is presented as follows:

1) *Plant Model Rule  $R_i$* :  $i = 1, \dots, r+2$

IF  $k(t)$  is  $M_i$  THEN

$$f^{(5)}(t) = -d_3^i f^{(4)}(t) - d_2^i f^{(3)}(t) - d_1^i f''(t) - d_0^i f'(t) + c_1^i u''(t) + c_0^i u'(t) \quad (24)$$

where  $R_i$  is denoted as the  $i$ -th fuzzy inference rule,  $r$  is the number of inference rules, and  $M_i$  ( $i \in \Omega$ ) is the fuzzy set.

The PDC design approach creates a control rule based on the same presumption with each rule in the T-S fuzzy plant model. A numerator component and a denominator part of the control signal are the two consequents each control rule of the fuzzy logic controller has in its consequent section [19] part and a denominator part of the control signal [20-21].

2) *Control Rule  $R_p$* ,  $i = 1, \dots, r+2$

IF  $k(t)$  is  $M_i$  THEN

$$\begin{aligned} num u'(t) &= -K_p^i c_1^i f''(t) - (K_p^i c_0^i + K_I^i c_1^i) f'(t) - K_I^i c_0^i f(t) + K_I^i c_0^i f_r - c_1^i u''(t) \\ den u'(t) &= c_0^i \end{aligned} \quad (25)$$

where  $num u'(t)$  is the numerator of  $u'(t)$  and  $den u'(t)$  is the denominator of  $u'(t)$ .

The closed loop system characteristic function is:

$$s^5 + \sum_{i=1}^r h_i(k) d_3^i s^4 + \sum_{i=1}^r h_i(k) d_2^i s^3 + \sum_{i=1}^r h_i(k) (d_1^i + K_p^i c_1^i) s^2 + \sum_{i=1}^r h_i(k) (d_0^i + K_p^i c_0^i + K_I^i c_1^i) s + \sum_{i=1}^r h_i(k) K_I^i c_0^i = 0 \quad (26)$$

where  $h_i(k)$  is the normalized membership function. It should be noted that the normalized membership functions satisfy the following convex sum property:

$$0 \leq h_i(k) \leq 1, \quad i = 1, \dots, r+2, \quad \sum_{i=1}^{r+2} h_i(k) = 1 \quad (27)$$

The left-hand side of (27) is the characteristic polynomial, and it is actually the polytope of polynomials:

$$H(s) = \sum_{i=1}^{r+2} h_i(z) H_i(s), \quad (28)$$

$$0 \leq h_i(z) \leq 1, \quad i = 1, \dots, r+2, \quad \sum_{i=1}^{r+2} h_i(k) = 1$$

where:

$$H_i(s) = s^5 + d_3^i s^4 + d_2^i s^3 + (d_1^i + K_p^i c_1^i) s^2 + (d_0^i + K_p^i c_0^i + K_I^i c_1^i) s + K_I^i c_0^i \quad (29)$$

Therefore, system stability analysis can be conducted by using the robust stability criteria [18-19], derived in [14].



## 3) Theorem

The fuzzy closed loop system (21), (22) is stable if and only if:

- The polynomials  $H_i(s)$ ,  $i=1, \dots, r+2$  are stable.
- All  $\frac{(r+1)(r+2)}{2}$  plots,  $z_{ij}(\omega) = \frac{H_i(j\omega)}{H_j(j\omega)}$ ,  $i, j=1, \dots, r+2$ ,  $i < j$  do not intersect the negative real semi axis.

## V. SIMULATION RESULTS

The system simulation model simulated in MATLAB R2019a includes three parts, i.e. the open-loop system (no control), one system using the robust PI controller, and one using the new T-S fuzzy controller. The control construct of contact force regulation of the pantograph catenary based on modified T-S Fuzzy models is shown in Figure 4.

## 1) Case 1. Robust PI Controller

For an example of the given PCS, the robust stability region of PI controller is found in Figure 3. The typical PI controller

parameters in this region can be determined as:  $K_p = 0.8$ ,  $K_I = 4$

## 2) Case 2. Modified T-S Fuzzy Controller

First, the interval  $[k_{min}, k_{max}]$  is divided into  $r=4$  overlapping subintervals with triangular membership functions as pointed out in Figure 5.

By using the proposed D-partition method, the robust stability regions for each subinterval can be constructed, and the local PI controller parameters can be chosen inside those regions:

$$K_p^1 = K_p^2 = K_p^3 = K_p^4 = K_p^5 = K_p^6 = 0.1$$

$$K_I^1 = 4.968, K_I^2 = 4.241, K_I^3 = 3.7, K_I^4 = 3.281, K_I^5 = 2.947, K_I^6 = 2.675.$$

Next, the two conditions in the given theorem are verified by using 6 Mikhailov's graphs  $H_i(j\omega)$  (Figure 6) and 15  $z_{ij}(j\omega)$  graphs as shown in Figure 7.

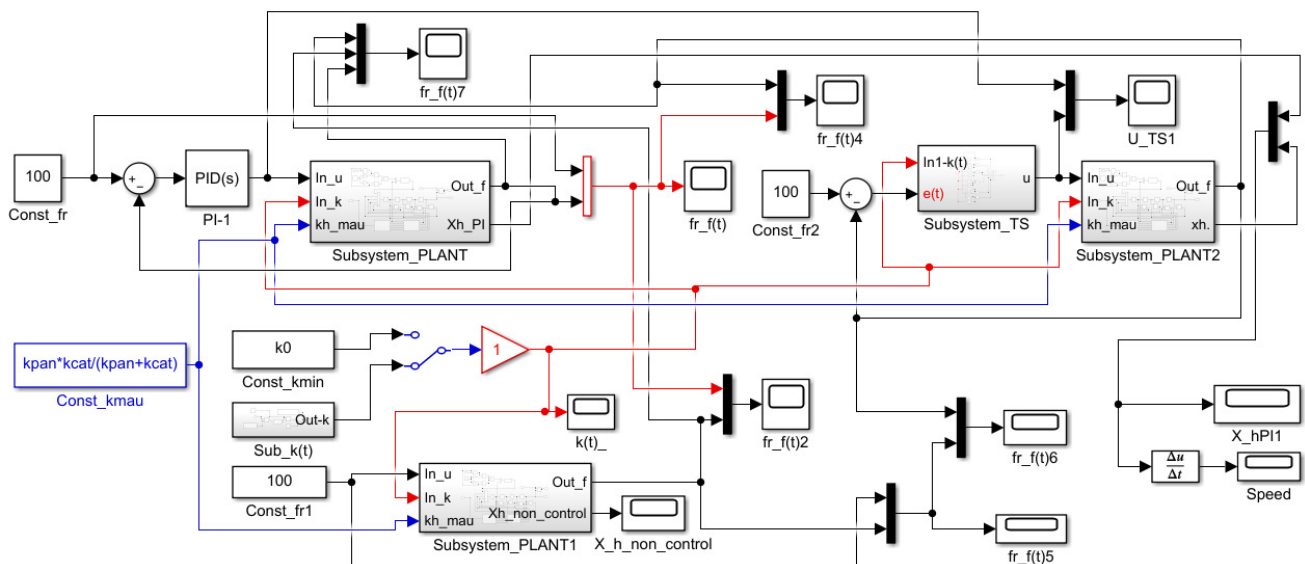


Fig. 4. The control construct of contact force regulation of pantograph catenary based on modified Takagi-Sugeno Fuzzy model.

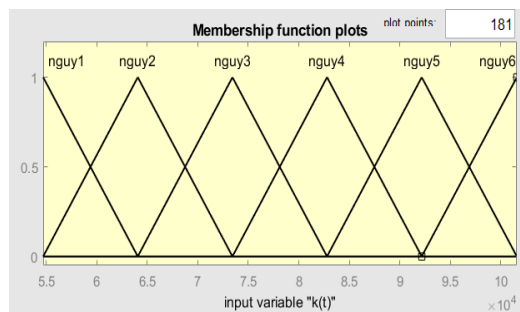


Fig. 5. The triangular membership functions.

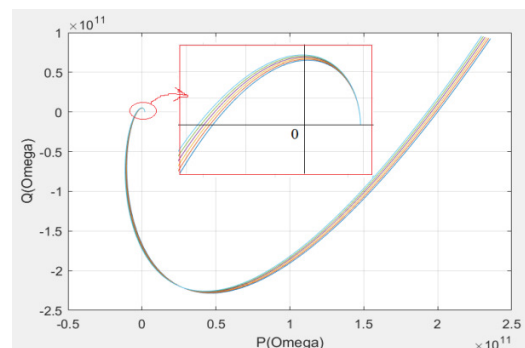


Fig. 6. Mikhailov's graphs.

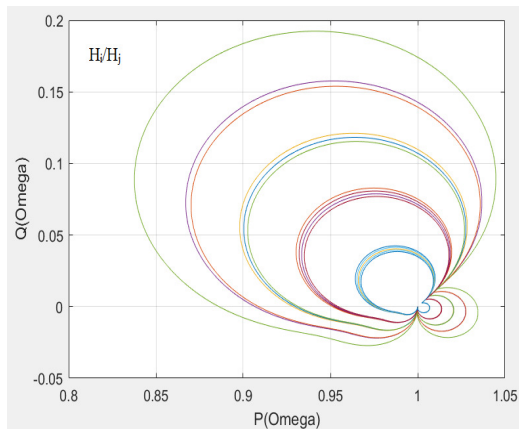
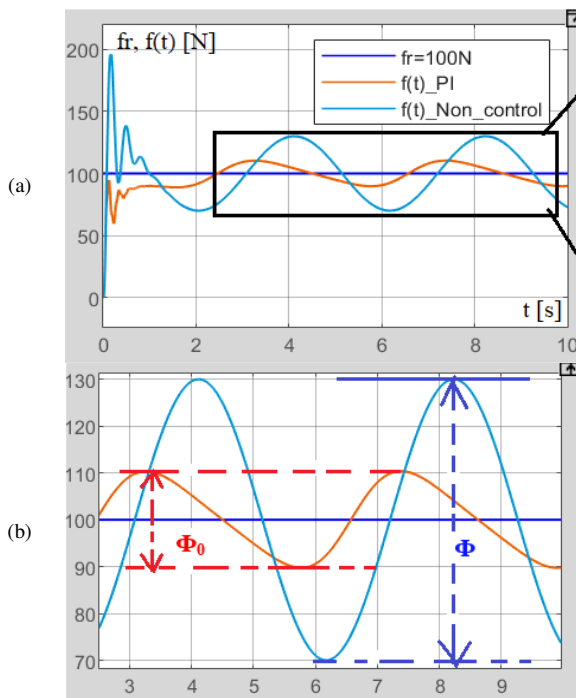
Fig. 7. Graphs  $z_{ij}(j\omega)$ .

Fig. 8. Contact force regulation performance of the robust PI algorithm.

Now let us consider a closed-loop control system, which consists of the plant described by (5) and the fuzzy controller of (25). In order to verify the regulation performance, the two control algorithms proposed in this paper have been analyzed and compared with passive control and MZL algorithm in [12], or each with other using the vibration suppression efficiency (VSE) [12]:

$$VSE \triangleq \left(1 - \frac{\Phi}{\Phi_0}\right) 100\% \quad (27)$$

where  $\Phi$  is the amplitude of the steady contact force value and  $\Phi_0 = 2\alpha f_r(t)$  is the amplitude of the contact force by the traditional passive control.

By applying the robust PI controller, the graph of contact force shown in Figure 8 is defined with  $VSE=65.6617\%$ . When using the proposed modified T-S fuzzy controller, the graph of contact force presented in Figure 9 is determined with  $VSE=97.68\%$ .

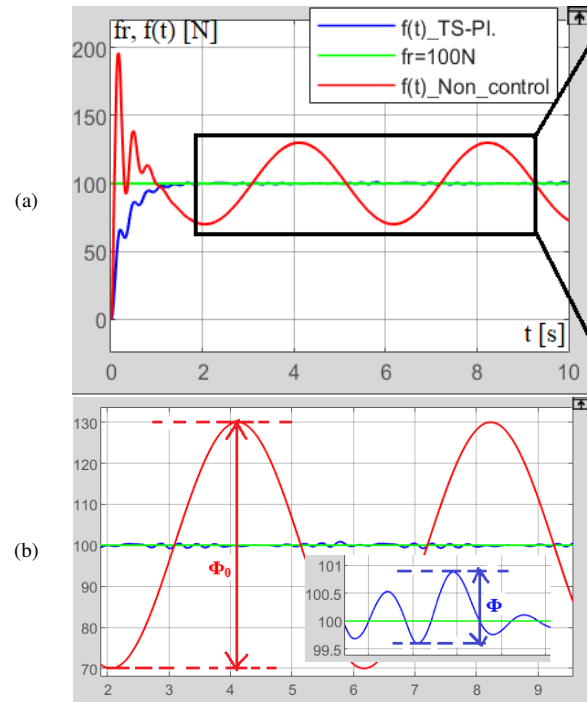
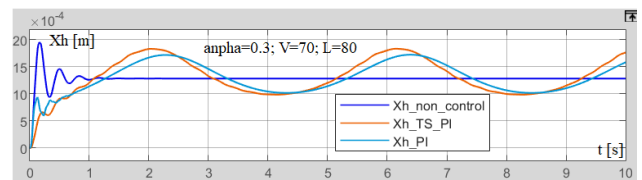
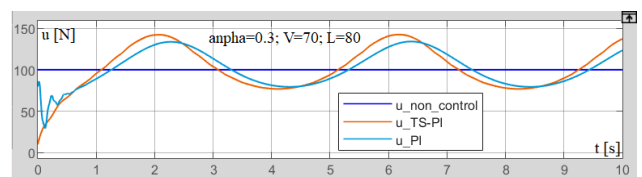


Fig. 9. The contact force regulation performance for the T-S fuzzy algorithm.

By comparing with the best MZL algorithm in [12], where  $VSE=83.59\%$ , it can be found that the contact force regulation performance of the robust PI algorithm is the lowest one, and the contact force regulation performance of the modified T-S fuzzy algorithm is the best one. It should be noted that the time to obtain the steady state is the same, about 1.7s for all three control algorithms. The vertical position of the pantograph head  $x_h(t)$  and uplift force  $u(t)$  are illustrated in Figures 10-11.

Fig. 10. Vertical position of the pantograph head  $x_h(t)$ .Fig. 11. Uplift force  $u(t)$ .

It can be seen that the T-S fuzzy algorithm uses a little larger uplift force than the robust PI algorithm. Some simulations were carried out to test the robustness of the proposed control algorithms. The cases in Figures 12-16 are first considered for the variations of the speed of train  $V$ , and the length in a span  $L$ .

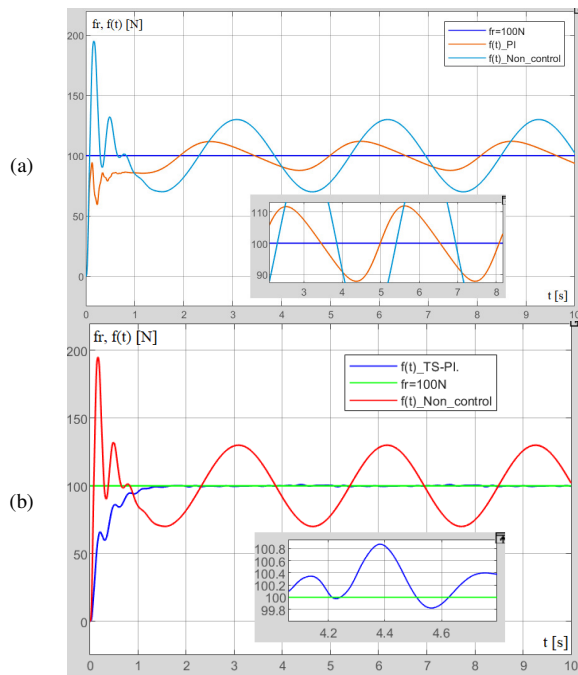


Fig. 12. Contact force relation performance for  $V=70$ ,  $L=60$ : (a) Robust PI controller, (b) modified T-S fuzzy controller.

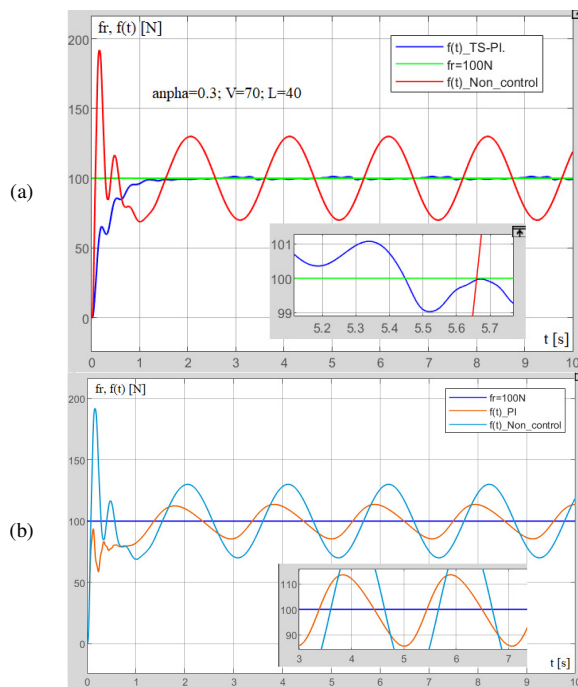


Fig. 13. Contact force regulation performance for  $V=70$ ,  $L=40$ : (a) Robust PI controller, (b) modified T-S fuzzy controller.

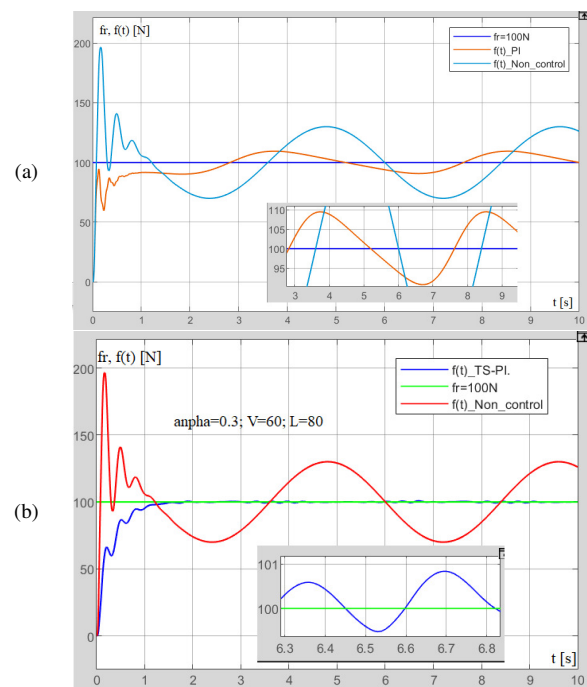


Fig. 14. Contact force relation performance for  $V=60$ ,  $L=80$ : (a) Robust PI controller, (b) modified T-S fuzzy controller.

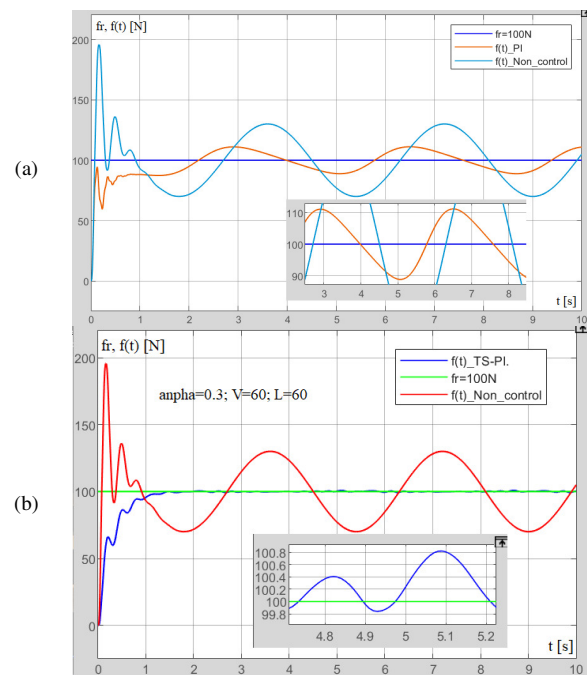


Fig. 15. Contact force regulation performance for  $V=60$ ,  $L=60$ : (a) Robust PI controller, (b) modified T-S fuzzy controller.

The obtained results are given in Table I. They show that when the speed  $V$  and the length  $L$  vary, the vibration suppression efficiency for the modified T-S fuzzy controller is very high and stable. The reason is that although these parameters can vary, the boundary of  $k \in [k_{min}, k_{max}]$  is not changed.

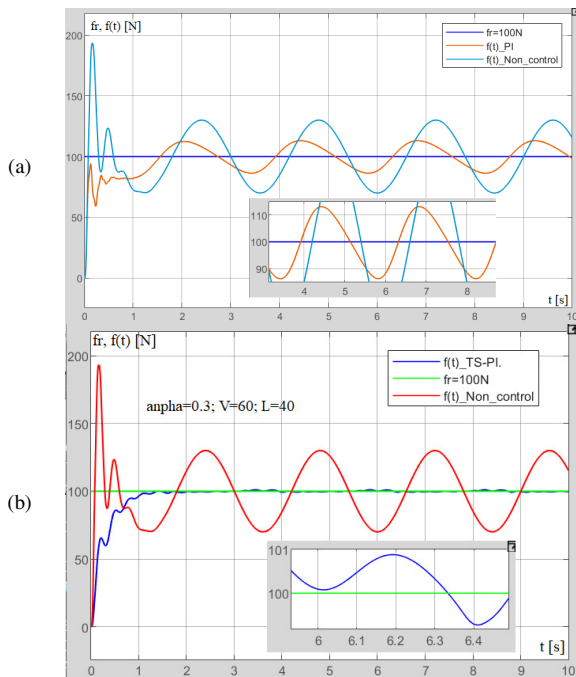


Fig. 16. Contact force regulation performance for  $V=60$ ,  $L=40$ : (a) Robust PI controller, (b) modified T-S fuzzy controller.

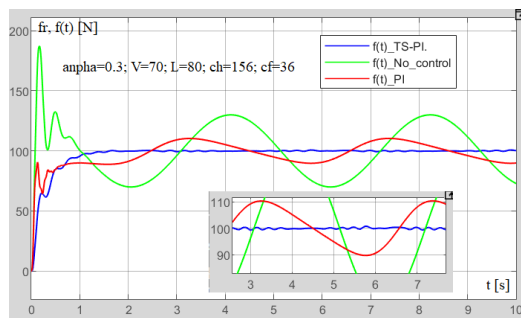


Fig. 17. Simulated parameters of  $V = 70$ ,  $L = 80$ ,  $c_h + 20\%$ ,  $c_f + 20\%$ .

The performance of vibration suppression and contact force regulation, the stiffness of the viscous damping of the pan-head suspension  $c_h$ , and the frame suspension  $c_f$  were transformed by

80% to 120% of their nominal values given in the above simulation. The results are given in Figures 17-20.

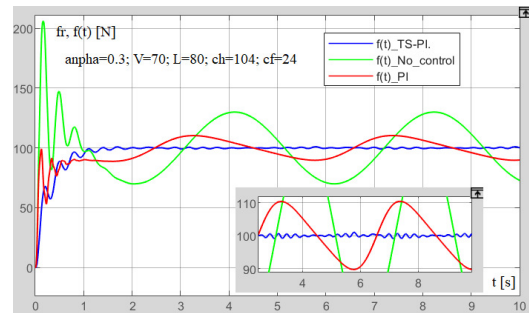


Fig. 18. Simulated parameters of  $V = 70$ ,  $L = 80$ ,  $c_h + 20\%$ ,  $c_f + 20\%$ .

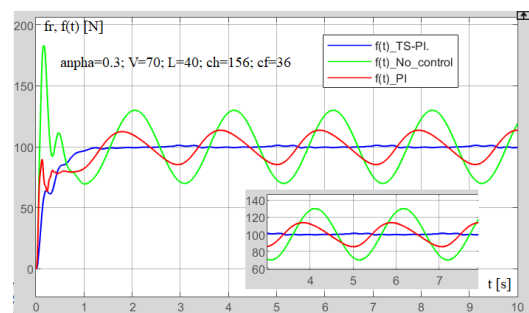


Fig. 19. Simulated parameters of  $V = 70$ ,  $L = 40$ ,  $c_h + 20\%$ ,  $c_f + 20\%$ .

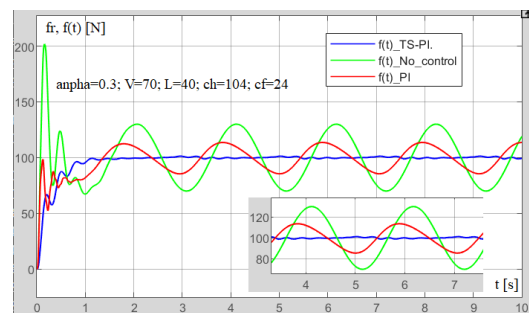


Fig. 20. Simulated parameters of  $V = 70$ ,  $L = 40$ ,  $c_h - 20\%$ ,  $c_f - 20\%$ .

TABLE I. COMPARATIVE RESULTS OF THE VIBRATION SUPPRESSION EFFICIENCY FOR  $a=0.3$

Time varying stiffness	$a$	0.3					
	$V$	60			70		
	$L$	40	60	80	40	60	80
PI	VSE (%)	55.39	63.01	68.80	53.16	59.80	65.66
MZL	VSE (%)	78.93	82.31	84.85	75.95	81.22	83.59
TS-PI	VSE (%)	97.35	98.36	97.74	96.58	98.25	97.68

TABLE II. TEST OF PERFORMANCE ROBUSTNESS

Item			VSE ( $a = 0.3$ , $V = 70$ , $L = 80$ )	VSE ( $a = 0.3$ , $V = 70$ , $L = 40$ )
Viscous damping coefficient $c_h$ , $c_f$	TS-PI	+20%	98.0967%	96.8567%
		-20%	97.5241%	96.2175%
	PI	+20%	65.6442%	53.0810%
		-20%	65.6805%	53.2043%

## VI. CONCLUSION

In this paper, a new robust intelligent control scheme has been successfully presented to suppress the vibration between the pantograph and the catenary by regulating the contact force to a reference value, hence, achieving a stable current collection. The proposed control method has some advantages. The T-S fuzzy model [15] used in this paper has some differences in comparison with the traditional T-S fuzzy model: differential equations are applied instead of the state model, and the consequents of the control rule have two parts, the numerator and the denominator of the control signal. This allows the description of the closed-loop T-S fuzzy control system as a polytope of linear systems so that the system stability analysis can be done, which is the main difficulty in many intelligent control methods, including fuzzy control.

Furthermore, the simple PID controllers are combined with fuzzy logic rules using the PDC design concept. The comparative simulation results have shown that the vibration of the contact force resulting from the time-varying stiffness of the catenary can be effectively suppressed. In particular, the proposed modified T-S fuzzy control algorithm is more efficient than the new intelligent control algorithm represented in [12] and the simple PI structure algorithm. Moreover, once the system stability is guaranteed, the proposed fuzzy controller is robust against possible changes in many system parameters: operational speed of train  $V$ , the length in a span  $L$ , and the stiffness of the viscous damping.

## ACKNOWLEDGMENT

This research is supported by the University of Transport and Communications, Hanoi, Vietnam.

## REFERENCES

- [1] T. X. Wu and M. J. Brennan, "Active vibration control of a railway pantograph," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 211, no. 2, pp. 117–130, Mar. 1997, <https://doi.org/10.1243/0954409971530969>.
- [2] W. Shudong, G. Jingbo, and G. Guosheng, "Research of the active control for high-speed train pantograph," in *2008 IEEE Conference on Cybernetics and Intelligent Systems*, Chengdu, China, Sep. 2008, pp. 749–753, <https://doi.org/10.1109/ICCIS.2008.4670754>.
- [3] H. Wei, C. D. Xu, and K. W. E. Cheng, "A short review on pantograph-catenary arcing issue in high-speed railway systems," in *The 11th IET International Conference on Advances in Power System Control, Operation and Management (APSCOM 2018)*, Hong Kong, China, Aug. 2018, pp. 1–4, <https://doi.org/10.1049/cp.2018.1808>.
- [4] M. Simarro, S. Postigo, M. Prado-Novoa, A. Pérez-Blanca, and J. J. Castillo, "Analysis of contact forces between the pantograph and the overhead conductor rail using a validated finite element model," *Engineering Structures*, vol. 225, Dec. 2020, Art. no. 111265, <https://doi.org/10.1016/j.engstruct.2020.111265>.
- [5] J. Baillieul and T. Samad, Eds., "Proportional-Integral-Derivative Control," in *Encyclopedia of Systems and Control*, London, UK: Springer, 2015, pp. 1103–1103, [https://doi.org/10.1007/978-1-4471-5058-9\\_100019](https://doi.org/10.1007/978-1-4471-5058-9_100019).
- [6] P. Mahajan, R. Garg, V. Gupta, and P. Kumar, "Design of controller for pantograph-catenary system using reduced order model," in *2014 6th IEEE Power India International Conference (PIICON)*, Delhi, India, Sep. 2014, <https://doi.org/10.1109/POWERI.2014.7117725>.
- [7] Z. Xiaodong and F. Yu, "Active self-adaptive control of high-speed train pantograph," in *2011 IEEE Power Engineering and Automation Conference*, Wuhan, China, Sep. 2011, vol. 3, pp. 152–156, <https://doi.org/10.1109/PEAM.2011.6135036>.
- [8] Z. Liu, Y. Liu, N. Zhou, D. Zou, and H. Tu, "Backstepping Controller Design for Pantograph-Catenary System," *IOP Conference Series: Materials Science and Engineering*, vol. 428, no. 1, Jun. 2018, Art. no. 012045, <https://doi.org/10.1088/1757-899X/428/1/012045>.
- [9] A. Ioan Chiriac, S. Constantin Olteanu, and D. Popescu, "Model Predictive Control Approach for a Pantograph-Catenary System (PAC) Described by a Transfer Function Model," in *2020 24th International Conference on System Theory, Control and Computing (ICSTCC)*, Sinaia, Romania, Jul. 2020, pp. 733–738, <https://doi.org/10.1109/ICSTCC50638.2020.9259700>.
- [10] N. Mokrani and A. Rachid, "A robust control of contact force of pantograph-catenary for the high-speed train," in *2013 European Control Conference (ECC)*, Zurich, Switzerland, Jul. 2013, pp. 4568–4573, <https://doi.org/10.23919/ECC.2013.6669668>.
- [11] Y. J. Huang and T. C. Kuo, "Discrete pantograph position control for the high speed transportation systems," in *IEEE International Conference on Networking, Sensing and Control, 2004*, Taipei, Taiwan, Mar. 2004, vol. 2, pp. 932–936 Vol.2, <https://doi.org/10.1109/ICNSC.2004.1297072>.
- [12] T.-C. Lin, C.-W. Sun, Y.-C. Lin, and M. M. Zirkohi, "Intelligent Contact Force Regulation of Pantograph-Catenary Based on Novel Type-Reduction Technology," *Electronics*, vol. 11, no. 1, Jan. 2022, Art. no. 132, <https://doi.org/10.3390/electronics11010132>.
- [13] M. Moradi Zirkohi and T.-C. Lin, "An efficient non-iterative method for computing the centroid of an interval type-2 fuzzy set," *Journal of Intelligent & Fuzzy Systems*, vol. 41, no. 2, pp. 2879–2889, Jan. 2021, <https://doi.org/10.3233/JIFS-202913>.
- [14] N. Zerroug, K. Behih, Z. Bouchama, and K. Zehar, "Robust Adaptive Fuzzy Control of Nonlinear Systems," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8328–8334, Apr. 2022, <https://doi.org/10.48084/etasr.4781>.
- [15] L. H. Lan, P. Van Lam, and N. Van Hai, "An Approach to the Analysis and Design of Fuzzy Control System," in *2020 3rd International Conference on Robotics, Control and Automation Engineering (RCAE)*, Chongqing, China, Aug. 2020, pp. 36–40, <https://doi.org/10.1109/RCAE51546.2020.9294179>.
- [16] C. Hwang, L. Hwang, and J. Hwang, "Robust D-partition," *Journal of the Chinese Institute of Engineers*, vol. 33, no. 6, pp. 811–821, Sep. 2010, <https://doi.org/10.1080/02533839.2010.9671671>.
- [17] I. H. Hamad, A. Chouchaine, and H. Bouzaouache, "A Takagi-Sugeno Fuzzy Model for Greenhouse Climate," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7424–7429, Aug. 2021, <https://doi.org/10.48084/etasr.4291>.
- [18] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 1, pp. 116–132, Jan. 1985, <https://doi.org/10.1109/TSMC.1985.6313399>.
- [19] H. O. Wang, K. Tanaka, and M. F. Griffin, "An approach to fuzzy control of nonlinear systems: stability and design issues," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 1, pp. 14–23, Oct. 1996, <https://doi.org/10.1109/91.481841>.
- [20] L. K. Wong, F. H. F. Leung, and P. K. S. Tam, "Design of fuzzy logic controllers for Takagi-Sugeno fuzzy model based system with guaranteed performance," *International Journal of Approximate Reasoning*, vol. 30, no. 1, pp. 41–55, May 2002, [https://doi.org/10.1016/S0888-613X\(02\)00062-2](https://doi.org/10.1016/S0888-613X(02)00062-2).
- [21] M. Fu and B. R. Barmish, "Polytopes of Polynomials with Zeros in a Prescribed Region," in *1988 American Control Conference*, Atlanta, GA, USA, Jun. 1988, pp. 2461–2464, <https://doi.org/10.23919/ACC.1988.4790138>.



# Buckling and Vibration Estimation of Girder Steel Portal Frames using the Bayesian Updating Methods

**Fahmi H. Fahmi**

Department of Civil Engineering, College of Engineering, University of Baghdad, Iraq  
f.fahmi1901m@coeng.uobaghdad.edu.iq  
(corresponding author)

**Salah Al-Zaidee**

Department of Civil Engineering, College of Engineering, University of Baghdad, Iraq  
salah.r.al.zaidee@coeng.uobaghdad.edu.iq

Received: 26 October 2022 | Revised: 5 November 2022 | Accepted: 7 November 2022

## ABSTRACT

In practice, structural systems are complicated. When dealing with such systems, the use of analytical simulations is or may be impossible to apply, since the problem includes many variables with partial differential formulations. The stability and dynamic response of structures is an important aspect that must be paid particular attention to in order to ensure safety against collapse. A physical phenomenon of a reasonably straight, slender member bending laterally (usually abruptly) from its longitudinal position due to compression is referred to as buckling. Two kinds of buckling can be distinguished: (1) bifurcation-type buckling and (2) deflection-amplification-type buckling. In vibration, frequencies and mode shapes must be known in order to use some methods of dynamic response analysis. To calculate frequencies and modes, an eigenvalue problem is solved in algebraic form. In structural dynamic problems, only the lowermost eigenpairs are of important interest. The highest eigenpairs are not needed and are not accurate due to discretization errors.

**Keywords**-interior supporting frame; Bayesian method; sample data; prior data; posterior data; first order; MCs

## I. INTRODUCTION

The majority of engineering designs are based on deterministic variables and often do not consider the variations in the material properties and the geometry of the structure [1]. Uncertainty is vital to the analysis and design of an engineering system and the effects of uncertainties must be quantified and propagated. Increasing amounts of data in engineering systems are collected and stored. This information can and should be used to decrease the uncertainty in engineering models and optimize the management of these systems. A coherent and effective way for merging new with existing information is provided by the Bayesian method, in which prior probabilistic information is updated with new data and observations. The Bayesian framework enables the combination of uncertain and incomplete information with models from different sources and provides probabilistic information on the accuracy of the updated model [2]. The current study emphasizes the analysis of the buckling and vibration of steel portal frames, where new data have been acquired for the independent variables of the system. Subsequently, the Bayesian method presents an efficient way to update the prior knowledge regarding the

stability and natural frequency of the frame. The first-order approximation method and Monte Carlo (MC) simulation have been adopted to investigate the statistical properties of the buckling and vibration.

## II. THE BAYESIAN METHOD

In engineering, one often needs to use whatever information is available in formulating a sound basis for making decisions. This may include observed data (field or experimental), information derived from theoretical models, and judgments based on experience. Moreover, the available information may need to be updated as new information or data are acquired [3]. The proper tool for combining and updating the available information is embodied in the Bayesian approach. Parameter estimation in the Bayesian approach is based on the updating formula:

$$f(\theta) = cL(\theta)p(\theta) \quad (1)$$

where  $p(\theta)$  is the prior Probability Density Function (PDF) representing the initial state of knowledge about the unknown parameters  $\theta$ ,  $L(\theta)$  is the likelihood function representing the

knowledge gained from a set of observations, and the constant  $c$  is a normalizing factor. The posterior PDF  $f(\theta)$  represents the updated state of knowledge regarding the parameters  $\theta$ .

#### A. Prior Distribution

The prior distribution represents the distribution of possible parameter values from which the parameter has been drawn. The selection of the prior distribution is an important matter in Bayesian modeling. Since the choice of the prior distribution has a major effect on the resulting inference, this choice must be conducted with the utmost care [4].

#### B. Posterior Distribution

The current information about the parameter is contained in the posterior probability distribution. It merges the information of the prior distribution and the likelihood. This results in a strong representation of the information than separate sources of information [5]. The prior and posterior densities are presented in Table I and the mean and variance of the posterior parameters in Table II.

TABLE I. RANDOM VARIABLES AND THEIR PRIOR AND POSTERIOR DENSITY DISTRIBUTIONS [3]

Basic random variables	Parameter	Prior and posterior distribution
Normal $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$	$\mu$	Normal $\mu'' = \frac{\mu'(\sigma'^2/n) + \bar{x}\sigma'^2}{\sigma'^2/n + (\sigma')^2}$
Lognormal $f_X(x) = \frac{1}{\sqrt{2\pi}\xi x} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \lambda}{\xi}\right)^2\right]$	$\lambda$	Normal $f_\lambda(\lambda) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{\lambda - \mu}{\sigma}\right)^2\right]$

TABLE II. MEAN AND VARIANCE OF THE PARAMETERS AND THEIR POSTERIOR STATISTICS [3]

Mean and variance of parameters	Posterior statistics
$E(\mu) = \mu_\mu$	$\mu'' = \frac{\mu'(\sigma'^2/n) + \bar{x}\sigma'^2}{\sigma'^2/n + (\sigma')^2}$
$Var(\mu) = \sigma_\mu^2$	$\sigma'' = \sqrt{\frac{(\sigma')^2(\sigma'^2/n)}{(\sigma')^2 + \sigma'^2/n}}$
$E(\lambda) = \mu$	$\mu'' = \frac{u'(\xi^2/n) + \sigma^2 \ln \bar{x}}{\xi^2/n + \sigma^2}$
$Var(\lambda) = \sigma^2$	$\sigma'' = \sqrt{\frac{\sigma^2(\xi^2/n)}{\sigma^2 + \xi^2/n}}$

### III. STATISTICAL CHARACTERISTICS OF THE INDEPENDENT VARIABLES

This study considers the applied loads, cross-section dimensions, modulus of elasticity, girder span, plastic section modulus, and yield strength as random variables. The sample data of the independent variables were summarized and are presented in Table III. The prior statistical characteristics of these variables have been collected from the literature and can be seen in Table IV.

The posterior statistical data of the input variables have been calculated from the updating process of the prior data by

the Bayesian method. The modulus of elasticity of steel  $E_s$  has been considered as an example of this process. Updating mean, standard deviation, and coefficient of variance have been estimated based on the relations in (1)-(3). The posterior statistical characteristics of the independent variables are summarized in Table V.

$$\mu'' =$$

$$\frac{\left[\frac{\bar{x}}{(\frac{\sigma'}{\sqrt{n}})^2} + \left[\frac{\mu'}{(\sigma')^2}\right]\right]}{\left[\frac{1}{(\frac{\sigma'}{\sqrt{n}})^2} + \left[\frac{1}{(\sigma')^2}\right]\right]} = \frac{\bar{x}(\sigma')^2 + \mu'(\frac{\sigma'^2}{n})}{(\sigma')^2 + (\frac{\sigma'^2}{n})} = \frac{(205)(16) + (200)(\frac{151.3}{5})}{(16) + (\frac{151.29}{5})} = 202 \text{GPa} \quad (1)$$

$$\sigma'' = \sqrt{\frac{(\sigma')^2(\frac{\sigma'^2}{n})}{(\sigma')^2 + (\frac{\sigma'^2}{n})}} = \sqrt{\frac{(16)(\frac{151.3}{5})}{(16) + (\frac{151.3}{5})}} = 3.24 \text{GPa} \quad (2)$$

$$cov'' = \frac{\sigma''}{\mu''} = \frac{3.24}{202} = 0.016 \quad (3)$$

TABLE III. SAMPLE STATISTICAL DATA OF THE VARIABLES

Variable	Mean Nominal	COV	Distribution type	Reference
Dead load $P_D$	1.03	0.08	Normal	[6]
Live load $P_L$	1.00	0.1	Gumbel	[7]
Moment of inertia $I_x$	0.96	0.03	Normal	Local survey
Modulus of elasticity $E_s$	1.025	0.05	Normal	[8]
Modulus of elasticity $E_c$	0.98	0.07	Lognormal	[9]
Member span $L$	1.00	0.07	Lognormal	[10]
Plastic section modulus $Z_x$	1.04	0.05	Lognormal	[11]
Yield strength of steel $F_y$	1.10	0.07	Lognormal	[11]

TABLE IV. PRIOR STATISTICAL DATA OF THE VARIABLES

Variable	Mean Nominal	COV	Distribution type	Reference
Dead load $P_D$	1.03	0.08	Normal	[6]
Live load $P_L$	1.00	0.1	Gumbel	[7]
Moment of inertia $I_x$	0.96	0.05	Normal	[12]
Modulus of elasticity $E_s$	1.00	0.02	Normal	[11]
Modulus of elasticity $E_c$	1.00	0.03	Lognormal	[13]
Member span $L$	1.00	0.004	Lognormal	[14]
Plastic section modulus $Z_x$	1.00	0.04	Lognormal	[15]
Yield strength of steel $F_y$	1.10	0.06	Lognormal	[12]

TABLE V. POSTERIOR STATISTICAL DATA OF THE VARIABLES

Variable	Mean Nominal	COV	Distribution type
Dead load $P_D$	1.03	0.08	Normal
Live load $P_L$	1.00	0.1	Gumbel
Moment of inertia $I_x$	0.96	0.005	Normal
Modulus of elasticity $E_s$	1.01	0.016	Normal
Modulus of elasticity $E_c$	1.00	0.02	Lognormal
Member span $L$	1.00	0.003	Lognormal
Plastic section modulus $Z_x$	1.00	0.03	Lognormal
Yield strength of steel $F_y$	1.10	0.04	Lognormal

### IV. PARAMETRIC STUDY

A steel single-story warehouse, as shown in Figure 1, with 45m total length and 15m width, consisting of 5 bays, has been considered in this parametric study. The superimposed dead

load has been determined based on the assumption of a flooring system of a concrete slab with a thickness of 100mm and a metal deck of 7.5mm, supported on IPE 300 floor beams. The first inner frame shown in Figure 2 has been considered the most critical, as it includes the exterior face of the first interior support where the shear force is about 15% greater than the average value. The frame consists of IPE 600 steel girder and columns. The applied live loads on the girder have been determined based on the ASCE 7 specifications. The self-weight of the girder and columns have been determined depending on the cross-section dimensions and material density.

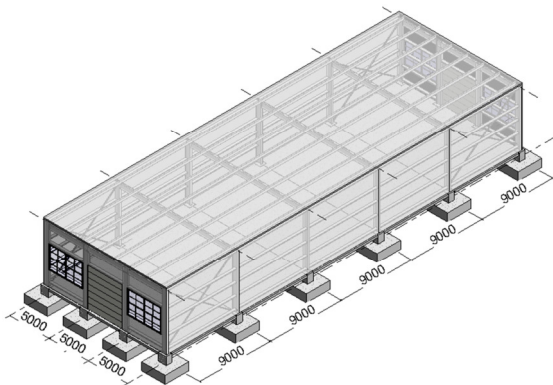


Fig. 1. Steel single-story warehouse building.

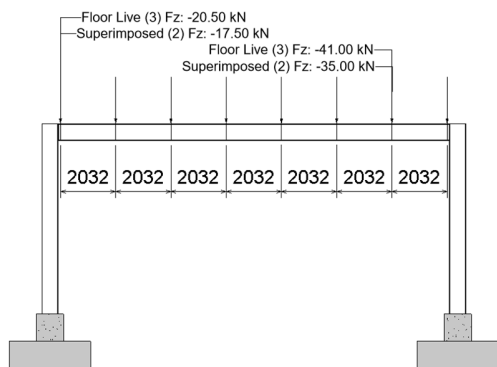


Fig. 2. Summary of the dead and live load on the interior supporting frame.

## V. DERIVED AND SIMULATED STATISTICAL PROPERTIES FOR THE BUCKLING AND VIBRATION OF THE GIRDER

Two stages have been considered in the first-order approximation method and the MC simulation method. The randomness of the dependent variables, in the first stage, has been deduced based on the prior knowledge of Table IV, while in the second stage, it has been deduced based on the posterior knowledge of Table V.

### A. First-Order Approximation Method

It is possible to expand the original model into an infinite Taylor Series (TS) around the mean values:

$$g(X) = g(\mu_X) + (X - \mu_X) \frac{dg}{dX} + \frac{1}{2} (X - \mu_X)^2 \frac{d^2g}{dX^2} + \dots + \frac{1}{n!} (X - \mu_X)^n \frac{d^ng}{dX^n} \quad (4)$$

where the function and derivatives are evaluated at  $\mu_X$ . It is common to include only linear terms, assuming that random input variables are independent. A function  $g(X)$  of  $N$  independent random variables can be approximated by linear terms of the TS, which are as follows [16]:

$$E(Y) \approx g(\mu_X) \quad (5)$$

and

$$\text{Var}(Y) \approx \text{Var}(X - \mu_X) \left( \frac{dg}{dX} \right)^2 = \text{Var}(x) \left( \frac{dg}{dX} \right)^2 \quad (6)$$

### 1) Buckling Analysis

The buckling  $P_{cr}$  is a function of the span  $L$ , the modulus of elasticity  $E$ , and the moment of inertia  $I$ . By refereeing to (4) and (5) the mean and variance of the buckling can be estimated by:

$$E(P_{cr}) = \frac{1}{K^2} \frac{\pi^2 \times \bar{E} \times \bar{I}_{Major}}{(L)^2} \quad (7)$$

$$\text{Var}(P_{cr}) = \text{Var}(I) \left( \frac{\partial P_{cr}}{\partial I} \right)^2 + \text{Var}(L) \left( \frac{\partial P_{cr}}{\partial L} \right)^2 + \text{Var}(E) \left( \frac{\partial P_{cr}}{\partial E} \right)^2 \quad (8)$$

Based on the statistical characteristics of the independent variables illustrated in Tables IV and V, the prior and posterior statistical characteristics for  $P_{cr}$  have been estimated and are presented in Tables VI and VII.

TABLE VI. PRIOR STATISTICAL CHARACTERISTICS OF THE BUCKLING OF THE GIRDER

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$P_{cr}$	44198	42072	2172	0.051

TABLE VII. POSTERIOR STATISTICAL CHARACTERISTICS OF THE BUCKLING OF THE GIRDER

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$P_{cr}$	44198	42493	750	0.017

### 2) Vibration Analysis

The vibration  $f_n$  is a function of the load  $P$ , span  $L$ , modulus of elasticity  $E$ , and moment of inertia  $I$ . The mean and variance of the vibration are calculated by:

$$E(f_n) = 0.18 \frac{1}{\sqrt{f_k}} \sqrt{\frac{g}{\Delta_{girder}}} \Rightarrow 1.247 \sqrt{\frac{g}{f_k}} \sqrt{\frac{EI}{PL^3}} \quad (9)$$

$$\text{Var}(f_n) = \text{Var}(I) \left( \frac{\partial f_n}{\partial I} \right)^2 + \text{Var}(L) \left( \frac{\partial f_n}{\partial L} \right)^2 + \text{Var}(E) \left( \frac{\partial f_n}{\partial E} \right)^2 + \text{Var}(P) \left( \frac{\partial f_n}{\partial P} \right)^2 \quad (10)$$

Based on Tables IV and V, the prior and posterior statistical characteristics for vibration are presented in Tables VIII and IX.

TABLE VIII. PRIOR STATISTICAL CHARACTERISTICS

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$f_n$	3.32	7.23	0.214	0.029

TABLE IX. POSTERIOR STATISTICAL CHARACTERISTICS

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$f_n$	3.32	7.267	0.103	0.014

### B. Monte Carlo Simulation

The MC method is considered one of the most powerful and accurate simulation tools and can be applied to many practical problems allowing the direct consideration of any type of probability distribution for the random variables [17]. The MC method has been adopted to achieve the simulation processes of the buckling, and vibration for the interior supporting frame. MATLAB code has been used to generate pseudo-random sampling with a size of 1000 for each input variable.

#### 1) Buckling Analysis

The prior statistical characteristics of the buckling are presented in Figure 3 and are summarized in Table X.

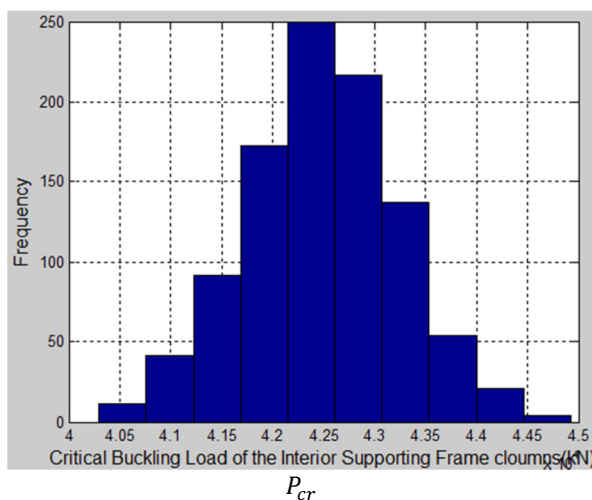


Fig. 3. Histogram of the prior buckling of a typical interior girder.

TABLE X. PRIOR STATISTICAL CHARACTERISTICS OF THE BUCKLING OF THE GIRDER

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$P_{cr}$	41999	2279	0.054	Normal

The posterior statistical characteristics of the buckling were estimated and are presented in Figure 4 and Table XI.

TABLE XI. POSTERIOR STATISTICAL CHARACTERISTICS OF THE BUCKLING OF THE GIRDER

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$P_{cr}$	42481	756	0.017	Normal

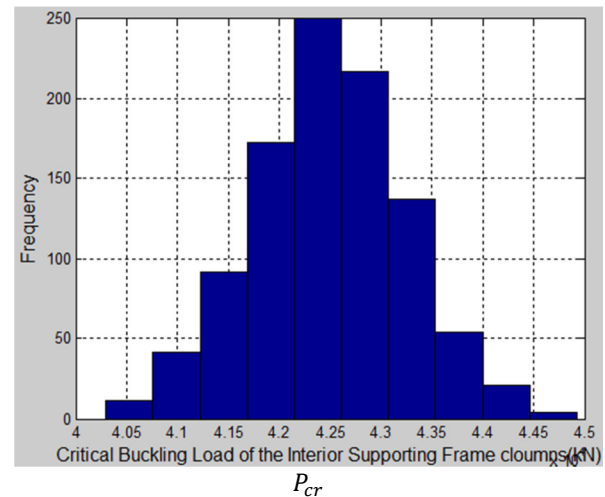


Fig. 4. Histogram of the posterior buckling of a typical interior girder.

#### 2) Vibration Analysis

The prior statistical characteristics of the natural frequency are presented in Figure 5 and Table XII.

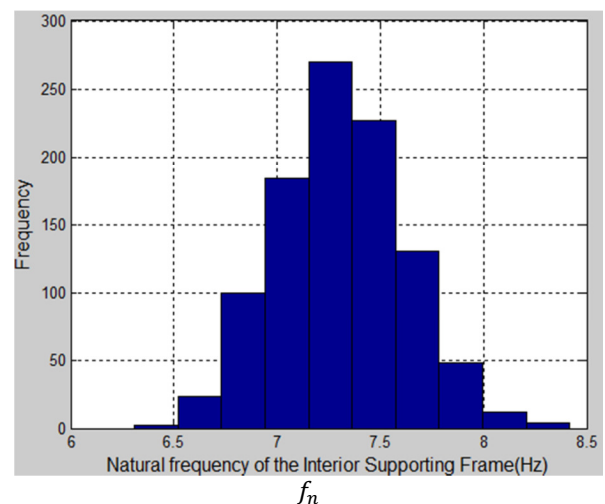


Fig. 5. Histogram of the prior natural frequency of a typical interior girder.

TABLE XII. PRIOR STATISTICAL CHARACTERISTICS OF THE NATURAL FREQUENCY OF THE GIRDER

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$f_n$	7.30	0.311	0.042	Normal

The estimated posterior statistical characteristics of the natural frequency are presented in Figure 6 and Table XIII.

TABLE XIII. POSTERIOR STATISTICAL CHARACTERISTICS OF THE NATURAL FREQUENCY OF THE GIRDER

Random variable	Nominal (kN)	Mean (kN)	Standard deviation (kN)	COV
$f_n$	7.36	0.26	0.035	Normal

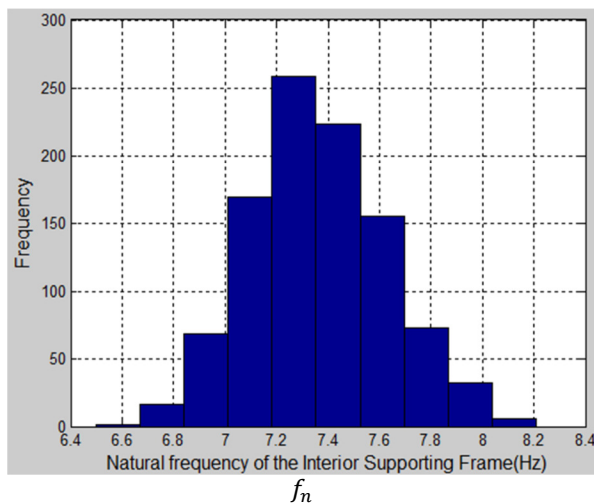


Fig. 6. Histogram of the posterior natural frequency of a typical interior girder.

## VI. CONCLUSION

In this paper, buckling and vibration analysis was performed for the girder of the first inner frame for the prior and posterior statistical characteristics. Through the application of the Bayesian method, and from the results of the prior and posterior analysis, it was shown that the greater the knowledge gained about the independent parameters, the less randomness is in the dependent parameters, and thus the analysis and design of the system is enhanced. In the first order analysis, the covariance of the critical buckling load decreased from 0.051 to 0.017, and for the natural frequency from 0.029 to 0.014, while in Monte Carlo simulation, the covariance results were from 0.054 to 0.017 for the critical buckling load, and from 0.042 to 0.035 for the natural frequency. The decrement of the covariance of the dependent parameters is due to the decrement of the values of the covariance of the independent parameters because of the increased knowledge of the independent parameters.

## REFERENCES

- [1] N. L. Tran and T. H. Nguyen, "Reliability Assessment of Steel Plane Frame's Buckling Strength Considering Semi-rigid Connections," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5099–5103, Feb. 2020, <https://doi.org/10.48084/etasr.3231>.
- [2] D. Straub and I. Papaioannou, "Bayesian Updating with Structural Reliability Methods," *Journal of Engineering Mechanics*, vol. 141, no. 3, Mar. 2015, Art. No. 04014134, [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0000839](https://doi.org/10.1061/(ASCE)EM.1943-7889.0000839).
- [3] A. H.-S. Ang and W. H. Tang, *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, 2e Instructor Site: Emphasis on Applications to Civil and Environmental Engineering, 2nd ed. New York, NY, USA: Wiley, 2006.
- [4] A. Belay, E. O'Brien, and D. Kroese, "Truck fleet model for design and assessment of flexible pavements," *Journal of Sound and Vibration*, vol. 311, no. 3, pp. 1161–1174, Apr. 2008, <https://doi.org/10.1016/j.jsv.2007.10.019>.
- [5] A. E. Mansour, H. Y. Jan, C. I. Zigelman, Y. N. Chen, and S. J. Harding, "Implementation of Reliability Methods to Marine Structures," *Transactions - Society of Naval Architects and Marine Engineers*, vol. 92, pp. 353–382, 1984.
- [6] M. S. Darmawan, A. N. Refani, M. Irmawan, R. Bayuaji, and R. B. Anugraha, "Time Dependent Reliability Analysis of Steel I Bridge Girder Designed Based on SNI T-02-2005 and SNI T-3-2005 Subjected to Corrosion," *Procedia Engineering*, vol. 54, pp. 270–285, Jan. 2013, <https://doi.org/10.1016/j.proeng.2013.03.025>.
- [7] S. G. Buonopane and B. W. Schafer, "Reliability of Steel Frames Designed with Advanced Analysis," *Journal of Structural Engineering*, vol. 132, no. 2, pp. 267–276, Feb. 2006, [https://doi.org/10.1061/\(ASCE\)0733-9445\(2006\)132:2\(267\)](https://doi.org/10.1061/(ASCE)0733-9445(2006)132:2(267)).
- [8] A. Ghabussi, J. Asgari Marnani, and M. S. Rohanimanesh, "Improving seismic performance of portal frame structures with steel curved dampers," *Structures*, vol. 24, pp. 27–40, Apr. 2020, <https://doi.org/10.1016/j.istruc.2019.12.025>.
- [9] P. Geyskens, A. D. Kiureghian, and P. Monteiro, "Bayesian Prediction of Elastic Modulus of Concrete," *Journal of Structural Engineering*, vol. 124, no. 1, pp. 89–95, Jan. 1998, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1998\)124:1\(89\)](https://doi.org/10.1061/(ASCE)0733-9445(1998)124:1(89)).
- [10] H. Q. Jebur and S. R. Al-Zaidee, "Non-deterministic Approach for Reliability Evaluation of Steel Portal Frame," *Civil Engineering Journal*, vol. 5, no. 8, pp. 1684–1697, Aug. 2019, <https://doi.org/10.28991/cej-2019-03091363>.
- [11] K. Atua, I. Assakkaf, and B. M. Ayyub, "Statistical Characteristics of Strength and Load Random Variables of Ship Structures," in *Proceedings of the Seventh Specialty Conference on Probabilistic Mechanics & Structural Reliability*, Worcester, MA, USA, 1996, pp. 106–109, Accessed: Nov. 19, 2022.
- [12] H. Zhang, H. Liu, B. R. Ellingwood, and K. J. R. Rasmussen, "System Reliabilities of Planar Gravity Steel Frames Designed by the Inelastic Method in AISC 360-10," *Journal of Structural Engineering*, vol. 144, no. 3, Mar. 2018, Art. no. 04018011, [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0001991](https://doi.org/10.1061/(ASCE)ST.1943-541X.0001991).
- [13] ACI Committee 318, *ACI 318-19 Building Code Requirements for Structural Concrete*. Farmington Hills, MI, USA: American Concrete Institute, 2019.
- [14] *IS 7215: Tolerances for fabrication of steel structures*. Bureau of Indian Standards, 1974.
- [15] P. E. Hess, D. Bruchman, I. A. Assakkaf, and B. M. Ayyub, "Uncertainties in Material and Geometric Strength and Load Variables," *Naval Engineers Journal*, vol. 114, no. 2, pp. 139–166, 2002, <https://doi.org/10.1111/j.1559-3584.2002.tb00128.x>.
- [16] D. Mohammed and S. R. Al-Zaidee, "Deflection Reliability Analysis for Composite Steel Bridges," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9155–9159, Oct. 2022, <https://doi.org/10.48084/etasr.5146>.
- [17] M. S. M. Noori and R. M. Abbas, "Reliability Analysis of an Uncertain Single Degree of Freedom System Under Random Excitation," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9252–9257, Oct. 2022, <https://doi.org/10.48084/etasr.5193>.



# Development of Two UWB Multiband MIMO Antennas with Enhanced Isolation and Cross-Correlation

**Shrenik Sarade**

Electronics Engineering Department, Walchand College of Engineering, India  
shreniks2k7@rediffmail.com  
(corresponding author)

**Sachin Ruikar**

Electronics Engineering Department, Walchand College of Engineering, India  
sachin.ruikar@walchandsangli.ac.in

*Received: 17 October 2022 | Revised: 4 November 2022 and 17 November 2022 | Accepted: 19 November 2022*

## ABSTRACT

An Ultra-Wide Band (UWB) MIMO antenna working at millimeter-wave is proposed in this research. It is composed of eight radiating elements with various shapes. It is designed with a rectangular structure and various carved slots. The carved slots are used to increase the antenna's bandwidth. The antenna's radiating elements are placed near one another, and thus, isolation is widespread. In order to improve the isolation of the MIMO antenna, parasitic elements and a defective ground structure are used. Antenna parameters such as Correlation Coefficient (CC), Envelope Correlation Coefficient (ECC), Diversity Gain (DG), and Total Active Reflection Coefficient (TARC) depend on the isolation. Parasitic elements with a rectangular form are positioned between the radiating patches. Rectangular-shaped ground structures with defects comprise the ground plane. An FR-4 substrate is used to fabricate the antenna. The analysis of the antenna shows that there is less than -14dB return loss, less than -40dB isolation, less than 0.0010 cross-correlation, less than 0.10 TARC, and higher than 500MHz bandwidth. The antenna uses a fractional bandwidth higher than 35% (UWB) for the 6GHz frequency and operates on a variety of bands. This antenna is suitable for many different wireless system applications.

**Keywords-**MIMO; UWB; DGS; TARC; CC; ECC; DG

## I. INTRODUCTION

Modern antenna technology relies heavily on Multiple Input-Multiple Output (MIMO) antennas with several antennas in both transmitter and receiver. MIMO antenna systems increase the poor data rate of the Single Input-Single Output (SISO) system. Depending on the number of patches in the antenna, the data rate of the MIMO antenna increases linearly. Return loss, isolation, Correlation Coefficient (CC), and Total Active Reflection Coefficient (TARC) represent a MIMO antenna's performance. Multiple radiating patches are arranged on the same substrate in a MIMO antenna [1]. The radiating elements are positioned near one another. The isolation and the CC are influenced by the close proximity of the rectangular radiating patches. The data rate is reduced by the isolation. The isolation, CC, and data rate of the antennas all need to be improved. Various techniques, such as Defective Ground Structure (DGS), Parasitic Elements (PEs), metamaterials, Decoupling Structures (DSs), and Electromagnetic Band Gap (EBG) are used to increase isolation and CC. In the ground plane, carved holes of different shapes are employed in DGS

methods [2-7]. In order to promote isolation, a decoupling structure is built that isolates the input of the radiating patches from the corresponding network arrangement in [8-12]. To increase the isolation between the patches, parasitic components are utilized in [13-18]. Using the literature mentioned above, it is necessary to create a MIMO antenna that operates on a multiband application with an ultra-wide bandwidth, improved TARC, and better cross correlation. In previous works, the same shaped radiating elements were used to design an MIMO antenna, whereas the cross correlation (CC and ECC) between the radiating elements was high.

In this study, a 2-MIMO antenna with a variety of shapes is proposed. The antenna is developed using radiating elements of different shapes. The designed antenna operates on a multiband and its bandwidth is UWB. For bandwidth improvement, a rectangular shaped cut slot is introduced between the radiating elements of the antenna. The isolation between the radiating elements is improved using DGS and PEs. Antenna parameters such as isolation, cross correlation (CC and ECC), and TARC depend on isolation.

The novelty of the current work is that the radiating elements are developed in different shapes, the rectangular shaped DGS cut slot and PEs are designed for the improvement of isolation, CC, ECC, TARC, and DG. For the improvement of the bandwidth, different rectangular cut slots are designed and introduced in the radiating elements.

## II. DESIGN OF A MASSIVE MIMO ANTENNA USING DEFECTIVE GROUND STRUCTURE (DGS)

In this section, the design of the massive MIMO antenna is explained in detail. The designed antenna comprises 8 rectangular radiating patches with distinct cut slots for the 6GHz frequency with the antenna dimensions, as shown in Figure 1. The antenna is analyzed and fabricated using an FR-4 substrate with 1.6mm height (h) and 4.4 dielectric constant. The radiating element's geometrical features, such as length

( $L_p$ ) and width ( $W_p$ ) have been created using the rectangular radiating patch transmission line model [21-23] (Table I). Table I shows the measurements of the 8 radiating elements with varying cut slots. A rectangular cut slot [21-22] is placed in the radiating elements  $P_1$  to  $P_8$ . Patch 1 has rectangular cut slots-1 and 2 (CS-1 and 2). Patches 2, 3, and 7 have the rectangular cut slot-2. Cut slot-3 is placed in patches 6 and 8. In patch 4, rectangular cut slot-1 is introduced. Patch 5 introduces rectangular cut slots-1 and 3 (CS-1 and 3). The radiating elements of the antenna are excited using an Inset Feed Line (IFL) [21-22]. The fabricated antenna structure is shown in Figure 3. The antenna is fabricated on the FR-4 substrate. The radiating elements of the antenna are placed close to each other. Therefore, the isolation between the radiating elements is high. Isolation of the antenna affects the performance parameters of the antenna [21-22].

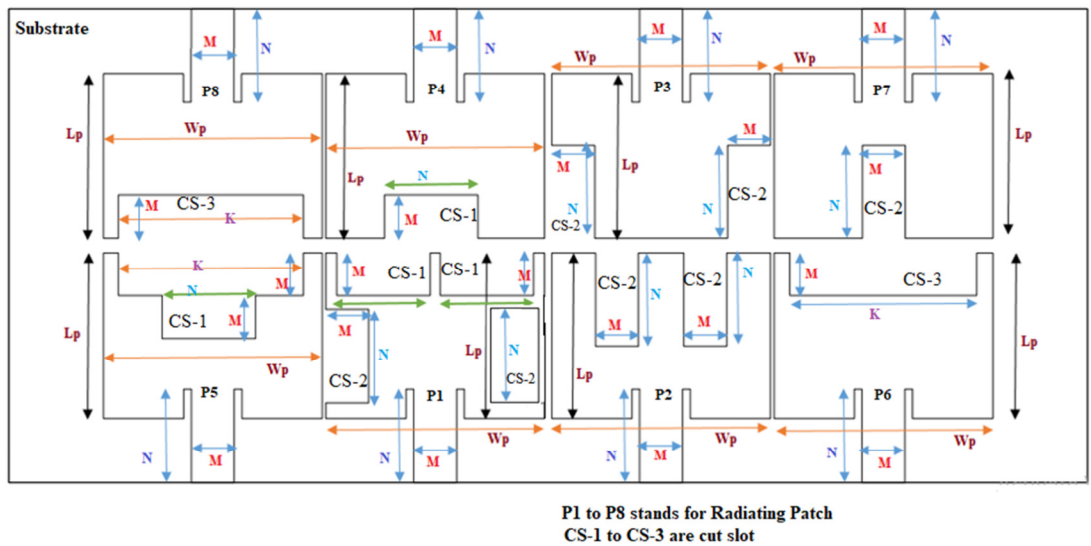


Fig. 1. The dimensions of a developed massive MIMO antenna (top side).

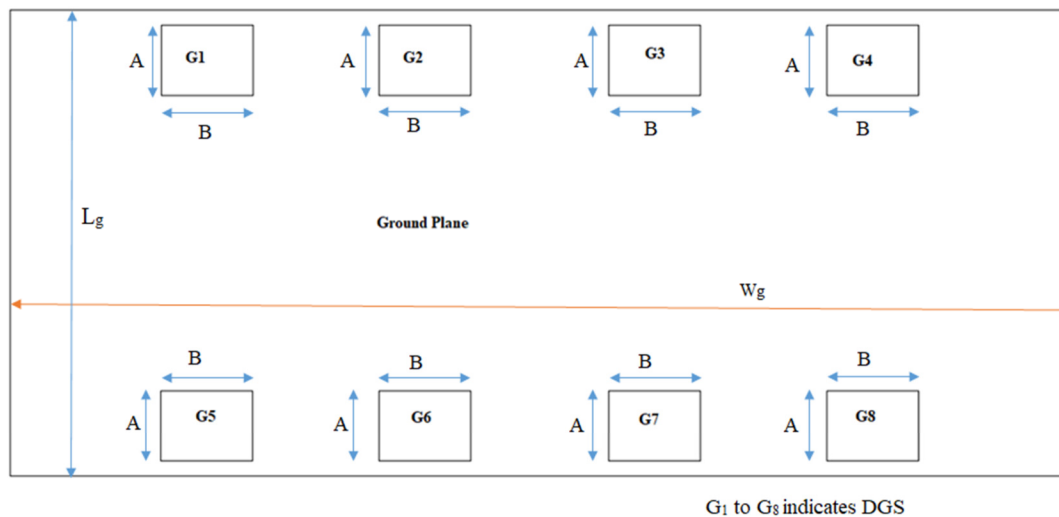


Fig. 2. The dimensions of a ground plane (back side) of the antenna.

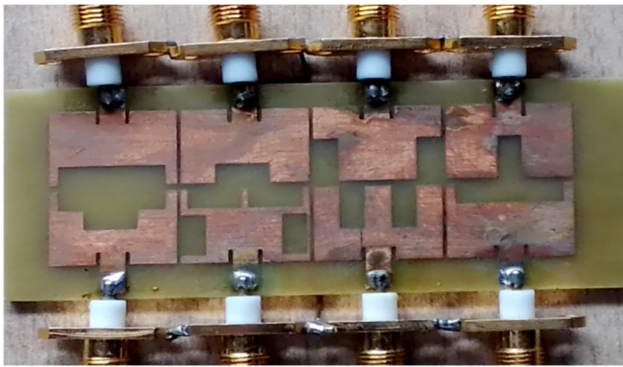


Fig. 3. Fabricated top side of a massive MIMO antenna.

TABLE I. MIMO ANTENNA DIMENSIONS

Parameter	Dimensions (mm)
Patch length ( $L_p$ )	11.38
Patch width ( $W_p$ )	15.21
Width of IFL	$N = 6.41$
Length of IFL	$M = 2.942$
Width of cut slot-1	$N = 6.41$
Length of cut slot-1	$M = 2.942$
Width of cut slot-2	$M = 2.942$
Length of cut slot-2	$N = 6.41$
Width of cut slot-3	$K = 12.69$
Length of cut slot-3	$M = 2.942$
Length of a substrate	$L_G = 32.58$
Width of a substrate	$W_G = 74$
Length of a DGS cut slot	$A = 5.884$
Width of a DGS cut slot	$B = 6.38$

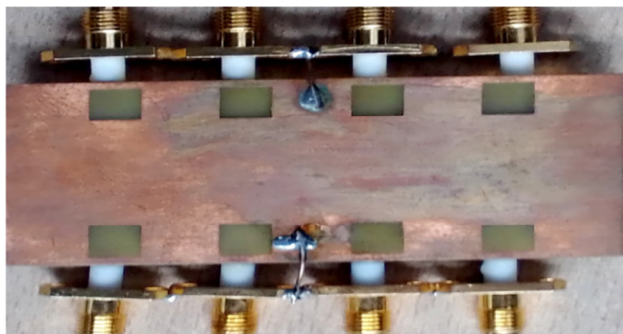


Fig. 4. Fabricated back side (ground plane) of a massive MIMO antenna.

Hence, there is a requirement to improve the isolation between the radiating elements. For isolation improvement, the DGS cut slots  $G_1$  to  $G_8$  [2-7] are introduced in the ground plane as shown in Figure 2, whereas the fabrication of the ground plane is shown in Figure 4. The dimensions of the ground plane with DGS are shown in Table I. The result analysis of the antenna (for radiating elements-1 and 4) is shown in Figures 5 and 6. The antenna parameters are tested in the Vector Network Analyser (VNA). The analysis shows -14.40dB return loss and -42.72dB isolation have been measured in the -6dB line of return loss. At the -6dB line of return loss, the antenna operates on the frequency band of 0.33GHz-7.5GHz. This antenna is operated on UWB. The Voltage Standing Wave Ratio (VSWR) is illustrated in Figure 7. The VSWR for the antenna is less than 3.65dB.

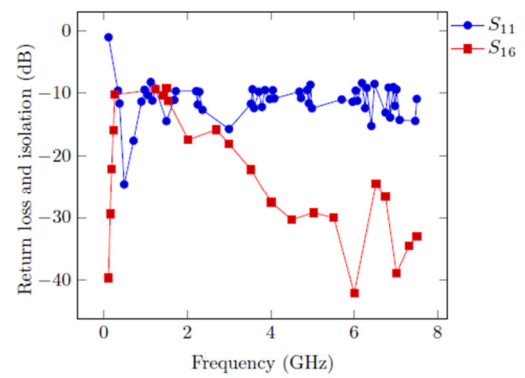
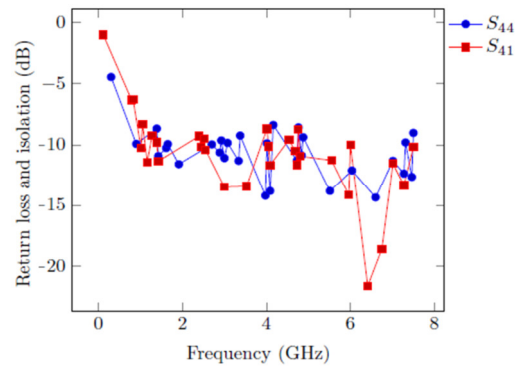
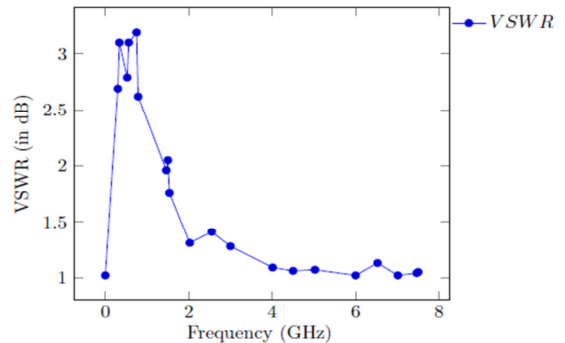
Fig. 5. Result of return loss  $S_{11}$  and isolation.Fig. 6. Result of return loss  $S_{44}$  and isolation.

Fig. 7. VSWR of radiating elements of the antenna.

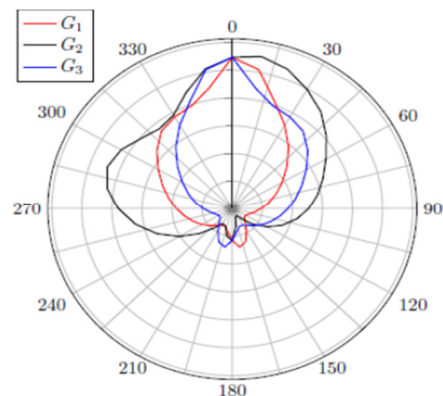


Fig. 8. Radiation pattern of the antenna.

Figure 8 shows the radiation pattern for the gain of the antenna. The gains  $G_1$ ,  $G_2$  and  $G_3$  of the antenna are measured at  $\phi=0^\circ$ ,  $\phi=90^\circ$  and  $\phi=180^\circ$ , respectively. The obtained gain of the antenna is 1.54dB. The antenna parameters, such as TARC, CC, ECC, and DG depend on the return loss and isolation [21-22]. The analysis of the antenna shows CC less than 0.153, ECC less than 0.0237, TARC less than 0.137 (-8.63dB), and 9.99 DG.

### III. DESIGN OF THE MASSIVE MIMO ANTENNA USING PARASITIC ELEMENTS

In this section, the massive MIMO antenna is designed using PEs for the frequency of 6GHz as shown in Figure 10. It consists of 8 radiating elements (Patch 1 to 8), which are designed using a rectangular shaped transmission line model [21-23]. Radiating elements 2, 4, 7, and 8 have rectangular cut slots. The radiating elements are placed close to each other and the isolation between them is high. For the isolation enhancement between the radiating elements, rectangular shaped PEs ( $P_1$  to  $P_{14}$ ) are placed among them [21-23]. The dimensions of the antenna are shown in Figure 9 and Table II. The PEs are used to negate parts of the coupled field between the radiating elements because they create the opposite coupling field, which lessens the original field and hence reduces the total coupling on the antenna. The PE structure and the radiating patch are not physically related [22]. Figure 9 exhibits the fabricated antenna (top side). This antenna is fabricated on the FR-4 substrate which has a 4.4 dielectric constant and 1.6mm height. The fabricated ground plane (back side) of the antenna is shown in Figure 11. The ground plane has 36.522mm length ( $L_{SUB}$ ) and 90mm width ( $W_{SUB}$ ). The result analysis of the antenna (for radiating elements-2 and 6) is shown in Figures 12 and 13. The antenna parameters are tested in the VNA. The measured return loss and isolation in the -6dB line of return loss are -13.96dB and -46.46dB, respectively. At the

-6dB line of return loss, the antenna operates on two frequency bands, namely, 0.82-1.76GHz and 3.03-7.5GHz. This antenna operates on UWB. Figure 14 shows the radiation pattern of the gains of the antenna. The gains  $G_1$ ,  $G_2$  and  $G_3$  of the antenna are measured at  $\phi=0^\circ$ ,  $\phi=90^\circ$ , and  $\phi=180^\circ$ , respectively. The obtained gain of the antenna is 2.73dB. The ECC is less than 0.001, CC is less than 0.01, DG is close to 10, and TARC is less than 0.015 (-8.23dB) [21-23].

TABLE II. DIMENSIONS OF A MASSIVE MIMO ANTENNA USING PES

Parameter	Dimensions (mm)
Patch length ( $L_{mp}$ )	11.34
Patch width ( $W_{mp}$ )	15.234
Width of IFL	$B = 6.41$
Length of IFL	$A = 2.942$
Width of cut slot	$B = 6.41$
Length of cut slot	$A = 2.942$
PE width ( $P_1$ to $P_{10}$ )	$D = 2.942$
PE length ( $P_1$ to $P_{10}$ )	$C = 11.34$
PE width ( $P_{11}$ to $P_{14}$ )	$E = 15.234$
PE length ( $P_{11}$ to $P_{14}$ )	$D = 2.942$
Length of substrate	$L_{SUB} = 36.522$
Width of substrate	$W_{SUB} = 90$

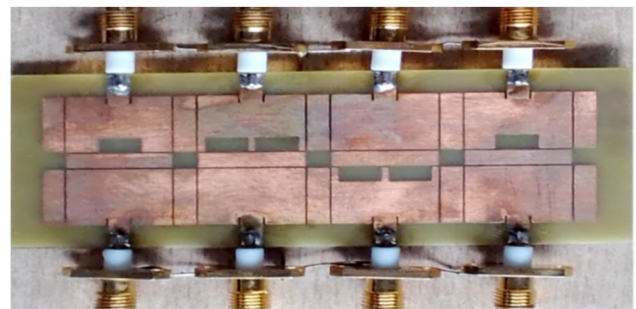
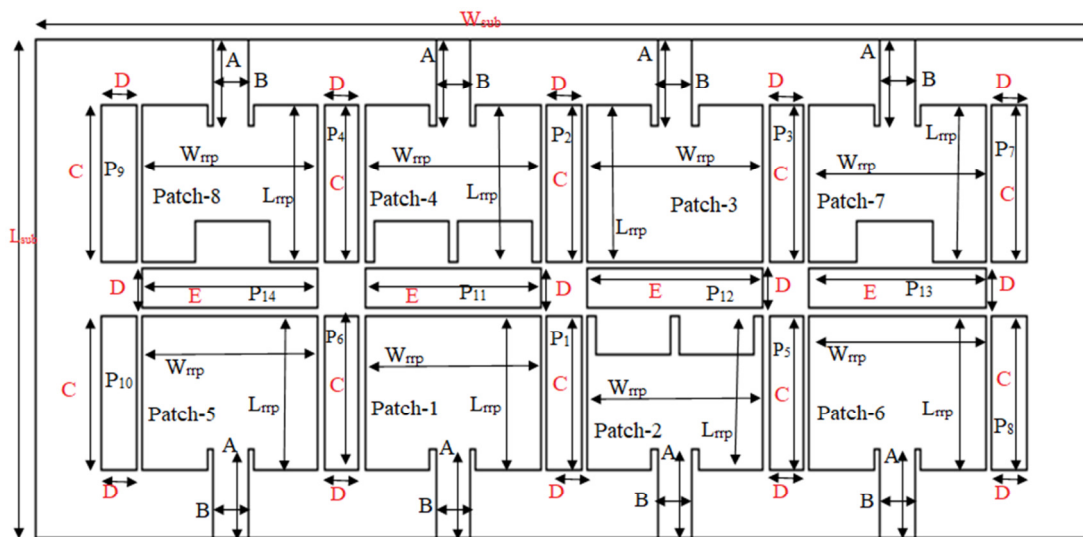


Fig. 9. The fabricated MIMO antenna (top side).



$P_1$  to  $P_{14}$  are indicates that parasitic elements

Fig. 10. Dimensions of the massive MIMO antenna using PEs.



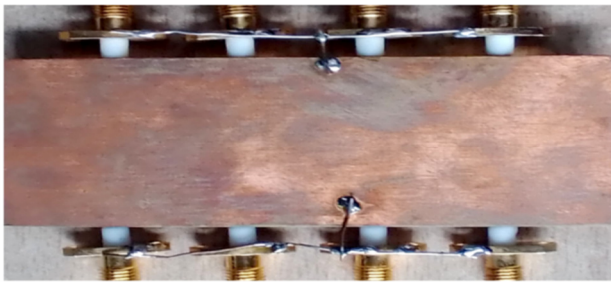


Fig. 11. The fabricated MIMO antenna (back side-ground plane).

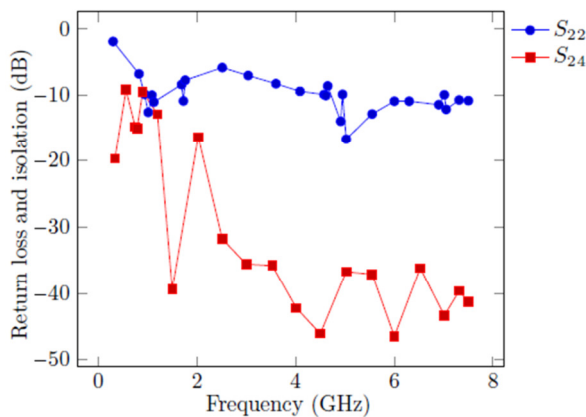
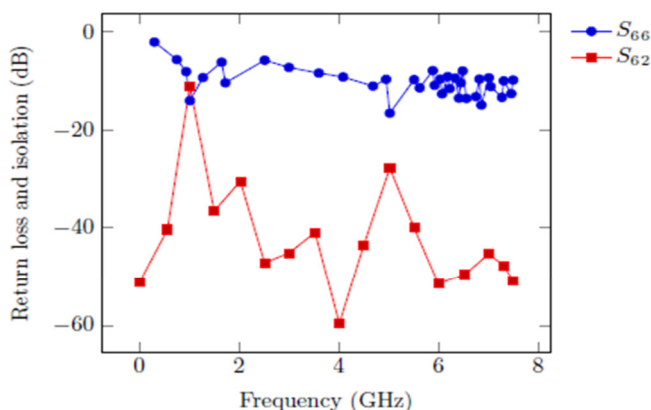
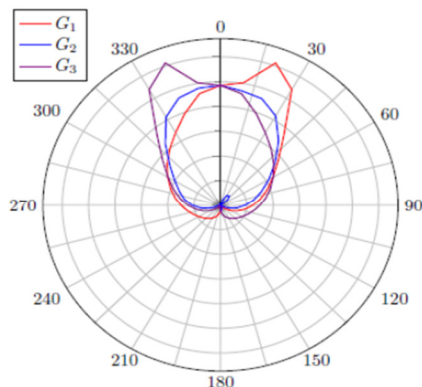
Fig. 12. Result analysis of return loss  $S_{22}$  and isolation.Fig. 13. Return loss  $S_{66}$  and isolation.

Fig. 14. Radiation pattern of the antenna.

#### IV. ANTENNA PERFORMANCE COMPARISON

The major contribution of the designed antenna is the different-shaped radiating elements. The antenna parameters are enhanced using two different techniques, namely, DGS and PEs. The designed antenna can operate on various UWB applications. The analysis shows that the cross correlation and TARC are better than the ones reported in previous research works. The comparison of the proposed antenna parameters with the other research works is shown in Table III. The proposed antenna is developed with 8 differently shaped radiating elements. It operates on a multiband frequency with a wide bandwidth.

TABLE III. COMPARISON OF THE PROPOSED WORK WITH OTHER WORKS

Ref	Method	Parameter values	Frequency band and bandwidth	Radiating patches
[2]	DGS	-10 dB RL, -12dB isolation	Single band and narrow band	8 same shaped radiating patches
[8]	DS	-20 dB RL, -15dB isolation	Dual band and narrow band	8 same shaped radiating patches
[13]	PE	-10 dB RL, -10dB isolation	Dual band and wide band	8 same shaped radiating patches
Proposed (section II)	DGS	-14.40 dB RL, -42.72dB isolation	0.33-7.5GHz and UWB with better cross correlation	8 radiating patches of different shapes
Proposed (section III)	PE	-13.96 dB RL, -46.46dB isolation	0.82-1.76GHz and 3.03-7.5GHz and UWB with better cross correlation	8 radiating patches of different shapes

#### V. CONCLUSION

In this paper, two MIMO antennas were designed. First, an antenna was developed using the DGS technique. The isolation, CC, ECC, TARC, DG, and channel capacity of the antenna were enhanced with the use of DGS. The DGS cut slots were introduced in the ground plane of the antenna. The analysis shows -14.40dB return loss and -42.72dB isolation measured in the -6dB line of return loss. At that line, the antenna operates on the frequency band of 0.33-7.5GHz. This antenna is operated on multiband applications with ultra-wide bandwidth.

Another antenna was developed using PEs. The antenna parameters CC, isolation, ECC, TARC, DG, and channel capacity were enhanced. The PEs were placed between the radiating elements. The analysis shows -13.96 dB return loss and -46.46dB isolation measured in the -6dB line of return loss, in which the antenna operates on two frequency bands, namely 0.82-1.76GHz and 3.03-7.5GHz.

#### REFERENCES

- [1] 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>.
- [2] Y. Li, C.-Y.-D. Sim, Y. Luo, and G. Yang, "High-Isolation 3.5 GHz Eight-Antenna MIMO Array Using Balanced Open-Slot Antenna

- Element for 5G Smartphones," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 6, pp. 3820–3830, Jun. 2019, <https://doi.org/10.1109/TAP.2019.2902751>.
- [3] A. A. Megahed, M. Abdelazim, E. H. Abdelhay, and H. Y. M. Soliman, "Sub-6 GHz Highly Isolated Wideband MIMO Antenna Arrays," *IEEE Access*, vol. 10, pp. 19875–19889, 2022, <https://doi.org/10.1109/ACCESS.2022.3150278>.
- [4] M. M. El-Wazzan, H. H. Ghous, S. K. El-Diasty, and M. A. Aboul-Dahab, "Compact and Integrated Microstrip Antenna Modules for mm-Wave and Microwave Bands Applications," *IEEE Access*, vol. 10, pp. 70724–70736, 2022, <https://doi.org/10.1109/ACCESS.2022.3187035>.
- [5] J. C. Dash and D. Sarkar, "Microstrip Patch Antenna System With Enhanced Inter-Port Isolation for Full-Duplex/MIMO Applications," *IEEE Access*, vol. 9, pp. 156222–156228, 2021, <https://doi.org/10.1109/ACCESS.2021.3128997>.
- [6] M. O. Dwairi, "Increasing Gain Evaluation of 2x1 and 2x2 MIMO Microstrip Antennas," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7531–7535, Oct. 2021, <https://doi.org/10.48084/etasr.4305>.
- [7] P. Kumar *et al.*, "Design of a Six-Port Compact UWB MIMO Antenna With a Distinctive DGS for Improved Isolation," *IEEE Access*, vol. 10, pp. 112964–112974, 2022, <https://doi.org/10.1109/ACCESS.2022.3216889>.
- [8] L. Cui, J. Guo, Y. Liu, and C.-Y.-D. Sim, "An 8-Element Dual-Band MIMO Antenna with Decoupling Stub for 5G Smartphone Applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 10, pp. 2095–2099, Jul. 2019, <https://doi.org/10.1109/LAWP.2019.2937851>.
- [9] M. Li and S. Cheung, "Isolation Enhancement for MIMO Dielectric Resonator Antennas Using Dielectric Superstrate," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 7, pp. 4154–4159, Jul. 2021, <https://doi.org/10.1109/TAP.2020.3044683>.
- [10] M. Li, M. Y. Jamal, L. Jiang, and K. L. Yeung, "Isolation Enhancement for MIMO Patch Antennas Sharing a Common Thick Substrate: Using a Dielectric Block to Control Space-Wave Coupling to Cancel Surface-Wave Coupling," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 4, pp. 1853–1863, Apr. 2021, <https://doi.org/10.1109/TAP.2020.3026897>.
- [11] Y. Luo, T. Yin, N. Yan, W. An, and K. Ma, "A Low-Cost Differentially Fed Dual-Mode Filtering MIMO Antenna With Enhanced Isolation Based on SISL Platform," *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 1, pp. 198–202, Jan. 2022, <https://doi.org/10.1109/LAWP.2021.3124970>.
- [12] T. Pei, L. Zhu, J. Wang, and W. Wu, "A Low-Profile Decoupling Structure for Mutual Coupling Suppression in MIMO Patch Antenna," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 10, pp. 6145–6153, Jul. 2021, <https://doi.org/10.1109/TAP.2021.3098565>.
- [13] N. O. Parchin *et al.*, "Eight-Element Dual-Polarized MIMO Slot Antenna System for 5G Smartphone Applications," *IEEE Access*, vol. 7, pp. 15612–15622, 2019, <https://doi.org/10.1109/ACCESS.2019.2893112>.
- [14] Y.-F. Tsao, A. Desai, and H.-T. Hsu, "Dual-Band and Dual-Polarization CPW Fed MIMO Antenna for Fifth-Generation Mobile Communications Technology at 28 and 38 GHz," *IEEE Access*, vol. 10, pp. 46853–46863, 2022, <https://doi.org/10.1109/ACCESS.2022.3171248>.
- [15] S. Yang, L. Liang, W. Wang, Z. Fang, and Y. Zheng, "Wideband Gain Enhancement of an AMC Cavity-Backed Dual-Polarized Antenna," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 12703–12712, Sep. 2021, <https://doi.org/10.1109/TVT.2021.3119643>.
- [16] A. E. Farahat, K. F. A. Hussein, and M. Abo. El-Hassan, "Design Methodology of Multiband Printed Antennas for Future Generations of Mobile Handsets," *IEEE Access*, vol. 10, pp. 75918–75931, 2022, <https://doi.org/10.1109/ACCESS.2022.3192548>.
- [17] Z. Chen, W. Zhou, and J. Hong, "A Miniaturized MIMO Antenna With Triple Band-Notched Characteristics for UWB Applications," *IEEE Access*, vol. 9, pp. 63646–63655, 2021, <https://doi.org/10.1109/ACCESS.2021.3074511>.
- [18] Z. Han, S. Shen, Y. Zhang, C.-Y. Chiu, and R. Murch, "A Pattern Correlation Decomposition Method for Analysis of ESPAR in Single-RF MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 4654–4668, Jul. 2022, <https://doi.org/10.1109/TWC.2021.3131612>.
- [19] 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>.
- [20] P. Bora, P. Pokkunuri, and B. T. P. Madhav, "The Design of Closed Square RR Loaded 2-Port MIMO for Dual Band Applications," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8382–8387, Apr. 2022, <https://doi.org/10.48084/etasr.4760>.
- [21] S. S. Sarade and S. D. Ruikar, "A Different Shaped Radiating Element Wide Band Multi-Band Massive MIMO Antenna for 5G/WLAN applications with Enhanced Correlation Coefficient," *IOP Conference Series: Materials Science and Engineering*, vol. 1187, no. 1, Jun. 2021, Art. no. 012017, <https://doi.org/10.1088/1757-899X/1187/1/012017>.
- [22] S. Sarade and S. R. Ruikar, "Development of a Wide Bandwidth Massive Eight Dissimilar Radiating Element Multiband MIMO Antenna for mm-Wave Application," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9166–9171, Oct. 2022, <https://doi.org/10.48084/etasr.5133>.
- [23] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. Hoboken, NJ, USA: Wiley, 2005.

## AUTHORS PROFILE



**Shrenik Sarade** has received the degree in Electronics Engineering from PVPIT, Budhagaon, Sangli, under the aegis of Shivaji University, Kolhapur, in 2007. He has received the postgraduate degree in Electronics Engineering from TKIET, Warananagar, Shivaji University, and Kolhapur, India in 2014. He is pursuing his PhD in Electronics Engineering under WCE, Sangli, affiliated to Shivaji University, Kolhapur.



**Sachin Ruikar** has received the degree in Electronics and Telecommunication from GCE, Aurangabad, in 1998. He has received a postgraduate degree in Electronics and Telecommunication Engineering from COEP, Pune University, India in 2002. He has completed his PhD in Electronics in the Shri Guru Gobind Singh Institute of Engineering Technology, SRTMU Nanded, in 2013. Presently, he is working as an Associate Professor in Electronics Engineering at WCE, Sangli, India. His research interests include image denoising with wavelet transforms, image fusion, image painting, and image superresolution.



# Space Syntax for Evaluating Attractivity and Visit Frequency

## A Comparative Study of Two Public Squares of Downtown Setif (Algeria)

**Abdelmalek Farouk Tedjari**

Habitat and Environment Laboratory, Ferhat Abbas University, Algeria  
fatedjari@univ-setif.dz  
(corresponding author)

**Messaoud Abbaoui**

Habitat and Environment Laboratory, Ferhat Abbas University, Algeria  
messaoud.abbaoui@univ-setif.dz

*Received: 27 October 2022 | Revised: 16 November 2022 | Accepted: 19 November 2022*

### ABSTRACT

The current study uses the Space Syntax innovative method for determining the causes of the attractivity and visit frequency gap between two open public spaces of downtown Setif (Algeria): Sahat El Istiklal square and Masjid Ibn Badis square. A design drawn from a map coming from the National Institute of Cartography and Teledetection (Algeria) is chosen for the configurative analysis carried out by Space Syntax through the DepthMap software. Visibility Graph Analysis (VGA) superimposed with the pedestrian real flow permits the visual integration of the analysis of each square. The analysis revealed relevant gaps in the syntactic measure values obtained for each square emphasizing their positive aspects and specific problems. Space Syntax allows the extension of the initial discussion about public squares in urban environments and to work out causality relationships between spatial configuration and human behavior.

*Keywords-Space Syntax; special configuration; public square; attractivity; visit frequency*

### I. INTRODUCTION

Public squares reflect the identity and culture of a city and are parts of its image and prestige. There is a growing demand to increase the entertainment, conviviality, and social interaction of city residents since the laws No. 2006-06 of 20 February 2006 were promulgated, and the amelioration of the the quality of the urban living environment became a priority for the Algerian State [1]. This social aspect is not a priority for the local authorities, which focus more on the development advantages. So, starting from the principle that public squares are optimally central sociability places we considered the case study of two squares of downtown Setif: Sahat El Istiklal and Masjid Ibn Badis squares, which have a gap of attractivity and visit frequency.

The current studies about the urban open space widely focus on the qualitative aspect and few studies concern the quantitative perspective. Thus, a Space Syntax-based study intends to quantitatively determine the essential causes of the inequality of attractivity and visit frequency between the two squares. In this way, we consider the accessibility to a place and the visibility degrees of its design as the indices determining social interactions and hence attractivity and visit frequency. A clear comprehension of the relationship between

the spatial configuration and popularity of a place is a must for better designs [2].

Previous researches focused on the social logic specific to the space, which impacts human behavior, using accessibility and visibility notions [3]. Authors in [4] define spatial configuration as the simultaneous relationship set existing between the parts and which design the whole, concluding that the spatial shape must be seen as a factor which contributes to the formation of the integration and segregation models in cities. Accessibility is the movement potentiality from one place to another or from a point to another. The event is perceived in the form of time, distance, and cost [5, 6]. Space Syntax theory presents three interaction forms between persons and space to observe movement, interactions with people, and watching the function of a visual field [7]. Moreover, while the visual potentialities are biological or technical, the choice modalities of what is seen or unseen depend on the social dimension [8].

For quantitatively apprehending the attractivity and visit frequency gap between the two considered urban squares, configurative analysis was carried out by Space Syntax through the DepthMap software. It carries out axial integration studies (global ( $r=n$ ) and local ( $r=3$ )) for describing visual move axes

and the Visibility Graph Analysis (VGA) for the evaluation of the integration and visual permeability levels. The results of the current research may contribute to innovative proposals for better designs in the development and planning of the urban open spaces.

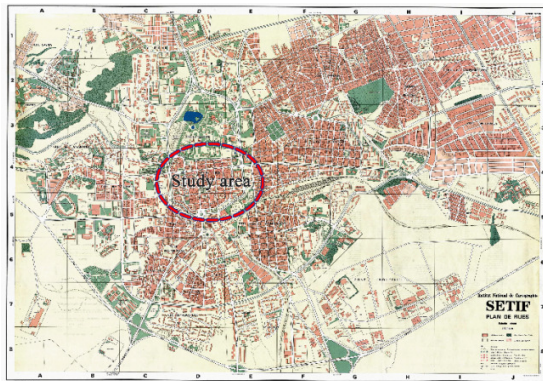


Fig. 1. The location of the study area in the city of Setif (original image © The National Institute of Cartography and Remote Sensing).

## II. METHODOLOGY

### A. Context and Case Study

Downtown Setif is an ex-nihilo creation of the French army [9]. Setif is an eastern city located at 300km from Algiers and at a height of 1100m in the Algerian highlands. Two central historical places, in a distance of 265m, are the object of the current study.



Fig. 2. Location of Sahal El Istiklal and Masjid Ibn Badis in downtown Setif. (Image © 2022 Maxar Technology, Google Earth, processed by the authors).

Sahal El Istiklal square (Figure 3A–C) is the first big historical square of downtown Setif, created in 1845 at the western entry of the city. It is distinguished by its location at the intersection of two big historical axes structuring the ancient downtown. Its shape results from the enlargement of 8 May 1945 Street constituting a 70m sided square, whose center contains a fountain monument surrounded by 4 high trees shading its space. Its framework contains Al Atik mosque, two

hotels, an administrative building, and other housing buildings with a commercial ground floor (cafeterias, restaurants and shops...) and in the upper floor some flats and professional spaces (physicians, lawyers, architects, etc.). After the works due to the tramway crossing in 2018, the square became a space reserved to pedestrians and so preventing the access to vehicles. The ground floor is paved with concrete elements and abounds in unified urban furniture with many public benches, flowers trays, and huge umbrellas.

The Ibn Badis mosque square (Figure 3B–D) is a historical square of the city, erected far from the big traffic, reinforcing its quiet space pattern [10]. It is located in the eastern side of Ben Boulaid Street, it is 22m wide and is used as a jewelry and craft commerce point. It is the result of the removal of two urban meshes, which replaced two blocks in the European district [11]. This gives a large space in the form of a rectangle whose length and width are 135 and 62m, respectively. This square is framed by communal schools and housing buildings with a ground floor limited to some small retail shops (shoes, one café, and two groceries) and some professionals at the upper floor. After the Independence, during the period 1962–1970, this square has known important modifications: The church became a mosque, the square space has been enclosed and finished as a garden, and a simple stele has replaced the memorial decorating the middle of the square. The urban furniture is almost absent in the square, except some benches and street lamps inside its garden.

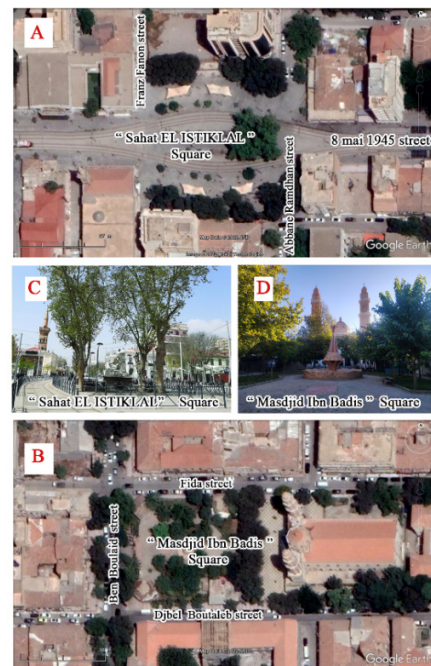


Fig. 3. A, C: The Sahal El Istiklal square, B, D: The Ibn Badis mosque square, (images A, B © 2022 Maxar Technology, Google Earth, processed by the authors).

### A. Analysis Method

As said above, public squares in urban environments are the study object of this research. The analysis of the attractivity

and the visit frequency in the considered urban punctualities of downtown Setif is based on Space Syntax as a method and theory developed by the Space Syntax Laboratory (SSL) in the seventies. Space Syntax can be summed up as a reflection on the rules and the principles of the urban organization and is based on the elements of configuration aspects, which compound the space as social behavior influencers while focusing on the questions linked to the persons' movement. In order to quantitatively explore the correlation between the spatial configuration and a social variable and using the tools given by Space Syntax, we focused on the fundamental variable of the number of persons crossing the square (pedestrian flow) considered as the real factor of attractivity and visit frequency. The person movement is figured from the axial map describing the mobility system of a given space. Through its decomposition in axes we can draw the traffic movement of open spaces of urban meshes [12]. A sketch drawn from the map given by the National Institute of Cartography and Teledetection (Algeria) is chosen for the syntactic study because it contains the whole historical downtown of Setif (urban colonial nucleus) and its two central historical squares as study topics. Its design will be completed by satellite photos from Google Earth which show the current places' state. Concerning the design of each square used for the visual analysis, it is reconstituted from the above design and from the surveys carried out on the field. A contour in dotted blue line drawn on the surrounding walls prints of the ancient city indicates the limits of the study area.

The study limit of the visual integration of each square is represented by the square space, the peripheral buildings, and its accessing streets. The design of each square will be drawn with Autocad. The global map of the whole study area and the partial maps of the public squares will be changed in DXF format then reproduced by the DepthMap software. After the evaluation of the syntactic properties and the calculation of the measures of the global ( $r=n$ ) and local ( $r=3$ ) integration plus the visual integration. DepthMap automatically assigns a series of colors with different shades: blue for the low values, yellow and green for the medium values, and red for the high values.

The axial map of the pedestrian traffic follows the pedestrians' walking which access the squares while the axial lines apply the rule of the longest and the less numerous. In the used method, the pedestrian flows systems are compounded of more or less accessible roads, which are then interpreted in function of their depth. The more accessible axes are qualified as less deep and integrated while the less accessible as deep and segregated. In theory, the most integrated lines are supposed to have a maximal flow of persons and so a greater attractivity and visit frequency. A space is considered as integrated if it tends to bring to it the other system spaces [13]. The axial map of the pedestrian traffic allowed us to compare the two squares with the medium value of the integration of all the axial elements of the global and local integration and the intelligibility of the urban system of the historical center could determine the role of each square in the urban system. In parallel, the values of the visual integration superimposed to the in situ observation (pedestrian real flow) allow us to compare the two squares and to explain the gap of the attractivity and visit frequency values.

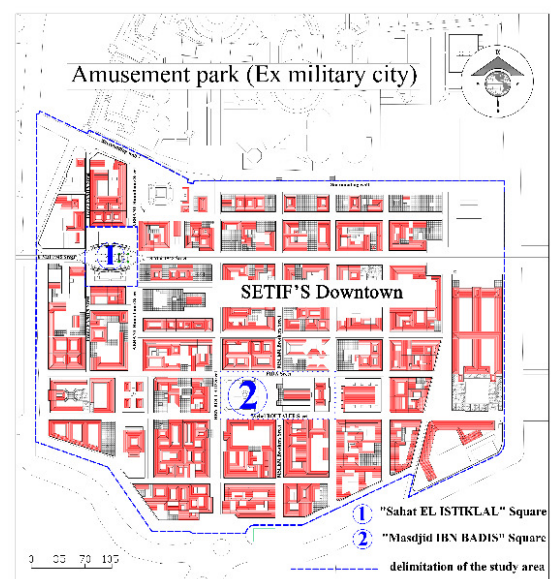


Fig. 4. Plan delimiting the study area and the two study squares. (Original image © The National Institute of Cartography and Remote Sensing. Processed by the authors).

#### B. Field Investigation: Counting and Following up the Pedestrian Real Flow

We are interested in the transit functions and the pedestrian follow up in order to measure the attractivity and the visit frequency of the considered squares. The pedestrian real flow has been counted from the numerical data and the statistical surveys. The used method is based on manual counting, which is preferable to the automatic counting [14]. It consists in counting and following up of the pedestrian flow carried out on each public square in the same time. This method can be based on the Geographical Information System (GIS) for exploring researches in the urban development fields [15]. The counting and following up were carried out by Architecture Master students. They worked as operators and were pre-trained about the counting program objectives, the data collection process, and the way of using the collect form and the follow up plans. Each operator had a chronometer and prepared investigation forms, containing all the information about the counting places and the pedestrians (hour, date, observator, direction, gender and age). Moreover, a detailed sketch of the square was given to help situate the pedestrian flow. This was necessary for the accuracy of the results during the data interpretation.

Standing at the access of the streets serving each square, one or two operators-in function of the importance of the way - carried out the counting of the persons entering the square and reported on the form the visually surveyed data. Persons leaving the square were not counted. Then, a sample of persons, determined from the counting results was followed from their entrance to the square until leaving it. The field visits and the counting development were carried out during the period between May and June 2020. For each counting the interval was not more than five minutes and ran during the working days of the week (Sunday and Monday) and the weekends (Friday and Saturday). The timing periods corresponded to rush hours: from 7.30 to 8.30 AM and 5.30 to



6.30 PM and between 8 to 9 PM. The counting results appear in Table I. The plans of the moves are exhibited by lines of different thicknesses and colors on plans containing all the ways converging to the square, which can be taken by pedestrians.

### III. RESULT ANALYSIS, AND DISCUSSION

#### A. Results, Analysis, and Discussion of the Axial Map

The axial integration map of Figure 5 obtained from the pedestrian flow shows a color difference of the axial lines of the most integrated to the system public square, which appears in dominant red color and the most segregated square in axial line mainly in blue tones. It also shows the analysis results of the global integration (HH) at the radius  $n$  and the local integration at radius 3, for the description of the move axes which intend to explore the most localized organizations like central public squares in the city. This axial map showcases the evident inequality in the integration values of the pedestrian traffic which crosses each square. So, Sahat El Istiklal square presents the highest local and global integration values (from 1.75889 to 6.05969) with maximum red colors lines. At the opposite, Ibn Badis mosque square presents a predominance of less high axial lines (from 2.48420 to 4.04988) with a blue tone primacy. It is relevant to notice in the urban system of the historical center the pertinence of important streets as integration axes of the pedestrian traffic serving each square. So, the map in Figure 5 shows that Abane Ramdane Street and 8 May 1945 Street which intersect in Sahat El Istiklal square, present the highest integration values (5.91332 and 6.06208) while Ben Boulaid Street adjacent to the western side of the Ibn Badis mosque square displays the less high integration value (4.20904). The streets' layout near each square influences the animation and their syntactic values. So we can deduce that the environment may have more effect on the square congestion than the square design itself.

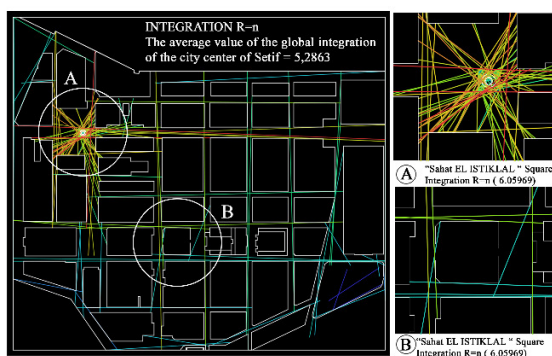


Fig. 5. Results of the axial pedestrian map analysis.

So, 8 May 1945 street being a highly integrated move axe, strongly impacts the peripheral sides of the square, especially the emergence of connection places, and the attractivity of the commercial areas located in the building ground floors framing the square, which then become multiplier spaces of attractivity and frequent visit. The fountain monument site, with its high trees shading its space at the axes intersection, presents the highest integration value. Thus, this is a highly

integrated connection space and an attractivity element giving Sahat El Istiklal square a good animation and a good user's frequent visit and permitting great mobility and a better access from its western entry. On the other hand, the wall enclosing the garden in the middle of the Ibn Badis mosque square, constitutes a barrier and a repulsive element for the pedestrians, preventing any movement.

If we put in correlation the global spatial integration and the connectivity of the map axes of the historic center urban network under a dispersion diagram form, we obtain an intelligibility value of  $R^2=0.893$ . This very high second degree measure shows that these two squares belong to a very intelligible urban system and defines the integration level of each square [17]. This can be demonstrated by the visibility of the historic center urban system and by the role of each square in terms of attractivity and frequent visit.

TABLE I. VALUES OF VISUAL INTEGRATION

Values of visual integration	Sahat El Istiklal	Masjid Ibn Badis
Minimum	5.18850	5.6545
Average	21.8004	11.1485
Maximum	36.8479	18.4607

TABLE II. COUNTS OF THE PERSONS COMING FROM THE STREETS SERVING THE SQUARE EVERY 5 MINUTES

People / 5 minutes		Sahat El Istiklal	Masjid Ibn Badis
	Street 1	328	108
	Street 2	281	46
	Street 3	29	58
	Street 4	53	28
	Street 5	14	64
	Street 6	8	45
Total		713	349
Observation		Very animated	Segregated

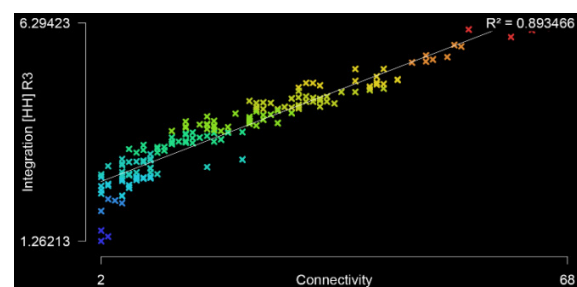


Fig. 6. Intelligibility diagram of the historic center of Setif City.

#### B. Results, Analysis, and Discussion of the Visual Integration

This part of the study deals with the square analysis compared to its own space. The Visibility Graph Analysis (VGA) maps present a chromatic differentiation of the visibility spaces from the most visually integrated spaces -in red- to the most visually segregated spaces, in blue. It shows the results of the visual integration analysis (HH) for the evaluation of the integration levels indicating the places with more or less big visual eventuality. The observation of the pedestrian flow concerned a sample of 500 persons coming from the streets of surrounding neighborhoods and entering the

square. The results of the visual integration of the syntactic study in Table II are compared to those of the counting and followed by the real move of the persons in the square. From the results of the pedestrian counting, it appears that greater pedestrian number passes from Sahat El Istiklal square, which makes it a very animated and visited space compared to the Ibn Badis mosque square. Ibn Badis mosque square was at first a church and had been designed as a quiet place. The counting results of Table II are in congruence with the visual integration values of Table I. The results of the comparison between the diagrams of the pedestrians real movement and the visual integration plus the square spatial configuration, show that the visual integration values are related with both the pedestrian movement and the spatial configuration as analyzed by the Space Syntax method (Figures 7, 8). Actually, the fact that Sahat El Istiklal square has been designed to be at the exact place of the enlargement of the most integrated axe of the urban system (8 May 1945 Street), enhances its attractivity and fluidity. The fountain monument and the dominant building (the mosque) as attractive elements are located at the eastern and western entry of the square. So, they permit good animation, balancing of the syntactic values of the walk in the middle of the square and all along its perimeter with an expressive reduction of the values around the fountain. These attractive elements of the square are positioned in the visibility axe of the most integrated move. The maximal values of the visual integration in Table I confirm that Sahat El Istiklal square offers good visibility and permits to clearly distinguish the borders of the peripheral convex buildings facilitating the interaction. In the visual integration map, the fountain monument erected on a stone pedestal and surrounded by four trees shading its space is considered as a barrier to the square visibility and has very low visual integration values (from 5.572 to 1.4654). Concerning Ibn Badis mosque square, due to its location adjacent to the eastern side of the Ben Boulaid Street, it cannot capture the pedestrian flow coming from this axe and presents low visual integration values. The building dominating the square was originally designed to be located in the visibility axes of the person movement and represented one of the most beautiful urban paintings decorating Setif City [18]. The central area of Sahat El Istiklal square concentrates the highest visual integration value (36.8479), gradually regressing as we approach the periphery of the buildings surrounding the square (Figure 7). The central area is the most visible area from which we can visually dominate the rest of the square, which makes it the most potential animated and visited space. On the other hand, for Ibn Badis mosque square, the buildings located in the square perimeter display low visual integration (9.09819), due to the construction of the garden enclosure wall. It contains a great number of high trees all around the square center constituting a physical barrier preventing continuity in visual integration. Moreover, in spite of the low visibility dominating this square, some parts inside the garden or at the mosque entry present moderately high integration values, and other parts situated at the intersection of the streets leading to the square present high values, which could imply a real possibility of animation. In the urban system of the historic center, each square is identified by its diversity and its functions which give vitality to urban life [19]. So the commercial frame gives Sahat El Istiklal square a role of very

animated and visited space, while the Ibn Badis mosque square is limited to some modest shops and is a very little attractive and less visited space. These in situ observation results are in correlation with the counting values of the pedestrian move (Table II).

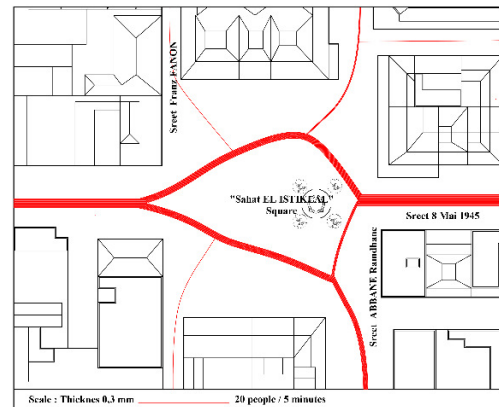
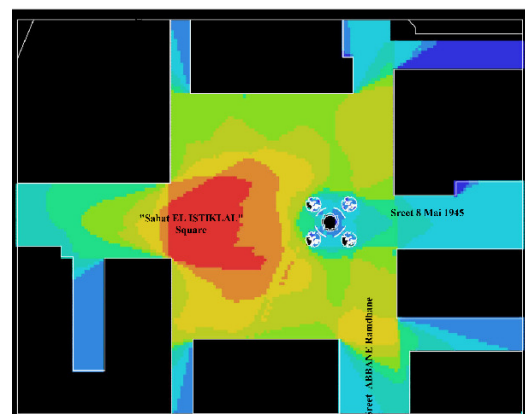


Diagram of a pedestrian pathway for 500people



Visual integration of the SAHAT EL ISTIKLAL square

Fig. 7. Juxtaposition of the pedestrian movement scheme and the visual integration plan in Sahat El Istiklal square.

#### IV. CONCLUSION

The spatial arrangement of downtown Setif City has a rectangular shape laying out ways on which the public squares are distributed. The methodology used for the comparative study of the two central squares, as spatial models through the Space Syntax method, focused on the study of supposed persons' flows as a necessary factor for real attractivity and visit frequency estimation. This permits to quantitatively estimate the attractivity and visit frequency differences for a future development of the squares. Other researches in the Space Syntax field, in different urban contexts, proved that the spatial integration level is a revelator of the pedestrian urban mobility. The inequalities in attractivity and visit frequency perceived in each square are the consequence of the configuration and the urban role originally devoted to each one.



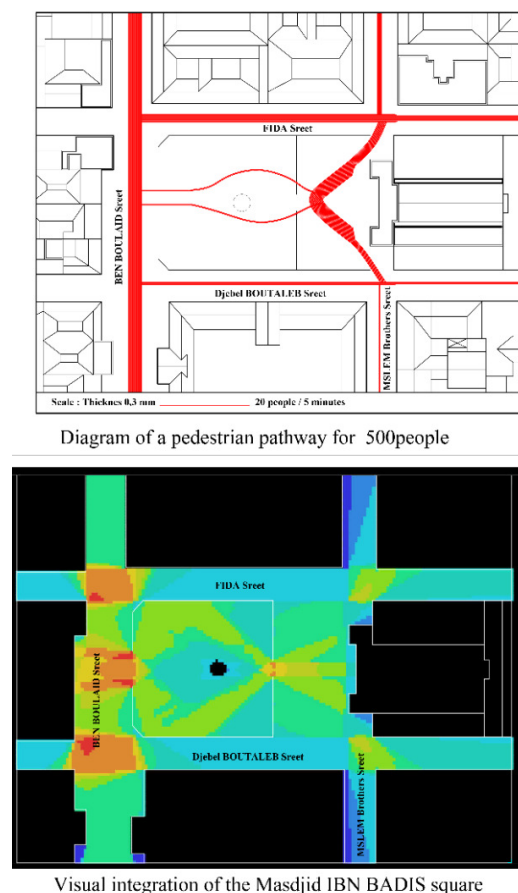


Fig. 8. Juxtaposition of the pedestrian movement scheme and the visual integration plan in Masjid Ibn Badis square.

So, Sahat El Istiklal square is destined to become the most animated and visited urban space of the city due to its design, location, and the number of shops surrounding its space. It has been designed in order to be the anchor link of the two most animated historical and commercial axes of the city. The syntactic measures confirm that the two axes, Abane Ramdane Street and 8 May 1945 Street are very integrated in the axial integration  $r=n$  of the global system. Thus, the square is highly visually and globally integrated and its attractiveness and visiting space are perfectly adapted to the pedestrians' mobility. Concerning Ibn Badis mosque square, even though it is situated in the eastern side of Ben Boulaid Street, it cannot catch its animated flow. This square has been created only for serving its dominant building (Sainte Monique, church now Ibn Badis mosque). Actually, it was intended to be used as a quiet space, situated apart from the great traffic and with a limited number of surrounding shops. The syntactic results prove that this square owns the lowest visual and global integration values. Its few attractive spaces seem more adapted to a mechanical traffic.

This animation and frequent visit gap does not absolutely constitute a factor proving the success of the first square compared to the second one. Actually, each one finds its originality in the social economic and political role assigned in the urban system of the center of Setif City. The obtained

results show that the connection to Sahat El Istiklal square is essentially assured by the square size, the peripheral buildings, the dominant building (the mosque), the monument and the urban context, which fully participate to its attractivity. Concerning the connection to Ibn Badis mosque square, it is more assured by the indoor spaces of the dominant equipment and the schools surrounding the square. This square remains segregated and less visited, especially in the night when the stores and schools are closed.

The development of downtown Setif goes through the increase of the global and visual integration values of the squares in the urban global context. Space Syntax, through DepthMap software, allowed us to demonstrate the attractivity and traffic inequalities. It allowed to extend the debate initiated about public squares in urban environments and to lay down causality relationships between spatial configuration and human behavior in order to contribute to the elaboration of innovating recommendations.

#### ACKNOWLEDGEMENTS

The authors would like to thank the Habitat and Environment Laboratory of Ferhat Abbas University for its contribution to this research.

#### REFERENCES

- [1] I. Kaouche and A. Boussoulaim, "À La Recherche Des Qualités Du Vécu Des Espaces Publics ; Cas Des Places Publiques En Algérie," *Courrier Du Savoir*, no. 20, pp. 109–118, Dec. 2015.
- [2] F. Fareh and D. Alkama, "The Effect of Spatial Configuration on the Movement Distribution Behavior: The Case Study of Constantine Old Town (Algeria)," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9136–9141, Oct. 2022, <https://doi.org/10.48084/etasr.5169>.
- [3] A. Rahmane and M. Abbaoui, "The Architectural Genotype Approach in Contemporary Housing (1995 to 2010): The Case Study of Setif, Algeria," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6810–6818, Feb. 2021, <https://doi.org/10.48084/etasr.4006>.
- [4] B. Hillier and L. Vaughan, "The city as one thing," *Progress in Planning*, vol. 67, no. 3, pp. 205–240, Apr. 2007.
- [5] C. Cauvin, H. Reymond, and R. Kleinschmager, *L'Espace géographique des villes. Pour une synergie multistrates*. Paris, France: Economica, 1998.
- [6] P. Merlin, *Méthodes Quantitatives Et Espace Urbain*. Paris, France: Persée - Portail des revues scientifiques en SHS, 1974.
- [7] B. Hillier, *Space is the machine: A configurational theory of architecture*. CreateSpace Independent Publishing Platform, 2015.
- [8] J.-N. Tardy, "Visibilité, invisibilité. Voir, faire voir, dissimuler," *Hypothèses*, vol. 10, no. 1, pp. 15–24, 2007, <https://doi.org/10.3917/hyp.061.0015>.
- [9] A. Prenant, "Facteurs du peuplement d'une ville de l'Algérie intérieure : Sétif," *Annales de Géographie*, vol. 62, no. 334, pp. 434–451, 1953, <https://doi.org/10.3406/geo.1953.13052>.
- [10] A. Picard, "Lotissements et colonisation : Algérie, 1830-1970," *Villes en Parallèle*, vol. 14, no. 1, pp. 214–237, 1989, <https://doi.org/10.3406/vilpa.1989.1062>.
- [11] A. Camborieu, *Sétif et sa région : essai de monographie historique, géographique et économique*. 1978.
- [12] B. Hillier, A. Penn, J. Hanson, T. Grajewski, and J. Xu, "Natural Movement: Or, Configuration and Attraction in Urban Pedestrian Movement," *Environment and Planning B: Planning and Design*, vol. 20, no. 1, pp. 29–66, Feb. 1993, <https://doi.org/10.1068/b200029>.
- [13] B. Hillier and J. Hanson, *The Social Logic of Space*. Cambridge, UK; New York, NY, USA: Cambridge University Press, 1989.

- 
- [14] *Méthodes de comptages piétons dans l'espace public*. Brussels, Belgium: Service Public Regional de Bruxelles, 2015.
- [15] K. Loumi and A. Redjem, "Integration of GIS and Hierarchical Multi-Criteria Analysis for Mapping Flood Vulnerability: The Case Study of M'sila, Algeria," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7381–7385, Aug. 2021, <https://doi.org/10.48084/etasr.4266>.
- [16] A. Hallil and A. Redjem, "Assessment of Urban Vulnerability to Flooding Using Multi-Criteria Analysis: The Case Study of El Bayadh City, Algeria," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8467–8472, Apr. 2022, <https://doi.org/10.48084/etasr.4828>.
- [17] A. R. Bouzgarrou, "Analyse des formes morpho-fonctionnelles urbaines: mise en place d'un indicateur de mutations paysagères de la ville de Monastir entre 1956 et 2013," Ph.D. dissertation, L'université De Bretagne Occidentale, Loire, France, 2019.
- [18] R. Unwin, *Étude pratique des plans de villes. Introduction à l'art de dessiner les plans d'aménagement et d'extension*. Paris, France: Éditions Parenthèses, 1981.
- [19] J. Jacobs, *The Death and Life of Great American Cities*. New York, NY, USA: Vintage, 1992.

# Cooley-Tukey FFT Algorithm based on GDFT for Phasor Estimation by PMU under Power Quality Disturbances

**Mohamed Abbaci**

Laboratory of Advanced Electronic Systems (LSEA), University of Medea, Algeria | PTAPC Ouargla, CRAPC, Algeria  
abbaci.mohamed@crapc.dz  
(corresponding author)

**Mohamed Ould Zmirli**

Laboratory of Advanced Electronic Systems (LSEA), University of Medea, Algeria  
m\_zmirli@yahoo.fr

*Received: 30 September 2022 | Revised: 13 November 2022 | Accepted: 19 November 2022*

## ABSTRACT

Due to the high penetration of renewable energy sources, such as photovoltaic panels and wind turbines, in addition to the use of different electric power supplies in the power grid, there are major disturbances in the forms of electric waves. These variations and disturbances must be monitored and controlled for the efficient management of transmission and distribution of electrical energy, safety, and electrical protection systems. Nowadays, Phasor Measurement Unit (PMU) technology is an essential tool to develop the supervision, protection, and control of the electrical power grid. PMUs measure the amplitude and angle of current and voltage waveforms on a Coordinated Universal Time (UTC) time scale and speedily measure the fundamental frequencies and their rates of change using fast and accurate estimation algorithms. This paper presents a phasor estimation using a Modified Cooley-Tukey Fast Fourier Transform algorithm based on the Generalized Discrete Fourier Transform (DFT) used in PMUs, using simulations in MATLAB. This algorithm was utilized to accelerate and simplify the computation of DFTs. To validate the performance under waveform disturbances, several tests with different waveforms and disturbances were simulated and interpreted according to the standard and compared with DFT.

**Keywords-**phasor measurement unit; rate of change of frequency; total vector error; DFT; GDFT; CTGDFT

## I. INTRODUCTION

Phasor Measurement Units (PMUs) are widely used for many years in electrical systems for control and monitoring, with accurate and fast fundamental component measurements [1, 2]. The phasor estimation is the main component of a PMU [3], so DFT is the most widely used algorithm [1, 4]. PMUs were introduced in response to the need for more efficient and secure monitoring devices in power grids. Since the developments in synchronized measurement technology, wide-area measurement systems are suitable to follow the dynamic behavior of electrical power systems using PMUs and PDC hierarchically [5, 6]. PMUs precisely measure the amplitudes and phase angles of the waveforms of the electrical network (current, voltage, and frequency) at various points [7, 8] using time stamping. Measured and calculated data are transmitted through a synchronized GPS transmission medium to a PDC [9]. Information from multiple PMUs located throughout the electrical power system is analyzed and processed at the PDC to detect electrical disturbances [10] and improve the protection

and control functions of an electric power system [6]. The IEEE 1344-1995 was the first published standard for conventions relating to original synchrophasor measurements [11, 12]. In 2005, IEEE C37.118-2005 was introduced and replaced the existing synchrophasor standard [1, 12]. In 2011, two new standards were published to address the measurement requirement (IEEE C37.118.1-2011) and synchrophasor communications (IEEE C37.118.2-2011) [11, 13, 14].

This study investigated the modified Cooley-Tukey Fast Fourier Transform (FFT) algorithm, based on a Generalized Discrete Fourier Transform (GDFT) framework, for PMU phasor estimation. The conventional Cooley-Tukey algorithm accelerates and simplifies the DFT calculation, but has disadvantages such as slow dynamics and sensitivity to frequency variations. The use of the Cooley-Tukey FFT algorithm based on the GDFT was proposed in [15] to overcome these drawbacks. Simulations were carried out to highlight the performance of the suggested algorithm over standard DFT techniques.

## II. PHASOR MEASUREMENT ARCHITECTURES

### A. Block Diagram of PMU

The functional diagrams of PMUs may differ from one manufacturer to another. Figure 1 shows a basic block diagram of a PMU [13]. A PMU generally consists of three main blocks: measurement, calculation, and communication [16].

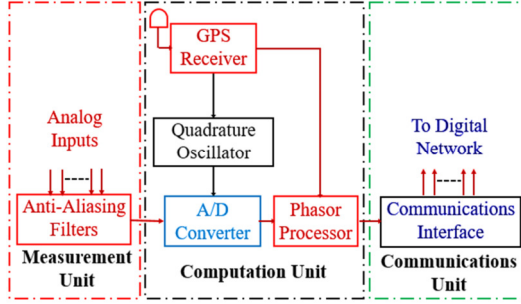


Fig. 1. Phasor Measurement Unit Components.

- The Measurement Unit receives analog data inputs such as three-phase voltages, three-phases, and neutral currents from the secondary measurement transformers (VT and CT), which are later filtered to eliminate and prevent aliasing errors by antialiasing filters [13, 17].
- The Computation Unit includes Analog-to-Digital Converters (A/D), Quadrature Oscillator, and Phasor Processor [18, 20].
- The computed synchrophasor and other information are transmitted to the PDC to be interchanged between other PMUs for monitoring, control applications, and system protection using communication units [13, 16].

### B. Wide Area Measurement Systems

The Wide Area Monitoring System (WAMS) collects real-time information on the status of the network at strategic points, with precise time stamping by GPS satellites. WAMS improves network analysis, integrating PMUs to detect any instability. The WAMS uses synchronized measurement data and modern communication systems to supervise, monitor, and analyze the current state of the wide-area power system and serves for the operation, control, and protection of the power system in real-time [18, 19]. WAMS systems mainly consist of three components [13]: Phasor Measurement Unit, Communication network, and Phasor Data Concentrator.

#### 1) Phasor Measurement Unit (PMU)

The PMU is a measuring device for electrical power systems, capable of performing phasor measurements of voltages and currents in large and distributed power system networks. The voltage and current waveforms of the power system are sinusoidal. A sinusoidal waveform is written as:

$$x(t) = X_m \cos(2\pi ft + \varphi) \quad (1)$$

where  $X_m$  is the amplitude,  $f$  is the frequency,  $\varphi$  is the angle, and  $t$  is the time. The representation of the phasor in the complex plane is:

$$X = \frac{X_m}{\sqrt{2}} e^{j\varphi} \quad (2)$$

$$X = \frac{X_m}{\sqrt{2}} (\cos \varphi + j \sin \varphi) = X_r + jX_i \quad (3)$$

where  $X_r$  and  $X_i$  are the real and the imaginary components of the complex form, and  $\frac{X_m}{\sqrt{2}}$  is the RMS value. Figure 2 shows the phasor representation of a sinusoidal input signal at the angular frequency in the complex plane [10, 17]. In the case of a phase current or voltage, the projection of the phase vector on the ordinate axis (imaginary axis) gives the value of the current or voltage at that instant.

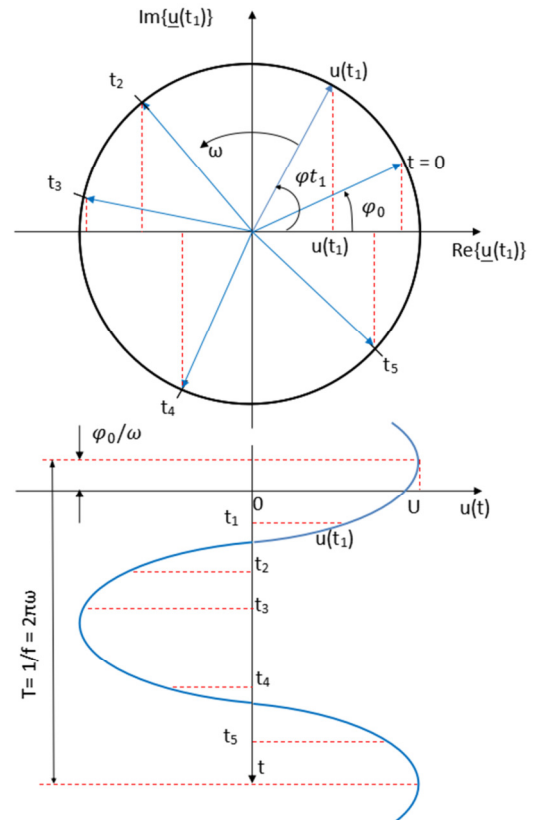


Fig. 2. Phasor representation of a sinusoidal signal.

#### 2) Communication Network

PMUs are usually located in substations that are geographically distant from the PDC. Communication networks are required to transmit information and measurements collected by PMUs from the substations to the control center [18]. Power grid monitoring and control require a high-speed communication framework that allows reliable, secure, and fast sharing of synchronized monitoring data between PMUs and PDC.

#### 3) Phasor Data Concentrator

Figure 3 shows the implementation of PDCs in a wide area monitoring system. The PDC collects discrete events and phasor data from PMUs, and probably other PDCs, and transmits them to other applications via a communications

protocol [10, 21]. Depending on the location of the PMU, the communication link between the PMU and the PDC may be a wired connection, a serial or Ethernet cable, or a wireless link.

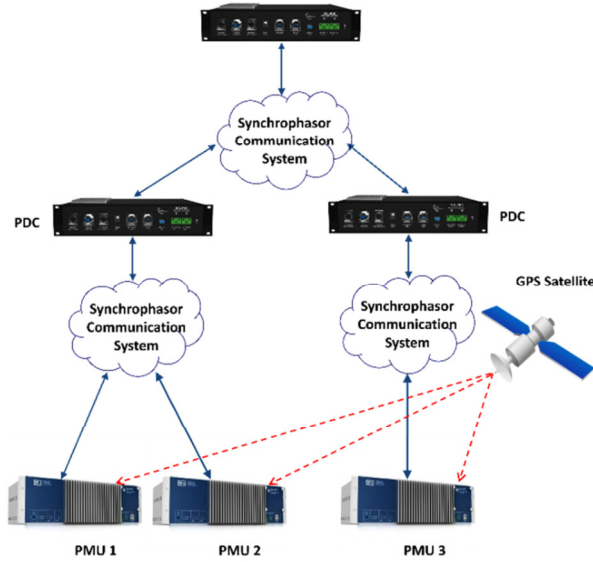


Fig. 3. Implementation of PDCs in WAMS.

Each PMU has two communication interfaces: One interface for communication of the PMUs via IEEE C37.118 and another to communicate with the digital control system. The central evaluation system collects the data, archives them, and displays them on a graphical user interface. This system can also be used to automatically monitor the electrical system and redistribute information to other PDCs or a numerical control system [10]. Communication of PMUs with PDCs, according to the IEEE C37.118 standard, is a client-server communication in which the PDC works as a client and the PMU as a server [10, 22].

### III. PERFORMANCE INDICES OF MEASUREMENT EVALUATIONS

#### A. Total Vector Error (TVE)

C37.118.1-2011 defines the total vector error to calculate the phasor estimation error and to evaluate the performance of the phasor estimation. TVE compares the theoretical value of an input signal with its estimated value at the same time [10]. TVE is calculated by considering the error of estimating the amplitude and phase angle [13]. The TVE at time  $n$  is defined as :

$$TVE(n) = \sqrt{\frac{(\hat{X}_r(n) - X_r(n))^2 + (\hat{X}_i(n) - X_i(n))^2}{(X_r(n))^2 + (X_i(n))^2}} \quad (4)$$

where  $\hat{X}_r(n)$  and  $\hat{X}_i(n)$  are the real and the imaginary components of the estimated phasor, and  $X_r(n)$  and  $X_i(n)$  are the theoretical values at time  $(n)$ . Figure 4 shows the TVE representation.

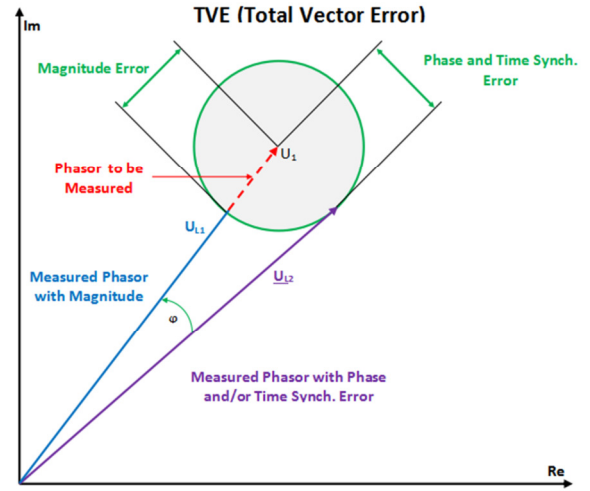


Fig. 4. Calculation of TVE.

#### B. Frequency Error

In addition to the synchrophasor, PMUs also calculate the power system's frequency. Frequency Error (FE) is the difference between the estimated and the theoretical frequency. The sinusoidal signal of (1) may be written as [18]:

$$x(t) = X_m \cos(\theta(t)) \quad (5)$$

The frequency of the signal (5) is determined as:

$$f(t) = \frac{1}{2\pi} \left( \frac{d\theta(t)}{dt} \right) \quad (6)$$

FE is defined as:

$$FE = |f - \hat{f}| \quad (7)$$

#### C. Rate Of Change Of Frequency Error (RFE)

ROCOF, measured by PMU at a given time instant, is determined by:

$$ROCOF(t) = \frac{d\left(\frac{1}{2\pi} \frac{d\theta(t)}{dt}\right)}{dt} = \frac{df(t)}{dt} \quad (8)$$

RFE is the difference between the theoretical and the estimated ROCOF value at a particular time instant [11]:

$$RFE = \left| \left( \frac{df}{dt} \right) - \left( \frac{d\hat{f}}{dt} \right) \right| \quad (9)$$

### IV. PMU ESTIMATION ALGORITHMS

Most phasor estimation algorithms utilized in commercial PMUs depend on the DFT estimation technique. The IEEE C37.118 standard does not recommend the use of a special algorithm or measurement in the PMU. Consequently, various estimation methods and algorithms are used in PMUs [12].

#### A. DFT Algorithm

A sinusoidal waveform  $x(t)$  of  $kf_0$  frequency with a Fourier series can be written as:

$$x(t) = a_k \cos(2\pi k f_0 t) + b_k \sin(2\pi k f_0 t) \quad (10)$$

$$x(t) = \left\{ \sqrt{a_k^2 + b_k^2} \right\} \cos(2\pi k f_0 t + \phi) \quad (11)$$



where  $\phi = \tan^{-1}(-b_k/a_k)$ . The phasor representation becomes:

$$X_k = \frac{1}{\sqrt{2}} \left\{ \sqrt{(a_k^2 + b_k^2)} \right\} e^{j\phi} \quad (12)$$

and the complex form of the phasor becomes:

$$X_k = \frac{1}{\sqrt{2}} (a_k - jb_k) \quad (13)$$

By applying the relationship of the Fourier series coefficients, the phasor formula is determined by:

$$X_k = \frac{1}{\sqrt{2}} \frac{2}{N} \sum_{n=0}^{N-1} x_n e^{-\frac{j2\pi kn}{N}} \quad (14)$$

and also:

$$X_k = \frac{\sqrt{2}}{N} \sum_{n=0}^{N-1} x_n \left\{ \cos\left(\frac{2\pi kn}{N}\right) - j \sin\left(\frac{2\pi kn}{N}\right) \right\} \quad (15)$$

where  $N$  is the number of samples,  $x_n$  is the input signal,  $n$  is the sample number and  $k$  is the harmonic index.

$$X_r = \frac{\sqrt{2}}{N} \sum_{n=0}^{N-1} x_n \cos\left(\frac{2\pi kn}{N}\right) \quad (16)$$

$$X_i = \frac{\sqrt{2}}{N} \sum_{n=0}^{N-1} x_n \sin\left(\frac{2\pi kn}{N}\right) \quad (17)$$

The phasor  $X_k$  becomes:

$$X_k = X_r + jX_i \quad (18)$$

The phasor magnitude of the phasor is given by:

$$|X_k| = \sqrt{X_r^2 + jX_i^2} \quad (19)$$

And finally, the phase angle of the phasor is defined by:

$$\phi = \tan^{-1} \frac{X_i}{X_r} \quad (20)$$

### B. Proposed Algorithm

This study investigated the Cooley-Tukey FFT algorithm, based on a Generalized DFT (CTGDFT), for phasor estimation. GDFT is described by:

$$X_k = \sum_{n=0}^{N-1} \omega_N^{(n+a)(k+b)} x_n \quad (21)$$

where  $\omega_N = e^{-\frac{2\pi j}{N}}$ , and  $a$  and  $b$  are two arbitrary complex numbers ( $a = b = 0$  gives the ordinary DFT) [23]. The second transform is presented by the Cooley-Tukey FFT algorithm to accelerate and simplify the computation of DFTs [24, 25]. This algorithm is based on the factorization of  $N$ , the length of the DFT as a product of a number less than  $N$ . Supposing that  $N$  is not a prime number, it is possible to write it as  $N = N_1 N_2$ , where the two factors are greater than 1. The DFT formula is:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{j2\pi kn}{N}} \quad (22)$$

This algorithm begins by dividing the range of integers from 0 to  $N-1$  in two different ways. For time index  $n$  or frequency index  $k$ , the division is divided into  $N_1$  intervals of  $N_2$  length each. The variables  $n$  and  $k$  are expressed as:

$$\begin{cases} n = n_1 + n_2 N_1; 0 \leq n_1 \leq N_1 - 1; 0 \leq n_2 \leq N_2 - 1 \\ k = k_1 N_2 + k_2; 0 \leq k_1 \leq N_1 - 1; 0 \leq k_2 \leq N_2 - 1 \end{cases} \quad (23)$$

The DFT Formula can be written in terms of  $n_1, n_2, k_1, k_2$  as:

$$X_{k_1 N_2 + k_2} = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x_{n_1 + n_2 N_1} e^{-\frac{j2\pi(n_1 + n_2 N_1)(k_1 N_2 + k_2)}{N}} \quad (24)$$

So, also:

$$X_{k_1 N_2 + k_2} = \sum_{n_1=0}^{N_1-1} \omega_N^{n_1 k_1} \omega_N^{n_1 k_2} \sum_{n_2=0}^{N_2-1} x_{n_1 + n_2 N_1} \omega_N^{n_2 k_2} \quad (25)$$

where  $\omega_N = e^{-\frac{j2\pi}{N}}$ ,  $\omega_{N_1} = e^{-\frac{j2\pi}{N_1}}$ , and  $\omega_{N_2} = e^{-\frac{j2\pi}{N_2}}$ . The inner summation was reduced to a standard DFT of  $N_2$  size. Furthermore, the splitting of the first exponential into two factors unveils that the outer summation is also a standard DFT of  $N_1$  size, of the product of the DFT of  $N_2$  size multiplied by the twiddle factors. Applying the composition  $N = N_1 N_2$ , the Cooley-Tukey re-indexings  $n_1 + n_2 N_1$  and  $k_1 N_2 + k_2$  can be used in (21). Then, the formula utilized in the modified Cooley-Tukey FFT algorithm based on GDFT is given by:

$$X_k = \omega_N^{ka} \sum_{n_1=0}^{N_1-1} \omega_{N_1}^{n_1 k_1} \omega_N^{n_1 k_2} \left( \sum_{n_2=0}^{N_2-1} x_n \omega_{N_2}^{n_2 k_2} \omega_N^{nb} \right) \quad (26)$$

where  $k = k_1 N_2 + k_2$  and  $n = n_1 + n_2 N_1$ .

## V. SIMULATION RESULTS OF DISTURBANCE FAULTS

### A. Simulation Parameters

The selected simulation parameters were:

- Nominal frequency:  $f_0 = 50\text{Hz}$
- Number of samples: 400 samples/cycle
- Sample rate: 20kHz

The input signal is generated from a two-phase voltage (line voltage) of a secondary Potential Transformer ( $U_{RMS} = 100\text{V}$ ) with a cosinusoidal form, i.e. the phase AB.

### B. Disturbance Faults and Evaluation Performance

To evaluate the operation reliability of the PMU, different disturbance faults should be examined with the application of the proposed algorithm. As the performance and evaluation of PMUs are tested under disturbed power system conditions, 3 different disturbance faults were examined. For each case, voltage input signal, phasor magnitude, phasor angle, TVE, estimated frequency, and ROCOF were simulated. This paper shows only the simulation results of the TVE and the estimated frequency to evaluate the performance of the PMU. The frequency is estimated according to (6) and tracking is carried out by iterative DFT with re-sampling, as presented in [26], with a modification in the first cycle of frequency calculation (3 phasors are calculated in the first cycle to obtain the measured frequency). The simulation results of the disturbance faults were as follows:

#### 1) Disturbance Faults 1

The idea of this steady-state test was to examine the effect of the input signal magnitude variations and its phase angle around its nominal values while keeping the frequency constant. This test signal was then generated with a modified magnitude of  $\pm 20\%$  of the nominal and a phase angle of  $45^\circ$ , as in [27], at  $t_1 = 0.2\text{s}$  and  $t_2 = 0.5\text{s}$ :

$$v(t) = 80\sqrt{2} \cos(2\pi f + \pi/4) \quad (27)$$

In the first disturbance fault, a variation of amplitude and phase were introduced to the voltage input signal. Figure 5 shows that the amplitude and phase values measured by the PMU were identical to the theoretical values, the evaluation parameters were in the norms, and TVE was almost 0%. Table I shows that the maximum value of the Frequency Error (FE) of the proposed algorithm was 0.0003Hz.

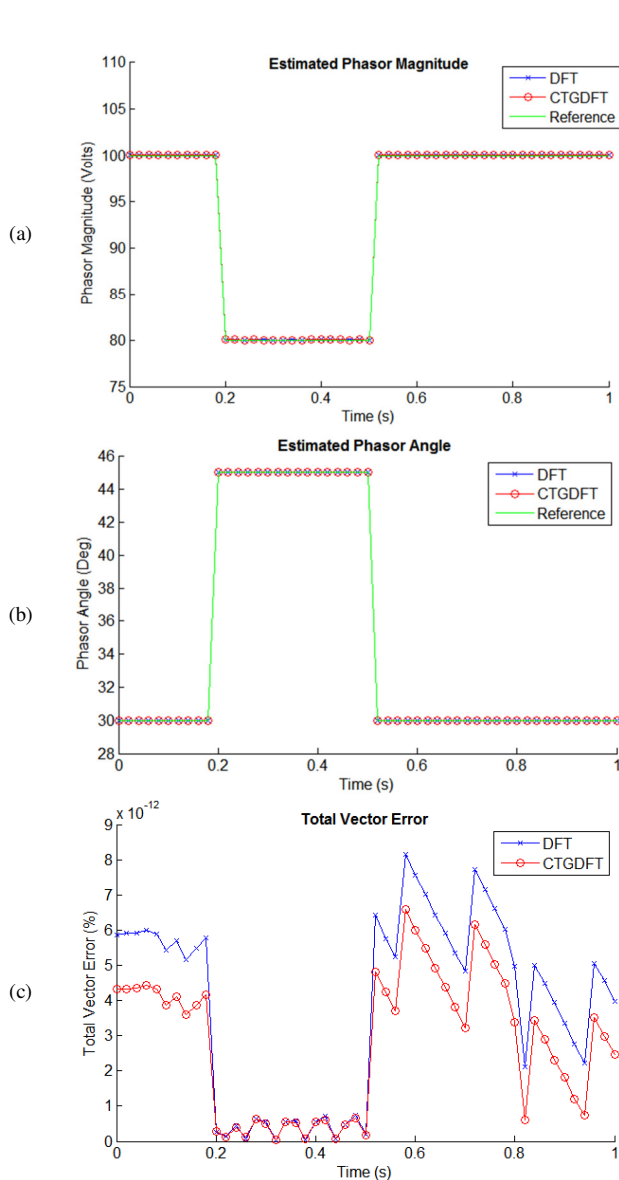


Fig. 5. Simulation results under Disturbance Faults 1: (a) estimated phasor magnitude, (b) estimated phasor angle, (c) total vector error.

TABLE I. ESTIMATED FREQUENCY, FREQUENCY ERROR, AND ROCOF ERROR UNDER DISTURBANCE FAULTS 1

Method	Frequency (Hz)		FE (Hz)	RFE (Hz/s)
	Min	Max		
DFT	50	50	0	0.00002
CTGDFT	49.9997	50	0.0003	0.00034

## 2) Disturbance Faults 2

This test examined the variation of the magnitude of the voltage input signal when the input signal frequency is 45Hz:

$$v(t) = 80\sqrt{2} \cos(2\pi f + \pi/6) \quad (28)$$

Figure 6 shows the efficiency of the results obtained. Table II shows that the maximum FE was on the order of 0.0003Hz.

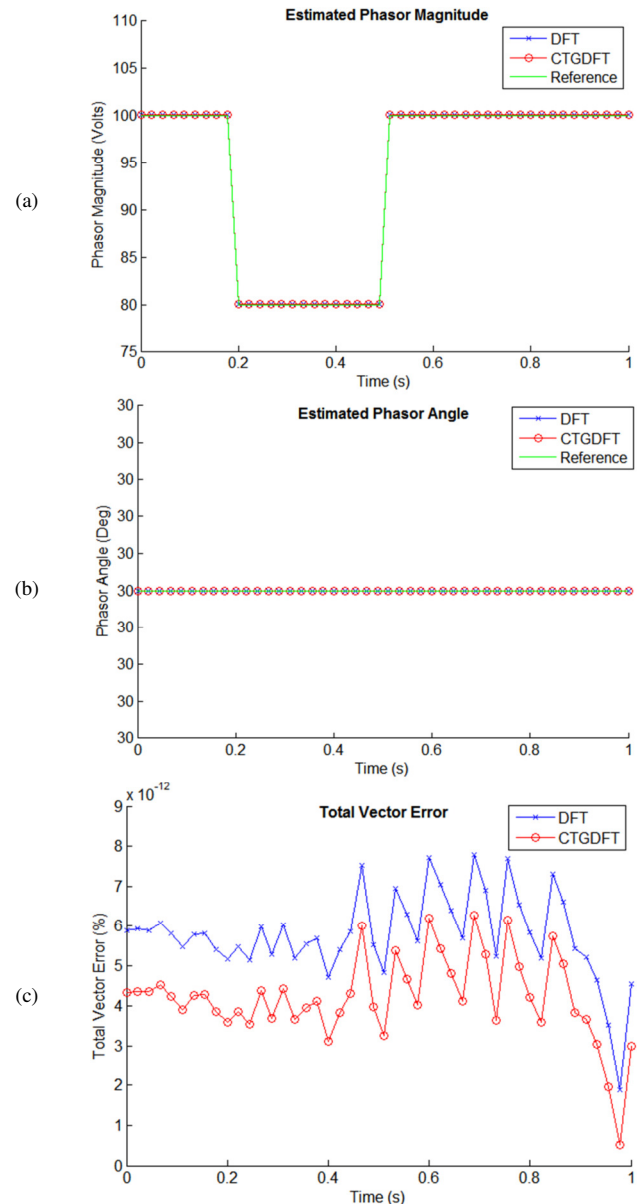


Fig. 6. Simulation results under Disturbance Faults 2: (a) estimated phasor magnitude, (b) estimated phasor angle, (c) total vector error.

TABLE II. ESTIMATED FREQUENCY, FREQUENCY ERROR, AND ROCOF ERROR UNDER DISTURBANCE FAULTS 2

Method	Frequency (Hz)		FE (Hz)	RFE (Hz/s)
	Min	Max		
DFT	45	45	0	0.00002
CTGDFT	44.9997	45	0.0003	0.00034

### 3) Disturbance Faults 3

Standard harmonic distortion tests require signals with 10% harmonics ranging from the second to the 50th harmonic. In this last test, only test signals containing 10% of the third and fifth harmonic components were generated, and a random noise was introduced into the voltage input signal with a slight variation in amplitude (-5%) and angle (+2°) at the  $t = 0.2s$ :

$$\begin{aligned} v(t) = & 95\sqrt{2} \cos(2\pi f + \pi/5.625) \\ & + 10\sqrt{2} \cos(6\pi f + \pi/3) \\ & + 10\sqrt{2} \cos(10\pi f + 5\pi/6) + e(t) \end{aligned} \quad (29)$$

where  $e(t)$  is the random noise. When the voltage input signal is superposed on a set of the third and fifth harmonic components and random noise with amplitude and phase variations, TVE was less than 0.15%, as shown in Figure 7.

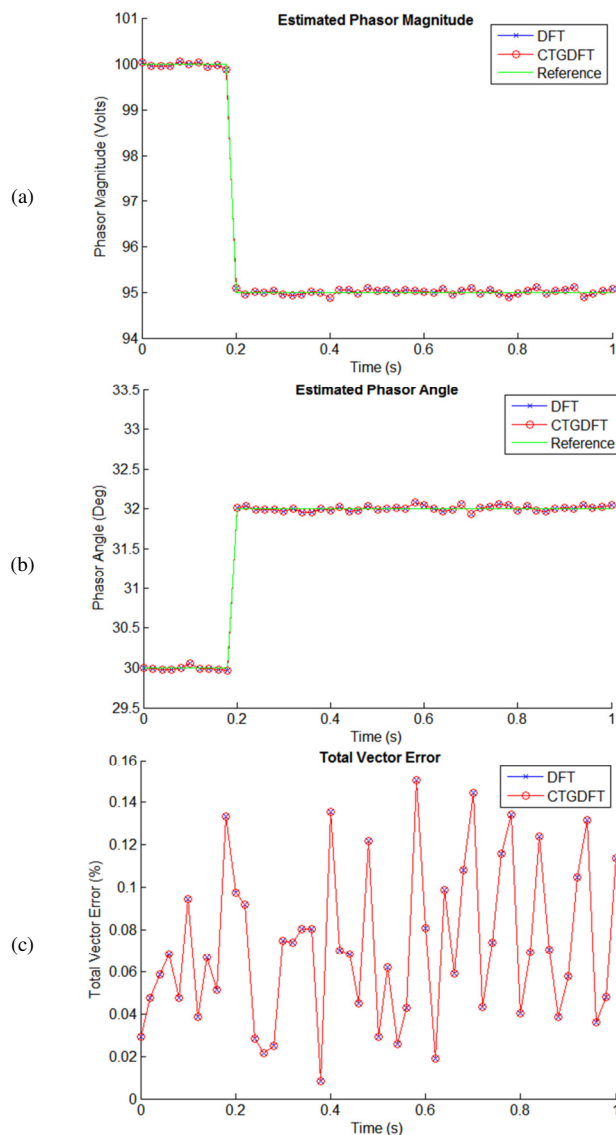


Fig. 7. Simulation results under Disturbance Faults 3: (a) estimated phasor magnitude, (b) estimated phasor angle, (c) total vector error.

Table III shows that the maximum FE of the proposed method was again in the order of 0.0003Hz.

TABLE III. ESTIMATED FREQUENCY, FREQUENCY ERROR, AND ROCOF ERROR UNDER DISTURBANCE FAULTS 3

Method	Frequency (Hz)		FE (Hz)	RFE (Hz/s)
	Min	Max		
DFT	49.9995	50	0.0005	0.00047
CTGDFT	49.9997	50	0.0003	0.00026

The IEEE C37.118.1-2011 standard specifies a TVE limit of 1%, an FE limit of 0.005Hz, and a limit of 0.01Hz/s for the ROCOF Error (RFE). The obtained TVE and FE limits were 0.15% and 0.0003Hz, respectively.

## VI. CONCLUSION

This study investigated the performance of the proposed Cooley-Tukey FFT algorithm based on GDFT for phasor estimation under conditions of power system disturbance. The performance of a PMU was examined and simulated by applying a voltage signal in different case studies. The performance and the evaluation of the results were analyzed according to the PMU standard (IEEE C37.118.1-2011) by calculating performance indices like TVE, Frequency Error (FE), and ROCOF Error (RFE). The performance of the proposed algorithm was tested under static and dynamic power system conditions, which is important for protection relay applications. The results obtained by the proposed method showed that the measurement evaluation parameters and PMU performance under power system disturbance conditions, such as integer harmonics component, frequency variation, input signal variation (amplitude and phase), and noise conditions, were in the limits according to the PMU standard (IEEE C37.118.1-2011). It should also be noted that the results of the PMU tests were very encouraging. Future work could investigate the implementation of the proposed algorithm in an embedded system as part of a PMU architecture.

## REFERENCES

- [1] R. B. Sharma and G. M. Dhole, "Wide Area Measurement Technology in Power Systems," *Procedia Technology*, vol. 25, pp. 718–725, Jan. 2016, <https://doi.org/10.1016/j.protcy.2016.08.165>.
- [2] S. Priyadarshini and C. K. Panigrahi, "Optimal Allocation of Synchrophasor Units in the Distribution Network Considering Maximum Redundancy," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6494–6499, Dec. 2020, <https://doi.org/10.48084/etasr.3862>.
- [3] P. Nanda, C. K. Panigrahi, and A. Dasgupta, "Phasor Estimation and Modelling Techniques of PMU- A Review," *Energy Procedia*, vol. 109, pp. 64–77, Mar. 2017, <https://doi.org/10.1016/j.egypro.2017.03.052>.
- [4] P. R. Bedse and N. N. Jangle, "Review on PMU using Recursive DFT Algorithm," in *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, Sep. 2018, pp. 375–377, <https://doi.org/10.1109/GUCON.2018.8675049>.
- [5] M. Kiruthika and S. Bindu, "Classification of Electrical Power System Conditions with Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5759–5768, Jun. 2020, <https://doi.org/10.48084/etasr.3512>.
- [6] A. G. Phadke, P. Wall, L. Ding, and V. Terzija, "Improving the performance of power system protection using wide area monitoring systems," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 3, pp. 319–331, Jul. 2016, <https://doi.org/10.1007/s40565-016-0211-x>.

- [7] T. B. Costa, R. O. Berriel, A. C. S. Lima, and R. F. S. Dias, "Evaluation of a Phase-Locked Loop Phasor Measurement Algorithm on a Harmonic Polluted Environment in Applications Such as PMU," *Journal of Control, Automation and Electrical Systems*, vol. 30, no. 3, pp. 424–433, Jun. 2019, <https://doi.org/10.1007/s40313-019-00450-5>.
- [8] R. Manam and S. R. Rayapudi, "Sensitive Constrained Optimal PMU Allocation with Complete Observability for State Estimation Solution," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2240–2250, Dec. 2017, <https://doi.org/10.48084/etasr.1542>.
- [9] M. U. Usman and M. O. Faruque, "Applications of synchrophasor technologies in power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 2, pp. 211–226, Mar. 2019, <https://doi.org/10.1007/s40565-018-0455-8>.
- [10] "SIPROTEC 5 Application - PMU functionality in SIPROTEC 5 devices, Edition 2," Siemens AG, APN-037, 2021.
- [11] A. G. Phadke and T. Bi, "Phasor measurement units, WAMS, and their applications in protection and control of power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 4, pp. 619–629, Jul. 2018, <https://doi.org/10.1007/s40565-018-0423-3>.
- [12] R. P. M. Silva, A. C. B. Delbem, and D. V. Coury, "Genetic algorithms applied to phasor estimation and frequency tracking in PMU development," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 921–929, Jan. 2013, <https://doi.org/10.1016/j.ijepes.2012.07.070>.
- [13] S. Das, "Algorithms to Improve Performance of Wide Area Measurement Systems of Electric Power Systems," Ph.D. dissertation, University of Western Ontario, London, Canada, 2014.
- [14] "IEEE Standard for Synchrophasor Data Transfer for Power Systems," *IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005)*, Sep. 2011, <https://doi.org/10.1109/IEEESTD.2011.6111222>.
- [15] H. Liu, H. Hu, H. Chen, L. Zhang, and Y. Xing, "Fast and Flexible Selective Harmonic Extraction Methods Based on the Generalized Discrete Fourier Transform," *IEEE Transactions on Power Electronics*, vol. 33, no. 4, pp. 3484–3496, Apr. 2018, <https://doi.org/10.1109/TPEL.2017.2703138>.
- [16] D. M. Laverty, J. Hastings, D. J. Morrow, R. Khan, K. McLaughlin, and S. Sezer, "A modular phasor measurement unit design featuring open data exchange methods," in *2017 IEEE Power & Energy Society General Meeting*, Chicago, IL, USA, Jul. 2017, pp. 1–5, <https://doi.org/10.1109/PESGM.2017.8273986>.
- [17] M. Abbaci and M. O. Zmirli, "Performance Evaluation of PMU using Modified Cooley-Tukey Algorithm Based on GDFT," in *2018 International Conference on Applied Smart Systems (ICASS)*, Medea, Algeria, Aug. 2018, pp. 1–6, <https://doi.org/10.1109/ICASS.2018.8651967>.
- [18] M. Shahraini, M. S. Ghazizadeh, and M. H. Javidi, "Co-Optimal Placement of Measurement Devices and Their Related Communication Infrastructure in Wide Area Measurement Systems," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 684–691, Jun. 2012, <https://doi.org/10.1109/TSG.2011.2178080>.
- [19] X. Zhao, D. M. Laverty, A. McKernan, D. J. Morrow, K. McLaughlin, and S. Sezer, "GPS-Disciplined Analog-to-Digital Converter for Phasor Measurement Applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 9, pp. 2349–2357, Sep. 2017, <https://doi.org/10.1109/TIM.2017.2700158>.
- [20] A. L. Kouzou, H. Bentarzi, R. D. Mohammadi, and M. Laoumer, "Optimal Placement of Phasor Measurement Unit in Power System using Meta-Heuristic Algorithms," *Electrotehnica, Electronica, Automatica*, vol. 67, no. 2, pp. 98–113, 2019.
- [21] G. S. Antonova, "Phasor Data Concentrator Definitions, Functions and Standards," presented at the Hands on Relay School, Mar. 2018.
- [22] R. Khan, K. McLaughlin, D. Laverty, and S. Sezer, "IEEE C37.118-2 Synchrophasor Communication Framework: Overview, Cyber Vulnerabilities Analysis and Performance Evaluation: 2nd International Conference on Information Systems Security and Privacy," *Proceedings of the 2nd International Conference on Information Systems Security and Privacy*, pp. 159–170, Feb. 2016, <https://doi.org/10.5220/0005745001670178>.
- [23] S. G. Johnson, "Modified Cooley-Tukey algorithms based on a generalized DFT framework," 2008.
- [24] H. Kim and S. Lekcharoen, "A Cooley-Tukey Modified Algorithm in Fast Fourier Transform," *Korean Journal of Mathematics*, vol. 19, no. 3, pp. 243–253, 2011, <https://doi.org/10.11568/kjm.2011.19.3.243>.
- [25] F. Qureshi, "Optimization of Rotations in FFTs," Ph.D. dissertation, Linköping University, 2012.
- [26] H. Li, "Frequency estimation and tracking by two-layered iterative DFT with re-sampling in non-steady states of power system," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, Art. no. 28, Feb. 2019, <https://doi.org/10.1186/s13638-018-1320-1>.
- [27] "Report of Task Force on Testing and Certification," North American Synchrophasor Initiative Task Force (NASPI TF), Oct. 2013.

# An Efficient Methodology for Detecting the Vertical Movement of Structures

**Khalid L. A. El-Ashmawy**

Department of Civil Engineering, College of Engineering and Islamic Architecture  
Umm Al-Qura University, Makkah, Saudi Arabia  
klashmawy@uqu.edu.sa  
(corresponding author)

Received: 2 November 2022 | Revised: 16 November 2022 | Accepted: 21 November 2022

## ABSTRACT

Details regarding the public safety of engineering structures can be gleaned from measurements and monitoring. The development of a methodology for monitoring and analyzing structures' vertical displacement is explained in this paper. The developed methodology aims to add a new dimension to geometric leveling, and leveling routing, by applying a least squares solution for level network adjustment and performing statistical analysis to assess the change in vertical displacement. To monitor and analyze the vertical deformation of a building in Cairo, Egypt, the proposed methodology was utilized. Twenty monitoring points, five auxiliary points, and three local reference stations were utilized. All the measurements were taken with a geodetic invar staff and an automatic level with an attachment of a parallel plate micrometer. The observations were made for an interval of 81 months. The least squares adjustment technique was applied to obtain the adjusted levels and observations and to generate the required statistical data. The results of the subsequent epochs were compared to the results of the first epoch to determine the vertical movement of the monitoring points for each epoch. In addition, the significance of the present displacement was ascertained by comparing the values of vertical displacement to the determined 95% corresponding confidence intervals. The findings demonstrated that the building remained stable throughout the monitoring period. The case study demonstrates how effectively geometric leveling with least square adjustment can be used to monitor the vertical displacement of structures.

**Keywords-**stability and safety of engineering structures; geometric leveling; vertical displacement monitoring; least squares adjustment; accuracy analysis

## I. INTRODUCTION

The aim of measuring a structure's movement is to decide whether the structure is safe and stable. Deformation can be subjected to more analysis to determine whether it is caused by seasonal factors, daily variations, or other factors and then utilize the data to estimate how the structure will move in the future [1-5]. This movement must be identified for the purposes of safety studies and the prevention of future disasters. Deformation monitoring has the advantages of improving the design process of structures for use in the future and increasing safety by lowering the possibility of structural damage [6]. It is a crucial risk management tool. Presently, there are many techniques for deformation monitoring. Modern techniques often use geometric leveling and total stations for three-dimensional (3D) control [7]. The above-mentioned techniques can be used in conjunction with a Terrestrial Laser Scanner (TLS) system [8-12] or Portable Digital Photogrammetric Stations (DPSs) [13-15]. In particular, TLS has evolved into a crucial tool for working with historical structures, supplying a point cloud that quickly generates a three-dimension model of a monument [8]. However, the majority of surveyors do not have access to the specialized equipment required by these remote

sensing technologies. It is costly to purchase or rent this equipment, and it could cost even more to train or hire employees who can manage and model point cloud data [9, 10]. A TLS system's data acquisition capacity also makes model creation challenging [10, 11].

Numerous software systems can help with this, but they consume time and computing resources and require an additional instrument that must be purchased or hired. As a result, these methods are severely constrained and can only be used for architectural heritage inspection [12, 16], where resources and budgets are significantly available. In addition, these methods have other issues, including the accuracy of the measurements they produce [17] despite the cost and complexity issues. Potential dimensional errors in the model can be caused by issues with close-range imaging [18], as well as variations in temperature and humidity [19], the reflectivity of various materials and colors, and the effects of incidence angle [20].

Comparing to direct measurements, significant errors in measuring vertical deformations can happen based on the above-mentioned conditions, sometimes exceeding one centimeter in magnitude [21, 22]. This emphasizes the



significance of evaluation techniques that allow calibrating laser scanners, particularly as high-precision surveys development in architectural heritage sites. Typical tests like deformation monitoring should be checked using direct measurement methods [12].

To overcome some of the issues listed above, this paper aims to:

- develop a cost-effective methodology based on a mathematical model for determining the accurate levels from measurements at individual monitoring points,
- adjust redundant measurements and get their precision using the least squares solution,
- develop software based on the derived mathematical model and the least squares solution for determining the adjusted levels and required statistical data of deformation points, and
- use the developed methodology for detecting the vertical displacement of a building as a case study at a particular time interval.

## II. THE PROPOSED METHODOLOGY

### A. Development of the Mathematical Model

Leveling is the method of measuring vertical distances, either directly or indirectly, to determine elevations. There are two approaches: geometric and trigonometric leveling [23]. Geometric leveling is more accurate, the required surveying instruments are more cost-effective, and the work in the field and office is easier. In geometric leveling, the difference in levels between two points can be obtained by taking the readings of the staffs, or rulers, placed on the points. A leveling instrument is used to measure the readings. The elevations of all points can be determined by knowing at least one point's elevation and the level differences between points. To apply the method of least squares in leveling adjustments, an observation equation is first written for any elevation difference [23]. Figure 1 illustrates the geometrical relationship for the elevation difference observed between two stations, A and B. The equation is written as:

$$\delta_{hAB} + v_{\delta hAB} = h_B - h_A \quad (1)$$

which can be rewritten as:

$$v_{\delta hAB} - h_B + h_A = -\delta_{hAB} \quad (2)$$

The observation equation (2) gives the relation between the unknown elevations of any two stations, A and B, the differential levelling observation  $\delta_{hAB}$  and its residual. This equation is essential in applying least squares adjustments of level nets. Equation (2) is linear and can be written directly in the form [24]:

$$V + B.A = \varepsilon \quad (3)$$

where  $A$  is the vector to the current values set for the unknowns (elevations of the deformation points) in the iterative solution,  $B$  is the matrix of the partial derivatives of (2) in relation to the unknowns,  $V$  is the vector of residuals, i.e. the correction vector

to the observations (height difference), and  $\varepsilon$  is the vector of discrepancies.

The solution of (3) using least squares method is [23,24]:

$$A = N^{-1}.C \quad (4)$$

where:

$$\left. \begin{aligned} N &= B'.W.B \\ C &= B'.W.\varepsilon \end{aligned} \right\} \quad (5)$$

The variance of unit weight is obtained by:

$$\hat{\sigma}_o^2 = V'.W.V / (N - U) \quad (6)$$

where  $\hat{\sigma}_o^2$  is the variance of unit weight,  $N$  is the number of observations,  $W$  is the weight matrix observations,  $U$  is the number of unknowns and equals to the number of deformation points, and  $(N-U)$  present the degrees of freedom.

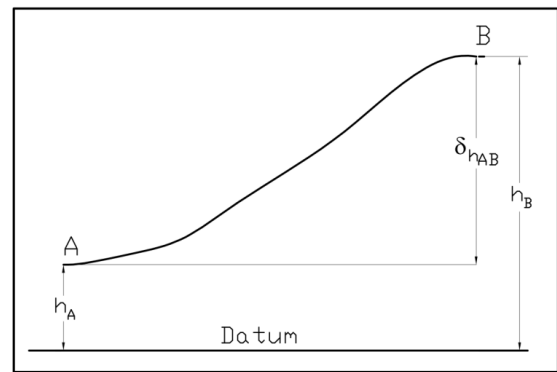


Fig. 1. Geometric depiction of the proposed model.

### B. Developing the Necessary Software

The software for determining the adjusted level and desired statistical data for each deformation point in every loop is being developed as part of the current research. The least squares method, as explained above, is used to determine the output of the developed software, which makes use of the data in the form of level differences and distances between network points in addition to the known levels of the benchmark(s). The variance of unit weight, adjusted levels of stations and their standard deviations (optional), residuals of observations, corrected observations, and their cofactor matrix (optional), and statistical data for gross error detection (optional) are all included in the software's ASCII file output. The software provides flexibility in level network adjustment as follows:

- Any number and distribution of network points is acceptable.
- The editing of input data specifications has remained flexible for practical reasons. If the input formats are consistent, they can be accepted.
- The point-numbering is another important practical feature. Point-numbering is not restricted in any way by the software. The point-numbering is easy and almost natural.

- Weighting observations is possible. From a theoretical standpoint, it is desirable to weight the observations because it greatly aids in the detection of gross errors.
- Determination of the necessary initial levels for unknown points in order to initiate the iterative process.
- An iterative least-squares solution that can show the results of each iteration or the final iteration.

The software makes use of effective methods like Data Structuring [25, 26], Random File Access, and Dynamic Memory Allocations [27] for the purpose of automatically processing and representing the data and the results. The software was created with window-driven user interfaces in mind to make it easier to use [28].

### C. Classification of Leveling Points

Three categories of points are proposed in the leveling line: i) monitoring points, ii) reference or benchmark points, and iii) auxiliary points. The monitoring points are selected to be placed on the building's exterior walls and required interior walls. To ensure that each building's wall is stable, the monitoring points are strategically placed. In order to determine only relative displacements, the reference points can be situated in the controlled area and undergo displacement. Absolute displacements can be obtained if the reference points are attached to solid foundation or other stationary structures outside that area. Even though leveling lines only require one reference point, experience suggests placing at least 3 of them in order to recognize unstable reference points. In order to ensure that they are unaffected by the building's movements, the reference points need to be positioned at a safe distance from it. Auxiliary points are positioned, for example, to connect independent sections of a levelling line.

### D. Fixing the Points

Regarding fixing the points, they are typically secured by utilizing drill and epoxy to seal concrete nails on a wall or floor. All monitoring points must be well connected to the structure, particularly the concrete columns otherwise the displacements might be more accurately described as displacements of the points rather than structural displacements.

### E. Leveling Instruments

It is necessary to use automatic levels, also known as optical levels, that have a built-in compensator that automatically makes the line of sight horizontal. Accuracy can be improved by mounting a parallel plate micrometer over the telescope objective [29]. The parallel plate micrometer permits estimated readings up to 0.01mm and direct readings up to 0.1mm on a staff of 1cm least count. Digital levels may also be used. Digital levels are automatic levels that have a digital image processing system built in. These levels allow for electronic recording and automatic reading of coded bars staff. The speed of leveling and the elimination of the errors brought on by human reading and recording are the main benefits of digital levels [30].

### F. Field Work

It is necessary to know the reference station levels before monitoring points can be observed. Otherwise, precise leveling can be used to obtain the local levels of reference stations [29]. Loops are used for the leveling of the monitoring points. Each loop begins at a reference station and ends at either the same or another reference station. At various epochs of a predetermined time interval, the levels of the monitoring points are obtained. After the reference and monitoring points have been selected and the leveling line route has been established, the necessary auxiliary points and an approximate location of the instrument in each leveling line must be marked in order to guarantee the precision required by having equal sections of leveling and by following almost the same route of line levelling. Starting and stopping leveling measurements at reference stations is the only way to control the quality of the measurements.

The misclosure ( $\Delta$ ) is the difference between the determined level and the known level, and it is calculated by :

$$\text{Misclosure } (\Delta) = \text{Reference}_{\text{Levelknown}} - \text{Reference}_{\text{Levelcomputed}} \quad (7)$$

Following the completion of the levelling work, the misclosure is computed. The fieldwork team is able to repeat the measurements at a lower cost and, most importantly, without altering the conditions of the structure by carrying out this data pre-processing in the field. The permissible misclosure can be calculated as [6]:

$$\text{Misclosure tolerance} = 0.9 \sqrt{n} \text{ (mm)} \quad (8)$$

where  $n$  is the number of set-ups.

All measurements ought to be repeated when the misclosure is more than the tolerance value. In case of permissible misclosure, the points' adjusted levels can be determined applying the least squares technique as explained above.

### G. Vertical Displacement Determination

The developed software is utilized for computing the adjusted levels of the monitoring points for each epoch. In order to compare the results of the subsequent epochs with those of the first epoch, the results of the first epoch are used as a reference. The vertical displacement of each monitoring point can be determined as follows:

$$dZ = Z_{\text{first epoch}} - Z_{\text{current epoch}} \quad (9)$$

where  $dZ$  is the vertical displacement of the monitoring point and  $Z$  the adjusted level of the monitoring point.

### H. Displacement Analysis

Comparing the level differences between the observations of the first epoch and the current epoch is the traditional method for determining the stability of reference and monitoring points, as shown in (9). In recent deformation modeling techniques, statistics are applied to each epoch of measurements to get the significance of point movements. Comparing the obtained displacements to their respective 95% confidence intervals is used to obtain the significance of point displacements [31]. In case the value of the vertical movement of a point  $K$  is classified as  $dZ_K$  and the maximum dimension of

the combined 95% confidence ellipse for point  $K$  is specified as  $E_K$ , two cases will arise [32]:

- In the first case if  $|dZ_K| < E_K$ , then no movement has happened in point  $K$  and the observed difference is due to measurement errors.
- If  $|dZ_K| > E_K$  then point displacement has occurred.

$E_K$  can be obtained as:

$$E_K = 1.96\sqrt{(m_{AK}^{i+1})^2 + (m_{AK}^i)^2} \quad (10)$$

where  $(m_{AK}^{i+1})$  is the standard error in level of point  $K$  in the current epoch and  $(m_{AK}^i)$  is the standard error in level of point  $K$  in the previous epoch.

### III. CASE STUDY

A building in Cairo, Egypt, was selected as a case study for monitoring the vertical displacement. The building is 110.0m long, 60.0m wide, and 20m high. There are various offices in this 3-story structure.

#### A. Choosing Points

The field has been identified in the initial step in order to carry out the vertical displacement monitoring of the studied building. As shown in Figure 2, 3 local reference stations and 5 auxiliary points were chosen. In order to ensure that the local reference stations were unaffected by the building's displacements, these points were fixed around it at a secure distance.

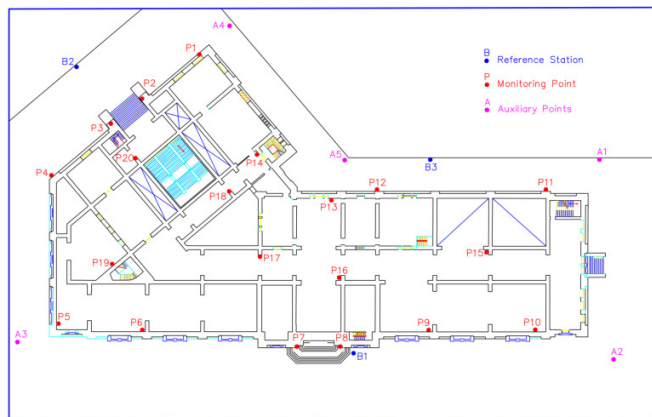


Fig. 2. Location of the case study points.

Twenty monitoring points were fixed on the first floor of the case study building as part of the vertical movement monitoring project to determine the structure's stability. Figure 2 depicts the positions of those monitoring points on the internal and external walls. The staff could easily hold on to the monitoring points on the building walls during levelling and observation because they were about 0.3m above the ground. Special nails were driven into the marked monitoring points to make them permanent.

#### B. Measurements

Measurements were acquired with a Leica NA2 automatic level, a GPLE3 geodetic invar staff with 1cm graduations, and an attachment for a Leica (10mm) GPM3 parallel plate micrometer. Before the monitoring points can be observed, it is necessary to know the levels of the reference stations B1, B2, and B3 as explained above. The level of point B1 is assumed to be 15.0m. The levels of points B2 and B3 were obtained by looping a level net from B1 to 5 auxiliary points, B2 and B3, and then closing back on B1. After determining the reference station levels, loops utilizing precise leveling were utilized to get the levels of monitoring points at 8 distinct 1-month interval epochs. Observations were made in the early morning. Each leveling line's misclosure did not exceed 0.4mm across all 8 observation epochs, which is an acceptable tolerance for the purposes of this work. The relatively flat topography of the study area, the experience of the observer, and the instruments used all contributed to the leveling's high accuracy.

#### C. Data Processing and Results

The developed software was used to process the data from the level networks. The necessary data and measurements such as the levels of the reference stations and their standard deviations, height differences, and distances between stations, were inserted into the software by editing the data file. The final results, for each epoch, were obtained from the adjustment of the level network. As previously mentioned, the variance of the unit weight, the adjusted levels of the monitoring points and their standard deviations, the residuals of the observations, and the corrected observations with their cofactor matrix comprised the received results. The high precision of the adjusted observations can be seen by the fact that the first to eighth leveling epochs had maximum variances of unit weight of  $1.02\text{m}^2$  and standard errors of 1.01mm. The high accuracy of the adjusted levels is demonstrated by the fact that their maximum standard error was 1.06mm across all epochs. All 8 epochs' computed standardized residuals for observations or levels were less than their rejection constants of the respective standardized residual which indicates that none of the observations were rejected due to gross errors or outliers.

The adjusted values of levels of monitoring points for each observation epoch and the differences in levels between the first and the subsequent epochs were respectively computed with (9) and are tabulated in Table I. Equation (10) was used to calculate the 95% confidence level differences between the first and respective confidence intervals of the subsequent epochs. The vertical displacement values of monitoring points and the confidence intervals that correspond to them at a 95% level of confidence are shown in Table II.

At a 95% confidence level, the magnitudes of the vertical movement of the monitoring points at each epoch were obtained and compared to their identical confidence intervals to decide whether the obtained displacements were actual (significant) displacements of the structure or caused by measurement errors. According to Tables I and II, all the evaluated movement values were less than their related confidence intervals, indicating that there was no vertical movement of the building within the monitoring period.

TABLE I. THE ADJUSTED LEVELS AND VERTICAL DISPLACEMENT OF THE MONITORING POINTS

Point	Epoch (1)		Epoch (2)		Epoch (3)		Epoch (4)		Epoch (5)		Epoch (6)		Epoch (7)		Epoch (8)	
	Level m	dZ mm	Level m	dZ mm	Level m	dZ mm	Level m	dZ mm	Level m	dZ mm	Level m	dZ mm	Level m	dZ mm	Level m	dZ mm
P1	15.3092	0	15.3093	-0.10	15.3095	-0.30	15.3094	-0.20	15.3093	-0.10	15.3097	-0.50	15.3097	-0.50	15.3095	-0.30
P2	15.4085	0	15.4086	-0.10	15.4082	0.30	15.4082	0.30	15.4086	-0.10	15.4083	0.20	15.4084	0.10	15.4086	-0.10
P3	15.3120	0	15.3116	0.40	15.3113	0.70	15.3114	0.60	15.3114	0.60	15.3112	0.80	15.3113	0.70	15.3114	0.60
P4	15.2960	0	15.2956	0.40	15.2957	0.30	15.2953	0.70	15.2954	0.60	15.2956	0.40	15.2954	0.60	15.2955	0.50
P5	15.8168	0	15.8166	0.20	15.8163	0.50	15.8162	0.60	15.8165	0.30	15.8163	0.50	15.8166	0.20	15.8167	0.10
P6	15.8097	0	15.8093	0.40	15.8092	0.50	15.8095	0.20	15.8096	0.10	15.8095	0.20	15.8098	-0.10	15.8096	0.10
P7	15.0963	0	15.0966	-0.30	15.0966	-0.30	15.0965	-0.20	15.0962	0.10	15.0963	0.00	15.0966	-0.30	15.0965	-0.20
P8	15.1254	0	15.1255	-0.10	15.1256	-0.20	15.1252	0.20	15.1254	0.00	15.1251	0.30	15.1253	0.10	15.1256	-0.20
P9	15.8458	0	15.8453	0.50	15.8451	0.70	15.8455	0.30	15.8457	0.10	15.8450	0.80	15.8455	0.30	15.8458	0.00
P10	15.6507	0	15.6507	0.00	15.6501	0.60	15.6505	0.20	15.6503	0.40	15.6500	0.70	15.6505	0.20	15.6502	0.50
P11	14.2996	0	14.2998	-0.20	14.2997	-0.10	14.2992	0.40	14.2995	0.10	14.2990	0.60	14.2993	0.30	14.2996	0.00
P12	14.2160	0	14.2154	0.60	14.2158	0.20	14.2157	0.30	14.2155	0.50	14.2158	0.20	14.2154	0.60	14.2158	0.20
P13	15.5106	0	15.5108	-0.20	15.5105	0.10	15.5107	-0.10	15.5103	0.30	15.5101	0.50	15.5105	0.10	15.51	0.60
P14	15.6034	0	15.6034	0.00	15.6032	0.20	15.6031	0.30	15.6036	-0.20	15.6033	0.10	15.6037	-0.30	15.6035	-0.10
P15	15.8085	0	15.8083	0.20	15.8087	-0.20	15.8086	-0.10	15.8083	0.20	15.8083	0.20	15.8085	0.00	15.8086	-0.10
P16	15.6054	0	15.6057	-0.30	15.6055	-0.10	15.6053	0.10	15.6051	0.30	15.6052	0.20	15.6055	-0.10	15.6056	-0.20
P17	15.4709	0	15.4705	0.40	15.4708	0.10	15.4703	0.60	15.4701	0.80	15.4703	0.60	15.4704	0.50	15.4705	0.40
P18	15.6615	0	15.6615	0.00	15.6616	-0.10	15.6613	0.20	15.6614	0.10	15.6612	0.30	15.6611	0.40	15.6612	0.30
P19	15.7997	0	15.7990	0.70	15.7992	0.50	15.7991	0.60	15.7995	0.20	15.7997	0.00	15.7990	0.70	15.7992	0.50
P20	15.6783	0	15.6782	0.10	15.6787	-0.40	15.6787	-0.40	15.6784	-0.10	15.6783	0.00	15.6781	0.20	15.6783	0.00

TABLE II. DETECTING THE VERTICAL DISPLACEMENT OF THE MONITORING POINTS

Point	Epoch (2)		Epoch (3)		Epoch (4)		Epoch (5)		Epoch (6)		Epoch (7)		Epoch (8)		Point displacement
	dZ  mm	E mm	dZ  mm	E mm	dZ  mm	E mm	dZ  mm	E mm	dZ  mm	E mm	dZ  mm	E mm	dZ  mm	E mm	
P1	0.10	1.67	0.30	1.97	0.20	2.01	0.10	2.41	0.50	1.39	0.50	1.89	0.30	1.81	None
P2	0.10	1.97	0.30	2.27	0.30	1.91	0.10	1.91	0.20	2.09	0.10	1.79	0.10	1.41	None
P3	0.40	2.07	0.70	2.17	0.60	1.81	0.60	2.31	0.80	1.99	0.70	1.79	0.60	1.51	None
P4	0.40	1.77	0.30	1.97	0.70	2.31	0.60	2.21	0.40	1.59	0.60	2.09	0.50	1.51	None
P5	0.20	1.37	0.50	2.17	0.60	2.01	0.30	2.01	0.50	1.99	0.20	1.59	0.10	1.51	None
P6	0.40	1.17	0.50	1.97	0.20	1.61	0.10	2.21	0.20	1.89	0.10	1.59	0.10	1.81	None
P7	0.30	1.57	0.30	1.87	0.20	2.01	0.10	2.61	0.00	1.69	0.30	1.59	0.20	1.71	None
P8	0.10	2.17	0.20	1.77	0.20	2.31	0.00	2.11	0.30	2.09	0.10	1.69	0.20	1.31	None
P9	0.50	1.97	0.70	2.07	0.30	1.51	0.10	2.11	0.80	2.49	0.30	2.09	0.00	1.31	None
P10	0.00	1.77	0.60	2.47	0.20	1.51	0.40	2.51	0.70	2.09	0.20	2.19	0.50	1.91	None
P11	0.20	2.07	0.10	1.97	0.40	2.41	0.10	2.01	0.60	2.29	0.30	1.59	0.00	1.31	None
P12	0.60	1.67	0.20	1.47	0.30	2.01	0.50	2.51	0.20	1.49	0.60	2.29	0.20	1.81	None
P13	0.20	1.97	0.10	2.17	0.10	1.71	0.30	2.71	0.50	1.99	0.10	1.49	0.60	2.11	None
P14	0.00	2.17	0.20	2.07	0.30	2.01	0.20	1.81	0.10	2.09	0.30	1.49	0.10	1.81	None
P15	0.20	1.67	0.20	1.47	0.10	2.01	0.20	2.61	0.20	1.79	0.00	1.69	0.10	1.51	None
P16	0.30	1.77	0.10	2.07	0.10	2.11	0.30	2.51	0.20	1.69	0.10	1.59	0.20	1.51	None
P17	0.40	1.87	0.10	1.57	0.60	2.41	0.80	2.51	0.60	1.59	0.50	2.29	0.40	2.11	None
P18	0.00	1.87	0.10	1.77	0.20	2.21	0.10	2.21	0.30	1.99	0.40	1.99	0.30	1.51	None
P19	0.70	1.67	0.50	1.67	0.60	2.01	0.20	1.91	0.00	1.59	0.70	2.59	0.50	1.41	None
P20	0.10	1.57	0.40	1.37	0.40	1.91	0.10	2.61	0.00	1.89	0.20	2.09	0.00	1.41	None

#### IV. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The aim of measuring a structure's movement is to decide whether there is actually a movement and whether the structure is safe and stable. Deformation monitoring has the advantages of improving the design process of structures for use in the future and increasing safety by lowering the possibility of structural damage. It is a crucial risk management tool.

This paper explains the development of a methodology for monitoring and analyzing structures' vertical displacement. The developed methodology aims to add a new dimension to the well-known ones for geometric leveling and leveling routing. The proposed methodology is accurate and effective in

determining the vertical displacement of structures. The case study demonstrates that geometric leveling is always effective for determining the vertical displacement of structures. For vertical displacement monitoring with a precision of less than 0.4mm, this approach is suggested. Geometric leveling is an accurate, precise, and cost-effective method of vertical displacement monitoring but it is not considered to be cutting-edge. Geometric leveling is recommended to be used extensively to monitor structures' deformation for the above-mentioned reasons.

The proposed method is subjected to limitations in its application areas. The proposed method can only be utilized to monitor displacements in the vertical plane, it cannot be

utilized to monitor displacements in the horizontal plane. The developed software and methodology are very adaptable, affordable, and of great interest to students and researchers. They provide a helpful tool for engineering firms to monitor the safety of structures. It is strongly recommended that engineering structures, especially high-rise buildings, should be monitored at a regular basis to check their stability and thereby increasing their safety.

## REFERENCES

- [1] H. O. Aghedo, *Deformation Monitoring of Ikpoba River Bridge in Benin City, Edo State, Using GPS*. Unpublished MSc Thesis of the Department of Surveying and Geoinformatics, Nnamdi Azikiwe University, Awka, 2016.
- [2] M. Javanpour and P. Zarfam, "Application of Incremental Dynamic Analysis (IDA) Method for Studying the Dynamic Behavior of Structures During Earthquakes," *Engineering, Technology & Applied Science Research*, vol. 7, no. 1, pp. 1338–1344, Feb. 2017, <https://doi.org/10.48084/etasr.902>.
- [3] C. Zhang, Y. Gao, and E. Duan, "The Influence of Environment on the Settlement of Historic Buildings in China," *KSCE Journal of Civil Engineering*, vol. 25, no. 6, pp. 1951–1963, Jun. 2021, <https://doi.org/10.1007/s12205-021-0690-9>.
- [4] M. Mackiewicz, J. R. Krentowski, P. Knyziak, and M. Wardach, "Consequences of excessive deformation of structural elements in precast buildings," *Engineering Failure Analysis*, vol. 137, Jul. 2022, Art. no. 106261, <https://doi.org/10.1016/j.engfailanal.2022.106261>.
- [5] J. Gao, L. Liu, Y. Zhang, and X. Xie, "Deformation mechanism and soil evolution analysis based on different types geogrid reinforced foundation," *Construction and Building Materials*, vol. 331, May 2022, Art. no. 127322, <https://doi.org/10.1016/j.conbuildmat.2022.127322>.
- [6] A. Ebeling, "Ground-Based Deformation Monitoring," Ph.D. dissertation, University of Calgary, Calgary, Canada, 2014, <https://doi.org/10.11575/PRISM/26325>.
- [7] K. Zschiesche, "Image Assisted Total Stations for Structural Health Monitoring—A Review," *Geomatics*, vol. 2, no. 1, pp. 1–16, Mar. 2022, <https://doi.org/10.3390/geomatics2010001>.
- [8] M. El-Tokhey, A. K. Abdel-Gawad, Y. Mogahed, and A. M. El-Maghraby, "Accuracy assessment of laser scanner in measuring and monitoring deformations of structures," vol. 26, no. 2, pp. 144–151, Jan. 2013, <https://doi.org/10.5829/idosi.wasj.2013.26.02.13461>.
- [9] S. Kalenjuk, W. Lienhart, and M. J. Rebhan, "Processing of mobile laser scanning data for large-scale deformation monitoring of anchored retaining structures along highways," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 6, pp. 678–694, 2021, <https://doi.org/10.1111/mice.12656>.
- [10] E. Kaartinen, K. Dunphy, and A. Sadhu, "LiDAR-Based Structural Health Monitoring: Applications in Civil Infrastructure Systems," *Sensors*, vol. 22, no. 12, Jan. 2022, Art. no. 4610, <https://doi.org/10.3390/s22124610>.
- [11] C. Cosarca, A. Joca, and A. Savu, "Analysis of error sources in Terrestrial Laser Scanning. RevCAD," *Journal of Geodesy and Cadastre*, vol. 9, pp. 115–125, 2009.
- [12] L. Fregonese, G. Barbieri, L. Biolzi, M. Bocciarelli, A. Frigeri, and L. Taffurelli, "Surveying and Monitoring for Vulnerability Assessment of an Ancient Building," *Sensors*, vol. 13, no. 8, pp. 9747–9773, Aug. 2013, <https://doi.org/10.3390/s130809747>.
- [13] X. Yu, Y. Ge, T. Zhao, G. Zhang, Y. Liu, and C. Yu, "The Dynamic Displacement Monitoring Method Using Unmanned Aerial Vehicles Based on Digital Close-Range Photogrammetry," *Mathematical Problems in Engineering*, vol. 2022, Sep. 2022, Art. no. e1949213, <https://doi.org/10.1155/2022/1949213>.
- [14] J. Hu, E. Liu, and J. Yu, "Application of Structural Deformation Monitoring Based on Close-Range Photogrammetry Technology," *Advances in Civil Engineering*, vol. 2021, Feb. 2021, Art. no. e6621440, <https://doi.org/10.1155/2021/6621440>.
- [15] M. E. Tjahjadi, L. A. Parsamardhani, and K. T. Suhari, "Bridge Structural Deformation Monitoring Using Digital Camera," *IOP Conference Series: Earth and Environmental Science*, vol. 1051, no. 1, Apr. 2022, Art. no. 012009, <https://doi.org/10.1088/1755-1315/1051/1/012009>.
- [16] J. A. Clarke and D. F. Laefer, "Systematic Approach for Large-Scale, Rapid, Dilapidation Surveys of Historic Masonry Buildings," *International Journal of Architectural Heritage*, vol. 8, no. 2, pp. 290–310, Mar. 2014, <https://doi.org/10.1080/15583058.2012.692849>.
- [17] R. M. Alkan and G. Karsidag, "Analysis of the Accuracy of Terrestrial Laser Scanning Measurements," in *FIG Working Week*, Rome, Italy, Aug. 2012.
- [18] A. Ekinci, T. Muturi, and P. M. V. Ferreira, "Aerial Close-Range Photogrammetry to Quantify Deformations of the Pile Retaining Walls," *Journal of the Indian Society of Remote Sensing*, vol. 49, no. 5, pp. 1051–1066, May 2021, <https://doi.org/10.1007/s12524-020-01275-5>.
- [19] G. Durán-Domínguez, A. Felicísimo, and M.-E. Polo, "3D study of cultural heritage for conservation. Reliability of the portable 3D laser scanners," presented at the International Congress on Science and Technology for the Conservation of Cultural Heritage, Seville, Spain, Jun. 2014.
- [20] A. Berényi, T. Lovas, and A. Barsi, "Terrestrial laser scanning in engineering survey: Analysis and application examples," presented at the ASPRS 2010 Annual Conference, San Diego, CA, USA, Jan. 2010.
- [21] A. Dumalski, "Evaluation of Possible Application of Terrestrial Laser Scanner - ScanStation in Vertical Displacement Measurements," *Technical Sciences / University of Warmia and Mazury in Olsztyn*, vol. 14, no. 1, pp. 33–43, 2011.
- [22] M. K. Villareal and A. F. Tongco, "Remote Sensing Techniques for Classification and Mapping of Sugarcane Growth," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6041–6046, Aug. 2020, <https://doi.org/10.48084/etasr.3694>.
- [23] C. D. Ghilani, *Adjustment Computations: Spatial Data Analysis, 6th Edition* | Wiley, 6th ed. Hoboken, NJ, USA: Wiley and Sons, 2017.
- [24] E. M. Mikhail, *Observations and Least Squares*. 1982.
- [25] D. S. Malik, *C++ Programming: Program Design Including Data Structures*, 8th ed. Australia: Cengage Learning, 2017.
- [26] B. Nethravathi, G. Amitha, A. Saruka, T. P. Bharath, and S. Suyagya, "Structuring Natural Language to Query Language: A Review," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6521–6525, Dec. 2020, <https://doi.org/10.48084/etasr.3873>.
- [27] H. Schildt, *Turbo C/C++: The Complete Reference*, 2nd ed. Berkeley, CA, USA: McGraw-Hill Osborne Media, 1992.
- [28] B. Perkins, J. V. Hammer, and J. D. Reid, *Beginning C# 7 Programming with Visual Studio 2017*, 1st ed. Indianapolis, IN, USA: Wrox, 2018.
- [29] K. L. A. El-Ashmawy, "Accuracy, time cost and terrain independence comparisons of levelling techniques," *Geodesy and Cartography*, vol. 40, no. 3, pp. 133–141, Jul. 2014, <https://doi.org/10.3846/20296991.2014.962727>.
- [30] V. Golubev, "Surveying," in *Springer Handbook of Geographic Information*, W. Kresse and D. Danko, Eds. Springer International Publishing, 2022, pp. 281–296, [https://doi.org/10.1007/978-3-030-53125-6\\_11](https://doi.org/10.1007/978-3-030-53125-6_11).
- [31] B. Bird, "Analysis of Survey Point Displacements Using Total Station Measurements," British Columbia Institute of Technology, Technical Report, Apr. 2009.
- [32] E. S. Okiemute, M. N. Ono, and O. O. Fatai, "Monitoring and Analysis of Vertical Deformation of Palm House Benin City Using Digital Level," *International Journal of Advances in Scientific Research and Engineering*, vol. 9, no. 4, pp. 6–16, Sep. 2022, <https://doi.org/10.31695/IJASRE.2018.32860>.



# Investigating the Slope Stability and Factor of Safety Properties of Soil Reinforced with Natural Jute Fibers under Different Rainfall Conditions

**Saurabh Kumar**

Department of Civil Engineering, National Institute of Technology Patna, India  
saurabhk.phd19.ce@nitp.ac.in  
(corresponding author)

**Lal Bahadur Roy**

Department of Civil Engineering, National Institute of Technology Patna, India  
lbroy@nitp.ac.in

*Received: 11 November 2022 | Revised: 19 November 2022 | Accepted: 21 November 2022*

## ABSTRACT

Heavy rainfall is often responsible for embankment failures. During intense rainfall, the embankment slope inclination is vital for slope stability. Some failures occur in the slope due to heavy rainfall and sudden change in the matric suction. Jute fiber is a reinforcing material that is added to improve soil strength. In this research, in order to explore the effects of slope inclination on soil stability, soil samples were collected and exposed to artificial rainfalls. This study presented various tests performed on the soil samples. Different tests like sieve analysis, permeability test, Direct Shear Test (DST), liquid limit, plasticity limit, and numerical modeling were conducted in the laboratory. The study's findings revealed that the failure is caused by a soil suction loss when the inclination of the slope is higher than the soil friction angle and the collapse is caused by the positive water pressure at the slope's toe when it is lower than the soil's friction angle. Furthermore, when the slope angle increases, the slopes are becoming increasingly vulnerable to rapid collapse. After that, jute fibers were combined with the soil to improve its performance. Samples of 2, 3, and 4 rows of jute fibers were tested. These jute fiber samples performed better than the ones without fibers under different rainfall conditions. The distribution of jute fibers had a favorable influence on both strength measurements and safety aspects. Utilizing the factor of safety and matric suction, the performance of jute fiber samples is superior to those without jute fibers. Consequently, by adding jute fibers the stabilization of the soil is significantly improved along with its factor of safety.

*Keywords-soil samples; slope stability, reinforced material; jute fibers; sieve analysis; factor of safety; matric suction; deformation*

## I. INTRODUCTION

Every year, thousands of landslides occur around the world, causing financial losses and even deaths. Soil stabilization is very important for slope embankment. To assure the ground soil's stability during construction, embankments are usually built in several phases. In order to quantify the gain in shear strength effectively, the foundation soil's undrained shear strength normally rises with consolidation [1]. The soil's top layer in an embankment slope is more sensitive to erosion during rainstorms. It provides the least erosion resistance and rapidly loses strength [2]. The undrained condition of soil stability analysis was calculated by First Order Reliability Method (FORM), and Monte Carlo Simulation (MCS) method

to calculate the Factor of Safety (FoS) in [3]. Ground water samples obtained from the river side embankment to check the drinking water quality and 12 parameters, namely, pH, EC, TDS, Ca, Mg, Na, F, SO<sub>4</sub>, K, Cl, and HCO<sub>3</sub> were used to compute the water quality based on the water quality index method [4]. The bearing capacity of a stone column is highly related to the diameter of the column, the properties of the soil, and its bearing capacity. It was found that when the diameter of the stone column increases, the bearing capacity of the soil increases exceptionally [5]. Rainfall infiltration reduces the soil matric suction and shear power and increases the soil loss susceptibility [6]. Displacements in the vertical and lateral planes were reduced and the bearing capacity was increased by

reinforced embankments [7]. To suit engineering requirements, soil stabilization is a technique of enhancing the physical, chemical, or biological qualities of natural soil. Shear strength quality, cohesiveness, and friction angle are determined utilizing a laboratory direct shear test [8]. The finite difference approach was utilized to perform numerical modeling and in the numerical modeling, the model scale was used to replicate the centrifuge tests [9].

Due to the high cost of creating and transporting free-draining coarse-grained materials, slope failure along highways is frequently repaired or restored using locally available soil with significant amounts of fines present (clay and silt) [10]. Due to the three-dimensional nature of the reinforcement, a geocell-reinforced slope acts as a beam in the soil. Furthermore, the moment of inertia is taken into consideration due to its bending features, resulting in enhanced bending strength, the reinforcement lowers slope displacement, and the slope's FoS is raised [11]. Jute fibers of predefined length and content are mixed extensively by hand with expansive soil until a fairly uniform distribution of jute fibers in the soil is achieved [12]. Fibers improve the soil composite mechanical properties. The normal pressures acting on the soil composite mobilize tensile resistance in the fibers, giving the soil higher shear strength [13]. The reinforcing effect location on the failure and deformation behavior of soil slope and superstructure by evaluating various geogrid layers is explored in [14]. Jute fiber is excellent for enhancing soil qualities since it is inexpensive, readily available, biodegradable, and environmentally friendly [15, 16]. A moisture barrier and altered recycled plastic pins were utilized to stabilize the pavement distresses and the rainfall-induced slope failure in [17]. The slopes' FoS or chance of failure were displayed in ASCII format in [18]. The matric suction losses caused by rainfall infiltration had only a minor impact on the reinforced sand slope's global and local FoS. The tensile reinforcement strengths are necessary to sustain FoS on reinforced clay and silt slopes [19].

Under relatively modest stress conditions, from a geotechnical standpoint, problematic soils are those that can disperse, collapse, have an excessive settlement, expand, or even fail. There is erosion or cracking in the soil, and consequently the soil may not be stable for the embankment. Due to this problem, the current paper presents various tests for testing slope soil stability. Here, rainfall intensity is set up to test the soil stability, because it generates cracks and slope failure may occur and reinforcing material like jute fibers is added, to strengthen the soil and increase the slope stability. This study is based on the reinforced soil stabilization process. It investigates various researches for slope stability. Authors in [20] proved the embankment slope stability for roads erected on soft terrain using Prefabricated Vertical Drains (PVDs). The slope stability was investigated using numerical analysis and modeling based on the FoS with an artificial neural network. A novel approach is proposed in [21], where the soil strength follows the nonlinear yield criterion to determine the FoS of soil slopes. First, the magnitude and coordinates of minor primary stress are determined using numerical simulation. The freeze-thaw cycles were investigated by the effect on the embankment slope stability of polypropylene fibers using a numerical analytic method [22]. The DST was built to measure

9 sets of specimens, shear strength metrics such as binding force and internal friction angle were considered along with 3 components (freeze-thaw cycle, fiber content, and fiber length). Authors in [23] investigated the failure and deformation behaviors of reinforced slopes with various geotextile layouts utilizing a set of centrifuge model drawdown experiments. Under drawdown conditions, the geotextile reinforcement was found to considerably raise the ductility and safety limit, change the slope failure feature, and minimize displacement. The injection approach was used to produce one untreated embankment and 3 small-scale bio-treated sand embankments in [24]. To investigate the embankment's slope collapse, rainfall scouring experiments were used.

## II. RESEARCH METHODOLOGY

Geotextile-reinforced model slopes are investigated with tests regarding the geotextiles' mechanism and the reinforcement effect. The soil's water content and dry density, and the slope's height and gradient, were kept constant throughout all of tests. The reinforced slopes' behavior influence is observed utilizing different geotextile layouts. Different centrifuge experiments were undertaken on unreinforced and reinforced slopes. For this experiment, disturbed and undisturbed soil samples were considered. During the sampling process, the soil's natural structure has been altered or destroyed, their structure is referred to as disturbed samples. The non-representative samples are soils from different mixed strata. The deadweight of the soil slope, which was applied to the entire slope, and the evenly distributed loading, which was applied to replicate the vehicle load are the two types of loadings that were used to the finite element model in the slope stability analysis. Using the proposed slope stability analysis approach, the maximum plastic strain and the slope's maximum horizontal displacement were respectively computed. The flow diagram of the proposed method is depicted in Figure 1.

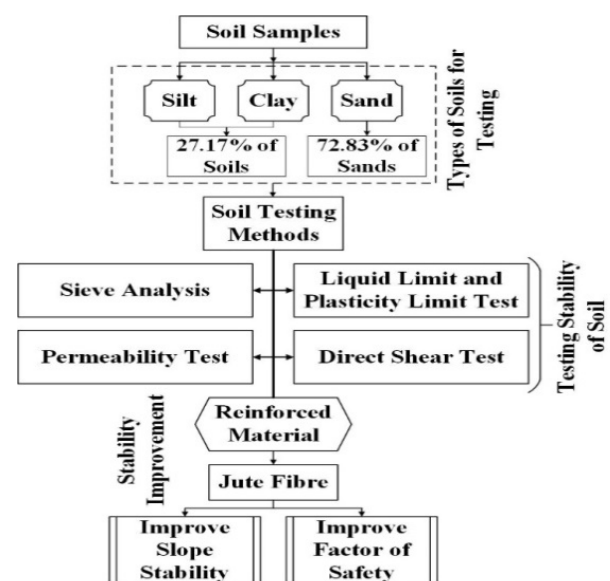


Fig. 1. Flow diagram of the research methodology.

### A. Material Collection

This study studies soil samples to determine the soil embankment stability. The soil samples taken for this study have 27.17% silt and clay soil and 72.83% sand. Multiple, smaller samples known as cores, compose an accurate soil sample. The majority of their nutrients are derived from plants or grasses that individual cores contain from the soil surface. In a variety of locations throughout the area, a minimum of 10 soil cores was taken. Several tests were used to measure the stability of the soil slope. Soil samples were powdered, sieved, and dried before being analyzed. For analysis, a homogenous mixture was ensured by grinding and sifting steps. Soil samples were dried in cardboard boxes at 50°C. The soil was ground in a mechanical mortar and pestle and passed via a 12-mesh, roughly 2mm screen.

Mineral particles, water, air, and organic matter are present in the samples. Chemistry, color, structure, porosity, and soil texture are determined by combinations of these factors. Table I illustrates the plastic limit, liquid limit, plasticity index, coefficient of permeability, dry unit weight, saturated unit weight, cohesion, angle of internal friction, Poisson's ratio, and Young's modulus of elasticity. These are the soil properties utilized in the experiment. To test the soil stability, various soil testing methods were utilized. Specified analytical equipment is utilized for the analysis. Sample preparation included sample cleaning up, sample pre-concentration, alkali or acid-based chemical digestion, sample extraction, crushing and dissolution. In Table I, the coefficient of permeability ( $K_{sat}$ ) was obtained by the constant head permeability test and the modulus of elasticity ( $E$ ) and the coefficient of internal friction ( $\phi$ ) were obtained by direct shear test in the laboratory.

TABLE I. SOIL SAMPLE PROPERTIES

Soil properties	Values
Plastic limit $W_p$	15%
Liquid limit $W_L$	21%
Plasticity Index PI	6%
Soil classification	SM-SC
Coefficient of permeability $K_{sat}$	$8.27 \times 10^{-4} \text{cm/sec}$
Dry unit weight $\gamma_{dry}$	$14.5 \text{kN/m}^3$
Saturated unit weight $\gamma_{sat}$	$18.67 \text{kN/m}^3$
Cohesion $C$	5kPa
The angle of internal friction $\phi$	$26.8^\circ$
Young's modulus of elasticity $E$	5110kPa
Poisson's ratio $\nu$	0.25

### B. Reinforced Slope with Jute Fibers for Slope Stability

By adding a surface cover, slopes can be stabilized. The slope reinforcement and the geometry of the slope can be modified with support structures. Continuous fibers are woven, stitched, braided, or knitted into the fabric from twisted and plied yarn to create multidirectional reinforcements. Almost any reinforcing fiber can be utilized to make fabrics. To enhance its resiliency and strength, incorporating reinforced material into the soil is involved in the method of soil stabilization. Jute fibers are widely known for their ability to be spun into coarse and strong threads. Jute fibers are noted for their glossy, long, and soft nature. Jute fibers can be utilized to strengthen soil to improve its engineering qualities. To endure

rot and heat, jute fibers are utilized because of their high tensile strength, superior durability, and excellent drainage and filtration provided by its porous texture. Jute fibers are also inexpensive, readily available, biodegradable, and environmentally beneficial. Lateral deformation prevention, settlement, ductility, bearing capacity, and strength are improved by the soil reinforcement. Jute fibers have the highest tensile strength among all natural fibers and can withstand rolling and heat. The fibers provide a high level of fineness and flexibility. Jute fibers were utilized for the slope stability improvement (Figure 2). In terms of fatigue and mechanical qualities, woven jute hybrids were found to be environmentally benign, cost-effective, and lightweight.



Fig. 2. Jute fibers.

TABLE II. JUTE FIBERS MECHANICAL PROPERTIES

Characteristics	Values
Density $\rho$	1.35gm/cc
Young's Modulus $E$	20GPa
Poisson's ratio $\nu$	0.38
Shear Modulus $G$	7.24GPa
Tensile Strength $T$	393MPa

### III. EXPERIMENTATION AND RESULT DISCUSSION

All slope models were created in a specially designed chamber. The initial ratio of the slope is 2:1, and it contains rainfall nozzle, frame, inlet pipe, outlet pipe, rain gauge, and a digital tensiometer. Initially, the soil samples are set up in slope format. Here, the rainfall intensity is fixed in the experiment to test the slope soil stability. The rainfall nozzle, inlet pipe, and outlet pipe are utilized for rainfall intensity. After the rainfall intensity, cracks are appearing on the soil and slope failure occurs.

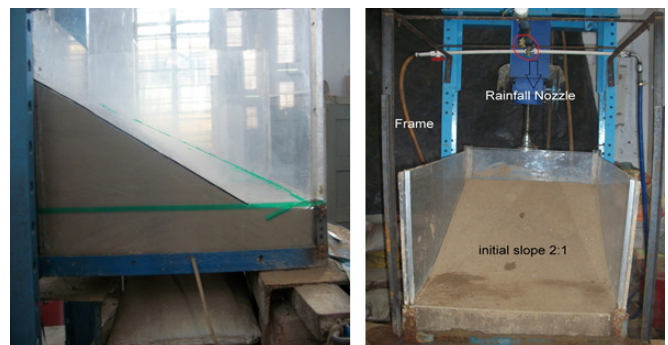


Fig. 3. Experimental setup.

For the rainfall measurement, a standard rain gauge instrument was used. This is a fundamentally 203mm diameter

circular funnel that collects rain into a calibrated and graduated cylinder. Precipitation up to 25mm can be recorded by the measuring cylinder. The free energy shifts in a unit volume of water, known as matric suction. The digital tensiometer was used to measure the matric suction of the soil. Figure 5 represents the experimental results of saturated slope deformation without utilizing jute fibers. This depicts the original slope as 2.25:1, whereas the embankment slopes are 2:1, 1.5:1, and 1.25:1.

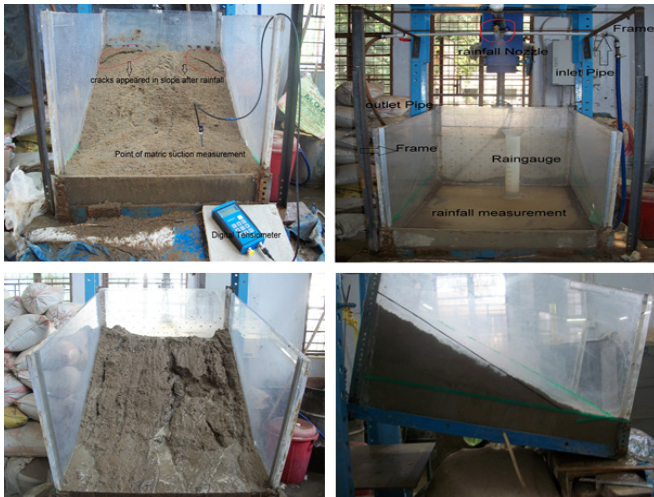


Fig. 4. Measurement of matric suction, rainfall, and slope failures.



Fig. 5. Experimental results of saturated slope deformation without jute reinforcement.

#### A. Numerical Modeling using the Geo-Studio 2021 Software

Geo-studio software is generally used for slope analysis, seepage analysis, and deformation analysis. It is based on the finite element method. By using this, the FoS of the slope can be determined at a different effective angle. By using SIGMA/W 2021, the maximum deformation of slope can be calculated and after that it can be compared with the experimental results. The FoS of the slope can be calculated using SLOPE/W 2021. Here, initially seepage analysis was done by using SEEP/W 2021 and after that, it is utilized as a parent analysis in SLOPE/W and SIGMA/W for calculating the deformation vector, the deformation mesh, and FoS.

$$\text{FoS} = \frac{\text{Ultimate Load (Strength)}}{\text{Allowable Load (Stress)}} \quad (1)$$

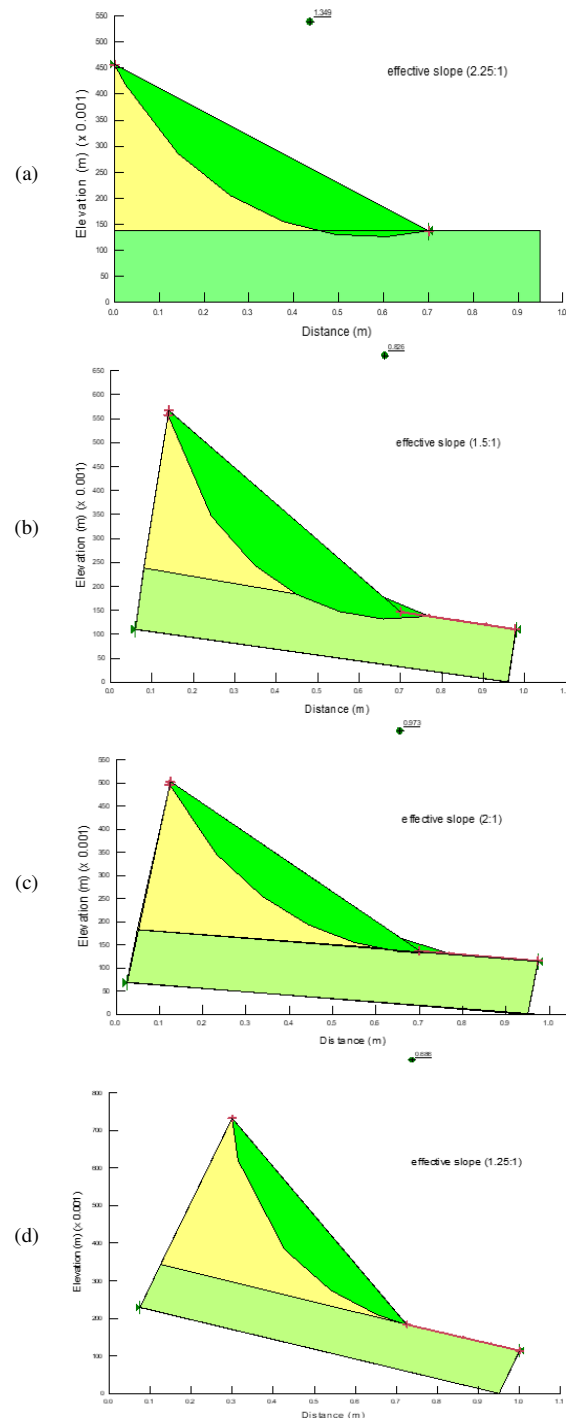


Fig. 6. FoS of original slope and different embankment slopes. (a) 2.25:1, (b) 1.5:1, (c) 2:1, (d) 1.25:1.

TABLE III. FOS AT DIFFERENT EMBANKMENT SLOPES

Embankment slope	FoS
2.25:1	1.349
2:1	0.973
1.5:1	0.826
1.25:1	0.686

Figure 6 exhibits the FoS of original and embankment slopes. The initial Figure depicts the original slopes of 2.25:1, next figure reveals the factor of safety for the embankment slopes of 2:1. Then the factor of safety for the embankment slopes 1.5:1 and 1.25:1 is demonstrated in the last two figures.

#### B. Deformation Analysis by using Geo-Studio SIGMA/W 2021

To perform deformation and stress analysis of earth structures, the finite element CAD software SIGMA/W was utilized. Deformation examination for mesh deformation and vectors has been completed at different embankment slopes. Maximum deformation was recorded and its value was compared with the experimental data. Figure 7 depicts the vector deformation of different embankment slopes of 2:1, 1.5:1, and 1.25:1. Vector deformation defined the deformation pattern during slope failure of each soil particle present inside the embankment slope after heavy rainfall.

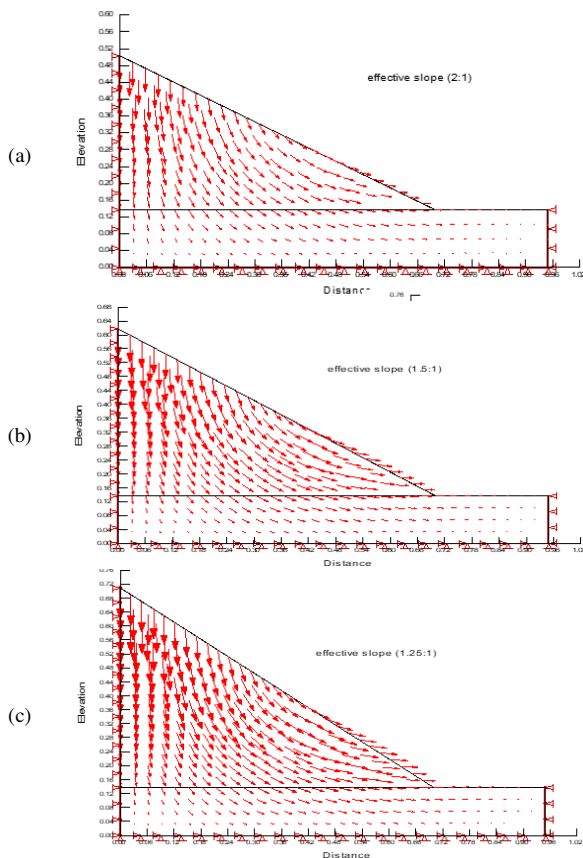


Fig. 7. Vector deformation of different embankment slopes.

Figure 8 depicts the matric suction comparison with and without 3 rows of jute for 50 mm/hr and 100 mm/hr rainfall at the middle of the slope. It shows that samples with jute perform better than the ones without jute. Figure 9 represents the variation of matric suction with time for the top, middle, and bottom of the slope. It is noticed that the matric suction variation for the top layer is higher than the matric suction variation of the middle and bottom layers.

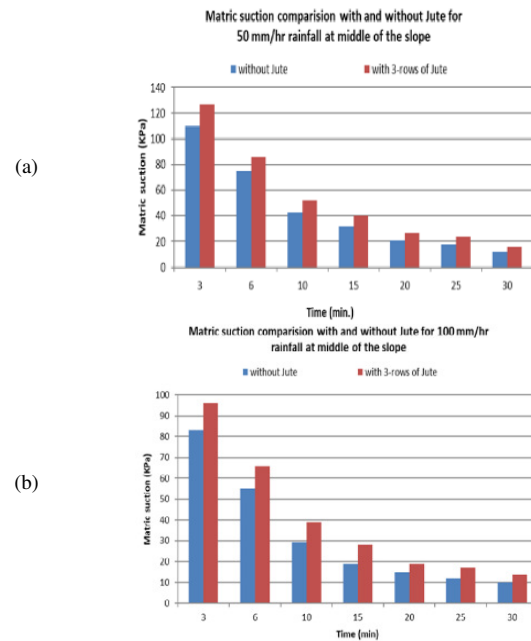


Fig. 8. Matric suction comparison with and without jute fibers.

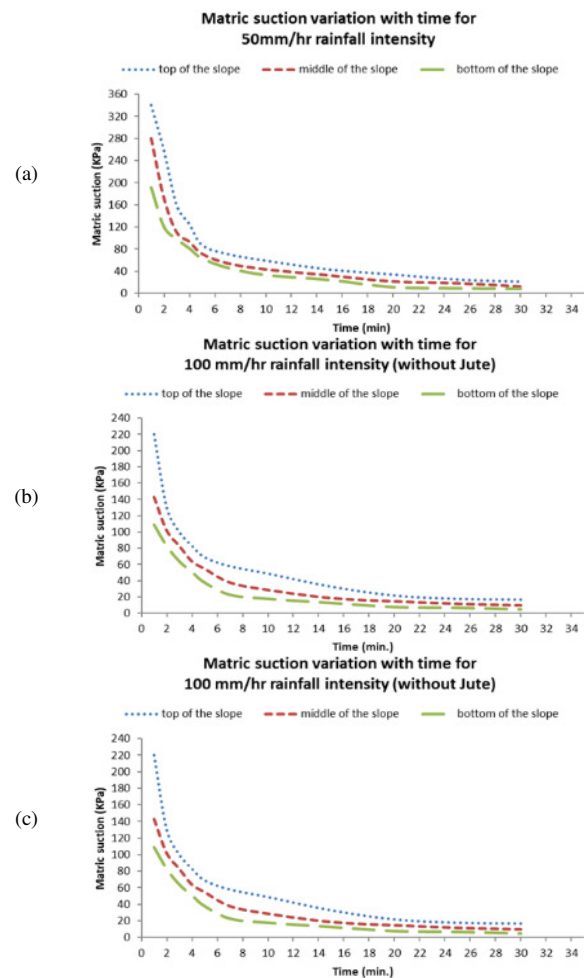


Fig. 9. Matric suction variation with time.



TABLE IV. EXPERIMENTAL AND NUMERICAL VALUES OF USING JUTE FIBERS

Emb. slope	No jute (mm)		2 rows of jute (mm)		3 rows of jute (mm)		4 rows of jute (mm)	
	Exp. value	Num. value	Exp. value	Num. value	Exp. value	Num. value	Exp. value	Num. value
2:1	12.5	9.8	10.6	7.6	8.5	5.6	4.4	2.3
1.5:1	17.7	14.6	16.6	13.7	13.4	7.8	9.6	4.6
1.25:1	33.8	25.8	29.2	22.4	24.8	14.6	18.5	11.8

Table IV reveals the deformation values for different embankment slopes, using 2, 3, and 4 rows of jute and without jute. It can be seen that the deformation when using jute is very lower than when not using jute. The performance of using 4 rows of jute is higher than the others.

Figure 10 demonstrates the FoS variation with and without using jute for the rainfall intensity of 50, 100, and 150mm/hr. The Figure clearly shows that the performance of 4-row jute samples is better than the others'.

#### IV. CONCLUSIONS

Soil stabilization is very important for embankment, but soils like clay and silt are weak for slope embankment. During heavy rainfall, these soils produce cracks and slope failures. To overcome these problems, jute fibers are combined with the soil to improve its stability. This research attempts to experimentally test slope stability. The rainfall intensity is used to assess the slope's stability and several soil tests were carried out. During the experimental and numerical analysis of the slope stability, several cracks and deformations occurred. Due to this failure, the inclusion of jute fibers and its effectiveness to strengthen the soil slope under heavy rainfall condition was assessed. For the embankment slope 2, 3, and 4 rows of jute fibers were considered, and the factor of safety and matric suction were calculated. The main conclusions of the current study are:

- According to the sieve analysis, the soil samples are well-graded with 27.17% silt and clay soil and 72.83% sand.
- The liquid limit of the soil is 21%, the plasticity limit is 15%, and the plasticity index is 6%.
- The results from the soil permeability test gave a dry unit weight of  $14.5\text{kN/m}^3$  and a saturated unit weight of  $18.67\text{kN/m}^3$ .
- Regarding the direct shear test for sample test 1, the normal stress is  $0.5\text{kg/cm}^2$ , and the strain rate is  $0.25\text{mm/min}$ . The predicted shear load failure is  $0.093\text{kN}$ , the failure shear stress is  $0.29588\text{kg/cm}^2$ . For sample test 2, the normal stress is  $78.48\text{kN/m}^2$ , and the strain rate is  $0.25\text{mm/min}$ , the shear load and shear stress failure are  $0.133\text{kN}$  and  $42.97\text{kN/m}^2$ . For sample test 3, the normal stress and the strain rate are  $98.1\text{kN/m}^2$  and  $0.25\text{mm/min}$ , and the shear load and shear stress failure are  $0.160\text{kN}$  and  $51.993\text{kN/m}^2$ .
- The experimental results for slope deformation and factor of safety were determined, and the vector deformation and mesh deformation were expressed for embankment slopes of 2:1, 1.5:1, and 1.25:1. The deformation analysis was performed in Geo-studio SIGMA/W 2021.
- The experimental results utilizing jute fibers for factor of safety, mesh deformation, and matric suction are described in the test results section above. The overall performance of 4-row jute fiber samples is better than without using jute fibers or by using less rows of fibers.
- Utilizing jute fibers improves the stability of the soil, the factor of safety, and matric suction performance.

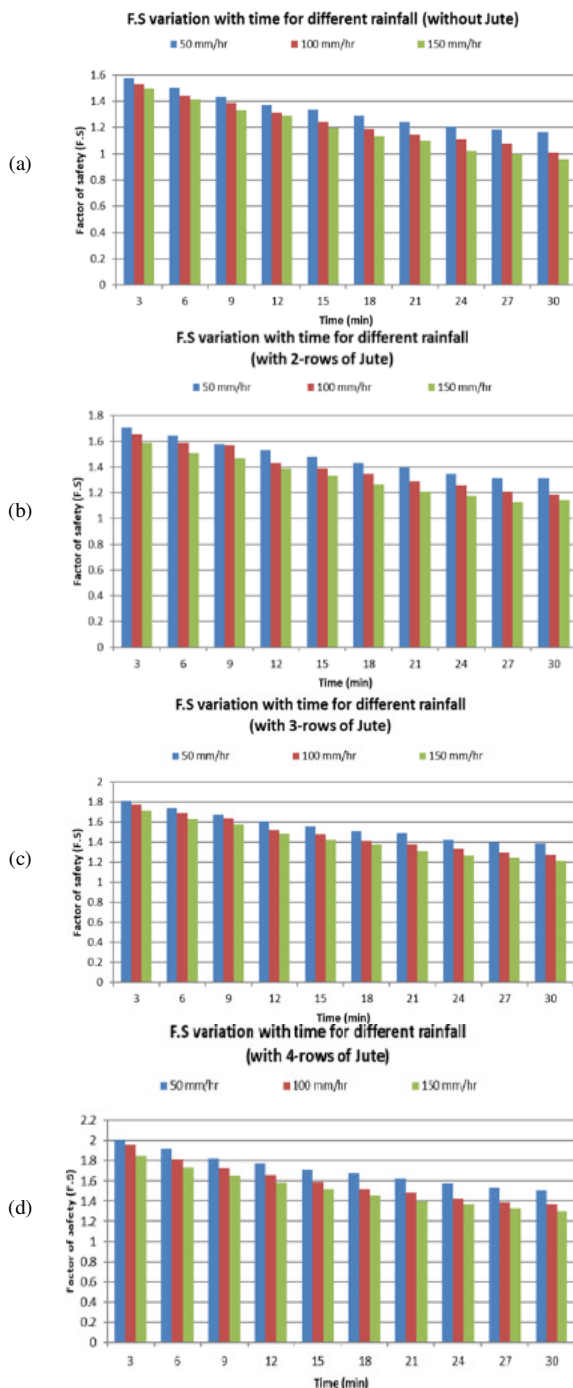


Fig. 10. FoS variation with (a) 2, (b) 3, (c) 4 rows of jute fibers and (d) without jute fibers.

## REFERENCES

- [1] F. Kassou, J. B. Bouziyane, A. Ghafiri, and A. Sabihi, "Slope Stability of Embankments on Soft Soil Improved with Vertical Drains," *Civil Engineering Journal*, vol. 6, no. 1, pp. 164–173, Jan. 2020, <https://doi.org/10.28991/cej-2020-03091461>.
- [2] M. U. Farooq, H. Mujtaba, K. Farooq, N. Sivakugan, and B. M. Das, "Evaluation of Stability and Erosion Characteristics of Soil Embankment Slope Reinforced with Different Natural Additives," *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, vol. 44, no. 1, pp. 515–524, Oct. 2020, <https://doi.org/10.1007/s40996-019-00340-5>.
- [3] S. S. Kar and L. B. Roy, "Probabilistic Based Reliability Slope Stability Analysis Using FOSM, FORM, and MCS," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8236–8240, Apr. 2022, <https://doi.org/10.48084/etasr.4689>.
- [4] K. Praveen and L. B. Roy, "Assessment of Groundwater Quality Using Water Quality Indices: A Case Study of Paliganj Distributary, Bihar, India," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8199–8203, Feb. 2022, <https://doi.org/10.48084/etasr.4696>.
- [5] S. Saxena and L. B. Roy, "The Effect of Geometric Parameters on the Strength of Stone Columns," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9028–9033, Aug. 2022, <https://doi.org/10.48084/etasr.5138>.
- [6] S. Kumar and L. B. Roy, "Rainfall Induced Geotextile Reinforced Model Slope Embankment Subjected to Surcharge Loading: A Review Study," *Archives of Computational Methods in Engineering*, vol. 29, no. 5, pp. 3203–3221, Aug. 2022, <https://doi.org/10.1007/s11831-021-09688-2>.
- [7] L. Li, F. Cui, P. Ferreira, H. Xiao, and H. Jie, "Experimental study of embankments with different reinforcement materials and spacing between layers," *Geotextiles and Geomembranes*, vol. 47, no. 4, pp. 477–482, Aug. 2019, <https://doi.org/10.1016/j.geotexmem.2019.03.003>.
- [8] S. Kundu, S. Sarkar, and P. Dhar, "Analysis of Slope Stability of Soil Using Fly-ash & Lime," in *Indian Geotechnical Society, Kolkata Chapter Advanced Construction and Computational Tools in Geotechniques -Practice to Theory*, Kolkata, India, Jul. 2018.
- [9] H. Zhou, G. Zheng, J. Liu, X. Yu, X. Yang, and T. Zhang, "Performance of embankments with rigid columns embedded in an inclined underlying stratum: centrifuge and numerical modelling," *Acta Geotechnica*, vol. 14, no. 5, pp. 1571–1584, Oct. 2019, <https://doi.org/10.1007/s11440-019-00825-7>.
- [10] D. G. Abebe, "Determination of Effect of Moisture Content and Density on Shear Strength Parameters and Slope Stability of Highly Plastic Silt Embankment Soil (the Case of Wozeka-Gidole Road)," *International Research Journal of Engineering and Technology*, vol. 7, no. 8, pp. 2333–2346, Aug. 2020.
- [11] B. B. Mehdipour, H. Hashemalhosseini, B. Nadi, and M. Mirmohamadsadeghi, "An investigation into a Geocell-reinforced Slope in The Unsaturated Numerical Model," *Structural Engineering & Geotechnic*, vol. 9, no. 1, [Online]. Available: <https://civilica.com/doc/1034216/>.
- [12] Y.-X. Wang *et al.*, "Laboratory Investigation on Strength Characteristics of Expansive Soil Treated with Jute Fiber Reinforcement," *International Journal of Geomechanics*, vol. 17, no. 11, Nov. 2017, Art. no. 04017101, [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000998](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000998).
- [13] R. Ramkrishnan, V. Karthik, M. R. Sruthy, and A. Sharma, "Soil Reinforcement and Slope Stabilization Using Natural Jute Fibres," in *New Solutions for Challenges in Applications of New Materials and Geotechnical Issues*, Cham, 2019, pp. 130–143, [https://doi.org/10.1007/978-3-319-95744-9\\_11](https://doi.org/10.1007/978-3-319-95744-9_11).
- [14] F. Luo, R. Huang, and G. Zhang, "Centrifuge modeling of the geogrid-reinforced slope subjected to differential settlement," *Acta Geotechnica*, vol. 15, no. 10, pp. 3027–3040, Oct. 2020, <https://doi.org/10.1007/s11440-020-01010-x>.
- [15] K. S. Kulhar and M. Raisinghani, "Shear Strength Performance of Sandy Soil Reinforced with Jute Fiber," *Journal of Basic and Applied Engineering Research*, vol. 4, no. 7, pp. 624–629, Oct. 2017.
- [16] N. A. Dar and D. G. Bhalla, "Stabilization of Soil Using Jute Fiber and Stone Dust," *International Journal of Scientific Development and Research*, vol. 5, no. 8, pp. 325–333, Aug. 2020.
- [17] A. Sapkota, A. Ahmed, P. Pandey, M. S. Hossain, and N. Lozano, "Stabilization of Rainfall-Induced Slope Failure and Pavement Distresses Using Recycled Plastic Pins and Modified Moisture Barrier," pp. 237–246, Mar. 2019, <https://doi.org/10.1061/9780784482070.023>.
- [18] S. Naidu, K. S. Sajinkumar, T. Oommen, V. J. Anuja, R. A. Samuel, and C. Muraleedharan, "Early warning system for shallow landslides using rainfall threshold and slope stability analysis," *Geoscience Frontiers*, vol. 9, no. 6, pp. 1871–1882, Nov. 2018, <https://doi.org/10.1016/j.gsf.2017.10.008>.
- [19] K.-H. Yang, J. N. Thuo, V. D. A. Huynh, T. S. Nguyen, and F. H. M. Portelinha, "Numerical evaluation of reinforced slopes with various backfill-reinforcement-drainage systems subject to rainfall infiltration," *Computers and Geotechnics*, vol. 96, pp. 25–39, Apr. 2018, <https://doi.org/10.1016/j.compgeo.2017.10.012>.
- [20] R. C. Mamat, A. M. Samad, A. Kasa, S. F. M. Razali, A. Ramli, and M. B. H. C. Omar, "Slope stability prediction of road embankment on soft ground treated with prefabricated vertical drains using artificial neural network," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 2, pp. 236–243, Jun. 2020, <https://doi.org/10.11591/ijai.v9.i2.pp236-243>.
- [21] Y. X. Li and X. L. Yang, "Soil-Slope Stability considering Effect of Soil-Strength Nonlinearity," *International Journal of Geomechanics*, vol. 19, no. 3, Mar. 2019, Art. no. 04018201, [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0001355](https://doi.org/10.1061/(ASCE)GM.1943-5622.0001355).
- [22] Y. Gong, Y. He, C. Han, Y. Shen, and G. Tan, "Stability Analysis of Soil Embankment Slope Reinforced with Polypropylene Fiber under Freeze-Thaw Cycles," *Advances in Materials Science and Engineering*, vol. 2019, Jan. 2019, Art. no. e5725708, <https://doi.org/10.1155/2019/5725708>.
- [23] F. Luo, G. Zhang, Y. Liu, and C. Ma, "Centrifuge modeling of the geotextile reinforced slope subject to drawdown," *Geotextiles and Geomembranes*, vol. 46, no. 1, pp. 11–21, Feb. 2018, <https://doi.org/10.1016/j.geotexmem.2017.09.001>.
- [24] Z. Wang, Q. Li, N. Zhang, Y. Jin, H. Qin, and J. Ding, "Slope failure of biotreated sand embankments under rainfall conditions: experimental investigation and numerical simulation," *Bulletin of Engineering Geology and the Environment*, vol. 79, no. 9, pp. 4683–4699, Nov. 2020, <https://doi.org/10.1007/s10064-020-01850-7>.

# An Analytical Solution for the Dynamics of a Functionally Graded Plate resting on Viscoelastic Foundation

**Dang Thu Thuy**

University of Transport and Communications, Vietnam  
dtthuy@utc.edu.vn

**Nguyen Ngoc Lam**

University of Transport and Communications, Vietnam  
nngoclamkc@utc.edu.vn  
(corresponding author)

**Dao Ngoc Tien**

Hanoi Architectural University, Vietnam  
tiendn@hau.edu.vn

**Hoang Van Thanh**

Vietnam Maritime University, Vietnam  
hoangvanthanh@vimaru.edu.vn

*Received: 16 October 2022 | Revised: 5 November 2022 | Accepted: 12 November 2022*

## ABSTRACT

This paper deals with the dynamic response of Functionally Graded Material (FGM) plates resting on a viscoelastic foundation under dynamic loads. The governing equations are derived by using Hamilton's principle using the classical plate theory and the higher-order shear deformation plate theory. Using state-space methods to find the closed-form solution of the dynamic response of functionally graded rectangular plates resting on a viscoelastic foundation. Numerical examples are given for displacement and stresses in the plates with various structural parameters and the effects of these parameters are discussed. The result of the numerical example shows a marked decrease in displacement and stresses as the coefficient of viscous damping is increased.

**Keywords-***dynamic; FGM plate; viscoelastic; viscoelastic foundation; analytical method*

## I. INTRODUCTION

Structural engineering problems such as in beams, frames, plates, and shells can be solved by experimental methods [1-3], analytical methods [4, 5], and numerical methods such as finite elements [6-11]. The dynamic problem of beams or plates on foundations has been investigated by many researchers [12-17]. When using analytical and semi-analytical methods, we depend on the boundary condition to choose the approximation function of the displacement fields. Usually, the Navier approach [18] and the Ritz method [19] are used for simple rectangular plates, and the Ritz method [20] and Fourier-Bessel series [21] for circular plates. The dynamic stability of the orthotropic plates subjected to an arbitrary dynamic load was studied with the Galerkin method in [22]. The dynamic response of plates [23] on an elastic foundation is subjected to

moving loads using the strip method. An asymptotic theory was used in [24] to study the dynamic response of anisotropic inhomogeneous and laminated plates. The Galerkin method was used to deal with the analysis of the nonlinear dynamic response of a laminated composite plate subjected to blast loading in [25]. The wave propagation of the rectangular FGM plates [26] with clamped supports under impulse load using the dispersion relation and integral transforms. The dynamic behavior of fiber-reinforced plastic sandwich plates with PVC foam core was analytically studied in [27]. The dynamic stiffness matrix was constructed in [28] for an infinite or semi-finite Timoshenko beam on the viscoelastic foundation to the harmonic moving load. The stochastic finite element method was utilized for the calculation of the variability of the eigenvalue coefficient of the function of the graded beam.

In this paper, the governing equation of the functionally graded plates resting on viscoelastic foundation is solved by Hamilton's principle. The sinusoidal function and the Navier approach are used to find the dynamic solution of functionally graded simply supported rectangular plates subjected to step loading.

## II. GOVERNING EQUATION OF THE FGM PLATE

We consider an FGM plate resting on a viscoelastic foundation. The geometry of the plate and the coordinate system is shown in Figure 1.

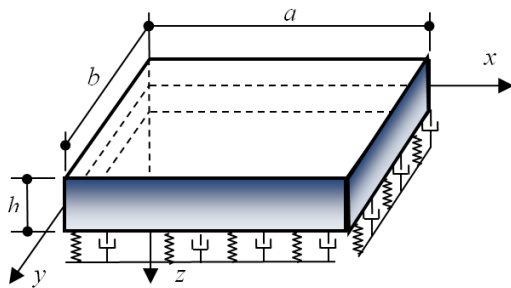


Fig. 1. The geometry of the FGM plate.

The viscoelastic foundation is modeled as a Winkler model with stiffness coefficient  $K$  and damping coefficient  $C$ . The displacement fields at an arbitrary point  $(x, y, z)$  in the plate are shown below. Using the classical plate theory we have:

$$\begin{aligned} U(x, y, z, t) &= u(x, y, z, t) - z \frac{\partial w}{\partial x} \\ V(x, y, z, t) &= v(x, y, z, t) - z \frac{\partial w}{\partial y} \\ W(x, y, z, t) &= w(x, y, t) \end{aligned} \quad (1)$$

Using higher-order shear deformation plate theory proposed by Shimpi [30] we have:

$$\begin{aligned} U(x, y, z, t) &= u(x, y, z, t) - z \frac{\partial w_b}{\partial x} + \left[ \frac{1}{4}z - \frac{5}{3}z \left( \frac{z}{h} \right)^2 \right] \frac{\partial w_s}{\partial x} \\ V(x, y, z, t) &= v(x, y, z, t) - z \frac{\partial w_b}{\partial y} + \left[ \frac{1}{4}z - \frac{5}{3}z \left( \frac{z}{h} \right)^2 \right] \frac{\partial w_s}{\partial y} \\ W(x, y, z, t) &= w_b(x, y, t) + w_s(x, y, t) \end{aligned} \quad (2)$$

The mechanical properties of FGM, such as Young's modulus  $E$  and mass density  $\rho$  are assumed as:

$$\begin{aligned} E(z) &= (E_c - E_m) \left( \frac{z}{h} + \frac{1}{2} \right)^n + E_m \\ \rho(z) &= (\rho_c - \rho_m) \left( \frac{z}{h} + \frac{1}{2} \right)^n + \rho_m \end{aligned} \quad (3)$$

where the subscripts  $m$  and  $c$  represent the metallic and ceramic constituents, respectively, and  $n$  is the power index of the volume fraction. The linear constitutive relations of plate can be written as:

$$\begin{Bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{Bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 & 0 & 0 \\ Q_{12} & Q_{22} & 0 & 0 & 0 \\ 0 & 0 & Q_{44} & 0 & 0 \\ 0 & 0 & 0 & Q_{55} & 0 \\ 0 & 0 & 0 & 0 & Q_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{yz} \\ \gamma_{xz} \\ \gamma_{xy} \end{Bmatrix} \quad (4)$$

where:

$$\begin{aligned} Q_{11} &= \frac{E(z)}{1-\nu^2}, \quad Q_{22} = Q_{11}, \quad Q_{12} = \frac{\nu E(z)}{1-\nu^2} \\ Q_{44} &= G_{23}, \quad Q_{55} = G_{13}, \quad Q_{66} = G_{12} \end{aligned}$$

Using Hamilton's principle we obtain the equations of motion of a plate as follows:

- Using the Classical Plate Theory (CPT):

$$\begin{aligned} \frac{\partial N_x}{\partial x} + \frac{\partial N_{xy}}{\partial y} &= I_0 \ddot{u} - I_1 \frac{\partial \ddot{w}}{\partial x} \\ \frac{\partial N_{xy}}{\partial x} + \frac{\partial N_y}{\partial y} &= I_0 \ddot{v} - I_1 \frac{\partial \ddot{w}}{\partial y} \\ \frac{\partial^2 M_x}{\partial x^2} + 2 \frac{\partial^2 M_{xy}}{\partial x \partial y} + \frac{\partial^2 M_y}{\partial y^2} - K w - C \dot{w} + f &= I_0 \ddot{w} + I_1 \left( \frac{\partial \ddot{u}}{\partial x} + \frac{\partial \ddot{v}}{\partial y} \right) - I_2 \nabla^2 \ddot{w} \end{aligned} \quad (5)$$

- Using the Higher-order Shear Deformation Plate Theory (HSDT):

$$\begin{aligned} \frac{\partial N_x}{\partial x} + \frac{\partial N_{xy}}{\partial y} &= I_0 \ddot{u} - I_1 \frac{\partial \ddot{w}_b}{\partial x} - J_1 \frac{\partial \ddot{w}_s}{\partial x} \\ \frac{\partial N_{xy}}{\partial x} + \frac{\partial N_y}{\partial y} &= I_0 \ddot{v} - I_1 \frac{\partial \ddot{w}_b}{\partial y} - J_1 \frac{\partial \ddot{w}_s}{\partial y} \\ \frac{\partial^2 M_x^b}{\partial x^2} + 2 \frac{\partial^2 M_{xy}^b}{\partial x \partial y} + \frac{\partial^2 M_y^b}{\partial y^2} - K(w_b + w_s) &- C(\dot{w}_b + \dot{w}_s) + q = I_0(\ddot{w}_b + \ddot{w}_s) + I_1 \left( \frac{\partial \ddot{u}}{\partial x} + \frac{\partial \ddot{v}}{\partial y} \right) \\ -I_2 \nabla^2 \ddot{w}_b - J_2 \nabla^2 \ddot{w}_s \frac{\partial^2 M_x^s}{\partial x^2} + 2 \frac{\partial^2 M_{xy}^s}{\partial x \partial y} + \frac{\partial^2 M_y^s}{\partial y^2} &+ \frac{\partial Q_{xz}}{\partial x} + \frac{\partial Q_{yz}}{\partial y} - K(w_b + w_s) - C(\dot{w}_b + \dot{w}_s) + q \\ = I_0(\ddot{w}_b + \ddot{w}_s) + J_1 \left( \frac{\partial \ddot{u}}{\partial x} + \frac{\partial \ddot{v}}{\partial y} \right) - J_2 \nabla^2 \ddot{w}_b - K_2 \nabla^2 \ddot{w}_s \end{aligned} \quad (6)$$

where:

$$\begin{aligned} I_i &= \int_{-h/2}^{h/2} \rho(z) z^i dz, \quad J_i = -\frac{1}{4} I_i + \frac{5}{3h^2} I_{i+2} \\ K_2 &= \frac{1}{16} I_1 - \frac{5}{6h^2} I_4 + \frac{25}{9h^4} I_6 \\ \nabla^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \end{aligned}$$

The force and moment resultants of the plate are defined by:

$$\begin{aligned}
[N] &= \begin{Bmatrix} N_x \\ N_y \\ N_{xy} \end{Bmatrix} = \int_{-h/2}^{h/2} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{Bmatrix} dz \\
[M^b] &= \begin{Bmatrix} M_x^b \\ M_y^b \\ M_{xy}^b \end{Bmatrix} = \int_{-h/2}^{h/2} z \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{Bmatrix} dz \\
[M^s] &= \begin{Bmatrix} M_x^s \\ M_y^s \\ M_{xy}^s \end{Bmatrix} = \int_{-h/2}^{h/2} \hat{f} \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{Bmatrix} dz \\
\begin{Bmatrix} Q_x \\ Q_y \end{Bmatrix} &= \int_{-h/2}^{h/2} \hat{g} \begin{Bmatrix} \sigma_{xz} \\ \sigma_{yz} \end{Bmatrix} dz
\end{aligned} \quad (7)$$

### III. ANALYTICAL SOLUTION FOR THE DYNAMIC RESPONSE OF THE FGM PLATE

The sinusoidal function based on the Navier approach is chosen to satisfy all boundary conditions, as follows:

$$\begin{aligned}
u(x, y, t) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \{U_{mn}(t) \cos \alpha x \sin \beta y\} \\
v(x, y, t) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \{V_{mn}(t) \sin \alpha x \cos \beta y\} \\
w_b(x, y, t) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \{W_{bmn}(t) \sin \alpha x \sin \beta y\} \\
w_s(x, y, t) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \{W_{smn}(t) \sin \alpha x \sin \beta y\}
\end{aligned} \quad (8)$$

$$\text{where } \alpha = \frac{m\pi}{a}, \beta = \frac{n\pi}{b}.$$

Substituting (8) into (5) and (6) the forced vibration of the functionally graded plate can be written as follows:

- Using the CPT:

$$\begin{aligned}
\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{12} & s_{22} & s_{23} \\ s_{13} & s_{23} & s_{33} \end{bmatrix} \begin{Bmatrix} U_{mn} \\ V_{mn} \\ W_{mn} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{c}_{33} \end{bmatrix} \begin{Bmatrix} \dot{U}_{mn} \\ \dot{V}_{mn} \\ \dot{W}_{mn} \end{Bmatrix} \\
+ \begin{bmatrix} m_{11} & 0 & m_{13} \\ 0 & m_{22} & m_{23} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} \begin{Bmatrix} \ddot{U}_{mn} \\ \ddot{V}_{mn} \\ \ddot{W}_{mn} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ F_{mn} \end{Bmatrix}
\end{aligned} \quad (9)$$

- Using the HSDT:

$$\begin{aligned}
\begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{12} & s_{22} & s_{23} & s_{24} \\ s_{13} & s_{23} & s_{33} & s_{34} \\ s_{14} & s_{24} & s_{34} & s_{44} \end{bmatrix} \begin{Bmatrix} U_{mn} \\ V_{mn} \\ W_{bmn} \\ W_{smn} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{c}_{33} & 0 \\ 0 & 0 & 0 & \hat{c}_{44} \end{bmatrix} \begin{Bmatrix} \dot{U}_{mn} \\ \dot{V}_{mn} \\ \dot{W}_{bmn} \\ \dot{W}_{smn} \end{Bmatrix} \\
+ \begin{bmatrix} m_{11} & 0 & m_{13} & m_{14} \\ 0 & m_{22} & m_{23} & m_{24} \\ m_{13} & m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & m_{34} & m_{44} \end{bmatrix} \begin{Bmatrix} \ddot{U}_{mn} \\ \ddot{V}_{mn} \\ \ddot{W}_{bmn} \\ \ddot{W}_{smn} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ F_{mn} \\ F_{mn} \end{Bmatrix}
\end{aligned} \quad (10)$$

where:

$$F_{mn} = \frac{4}{ab} \int_0^b \int_0^a q \sin \alpha x \sin \beta y dx dy \quad (11)$$

We consider uniformly distribution loads as follows:

$$q = q_0 \hat{F}(t) \quad (12)$$

The function of time  $\hat{F}(t)$  of dynamic loadings is considered as step loading:

$$\hat{F}(t) = \begin{cases} 1 & 0 \leq t \leq t_1 \\ 0 & t \geq t_1 \end{cases} \quad (13)$$

In this work, we use the state-space method to solve (9) and (10), which must be rewritten in order to find a solution, as follows:

$$\dot{\mathbf{Z}} = \mathbf{AZ} + \mathbf{b} \quad (14)$$

where:

$$\begin{aligned}
\mathbf{Z} &= \{U_{mn} \quad V_{mn} \quad W_{bmn} \quad W_{smn} \quad \dot{U}_{mn} \quad \dot{V}_{mn} \quad \dot{W}_{bmn} \quad \dot{W}_{smn}\}^T \\
\mathbf{b} &= \{0 \quad 0 \quad 0 \quad 0 \quad b_1 \quad b_2 \quad b_3 \quad b_4\}^T
\end{aligned}$$

The solution of (14) is obtained as:

$$\mathbf{Z}(t) = e^{\mathbf{A}(t-t_0)} \mathbf{Z}(t_0) + \int_{t_0}^t e^{\mathbf{A}(t-\tau)} \mathbf{b}(\tau) d\tau \quad (15)$$

where  $t_0$  is the initial time,  $\mathbf{Z}(t_0)$  is the initial response, and  $e^{\mathbf{A}(t-t_0)}$  is the exponential matrix. This exponential matrix can be formulated in terms of the matrix of eigenvectors and eigenvalues associated with matrix  $\mathbf{A}$ .

### IV. NUMERICAL EXAMPLES

We consider a simply supported rectangular FGM plate with side-to-thickness ratio  $a/h=10$ , rectangular dimensions of  $a=0.3\text{m}$ ,  $b=0.5\text{m}$ , and power index  $p=3$ . The elastic moduli and mass density are chosen to be the same as in [31]:  $\rho_m = 2707\text{kg/m}^3$ ,  $E_m = 70\text{GPa}$ ,  $E_c = 380\text{GPa}$ ,  $\rho_c = 3800\text{kg/m}^3$ , with Poisson's ratio being 0.3. The uniform load is written as

$q_0 = 100 \left( \frac{\text{kN}}{\text{m}^2} \right)$  and the duration of load application time is  $t_1 = 0.002\text{s}$ . We note the normalized stiffness coefficient and the normalized damping coefficient of the foundation as follows:

$$k = \frac{Ka^4}{D_m}, \quad c = \frac{Ca^2}{D_m} \quad (16)$$

where  $D_m = \frac{E_m h^3}{12(1-\nu^2)}$  is the flexural rigidity of a full-metal plate.

The first two mode shapes are shown in Figure 2.



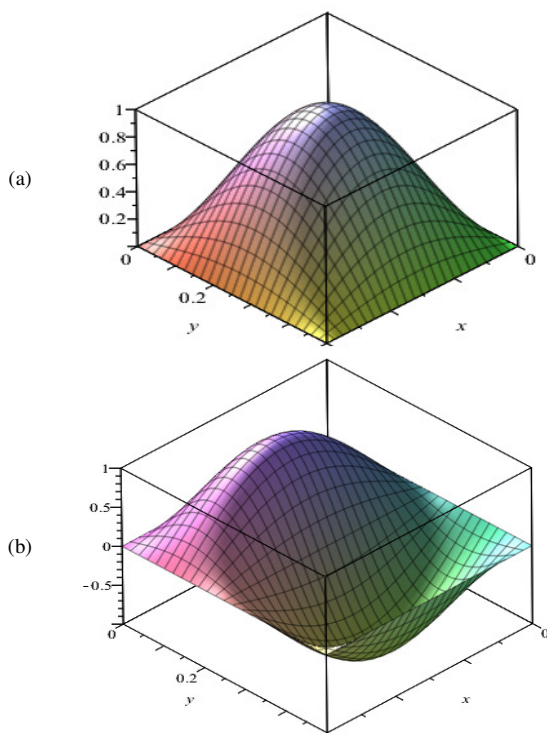


Fig. 2. (a) First and (b) second mode shapes of the simply supported rectangular FGM plate.

We consider two cases of viscoelastic foundation to investigate the effect of the damping coefficient of the viscoelastic foundation:

- Case 1:  $k = 5$ ,  $c = 0.05$
- Case 2:  $k = 5$ ,  $c = 0.1$

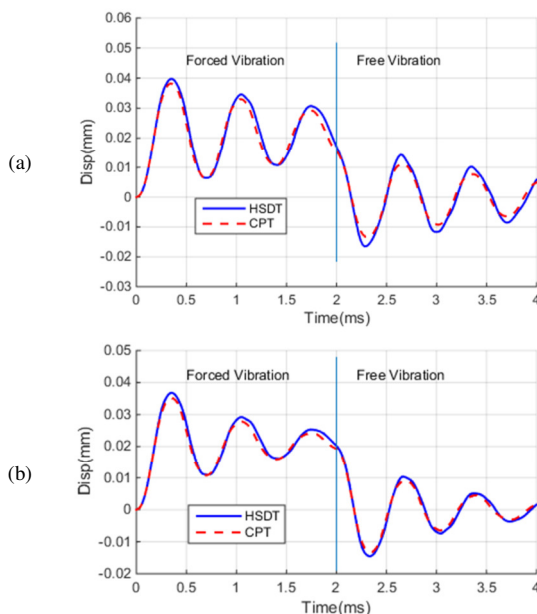


Fig. 3. The deflection response at the central plate. Cases (a) 1, (b) 2.

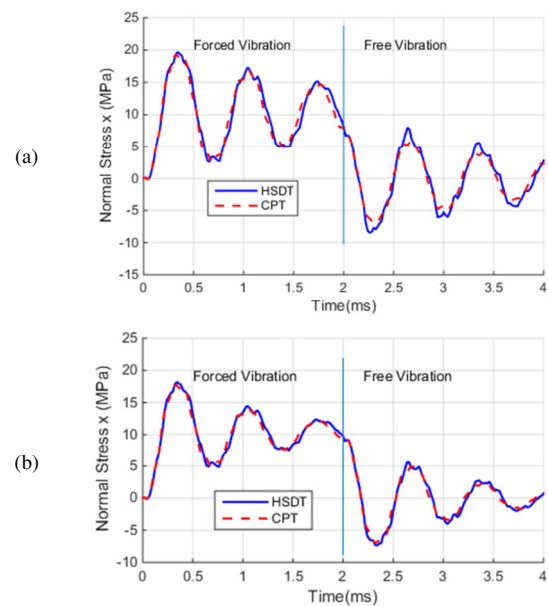


Fig. 4. The normal stress  $\sigma_x$  at the top of the central plate ( $z=h/2$ ). Cases (a) 1, (b) 2.

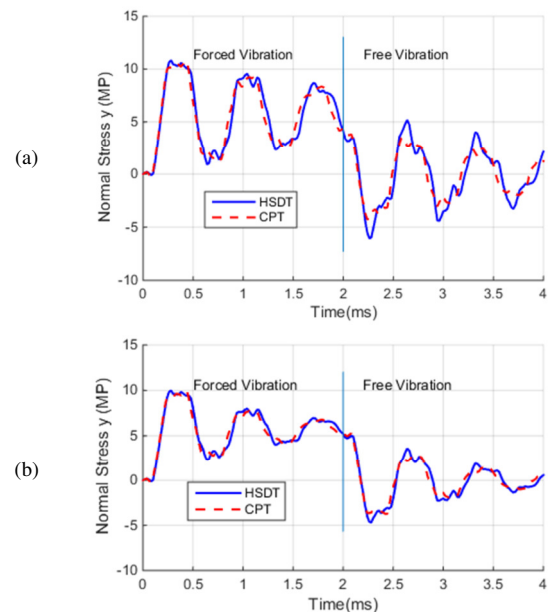


Fig. 5. The normal stress  $\sigma_y$  at the top of the central plate ( $z=h/2$ ). Cases (a) 1, (b) 2.

Figure 2 illustrates the transverse deflection due to step loading as functions of time. Figures 3–4 show the transient response of stress at the center of the plate for the two considered cases. The amplitude of dynamic responses decreases due to the influence of viscous damping. The displacements computed by HSDT are clearly larger than those using the CPT because the plate modeled by CPT is relatively stiffer. However, the deviation of stresses  $\sigma_x$ ,  $\sigma_y$  between CPT and higher order plate theories is small. So, the effect of shear deformation is more significant in predicting displacement than

predicting stresses. In the forced vibration regime, it can be seen in Figures 2–4 that the deflection and stresses predicted for functionally graded plates with  $c=0.05$  are moderately larger than for the plate with  $c=0.1$ . It is observed that when the damping coefficient increases, the deflection and stresses become smaller, as expected.

## V. CONCLUSIONS

This research computed the vibration of the functionally graded plate resting on the viscoelastic foundation by the analytical method, using both classical plate theory and higher-order shear deformation plate theory. The analytical solution for the dynamic functionally graded plate was solved to be a double sine series based on the Navier approach. Dynamic responses are considered for both forced and free vibrations. The results show that the damping coefficient of the foundation has a significant influence on the dynamic response of the functionally graded plate. The damping factor of the foundation dissipates energy, reducing the ambient vibrations in forced and free vibrations. This work is significant to structural engineering, and the results can be used in practical design structures.

## ACKNOWLEDGMENT

This research is funded by University of Transport and Communications (UTC) under grant number T2022-CT-023.

## REFERENCES

- [1] Z. H. Dakhel and S. D. Mohammed, "Castellated Beams with Fiber-Reinforced Lightweight Concrete Deck Slab as a Modified Choice for Composite Steel-Concrete Beams Affected by Harmonic Load," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8809–8816, Aug. 2022, <https://doi.org/10.48084/etasr.4987>.
- [2] A. N. Hassooni and S. R. A. Zaidee, "Behavior and Strength of Composite Columns under the Impact of Uniaxial Compression Loading," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8843–8849, Aug. 2022, <https://doi.org/10.48084/etasr.4753>.
- [3] Z. H. Abdulghafoor and H. A. Al-Baghdadi, "Static and Dynamic Behavior of Circularized Reinforced Concrete Columns Strengthened with Hybrid CFRP," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9336–9341, Oct. 2022, <https://doi.org/10.48084/etasr.5162>.
- [4] M. V. Shitikova and A. I. Krusser, "Force driven vibrations of nonlinear plates on a viscoelastic winkler foundation under the harmonic moving load," *International Journal for Computational Civil and Structural Engineering*, vol. 17, no. 4, pp. 161–180, Dec. 2021, <https://doi.org/10.22337/2587-9618-2021-17-4-161-180>.
- [5] H.-D. Ta and H.-C. Noh, "Analytical solution for the dynamic response of functionally graded rectangular plates resting on elastic foundation using a refined plate theory," *Applied Mathematical Modelling*, vol. 39, no. 20, pp. 6243–6257, Oct. 2015, <https://doi.org/10.1016/j.apm.2015.01.062>.
- [6] D. H. Duc, D. V. Thom, and P. M. Phuc, "Buckling Analysis of Variable Thickness Cracked Nanoplates considering the Flexoelectric Effect," *Transport and Communications Science Journal*, vol. 73, no. 5, pp. 470–485, Jun. 2022, <https://doi.org/10.47869/tcsj.73.5.3>.
- [7] V. T. A. Ninh, "Fundamental frequencies of bidirectional functionally graded sandwich beams partially supported by foundation using different beam theories," *Transport and Communications Science Journal*, vol. 72, no. 4, pp. 452–467, 2021.
- [8] P. C. Nguyen, "Nonlinear Inelastic Earthquake Analysis of 2D Steel Frames," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6393–6398, Dec. 2020, <https://doi.org/10.48084/etasr.3855>.
- [9] P. C. Nguyen, D. D. Pham, T. T. Tran, and T. Nghia-Nguyen, "Modified Numerical Modeling of Axially Loaded Concrete-Filled Steel Circular-Tube Columns," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7094–7099, Jun. 2021, <https://doi.org/10.48084/etasr.4157>.
- [10] T. D. Hien and P.-C. Nguyen, "Evaluation of Response Variability of Euler–Bernoulli Beam Resting on Foundation Due to Randomness in Elastic Modulus," in *International Conference on Sustainable Civil Engineering and Architecture*, Ho Chi Minh, Vietnam, Oct. 2019, pp. 1087–1092, [https://doi.org/10.1007/978-981-15-5144-4\\_105](https://doi.org/10.1007/978-981-15-5144-4_105).
- [11] T. D. Hien, B. T. Thanh, N. N. Long, N. Van Thuan, and D. T. Hang, "Investigation Into The Response Variability of A Higher-Order Beam Resting on A Foundation Using A Stochastic Finite Element Method," in *5th International Conference on Geotechnics, Civil Engineering Works and Structures*, 2020, pp. 117–122, [https://doi.org/10.1007/978-981-15-0802-8\\_15](https://doi.org/10.1007/978-981-15-0802-8_15).
- [12] P.-C. Nguyen, T. N. Van, and H. T. Duy, "Stochastic Free Vibration Analysis of Beam on Elastic Foundation with the Random Field of Young's Modulus Using Finite Element Method and Monte Carlo Simulation," in *6th International Conference on Geotechnics, Civil Engineering and Structures*, Ha Long, Vietnam, Oct. 2021, pp. 499–506, [https://doi.org/10.1007/978-981-16-7160-9\\_50](https://doi.org/10.1007/978-981-16-7160-9_50).
- [13] N. T. Nguyen, H. D. Ta, T. N. Van, and T. N. Dao, "Stochastic finite element analysis of the free vibration of non-uniform beams with uncertain material," *Journal of Materials and Engineering Structures*, vol. 9, no. 1, pp. 29–37, Apr. 2022.
- [14] D. X. Quy and N. V. Thi, "Static analysis of beam resting on elastic foundation by anisotropic beam-foundation element taking into account non-contact between beam and foundation," *Transport and Communications Science Journal*, vol. 72, no. 5, pp. 552–564, Jun. 2021, <https://doi.org/10.47869/tcsj.72.5.4>.
- [15] Q.-H. Pham, V. K. Tran, T. T. Tran, T. Nguyen-Thoi, P.-C. Nguyen, and V. D. Pham, "A nonlocal quasi-3D theory for thermal free vibration analysis of functionally graded material nanoplates resting on elastic foundation," *Case Studies in Thermal Engineering*, vol. 26, Aug. 2021, Art. no. 101170, <https://doi.org/10.1016/j.csite.2021.101170>.
- [16] T. T. Tran, P.-C. Nguyen, and Q.-H. Pham, "Vibration analysis of FGM plates in thermal environment resting on elastic foundation using ES-MITC3 element and prediction of ANN," *Case Studies in Thermal Engineering*, vol. 24, Apr. 2021, Art. no. 100852, <https://doi.org/10.1016/j.csite.2021.100852>.
- [17] Q.-H. Pham and P.-C. Nguyen, "Dynamic stability analysis of porous functionally graded microplates using a refined isogeometric approach," *Composite Structures*, vol. 284, Mar. 2022, Art. no. 115086, <https://doi.org/10.1016/j.compstruct.2021.115086>.
- [18] J. N. Reddy, *Mechanics of Laminated Composite Plates and Shells: Theory and Analysis*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2003.
- [19] T. D. Hien and B. T. Quang, "Analysis of isotropic rectangular plate resting on non-uniform elastic foundation using Ritz approach," *Materials Today: Proceedings*, vol. 19, pp. 158–160, Jan. 2019, <https://doi.org/10.1016/j.matpr.2019.06.631>.
- [20] X. H. Si, W. X. Lu, and F. L. Chu, "Modal analysis of circular plates with radial side cracks and in contact with water on one side based on the Rayleigh–Ritz method," *Journal of Sound and Vibration*, vol. 331, no. 1, pp. 231–251, Jan. 2012, <https://doi.org/10.1016/j.jsv.2011.08.026>.
- [21] Y. Wang, R. Q. Xu, and H. J. Ding, "Analytical solutions of functionally graded piezoelectric circular plates subjected to axisymmetric loads," *Acta Mechanica*, vol. 215, no. 1, pp. 287–305, Dec. 2010, <https://doi.org/10.1007/s00707-010-0332-7>.
- [22] T.-J. Chen, C.-S. Chen, and C.-W. Chen, "Dynamic response of fiber-reinforced composite plates," *Mechanics of Composite Materials*, vol. 47, no. 5, pp. 549–560, Nov. 2011, <https://doi.org/10.1007/s11029-011-9233-7>.
- [23] M.-H. Huang and D. P. Thambiratnam, "Dynamic Response of Plates on Elastic Foundation to Moving Loads," *Journal of Engineering Mechanics*, vol. 128, no. 9, pp. 1016–1022, Sep. 2002, [https://doi.org/10.1061/\(ASCE\)0733-9399\(2002\)128:9\(1016\)](https://doi.org/10.1061/(ASCE)0733-9399(2002)128:9(1016)).
- [24] T. Jiann-Quo and W. Yung-Ming, "An asymptotic theory for dynamic response of anisotropic inhomogeneous and laminated plates,"

- International Journal of Solids and Structures*, vol. 31, no. 2, pp. 231–246, Jan. 1994, [https://doi.org/10.1016/0020-7683\(94\)90052-3](https://doi.org/10.1016/0020-7683(94)90052-3).
- [25] Z. Kazanci, "Nonlinear transient response of a laminated composite plate under time-dependent pulses," in *4th International Conference on Recent Advances in Space Technologies*, Istanbul, Turkey, Jun. 2009, pp. 125–130, <https://doi.org/10.1109/RAST.2009.5158181>.
- [26] D. Sun and S.-N. Luo, "The wave propagation and dynamic response of rectangular functionally graded material plates with completed clamped supports under impulse load," *European Journal of Mechanics - A/Solids*, vol. 30, no. 3, pp. 396–408, May 2011, <https://doi.org/10.1016/j.euromechsol.2011.01.001>.
- [27] M. Meunier and R. A. Shenoi, "Dynamic analysis of composite sandwich plates with damping modelled using high-order shear deformation theory," *Composite Structures*, vol. 54, no. 2, pp. 243–254, Nov. 2001, [https://doi.org/10.1016/S0263-8223\(01\)00094-0](https://doi.org/10.1016/S0263-8223(01)00094-0).
- [28] Y.-H. Chen, Y.-H. Huang, and C.-T. Shih, "Response of an infinite timoshenko beam on a viscoelastic foundation to a harmonic moving load," *Journal of Sound and Vibration*, vol. 241, no. 5, pp. 809–824, Apr. 2001, <https://doi.org/10.1006/jsvi.2000.3333>.
- [29] N. V. Thuan and T. D. Hien, "Stochastic Perturbation-Based Finite Element for Free Vibration of Functionally Graded Beams with an Uncertain Elastic Modulus," *Mechanics of Composite Materials*, vol. 56, no. 4, pp. 485–496, Sep. 2020, <https://doi.org/10.1007/s11029-020-09897-z>.
- [30] R. P. Shimpi and H. G. Patel, "A two variable refined plate theory for orthotropic plate analysis," *International Journal of Solids and Structures*, vol. 43, no. 22, pp. 6783–6799, Nov. 2006, <https://doi.org/10.1016/j.ijsolstr.2006.02.007>.
- [31] A. Hasani Baferani, A. R. Saidi, and H. Ehteshami, "Accurate solution for free vibration analysis of functionally graded thick rectangular plates resting on elastic foundation," *Composite Structures*, vol. 93, no. 7, pp. 1842–1853, Jun. 2011, <https://doi.org/10.1016/j.compstruct.2011.01.020>.

# The Combined Effect of Current Boosting and Power Loss on Photovoltaic Arrays under Partial Shading Conditions

**Abdullahi Abubakar Mas'ud**

Department of Electrical Engineering, Jubail Industrial College, Saudi Arabia  
masud\_a@rcjy.edu.sa  
(corresponding author)

Received: 24 September 2022 | Revised: 23 October 2022 | Accepted: 25 October 2022

## ABSTRACT

This study proposes a novel technique for improving the performance of photovoltaic (PV) arrays under Partial Shading Conditions (PSCs). A 4×4 solar PV array with 16 panels was considered. Bridge-Linked (BL), Total Cross-Tied (TCT), Honey Comp (HC), One Cross-Link (OCL), and Two Cross-Link (TCL) were among the topologies of interest. First, the combined effect of connecting switches and partial shading on the PV array was studied. Then, the power loss/gain caused by reconfiguring the PV array structure from Series-Parallel (SP) to other schemes was investigated. Finally, a method of boosting current into the PV array is proposed to reduce PSCs-related power losses in the connecting switches. The results show that the number of connecting switches in the topology plays an important role in determining power gain or loss at different partial shading levels. TCT and HC outperformed the others in terms of power improvement when PSCs were considered without current boosting. This is true for different levels of solar irradiation exposure. The SP topology is optimal when the solar irradiation level is greater than 900W/m<sup>2</sup> or less than 200W/m<sup>2</sup>. TCT outperformed the others when the current was boosted in the PV array, with a power improvement of 108%, for certain PSCs.

**Keywords:** partial shading conditions; photovoltaic; total cross-tied; bridge coupled; honeycomb

## I. INTRODUCTION

It is widely accepted that the global electricity demand is increasing rapidly, but the amount of fossil fuels available for power production is limited [1, 2]. On the other hand, fossil fuel emissions contribute to environmental pollution [3, 4]. As a result, renewable energy sources for power generation are widely advocated. Among these, solar photovoltaic (PV) technology has advanced significantly during the recent years [5, 6]. So far, it is one of the most imaginative developments, with the highest probability of dominating the Renewable Energy (RE) market [7]. Because of their low maintenance, zero pollution, and low-noise characteristics, PV systems have a wide range of applications, including solar vehicles, streetlights, and other hybrid energy systems [8, 9]. On the other hand, the output characteristics of PV panels are primarily determined by temperature and irradiance, and they are frequently unpredictable [9, 11]. When PV modules experience partial shading, they are not exposed to the same amount of irradiance. This results in inconsistent module characteristics [11]. Partial Shading Conditions (PSCs) occur when some modules in a PV array are shaded by trees and buildings and this results in the reduction of the voltage and current of the afflicted panel [12-15]. The panel that generates the least amount of electricity will act as an electrical load drawing energy from the system's other modules [14, 15],

resulting in a decrease in the overall power output. Partial shading can be reduced by selecting appropriate PV topologies or physically reconfiguring PV modules, as well as employing Maximum Power Point Tracking (MPPT) technologies. The most common configurations for PV arrays are honeycomb (HC), Series-Parallel (SP), Total Cross-Tied (TCT), and Bridge-Link (BL) [20-21]. Under various PSCs, the TCT has the lowest mismatch losses and the highest power output. In [20], it has been shown that the TCT configuration does not deliver the maximum power possible when operating in partial shading and the extracted output power can be further increased [20].

Several researchers have investigated different techniques to reduce mismatching power losses in PV arrays [5, 17, 23-33]. Authors in [27] compared large interconnected SP array schemes operating under PSCs [27]. These combinations were modeled and analyzed in MATLAB/Simulink, and the results were validated experimentally. The result indicates that the Global Maximum Power point (GMPP) depends on the shading pattern and the architecture of the PV array. Authors in [22] developed a modified bypass circuit that effectively finds a compromise between the increased reliability and the increased complexity of the PV array. The results were validated experimentally. The bypass circuit worked well in lowering the hot spot temperature of the cryptic cell below the average

temperature of the module without significantly increasing the complexity of the system. Genetic Algorithm [33], Particle Swarm Optimization [34], Mine Blast Optimization [35], Cuckoo Search [36], and Ant Colony Optimization [34] are other examples of optimization techniques deployed for minimizing losses under PSCs.

Recently, authors in [23] proposed an improved technique to enhance the performance of the PV array technologies (PVAT) under PSCs. The study takes into account 8 shading patterns. To combat partial shading, 6 novel PVATs based on specific array configurations were presented. The simulation results were compared to those obtained with conventional PVATs, and the best topologies for different PSCs are determined. The TCT typically achieves maximum power extraction from PSCs. It also enhanced PVAT output power for short and wide PSC by 105%. The performance of the TCT was enhanced in another investigation employing the minimizing losses technique [25]. Different arrangements were introduced by reducing the number of connecting switches. The PV module configuration was altered to generate 5% more power in PSCs. If 50% or more of the area was partially shadowed, no reconfiguration was required. When the darkened region was less than 50%, reconfiguration significantly increased power.

Many studies have developed PV reconfiguration strategies that can reduce the impact of PSCs on PV arrays. However, one aspect that has yet to be investigated is the way PV arrays perform under PSCs, when both current boosting and switch losses are considered. A PV array reconfiguration, on the other hand, necessitates an increase in the number of switches, resulting in increased switching power loss and the need for more current to be boosted during PSCs. Therefore, in this paper, a novel strategy for minimizing losses in PV arrays under PSCs is presented. The approach involves boosting the current in the partially shaded sections of the PV array. To obtain a more robust result, current losses in connecting switches are also considered. A 4×4 PV array with 4 shading patterns was investigated, and all simulations were performed in MATLAB/SIMULINK. Configurations of interest include SP, TCT, HC, BL, OCL, and TCL.

## II. MATERIALS AND METHODS

### A. Modeling of the PV Array

The PV cell is the most important part of a PV system because it converts energy from the sun to electricity [38, 40]. The amount of electricity produced by a PV module or cell is primarily determined by temperature and solar irradiation. Because of its higher accuracy in obtaining the shunt and series resistances, the single diode model of a PV cell has been used in PV modeling [28, 42]. A PV module is a grouping of PV cells connected in series, whereas a PV array is a grouping of PV modules connected in series or parallel, as well as a slew of others [21, 43]. Equation (1) depicts the mathematical representation of the current-voltage (I-V) characteristics for an ideal PV cell model with a terminal current [29].

$$I = I_{PV} - I_0 \left[ \exp \left( \frac{q(V + R_s)}{A k_B T} \right) - 1 \right] - \frac{V + R_{SH}}{R_{SH}} \quad (1)$$

where  $I_0$  represents the saturation current of the diode,  $I_{PV}$  represents the current of the PV,  $V$  represents the PV module's terminal voltage,  $A$  represents the identity factor,  $k_B$  is the Boltzmann's constant,  $q$  represents the electronic charge,  $R_s$  represents the series resistance,  $T$  represents the junction temperature, and  $R_{SH}$  represents the shunt resistance.

The 4×4 PV module is designed in MATLAB/SIMULINK using a single diode PV model. Equation (2) describes the I-V characteristics of a PV array composed of  $N_p$  parallel and  $N_s$  series-connected modules [41].

$$I = I_{PV} N_p - I_r N_p \left[ \exp \left( \frac{q \left( V + R_s \frac{N_s}{N_p} \right)}{A k_B T} \right) - 1 \right] - \frac{V + R_s \frac{N_s}{N_p}}{R_{SH} \frac{N_s}{N_p}} \quad (2)$$

where  $I$  represents the diode reverse leakage current. The parameters of the 25W Blue solar PV module that was used to model the PV array schemes are listed in Table I.

TABLE I. PARAMETERS FOR THE 25W SOLARI PV MODULE

S/No	Parameters	Values
1	Maximum power of the PV	25W
2	Short circuit current of the PV ( $I_{sc}$ )	1.45A
3	Open circuit voltage of the PV ( $V_{oc}$ )	22.2V
4	Current at the maximum power	1.37A
5	The voltage at the maximum power	18.2V
6	The temperature coefficient of $V_{oc}$	-0.34%/°C
7	Cells per module	40
8	Cells per module	40
9	Temperature coefficient of $I_{sc}$	0.041%/°C
10	$R_{SH}$	1114.063Ω
11	$R_s$	0.9635Ω

### B. Proposed PV Module Configuration

To reduce the partial shading impact, PV array arrangement is critical. A proper array arrangement must be adopted to produce the necessary power output improvement from the PV array system. Some of the most commonly used PV array configurations in the literature are SP, TCT, BL, and HC [42]. Among them, the SP (see Figure 1(a)) is the most widely used. Under the PSCs, the SP scheme is the most effective against mismatch power losses in series strings. To minimize mismatch power losses in series strings, alternative tie connections are inserted between the parallel strings of a SP scheme. They are designed in the BL, HC, and TCT configurations. Figure 1(b) is the BL scheme. The PV modules resemble a bridge rectifier arrangement. One drawback of the BL arrangement is that it performs poorly under PSCs. The TCT configuration is shown in Figure 1(d). To design the TCT arrangement, cross ties must be connected between each module in the SP scheme. These cross-links connect the modules in the array arrangement.



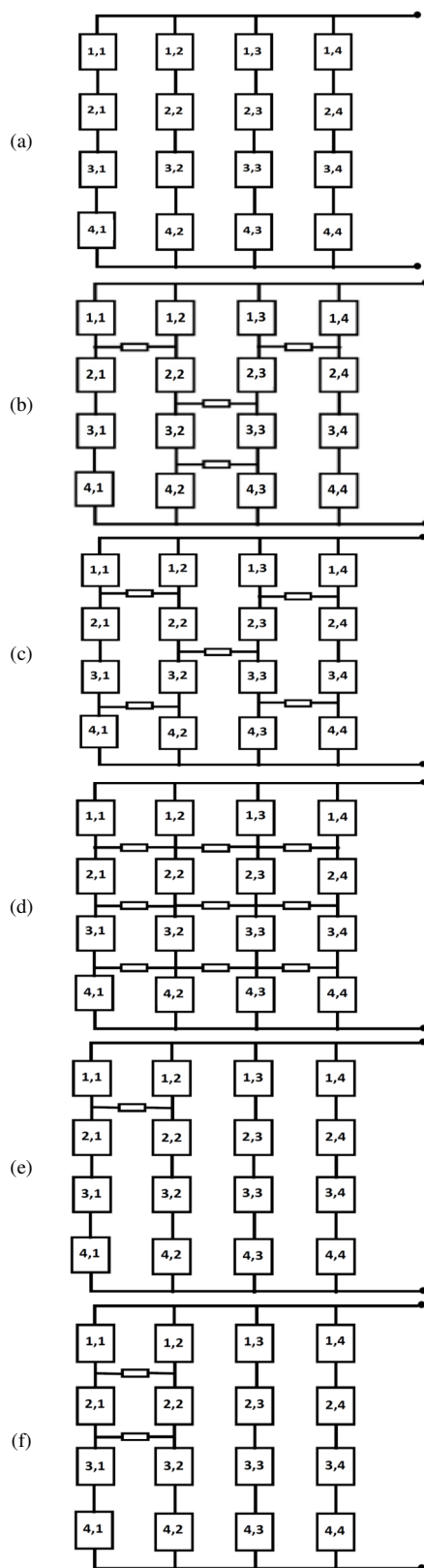


Fig. 1. PV array schemes considered: (a) BP, (b) BL, (c) HC, (d) TCT, (e) OCL, (f) TCL.

The cross ties help reduce mismatch losses in the shaded module's series string by providing an additional current channel. TCT, on the other hand, suffers from high initial installation costs and cable losses due to the extensive interconnections. To reduce these losses, some interconnections can be disabled and set to either HC or BL. In this study, two additional schemes will be considered, i.e. one and two crosslinks in the SP configurations (see Figure 1(e)-(f)). Previous research did not consider non-ideal switches between crosslinks, particularly during PV shading mitigation using the current boosting method [23]. Due to the fact that these switches can result in additional power losses during PSCs, the purpose of this study is to examine the possibility of incorporating crosslink switching into the current boosting procedure. In this instance, a more reliable result will be obtained.

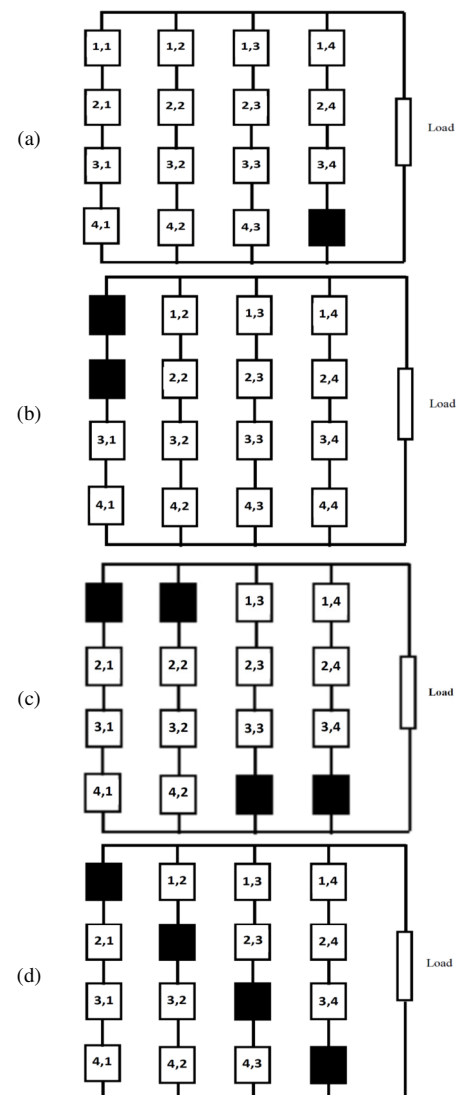


Fig. 2. Shading patterns on the PV scheme. Cases (a) 1, (b) 2, (c) 3, (d) 4.

### C. Proposed Partial Shading Schemes

When PSCs occur, portions of the PV array will not receive the required level of irradiance. As a result, the shading scenario is unpredictable. As shown in Figure 2, four shading conditions will be considered to analyze the partial shading effect on the 4x4 PV modules:

- Case 1: A single PV module is shaded.
- Case 2: Two PV modules are shaded.
- Case 3: Four PV modules are shaded
- Case 4: Four PV modules are shaded (diagonally).

As previously stated, this study investigates the effects of shading on PV modules at 40°C with varying amounts of irradiance. Under the same conditions, when the PV array is uniformly shaded, the irradiance value of all 16 modules is exactly 1000W/m<sup>2</sup>. Solar radiation reduction from 900W/m<sup>2</sup> to 100W/m<sup>2</sup> will be considered for all PSCs. Each model ran with a load of 13.2Ω.

### III. CURRENT BOOSTING PROCEDURE

A current boosting strategy has been proposed in [14] which improves PV power during PSCs. However, the current boosting procedure for different schemes has not been well investigated. A more in-depth analysis will be provided in this study. In this study, the current is directly boosted across the partially shaded panel, with a bypass diode connected across all PV modules to provide a low resistance path for the flow of any current, thereby improving the GMPP. Although using bypass diodes increases the cost and complexity of any PV array, it can significantly improve power during PSCs. Previous research has shown that under PSCs and with bypass diodes, the GMPP for SP, TCT, HC, and BL PV topologies is 21.54%, 52.74%, 16.09%, and 20.78% higher than what it would be in a scheme without bypass diodes [32]. The bypass diodes across a PV panel can cause peaks in the P-V characteristic curve but are not detrimental to the system.

In this study, the bypass diode was utilized. Figure 3 depicts the current boosting schemes for each of the considered topologies. Equation (3) is a generalized equation that shows the current produced by a PV module in a TCT architecture [14].

$$I_{ni} = \frac{G_{ni}}{G_{STC}} \times I_g, \quad 1 < i < d \quad (3)$$

where  $n$  denotes the number of rows and  $d$  the number of PV panels.  $G_{STC}$  denotes irradiation at standard temperature conditions, whereas  $G_{ni}$  denotes irradiation received by a PV panel.  $I_g$  denotes the maximum current generated by a PV panel under STCs. The total current generated in each row is the sum of the currents generated in each PV module in that row. The mismatch current in the TCT topology is the difference between the total current under PSCs and the peak value of the current without PSCs. The mismatch current will be boosted across the specific row to improve power during the current boosting procedure. The current is boosted across the shaded rows on the TCT topology at every instant, based on the partial

shading situation, as shown in Figure 3. The boosted current scheme cannot be implemented in the other topologies shown in Figure 3 (BL, HC, OCL, and TCL). In this case, regardless of the shaded PV panel in the topology, the current will only be boosted at specific locations. This is proposed to comply with circuit theory rules and to avoid boosting current in places where it is not required. Because there are fewer interlink connections between neighboring strings in these topologies than in the conventional TCT topology, mismatch losses are greater. However, the proposed scheme will only provide the best PV or I-V characteristic curve in TCT configuration.

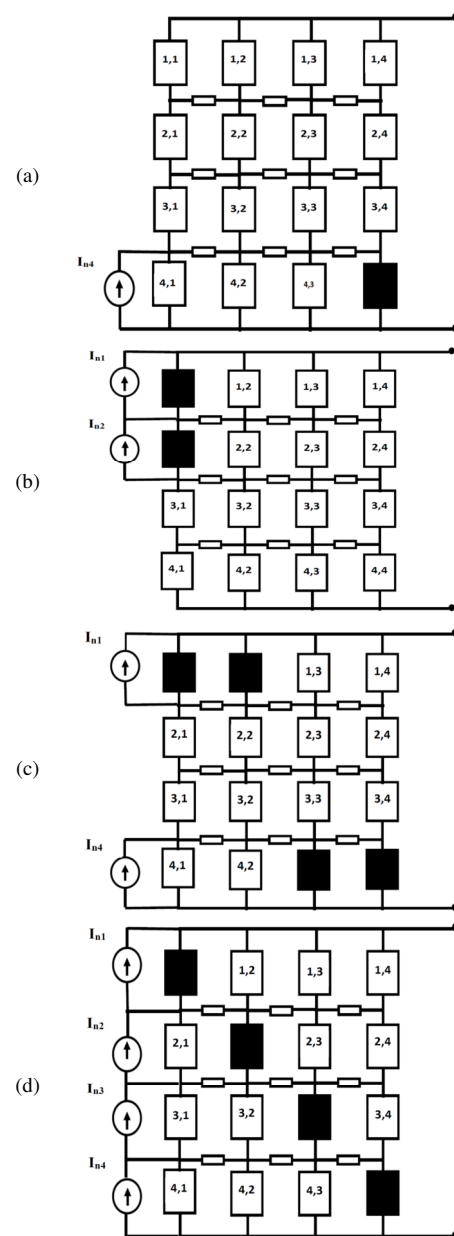


Fig. 3. Current boosting scheme for a TCT configuration under PSCs.

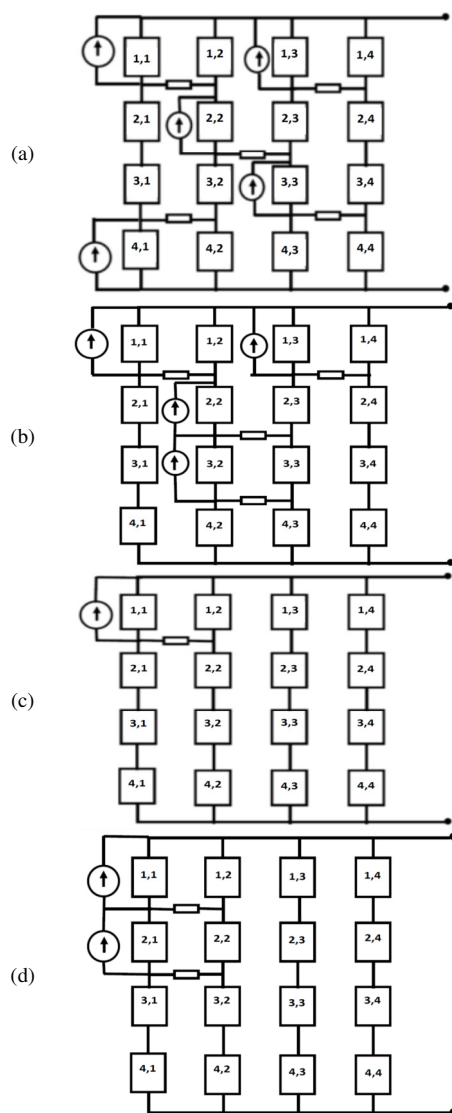


Fig. 4. Current boosting schemes with connecting switches under PSCs: (a) HC, (b) BL, (c) OCL, (d) TCL.

#### IV. RESULT AND DISCUSSIONS

MATLAB/Simulink was utilized to simulate a 4×4 PV array with 5 different configurations (i.e. SP, BL, HC, TCT, OCL, and TCL) and 4 different shading patterns (cases 1-4). First, the combined effect of connecting switches and partial shading on the power output of 5 PV topologies (BL, HC, TCT, OCL, and TCL) was investigated. Then, the power loss/gain caused by switching from SP to the aforementioned topologies was investigated. A current boosting technique was subsequently used to see if the performance of the PV array had improved. The ultimate goal was to use as few switches as possible in order to minimize switch loss. When partial shading occurs, this approach may be able to produce the best architecture in terms of peak output power. Table II shows a comparison of the proposed and existing PV topologies under 4 distinct PSCs. In contrast to the current topology, the proposed method employs the boosting current strategy.

TABLE II. THE OUTPUT POWER IMPROVEMENT OF ALL SCHEMES UNDER VARIOUS PSCS

Switch combinations	PV shading configuration	$G_{MPP}(W)$		Total boosted current (A)	Power Improvement (W)
		Existing scheme	Proposed scheme		
All switches closed (TCT)	Case 1	266.72	350.21	0.79	84.49
	Case 2	263.9	344.13	1.64	80.23
	Case 3	132.11	370.2	5.528	228.09
	Case 4	259.13	337.32	3.4	78.19
Switches 1,3,5,7,9 closed (HC)	Case 1	266.44	344.45	0.748	78.01
	Case 2	262.50	337.88	1.562	75.38
	Case 3	139.23	362.0	5	223
	Case 4	261.33	335	3.164	73.67
Switches 1,5,6,7 closed (BL)	Case 1	265.99	336.4	0.75	70.41
	Case 2	263.53	337.60	1.6	74.07
	Case 3	208.57	356.70	4.432	148.13
	Case 4	219.76	323	3.68	103.24
Switches 1,2 closed (OCL)	Case 1	274.65	330.33	0.69	55.68
	Case 2	264.87	340.989	0.77	76.12
	Case 3	228.20	326.50	1.1	98.3
	Case 4	212.32	360.17	1.254	147.85
Switch 1 closed (TCL)	Case 1	275.11	328.39	0.66	53.28
	Case 2	263.91	338.49	0.79	74.6
	Case 3	228.75	321.93	1.09	93.18
	Case 4	212.611	357.77	1.16	145.16

First, the shaded PV module is assumed to have a solar irradiation level of  $100W/m^2$ , implying very low PV shading. When one of the 16 PV panels is shaded, the HC, TCT, and BL configurations suffer more losses than the others, with the TCL arrangement appears to be the most efficient because the shaded PV module is at the end of the topology and the TCT has more interconnecting switches than the OCL and TCL. Furthermore, in Case 1, the more connecting switches there are, the greater the power loss. For Case 2, the power loss under PSCs appears to be nearly identical for all topologies, with TCL appearing to be the optimum. Because the shaded panels are in the upper left quadrant of the array, close to the two connecting switches, current may flow in a different direction during partial shading. Case 3 appears to favor OCL and TCL, though BL outperforms TCT and HC. This is an indication that the number of connecting switches increased the power loss of the complete system. This conclusion, however, does not apply to all PV shading situations. The TCT and HC appear to perform best in Case 4, where the shading appears in a diagonal form, demonstrating that the scheme's performance improves with more connecting switches. In general, it appears that the shading pattern in Case 3 causes more power loss than the others.

Figures 5–8 show the percentage of power loss or gain when converting from SP to HC, BL, TCT, OCL, and TCL configurations. Solar radiation is varied between 900 and  $100W/m^2$ , and each shading pattern is investigated. It should be noted that the percentage of power loss/gain for BP is zero in all the Figures because it was assumed to be the reference configuration. TCT is the best configuration when switching from BP to the others in Case 1 (Figure 5), followed by HC. At  $500W/m^2$ , TCT and HC improve power by 3% and 1.8%, respectively.

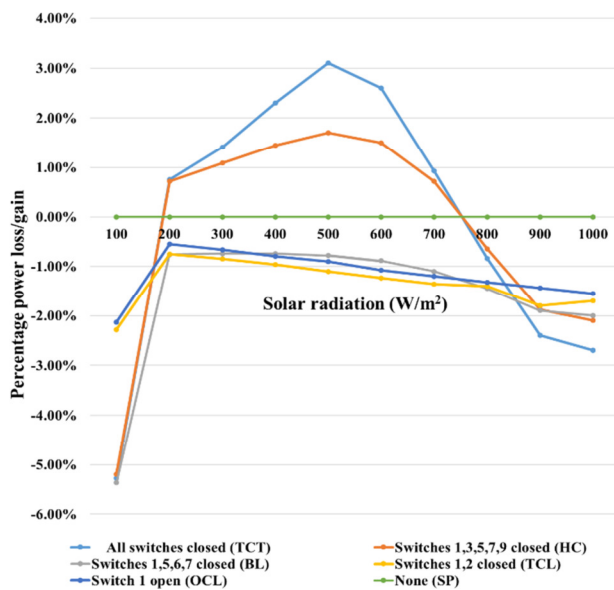


Fig. 5. The percentage of power loss/gain when switching from SP to the other configurations in Case 1.

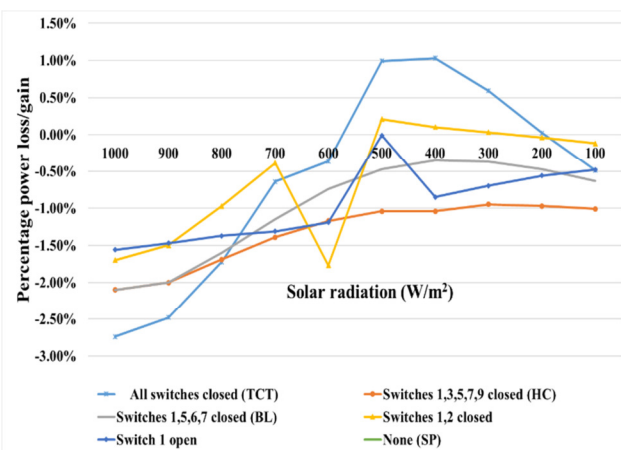


Fig. 6. The percentage of power loss/gain when switching from SP to the other configurations in Case 2.

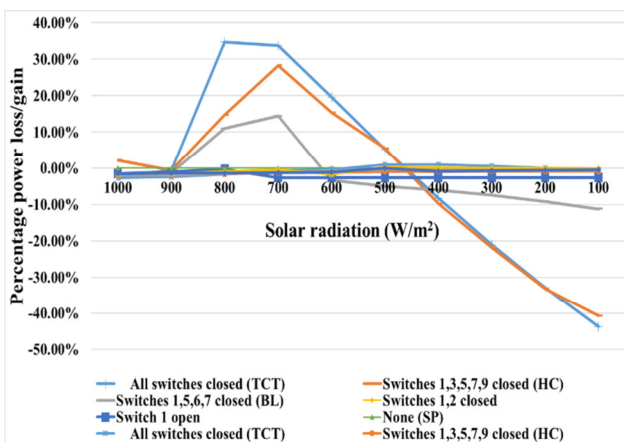


Fig. 7. The percentage of power loss/gain when switching from SP to the other configurations in Case 3.

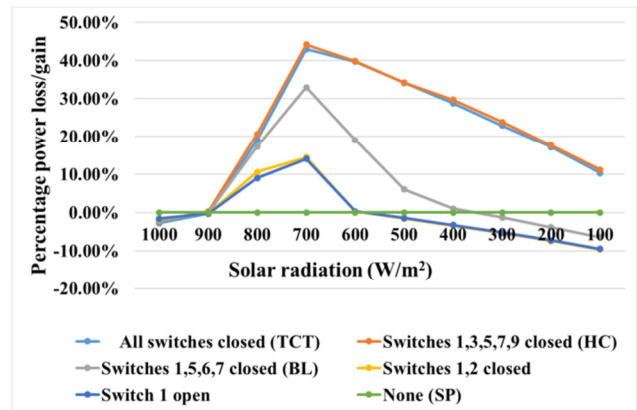


Fig. 8. The percentage of power loss/gain when switching from SP to the other configurations in Case 4.

The solar radiation in the shaded module for Case 1 should be between 200 and 700W/m<sup>2</sup> for any power improvement in the TCT and HC. TCT is the best configuration for Case 2 (Figure 6), followed by TCL. Power improvements of 1% have been made. In Case 3 (Figure 7), when reconfiguring from BP to TCT, the TCT appears to outperform the others. In this case, however, the HC performs nearly as well as the TCT. This performance is only noticeable when the solar radiation on the shaded module exceeds 500W/m<sup>2</sup>. However, reconfiguration from BP to the other schemes below 500W/m<sup>2</sup> results in a power loss of up to 50% for TCT and HC and 10% for the BL. As a result, when reconfiguring PV modules under PSCs, care must be taken because significant power losses are possible and thus detrimental at certain solar irradiation levels. In Case 4, the TCL and HC appear to outperform the others at nearly all irradiation levels. As a result, it appears that this is an appropriate scheme for this type of partial shading. It is important to note that the TCT can achieve maximum power improvements of 3%, 1%, 35%, and 42% for Case 1, Case 2, Case 3, and Case 4 shading configurations, respectively.

When the boosted current values in Table II are examined, it is clear that the TCT receives more boosted current than the others for all of the PV shading schemes considered. This is because a current is boosted in every row with mismatching power in the TCT, whereas in the other cases, a current is boosted in specific rows regardless of shading location. Despite the fact that the GMPP value will improve, one major disadvantage of non-uniform current boosting is the formation of numerous peaks in the characteristic curves. TCT appears to be the best strategy in terms of power improvement when considering the proposed technique (see Figure 9). Case 3 shading pattern can improve power by up to 228W. The TCT has more interconnecting switches than the others, and the mismatch current is enhanced in each row. Figure 9 shows the proposed and existing P-V and I-V curves for TCT in each of the 4 partial shading scenarios. The P-V curve of the proposed technique has fewer peaks than the existing technique, with power improvements of 31%, 30%, 108%, and 30.1% for Case 1, Case 2, Case 3, and Case 4, respectively. For Cases 3 and 4, the proposed technique has one peak, whereas the existing technique has two peaks.

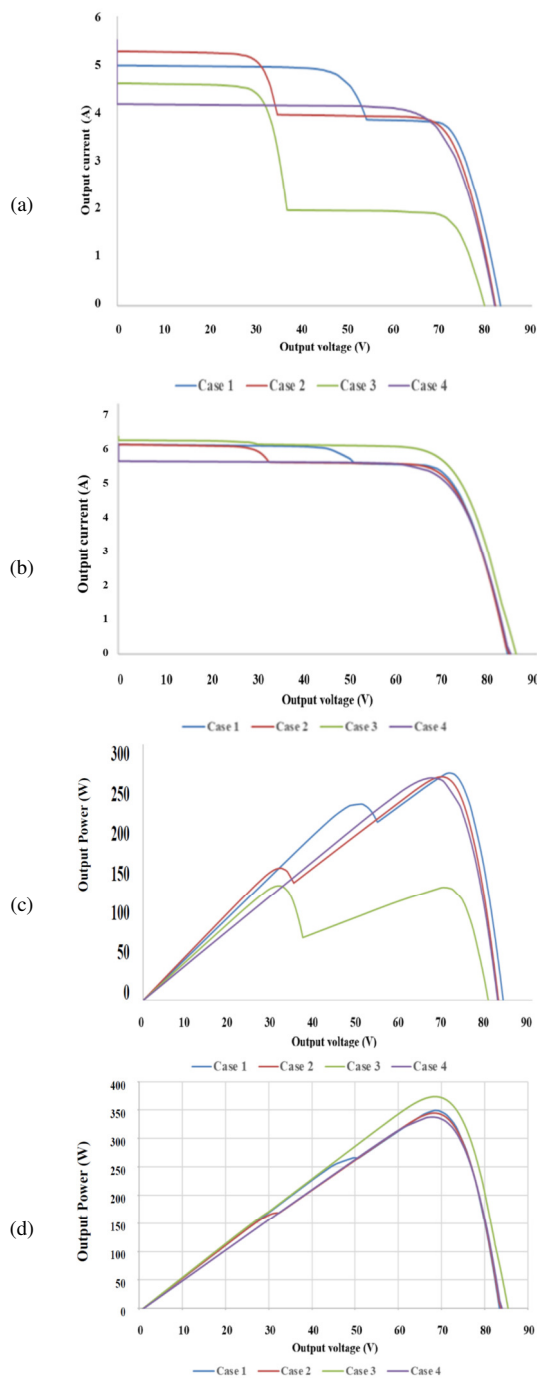


Fig. 9. I-V curve of the (a) existing PV scheme, (b) proposed TCT scheme. P-V curve of the (c) existing PV scheme, (d) proposed P-V scheme.

Despite a significant power improvement, the number of peaks in Cases 1 and 2 is the same for both the existing and proposed techniques. In this study, the P-V and I-V curves for the other switching configurations (BL, HC, OCL, and TCL) were not shown, because the number of peaks does not change even with current boosting, and only power improvement is visible. The current boosting scheme, as described in the literature, cannot be implemented in the SP configuration [14].

According to the results of this study, reconfiguration of the BP scheme under PSCs may not be required for certain solar irradiation levels. Furthermore, current boosting will result in power gains relative to any of the commonly used PV topologies, i.e. TCT, HC, BL, OCL, and TCT, and the power loss in the connecting switches plays a significant role in the amount of current boosted during PSCs.

## V. CONCLUSION

In this study, the effect of current boosting and power loss caused by connecting switches on the I-V and P-V characteristic curves of a PV array were investigated. MATLAB/Simulink was utilized to simulate 4x4 PV array under various degrees of shade and shading patterns. The shading patterns of 4, 2, and 1 in 16 PV panels were investigated. Power loss or gain as a result of reconfiguring PV panels from the BP to other configurations such as BL, TCT, HC, OCL, and TCL has been studied in depth. The level of solar irradiation at which the PV array could be reconfigured was also evaluated. The results show that the number of connecting switches in the topology is an important factor in determining power gain/loss at various shading levels. The more connecting switches there are, the more power is lost, and more current is required to compensate for the power mismatch. TCT and HC appear to outperform the others in terms of power improvement for the majority of the partial shading scenarios considered without current boosting. This is only true for certain solar irradiation levels. It has also been shown that, under certain shading patterns (case 4, diagonal shading), up to 50% of the power is lost when compared to the BP scheme. In the case of current boosting, TCT showed greater power improvement than others because, in TCT, current is boosted in every row with power mismatch, whereas in the other cases, current is boosted in specific rows regardless of the shading position. For a specific shading pattern, the proposed technique improved power by up to 108%. The results show that for the majority of the shading patterns considered, the BP scheme is the best for solar irradiation level higher than  $900\text{W/m}^2$  or lower than  $200\text{W/m}^2$ .

## ACKNOWLEDGMENT

The author wishes to acknowledge the support given by Jubail Industrial College during this research.

## REFERENCES

- [1] K. Chen, S. Tian, Y. Cheng, and L. Bai, "An Improved MPPT Controller for Photovoltaic System Under Partial Shading Condition," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 3, pp. 978–985, Jul. 2014, <https://doi.org/10.1109/TSTE.2014.2315653>.
- [2] D. K. Dhaked, Y. Gopal, and D. Birla, "Battery Charging Optimization of Solar Energy based Telecom Sites in India," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 5041–5046, Dec. 2019, <https://doi.org/10.48084/etasr.3121>.
- [3] J. Ma, D. Hong, K. Wang, Z. Bi, X. Zhu, and J. Zhang, "Analytical modeling and parameter estimation of photovoltaic strings under partial shading conditions," *Solar Energy Materials and Solar Cells*, vol. 235, Jan. 2022, Art. no. 111494, <https://doi.org/10.1016/j.solmat.2021.111494>.
- [4] O. Bingöl and B. Özkaya, "Analysis and comparison of different PV array configurations under partial shading conditions," *Solar Energy*, vol. 160, pp. 336–343, Jan. 2018, <https://doi.org/10.1016/j.solener.2017.12.004>.



- [5] N. Belhaouas *et al.*, "A new approach of PV system structure to enhance performance of PV generator under partial shading effect," *Journal of Cleaner Production*, vol. 317, Oct. 2021, Art. no. 128349, <https://doi.org/10.1016/j.jclepro.2021.128349>.
- [6] D. K. Dhaked and D. Birla, "Microgrid Designing for Electrical Two-Wheeler Charging Station Supported by Solar PV and Fuel Cell," *Indian Journal of Science and Technology*, vol. 14, no. 30, pp. 2517–2525, Sep. 2021, <https://doi.org/10.17485/IJST/v14i30.224>.
- [7] D. K. Dhaked and D. Birla, "Modeling and control of a solar-thermal dish-stirling coupled PMDC generator and battery based DC microgrid in the framework of the ENERGY NEXUS," *Energy Nexus*, vol. 5, Mar. 2022, Art. no. 100048, <https://doi.org/10.1016/j.nexus.2022.100048>.
- [8] H. Rezk, I. Tyukhov, and A. Raupov, "Experimental implementation of meteorological data and photovoltaic solar radiation monitoring system," *International Transactions on Electrical Energy Systems*, vol. 25, no. 12, pp. 3573–3585, 2015, <https://doi.org/10.1002/etep.2053>.
- [9] H. Rezk, I. Tyukhov, M. Al-Dhaifallah, and A. Tikhonov, "Performance of data acquisition system for monitoring PV system parameters," *Measurement*, vol. 104, pp. 204–211, Jul. 2017, <https://doi.org/10.1016/j.measurement.2017.02.050>.
- [10] M. H. Zafar *et al.*, "Group Teaching Optimization Algorithm Based MPPT Control of PV Systems under Partial Shading and Complex Partial Shading," *Electronics*, vol. 9, no. 11, Nov. 2020, Art. no. 1962, <https://doi.org/10.3390/electronics9111962>.
- [11] G. Sai Krishna and T. Moger, "Improved SuDoKu reconfiguration technique for total-cross-tied PV array to enhance maximum power under partial shading conditions," *Renewable and Sustainable Energy Reviews*, vol. 109, pp. 333–348, Jul. 2019, <https://doi.org/10.1016/j.rser.2019.04.037>.
- [12] Y. Kassem, H. Camur, and O. a. M. Abughinda, "Solar Energy Potential and Feasibility Study of a 10MW Grid-connected Solar Plant in Libya," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5358–5366, Aug. 2020, <https://doi.org/10.48084/etasr.3607>.
- [13] F. Belhachat and C. Larbes, "PV array reconfiguration techniques for maximum power optimization under partial shading conditions: A review," *Solar Energy*, vol. 230, pp. 558–582, Dec. 2021, <https://doi.org/10.1016/j.solener.2021.09.089>.
- [14] J. Ahmed and Z. Salam, "A critical evaluation on maximum power point tracking methods for partial shading in PV systems," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 933–953, Jul. 2015, <https://doi.org/10.1016/j.rser.2015.03.080>.
- [15] N. Priyadarshi *et al.*, "Performance Evaluation of Solar-PV-Based Non-Isolated Switched-Inductor and Switched-Capacitor High-Step-Up Cuk Converter," *Electronics*, vol. 11, no. 9, Jan. 2022, Art. no. 1381, <https://doi.org/10.3390/electronics11091381>.
- [16] R. K. Pachauri *et al.*, "Impact of Partial Shading on Various PV Array Configurations and Different Modeling Approaches: A Comprehensive Review," *IEEE Access*, vol. 8, pp. 181375–181403, 2020, <https://doi.org/10.1109/ACCESS.2020.3028473>.
- [17] B. Yang *et al.*, "Comprehensive overview of maximum power point tracking algorithms of PV systems under partial shading condition," *Journal of Cleaner Production*, vol. 268, Sep. 2020, Art. no. 121983, <https://doi.org/10.1016/j.jclepro.2020.121983>.
- [18] H. Braun *et al.*, "Topology reconfiguration for optimization of photovoltaic array output," *Sustainable Energy, Grids and Networks*, vol. 6, pp. 58–69, Jun. 2016, <https://doi.org/10.1016/j.segan.2016.01.003>.
- [19] G. Sai Krishna and T. Moger, "Reconfiguration strategies for reducing partial shading effects in photovoltaic arrays: State of the art," *Solar Energy*, vol. 182, pp. 429–452, Apr. 2019, <https://doi.org/10.1016/j.solener.2019.02.057>.
- [20] F. Belhachat and C. Larbes, "Modeling, analysis and comparison of solar photovoltaic array configurations under partial shading conditions," *Solar Energy*, vol. 120, pp. 399–418, Oct. 2015, <https://doi.org/10.1016/j.solener.2015.07.039>.
- [21] S. Sugumar, D. Prince Winston, and M. Pravin, "A novel on-time partial shading detection technique for electrical reconfiguration in solar PV system," *Solar Energy*, vol. 225, pp. 1009–1025, Sep. 2021, <https://doi.org/10.1016/j.solener.2021.07.069>.
- [22] S. Ghosh, V. K. Yadav, and V. Mukherjee, "Improvement of partial shading resilience of PV array through modified bypass arrangement," *Renewable Energy*, vol. 143, pp. 1079–1093, Dec. 2019, <https://doi.org/10.1016/j.renene.2019.05.062>.
- [23] D. Prince Winston, S. Kumaravel, B. Praveen Kumar, and S. Devakirubakaran, "Performance improvement of solar PV array topologies during various partial shading conditions," *Solar Energy*, vol. 196, pp. 228–242, Jan. 2020, <https://doi.org/10.1016/j.solener.2019.12.007>.
- [24] C. Saiprakash, A. Mohapatra, B. Nayak, and S. R. Ghatak, "Analysis of partial shading effect on energy output of different solar PV array configurations," *Materials Today: Proceedings*, vol. 39, pp. 1905–1909, Jan. 2021, <https://doi.org/10.1016/j.matpr.2020.08.307>.
- [25] C. Tubniyom, R. Chatthaworn, A. Suksri, and T. Wongwuttanasatian, "Minimization of losses in solar photovoltaic modules by reconfiguration under various patterns of partial shading," *Energies*, vol. 12, no. 1, Jan. 2019, <https://doi.org/10.3390/EN12010024>.
- [26] V. Balaraju and C. Chengaiah, "A Comprehensive Study on Re-arrangement of Modules Based TCT Configurations of Partial Shaded PV Array with Shade Dispersion Method," *Trends in Renewable Energy*, vol. 6, no. 1, pp. 37–60, Feb. 2020, <https://doi.org/10.17737/tre.2020.6.1.00111>.
- [27] H. Patel and V. Agarwal, "MATLAB-Based Modeling to Study the Effects of Partial Shading on PV Array Characteristics," *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 302–310, Mar. 2008, <https://doi.org/10.1109/TEC.2007.914308>.
- [28] B. Aljafari, P. R. Satpathy, and S. B. Thanikanti, "Partial shading mitigation in PV arrays through dragonfly algorithm based dynamic reconfiguration," *Energy*, vol. 257, Oct. 2022, Art. no. 124795, <https://doi.org/10.1016/j.energy.2022.124795>.
- [29] P. R. Satpathy, B. Aljafari, and S. B. Thanikanti, "Power losses mitigation through electrical reconfiguration in partial shading prone solar PV arrays," *Optik*, vol. 259, Jun. 2022, Art. new. 168973, <https://doi.org/10.1016/J.IJLEO.2022.168973>.
- [30] E. V. Paraskevadaki and S. A. Papathanassiou, "Evaluation of MPP Voltage and Power of mc-Si PV Modules in Partial Shading Conditions," *IEEE Transactions on Energy Conversion*, vol. 26, no. 3, pp. 923–932, Sep. 2011, <https://doi.org/10.1109/TEC.2011.2126021>.
- [31] I. Faye, A. Ndiaye, D. Kobor, M. Thiame, C. Sene, and L.-G. Ndiaye, "Evaluation of the impact of partial shading and its transmittance on the performance of crystalline silicon photovoltaic modules," *International Journal of Physical Sciences*, vol. 12, no. 21, pp. 286–294, Nov. 2017, <https://doi.org/10.5897/IJPS2017.4666>.
- [32] S. Rezazadeh, A. Moradzadeh, K. Pourhossein, B. Mohammadi-Ivatloo, and F. P. García Márquez, "Photovoltaic array reconfiguration under partial shading conditions for maximum power extraction via knight's tour technique," *Journal of Ambient Intelligence and Humanized Computing*, Feb. 2022, <https://doi.org/10.1007/s12652-022-03723-1>.
- [33] S. Daraban, D. Petreus, and C. Morel, "A novel global MPPT based on genetic algorithms for photovoltaic systems under the influence of partial shading," in *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, Vienna, Austria, Aug. 2013, pp. 1490–1495, <https://doi.org/10.1109/IECON.2013.6699353>.
- [34] L. L. Jiang, D. L. Maskell, and J. C. Patra, "A novel ant colony optimization-based maximum power point tracking for photovoltaic systems under partially shaded conditions," *Energy and Buildings*, vol. 58, pp. 227–236, Mar. 2013, <https://doi.org/10.1016/j.enbuild.2012.12.001>.
- [35] A. Fathy and H. Rezk, "A novel methodology for simulating maximum power point trackers using mine blast optimization and teaching learning based optimization algorithms for partially shaded photovoltaic system," *Journal of Renewable and Sustainable Energy*, vol. 8, no. 2, Mar. 2016, Art. no. 023503, <https://doi.org/10.1063/1.4944971>.
- [36] H. Rezk, A. Fathy, and A. Y. Abdelaziz, "A comparison of different global MPPT techniques based on meta-heuristic algorithms for photovoltaic system subjected to partial shading conditions," *Renewable*

- and *Sustainable Energy Reviews*, vol. 74, pp. 377–386, Jul. 2017, <https://doi.org/10.1016/j.rser.2017.02.051>.
- [37] Y. Zhang, J. Su, C. Zhang, Z. Lang, M. Yang, and T. Gu, "Performance estimation of photovoltaic module under partial shading based on explicit analytical model," *Solar Energy*, vol. 224, pp. 327–340, Aug. 2021, <https://doi.org/10.1016/j.solener.2021.06.019>.
- [38] S. Boubaker, S. Kamel, and M. Kchaou, "Prediction of Daily Global Solar Radiation using Resilient-propagation Artificial Neural Network and Historical Data: A Case Study of Hail, Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5228–5232, Feb. 2020, <https://doi.org/10.48084/etasr.3278>.
- [39] A. Kajihara and A. T. Harakawa, "Model of photovoltaic cell circuits under partial shading," in *2005 IEEE International Conference on Industrial Technology*, Hong Kong, China, Sep. 2005, pp. 866–870, <https://doi.org/10.1109/ICIT.2005.1600757>.
- [40] M. Z. Shams El-Dein, M. Kazerani, and M. M. A. Salama, "Optimal photovoltaic array reconfiguration to maximize power production under partial shading," in *2012 11th International Conference on Environment and Electrical Engineering*, Venice, Italy, Feb. 2012, pp. 255–260, <https://doi.org/10.1109/EEEIC.2012.6221584>.
- [41] M. G. Villalva, J. R. Gazoli, and E. R. Filho, "Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays," *IEEE Transactions on Power Electronics*, vol. 24, no. 5, pp. 1198–1208, Feb. 2009, <https://doi.org/10.1109/TPEL.2009.2013862>.
- [42] Y. J. Wang and P. C. Hsu, "An investigation on partial shading of PV modules with different connection configurations of PV cells," *Energy*, vol. 36, no. 5, pp. 3069–3078, 2011, <https://doi.org/10.1016/J.ENERGY.2011.02.052>.

## AUTHORS PROFILE



**Abdullahi A. Masud** received his BE Degree in Electrical Engineering from Ahmadu Bello University (ABU) in 1999. Subsequently, he pursued his Master's degree in the same discipline at ABU and graduated in 2003. He has also obtained a PhD in Electrical Engineering in 2013 from the Glasgow Caledonian University, Scotland, UK. In 2002 he became an assistant lecturer in the Faculty of Engineering at the Ahmadu Bello University, Zaria, Nigeria. In 2013 he

joined Jubail Industrial College, Saudi Arabia, and is currently an Associate Professor in the Department of Electrical Engineering. He has several publications in high impact factor journals and conference proceedings in high voltage partial discharge and renewable energy. He is a Chartered Engineer, a Member of the IET, and a registered Engineer (COREN) Nigeria.

# ANN and GRNN-Based Coupled Model for Flood Inundation Mapping of the Punpun River Basin

**Shashi Ranjan**

Department of Civil Engineering, National Institute of Technology Patna, India  
shashi.ce18@nitp.ac.in  
(corresponding author)

**Vivekanand Singh**

Department of Civil Engineering, National Institute of Technology Patna, India  
vsingh@nitp.ac.in

*Received: 13 November 2022 | Revised: 22 November 2022 | Accepted: 25 November 2022*

## ABSTRACT

The Punpun River is primarily a rain-fed river. Forecasting rainfall accurately would enable an early evaluation of drought and flooding conditions. Therefore, having a flawless model for predicting rainfall is important for the hydrological analysis of any river basin. In this study, Artificial Neural Network (ANN)-based models were developed to predict rainfall and discharge in the basin. During the rainy season, water is spread in and around the area of the watershed, thus a General Regression Neural Network (GRNN)-based model was proposed for fast estimation of the inundation area during the flood taking as input cross-section, rainfall, and discharge. The proposed ANN-GRNN coupled model is the first of its kind for this study area. The assessment of the results shows that the proposed GRNN-based model is capable of estimating the water-spreading area.

**Keywords-**ANN; GRNN; rainfall; flood plain; Punpun river basin

## I. INTRODUCTION

Rainfall is recognized as one of the most significant elements of the hydrological process [1]. If there is no evaporation, runoff, or infiltration, rainfall is the amount of water that falls on a flat surface over a specific period and is measured in millimeters (mm) above the horizontal surface. Forecasting rainfall accurately enables the early evaluation of drought and flooding conditions. Therefore, it is important to have a flawless rainfall prediction model. In a country like India, where most farmers rely on monsoons, good quality and quantity of water are very necessary for their crops. So, it is necessary to have advanced knowledge of the real amount of rainfall [2-5]. As several states in India are experiencing drought at the same time that many other states experience flooding, it is necessary to use an accurate and effective rainfall forecast model. This rainfall forecasting model could improve the handling of flood and drought problems. The ability of this model to anticipate rainfall in advance also provides sufficient time to plan transportation, lifesaving measures, and food and medication supply [6].

Time-series data are provided by scientific research, meteorological stations, GPS, sensor networks, etc. Time-series data have high dimensions and large volumes, and are

continuously updated. Data analysis research is interested in the use of time series data for forecasting, pattern recognition, anomaly detection, pattern identification, clustering, classification, and segmentation [7]. Artificial Neural Networks (ANNs) are a well-established method for simulating complicated nonlinear and dynamic systems. ANNs are useful in creating an appropriate model, especially when the physical process relationship is unclear or the nature of the event exhibits chaotic qualities. Although prior knowledge of the system is necessary, ANNs lessen their reliance on this knowledge. As a result, there is no longer a requirement for a precise specification of the relationship's actual functional form, which the model aims to depict. Several studies utilized a variety of strategies to analyze time-series data, and among them, Generalized Regression, Pearson Coefficient, Fuzzy Inference System, Focused Time Daley Neural Network, and other Neural Networks (NNs) were found to perform sufficiently. Time series data have been analyzed in a variety of ways [8-10].

This study aimed to develop a rainfall prediction model for the Punpun river basin using ANN and GRNN-based models for fast estimation of the inundation area, taking as input cross-section, rainfall, and discharge. The three-layer Feed Forward Network (FFN) structure was used to construct the models and

they were trained using back-propagation. The daily meteorological data of the Punpun river basin were considered for the development of the daily rainfall prediction models. The performance of the models was evaluated using various statistical indices.

## II. STUDY AREA

Figure 1 shows a map of the Punpun river basin, which originates in the highlands of the Palamu district of the Indian state of Jharkhand at an elevation of 300 meters, 24°11' north and 84°9' east. Punpun merges with the river Ganga at Fatuha, 25Km downstream of Patna. It has no flow during dry months, but it has a significant discharge during monsoons which causes floods and submerges the nearby area. The water management of the basin is affected by both high discharge and no flow. From August to October, there is a lot of rain and the humidity of the region is moderate. The average annual rainfall ranges from 99 to 134cm, the highest at the upper levels (Palamu district). 80–87% of the yearly rainfall in the Punpun basin falls during the monsoon. The lower portion of the basin experiences uniform precipitations that do not change frequently. The Punpun river basin research area is 7,055Km<sup>2</sup>. The region's geology is diverse, ranging from recent alluvium in the plains to granite, gneiss, and charnokites in the highlands.

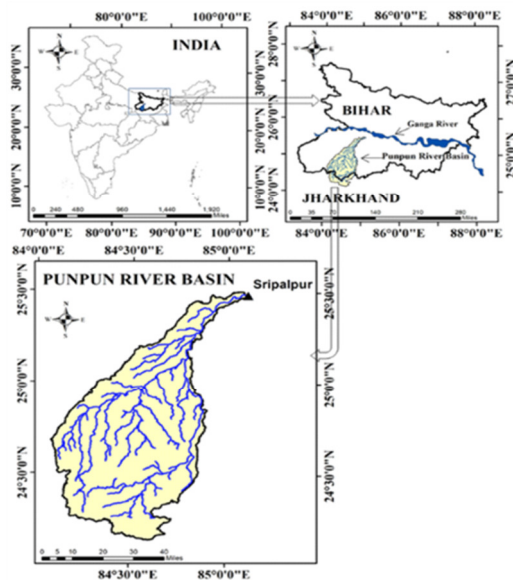


Fig. 1. The study area of the Punpun river basin.

## III. METHODOLOGY

### A. Data Collection

The daily observed rainfall data for 39 years (1 Jan 1982 to 31 December 2020) were collected from NASA's Agro Climatic Data (NASA POWER data access service (<http://power.larc.nasa.gov/data-access-viewer/>)) for the Punpun river basin in India. Meteorological parameters, i.e., daily precipitation, relative humidity, wind speed, and maximum and minimum temperature, were used as input for the models to forecast daily rainfall. These data involved daily values of

rainfall, temperature (min, max), relative humidity, wind speed, and solar radiation and were subjected to preanalysis and formulation of the database. 70% of the data were used for model calibration and the remaining 30% was used for validation.

### B. Rainfall Prediction Modeling using ANN

This study used ANN models to predict daily rainfall in the Punpun River Basin, India. The combination of meteorological parameters as input for training ANNs was found as the most satisfactory for forecasting [11]. Therefore, the observed daily time series of temperature ( $T_{max}$ ,  $T_{min}$ ), relative humidity ( $RH$ ), wind speed ( $W$ ), and solar radiation ( $S_r$ ) were considered. As the rainfall process is dynamic, current-day rainfall ( $P_{ij}$ ) is considered to be a function of temperature, relative humidity, wind speed, and solar radiation. The functional form of rainfall modeling can be expressed as follows:

$$P_{ij} = f(T_{max}, T_{min}, RH, W, S_r) \quad (1)$$

ANNs have been widely used to model basin hydrology and develop a non-linear relationship between the variables. ANNs have no fixed method for determining the pairs of input and output data. To fill this void, the number of data pairs used for training should be equal to or greater than the number of parameters (weights) in the network. This study used 14244 input-output datasets, separated into training from January 1982 to December 2014 and prediction/validation from Jan-2015 to Dec 2020. Rainfall prediction was modeled using the ANN tool in MATLAB. Figure 2 shows the approach to predicting daily rainfall using an ANN with LM training methods. One hidden layer was employed to mimic the rainfall in addition to the input and output layers. The model architecture used a feedforward neural network and backpropagation and was made up of several neurons dispersed across and related to each other. At first, the network was trained on a set of paired data to manage the input-output meaning. Then the weights of the connections between neurons were fixed, and the network was used to determine the classification of new data.

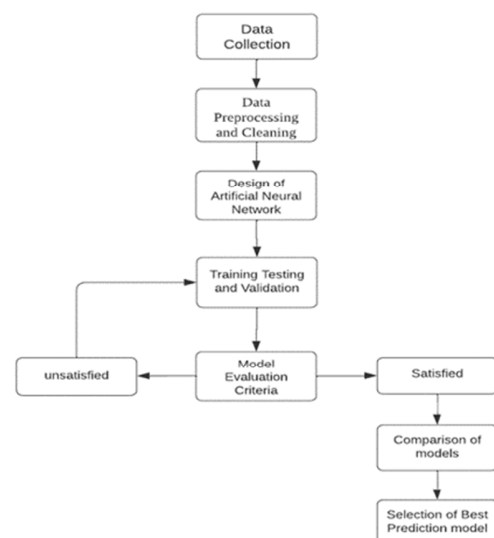


Fig. 2. Method for the prediction of daily rainfall in Punpun river basin.

This study used the most suitable algorithm based on the model evaluation criteria for forecasting, the feed-forward method with back-propagation [12]. The input data for the system were the daily meteorological data of the Punpun river basin. The training process used 70% of the time-series data, while each of the testing and validation processes used 15%. The purpose of the training was to identify the connections, weights, and biases that allow the neural network to estimate outputs that are reasonably similar to the measured outputs. The training datasets were also used to reduce errors. By changing the weights and biases in the ANN, different training methods can be used to train the input parameters to produce an output with a smaller global error [13]. The Levenberg-Marquardt (LM) [14] training algorithm was used in the training process of the developed ANNs. Table I shows the training parameters.

TABLE I. DETAILS OF TRAINING PARAMETERS

Parameters used for training ANN	
Network type	Feed Forward Neural Network with Back-Propagation
Training function	Levenberg Marquardt (LM)
Adoption learning function	Learn GDM
Performance function	Mean Square Error (MSE); Regression (r)
Transfer function	Hidden layer-Tansigmoid; Output layer- Linear
No. of neurons used in the hidden layer	5-45

### C. Levenberg-Marquardt (LM) Training Algorithm

The LM training algorithm is a standard and widely applied iterative method that, in most cases, deals with curve-fitting arrangements and trains ANNs much more quickly than the typical back-propagation algorithm. The LM method finds the base of a multivariable function and expresses it as the sum of squares of nonlinear genuine valued functions. This approach combines the advantages of two strategies by switching the parameter updates between the Gauss-Newton and the gradient descent update. The LM algorithm alters the conventional Newton algorithm to determine the best answer to a minimization issue, as:

$$X_{n+1} = X_n - \{J^T J + \mu I\}^{-1} J^T e \quad (2)$$

where  $x$  is the weight of NNs,  $J$  is the Jacobian matrix,  $e$  is the residual error vector, and  $\mu$  represents the scaler that controls the learning process. As the LM algorithm is computationally and memory intensive, it is best suited for small networks.

### D. General Regression Neural Network (GRNN)

The General Regression Neural Network (GRNN) model was used to estimate the river water spread by using rainfall and discharge as estimated by the ANN model and river-cross section. GRNN is a neural network-based function predicting algorithm, more desirable than other neural network models as it doesn't require any data for the iterative training process [15, 16]. GRNN is based on nonlinear regression and can utilize the training data to approximate any arbitrary continuous nonlinear function. For GRNN, a weighted average of the training dataset's outputs is used to calculate the new output. The Euclidean distance between a given pattern and the training

dataset is used to estimate the weight of that pattern. A smaller distance results in a higher weight for the pattern while a larger distance results in a lower. A GRNN model has four fundamental layers: input, pattern, summation, and output [17]. Figure 3 shows the GRNN network model. The input layer contains all input data from the simulation. The pattern layer determines the relevant weight and computes the Euclidean distance. The inputs' Euclidean distances are used as the basis for determining how much weight to give to each pattern. Equation 3 calculates the new output based on the new input and training datasets.

$$Y(X) = \frac{\sum y_i e^{-\left(\frac{d_i^2}{2\sigma^2}\right)}}{\sum e^{-\left(\frac{d_i^2}{2\sigma^2}\right)}} \quad (3)$$

where  $d_i$  is the Euclidean distance or spread between the new ( $X$ ) and the training input ( $X_i$ ), the higher the value of, the less likely it is that the output will be close to the inputs' extreme values and vice versa. When the distance  $d_i$  value is small, the weight term returns a relatively large value and vice versa. If  $d_i$  is zero, the weight term gives a result of 1, meaning that the output of the test data will be the same as the output of the training sample.

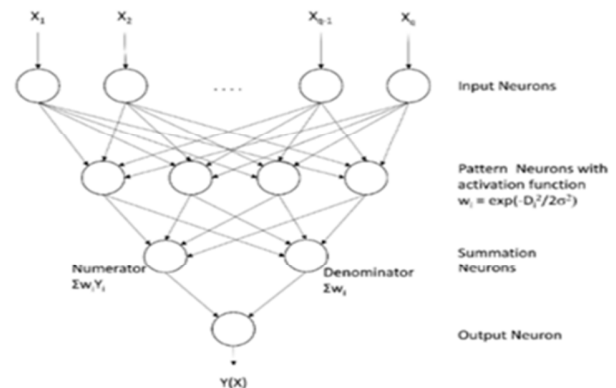


Fig. 3. The basic network diagram of a GRNN model.

The spread value  $\sigma$  is the only parameter that needs to be estimated for the GRNN model. The training process involves determining the ideal value of  $\sigma$ . The best method to determine the spread value is to use an optimization algorithm to reduce Mean Squared Error (MSE). The entire dataset is split into two parts, training and testing, to obtain the ideal value of the spread. The best value of the spread is then obtained by minimizing MSE after GRNN is applied to the test data based on the training data. As stated previously, the model is created to estimate the extent of water spreading, particularly during the monsoon. A possible representation of an input pattern ( $k$ ) is as follows:

$$X_k = [Q_k^t, Q_k^{t-1}, Q_k^{t-2}, S_k^t, S_k^{t-1}, S_k^{t-2}, P_k^t, P_k^{t-1}, P_k^{t-2}] \quad (4)$$

where  $Q$  represents the discharge,  $P$  is precipitation, and  $S$  is the river cross-section. The corresponding output pattern is given by:

$$Y_k = [SW_k^{t+1}] \quad (5)$$



Figure 4 shows the GRNN model applied in this study. Any arbitrary pattern  $k$  for the Punpun river basin, where recorded water level data is available, is represented by the input pattern  $X_k^i$ . The pattern layers calculate the weight ( $w_n^i$ ) of each individual pattern as well as the distance ( $d_n^i$ ) between the input and training patterns. There are  $N$  different training pattern options.  $N$  neurons will therefore be present in the pattern layer. The weight of the calculated pattern is multiplied by the corresponding output in the summation layer, the result is summed at the numerator neuron, and all the weights are added up at the denominator neuron. The output neuron uses (3) to calculate the output for the arbitrary vector. The spread value  $\sigma$ , which is the only parameter in the model, can be calculated using an optimization algorithm. The model can now be used to estimate the extent of water spreading after it has been trained or after the optimal spread value has been determined. This model was constructed for the  $i^{th}$  cross-section of the watershed along the Punpun river line. Thus, the model was named  $GRNN_i$ .

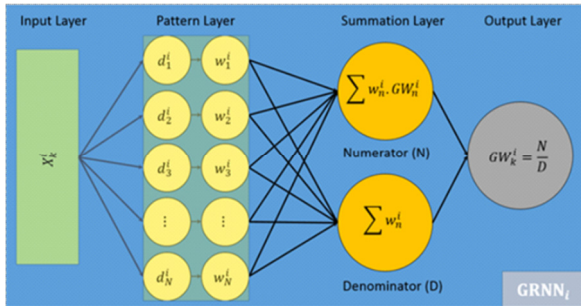


Fig. 4. GRNN architecture of the model.

#### E. Model Performance Evaluation

The prediction performance of each model was determined by the statistical and hydrological indices and quantitative comparison between the observed and predicted values.

##### 1) Statistical Indices

###### a) Mean Square Error (MSE)

MSE is used to compare differences between observed and predicted values. MSE is zero for a perfect fit, and increased values indicate a higher deviation between predicted and observed values. MSE is determined using:

$$MSE = \frac{1}{n} \sum_{i=1}^n (P_{observed} - P_{predicted})^2 \quad (6)$$

###### b) Correlation Coefficient

The correlation coefficient ( $r$ ) is an indicator of the degree of closeness between the observed and predicted values and provides the level of explained variance between them [18]. The correlation coefficient is given by:

$$r = \frac{\sum_{i=1}^n (P_{observed} - \bar{P}_{observed})(P_{predicted} - \bar{P}_{predicted})}{\sqrt{(\sum_{i=1}^n (P_{observed} - \bar{P}_{observed})^2) \sum_{i=1}^n (P_{predicted} - \bar{P}_{predicted})^2}} \quad (7)$$

##### 2) Nash-Sutcliffe Efficiency (NSE) Hydrological Indice

The NSE was developed by Nash and Sutcliffe in 1970. It is also called the coefficient of efficiency. It provides the proportions of variance of the observation for a model and is

widely used in hydrology [19]. NSE has a range between  $-\infty$  to 1, and a value of 1 indicates a perfect match. It is determined by:

$$NSE = 1 - \frac{\sum_{i=1}^n (P_{observed} - P_{predicted})^2}{\sum_{i=1}^n (P_{observed} - \bar{P}_{observed})^2} * 100 \quad (8)$$

where  $P_{observed}$ ,  $P_{predicted}$ , and  $\bar{P}_{observed}$  are observed, predicted, and mean rainfall values at the Punpun River basin, and  $n$  is the number of observed datasets.

#### IV. RESULTS AND DISCUSSION

The performance of the models was qualitatively and quantitatively evaluated by visual observation and various statistical and hydrological indices.

##### A. ANN-Based Models for Rainfall Prediction during Training and Testing

The model was trained by varying the number of neurons in the hidden layer. Each NN was identified by the number of input, hidden, and output nodes. Table II shows the performance evaluation indices for rainfall estimation during training and testing. The networks having lower MSE and higher  $r$  and NSE values were chosen as the best prediction models. Table II shows that the performance of model 5-15-1 was better during training and testing. The  $r$  values for the ANN networks ranged from 0.76 to 0.93 and from 0.82 to 0.97 during training and testing, respectively. MSE values ranged from 7.08 to 20.7 and from 3.35 to 16.7 during training and testing, respectively.

TABLE II. PERFORMANCE OF VARIOUS ANN MODELS

Training				Testing			
Network	MSE	r	NSE	Network	MSE	r	NSE
5-5-1	12.45	0.83	81	5-15-1	10.45	0.87	88
5-10-1	16.78	0.81	78	5-19-1	13.78	0.85	86
5-15-1	<b>7.08</b>	<b>0.93</b>	<b>86</b>	5-15-1	<b>3.35</b>	<b>0.97</b>	<b>94</b>
5-20-1	10.6	0.85	82	5-20-1	8.6	0.91	9
5-25-1	8.47	0.9	83	5-25-1	5.47	0.95	92
5-30-1	13.48	0.82	8	5-30-1	11.48	0.86	87
5-35-1	19.48	0.77	73	5-35-1	15.48	0.83	85
5-40-1	17.33	0.8	77	5-40-1	14.33	0.84	86
5-45-1	20.7	0.76	7	5-45-1	16.7	0.82	84

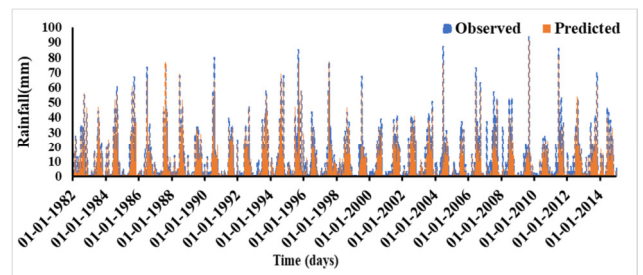


Fig. 5. Observed and predicted rainfall during training period.

Figure 5 shows the observed and predicted values of rainfall during the training period (1982-2014) using the 5-15-1 ANN model, while Figure 6 shows the same values for the testing period (2015-2020). These figures show that there is close or nearly close agreement between the observed and predicted rainfall, and the overall shape of the predicted rainfall

curve is similar to the observed rainfall. The quantitative evaluation of the model, based on NSE, shows a significant relationship between observed and predicted rainfall. The NSE values for the 5-15-1 ANN were 84% and 94% during training and testing, respectively. Thus, the 5-15-1 ANN was selected as the best for rainfall prediction. Figure 7 shows the correlation between the observed and predicted rainfall for the 5-15-1 ANN during training and testing.

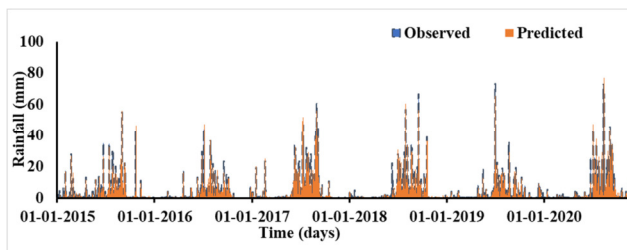


Fig. 6. Observed vs predicted rainfall for the 5-15-1 ANN model during the testing period.

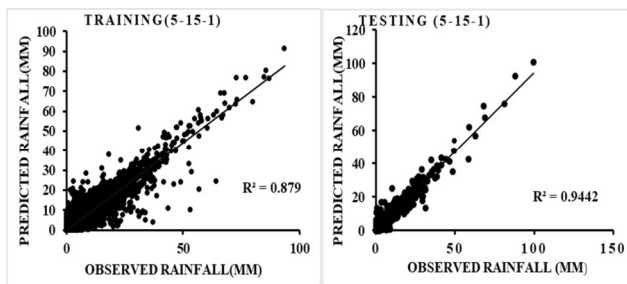


Fig. 7. Correlation of observed and predicted rainfall for the 5-15-1 ANN during training (left) and testing (right).

The quantitative and qualitative performance during training and testing was found satisfactory, as these models predicted rainfall with acceptable accuracy. Based on qualitative and quantitative performance evaluation, the ANN models gave satisfactory results for the study area.

#### B. GRNN Model for Punpun River Flood Mapping

A scattered plot was plotted for the surface water spread between the simulated and the actual data at different cross-sections, and a map was created in ArcGIS for better visualization [20]. Table III and Figure 8 show that the variation of the water spread in form of fluctuation is satisfactory and the  $R^2$  value varied from 0.69 to 0.91.

TABLE III. CORRELATION OF ACTUAL AND PREDICTED WATER SPREAD

Cross section name	$R^2$ value b/w actual and GRNN predicted
1	0.696
2	0.769
3	0.672
4	0.619
5	0.711
6	0.699
7	0.912
8	0.811
9	0.678
10	0.745

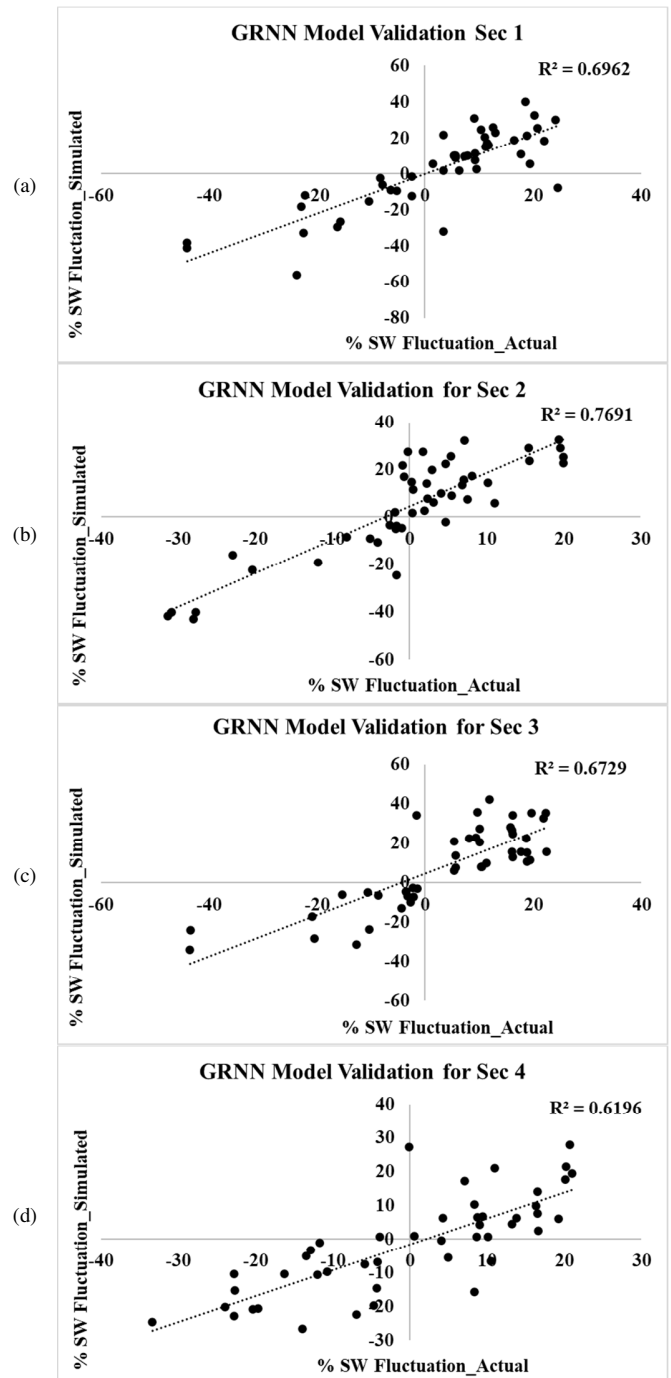


Fig. 8. Scatterplots between actual and predicted spread water fluctuation by the GRNN model of section 1, 2, 3, and 4 cross-sections, respectively.

Figure 9 shows a water spreading map, where there is not much water inundation in the Punpun river basin. The map shows the width of the river to the water spread area of that river section. The main reason behind this is that Punpun is a rainfed river. At some cross sections, mainly in the middle part of the river basin, the water is spreading away from the main streamline to the flood plain.

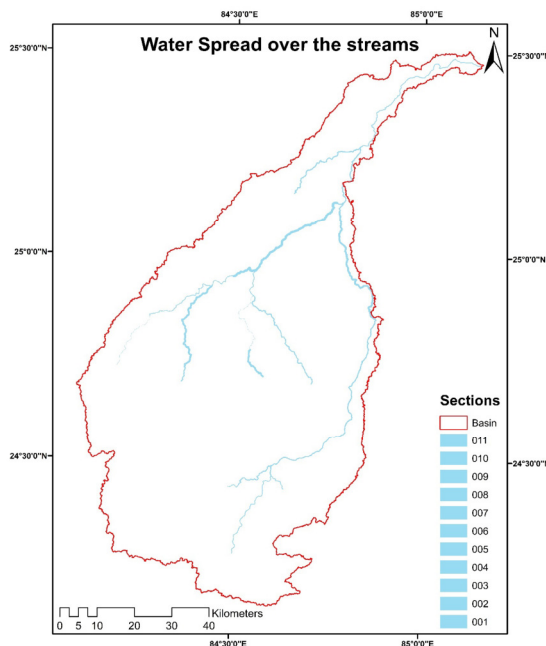


Fig. 9. Actual spread water fluctuation along the river cross-sections.

## V. CONCLUSIONS

This study showed that feedforward backpropagation networks using the LM training algorithm produced better results compared to others. Rainfall is more influenced by factors such as minimum and maximum temperature, relative humidity, wind speed, and solar radiation, so these characteristics cannot be ignored in predicting. Instead of pre- and post-monsoon months, mid-monsoon months provide more accurate results for predicting rainfall. The initial flooded area map was created using GRNN, a popular machine-learning method with strong learning performance. The initial flood map projected by the GRNN was subjected to easy-to-implement two-stage post-processing to further reduce missed and false water spread detections and increase the accuracy and reliability of flood mapping [20]. Therefore, the proposed GRNN model was capable of producing the water spread fluctuation over the cross-section. The limitation of the coupled model was that this study area is mainly a rainfed river system and the coupled model can be used in other river basin systems such as the Himalayan River system to get more satisfactory results.

## REFERENCES

- [1] S. Rana, V. Deoli, and S. R. Chavan, "Detection of abrupt change in trends of rainfall and rainy day's pattern of Uttarakhand," *Arabian Journal of Geosciences*, vol. 15, no. 7, Mar. 2022, Art. no. 618, <https://doi.org/10.1007/s12517-022-09883-w>.
- [2] V. Deoli and S. Rana, "Seasonal trend analysis in rainfall and temperature for Udaipur district of Rajasthan," *Current World Environment*, vol. 14, no. 2, pp. 312–319, 2019.
- [3] A. Towheed and R. Thendiyath, "Spatiotemporal Rainfall Dynamics in Kosi Basin Using Wavelet Analysis," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7578–7584, Oct. 2021, <https://doi.org/10.48084/etasr.4372>.
- [4] A. Towheed and T. Roshni, "Linking climate change to soil loss estimation in the Kosi river basin, India," *Journal of Water and Climate Change*, vol. 12, no. 6, pp. 2338–2363, Mar. 2021, <https://doi.org/10.2166/wcc.2021.259>.
- [5] K. Praveen and L. B. Roy, "Assessment of Groundwater Quality Using Water Quality Indices: A Case Study of Paliganj Distributary, Bihar, India," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8199–8203, Feb. 2022, <https://doi.org/10.48084/etasr.4696>.
- [6] A. Rawat, P. Kumar, and V. Deoli, "Daily Monsoon Rainfall Prediction using Artificial Neural Network (ANN) for Parbhani District of Maharashtra," *International Journal of Current Microbiology and Applied Sciences*, vol. 8, no. 12, pp. 1949–1963, Dec. 2019, <https://doi.org/10.20546/ijcmas.2019.812.233>.
- [7] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys*, vol. 45, no. 1, Sep. 2012, Art. no. 12, <https://doi.org/10.1145/2379776.2379788>.
- [8] Y. Dash, S. K. Mishra, and B. K. Panigrahi, "Rainfall prediction for the Kerala state of India using artificial intelligence approaches," *Computers & Electrical Engineering*, vol. 70, pp. 66–73, Aug. 2018, <https://doi.org/10.1016/j.compeleceng.2018.06.004>.
- [9] A. Raja and G. T. "Drought Prediction and Validation for Desert Region using Machine Learning Methods," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 7, pp. 47–53, 2022, <https://doi.org/10.14569/IJACSA.2022.0130707>.
- [10] 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>.
- [11] N. Q. Hung, M. S. Babel, S. Weesakul, and N. K. Tripathi, "An artificial neural network model for rainfall forecasting in Bangkok, Thailand," *Hydrology and Earth System Sciences*, vol. 13, no. 8, pp. 1413–1425, Aug. 2009, <https://doi.org/10.5194/hess-13-1413-2009>.
- [12] L. Zhao, F. E. Hicks, and A. R. Fayek, "Applicability of multilayer feed-forward neural networks to model the onset of river breakup," *Cold Regions Science and Technology*, vol. 70, pp. 32–42, Jan. 2012, <https://doi.org/10.1016/j.coldregions.2011.08.011>.
- [13] A. Ehret, D. Hochstuhl, D. Gianola, and G. Thaller, "Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle," *Genetics Selection Evolution*, vol. 47, no. 1, Mar. 2015, Art. no. 22, <https://doi.org/10.1186/s12711-015-0097-5>.
- [14] D. Yadav, R. Naresh, and V. Sharma, "Stream flow forecasting using Levenberg-Marquardt algorithm approach," *International Journal of Water Resources and Environmental Engineering*, vol. 3, no. 1, pp. 30–40, Jan. 2011.
- [15] D. Kumar and R. K. Bhattacharjya, "GRNN Model for prediction of groundwater fluctuation in the state of Uttarakhand of India using GRACE data under limited bore well data," *Journal of Hydroinformatics*, vol. 23, no. 3, pp. 567–588, Mar. 2021, <https://doi.org/10.2166/hydro.2021.108>.
- [16] 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>.
- [17] S. A. Hannan, R. R. Manza, and R. J. Ramteke, "Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis," *International Journal of Computer Applications*, vol. 7, no. 13, pp. 7–13, Oct. 2010, <https://doi.org/10.5120/1325-1799>.
- [18] K. R. K. Reddy and V. Singh, "Numerical Model to Simulate Soil Wetting Pattern under Drip Irrigation System," *Agricultural Science Digest*, Aug. 2022, <https://doi.org/10.18805/ag.D-5619>.
- [19] S. Ranjan and V. Singh, "HEC-HMS based rainfall-runoff model for Punpun river basin," *Water Practice and Technology*, vol. 17, no. 5, pp. 986–1001, Apr. 2022, <https://doi.org/10.2166/wpt.2022.033>.
- [20] Q. Zhang, P. Zhang, and X. Hu, "Unsupervised GRNN flood mapping approach combined with uncertainty analysis using bi-temporal Sentinel-2 MSI imageries," *International Journal of Digital Earth*, vol. 14, no. 11, pp. 1561–1581, Nov. 2021, <https://doi.org/10.1080/17538947.2021.1953160>.

# Implementation of Crow Search Algorithm for Achieving Optimal Control of a Single-Link Flexible Manipulator

**Venkata Satya Durga Manohar Sahu**

School of Electrical Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, India

manohar17.me@gmail.com

(corresponding author)

**Padarbinda Samal**

School of Electrical Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, India

padarbindasamal87@gmail.com

**Chinmoy Kumar Panigrahi**

School of Electrical Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, India

panigrahichinmoy@gmail.com

*Received: 30 September 2022 | Revised: 29 October 2022 | Accepted: 16 November 2022*

## ABSTRACT

A metaheuristic optimizer based on the cognitive behavior of crows, called the Crow Search Algorithm (CSA), is suggested in this paper for the optimization of the performance of the single link flexible manipulator as a constrained optimal control problem, which comprises objective functions, constraints, and decision variables. The CSA outperformed other algorithms, such as Differential Evolution, Grey Wolf Optimization, and Particle Swarm Optimization. Parameter setting is another highlight of the current paper. In conclusion, CSA yields more promising results than other approaches.

*Keywords-optimal control system; crow search algorithm; single link flexible manipulator*

## I. INTRODUCTION

Optimization has played a critical part in a wide range of problems during the last few decades. Engineers have relied on conventional search methods for decades to solve design problems. Despite the fact that these approaches give promising outcomes, they may fail in more complicated design challenges. Optimization problems are often complicated due to their complex objective functions, a substantial amount of decision variables, and many constraints. So, traditional optimization strategies often have limited effectiveness. The decision parameters may be too many and their influence on the objective function can be complex in real-world design challenges [1]. This means that the objective function might have multiple local optimal solutions while the design is only interested in the optimal global solution. These challenges can't be solved by conventional methods that only discover local optima in a specific location. We need efficient optimization methods in these situations. Consequently, authors in [2]

developed a new field of research called Swarm Intelligence (SI) in the late 1980s, which is still being explored today. Today's trend is to use natural metaheuristic algorithms [3, 4] to solve challenging issues, and metaheuristics have been proved to be surprisingly efficient. Metaheuristic algorithms have shown promising performance in handling the most severely nonlinear and multimodal real-world optimization problems. Particular randomization and local search are used by all metaheuristic algorithms [5]. These methods can find good answers to challenging optimization problems, but optimum solutions cannot be found. However, it is hoped that these methods will work in most cases. It is possible to use metaheuristic algorithms for global optimization.

Single-link flexible manipulators are often formulated as optimization problems [6-9] in order to optimize the performance index while meeting various constraints. As it has the potential advantage of lower cost, larger work volume, higher operational speed, greater payload-to-manipulator-

weight ratio, smaller actuators, lower energy consumption, better transportability, and safer operation due to reduced inertia over traditional heavy and bulky robots. The main drawback of these manipulators is the vibration problem caused by their poor rigidity [2]. The Lagrangian-assumed modes technique [10] is used to create a dynamic model of a single-link flexible manipulator. Controlling was conducted by various law [11] and intelligent techniques [12, 13]. An overall goal is to maximize system characteristics [14] for accessibility, which is represented by a vector. Particle Swarm Optimization (PSO) algorithm was used to model a flexible manipulator system in [15]. An optimization technique begins by creating an objective function capable of generating the problem objectives without removing any constraints in order to identify the best feasible solution. According to the Glover Convention, all current procedures influenced by nature are considered metaheuristics [16]. Differential Evolution (DE) algorithm [9, 17], based on natural selection and PSO which is based on bird flock and fish schooling social behavior [18], are some of the best-known metaheuristic algorithms. The Crow Search Algorithm (CSA) [19] was presented in 2016. The social intelligence and food-collecting mechanism of a crow group are included in the design. Crows are a bird species that has expanded far across the globe and are considered very intelligent. They can memorize faces, utilize tools, communicate, and conceal food in various complicated ways.

The main objectives of the work at hand are:

- Parameter setting is done with the manipulator model using the proposed approach.
- Performance comparison of the proposed CSA with the existing DE, PSO, and Grey Wolf Optimizer (GWO).
- The CSA's ability to identify an appropriate alternative strategy for addressing complex engineering optimization problems is investigated.

## II. PROBLEM FORMULATION

An optimization problem [20] is the objective of identifying the optimal solution among all possible options. Optimization problems can be classified as either unconstrained or constrained based on whether the variables are discrete or continuous. This section uses the SLFM to formulate the optimal control problem [9] described in detail.

### A. Modeling of the Dynamic System

Regarding the modeling of the dynamic system, assume the following set of state equations:

$$\dot{z} = g(z(t), v(t), t) \quad (1)$$

where  $z$  is a state vector,  $v$  is a manipulator control torque input vector,  $t$  is continuous-time, and  $g$  is a vector function.

### B. Performance Index

An objective function or its inverse (referred to as a profit function, fitness function, utility function, or a reward function in some disciplines) is an objective function that should be maximized. It is referred to as:

$$J(v) = \varphi(z(t_f), t_f) + \int_{t_0}^{t_f} L(z(t), v(t), t) dt \quad (2)$$

$L$  is a real-valued state function and the input vector that may or may not be explicitly time-dependent depending on the situation. The cost function is a term that is often used to describe this function. The function is a real-valued function of the vector of the final state and its time vector, with the final state vector serving as the input. When it comes to quality, a performance index should be constructed, so that it is constrained from below, and the higher the index, the poorer the controller's performance. For example, ensuring that the first term in (2) and the integrand in the second term are positive for all values of  $z(t)$ ,  $v(t)$  and  $t$ , may be achieved by requiring. The first term in (2) denotes a constraint on the state vector's final or terminal value. The value of the performance index decreases as the end state vector approaches the desired value. After specifying the cost function and constraint, the objective is to find the optimal controller, that is, the value of the control signal  $v(t)$  for the time period  $t_0 \leq t \leq t_f$  that yields the lowest values  $J$  under the assumption that the state vector obeys the state equation (1).

### C. Constraints

When solving an optimization problem, a constraint is a requirement that the solution must meet. There are numerous different types of constraints, the most common of which are equality constraints, inequality constraints, and integer constraints.

- Inequality constraints exist when a function has a border or limitations, e.g. less than equal to, less than, greater than, greater than equal to, and is expressed as:

$$G(z(t)) \leq 0 \quad (3)$$

where  $G(\cdot)$  denotes the limitations imposed by inequality on the vector mapping of the state variables.

- Equality restrictions exist when a function is just the equal to the requirements, i.e. it perfectly matches the value of the resource, and they are expressed as:

$$\Psi(z(t)) = 0 \quad (4)$$

where  $\Psi(\cdot)$  represents equality constraints of the state variables. It's possible that this variable is a state or a control variable. It may alter based on the problem's design and demand. This variable may have limits on rare occasions. It's referred to as:

$$v_{\min} \leq v(t) \leq v_{\max} \quad (5)$$

where  $v(\cdot)$  is a control variable with lower and upper limits. Correspondingly, state variables may be specified based on the design and demand, but it is more complicated to identify boundary values, although a set of minimum and maximum values are used for estimations. Our ultimate objective is, therefore, to achieve the best value possible.



#### D. Optimal Control Problem (OCP) Formulation

Consider the following description of an OCP [21] to be designed:

$$\min_{v(k)} J = \varphi(z(t_f), t_f) + \int_{t_0}^{t_f} L(z(t), v(t), t) dt$$

subjected to

$$G(z(t)) \leq 0$$

$$v_{\min} \leq v(t) \leq v_{\max}$$
(6)

where  $J$  is the performance index, with state constraints  $G(\cdot)$  and control constraints  $v(t)$ . The dynamic model of the SLFM assessed the performance index as well as inequality and equality constraints. For example, in [7, 8, 21], we determine the best solution to the given issue using the CSA and the constrained performance index.

#### E. Example

##### 1) Single-Link Flexible Manipulator Model

Assume a flexible manipulator model [6, 22, 23], as shown in Figure 1.

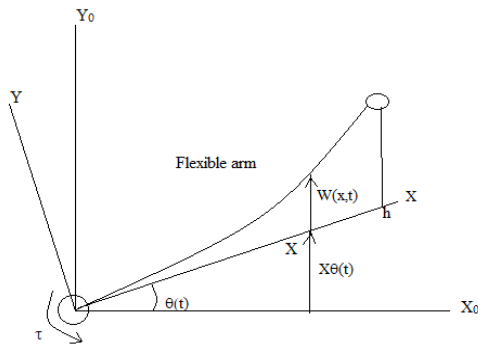


Fig. 1. Flexible link manipulator.

$$(J_h + J_l) \ddot{\theta} + J_l \ddot{\alpha} - mgh \sin(\theta + \alpha) = \tau$$

$$J_l \ddot{\theta} + J_l \ddot{\alpha} + K_s \alpha - mgh \sin(\theta + \alpha) = 0$$
(7)

Let us assume the above equation by considering:

$$z_1 = \theta, z_2 = \alpha, z_3 = \dot{\theta}, z_4 = \dot{\alpha}$$
(8)

and the above equation can be rewritten as follows:

$$\dot{z}_1 = z_3$$

$$\dot{z}_2 = z_4$$

$$\dot{z}_3 = \frac{k_s}{J_h} z_2 - \frac{k_m^2 k_g^2}{R_m J_h} z_3 + \frac{k_m k_g}{R_m J_h} v$$

$$\dot{z}_4 = -\frac{k_s}{J_h} z_2 + \frac{k_m^2 k_g^2}{R_m J_h} z_3 - \frac{k_m k_g}{R_m J_h} v \dots$$

$$\dots + \frac{mgh}{J_1} \sin(z_1 + z_2) - \frac{k_s}{J_1} z_2$$
(9)

where  $m$  is the arm end mass,  $J_l$  is the inertia of the arm,  $g$  is the gravity acceleration,  $B_m$  is the rotational friction,  $J_h$  is the moment of inertia of the motor,  $r$  is the radius,  $z_1$  is the arm end position (i.e. the angular position of the motor  $\theta$ ),  $z_2$  is the arm axis velocity (i.e. the angular displacement of the flexible joint  $\alpha$ ),  $z_3$  is the arm position,  $z_4$  is the arm end velocity. The parameter values are taken from [6].

##### 2) Control Requirements

Initially, consider the manipulator was at  $z = [0.09 \ 0.09 \ 0.09 \ 0.09]^T$  and we wish to move it into the desired position, e.g.  $z = [0 \ 0.4 \ 0 \ 0]^T$  at  $t = 0.78$ s with the following constraints on the velocity and control signal:

$$z_2(t) \leq 0.4, \quad t \in [0, 0.78]$$

$$-10 \leq v(t) \leq 10, \quad t \in [0, 0.78]$$
(10)

##### 3) Formulation as a Constrained Optimal Control Problem

The OCP is defined and formulated as:

$$J = \int_{t_0}^{t_f} z_1^2 + z_2^2 + v^2 dt$$
(11)

Subject to:

$$\begin{cases} \dot{z}_1 = z_3 \\ \dot{z}_2 = z_4 \\ \dot{z}_3 = \frac{k_s}{J_h} z_2 - \frac{k_m^2 k_g^2}{R_m J_h} z_3 + \frac{k_m k_g}{R_m J_h} v \\ \dot{z}_4 = -\frac{k_s}{J_h} z_2 + \frac{k_m^2 k_g^2}{R_m J_h} z_3 - \frac{k_m k_g}{R_m J_h} v \\ \quad + \frac{mgh}{J_1} \sin(z_1 + z_2) - \frac{k_s}{J_1} z_2 \\ z_2(t) - 0.4 \leq 0 \\ -10 \leq v(t) \leq 10 \end{cases}$$
(12)

### III. SOLUTION METHODOLOGY

CSA [19, 22], is based on the foraging behaviors of crows. Crows have various distinct characteristics that they display while hunting for food. When compared to other bird species, they have a large memory for remembering the locations of potential food stocks. Crows steal their owner's food supply after following other birds to food hiding locations. The crow moves the food source from its current site to one that is more barren as it scavenges the available food in order to protect it from other crows. Each point in the search space indicates a solution in the context of CROW. Crows have been observed to watch other birds, noting where they keep their food and then stealing it once the owner has left. If it has previously been a victim of theft, a crow will take extra precautions, such as changing hiding places, to avoid being a victim again. The previous behavior of an individual may be used to predict the future behavior of a thief and find the safest strategy to safeguard their stockpiles from theft. Based on the above-

mentioned intelligent behaviors, the population-based metaheuristic algorithm, CSA, is implemented in this study. To attain the objectives, we must follow specific principles:

- Crows live in groups of individuals (flocks).
- Crows memorize their hiding places.
- Crows congregate in order to steal.
- Crows use measures to protect their caches against stealing.

Each crow is required to make a move based on 1 of the 2 fundamental CSA rules: either protect its own hiding place or find the locations of other members. Imagine, for instance, that a crow decides on iteration  $iter$  to steal food from a hiding spot. Depending on the situation, either (1) the individual crow is not aware that it is being watched, or (2) the individual crow recognizes the presence of the plunderer and acts dishonestly. According to the first scenario, the crow may discover and pillage the hiding location. The crow can be relocated as follows:

$$z^{i,iter+1} = z^{i,iter} + r_i \times fl^{i,iter} \times (m^{i,iter} - z^{i,iter}) \quad (13)$$

where  $r_i$  is a random number with an even distribution between 0 and 1 and  $fl^{i,iter}$  signifies the flight length of the crow  $i$  at iteration  $iter$ .

It's worth noting that  $fl^{i,iter}$  is a parameter of the algorithm and can affect the method's search capabilities. Assume that smaller  $fl$  values result in a local search near  $z^{i,iter}$ , whereas more significant values of  $fl$  expand the search space. In terms of optimization, smaller  $fl$  values contribute to the intensification of the results, whereas more significant  $fl$  values contribute to the diversification of the results. Both high-intensification and high-diversification are desirable properties of an efficient optimization technique.

Alternatively, the  $j^{th}$  crow might think a fellow flock member is chasing it (say the  $i^{th}$  crow). The  $j^{th}$  crow flew erroneously over a non-hideout area to guard its food source. If the  $i^{th}$  crow is arbitrarily placed in the  $d$ -dimensional option space, the CSA can repeat the operation. Therefore, for the 2 conditions mentioned above, the crows' tailing motion may be explained as follows:

$$z^{(i,iter+1)} = \begin{cases} z^{i,iter+1} = z^{i,iter} + \dots \\ \dots r_i \times fl^{i,iter} \times (m^{i,iter} - z^{i,iter}) & \text{rand() } \geq AP \\ \text{random crow position} & \text{otherwise} \end{cases} \quad (14)$$

In metaheuristic algorithms, diversity and intensity should be well-balanced. When it comes to CSA, the system is essentially deepened and diversified by the Awareness Probability (AP) component. When a suitable solution already exists in place, CSA is more likely to look locally by lowering the AP. Low AP levels can be used to boost intensity. However, when knowledge grows, there may be a chance to find a workable solution, and CSA often expands the search

area globally (randomization). The use of high AP values enables the achievement of greater variation

#### A. CSA Implementation for Optimization

The step-wise procedure for the implementation of CSA is given in this section.

Step 1: Initialize parameters AP, flight length  $fl$ , randomly initialize the control variable.

Step 2: Initialize the crows' position and memory.

Step 3: Estimate fitness (objective) function.

Step 4: Generate new position using (14).

Step 5: Inspect the possibility of new positions.

Step 6: Update fitness of new positions

Step 7: Update the memory

Step 8: Verify that the end criteria has been met

Repeat steps 4–7 as many times as necessary until the maximum number of iterations is reached. After that, an optimal memory position in terms of the objective function's value is provided to solve an optimization problem.

#### B. Crow search Algorithm Pseudocode

```
Initialize the parameters
Randomly initialize the position of N
crows flows in the search space
Assess position, performance, and fitness
State the crow memory
while Termination conditions not satisfied
for N of the flock's crows
Pick a crow at random from the group
State awareness probability
if rand() > AP ,
Update the position using (13)
else
Update the position randomly
end if
end for
Assess the new performance and fitness
with the new crow's position
Update the crow's memory
end while
```

## IV. SIMULATION RESULTS AND DISCUSSION

Optimal control problems can't be solved by a single search algorithm. Another way to put it is that one algorithm may answer some problems better and others worse than other algorithms. For a fair evaluation of the proposed CSA, problems are discussed and solved. CSA has been run successfully on a laptop with an i5 processor and 16GB of RAM in MATLAB environment. The flowchart of CSA implementation can be seen in [24].

#### A. CSA Parameter Setting for Implementation Optimization

The CSA's parameter settings have been optimized in a number of ways for 50 runs and 450 iterations. According to

Tables I-III, the best performance index obtained by CSA is 0.001985 and 0.001982 for the single link flexible manipulator. Fair AP values produce better results, nevertheless, bigger values of AP are suggested to avoid becoming trapped in local optima. If AP remains constant ( $AP = 0.1$ ) and the value of  $fl$  is increased from 0.3 to 4, it is projected that the overall performance of CSA will be improved on average.

In Table I, the impact of the AP is shown to assess the CSA's performance. AP and  $fl$  have been set to 0.1 and 2, respectively, in CSA. CSA is also analyzed to see how different parameter settings affect its performance. AP and  $fl$  values influence the results for the SLFM, as shown in Tables I and II. Table I shows that  $AP = 0.25$  results in CSA poor performance. Figure 2 shows its impact. In Table II, it can be seen that when

the value of AP is adjusted to 0.1, and the value of  $fl$  is increased from 0.3 to 4, the overall performance of the CSA improves. As a result, fine tuning CSA, like other optimization methods, is a process best accomplished by trial and error. The impact of the flight length can be observed in Figure 3.

TABLE I. IMPACT OF VARYING AP VALUES FOR  $fl=2$ 

Performance index (J)	AP = 0.1	AP = 0.25	AP = 0.5	AP = 0.75
Best value	0.00200	0.002645	0.001985	0.001998
Worst value	5.1851	18.591	17.1683	8.839
Mean	0.0967	0.3692	0.1079	0.038
Standard deviation	0.4213	1.4912	0.8261	0.5583

TABLE II. THE EFFECT OF  $fl$  ON THE PERFORMANCE OF CSA

Index	$fl = 0.3$ AP = 0.1	$fl = 0.5$ AP = 0.1	$fl = 1$ AP = 0.1	$fl = 1.5$ AP = 0.1	$fl = 1.75$ AP = 0.1	$fl = 2$ AP = 0.1	$fl = 3$ AP = 0.1	$fl = 4$ AP = 0.1
Best value	0.0073622	0.007324	0.016931	0.00496991	0.00883244	0.00200153	0.00256202	0.00250375
Worst value	10.1884	11.9041	5.5145	15.0302	8.1574	5.1851	21.1613	10.3002
Mean	0.9474	0.2573	0.3259	0.0733	0.1488	0.0967	0.129	0.1551
Standard deviation	1.6151	1.1226	0.7193	0.6806	0.7486	0.4213	1.3639	0.7575

TABLE III. THE EFFECT OF ITERATION NUMBER ON THE PERFORMANCE OF CSA

Index	500	1000	1500	2000	2500	3000	3500	4000	5000
Best Value	0.002002	0.001982	0.002217	0.001985	0.00198	0.001989	0.002019	0.001994	0.001988
Worst Value	5.1851	12.3677	16.9836	22.2053	9.3461	16.9836	8.2768	8.2161	16.9836
Mean	0.0967	0.1526	0.0845	0.08	0.0196	0.0433	0.0181	0.0316	0.0268
Std	0.4213	1.0856	0.8877	0.777	0.2444	0.636	0.3138	0.2933	0.493

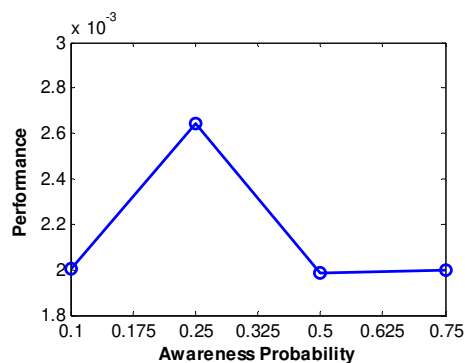


Fig. 2. The impact of varying AP values on CSA performance.

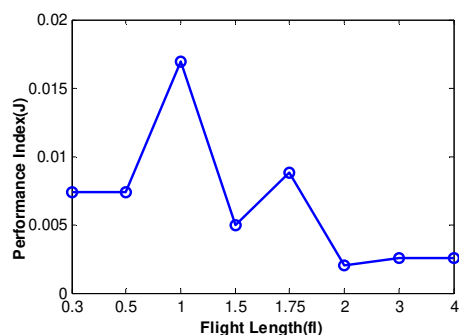
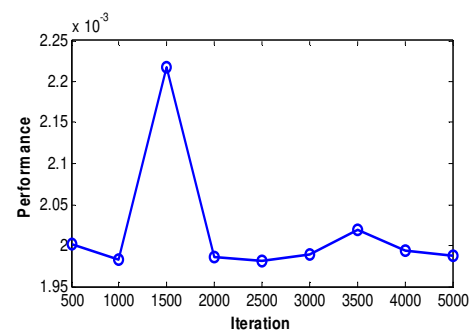
Fig. 3. Effect of  $fl$  on the performance of the CSA.

Fig. 4. Effect of iteration number on the performance of CSA.

From Table III we can note that by increasing the iteration number more than 2000 and by keeping the AP as 0.1 and  $fl$  as 2, the performance remains almost the same, showing poor performance which can be clearly seen in Figure 4.

#### B. Comparison Analysis of DE, PSO, GWO, and CSA

Optimal control problem can be solved by using the CSA parameters as shown in Table IV. The statistical findings produced using CSA are compared in Table V. In comparison to the other methods on the SLFM, the results suggest that CSA delivers promising results. When it comes to the best index, CSA surpasses DE, PSO, and GWO. In this example, CSA's results are roughly 45% more accurate than the other methods used in terms of performance.

TABLE IV. UTILIZED CSA PARAMETERS

Parameter	Value
Flight Length	2
Awareness probability	0.1
Number of population	50
Number of iteration	450

TABLE V. SLFM RESULT COMPARISON ACHIEVED BY CSA AND OTHER ALGORITHMS

Algorithm	Worst	Mean	Best	Std.	Convergence iteration
DE	1.3561	0.039259	0.0041536	0.15383	435
PSO	1.8147	0.040232	0.003053	0.19363	363
CSA	5.1851	0.0967	0.002109	0.4213	261
GWO	0.0709	0.0515	0.042707	0.0073	448

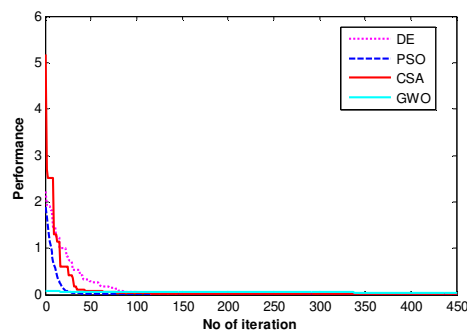


Fig. 5. Algorithm comparison graph.

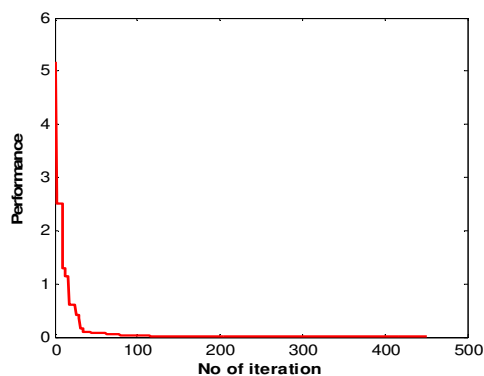
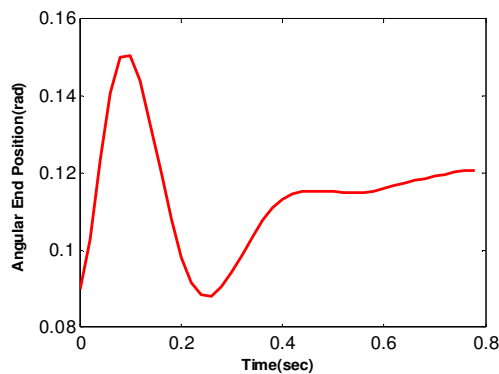


Fig. 6. Convergence graph of CSA of SLFM.

Fig. 7. Arm axis position ( $z_1$ ) versus time (s) of CSA.

CSA shows encouraging performance in terms of worst and mean indices. The std indicates the CSA's substantial flexibility. The CSA's performance index value is minimum. The convergence rate of the CSA for identifying the optimum SLFM solution is shown in Figures 5 and 6. The CSA algorithm exhibits an excellent convergence rate.

Figure 7 explains the arm end position of the SLFM which became stable 0.12 at nearly 0.7s and minimum position 0.088 at nearly 0.26s. Figure 8 shows the comparison graph between the CSA, DE, PSO, and BWO and it is clearly observed that there is more disturbance in GWO, which is unstable.

TABLE VI. COMPARISON OF DE, PSO, GWO, AND CSA FOR ARM END POSITION

Value	CSA	DE	GWO	PSO
Maximum	0.1506	0.1238	1.4439	0.1239
Minimum	0.0882	5.5226e-4	-0.2679	4.9009e-4
Mean	0.1143	0.0293	0.5935	0.0290

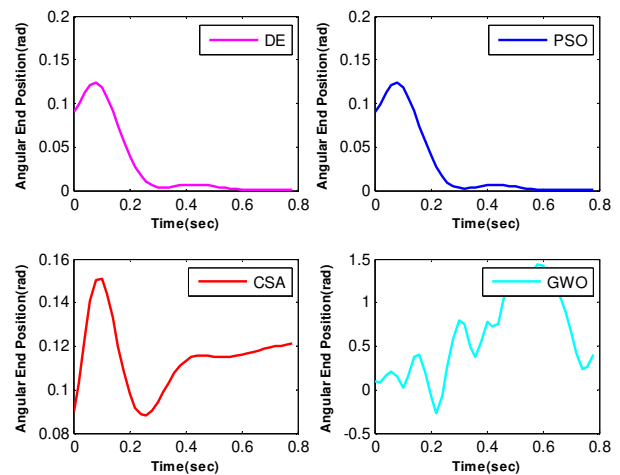


Fig. 8. Comparison graph of arm axis position for DE, PSO, GWO, and CSA.

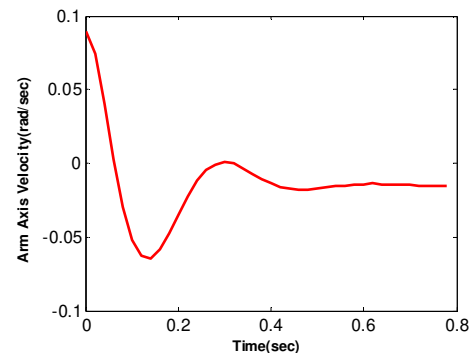
Fig. 9. Arm axis position ( $z_2$ ) versus time (s) of CSA.

Table VI shows the minimum, maximum, and mean positions attained by the algorithm comparison. DE, PSO, and CSA gave almost the same results of attained angular position, which can be observed clearly in Figure 8.

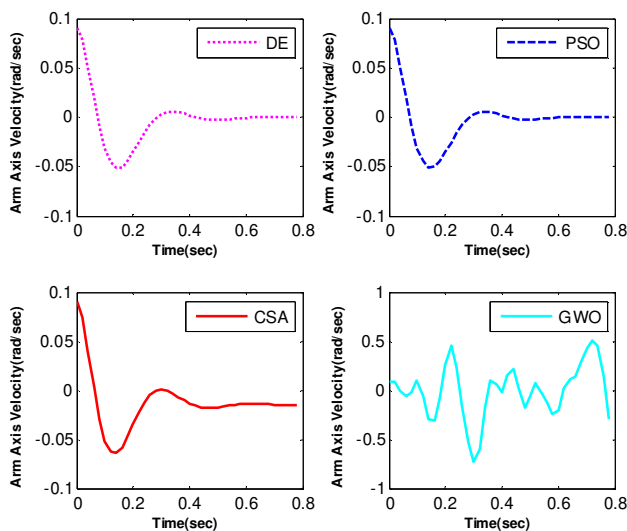


Fig. 10. Comparison graph of arm axis velocity for DE, PSO, GWO, and CSA.

Figure 9 explains the arm axis velocity of the SLFM. It can be seen that less velocity was required to move the SLFM. Figure 10 shows the comparison graph between the CSA, DE, PSO, and GWO. Table VIII shows the minimum, maximum, and mean velocity attained by the algorithm comparison.

TABLE VII. COMPARISON OF DE, PSO, GWO, AND CSA FOR ARM AXIS VELOCITY

Value	CSA	DE	GWO	PSO
Maximum	0.0900	0.0900	0.5039	0.0900
Minimum	-0.0644	-0.0515	-0.7292	-0.0510
Mean	-0.0127	-0.0018	-0.0070	-0.0017

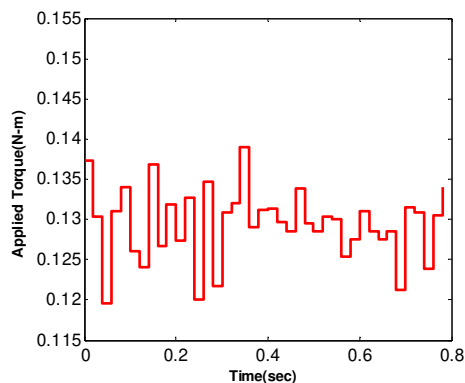


Fig. 11. Applied torque versus time curve.

TABLE VIII. COMPARISON OF DE, PSO, GWO, AND CSA FOR THE APPLIED TORQUE

Value	CSA	DE	GWO	PSO
Maximum	0.1389	0.0018	10	0.0528
Minimum	0.1196	-0.1874	-9.9246	-0.1862
Mean	0.1295	-0.0398	0.0805	-0.0385

Table VIII compared the minimum, maximum, and mean torque applied to SLFM by the considered algorithms. Figure 11 gives an idea of how much torque is applied to SLFM and

we can see that for a small torque driven by SLFM, it attains its position. Figure 12 shows the comparison graph between the CSA, DE, PSO, and GWO, in the presence of negative torque which leads the manipulator to brake on in order to move to the desired position. At the same time, CSA has a positive torque which leads the manipulator to move to the desired position during the simulated period of time.

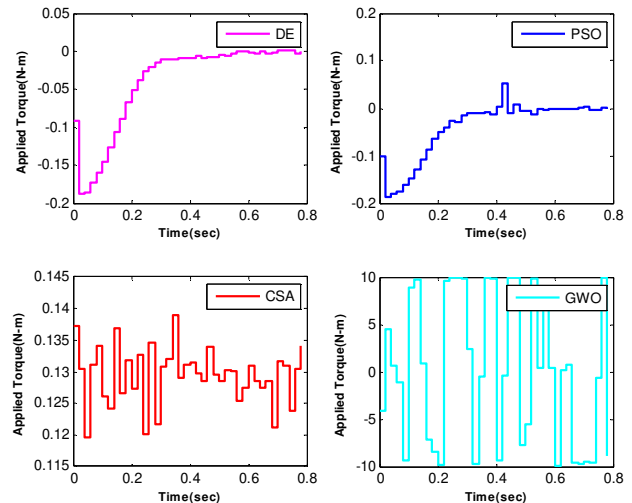


Fig. 12. Comparison of DE, PSO, GWO, and CSA for the applied torque and time.

## V. CONCLUSION

In this paper, Crow Search Algorithm (CSA), is proposed to solve OCP, i.e. SLFM. The two regulating factors in CSA are flight duration and awareness probability. CSA is simpler to utilize and has fewer factors to adjust than DE, PSO, and GWO. The simulation results show that the CSA performance is promising since it provided results that were competitive with those of the other methods. Despite the reputation of the PSO as a quick approach among population-based algorithms, CSA's performance surpasses PSO along with the other algorithms.

## REFERENCES

- [1] M. I. Ullah, S. A. Ajwad, M. Irfan, and J. Iqbal, "Non-linear Control Law for Articulated Serial Manipulators: Simulation Augmented with Hardware Implementation," *Elektronika ir Elektrotechnika*, vol. 22, no. 1, pp. 3–7, Feb. 2016, <https://doi.org/10.5755/j01.eec.22.1.14094>.
- [2] G. Beni and J. Wang, "Swarm Intelligence in Cellular Robotic Systems," in *Robots and Biological Systems: Towards a New Bionics?*, Berlin, Heidelberg, Germany, 1993, pp. 703–712, [https://doi.org/10.1007/978-3-642-58069-7\\_38](https://doi.org/10.1007/978-3-642-58069-7_38).
- [3] X.-S. Yang, "Metaheuristic Optimization," *Scholarpedia*, vol. 6, no. 8, Aug. 2011, Art. no. 11472, <https://doi.org/10.4249/scholarpedia.11472>.
- [4] K. Hussain, M. N. Mohd Salleh, S. Cheng, and Y. Shi, "Metaheuristic research: a comprehensive survey," *Artificial Intelligence Review*, vol. 52, no. 4, pp. 2191–2233, Dec. 2019, <https://doi.org/10.1007/s10462-017-9605-z>.
- [5] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Computing Surveys*, vol. 35, no. 3, pp. 268–308, Jun. 2003, <https://doi.org/10.1145/937503.937505>.



- [6] N. Razmjoo, M. Ramezani, and A. Namadchian, "A New LQR Optimal Control for a Single-Link Flexible Joint Robot Manipulator Based on Grey Wolf Optimizer," *Majlesi Journal of Electrical Engineering*, vol. 10, no. 3, Oct. 2016, Art. no. 53.
- [7] B. Farzanegan, S. Dehghan Banadaki, and M. B. Menhaj, "Direct artificial neural network control of single link flexible joint," in *2016 4th International Conference on Control, Instrumentation, and Automation (ICCIA)*, Qazvin, Iran, Jan. 2016, pp. 131–135, <https://doi.org/10.1109/ICCIAutom.2016.7483149>.
- [8] K. Groves and A. Serrani, "Modeling and Nonlinear Control of a Single-link Flexible Joint Manipulator," 2004.
- [9] V. S. Durga Manohar Sahu, P. Samal, and C. K. Panigrahi, "Application of Differential Evolution Algorithm to Optimal Control Problem with State Variable Constraints," in *2020 IEEE 17th India Council International Conference (INDICON)*, New Delhi, India, Sep. 2020, <https://doi.org/10.1109/INDICON49873.2020.9342273>.
- [10] N. Mishra, S. P. Singh, and B. C. Nakra, "Dynamic analysis of a single link flexible manipulator using Lagrangian-assumed modes approach," in *2015 International Conference on Industrial Instrumentation and Control (IIC)*, Pune, India, Feb. 2015, pp. 1144–1149, <https://doi.org/10.1109/IIC.2015.7150920>.
- [11] J. Iqbal, "Modern Control Laws for an Articulated Robotic Arm: Modeling and Simulation," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 4057–4061, Apr. 2019, <https://doi.org/10.48084/etasr.2598>.
- [12] H. Medjoubi, A. Yassine, and H. Abdelouahab, "Design and Study of an Adaptive Fuzzy Logic-Based Controller for Wheeled Mobile Robots Implemented in the Leader-Follower Formation Approach," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6935–6942, Apr. 2021, <https://doi.org/10.48084/etasr.3950>.
- [13] M. Fouzia, N. Khenfer, and N. E. Boukezzoula, "Robust Adaptive Tracking Control of Manipulator Arms with Fuzzy Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6131–6141, Aug. 2020, <https://doi.org/10.48084/etasr.3648>.
- [14] K. L. Teo, "A unified computational approach to optimal control problems," in *A unified computational approach to optimal control problems*, De Gruyter, 2011, pp. 2763–2774, <https://doi.org/10.1515/9783110883237.2763>.
- [15] M. S. Alam and M. O. Tokhi, "Dynamic Modelling of a Single-Link Flexible Manipulator System: A Particle Swarm Optimisation Approach," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 26, no. 1, pp. 57–72, Mar. 2007, <https://doi.org/10.1260/026309207781487466>.
- [16] F. Glover, "Future paths for integer programming and links to artificial intelligence," *Computers & Operations Research*, vol. 13, no. 5, pp. 533–549, Jan. 1986, [https://doi.org/10.1016/0305-0548\(86\)90048-1](https://doi.org/10.1016/0305-0548(86)90048-1).
- [17] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, Dec. 1997, <https://doi.org/10.1023/A:1008202821328>.
- [18] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia, Aug. 1995, vol. 4, pp. 1942–1948, <https://doi.org/10.1109/ICNN.1995.488968>.
- [19] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm," *Computers & Structures*, vol. 169, pp. 1–12, Jun. 2016, <https://doi.org/10.1016/j.compstruc.2016.03.001>.
- [20] G. Lindfield and J. Penny, *Introduction to Nature-Inspired Optimization*. Elsevier Academic Press, 2017, <https://doi.org/10.1016/C2015-0-00099-5>.
- [21] B. Subudhi and A. S. Morris, "Soft computing methods applied to the control of a flexible robot manipulator," *Applied Soft Computing*, vol. 9, no. 1, pp. 149–158, Jan. 2009, <https://doi.org/10.1016/j.asoc.2008.02.004>.
- [22] V. S. D. M. Sahu, P. Samal, and C. K. Panigrahi, "Application of Crow Search Algorithm to Solve Discrete Optimal Control Problem with Constraints for a Flexible manipulator," in *2022 3rd International Conference for Emerging Technology (INCET)*, Belgaum, India, Feb. 2022, <https://doi.org/10.1109/INCET54531.2022.9825392>.
- [23] V. Satya Durga Manohar Sahu, P. Samal, and C. Kumar Panigrahi, "Modelling, and control techniques of robotic manipulators: A review," *Materials Today: Proceedings*, vol. 56, pp. 2758–2766, Jan. 2022, <https://doi.org/10.1016/j.matpr.2021.10.009>.
- [24] B. Zolghadr-Asli, O. Bozorg-Haddad, and X. Chu, "Crow Search Algorithm (CSA)," in *Advanced Optimization by Nature-Inspired Algorithms*, O. Bozorg-Haddad, Ed. Singapore: Springer, 2018, pp. 143–149, [https://doi.org/10.1007/978-981-10-5221-7\\_14](https://doi.org/10.1007/978-981-10-5221-7_14).

# Static and Seismic Stability of a Slope Reinforced with Two Rows of Piles

**Lokmane El Hakim Chekroun**

Department of Civil Engineering, Faculty of Technology, EOLE Research Laboratory, University of Tlemcen, Algeria  
lokmaneelhakim.chekroun@univ-tlemcen.dz  
(corresponding author)

**Nadir Boumechra**

Department of Civil Engineering, Faculty of Technology, EOLE Research Laboratory, University of Tlemcen, Algeria  
nadir.boumechra@univ-tlemcen.dz

*Received: 4 November 2022 | Revised: 24 November 2022 | Accepted: 27 November 2022*

## ABSTRACT

The use of piles for slide stabilization is considered among the most important innovative reinforcement techniques. Piles have been successfully used in many situations to stabilize slopes or as parts of a stability improvement. Our case study is in the Tlemcen section. The left side of the roadway collapsed following the slide of the downstream side. Inclinator readings showed signs of instability with a slip depth of about 9m near the motorway platform. The likely causes of the instability were the removal of the lower abutment of the embankment upstream of the road and the establishment of an earth deposit which overloaded the embankment and disrupted the flow of water downstream. The stabilization study is based on the installation of two rows of anti-slip piles. Stability analysis study was carried out under static and dynamic loads and highlights that this solution is advantageous and effective.

*Keywords-pile behavior; dynamic study; finite elements; slope; highway*

## I. INTRODUCTION

Many methods have been developed for the analysis of piles [1-4]. The limit equilibrium method was used in [3] to deal with the problem of stability of embankments containing piles. The safety factor of a slope reinforced by piles was defined as the ratio of the moment of resistance to the moment of overturning (engine) acting on the mass of the potentially unstable ground. The time of resistance consists of two components: the moment due to the resistance of the ground to the shear along the sliding surface and the moment provided by the reaction force piles. The driving moment and the moment of soil shear resistance were obtained by the simple slice method. The study of a slope includes, in addition to the recognition of the site and the choice of the mechanical characteristics of the soils, stability calculation in order to determine not only the breaking curve along which the risk of landslide is the highest, but also the corresponding value of the safety coefficient [5]. It should be noted that ground movements are very varied in nature and size. The problems of slope stability are recorded in a general way in the realization of roads, dams, and natural slopes. The landslide passes through several chronological stages of activity. There are major factors that control the type and rate of mass movement that could occur.

The methods of slope stabilization by the so-called anti-slide piles have attracted the interest of several researchers and many studies have been conducted on slopes introducing stabilization piles. We distinguish 2 methods of calculation: limit equilibrium methods and numerical methods. The kinematic approach of treating the slope with one or more rows of piles called "passive piles" has the advantage of improving its stability [1]. We note that the major problem in the design of the slope-pile system is the determination of the appropriate and effective location of the piles. The influence of pile reinforcement on the stability of the slope was studied in [6]. A 2D numerical modeling was used to validate the proposed approach by comparing the numerical results with the field measurements. The effect of pile position as a function of soil shear parameters was also studied. The results indicate that the ideal position for this type of stabilization piles is in the average height of the slope. The slope should be between 30° and 60°.

It should be noted that this study is focused on a problem that occurred at the section of PK 210+480 to 210+800 of the East-West Algerian highway near the Didouche Mourad, Constantine. The results of this study confirm that soil suction has a significant effect on slope stability. Negative pressures from pore water in the slope must be taken into account. It is

imperative that an appropriate drainage system is in place and that it is properly maintained.

Authors in [7] conducted an experimental and numerical study of the stability of a slope reinforced with a concrete pile. This study showed that when the pile is located in the middle of the slope, the soil structure collapses under a pressure of about 10.9kPa, which is the highest pressure that caused the instability of the slope consisting of reinforced sand. However, when the pile is located upstream or downstream, slope failure occurs under a pressure of 7.8 or 3.12kPa, respectively. Therefore, his work showed that a pile located in the middle of the slope can improve the reinforcement of the soil, providing it with optimal stability. Authors in [8] investigated the stability of a slope under 3 configurations, i.e. without protection, with one row of piles, and with two rows of piles. This study focused on the analysis of the sliding mode, the distribution of bending moments, and the factor of safety of the slope. For the single-row stabilization pile reinforced slope, the optimal location of the piles is in the lower and middle part of the slope where the factor of safety takes a maximum value of 1.26. For the two-row stabilization pile reinforced slope, the appropriate location of the piles is in the lower and lower middle part of the slope. The factor of safety in this case has a maximum value of 1.38. The bending moment and thrust on the double-row piles are more reasonable than the other configurations.

## II. PRESENTATION OF THE LANDSLIDE OF HIGHWAY AT KILOMETER POINT 52

In this article, we study the landslide that occurred on March 2, 2014 on a section of the East-West highway near the city of Tlemcen (north western of Algeria). The platform at PK 52+040 to PkK52+220 underwent major deformations, with the rear wing moving vertically, i.e. by 3 to 4m horizontally. As a result, the left side of the roadway was completely closed to traffic and the outer part was completely deformed, being subject to sliding, as shown in Figure 1.



Fig. 1. Deformation of the pavement of highway at Pk52.

According to the on-site observation, the landslide was rapidly developed while dense cracks were found in the body of the landslide not far from the river and there were clear deformations in the form of shearing. The vertical displacement of the slide platform was approximately 2m on an embankment section. Observation of the core samples showed a highly

variable lithology, with high pebble content in the boreholes taken. According to the order of destruction of the inclinometer boreholes, the borehole next to the bank of the river was damaged first and the destroyed borehole next to the platform was next, so the slip was caused by a bottom-up pull. Furthermore, there is no evidence of deformation in the hole on the right side of the route. It should also be noted that the scouring of the river water on the front edge of the landslide quickly caused the landslide to appear. One day before, i.e. March 1, 2014, torrential rains that fell on the region caused water infiltration and the rise in the level of the underground water table in the body of the slope. This lubricating action accelerated the mechanism of the landslide phenomenon. Figure 2 shows the evolution of displacements over time.

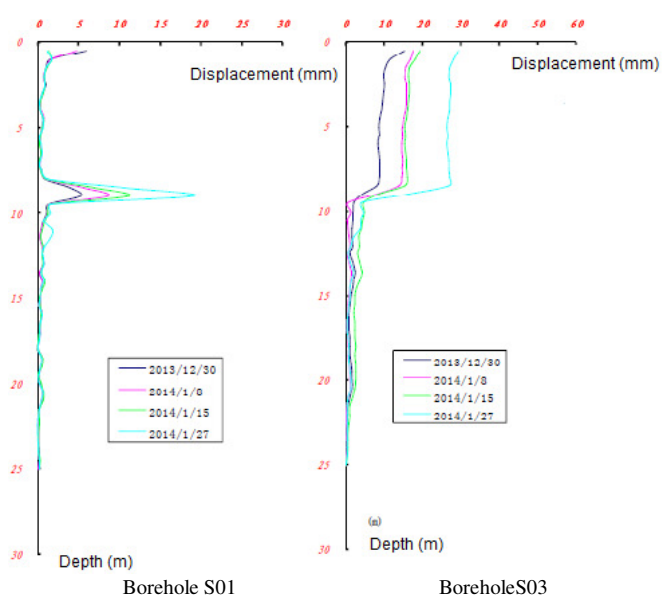


Fig. 2. The evolution of movements over time for boreholes S01 and S03.

## III. GEOLOGICAL AND GEOTECHNICAL PROFILE OF THE SITE

The geotechnical profile of the site was determined by reconnaissance, with 6 boreholes to a total depth of 127.4m. The determined geotechnical profile of the site shown in Figure 3 shows that the area has a sloping profile. The soil layers are made up of a bedding fill of medium density, followed by a layer of brownish to yellowish silty clay, distributed mainly in the left embankment of the route. The soil is homogeneous, interspersed with a few pebbles and sand, very plastic, but not locally. Overlying a layer of completely weathered sandstone distributed mainly on the bank of the river, at the foot of the left embankment and at the foot of the right embankment, a layer of completely altered marl of yellow-green color with a clay structure, interspersed with very plastic sand. The water level varies in depth from 13 to 24.5m in the various boreholes drilled, with a tendency to flow in the direction of the slope. Analysis of the inclinometer reading data showed that the slip surface was at 9m, 7m, and 3m as shown in Table I.

TABLE I. VALUES OF THE INCLINOMETER READING [9]

Borehole number	Start date of measurement	End date of measurement	Slip depth (m)	Final displacement (mm)
S01	13-12-2013	29-12-2014	9	26.99
S02	18-12-2013	30-12-2013	7	-
S03	25-12-2013	02-03-2014	9	58.81

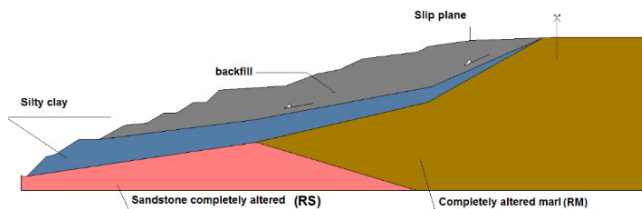
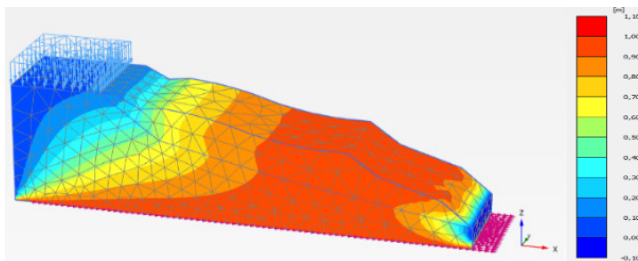


Fig. 3. Geotechnical profile of the site.

The results of the displacements recorded at the foot of the slope on the right side of the platform, showed no signs of significant deformation for 3 months after the event day.

#### IV. NUMERICAL MODEL OF THE LANDSLIDE

The static calculations led to the assumption that the main cause of the landslide, in addition to the local geomorphological and geotechnical conditions of the site and the presence of the water table, is the head loading due to the motorway traffic. This can be neglected or at least not properly taken into account in the pre-dimensioning. Figure 4 shows the distribution of horizontal displacements in the model. They are maximal at the bottom of the slope with values of 1.00m.

Fig. 4. Horizontal displacement fields ( $U_x=1.00m$ ).

##### A. Static Analysis

The landslide section for the solution chosen for the treatment [10-12] consists of the installation of 11 piles in a single row at the edge of the landslide (in the middle of the alignment) in order to ensure the safety of the motorway users, during the emergency works, and then the realization of a row of 60 piles on the left side of the section PK52+040~PK52+220, with a beam connecting them at the head.

##### 1) Static Analysis without the Effect of Rainfall

The model area of the massif extends laterally for 30m, and to a depth of 40m. For the boundary conditions, the vertical and horizontal displacements at the model boundaries are assumed to be zero. The soil-pile system is discretized using the 3-dimensional mesh options of the software into 23010 elements. Each finite element has 10 nodes with 3 degrees of freedom for

a total of 35295 nodes. The average element size is 2.797m [3-5, 7-19]. It should be noted that the bored piles are modelled by a linear elastic behavior law. The soil is modelled by an elastoplastic behavior law of the Mohr-Coulomb type. Figure 5 shows this proposed treatment of slip which was statically modelled in plane deformation using Plaxis 3D software. The piles have a diameter of 1.20m while the depth is 15m for the first row of 11 piles and 19-23m for the second row of 60 piles spaced at 3m. The nominal strength of the concrete used for the piles is 35MPa while the modulus of elasticity is  $E=35982\text{MPa}$ . The density of the concrete is  $25\text{kN/m}^3$ .

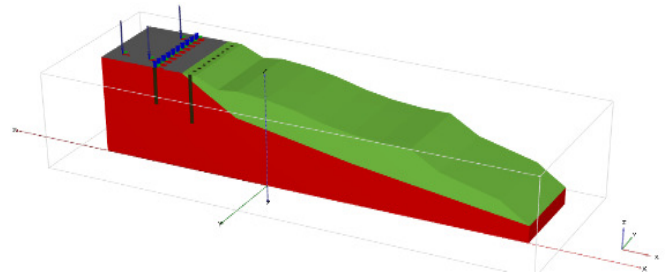


Fig. 5. Numerical model after the introduction of 2 rows of piles.

The analysis of the stability of the slide after the introduction of the 2 rows of anti-slip piles clearly shows that the slope is more stable than the initial configuration with an increase in the safety coefficient. Therefore, the pavement is now in a safe condition.

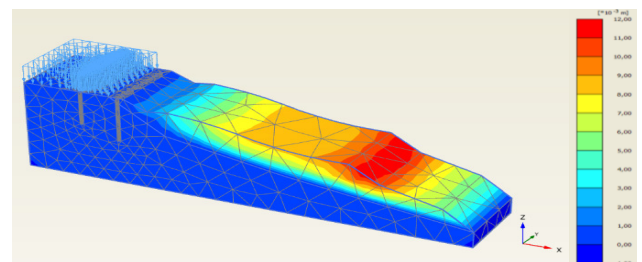
Fig. 6. Displacement fields  $U_x$  ( $U_x \text{ max} = 0.01186m$ ,  $F_s=1.750$ ).

Figure 6 shows that the model was really stabilized after the introduction of the two rows of piles and the displacements are almost zero. The largest displacements are of the order of only 1cm at the bottom of the slope.

##### 2) Static Analysis with the Effect of Rainfall

Four precipitation values were taken as follows:  $Q = 0$ , 0.53, and  $5.30\text{m}^3/\text{day}$ . This last value corresponds to the rainfall recorded during January in Tlemcen (i.e. 62mm). We note that the section of PK 52 that has undergone the slip makes a section of  $2700\text{m}^2$ , or 27% of the area taken in the rainfall estimate. Indeed, a 1mm represents the rainfall of  $10.000\text{m}^2$  on an area of  $100\text{m} \times 100\text{m}$  (1 hectare). Figure 7 shows an example showing the introduction of rainfall in the calculation code used: 3 simulations were carried out by changing the precipitation value in order to inspect the effect of this parameter in the evolution of the stresses. The Figure shows their changes according to the calculation steps, in a point B located at the bottom of the slope.

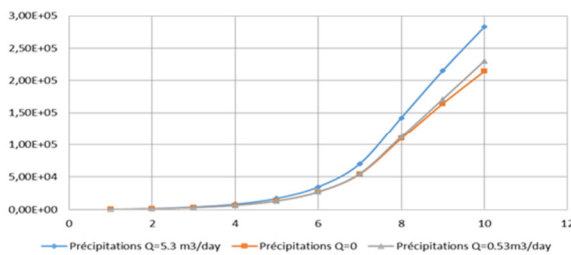


Fig. 7. Evolution of stress  $\sigma_{xx}$  according to the rainfall flow.

Figure 7 shows the rainfall which significantly influences the evolution of stresses under static loading. The recorded increase in stresses is 32%, passing from zero precipitation to  $5.30\text{m}^3/\text{day}$ .

### B. Seismic Analysis

The seismic loading, defined by an accelerogram that was taken into account for the dynamic analysis, is that of the Boumerdes earthquake of May, 21 2003, characterized by a magnitude  $M_w$  of 6.8 on the Richter scale. Strong seismic aftershocks were recorded, in particular at the Keddara site, located south-east of the capital Algiers and 20km from the epicenter. The proposed landslide treatment was also the subject of dynamic modeling in plane deformation using Plaxis [20]. The landslide was subjected to a combination of loads (permanent load, operational load, and seismic load from the Boumerdes 2003 earthquake illustrated in Figure 8).

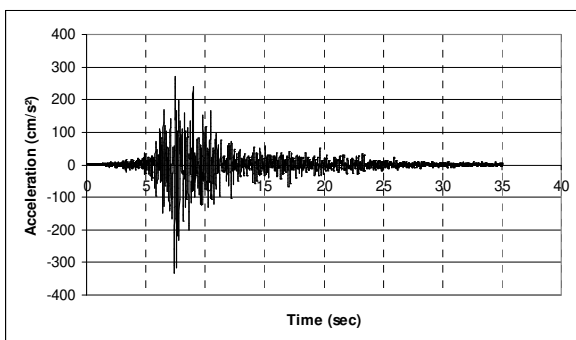


Fig. 8. The applied accelerogram of the Boumerdes 2003 earthquake.

The Peak Ground Acceleration (PGA) reached  $0.58g$  in the East-West component (Keddara2 station) and  $0.35g$  (Keddara1 station) in the North-South component of the horizontal direction, while it was  $0.22g$  in the vertical component [15, 21, 22]. The data for these events were obtained from the network records of the National Center of Seismic Engineering (CGS) in Algiers. Although various simulations were performed, we only present the deformation of the configuration at the end of the seismic loading ( $t = 35s$ ). We also note that several runs were carried out, changing the model configuration each time.

#### 1) Model with a Single Row of Piles with Length of 15m

In order to analyze the influence and impact of the pile string located at the edge of the landslide (in the middle of the alignment), only one pile string was introduced [14]. It should be noted that this row was planned as an emergency solution.

Figure 9 shows the location of this pile string at the level of the highway platform.

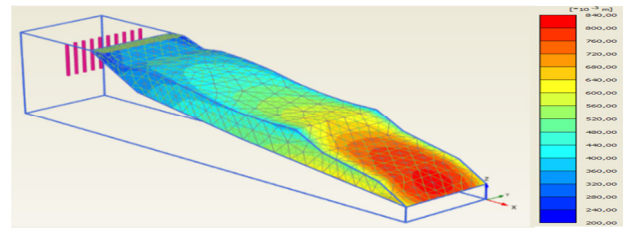


Fig. 9. The introduced pile row (emergency treatment) ( $\text{Max } U_x = 0.8216\text{m}$ ).

#### 2) Model with Two Rows of Piles

The analysis of the slip stability was conducted after the introduction of 2 rows of anti-slip piles subjected to dynamic loading. Figure 10 shows the deformation of the model of the pile system.

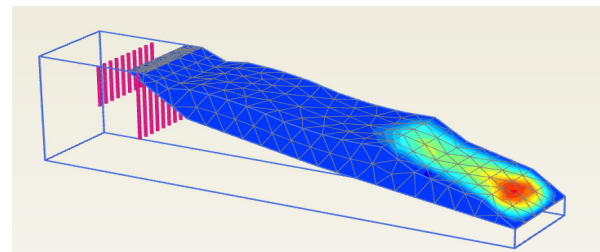


Fig. 10. Representation of the system model with two piles.

The dynamic study of this model shows that it is the second row of piles, with a length of  $l = 20.00\text{m}$ , that provides significant stabilization of the slide. This results in a noticeable decrease in displacement, mainly at the motorway section. Note that the maximum value, which is  $16.21\text{cm}$ , was recorded at the lower part of the model as shown in Figure 10.

#### 3) Numerical Model with Rock at the Foot of the Slope

Although rock was not used, this modeling was initiated in order to see the rock impact on the overall behavior of the model. The new model configuration is shown in Figure 11.

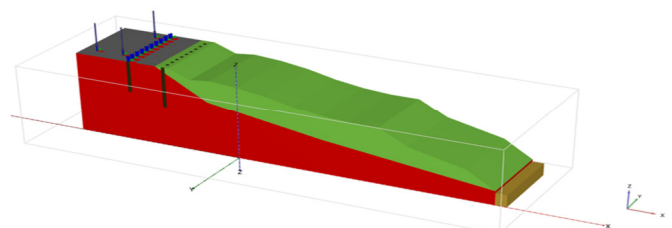


Fig. 11. Numerical model after the introduction of rock at the foot of the slope.

#### 4) Effect of Rainfall

A parametric study was carried out, by varying the rainfall rate, the considered values were 0, 0.53, 2.4, and  $5.40\text{m}^3/\text{day}$



[18]. The seismic loading accelerogram taken into account for the dynamic analysis is that of the Boumerdes earthquake 2003. The main objective of the study of this parameter lies in the estimation of the stresses exerted on the stabilizer piles.

## V. RESULTS AND DISCUSSION

### A. Under Static Loading

Figure 12 shows the evolution of the displacements under static loading from the exploitation load. His value is equal to  $10\text{ kN/m}^2$ . Figure 13 shows clearly the very significant decrease in the stresses under the effect of the 2 rows of piles, after having reached a maximum value of  $2.10^5\text{ kN/m}^2$  during the landslide.

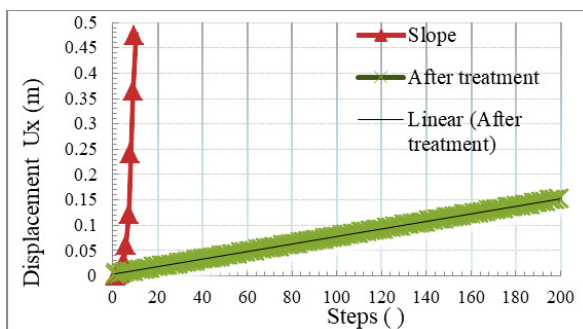


Fig. 12. Evolution of displacements  $U_x$  in a point A of the highway's platform.

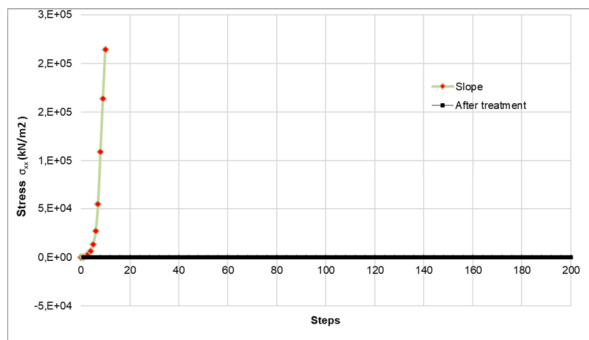


Fig. 13. Stresses  $\sigma_{xx}$  in a point A of the highway platform.

### B. Under Seismic Loading with the Effect of Rainfall

The two graphic representations in Figure 14 and 15 show that under dynamic loading the stresses have a very short duration. The evolution of the stresses, under the effect of the rainfall is not significant and resulted in very close curves for the case of two distinct points.

TABLE II. SIMULATION RESULT SUMMARY

Model configurations	Displacement $U_x$ (cm)	Stress $\sigma_{xx}$ ( $\text{kN/m}^2$ )	Safety factor ( $F_s$ )
Before treatment	100.00	$2.10 \text{ e}+5$	$<1.00$
1 row of piles (seismic case)	82.16	/	1.00
2 rows of piles (seismic case)	16.21	0	1.20
2 rows of piles (static case)	1.186	/	1.75

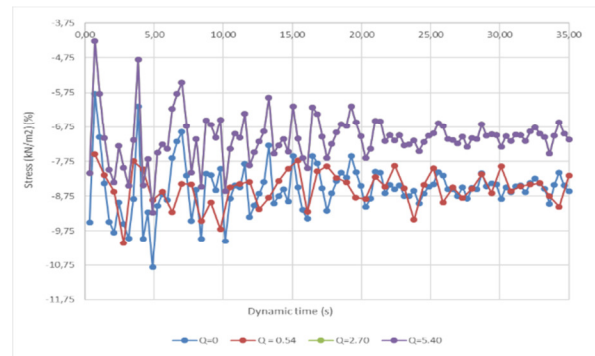


Fig. 14. Evolution of stresses  $\sigma_{xx}$  in point Q at the first row of piles.

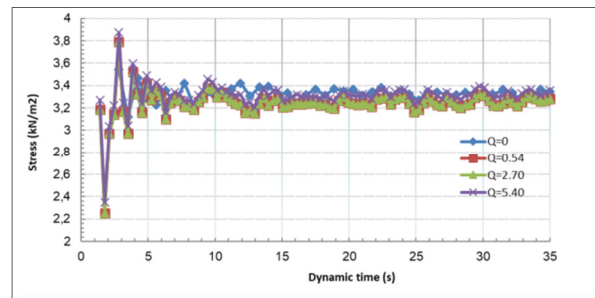


Fig. 15. Evolution of stresses  $\sigma_{xx}$  at the head of the second row of piles.

Table IV summarizes the results found in terms of displacements, stresses, and safety coefficients for the different configurations of the numerical model.

## VI. CONCLUSION

In this article, the stability study of the landslide that occurred on March, 2, 2014 on a section of the East-West highway near the city of Tlemcen (North-West Algeria) was carried out. Vertical piles were used to stabilize the highway on the right-hand side of the PK 52 section in the boundary between the cities of Tlemcen and Sidi Bel-Abbes.

It should be noted that the company carrying out the treatment work did not rely on the analysis when establishing the treatment of this slide. Nevertheless, the calculations carried out in this work show that the treatment solution presents more stability compared to the initial state of the section that suffered the slide. This is justified by the maximum value of the displacement recorded, which is 16.21cm under dynamic loading and almost zero under static loading.

We would like to report that it is the second row of 61 piles that brings more stability to the slip section. The first one, consisting of 11 piles, was planned as an emergency solution. On the other hand, the introduction of rock at the bottom of the slope provides optimum stability for the model, essentially by reducing both stresses and displacements. It should be noted that this solution was not foreseen in the field. We note that the dynamic analysis based on the accelerogram of a very strong seismic movement of Boumerdes 2003, of the PK 52 slope shows the effectiveness of the slope stabilization technique by piles. This method is able to solve in a permanent way the phenomenon of the slides.

Rainfall remains a crucial parameter in the treatment of sliding stability, essentially under static loading. Particular attention must be paid to this parameter in order to have a correct digital model that represents the real case. This technique is mainly applicable to soils resting on coherent, sometimes soft or sensitive soil.

## REFERENCES

- [1] M. Ashour and H. Ardalan, "Analysis of pile stabilized slopes based on soil-pile interaction," *Computers and Geotechnics*, vol. 39, pp. 85–97, Jan. 2012, <https://doi.org/10.1016/j.compgeo.2011.09.001>.
- [2] L. T. Chen and H. G. Poulos, "Piles Subjected to Lateral Soil Movements," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 123, no. 9, pp. 802–811, Sep. 1997, [https://doi.org/10.1061/\(ASCE\)1090-0241\(1997\)123:9\(802\)](https://doi.org/10.1061/(ASCE)1090-0241(1997)123:9(802)).
- [3] R. Kourkoulis, F. Gelagoti, I. Anastasopoulos, and G. Gazetas, "Slope Stabilizing Piles and Pile-Groups: Parametric Study and Design Insights," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 137, no. 7, pp. 663–677, Jul. 2011, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0000479](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000479).
- [4] R. Liang and S. Zeng, "Numerical Study of Soil Arching Mechanism in Drilled Shafts for Slope Stabilization," *Soils and Foundations*, vol. 42, no. 2, pp. 83–92, Apr. 2002, [https://doi.org/10.3208/sandf.42.2\\_83](https://doi.org/10.3208/sandf.42.2_83).
- [5] W. Gong, H. Tang, H. Wang, X. Wang, and C. H. Juang, "Probabilistic analysis and design of stabilizing piles in slope considering stratigraphic uncertainty," *Engineering Geology*, vol. 259, Sep. 2019, Art. no. 105162, <https://doi.org/10.1016/j.enggeo.2019.105162>.
- [6] K. Boulfoul, F. Hammoud, and K. Abbeche, "Numerical study on the optimal position of a pile for stabilization purpose of a slope," *Geomechanics and Engineering*, vol. 21, no. 5, pp. 401–411, 2020, <https://doi.org/10.12989/gae.2020.21.5.401>.
- [7] M. Hajiazizi, M. Bavali, and A. Fakhimi, "Numerical and Experimental Study of the Optimal Location of Concrete Piles in a Saturated Sandy Slope," *International Journal of Civil Engineering*, vol. 16, no. 10, pp. 1293–1301, Oct. 2018, <https://doi.org/10.1007/s40999-017-0155-1>.
- [8] H. Lei, X. Liu, Y. Song, and Y. Xu, "Stability analysis of slope reinforced by double-row stabilizing piles with different locations," *Natural Hazards*, vol. 106, no. 1, pp. 19–42, Mar. 2021, <https://doi.org/10.1007/s11069-020-04446-2>.
- [9] S. M. E.-A. Bourdim, L. E.-H. Chekroun, A. Benanane, and A. Bourdim, "Treatment of a Landslide by Using Piles System, Case Study of the East-West Highway of Algeria," in *GeoMEast 2017: Soil Dynamics and Soil-Structure Interaction for Resilient Infrastructure*, 2018, pp. 16–24, [https://doi.org/10.1007/978-3-319-63543-9\\_2](https://doi.org/10.1007/978-3-319-63543-9_2).
- [10] T. Ito and T. Matsui, "Methods to Estimate Lateral Force Acting on Stabilizing Piles," *Soils and Foundations*, vol. 15, no. 4, pp. 43–59, Dec. 1975, [https://doi.org/10.3208/sandf1972.15.4\\_43](https://doi.org/10.3208/sandf1972.15.4_43).
- [11] M. A. Soomro, K. F. Memon, M. A. Soomro, A. Memon, and M. A. Keerio, "Single Pile Settlement and Load Transfer Mechanism due to Excavation in Silty Clay," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2485–2492, Feb. 2018, <https://doi.org/10.48084/etasr.1666>.
- [12] H. G. Poulos, "Design of reinforcing piles to increase slope stability," *Canadian Geotechnical Journal*, vol. 32, no. 5, pp. 808–818, Oct. 1995, <https://doi.org/10.1139/t95-078>.
- [13] A. Bouafia, "Comportement des pieux isolés chargés latéralement - Problèmes résolus," in *Calcul pratique des fondations et des soutènements*, 3rd ed., Algiers, Algeria: Editions OPU, 2017, pp. 163–181.
- [14] A. Firoozfar, A. Rostami, H. Ghaderi, H. Zamani, and A. Rostamkhani, "Assessing the Effects of Length, Slope and Distance between Piles on the Bearing Capacity of a Pile Group under Axial Loading in Granular Soil," *Engineering, Technology & Applied Science Research*, vol. 7, no. 5, pp. 1894–1899, Oct. 2017, <https://doi.org/10.48084/etasr.1352>.
- [15] N. Laouami, A. Slimani, Y. Bouhadad, J.-L. Chatelain, and A. Nour, "Evidence for fault-related directionality and localized site effects from strong motion recordings of the 2003 Boumerdes (Algeria) earthquake: Consequences on damage distribution and the Algerian seismic code," *Soil Dynamics and Earthquake Engineering*, vol. 26, no. 11, pp. 991–1003, Nov. 2006, <https://doi.org/10.1016/j.soildyn.2006.03.006>.
- [16] N. Mangi, D. K. Bangwar, H. Karira, S. Kalhor, 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>.
- [17] V. Maralapalle and R. A. Hegde, "An Experimental Study on the Socketed Pile in Soft Rock," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9665–9669, Dec. 2022, <https://doi.org/10.48084/etasr.5338>.
- [18] T. V. Tran, M. T. Trinh, G. Lee, S. Oh, and T. H. V. Nguyen, "Effect of Extreme Rainfall on Cut Slope Stability: Case Study in Yen Bai City, Viet Nam," *Journal of the Korean GEO-environmental Society*, vol. 16, no. 4, pp. 23–32, 2015, <https://doi.org/10.14481/jkges.2015.16.4.23>.
- [19] Y.-S. Chung, "Etude numérique de l'interaction sol-pieu-structure sous chargement sismique," Ph.D. dissertation, Université des Sciences et Technologies de Lille, France, 2000.
- [20] *PLAXIS 3D Tutorial Manual*. 2013.
- [21] S. M. E.-A. Bourdim, N. Boumechra, A. Djedid, and H. Rodrigues, "Effect of spatio-temporal variability of the seismic signal on the dynamic pressure behind retaining walls," *Innovative Infrastructure Solutions*, vol. 7, no. 1, Oct. 2021, Art. no. 35, <https://doi.org/10.1007/s41062-021-00626-6>.
- [22] S. M. A. Bourdim, N. Boumechra, and A. Djedid, "Calibrage du modèle numérique, cas du comportement dynamique d'un système sol-soutènement (Numerical model calibration, case of dynamic behavior of a soil-retaining wall system)," *Journal of Materials and Environmental Science*, vol. 7, no. 3, pp. 1048–1055, 2016.

# A Real-Time Analytic Face Thermal Recognition System Integrated with Email Notification

**Ranjit Singh Sarban Singh**

Centre for Telecommunication Research & Innovation (CeTRI), Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Malaysia  
ranjit.singh@utem.edu.my  
(corresponding author)

**T. Joseph Sahaya Anand**

Sustainable and Responsive Manufacturing, Fakulti Teknologi Kejuruteraan Mekanikal dan Pembuatan, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia  
anand@utem.edu.my

**Siti Aisyah Anas**

Centre for Telecommunication Research & Innovation (CeTRI), Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Malaysia  
aisyah@utem.edu.my

**Biswaranjan Acharya**

Department of Computer Engineering-AI & BD, Marwadi University, India  
biswaacharya@ieee.org

*Received: 21 October 2022 | Revised: 16 November 2022 | Accepted: 19 November 2022*

## ABSTRACT

COVID-19 is a contagious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The disease has spread worldwide, leading to an ongoing pandemic. The most common symptom of COVID-19 is fever which can be detected using various manual screening techniques that have the risk of exposing the personnel. Since the virus has globally spread, a reliable system to detect COVID-19-infected people, especially before entering any premises and buildings, is in high demand. The most common symptom that can be detected is fever, even though people with fever might not have COVID-19. Thus, a real-time analytic face thermal recognition system integrated with email notification that has the capability to scan the person's temperature and simultaneously analyze the measured temperature with the recorded/stored information/data is presented in this paper. The proposed system is also able to send an email notification to the relevant authorities during the real-time analytical process. Besides that, this information is also recorded in the system database for continuous monitoring of the respective person's health status. The development of the proposed system is integrated with a Thermal Module AMG8833, Pi camera, and Raspberry Pi Zero Wireless. The proposed system has been tested and the captured results successfully accomplished the development objectives.

**Keywords-***Covid-19; face recognition; temperature; real-time analysis; email notification*

## I. INTRODUCTION

There is a high demand in face thermal recognition systems due to the COVID-19 pandemic. There are many off-the-shelf face thermal recognition systems available in the market, everyone with its abilities and advantages. To enable effective monitoring which also complies with the Standard Operating

Procedure (SOP), a real-time analytic face thermal recognition system integrated with email notification and having the capability to simultaneously analyze the measured person's temperature and update recorded/stored information/data is presented in this paper. Prior to the development of the proposed system, off-the-shelf systems related to detecting possible symptom of COVID-19 are studied. The studied

systems investigate the aspect of simplicity in terms of product development as well as quick and fast delivering the analyzed and processed information to the relevant authorities. Although the developed system is not able to prevent the COVID-19 pandemic, it can assist in controlling it more effectively and efficiently. Generally, infrared and wireless thermometers are deployed as general medical tools, especially at the entries and exits of buildings. The infrared and wireless thermometers generally use body temperature sensing [1]. Designing and developing systems able to assist the early detection of any fever-related illness so that it can be further examined it is CoVID-19 or something else is a scientific topic of major interest.

In [2], fever is mentioned as the most common symptom of patients with COVID-19. The off-the-shelf studied systems [3-5] use the Pi Camera to visualize the thermal data, where the temperature values are mapped into a color gradient, ranging from blue to red. The blue color is mapped to value 0 representing the lowest temperature, while red color is mapped to value 1023 representing the highest temperature. A real image of the respective person is blended with the saved image to produce the read thermal temperature value. The produced thermal temperature value is displayed in real-time and an email is sent to the person in charge at the monitoring station if any abnormality is detected. In [6], an IoT smart health monitoring system that analyzes the human body temperature, pulse and SpO2 has been developed and integrated into a mobile application for COVID-19 detection. With the integrated system, the patient can immediately seek medical attention if a physician is not available. In [7], a smart helmet is developed to innovatively detect and monitor COVID-19 using a thermal imaging system attached onto the smart helmet. The smart helmet detects high body temperature and sends the measured value into a mobile application, assisting the early detection of COVID-19. Thermal face recognition systems have been proven to be effective in detecting possible COVID-19 affected persons besides from been used to authorize the person's authentication [5]. Authors in [8] utilized the smart home concept to detect if a visiting person is infected with COVID-19. Infrared technology was used to detect the person's temperature and if it was high, an integrated thermal face camera was used to scan and monitor the person's condition at the door. Authors in [9] mentioned that CT-scans and X-ray images can also be used for COVID-19 early detection. The paper also explains that the use of deep learning models [10, 11] could give more precise COVID-19 early detection, but systems with deep learning models is more expensive. Thermal detection is a popular method to screen people during the outburst of the COVID-19 pandemic. The infrared thermal scanning method is also widely used at many places as an early precaution step. Such systems are often able to make buzzer sounds when a person with high temperature is detected during the scanning process.

After reviewing some of the developed systems that use facial thermal measuring for early COVID-19 detection, and to the best of our knowledge, features such as real-time identification/authentication and email notification are not applicable. Hence, the current research project implements real-time facial identification/authentication, then analyzes the

temperature reading and compares it with the preset temperature value before sending an email notification if the temperature is above the preset value. The email notification notifies the relevant authorities about the person's condition, or records the person absence from work. Phyton programming language was used to develop the proposed research project. Thonny application was selected as Python IDE because it is free and open-source and it provides significant features such as code inspection and debugging. The standard libraries used in this project are OpenCV and DLIB. The algorithm development begins with facial detection and face recognition. Then, the AMG8833 Thermal Camera Sensor for facial temperature measuring, recording, and comparing is activated.

The proposed project continuously processes the recorded/stored information/data and automatically updates the stored data into the developed system. Also, the developed and deployed system tends to reduce the risk of infection of manual screening.

## II. RESEARCH METHODOLOGY

The proposed development methodology is divided into two phases. Phase one focuses on the hardware development which consists of 1) the integration of the Pi Camera to perform the facial detection and recognition and 2) the integration of the thermal camera sensor to measure facial temperature. The integration of the Pi Camera is conducted to perform the face detection and recognition via existing images that have been preloaded into the Raspberry Pi Zero Wireless SD-Card storage. Upon the success of the face detection and recognition, the thermal camera sensor is activated to measure the human facial temperature. Then, the recognized person and its measured temperature details are stored into the Raspberry Pi Zero Wireless SD-Card. The second phase focuses on the embedded software development for 1) facial detection and recognition, 2) measurement of the facial temperature, and 3) configuration of the email notification. The facial detection and recognition perform real-time face recognition based on the preloaded images. Upon detection of a human face, the thermal camera sensor measures the respective human facial's temperature and stores it into the SD-Card. While the recognized face and human body temperature information is recorded, the process of comparing the measured temperature with the preset temperature value is performed to detect normal and high temperature. If the measured temperature is above the preset temperature value, an email notification is delivered to the relevant authorities (security, managers, internal clinic, COVID-19 screening team, etc.).

### A. IoT Face Thermal Recognition System Integration

- Pi Camera Integration

In the first phase, the Pi Camera performs facial detection and recognition. The integration is shown in Figure 1. The Pi Camera is connected to the Raspberry Pi Zero Wireless via the Camera Serial Interface (CSI) connector. The integrated Pi Camera has the capability to capture an 8 Megapixel real-time image which will be used to perform the facial detection and recognition with the images stored in the Raspberry Pi Zero Wireless SD-Card Storage.

### • Integration of the AMG8833 Thermal Camera Sensor

This section describes the AMG8833 Thermal Camera Sensor connectivity with the Raspberry Pi Zero Wireless as shown in Figure 2. The General-Purpose Input/Output (GPIO) port 2 and port 3 are respectively connected to the Serial Information (SDA) and Serial Clock Line (SCL). The power source (VCC) is connected to 3.3V at port 1 and the ground is connected to port 9. The AMG8833 Thermal Camera Sensor is activated when the facial detection and recognition is completed.

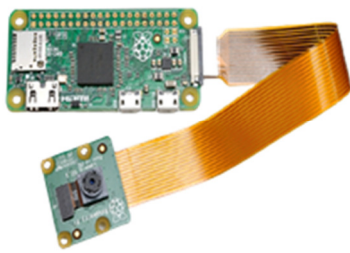


Fig. 1. Pi Camera and Raspberry Pi Zero Wireless connection.

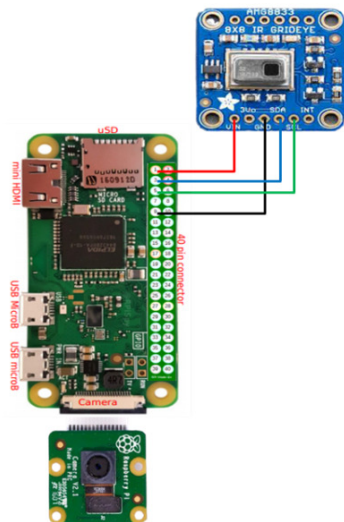


Fig. 2. AMG8833 Thermal Camera Sensor connection with the Raspberry Pi Zero Wireless.

### B. Software Design and Development

#### • Embedded facial detection and recognition algorithm

To perform the facial detection and recognition tasks, a suitable algorithm has been implemented (Figure 3). When the integrated Pi Camera is active, it will acquire a real-time image which will be compared with a preloaded image. Upon confirming the acquired real-time image, the embedded software algorithm will perform facial recognition to confirm the person's identity and will provide the recognized person's name. Once the name has been confirmed, it will be recorded into the attendance list.

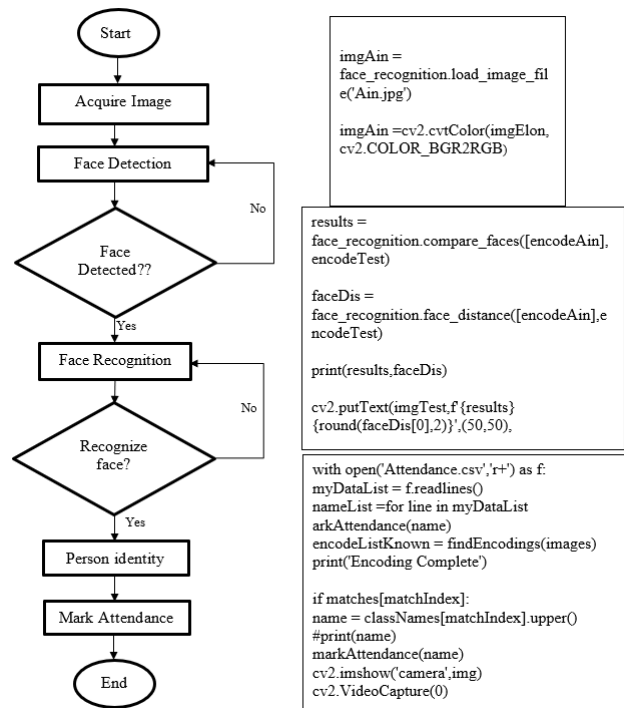


Fig. 3. Embedded face detection and recognition algorithm.

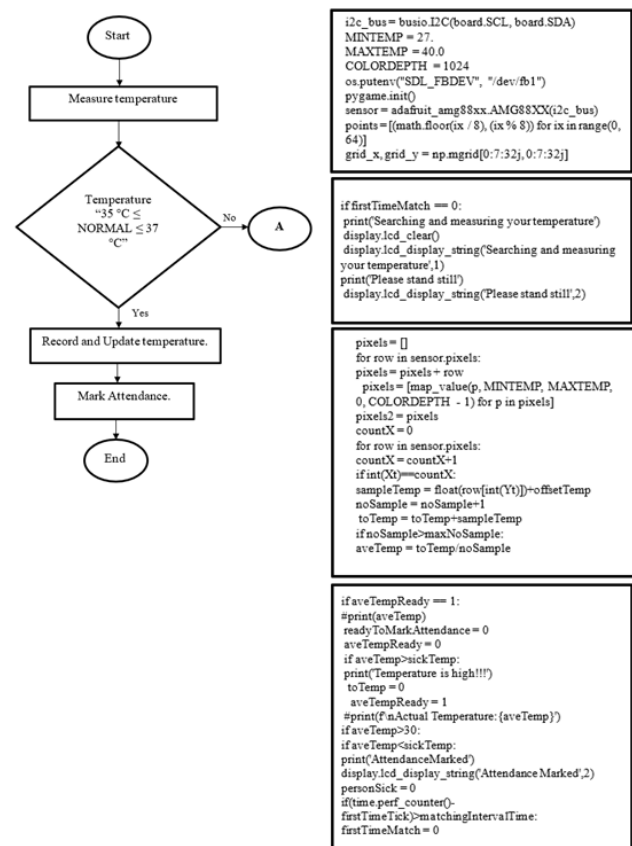


Fig. 4. Embedded AMG8833 thermal sensor algorithm.



- Embedded algorithm for the AMG8833 thermal camera sensor and email notification

The AMG8833 thermal camera sensor is integrated to measure the real-time human temperature. The AMG8833 sensor reads a real-time temperature value when a person's face approaches the camera and the sensor. Hence, when the face is recognized, the AMG8833 thermal camera sensor is activated to measure the person's facial temperature and match the measured temperature with the person's name that is stored in the attendance list. If the person's name is matched, the name and measured temperature will be displayed at the integrated Liquid Crystal Display (LCD). Otherwise, the embedded software continues to perform the person's name search to record the measured temperature value. Also, during the temperature measuring, the embedded software algorithm for the AMG8833 thermal camera sensor categorizes the measured temperature into two classes, i.e. " $35^{\circ}\text{C} \leq \text{NORMAL} \leq 37^{\circ}\text{C}$ " or " $37^{\circ}\text{C} < \text{HIGH} \leq 40^{\circ}\text{C}$ ". If the measured temperature is Normal, its value is displayed on the LCD, otherwise an email notification will be sent to the authorities to inform about the person's condition as depicted in Figures 4 and 5.

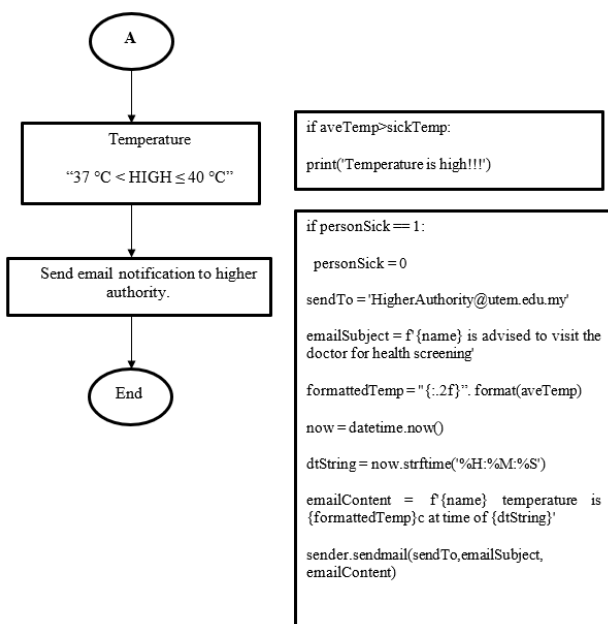


Fig. 5. Embedded real-time email processing and notification algorithm.

### III. RESULTS AND DISCUSSION

#### A. Hardware System Design and Development

Figure 6 shows the arrangement of AMG8833 thermal camera sensor, Pi Camera, and LCD for the hardware system design and development. The AMG8833 and the Pi Camera are placed side by side to allow simultaneous facial detection and recognition and facial temperature measurement upon authentication. Details such as name and temperature of the scanned face are displayed on the integrated LCD.

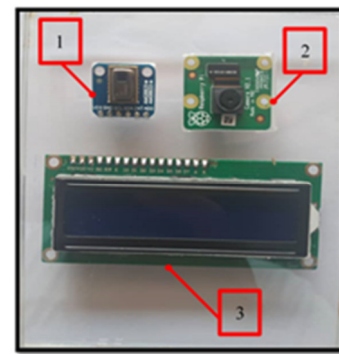


Fig. 6. Placement of integrated Pi Camera, AMG8833 Thermal Camera Sensor, and LCD - IoT face thermal recognition system.



Fig. 7. Illustration of face recognition in a rectangular box.

Another reason of placing the AMG8833 sensor and the Pi Camera side by side is because this way the AMG8833 thermal camera sensor picture size of  $480 \times 480$  pixels is centered at the Pi Camera picture size of  $640 \times 480$  pixels as shown in Figure 7. This allows the AMG8833 Thermal Camera Sensor to effectively scan the center of the forehead for temperature measuring as can be seen in Figure 7. Figure 8 shows the complete integration of Pi Camera, AMG8833 thermal camera sensor, and LCD into the Raspberry Pi Zero Wireless.

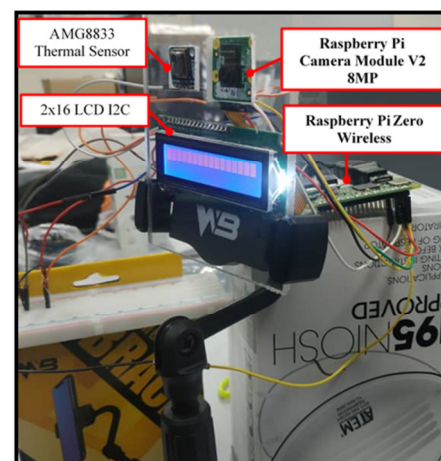


Fig. 8. Integration of the Pi Camera, AMG8833 Thermal Camera Sensor, and LCD in the Raspberry Pi Zero Wireless.

#### B. Validation of the Developed Software

This section validates the developed software algorithm which integrates the Pi Camera and AMG8833 Thermal Camera Sensor in Raspberry Pi Zero Wireless for face

detection, face recognition, and face temperature reading and measurement. To perform face detection and recognition a suitable algorithm (Figure 4) was implemented. The system starts with the embedded algorithm for face detection and recognition. When the integrated Pi Camera is activated it will acquire a real-time image, which will be compared with the stored image in the SD-Card. Upon confirming the acquired real-time image, the embedded software algorithm will perform face recognition to confirm the person's identity. Once the person's face is recognized, the AMG8833 thermal camera sensor will be activated to measure the person's facial temperature which will be evaluated based on the two conditions shown in Figures 4 and 5. These conditions will be matched while real-time facial temperature is measured and the information is stored into the Raspberry Pi Zero Wireless SD-Card. The person attendance status is also updated upon completing the temperature measuring process.

Figure 9 shows the captured result from the Python shell in the Raspberry Pi Zero Wireless when the program starts to operate. After the system successfully acquires the preloaded image, the person's face detection and recognition are executed. During the execution, the person's name is validated and the proposed system continues to measure the facial temperature by using the AMG8833 sensor and stores the measured temperature value into the SD-card. Figure 9 shows the name and temperature of the person after scanning, using the Pi Camera and the AMG8833 thermal camera sensor.

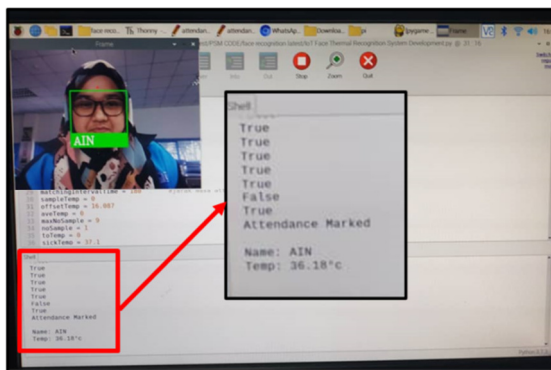


Fig. 9. Name and captured temperature– Python shell - Raspberry Pi Zero Wireless.



Fig. 10. Person's name and measured temperature are displayed at the integrated LCD screen.

The AMG8833 thermal camera sensor displays a real-time temperature value when a person's face approaches the Pi Camera. Hence, once the face is recognized, the AMG8833 thermal camera sensor is activated to measure the person's facial temperature and match the measured temperature with the person's name before storing the information into the SD-card and records the person's attendance. Figure 9 also shows the Python shell output which validates the information that has been stored into the SD-card. The output of the Python shell is also shown on the LCD as shown in Figure 10.

### C. Conditions of the Face Thermal Recognition System

When the face detection and recognition are successfully validated, the program reads the temperature of the person. As mentioned above, fever is the most common symptom of COVID-19. This stage is important to validate the system's operation. If the system is unable to distinguish between a low-risk and a high-risk individual, this system failure will cause possible spread of the COVID-19 disease. According to the World Health Organization (WHO), normal human body temperature ranges from 35.0°C to 37.1°C as shown in Figure 4 and people with this range of facial or body temperature are considered healthy and fit to work. A person with a fever is defined as someone who has a body temperature greater than 37.1°C (Figure 5) [12-14]. So, the temperature scanned values are classified as Normal and High for the respective temperatures as shown in Table I.

TABLE I. CONDITIONS AND TEMPERATURES

Conditions	Temperature (°C)
Normal	35.0 – 37.0
High	37.1 – 40.0

#### • Condition 1

If the measured temperature is between 35°C and 37°C (Normal), the temperature value will be displayed on the LCD when the person's name is matched (Figure 11). Once the person's name has been confirmed and the AMG8833 thermal camera sensor has successfully read the normal temperature of the user, the name and the temperature will be recorded into the attendance list accordingly. Similarly, the person's data are also recorded into the system's database to provide real-time health monitoring update of the person. This attendance record can also be sent to relevant authorities as shown in Figure 11.

#### • Condition 2

If the measured temperature is between 37°C and 40°C, then the person is not allowed to enter the premise or building. The person's information/data are sent through an email notification to the competent authorities. Also, during the temperature scanning, the status "Temperature High" is displayed on the LCD and at the Python shell as shown in Figure 13. Figure 14 shows an email notification containing information about the scanned person's temperature.

### D. Summary

The presented results successfully validated the proposed real-time analytic face thermal recognition system integrated with email notification. The system did successfully perform the face detection and recognition tasks, measured the facial

temperature of the scanned person, and sent an email notification when required. Therefore, the proposed system's design and integration method can be implemented into premises or buildings to assist the early detection of any fever-related illness.



Fig. 11. Normal temperature condition.

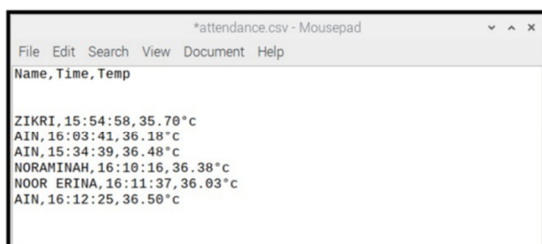


Fig. 12. File CSV- individual name and temperature - attendance list.



Fig. 13. High temperature condition.

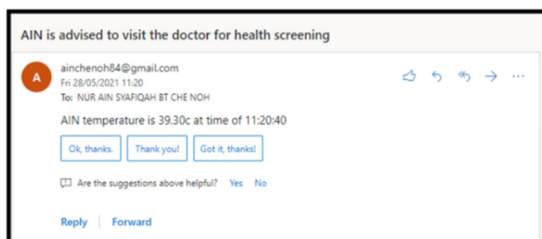


Fig. 14. Email notification.

#### IV. CONCLUSION

A real-time analytic face thermal recognition system integrated with email notification was developed and deployed in order to assist the early detection of any fever-related illness, especially COVID-19. Due to the severity of the COVID-19 disease, the proposed system has integrated some features that may assist the authorities and keep the premises or buildings safe. Also, the integrated hardware and software successfully validated the objectives of the real-time analytic face thermal recognition system integrated with email notification. In terms of system deployment, the system can be placed in any open space, where the respective person is needed to perform temperature scanning before entering a building or premise.

#### ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Centre of Telecommunication Research & Innovation (CeTRI), Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka and the Ministry of Higher Education, Malaysia.

#### REFERENCES

- [1] R. Saher and M. Anjum, "Role of technology in COVID-19 pandemic," *Researches and Applications of Artificial Intelligence to Mitigate Pandemics*, vol. 2021, pp. 109–138, 2021, <https://doi.org/10.1016/B978-0-323-90959-4.00005-5>.
- [2] A. Barnawi, P. Chhikara, R. Tekchandani, N. Kumar, and B. Alzahrani, "Artificial intelligence-enabled Internet of Things-based system for COVID-19 screening using aerial thermal imaging," *Future Generations Computer Systems*, vol. 124, pp. 119–132, Nov. 2021, <https://doi.org/10.1016/j.future.2021.05.019>.
- [3] R. E. Tabiah, M. I. Solihin, and A. Machmudah, "Developing Affordable IoT-based Body Temperature Screening," *Journal of Sustainable Engineering and Built Environment*, pp. 1–5, 2020.
- [4] V. Manuel Ionescu and F. Magda Enescu, "Low cost thermal sensor array for wide area monitoring," in *2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Bucharest, Romania, Jun. 2020, <https://doi.org/10.1109/ECAI50035.2020.9223193>.
- [5] M. Krišto and M. Ivasic-Kos, "An overview of thermal face recognition methods," in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, Feb. 2018, pp. 1098–1103, <https://doi.org/10.23919/MIPRO.2018.8400200>.
- [6] M. M. Khan, S. Mehnaz, A. Shaha, M. Nayem, and S. Bourouis, "IoT-Based Smart Health Monitoring System for COVID-19 Patients," *Computational and Mathematical Methods in Medicine*, vol. 2021, Nov. 2021, Art. no. e8591036, <https://doi.org/10.1155/2021/8591036>.
- [7] Assoc. Prof. Dr. M. Abdulrazaq, H. Zuhriyah, S. Al-Zubaidi, S. Karim, R. Ramli, and E. Yusuf, "Novel Covid-19 Detection and Diagnosis System Using Iot Based Smart Helmet," *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 7, pp. 2296–2303, Mar. 2020.
- [8] A. J. Ismael, H. A. Rahim, R. Ghazali, and K. N. Mezher, "Smart Home and Thermal Imaging Technology for COVID 19 Detection Using the Internet of Things (IoT)," *ELEKTRIKA- Journal of Electrical Engineering*, vol. 20, no. 2–2, pp. 87–92, Sep. 2021.
- [9] N. Kumar, A. Hashmi, M. Gupta, and A. Kundu, "Automatic Diagnosis of Covid-19 Related Pneumonia from CXR and CT-Scan Images," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 7993–7997, Feb. 2022, <https://doi.org/10.48084/etasr.4613>.
- [10] H. A. Owida, A. Al-Ghraibah, and M. Altayeb, "Classification of Chest X-Ray Images using Wavelet and MFCC Features and Support Vector Machine Classifier," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7296–7301, Aug. 2021, <https://doi.org/10.48084/etasr.4123>.

- 
- [11] S. A. A. Biabani and N. A. Tayyib, "A Review on the Use of Machine Learning Against the Covid-19 Pandemic," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8039–8044, Feb. 2022, <https://doi.org/10.48084/etasr.4628>.
  - [12] X. Fu, "Design of Facial Recognition System Based on Visual Communication Effect," *Computational Intelligence and Neuroscience*, vol. 2021, Dec. 2021, Art. no. e1539596, <https://doi.org/10.1155/2021/1539596>.
  - [13] K. H. Rahouma and A. Z. Mahfouz, "Design and Implementation of a Face Recognition System Based on API mobile vision and Normalized Features of Still Images," *Procedia Computer Science*, vol. 194, pp. 32–44, Jan. 2021, <https://doi.org/10.1016/j.procs.2021.10.057>.
  - [14] V. K. Sharma, "Designing of Face Recognition System," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, Madurai, India, Feb. 2019, pp. 459–461, <https://doi.org/10.1109/ICCS45141.2019.9065373>.



# Control Design for the Ward–Leonard System in Wind Turbines

**Huy Hai Bui**

Faculty of Electronic Engineering, Technology University of Economics - Technology for Industries, Vietnam

bhhai@uneti.edu.vn

(corresponding author)

Received: 20 October 2022 | Revised: 3 November 2022 and 13 November 2022 | Accepted: 15 November 2022

## ABSTRACT

A robust optimal controller for the Ward-Leonard system in a wind turbine is used to meet the performance and stability requirements when the system parameters change. However, designing according to the robust optimal method often leads to a high-order controller. This study investigated the application of order reduction algorithms to simplify the controller and help it better meet the real control problem. Comparative evaluation of the order reduction controller methods showed that the second-order reduction controller according to Moore's balanced truncation algorithm was the most suitable to replace the higher-order controller. The step response quality of the system was better when using a second-order reduction controller than a higher-order controller.

**Keywords-model order reduction; Ward–Leonard system; wind tube system; robust optimal control**

## I. INTRODUCTION

Wind energy is a natural form of energy produced by the movement of air in the atmosphere. Also, it is an indirect form of solar energy. Wind energy is the process where wind uses its movement to generate mechanical energy. The research on the potential application of wind energy is quite diverse [1-9], but most researchers believe that the prospect of wind energy development is an inevitable future trend, as clean energy sources will gradually replace current fossil energy sources. This share is determined based on the countries' energy policies and the roadmap to reduce adverse environmental impact. Wind turbines convert the kinetic energy of the air into mechanical energy. For maximum efficiency, the wind turbine blades must direct the wind with an optimum angle, must be flexible and strong to withstand disturbances and high-speed rapid oscillations, and require a suitable control system [10].

The Ward-Leonard DC motor speed control system is used when the speed control requirements are very wide and very sensitive. It is a suitable system for wind turbine control and antenna tracking systems [11]. The system contains a DC generator that serves as a power amplifier for the DC control signal. The generator is rotated at a constant speed from the main motor and an output voltage is generated and supplied to the DC motor, while an external DC power source powers the motor's inductor. The wind turbine control problem requires the controller to have good performance and strong stability when the parameters of the model change. A robust optimal controller is the most suitable to control the Ward–Leonard system in a wind turbine system. However, the design method of the robust optimal controller  $H^\infty$  [12] often leads to a high-

order controller. A high-order controller has many disadvantages, such as complicated programming and long computation time, resulting in slow system responses. Reducing the order of a high-order robust controller while ensuring quality has a high practical significance. Two basic methods are used to obtain a low-order robust optimal controller. The first method designs a high-order controller according to a stable optimal control algorithm and then applies model order reduction algorithms to obtain a lower-order stable optimal controller [13-14], while the second method uses optimization algorithms to design a low-order controller that satisfies the requirements of sustainable optimal control [15].

## II. THE WARD-LEONARD SYSTEM IN WIND TURBINES AND HIGH-ORDER ROBUST CONTROLLER

The Ward-Leonard system contains a DC generator that serves as a power amplifier of the control signal. The DC generator is rotated at a constant speed by the main motor, and its output voltage powers the DC motor, while the field winding of the motor is supplied by a separate DC source. In [10], the Ward-Leonard system was used to control a wind turbine system, and its modeling gave the following results:

$$G(s) = \frac{300}{s(s^3 + 184s^2 + 760.5s + 162)} \quad (1)$$

The wind turbine control problem requires the controller to have good performance and strong stability when the parameters of the model change. A robust optimal controller with the following results was designed in [10]:

$$R(s) = \frac{8.967s^5 + 1663s^4 + 9154s^3 + 1.159 \cdot 10^4 s^2 + 4096s + 436.2}{s^6 + 189.9s^5 + 1852s^4 + 6969s^3 + 1.302 \cdot 10^4 s^2 + 9566s + 1664} \quad (2)$$



The 6<sup>th</sup>-order robust controller has many disadvantages when used in practice, such as slow response time and complex programming. Therefore, it is necessary to reduce the order of this 6<sup>th</sup>-order robust controller to increase the response speed of the system and better meet the requirements of robust control.

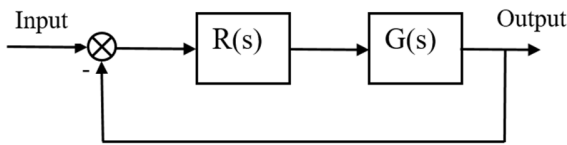


Fig. 1. Block diagram of the control system.

### III. REDUCING THE HIGH-ORDER ROBUST CONTROLLER

All the poles of the 6<sup>th</sup>-order controller have negative real parts, so the 6<sup>th</sup>-order controller is a stable model. Different order reduction algorithms were used to reduce the order of the 6<sup>th</sup>-order controller, and the different resulting controllers were compared to choose the most suitable low-order controller. This study used Moore's balanced truncation algorithm [16], LQG balancing algorithm [17], and Singular Perturbation Approximation (SPA) [18] to reduce the order of the controller. The results of the step reduction are shown in Tables I-III.

TABLE I. ORDER REDUCTION OF THE 6<sup>TH</sup>-ORDER CONTROLLER ACCORDING TO MOORE'S BALANCED TRUNCATION ALGORITHM [16]

Order	$R_r(s)$	$\ R - R_c\ _\infty$
3	$\frac{8.968s^2 + 45.54s + 10.7}{s^3 + 9.531s^2 + 29.82s + 40.12}$	0.0047
2	$\frac{8.881s + 1.978}{s^2 + 4.501s + 7.927}$	0.017
1	$\frac{8.895}{s + 4.493}$	1.1718

TABLE II. ORDER REDUCTION OF THE 6<sup>TH</sup>-ORDER CONTROLLER ACCORDING TO LQG BALANCING ALGORITHM [17]

Order	$R_r(s)$	$\ R - R_c\ _\infty$
3	$\frac{8.967s^2 + 46.1s + 10.71}{s^3 + 9.596s^2 + 30.11s + 40.47}$	0.0043
2	$\frac{8.831s + 2.052}{s^2 + 4.565s + 7.992}$	0.0324
1	$\frac{8.945}{s + 4.552}$	1.707

TABLE III. ORDER REDUCTION RESULTS OF THE 6<sup>TH</sup>-ORDER CONTROLLER ACCORDING TO SINGULAR PERTURBATION APPROXIMATION [18]

Order	$R_r(s)$	$\ R - R_c\ _\infty$
3	$\frac{8.967s^2 + 45.79s + 10.62}{s^3 + 9.561s^2 + 29.95s + 40.2}$	0.0044
2	$\frac{8.953s + 2.051}{s^2 + 4.574s + 7.978}$	0.0335
1	$\frac{8.956}{s + 4.548}$	1.7029

To evaluate the reduced-order controllers, their step and bode responses were compared with the original. Figures 2-7 show the results.

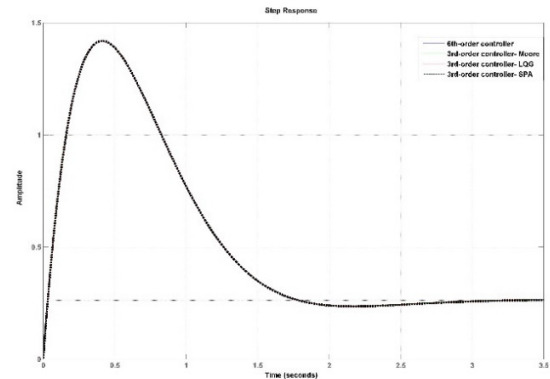


Fig. 2. The step response of the 3<sup>rd</sup> and the 6<sup>th</sup>-order controllers.

The step response of the 3<sup>rd</sup>-order controller is exactly the same as that of the 6<sup>th</sup>-order.

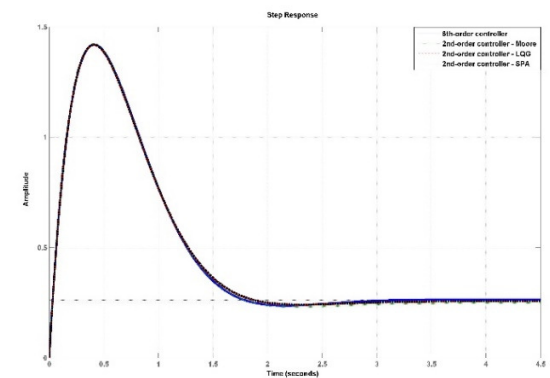


Fig. 3. The step response of the 2<sup>nd</sup> and the 6<sup>th</sup>-order controllers.

The step responses of the 3<sup>rd</sup> and the 2<sup>nd</sup>-order controllers are exactly the same as that of the 6<sup>th</sup>-order.

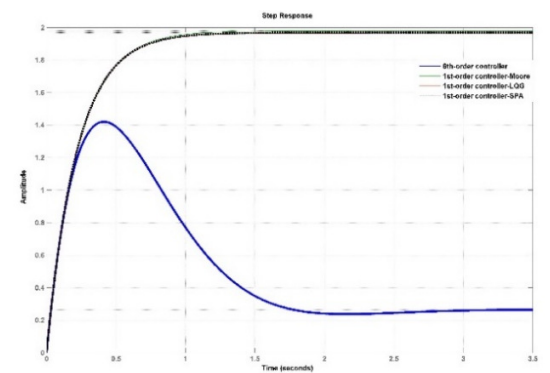


Fig. 4. The step response of the 1<sup>st</sup> and the 6<sup>th</sup>-order controllers.

The step response of the 1<sup>st</sup>-order controller is different from that of the 6<sup>th</sup>-order controller. The step response of the 1<sup>st</sup>-order controllers, according to the different order reduction algorithms, is completely coincident.

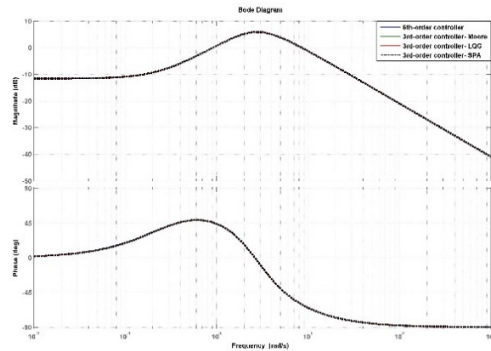


Fig. 5. Bode response of the 3<sup>rd</sup> and 6<sup>th</sup>-order controllers

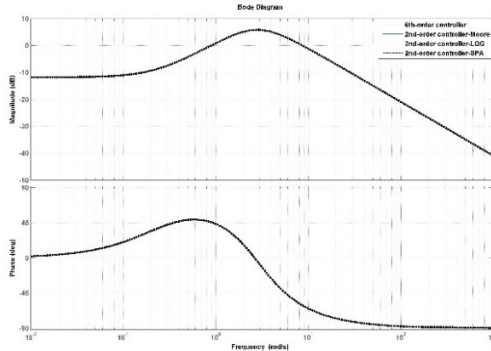


Fig. 6. Bode response of the 2<sup>nd</sup> and 6<sup>th</sup>-order controllers.

The bode response of the 3<sup>rd</sup> and the 2<sup>nd</sup>-order controllers completely coincides with the frequency response of the 6<sup>th</sup>-order controller.

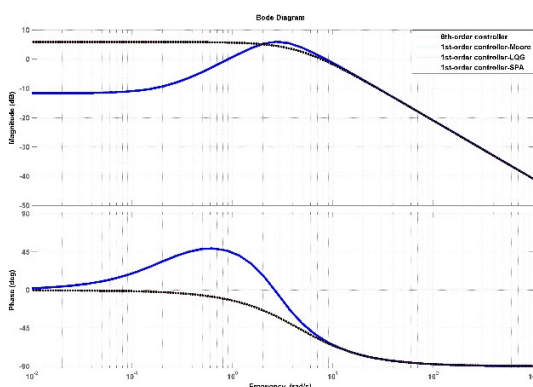


Fig. 7. Bode response of the 1<sup>st</sup> and 6<sup>th</sup>-order controllers.

The frequency response of the 1<sup>st</sup>-order controller, according to the algorithms, is coincident. In the frequency range  $\omega < 29.8$  rad/s, the frequency response of the 1<sup>st</sup>-order controller is different from that of the 6<sup>th</sup>-order. In the frequency range over 29.8 rad/s, the frequency response of the 1<sup>st</sup>-order controller coincides with that of the 6<sup>th</sup>-order. The 2<sup>nd</sup>-order controllers can all replace the 6<sup>th</sup>-order, where the 2<sup>nd</sup>-order controller according to Moore's balanced truncation algorithm is the most suitable controller to replace the 6<sup>th</sup>-order (due to the smallest order reduction error). The 1<sup>st</sup>-order controllers can not substitute the 6<sup>th</sup>-order controllers.

#### IV. USING A REDUCED ORDER CONTROLLER TO CONTROL THE WARD - LEONARD SYSTEM IN A WIND TURBINE SYSTEM

The results of using the 2<sup>nd</sup>-order controllers in Tables I, II, and III and the 6<sup>th</sup>-order controller for the Ward-Leonard system in a wind turbine are shown in Figures 8-10 and Table IV.

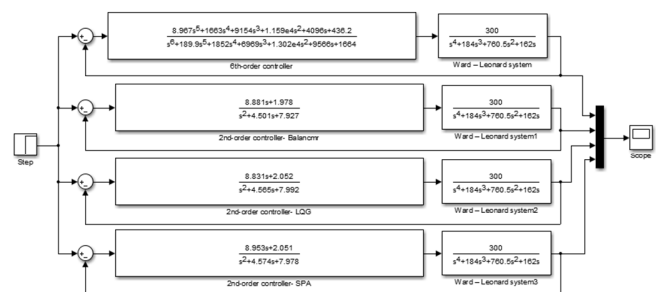


Fig. 8. Simulink diagram of the control system for the Ward-Leonard system in a wind turbine using 6<sup>th</sup>-order and 2<sup>nd</sup>-order controllers.

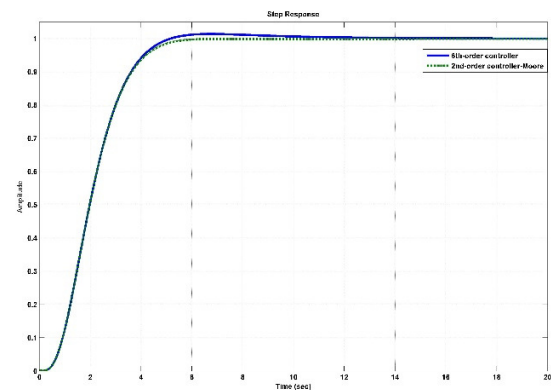


Fig. 9. Response of the control system for the Ward-Leonard system in a wind turbine using 6<sup>th</sup>-order and 2<sup>nd</sup>-order controllers according to Moore's balanced truncation algorithm.

TABLE IV. RESPONSE RESULTS OF CONTROL SYSTEM USING 6<sup>TH</sup> AND 2<sup>ND</sup>-ORDER REDUCTION CONTROLLERS

Response	6 <sup>th</sup> -order controller	2 <sup>nd</sup> -order controller - Moore	2 <sup>nd</sup> -order controller - LQG	2 <sup>nd</sup> -order controller - SPA
Response time	3.61	3.633	3.62	5.75
Settling time (0.5%)	11	5.25	9.08	7.7
Overshoot	1.4%	0%	0.87%	0.67%

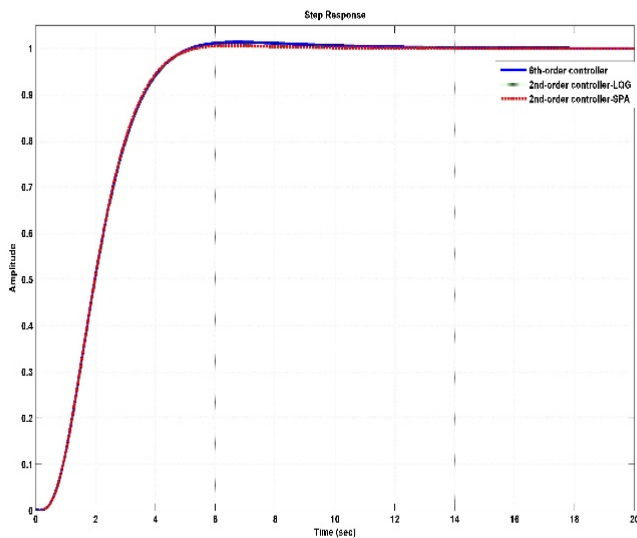


Fig. 10. Response of the control system for a Ward-Leonard system in a wind turbine using 6<sup>th</sup> and 2<sup>nd</sup>-order controllers according to LQG truncation algorithm and SPA.

Compared to the 6<sup>th</sup>-order controller, the Ward-Leonard system in the wind turbine gives better response quality (smaller settling time and overshoot) when using reduced 2<sup>nd</sup>-order controllers. Using a 2<sup>nd</sup>-order controller according to Moore's balanced truncation algorithm gives the best response quality (minimum settling time and overshoot -0%). These results show that the 2<sup>nd</sup>-order controller according to Moore's balanced truncation algorithm is the most suitable controller to replace the 6<sup>th</sup>-order controller.

## V. CONCLUSIONS

Controlling the Ward-Leonard system in a wind turbine according to a sustainable optimization method helps the system to perform well and maintain strong stability when the parameters of the model change. The disadvantage of the controller design, according to the robust optimization method, is that the order of the controller is high (6<sup>th</sup>-order). Three different order reduction algorithms were used in this paper to simplify high-order controllers. Comparison and evaluation of the step-down controllers showed that 2<sup>nd</sup>-order controllers are the most suitable to replace the 6<sup>th</sup>-order controller. Among the 2<sup>nd</sup>-order controllers, Moore's balanced truncation algorithm provided the controller with the best response quality. Future work could investigate other order reduction algorithms to reduce the order of the 6<sup>th</sup>-order controller of the Ward-Leonard system in a wind turbine.

## REFERENCES

- [1] E. A. Al-Ammar, N. H. Malik, and M. Usman, "Application of using Hybrid Renewable Energy in Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 1, no. 4, pp. 84–89, Aug. 2011, <https://doi.org/10.48084/etasr.33>.
- [2] B. Memon, M. H. Baloch, A. H. Memon, S. H. Qazi, R. Haider, and D. Ishak, "Assessment of Wind Power Potential Based on Raleigh Distribution Model: An Experimental Investigation for Coastal Zone,"

*Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3721–3725, Feb. 2019, <https://doi.org/10.48084/etasr.2381>.

- [3] M. Hussain, M. H. Baloch, A. H. Memon, and N. K. Pathan, "Maximum Power Tracking System Based on Power Electronic Topology for Wind Energy Conversion System Applications," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3392–3397, Oct. 2018, <https://doi.org/10.48084/etasr.2251>.
- [4] M. Aien, R. Ramezani, and S. M. Ghavami, "Probabilistic Load Flow Considering Wind Generation Uncertainty," *Engineering, Technology & Applied Science Research*, vol. 1, no. 5, pp. 126–132, Oct. 2011, <https://doi.org/10.48084/etasr.64>.
- [5] J. B. V. Subrahmanyam, P. Alluvada, Bandana, K. Bhanupriya, and C. Shashidhar, "Renewable Energy Systems: Development and Perspectives of a Hybrid Solar-Wind System," *Engineering, Technology & Applied Science Research*, vol. 2, no. 1, pp. 177–181, Feb. 2012, <https://doi.org/10.48084/etasr.104>.
- [6] S. H. E. Osman, G. K. Irungu, and D. K. Murage, "Application of FVSI, Lmn and CPF Techniques for Proper Positioning of FACTS Devices and SCIG Wind Turbine Integrated to a Distributed Network for Voltage Stability Enhancement," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4824–4829, Oct. 2019, <https://doi.org/10.48084/etasr.3101>.
- [7] M. F. Basar, A. M. Norazizi, I. Mustafa, C. T. Colin, S. N. S. Mirin, and Z. Jano, "Investigation on the Performance of a Portable Power Generation System with a Low-Cost Vertical Axis Wind Turbine," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7809–7813, Dec. 2021, <https://doi.org/10.48084/etasr.4454>.
- [8] S. M. Ghania, K. R. M. Mahmoud, and A. M. Hashmi, "A Reliability Study of Renewable Energy Resources and their Integration with Utility Grids," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9078–9086, Oct. 2022, <https://doi.org/10.48084/etasr.5090>.
- [9] M. A. Khelifi, A. Alkassem, and A. Draou, "Performance Analysis of a Hybrid Microgrid with Energy Management," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8634–8639, Jun. 2022, <https://doi.org/10.48084/etasr.4873>.
- [10] E. M. Mekheal, M. Zakaria, R. A. Hamdy, and Medhat. I. El Singaby, "Design of Low Order Robust Controller Using the Artificial Bee Colony Algorithm," in *2019 10th International Renewable Energy Congress (IREC)*, Sousse, Tunisia, Mar. 2019, pp. 1–6, <https://doi.org/10.1109/IREC.2019.8754561>.
- [11] E. Mekheal, M. i. El-Singaby, and A. Khalil, "Robust Controller Design Using H/sub ∞/Loop-Shaping and Method of Inequalities," in *2006 IEEE International Symposium on Industrial Electronics*, Montreal, Canada, Jul. 2006, vol. 1, pp. 118–123, <https://doi.org/10.1109/ISIE.2006.295578>.
- [12] K. Zhou and J. C. Doyle, *Essentials of Robust Control*, 1st edition. Upper Saddle River, NJ, USA: Prentice Hall, 1997.
- [13] N. K. Vu and H. Q. Nguyen, "Balancing Control of Two-Wheel Bicycle Problems," *Mathematical Problems in Engineering*, vol. 2020, Jul. 2020, Art. no. e6724382, <https://doi.org/10.1155/2020/6724382>.
- [14] V. N. Kien, N. H. Trung, and N. H. Quang, "Design Low Order Robust Controller for the Generator's Rotor Angle Stabilization PSS System," *Emerging Science Journal*, vol. 5, no. 5, pp. 598–618, Oct. 2021, <https://doi.org/10.28991/esj-2021-01299>.
- [15] B. T. Thanh and M. Parnichkun, "Balancing Control of Bicyrobo by Particle Swarm Optimization-Based Structure-Specified Mixed H2/H∞ Control," *International Journal of Advanced Robotic Systems*, vol. 5, no. 4, p. 39, Nov. 2008, <https://doi.org/10.5772/6235>.
- [16] V. N. Kien, N. H. Quang, and N. K. Trung, "Application of model reduction for robust control of self-balancing two-wheeled bicycle," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 1, pp. 252–264, Feb. 2021, <https://doi.org/10.12928/telkomnika.v19i1.16298>.
- [17] E. Jonckheere and L. Silverman, "A new set of invariants for linear systems—Application to reduced order compensator design," *IEEE Transactions on Automatic Control*, vol. 28, no. 10, pp. 953–964, Oct. 1983, <https://doi.org/10.1109/TAC.1983.1103159>.

- [18] Yi Liu and Brian O. D. Anderson, "Singular perturbation approximation of balanced systems," *International Journal of Control*, vol. 50, no. 4, pp. 1379–1405, Oct. 1989, <https://doi.org/10.1080/00207178908953437>.

# Bridging the Effect of the Impactor Head Shape to the Induced Damage during Impact at Low Velocity for Composite Laminates

**Mustapha Rabouh**

Laboratory of Development in Mechanics and Materials (LDMM), University of Djelfa, Algeria  
1700ham@gmail.com  
(corresponding author)

**Khelifa Guerraiche**

Mechanical Engineering Department, Faculty of Technology, University of Batna 2, Algeria | NMISII  
Laboratory, Faculty of Science and Technology, Biskra University, Algeria  
guer.khelifa@yahoo.com

**Kamel Zouggar**

Laboratory of Mechanics of Structures and Solids (LMSS), Faculty of Technology, Djillali Liabès  
University of Sidi Bel-Abbès, Algeria  
zouggar-kamel.zk@gmail.com

**Djemaa Guerraiche**

Applied Energy Physics Laboratory (LPEA), Department of Physics, Faculty of Matter Sciences,  
University of Batna 1, Algeria  
d.guerraiche@gmail.com

Received: 27 October 2022 | Revised: 17 November 2022 | Accepted: 26 November 2022

## ABSTRACT

The present paper presents an investigation and analysis study of the effect of the head shapes of the impactor on the damages observed during low-velocity impact on T700/Epoxy composite laminate. Two types of impactors were investigated: hemispherical and flat-face. A new criterion based on the LARC05 damage model was linked as a LARC\_VUMAT subroutine to simulate the impact and explore the effects of the head form shape through a three-dimensional finite element model. To properly analyze the problem, the history time of the mechanical responses, such as impact forces, bending, principal, and residual stresses, are highlighted and assessed. Additionally, a comparison with the experimental data found in the literature was performed to check the validity and accuracy of the considered finite element model. The damage occurring in the T700/Epoxy plates is illustrated for each impactor head shape. The mechanical response curves and all kinds of damage of the presented simulations are in perfect agreement with the experiments. The proposed VUMAT is efficient in the prediction of fiber kinking, matrix cracking, fiber splitting, and fiber tension of a laminate, and more importantly, it is easy to implement for other types of materials and the reproducibility of the analysis is assured.

**Keywords-**composites; impact; VUMAT; LARC damage

## I. INTRODUCTION

Composite structures are increasingly used in the manufacturing of aircraft and automobile structural components, mainly due to the balanced ratio of high strength to low weight [1]. However, they remain vulnerable to defects induced by the non-miscible and inhomogeneous composition of the material. Moreover, the thin thickness of composite

structures makes them susceptible to damage, especially under impact loading [2]. For such a loading case, the damage may barely be visible, but it might nevertheless considerably reduce strength. Accordingly, a better understanding of the behavior under impact loading is required for composite structure [3]. To meet this need, considerable efforts have been made to study and optimize the impact performance of composite laminates



[4, 5]. Many researchers use experimentation to give a clear explanation of the impact behavior of composite laminates, although relatively few comparative studies have been reported on the impactor head shape under low-velocity impact conditions. Impacts commonly occur accidentally during manufacturing, maintenance operations, etc. Up to now, research has fallen beneath three categories: impact response of composite laminates, damage characterization, and contact mechanisms in impact-loaded composite structures [6].

Authors in [7, 8] investigated and experimentally emphasized the impact of impactor head-shape on the impact response of thin woven carbon/epoxy laminates. They report that different impactor shapes have a large effect on the damage mechanism. Authors in [9] reported that different impactor head-shapes produce different degrees of different types of damage, such as fiber breakage, matrix cracking, and delamination, which can affect the residual properties of the composite. Authors in [10] were interested in the effect of impactor diameter. They showed that this parameter significantly affects the impact and post-impact compression response of composite materials. Authors in [11], from impact tests on woven glass and carbon fiber composites at low impact velocities using a drop tower having a hemispherical impactor, found that impact damage was primarily a function of impact energy rather than velocity or mass separately. Furthermore, regardless of the mass of the impactor used, the measured peak forces reveal a correlation with the impact energy, indicating that the low-velocity impact of these samples is a quasi-static process [12]. However, according to the results of their tests, the relationship between peak load and the size of the damage was independent of the diameter of the impactor. In most of the mentioned studies, impact tests or simulations were performed using hemispherical impactors according to ASTM D7136 [13]. Since different shaped objects might collide with the composite structures, the head shape of impactors must be considered in the experiments [14].

Recently, it has been pointed out that the contact area between the impactor and the composite test sample is a crucial factor in the response of the impacted composites. In addition, impactors with a larger contact area yielded a higher initial peak load, highest maximum load, shortest contact time for the impact event timescale, and more extensive delamination [15, 16]. Despite this, the overall relationship between internal and external damage to impacted composite laminates is difficult to describe in a simple way [17, 18], especially at various impact conditions. Developing of analytical and numerical models is another alternative for investigating the impact response of composite laminates. These techniques provide a better understanding of the physics of impact effects and are more efficient computationally since they are comparatively processable with simple geometries and boundary constraints [19]. For example, to simulate the behavior of composite laminates in high-mass, low-velocity impacts, a simplified analytical model for predicting the onset and propagation of impact damage in composites was proposed in [20, 21]. Authors in [22] distinguished the relation between analytical and numerical models for finite element simulation. As a result, they report that analytical models are mainly used to model the overall response to an impact. In [23], two analytical design

methods are distinguished. One based on energy equilibrium, where the behavior remains quasi-static at low energy and low velocity [24] and another that uses the notion of semi-infinite space which only takes into account the plate deflection. The effectiveness of this model is established when the strain front does not reach the plate boundaries. Otherwise, the model is no longer suitable, as the dynamics of the plate must be considered [25].

The literature information on the effects of impactor head shapes upon the impact response of laminated composite plates is rather poor. The effect of the impactor head shape is an essential parameter to be studied in order to determine the response of laminates under different impactor shapes as well as the induced damage. So, the main objective of the current study is to numerically determine the influence of impactor head shapes on the low-velocity impact performance of composite laminate plates and compare them to the results of an experimental study found in the literature [6]. For this purpose, a three-dimensional non-linear finite element model was developed to simulate the effect of the impactor head shapes on the impact behavior of a composite laminate plate associated to a user subroutine based on LARC05 criterion and was implemented under the commercial software ABAQUS/Explicit. In terms of impact forces-time, deflections-time, predicted stresses, deformation plots, and failure damages ascertainment in the composite plate, the finite element analysis findings agree well with the experimental results of [6]. In the predicted stress results, the stresses in the three main directions ( $S_{11}$ ,  $S_{22}$ , and  $S_{33}$ ) are illustrated as a replacement for the commonly used von Mises stresses, to show the stress state in each layer according to the fiber direction as well as in the strain illustration case.

## II. NUMERICAL MODELLING TECHNIQUES

In order to maintain the same experimental protocol conducted by [1], all the simulated composite specimens possessed the same lay-up of  $[+45_3/0_3/-45_3/0_3]_S$ . The simulated plates are made of T700/Epoxy pre-pregs. The dimensions of the composite specimens were  $150 \times 100 \times 3 \text{ mm}^3$  according to the ASTM D7136 standard [2]. Two impactor head shapes were taken into consideration: An impactor with a Hemispherical head end (HemI) and a Flat-Face end Impactor (FFI). The two impactors were made of steel. All mechanical data concerning the plate and the impactors are shown in Table I. The modeling techniques for the composite failure and the impactor head shapes were investigated and ABAQUS/Explicit was used for the impact simulation. The composite plate was modeled using 337296 C3D8I elements and for the impactor R3D4 elements were used as follows: 3953 elements for the HemI and 3638 for the FFI. An investigation of mesh convergence was conducted and the optimal size of 1mm was considered. All details are illustrated in Figure 1(a).

The boundary conditions for the constraints applied during the simulations are similar to those used in the experiments of [6]. More details are presented in Figure 1(b) and Table II.

TABLE I. MECHANICAL PROPERTIES OF COMPOSITE LAMINATES T700/EPOXY CALCULATED FROM MICROMECHANICS

Properties T700/Epoxy			
Fiber $u_f$	50%	$X^T$ (MPa)	$+1.950.10^{+3}$
Layer thickness $h$ (mm)	$1.25.10^{-1}$	$Y^T$ (MPa)	$+7.5.10^{+1}$
$\epsilon_{11}$ (MPa)	$13.10^{+4}$	$X^C$ (MPa)	$-1.0150.10^{+3}$
$\epsilon_{22}$ (MPa)	$7.7.10^{+3}$	$Y^C$ (MPa)	$-2.20.10^{+2}$
$\epsilon_{33}$ (MPa)	$7.7.10^{+3}$	$\tau_{xy}$ (MPa)	$+1.50.10^{+2}$
$G_{12}$ (MPa)	$4.8.10^{+3}$	$\epsilon_{xx}^T$ (%)	$+2.16867.10^{-3}$
$G_{13}$ (MPa)	$4.8.10^{+3}$	$\epsilon_{yy}^T$ (%)	$+6.91057.10^{-3}$
$G_{23}$ (MPa)	$3.8.10^{+3}$	$\epsilon_{xx}^C$ (%)	$-1.15663.10^{-2}$
$n_{12}$	$3.0.10^{-1}$	$\epsilon_{yy}^C$ (%)	$-3.33333.10^{-3}$
$n_{13}$	$3.0.10^{-1}$	$\epsilon_{xy}^T$ (%)	$+2.16867.10^{-3}$
$n_{23}$	$3.3.10^{-1}$	$\rho$ (g/mm <sup>3</sup> )	$1.6048.10^{-3}$
$S^L$ (MPa)	150	$\alpha_0$	$51^\circ$
$S^T$ (MPa)	150	$\varphi_0$	$53^\circ$
$\eta_L$	0.1	$\eta_T$	0.1

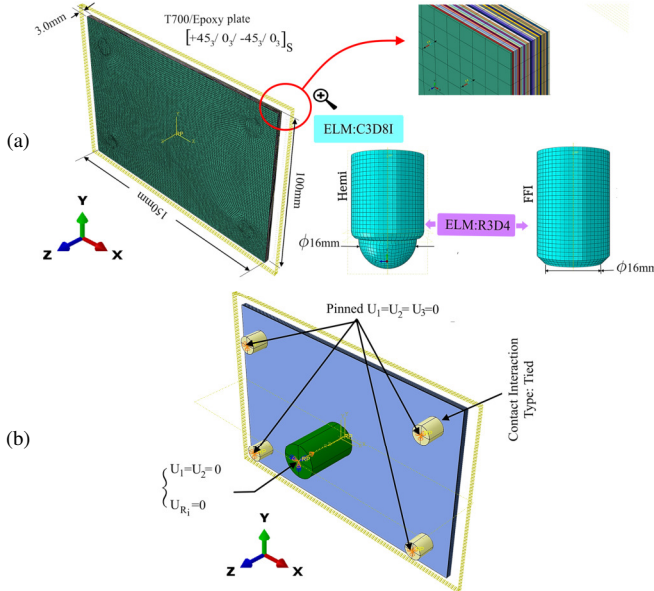


Fig. 1. Modeling details of impact tests. (a) Plate and the two head shape impactor geometrical dimensions and finite element modeling, (b) Illustration of plate and impactor assembly with BCs.

TABLE II. STEPS OF THE PROPOSED MODEL

Impact model	
Boundary Conditions (BCs)	<ul style="list-style-type: none"> <li>× The z-direction of the supports of the plate is pinned.</li> <li>× The z-direction of the impactors is unconstrained.</li> <li>✓ Ensure impactor-to-plate contact</li> <li>✓ Interaction conditions were attributed using standard contact algorithm</li> </ul>
Loading	<ul style="list-style-type: none"> <li>× A predefined field in terms of velocity was applied into the impact: <math display="block">\begin{cases} V_{IHSE} = 2.4 \text{ m/s} \\ V_{IFFE} = 2.39 \text{ m/s} \end{cases}</math> </li> <li>✓ In addition, for the Heml and FFI, a point inertia mass of 5.20kg and 5.25Kg was assigned.</li> </ul>
Simulation time	<ul style="list-style-type: none"> <li>× Time of impact is calculated by: <math display="block">\frac{\text{Contact time in experiments}}{\text{number of increments}}</math> </li> </ul>

Furthermore, considering the damages, a VUMAT user subroutine modeling based on the LARC05 criterion [26] was implemented for the explicit scheme. This failure criterion allows the distinction of 4 different damage initiation mechanisms: matrix cracking, fiber kinking, fiber splitting, and fiber tension, as follows:

## 1) Matrix Cracking

$$F_m^{crack} = \sqrt{\left(\frac{\tau_T}{S_T - \eta_T \sigma_N}\right)^2 + \left(\frac{\tau_L}{S_L - \eta_L \sigma_N}\right)^2 + \left(\frac{\langle \sigma_N \rangle_+}{Y_T}\right)^2} \quad (I.1)$$

$$\begin{cases} \sigma_N = \sigma_{22} \cos^2 \alpha + \sigma_{33} \sin^2 \alpha + \sigma_{23} \sin(2\alpha) \\ \tau_T = \frac{1}{2}(\sigma_{33} - \sigma_{22}) \sin(2\alpha) + \sigma_{23} \cos(2\alpha) \\ \tau_L = \sigma_{12} \cos \alpha + \sigma_{31} \sin \alpha \end{cases} \quad (I.2)$$

The angle  $\alpha$  represents the direction of the critical plane perpendicular to the local 23-plane. The  $\alpha$  value is numerically found such that it maximizes  $F_m^{crack}$

2) Fiber Tension ( $\sigma_{11} \geq 0$ )

$$F_f^{tens} = \frac{\sigma_{11}}{X_T} \quad (I.3)$$

3) Fiber Splitting ( $-\frac{X_C}{2} \leq \sigma_{11} \leq 0$ )

$$F_f^{split} = \sqrt{\left(\frac{\tau_{23}^m}{S_T - \eta_T \sigma_2^m}\right)^2 + \left(\frac{\tau_{12}^m}{S_L - \eta_L \sigma_2^m}\right)^2 + \left(\frac{\langle \sigma_2^m \rangle_+}{Y_T}\right)^2} \quad (I.4)$$

4) Fiber Kinking ( $\sigma_{11} \leq -\frac{X_C}{2}$ )

$$F_f^{kink} = \sqrt{\left(\frac{\tau_{23}^m}{S_T - \eta_T \sigma_2^m}\right)^2 + \left(\frac{\tau_{12}^m}{S_L - \eta_L \sigma_2^m}\right)^2 + \left(\frac{\langle \sigma_2^m \rangle_+}{Y_T}\right)^2} \quad (I.5)$$

Within the above fiber splitting and kinking equations, the stresses within the misaligned fibers are calculated as follows:

$$\begin{cases} \sigma_2^m = \sin^2 \varphi \sigma_{11} + \cos^2 \varphi \sigma_{22} - 2 \sin \varphi \cos \varphi \tau_{12}^\psi \\ \tau_{12}^m = \sin \varphi \cos \varphi (\sigma_{22} - \sigma_{11}) + (\cos^2 \varphi - \sin^2 \varphi) \tau_{12}^\psi \\ \tau_{23}^m = \tau_{23}^\psi \cos \varphi - \tau_{31}^\psi \sin \varphi \end{cases} \quad (I.6)$$

where  $\sigma_{ij}^\psi$  represents the stresses in the fiber-kinking plane:

$$\begin{cases} \sigma_{22}^\psi = \cos^2 \psi \sigma_{22} + \sin^2 \psi \sigma_{33} + 2 \sin \psi \cos \psi \tau_{23} \\ \tau_{12}^\psi = \tau_{12} \cos \psi + \tau_{31} \sin \psi, \quad \tau_{31}^\psi = \tau_{31} \cos \psi - \tau_{12} \sin \psi \\ \tau_{23}^\psi = \sin \psi \cos \psi (\sigma_{33} - \sigma_{22}) + (\cos^2 \psi - \sin^2 \psi) \tau_{23} \end{cases} \quad (I.7)$$

Terms  $\tau_{23}$ ,  $\tau_{12}$ , and  $\sigma_{22}$  are the traction components of the fracture plane. In-situ strengths are denoted as  $Y_T^{is}$ ,  $S_T^{is}$ ,  $S_L^{is}$ , and  $S_L^{is}$  represent the transverse tensile strength, transverse shear strength, and longitudinal shear strength respectively. The coefficients  $\eta_T$  and  $\eta_L$  are the friction coefficient under longitudinal shear stress and transverse shear stress, are calculated according to (1.8).

$$\begin{cases} S_T = \frac{Y_C}{2 \tan(\varphi_0)} \\ \eta_L = \frac{S_L \cos(2\varphi_0)}{Y_C \cos^2(\varphi_0)} = S_L \frac{\eta_T}{S_T} \\ \eta_T = -\frac{1}{\tan(2\varphi_0)} \end{cases} \quad (I.8)$$

where  $\varphi_0$  is the direction of thickness under transverse compressive load, which is the experimentally measurable material property and its value is usually  $53^\circ$ .

Through the VUMAT, the signs of  $\sigma_{II}$  for all the above fiber failure mechanisms must be firstly checked. If  $\sigma_{II}$  is positive, then the fiber tension criterion  $F_f^{tens}$  is calculated. Else, for both fiber splitting and kinking, the stresses are rotated to the kinking plane of the fibers, and subsequently these resultant stresses are rotated to the misalignment plane of fibers. Through this process, the angles  $\psi$  and  $\varphi$  are varied together until a peak value of  $F_f^{split}$  or  $F_f^{kink}$  is reached. More details are available in Appendix C.

### III. RESULTS AND DISCUSSION

#### A. Impact Forces and Bending

The typical profile of the forces and bending curves of the simulated composite samples obtained from the two shapes were compared to the experiments of [1] and are shown in Figure. 2. For the same energy level of impact of 15J, it was found that:

- In each impact force ( $F^{imp}$ ) versus time curve, the maximum forces are smaller in the order of HemI and FFI, and the contact time is longer in the same order. These observations are reversed when the maximum bending values are recorded, i.e. the maximum bending values are smaller in the order of FFI and HemI.
- On average, a peak impact force of 6.92kN was recorded for the flat head shape impactor upon contacting the plate. Compared to the HemI, where the average maximum impact load was 4.86kN, this is roughly 70% higher, see Figure 2(a).
- Regarding contact time, the time duration of the impact event changes for each head shape of impactor. Accordingly, the FFI struck the plate with an average contact time of 6.4ms, a reduction of 0.8ms (or approximately 12.50%) compared to the HemI impactor, where the average contact time of 7.2 was registered.
- Observing the records of bending shown on Figure 2(b), the highest values are recorded at the mid-contact time and are 5.26 and 4.48mm for the HemI and FFI respectively. The reduction in bending between the two shapes results from the load distribution over the contact area. For the HemI, the motion process of the impacts is associated with the formation of a spherical cavity in the medium. However, for the FFI this process is eliminated, and a large contact area is put at stake.
- Compared to the experiments of [1], a delay in all plotted numerical results is observed, which can be explained by the configuration of boundary conditions and interactions set in the finite element model.
- Finally, it can be clearly seen that the simulation results of the impact forces and bending for the two suggested shapes are in good agreement with the experimental results. The FE model accurately predicts the effects of the impactor

head shape on the recorded maximum forces and bending of the impacted composite plate and the impact times.

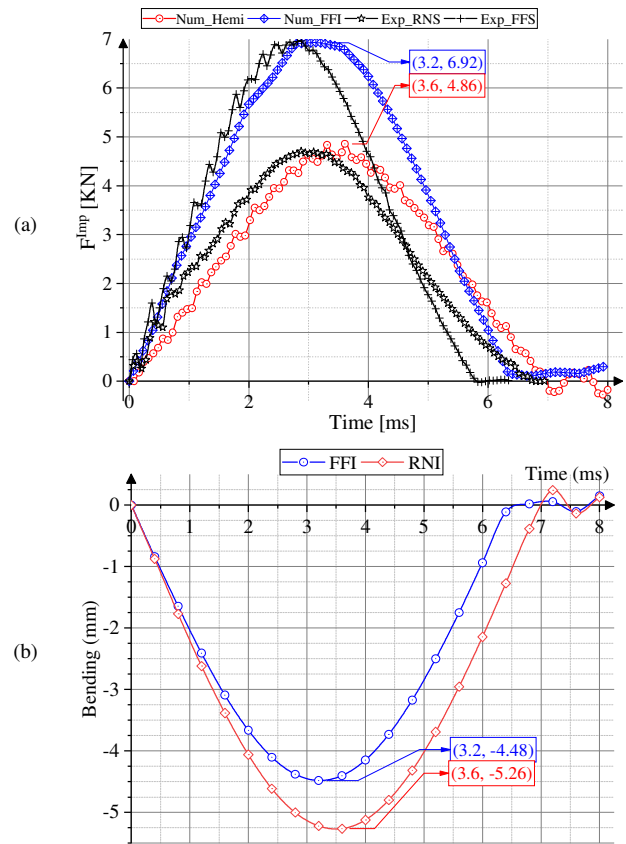


Fig. 2. (a) Impact force and (b) bending curves during contact time. Numerical results versus the experimental data of [1].

#### B. Main Stresses

Compared to experiments that require stress gauges to measure stresses, the numerical model presented in this study provides an easy way to pick up stresses. Because the presented composite plate is balanced, only the stress levels of the upper layers of the composite are presented. Figure 3 shows the principal stress distribution in the three plate directions for each layer. The main observations are:

- A detailed overview of the stress status of each layer is available within the model.
- The stresses have a quasi-symmetric distribution between the upper and lower layers of the composite.
- All the impact resistance is absorbed by the two upper layers. The same findings will be observed but conversely in layers  $L_{11}^S$  and  $L_{12}^S$ .
- As the impacts reaches their maximum, both inner layers  $L_{11}$  and  $L_{12}$  undergo traction, while the others are subjected to compression in both X and Y-directions.

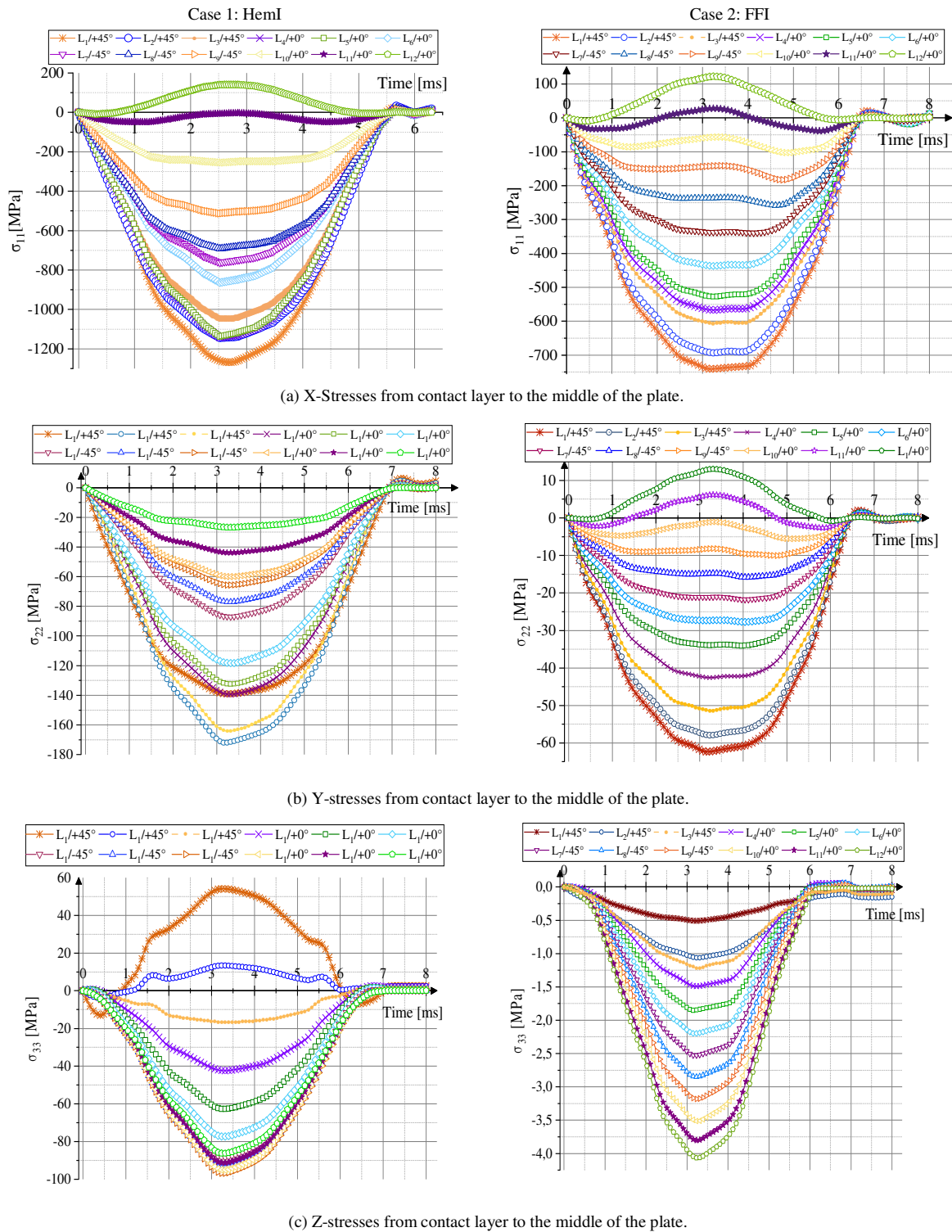


Fig. 3. Pricipal stress evolution during impact for HemI and FFI and energy level of 15J.

- The peak value of stresses for  $\sigma_{11}$  and  $\sigma_{22}$  is registered through the first layer ( $L_1/+45^\circ$ ) in contact with impactors. However, for  $\sigma_{33}$  the peak value exists in the inner-layer of the composite plate ( $L_{12}/+0^\circ$ ).
- On average, the peak values are respectively for  $\sigma_{11}$ ,  $\sigma_{22}$  and  $\sigma_{33}$  equal to 750MPa, 62.35MPa, and 4.04MPa obtained from the simulated impact of the FFI impactor.



For the HemI, the recorded values are 1147MPa, 172MPa, and 96.5MPa.

- The values of  $\sigma_{11}$ ,  $\sigma_{22}$  are doubling when the contact surface changes from flat to curved, which is confirmed by the high increase in the value of  $\sigma_{33}$  in the case of the HemI simulation.
- In each stress versus time curve, the maximum stresses are smaller in the order of FFI and HemI, and the contact time is longer in the same order.
- An interaction between the main stresses is shown in Figure 4.

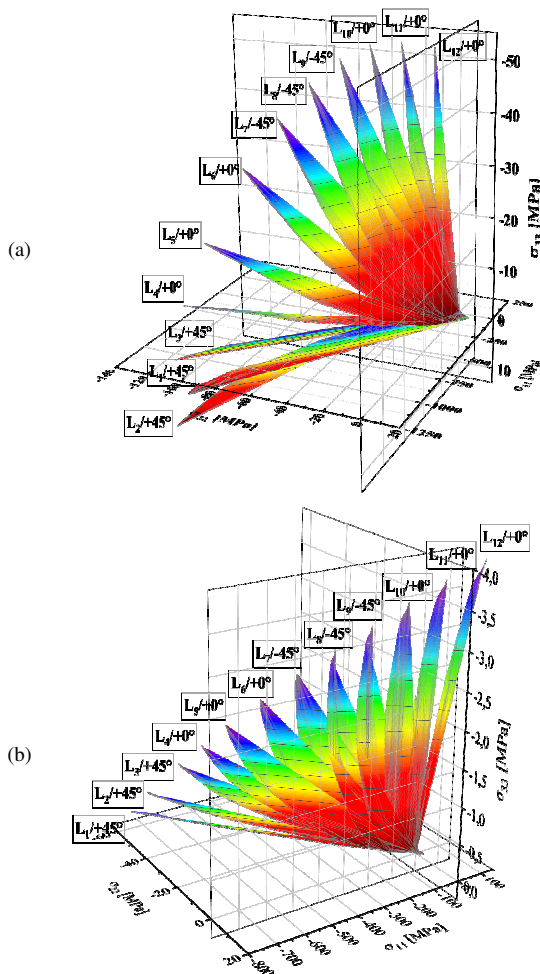


Fig. 4. A 3D-illustration of the interaction of the main stresses. (a) HemI, (b) FFI.

### C. Residual Stresses

After the separation of the impactor from the plate, residual stresses ( $\sigma^{res}$ ) remain. In Figure 5, the residual stress contours are presented according to the FE simulation on the T700/Epoxy [+45<sub>3</sub>/0<sub>3</sub>/−45<sub>3</sub>/0<sub>3</sub>]<sub>S</sub> composite plate after the impact using both proposed impactor head shapes, through which it can be clearly observed that:

- The impactor with the larger contact area provides very high stress waves across the whole area. Nevertheless, similar stress wave patterns are concentrated in the area around the contact area and the supports for both HemI and FFI.
- The highest values of  $\sigma^{res}$  recorded for the FFI and HemI are respectively 59.76MPa and 40.96MPa along the largest dimension of the plate (X-direction, see Figure 5(a)). This reflects the fact that the bending stresses are more dominant and important at this direction. On the other hand, through the depth of the plate, the values of  $\sigma^{res}$  decreased as a result of the indentation mechanism, as shown on Figure 7. In fact, the occurrence of this phenomenon can be explained by the storage of a portion of the impact energy by strain mechanisms in the matrix (indentation) and fiber breakage dissipation.

### D. Expected Strain

In Figure 6, the results for the strain distribution contours across thickness  $\epsilon_{33}$  are illustrated. From the simulations at the peak displacement along the focused mid-zone we notice:

- For both HemI and FFI, it is found that there is a good correlation between the strain distribution maps  $\epsilon_{33}$  and the extent and location of intralaminar matrix damage in the specimens, as shown in Figures 5-7.
- In the case of the HemI, the strain, which was initially localized and located at the center region under the impactor, was extended to the peripheral regions during the interaction between the impactor and the plate, while the FFI has initiated the damage in the composite plate at the region in contact with the ring edge of this impactor, as shown in Figure 5(c).
- In addition, it is noticeable that the curvature is less in the central region under the FFI compared to the HemI.

### E. Damage Analysis

For the sake of brevity, only the damages caused by the HemI are presented in this section. However, in Appendix B, the reader can find detailed information on the FFI damages. All observations about this impactor will be highlighted during the damage results discussion.

In Figure 7, the typical damages generated using the implemented VUMAT\_LARC subroutine of the LARC05 damage criterion in Abaqus/Explicit for both the energy level and the impactor head shape under consideration are illustrated. Based on the outcomes, it is found that:

- The model displays the most important mechanisms of damage that can be observed when a composite plate is impacted. Moreover, compared to the existing damage criteria such as those of [4, 27, 28], the implemented VUMAT\_LARC gives a large insight of the kinking and the slitting of fibers in the zone under impactors (see Figures 7(a),(d)) as well as the extend of the matrix cracking (see Figure 7(b)), which cannot be identified using the previously mentioned criteria.



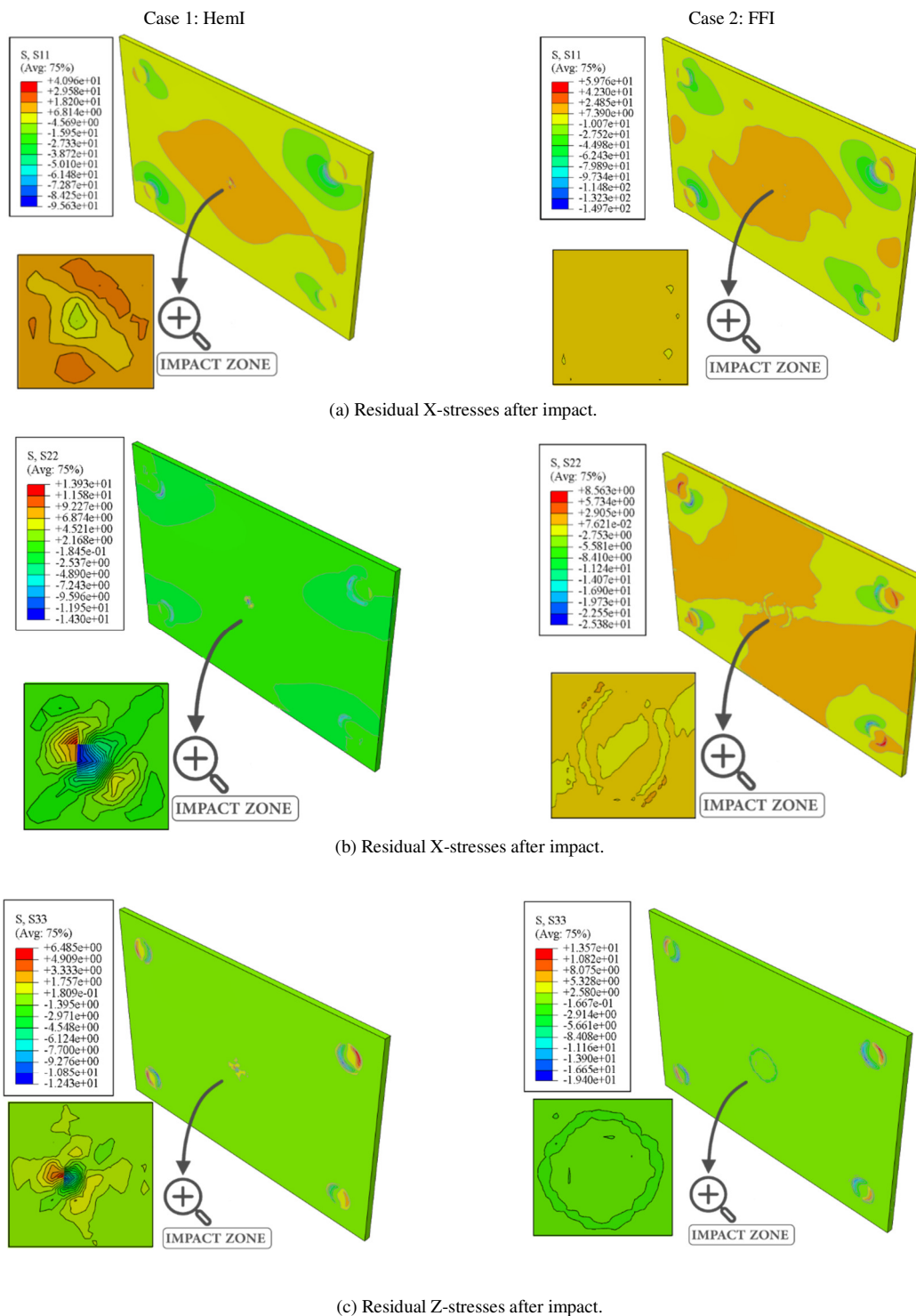


Fig. 5. Contours of residual stresses after impact.

- During the analysis of the impact results, damages are observed on both the front and back surfaces. This observation is highly observed for fiber traction in Figure 7(b), which supports the experimental findings.
- The visible damages on the frontal face impacted of the plate are only impact dents of different sizes. The damage differences in the frontal surface mainly occur due to the difference in contact sizing under different impact conditions (impactor head shapes, interactions, BCs, etc.).

- Indeed, in the case of the HemI, damages like fiber kinking and matrix cracking, are initially localized at the center region under the impactor. Then they are extended to the peripheral regions during the interaction between the impactor and the plate. Moreover, it is observed that the local damages are growing through the plate depth, as shown in Figures 7(a),(c). On the other hand, the FFI damages are initiated at the contact region evolving in a ring shape on the edge of this impactor, as shown in Figure 8. Accordingly, it is found that the FFI damages are less pronounced than the HemI damages.

- It is also noticed that the weakest damage mechanism is the fiber splitting, compared to the other damage mechanisms which are, in decreasing order, matrix cracking, fiber kinking, and fiber tension.

As a final note, some key results from the proposed modeling predictions are summarized in Table III, i.e. maximum impact forces  $F^{Imp}$ , maximum bending, duration of contact with the impactor, and the CAREAs. The findings clearly demonstrate the capabilities of the model and the power of implementation via VUMAT\_LAR to accurately and efficiently capture the impact responses and the damaged area among the composite specimens.

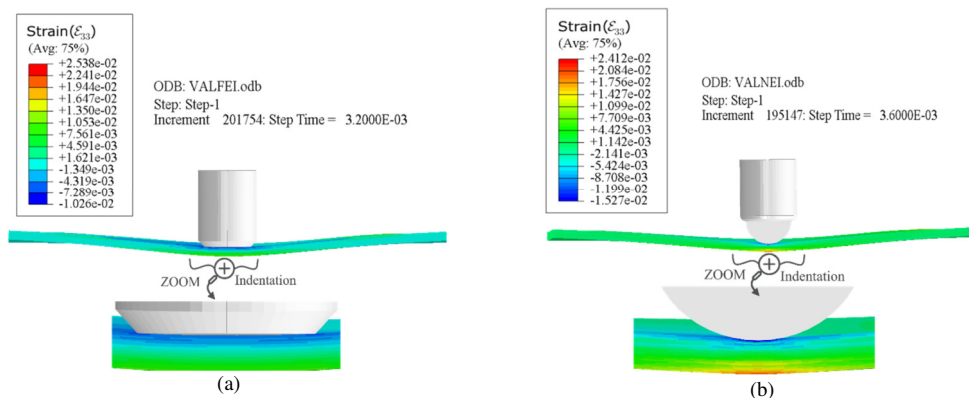


Fig. 6. Residual Z-strain during impact, at the mid-contact time. (a) Case 1 – HemI, (b) Case 2 – FFI.

TABLE III. KEY RESULT COMPARISON BETWEEN THE NUMERICAL MODELING AND THE EXPERIMENTS OF [6]

Variable	Numerical		Numerical [6]		Experimental		Aug. (+)/red. (–) (%)	
	HemI	FFI	HemI	FFI	HemI	FFI	HemI	FFI
Contact Time (ms)	7.20	6.40	6.3	5.7	6.7	5.8	+7.46%	+10.34%
$F^{Imp}$ (kN)	4.86	6.92	4.5	7.0	4.7	7.1	+3.40%	–02.60%
Bending (mm)	5.26	4.48	5.9	4.6	6.0	4.7	–12.34%	–04.68%
CAREA (mm <sup>2</sup> )	1106	957	1180	980	1090	1060	+01.47%	–09.72%

#### IV. CONCLUSION

In an effort to better understand the effects of the impactor head geometry upon low velocity impact, a three-dimensional finite element model was developed and linked to a VUMAT subroutine to simulate the impact response of the T700/Epoxy composite laminates when struck by the two suggested impactors. Based on the results of these studies, the following conclusions can be made [29]:

- The integration of the LARC damage model for material response has shown its effectiveness in determining very accurately the predicted key parameters of the impact event, such as (a) peak impact forces and bending, (b) duration of the impact event, (c) the stresses state, (d) the different mechanisms of damage, and (e) contact area shape (CAREA).
- The composite plate impacted by the Flat-Face Impactor (FFI) showed a higher impact force peak during impact, whilst suffering a lower out-of-plane bending than the Hemispherical Impactor (HemI).
- The HemI's contact time is considerable and may affect the extent of the damages.
- The analysis of the stress state allows clarification of the extent of the damage and how it is caused. Furthermore, the analysis reveals that the material has accumulated a proportion of the impact energy after the impact, which is highlighted through the existence of the residual stresses.
- The extent of the damaged area was slightly greater when using a HemI. Indeed, as the contact area increases, the depth of impact decreases, which implies the shape of the impactor head. For a small contact area impactor, the visual impact damages are worse on the frontal face than on a large face in the case of an equienergetic impact, as well as on the rear face.
- The determined impact responses and damage characteristics are an important factor to assess the effect of the impactor head shape on the low-velocity impact and how they correlate.

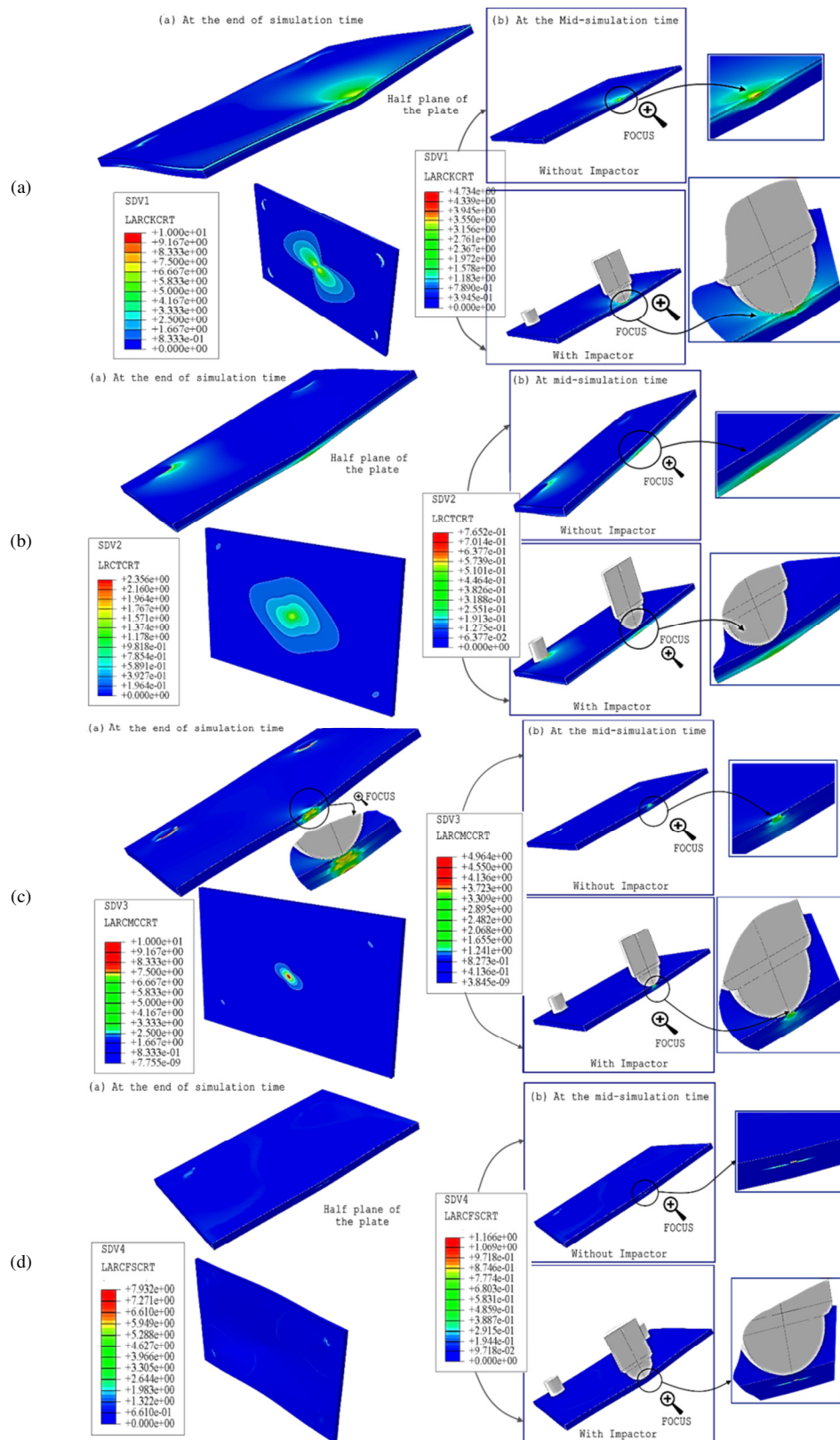


Fig. 7. Larc05 damage criterion contour illustration using VUMAT abilities (Heml). (a) Contours of fiber kinking criterion: SDV1 from VUMAT, (b) contours of fiber tension criterion: SDV2 from VUMAT, (c) contours of matrix cracking criterion: SDV3 from VUMAT, (d) contours of fiber splitting criterion: SDV4 from VUMAT.

- Finally, it can be said that the implemented model strongly agrees with the experimental findings of [1].

Our investigations have enhanced the understanding of the behavior of laminates under low velocity impact and also enabled the elaboration and validation of a three-dimensional finite element model based on the LARC criterion to predict the response of the T700/Epoxy laminate [30]. The proposed model will be important for future research into the predictability and the enhancement of the impact behavior of composite laminates.

## APPENDIX A

### ACRONYMS AND ABBREVIATIONS USED

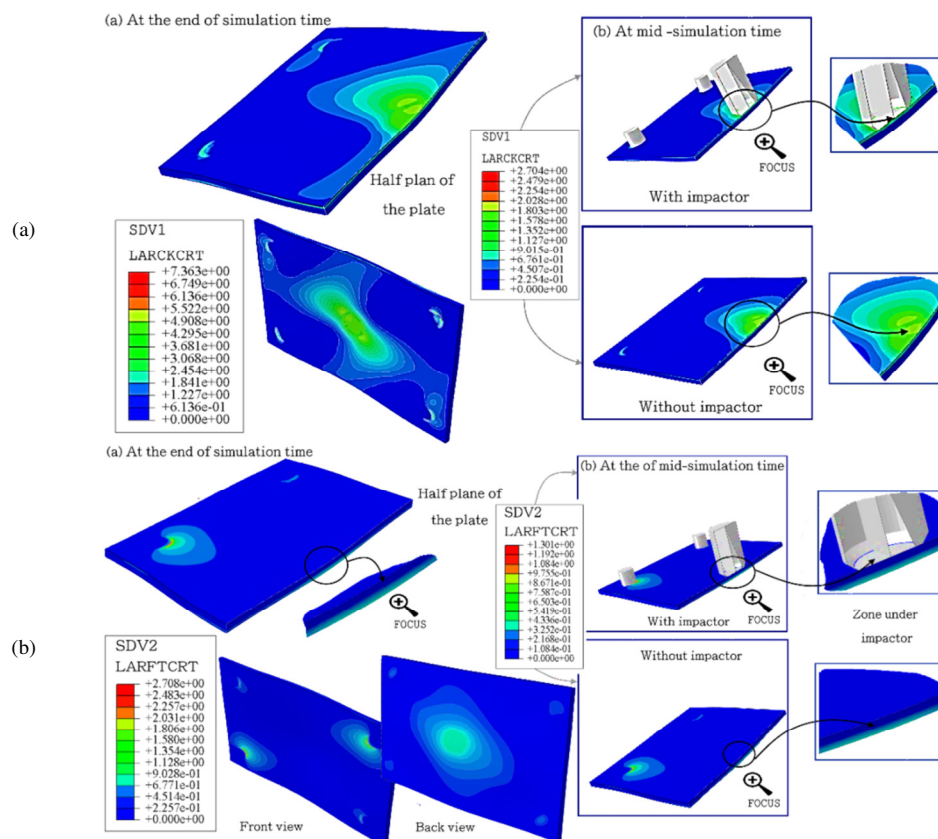
FEM	Finite Element Model
FFI	Flat Face head end Impactor
HemI	Hemispherical head end Impactor
IMP	Impact
Num	Numerical
EXP	Experimental
$\sigma^{res}$	Residual stress
VUMAT	User Material in Explicit case
CAREA	Contact Area
SDV	State Damage Variable
BCs	Boundary conditions

### CONSIDERED PHYSICAL QUANTITIES

Symbols	Physical quantity	Unit
$v$	Velocity	m/s
$H$	Layer thickness	mm
$L_i$	Layers indice	/
$\epsilon_{11}$	Young's modulus in X-direction	MPa
$\epsilon_{22}$	Young's modulus in Y-direction	MPa
$\epsilon_{33}$	Young's modulus in Z-direction	MPa
$G_{12}$	Shear modulus in XY- plane direction	MPa
$G_{13}$	Shear modulus in XZ- plane direction	MPa
$G_{23}$	Shear modulus in YZ- plane direction	MPa
$X^T$	Tensile strengths in X-direction	MPa
$Y^T$	Tensile strengths in Y-direction	MPa
$X^C$	Compressive strengths in X-direction	MPa
$Y^C$	Compressive strengths in Y-direction	MPa
$\tau_{xy}$	Transverse to fiber direction shear strengths ( $S_{12}$ )	MPa
$\epsilon_{xx}^T$	Tensile strains in X-direction	(%)
$\epsilon_{yy}^T$ (%)	Tensile strains in Y-direction	(%)
$\epsilon_{xx}^C$ (%)	Compressive strains in X-direction	(%)
$\epsilon_{yy}^C$ (%)	Compressive strains in Y-direction	(%)
$\epsilon_{xy}$ (%)	Transverse to fibre direction shear strains	(%)
$\nu_{12}$	Poisson's ratio modulus in XY- plane direction	
$\nu_{13}$	Poisson's ratio modulus in XZ- plane direction	
$\nu_{23}$	Poisson's ratio modulus in YZ- plane direction	
$v_f$	Volumetric fraction of fibers	(%)
$\rho$	Density	g/mm <sup>3</sup>
$S_{11}, S_{22}, S_{33}$	Main stresses in directions X, Y, and Z	MPa

## APPENDIX B

The FFI damages for the presented criterion mentioned in Section V are presented below.



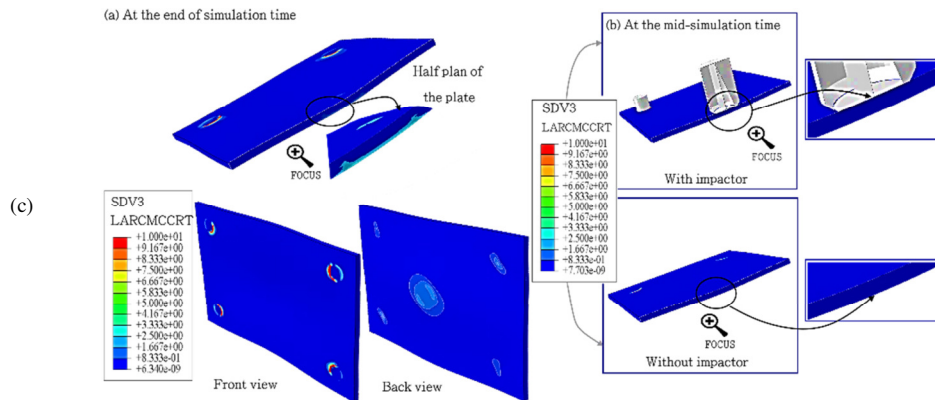
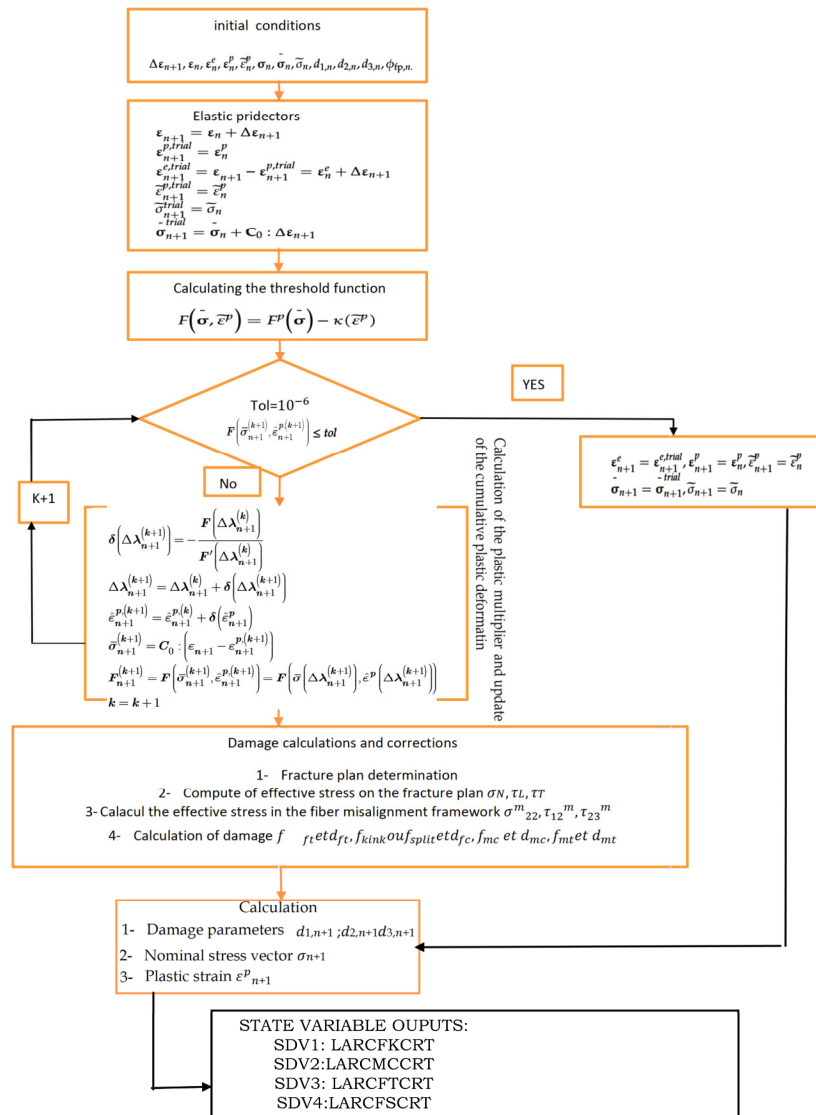


Fig. 8. Larc05 Damage Criterion contours illustration using VUMAT abilities in explicit analysis (FFI). (a) Contours of fiber kinking criterion: SDV1 from VUMAT, (b) contours of fiber tension criterion: SDV2 from VUMAT, (c) contours of matrix cracking criterion: SDV3 from VUMAT.

### APPENDIX C

The organigram of the LARC\_VUMAT subroutine used during simulations is presented below.





## REFERENCES

- [1] S. Agrawal, K. K. Singh, and P. Sarkar, "Impact damage on fibre-reinforced polymer matrix composite – A review," *Journal of Composite Materials*, vol. 48, no. 3, pp. 317–332, Feb. 2014, <https://doi.org/10.1177/0021998312472217>.
- [2] S. Hasebe, R. Higuchi, T. Yokozeki, and S. Takeda, "Internal low-velocity impact damage prediction in CFRP laminates using surface profiles and machine learning," *Composites Part B: Engineering*, vol. 237, May 2022, Art. no. 109844, <https://doi.org/10.1016/j.compositesb.2022.109844>.
- [3] K. Zouggar, F. B. Boukhoulad, B. Haddag, and M. Nouari, "Numerical and experimental investigations of S-Glass/Polyester composite laminate plate under low energy impact," *Composites Part B: Engineering*, vol. 89, pp. 169–186, Mar. 2016, <https://doi.org/10.1016/j.compositesb.2015.11.021>.
- [4] H. Cao *et al.*, "Experimental Investigation of Impactor Diameter Effect on Low-Velocity Impact Response of CFRP Laminates in a Drop-Weight Impact Event," *Materials*, vol. 13, no. 18, Jan. 2020, Art. no. 4131, <https://doi.org/10.3390/ma13184131>.
- [5] H. Liu, B. G. Falzon, and W. Tan, "Experimental and numerical studies on the impact response of damage-tolerant hybrid unidirectional/woven carbon-fibre reinforced composite laminates," *Composites Part B: Engineering*, vol. 136, pp. 101–118, Mar. 2018, <https://doi.org/10.1016/j.compositesb.2017.10.016>.
- [6] H. Liu *et al.*, "Effects of Impactor Geometry on the Low-Velocity Impact Behaviour of Fibre-Reinforced Composites: An Experimental and Theoretical Investigation," *Applied Composite Materials*, vol. 27, no. 5, pp. 533–553, Oct. 2020, <https://doi.org/10.1007/s10443-020-09812-8>.
- [7] T. Mitrevski, I. H. Marshall, R. Thomson, R. Jones, and B. Whittingham, "The effect of impactor shape on the impact response of composite laminates," *Composite Structures*, vol. 67, no. 2, pp. 139–148, Feb. 2005, <https://doi.org/10.1016/j.compstruct.2004.09.007>.
- [8] T. Mitrevski, I. H. Marshall, and R. Thomson, "The influence of impactor shape on the damage to composite laminates," *Composite Structures*, vol. 76, no. 1, pp. 116–122, Oct. 2006, <https://doi.org/10.1016/j.compstruct.2006.06.017>.
- [9] E. Sevkat, B. Liaw, and F. Delale, "Drop-weight impact response of hybrid composites impacted by impactor of various geometries," *Materials & Design (1980-2015)*, vol. 52, pp. 67–77, Dec. 2013, <https://doi.org/10.1016/j.matdes.2013.05.016>.
- [10] B. M. Icten, B. G. Kiral, and M. E. Deniz, "Impactor diameter effect on low velocity impact response of woven glass epoxy composite plates," *Composites Part B: Engineering*, vol. 50, pp. 325–332, Jul. 2013, <https://doi.org/10.1016/j.compositesb.2013.02.024>.
- [11] J. A. Artero-Guerrero, J. Pernas-Sanchez, J. Lopez-Puente, and D. Varas, "Experimental study of the impactor mass effect on the low velocity impact of carbon/epoxy woven laminates," *Composite Structures*, vol. 133, pp. 774–781, Dec. 2015, <https://doi.org/10.1016/j.compstruct.2015.08.027>.
- [12] K. Zouggar, K. Guerraiche, and A. Lousdad, "Numerical and predictive analysis of the low-velocity impact response of UD composite plate under a controlled environment," *Composite Structures*, vol. 299, Nov. 2022, Art. no. 116053, <https://doi.org/10.1016/j.compstruct.2022.116053>.
- [13] ASTM D7136/D7136M-15(2015), *Standard Test Method For Measuring The Damage Resistance Of A Fiber-Reinforced Polymer Matrix Composite To A Drop-Weight Impact Event*. West Conshohocken, PA, USA: ASTM International, 2015.
- [14] L. Raimondo, L. Iannucci, P. Robinson, and P. T. Curtis, "A progressive failure model for mesh-size-independent FE analysis of composite laminates subject to low-velocity impact damage," *Composites Science and Technology*, vol. 72, no. 5, pp. 624–632, Mar. 2012, <https://doi.org/10.1016/j.compscitech.2012.01.007>.
- [15] A. Kursun, M. Senel, H. M. Enginsoy, and E. Bayraktar, "Effect of impactor shapes on the low velocity impact damage of sandwich composite plate: Experimental study and modelling," *Composites Part B: Engineering*, vol. 86, pp. 143–151, Feb. 2016, <https://doi.org/10.1016/j.compositesb.2015.09.032>.
- [16] A. Ahmed and L. Wei, "The Low-velocity Impact Damage resistance of the Composite Structures - a Review," *Reviews on Advanced Materials Science*, vol. 40, pp. 127–145, Apr. 2015.
- [17] A. Kursun, M. Senel, and H. M. Enginsoy, "Experimental and numerical analysis of low velocity impact on a preloaded composite plate," *Advances in Engineering Software*, vol. 90, pp. 41–52, Dec. 2015, <https://doi.org/10.1016/j.advengsoft.2015.06.010>.
- [18] P. Gaudenzi, M. Bernabei, E. Dati, G. De Angelis, M. Marrone, and L. Lampani, "On the evaluation of impact damage on composite materials by comparing different NDI techniques," *Composite Structures*, vol. 118, pp. 257–266, Dec. 2014, <https://doi.org/10.1016/j.compstruct.2014.07.048>.
- [19] C. Bouvet, S. Rivallant, and J. J. Barrau, "Low velocity impact modeling in composite laminates capturing permanent indentation," *Composites Science and Technology*, vol. 72, no. 16, pp. 1977–1988, Nov. 2012, <https://doi.org/10.1016/j.compscitech.2012.08.019>.
- [20] R. Olsson, "Analytical model for delamination growth during small mass impact on plates," *International Journal of Solids and Structures*, vol. 47, no. 21, pp. 2884–2892, Oct. 2010, <https://doi.org/10.1016/j.ijsolstr.2010.06.015>.
- [21] R. Olsson, "Analytical prediction of damage due to large mass impact on thin ply composites," *Composites Part A: Applied Science and Manufacturing*, vol. 72, pp. 184–191, May 2015, <https://doi.org/10.1016/j.compositesa.2015.02.005>.
- [22] G. A. O. Davies and R. Olsson, "Impact on composite structures," *The Aeronautical Journal*, vol. 108, no. 1089, pp. 541–563, Nov. 2004, <https://doi.org/10.1017/S0001924000000385>.
- [23] S. Abrate, *Impact Engineering of Composite Structures*. New York, NY, USA: Springer, 2011.
- [24] J. Abd and I. K. Ahmed, "The Effect of Low Velocity Impact Loading on Self-Compacting Concrete Reinforced with Carbon Fiber Reinforced Polymers," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7689–7694, Oct. 2021, <https://doi.org/10.48084/etasr.4419>.
- [25] E. Troussset, "Prevision des dommages d'impact basse vitesse et basse energie dans les composites a matrice organique stratifies," Ph.D. dissertation, Ecole nationale superieure d'arts et metiers - ENSAM, Paris, France, 2013.
- [26] "Abaqus Scripting Reference Guide (6.13)." <http://130.149.89.49:2080/v6.13/books/ker/default.htm>.
- [27] K. Zouggar, "Etude du comportement mecanique d'une plaque composite de type verre/polyester soumise a un impact a faible energie," Ph.D. dissertation, Universite Djillali Liabes Sidi Bel Abbes, Sidi Bel Abbes, Algeria, 2016.
- [28] S. Li, E. Sitnikova, Y. Liang, and A.-S. Kaddour, "The Tsai-Wu failure criterion rationalised in the context of UD composites," *Composites Part A: Applied Science and Manufacturing*, vol. 102, pp. 207–217, Nov. 2017, <https://doi.org/10.1016/j.compositesa.2017.08.007>.
- [29] H. Q. Abbas and A. H. Al-Zuhairi, "Flexural Strengthening of Prestressed Girders with Partially Damaged Strands Using Enhancement of Carbon Fiber Laminates by End Sheet Anchorages," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8884–8890, Aug. 2022, <https://doi.org/10.48084/etasr.5007>.
- [30] J. A. Martins and E. C. Romao, "The Importance of Accurate Boundary Condition in Obtaining Reliable Shearing Stresses on a Torsional Finite Element Simulation," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8482–8487, Jun. 2022, <https://doi.org/10.48084/etasr.4708>.

# Stochastic Higher-order Finite Element Model for the Free Vibration of a Continuous Beam resting on Elastic Support with Uncertain Elastic Modulus

**Ta Duy Hien**

University of Transport and Communications, Vietnam  
tdhien@utc.edu.vn

**Nguyen Dang Diem**

Campus in Ho Chi Minh City, University of Transport and Communications, Vietnam  
diemnd\_ph@utc.edu.vn

**Giap Van Tan**

Hanoi Architectural University, Vietnam  
tangv@hau.edu.vn

**Vu Van Hiep**

University of Transport and Communications, Vietnam  
vvhiep@utc.edu.vn

**Nguyen Van Thuan**

Nha Trang University, Vietnam  
thuannv@ntu.edu.vn  
(corresponding author)

Received: 28 October 2022 | Revised: 3 December 2022 | Accepted: 4 December 2022

## ABSTRACT

This paper deals with a continuous beam resting on elastic support with elastic modulus derived from a random process. Governing equations of the stochastic higher-order finite element method of the free vibration of the continuous beam were derived from Hamilton's principle. The random process of elastic modulus was discretized by averaging random variables in each element. A solution for the stochastic eigenvalue problem for the free vibration of the continuous beam was obtained by using the perturbation technique, in conjunction with the finite element method. Spectral representation was used to generate a random process and employ the Monte Carlo simulation. A good agreement was obtained between the results of the first-order perturbation technique and the Monte Carlo simulation.

**Keywords-***SFEM; free vibration; continuous beam; random field*

## I. INTRODUCTION

Beams and frames are used in many engineering applications, such as civil engineering [1-3], bridge engineering, and aerospace engineering. The importance of dynamic problems has been the subject of many research studies [4-9]. In civil engineering, the structures resting on foundations are very important [10-15]. Considering continuous beams resting on elastic foundations, elastic support

is a common problem. Many practical problems are idealized to continuous beams resting on elastic foundations, elastic support such as rails in the railway, and strip foundations on the soil. The nonlinear dynamic of plates on a viscoelastic Winkler foundation under harmonic moving load is solved using the multiple time scales method [16]. Subway dynamic problem under loads from the ground and underground transport is modeled using finite element software in [17]. The dynamic of sandwich beams with a viscoelastic core subjected to a moving

load is investigated using the finite element method in [18]. The natural frequencies of non-uniform axially functionally graded beams were investigated using the Chebyshev collocation method in [19].

For stochastic problems, material properties, loads, and geometrical dimensions are assumed to be stochastic. Stochastic dynamics problems have been the subject of many studies [20-24]. There are several types of Stochastic Finite Element Methods (SFEM), e.g. the probabilistic finite element method [25, 26], the Spectral SFEMs (SSFEMs) using Karhunen-Loève expansion series [27], the homogeneous chaos expansion method [28] for the representation of random fields, and the SFEMs that use weighed integration techniques [29, 30]. The effect of randomness in elastic modulus on the stochastic free vibration of non-uniform beams is investigated by using an SFEM in [31]. Authors in [32] dealt with the stochastic dynamics of an infinite double beam resting on a random elastic foundation subjected to a moving load. Authors in [33] investigated a beam on a random foundation using the SFEM. The stochastic buckling of non-uniform columns was analyzed using stochastic finite elements in [34].

In the present work, the SFEM is used in the study of the free vibration of a continuous beam resting on an elastic support. The results are compared with the ones of the Monte Carlo simulation.

## II. STOCHASTIC FINITE ELEMENT FORMULATION FOR CONTINUOUS BEAM RESTING ON ELASTIC SUPPORT

Consider a continuous beam resting on elastic support with length  $L$  as shown in Figure 1.

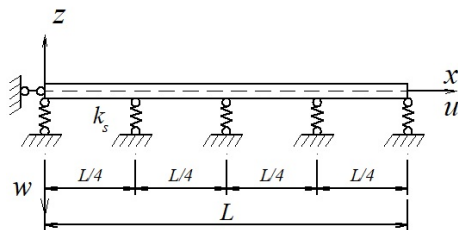


Fig. 1. Continuous beam resting on elastic support.

The displacement fields using Reddy's beam theory [35] present as formulation are:

$$u(x, z) = z\psi(x) + z^3 \left( -\frac{4}{3h^2} \right) \left( \psi(x) + \frac{\partial w_0}{\partial x} \right) \quad (1)$$

$$w(x, z) = w_0(x)$$

The strain formulation is calculated from (1) as follows:

$$\begin{aligned} \varepsilon_x &= \frac{\partial \psi}{\partial x} \left( z - \frac{4z^3}{3h^2} \right) - \frac{4z^3}{3h^2} \frac{\partial^2 w_0}{\partial x^2} \\ \gamma_{xz} &= \psi \left( 1 - \frac{4z^2}{h^2} \right) + \frac{\partial w_0}{\partial x} \left( 1 - \frac{4z^2}{h^2} \right) \end{aligned} \quad (2)$$

The displacement fields are approximated by an interpolation functions as follows:

$$\begin{cases} w_0(x) = N_1 w_1 + N_2 \theta_1 + N_3 w_2 + N_4 \theta_2 \\ \psi(x) = N_1^L \psi_1 + N_2^L \psi_2 \end{cases} \quad (3)$$

The interpolation functions are defined as follows:

$$\begin{cases} N_1 = 1 - 3\frac{x^2}{L^2} + 2\frac{x^3}{L^3} & N_3 = 3\frac{x^2}{L^2} - 2\frac{x^3}{L^3} \\ N_2 = x \left( 1 - 2\frac{x}{L} + \frac{x^2}{L^2} \right) & N_4 = x \left( -\frac{x}{L} + \frac{x^2}{L^2} \right) \\ N_1^L = 1 - \frac{x}{L} & N_2^L = 1 - \frac{x}{L} \end{cases} \quad (4)$$

The displacement vector of finite element is:

$$\{q\}_e = \{\psi_1 \quad w_1 \quad \theta_1 \quad \psi_2 \quad w_2 \quad \theta_2\}^T \quad (5)$$

The strain energy of beam element is given as:

$$U_e = \int_0^L \int_A \left\{ \frac{1}{2} E \left[ y \frac{\partial \phi_x}{\partial x} - \frac{4}{3h^2} y^3 \left( \frac{\partial \psi}{\partial x} + \frac{\partial^2 w_0}{\partial x^2} \right) \right]^2 + \frac{1}{2} G \left[ \left( 1 - \frac{4}{h^2} \right) \left( \psi + \frac{\partial w_0}{\partial x} \right) \right]^2 \right\} dx \quad (6)$$

The potential energy of elastic support is defined as:

$$U_s = \frac{1}{2} \sum k_s [w_0(x)]^2 \quad (7)$$

One dimensional homogeneous random field of elastic modulus is assumed:

$$E(x) = E_0 [1 + r_E(x)] \quad (8)$$

where  $r_E(x)$  is one dimensional homogeneous random field with zero mean. The autocorrelation function of the random field is:

$$R(\xi) = \sigma^2 \exp \left( -\frac{|\xi|}{d} \right) \quad (9)$$

where  $\sigma$ ,  $d$ , are the coefficient of variation and the correlation distance for the random field of an elastic modulus, and the relative distance vector  $\xi$  is defined as  $\xi = \mathbf{x}_j - \mathbf{x}_i$ .

The randomness in the elastic modulus of the beam element is illustrated in Figure 2.

$$E_i = E_0 (1 + r_i) \quad (10)$$

The random field of elastic modulus in the element is approximated by averaging random variables  $r_i$  at  $n$  points in the element:

$$E^e \approx E_0 \left[ 1 + \frac{r_1 + r_2 + \dots + r_n}{n} \right] \quad (11)$$

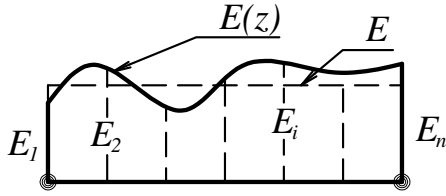


Fig. 2. Beam element with randomness in elastic modulus.

The stochastic stiffness matrix of the element is:

$$[K]^e = \int_{\Omega} [B][D][B] d\Omega$$

$$\approx \int_{\Omega} \left\{ [B][D]_0 \left[ 1 + \frac{r_1 + r_2 + \dots + r_n}{n} \right] [B] \right\} d\Omega \quad (12)$$

The stiffness matrix of the structures consists of the stiffness matrix of the beam element and the stiffness of the elastic foundation:

$$[K] = [K]_{beam} + [K]_{sup port} \quad (13)$$

The free vibration equation is:

$$([K] - \omega^2 [M])\{X\} = \{0\} \quad (14)$$

The eigenvalue of the free vibration is  $\lambda = \omega^2$ . The stiffness matrix, the eigenvector, and the eigenvalue are expanded to Taylor series as:

$$[K] = [K]_0 + \sum_{i=1}^{Nr} \frac{\partial [K]}{\partial r_{ei}} r_{ei} + \frac{1}{2} \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial^2 [K]}{\partial r_{ei} \partial r_{ej}} r_{ei} r_{ej} + \dots$$

$$\lambda = \lambda_0 + \sum_{i=1}^{Nr} \frac{\partial \lambda}{\partial r_{ei}} r_{ei} + \frac{1}{2} \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial^2 \lambda}{\partial r_{ei} \partial r_{ej}} r_{ei} r_{ej} + \dots \quad (15)$$

$$\{X\} = \{X\}_0 + \sum_{i=1}^{Nr} \frac{\partial \{X\}}{\partial r_{ei}} r_{ei} + \frac{1}{2} \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial^2 \{X\}}{\partial r_{ei} \partial r_{ej}} r_{ei} r_{ej} + \dots$$

and by substituting the series expressions in (6) into (5):

$$\left( [K]_0 + \sum_{i=1}^{Nr} \frac{\partial [K]}{\partial r_{ei}} r_{ei} + \frac{1}{2} \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial^2 [K]}{\partial r_{ei} \partial r_{ej}} r_{ei} r_{ej} + \dots \right) \times$$

$$\left( -\left\{ \lambda_0 + \sum_{i=1}^{Nr} \frac{\partial \lambda}{\partial r_{ei}} r_{ei} + \frac{1}{2} \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial^2 \lambda}{\partial r_{ei} \partial r_{ej}} r_{ei} r_{ej} + \dots \right\} [M] \right) \quad (16)$$

$$\left\{ \{X\}_0 + \sum_{i=1}^{Nr} \frac{\partial \{X\}}{\partial r_{ei}} r_{ei} + \frac{1}{2} \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial^2 \{X\}}{\partial r_{ei} \partial r_{ej}} r_{ei} r_{ej} + \dots \right\} = \{0\}$$

Equation (16) is solved by comparing the terms of the random variables to obtain the solution:

$$[[K]_0 - \lambda_0 [M]] X_0 = \{0\} \quad (17)$$

and

$$\frac{\partial \lambda}{\partial r_{ei}} = (X_0)^T \frac{\partial [K]}{\partial r_{ei}} X_0 \quad (18)$$

The first-order approximation of the variance of the eigenvalues is:

$$Var[\lambda(r_1, r_2, \dots, r_{Nr})]$$

$$\cong E[(\lambda - E[\lambda])(\lambda - E[\lambda])] \quad (19)$$

$$= \sum_{i=1}^{Nr} \sum_{j=1}^{Nr} \frac{\partial \lambda}{\partial r_{ei}} \frac{\partial \lambda}{\partial r_{ej}} R_E(\xi_{ij})$$

The variability of response COV of the eigenvalue is defined as:

$$COV = \frac{\sqrt{Var(\lambda)}}{E(\lambda)} \quad (20)$$

### III. NUMERICAL EXAMPLES

In this section, a continuous beam on 5 elastic supports, as shown in Figure 1, is investigated. The geometrical and material parameters of the beam are: length  $L=10\text{m}$ , height  $h=0.6\text{m}$ , mean of Young's modulus  $E=30 \times 10^5 \text{MPa}$ , coefficient of variation of a random field of elastic modulus  $\sigma=0.1$ , and mass density  $\rho=2400 \text{kg/m}^3$ . To valid the proposed approach, it is compared with Monte Carlo simulation. The random field of stiffness of the elastic modulus is generated using the spectral representation method [36, 37]:

$$r_E(x) = \sqrt{2} \sum_{n=0}^{N-1} A_n \cos(\omega_n x + \phi_n)$$

$$A_n = \sqrt{2 S_{ff}(\omega_n) \Delta \omega} \quad (21)$$

$$\Delta \omega = \frac{\omega_u}{N}$$

$$\omega_n = n \Delta \omega, n = 0, 1, 2, \dots, N-1$$

where  $\omega_u$  denotes the upper cut-off frequency beyond with the power spectral density function  $S_{ff}(\omega_n)$  of the random field of the stiffness of the elastic modulus.

The sample value of random field  $r_E(x)$  is substituted to (11) and the stiffness matrix in (12) is deterministic. The free vibration equation (14) is repeatedly solved 10000 times to get the eigenvalue. Figure 3 shows the first and second mode shapes with samples of 3 cases of stiffness of elastic support  $K_s = 10^6, 10^7, 10^8 \text{N/m}$ . It clearly shows that the first mode represents a symmetric mode shape and the second mode represents an antisymmetric mode shape. The mode shapes on Figure 3 clearly show the effect of the stiffness of support on mode shapes.

The effect of the correlation distance  $d$  of the random field on the variability of the eigenvalue is shown in Figure 4, where the results of the proposed formulation were compared with those of the Monte Carlo simulation using 10,000 samples with two cases of stiffness of elastic support, i.e.  $10^7$  and  $10^8 \text{N/m}$ . As seen in Figure 4, the COV varies depending on the correlation distance and the coefficient of stiffness of the elastic support. In all other cases, the obtained COV has a small value

of the assumed standard deviation of the random field of elastic modulus. Also, the result on Figure 4 shows the good agreement between the proposed approach and Monte Carlo simulation.

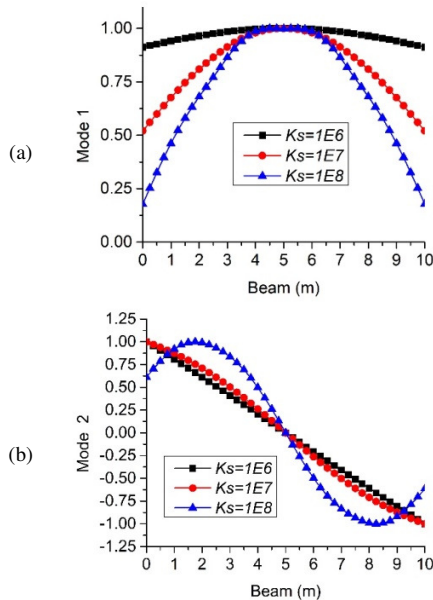


Fig. 3. Mode shapes. (a) First mode, (b) second mode.

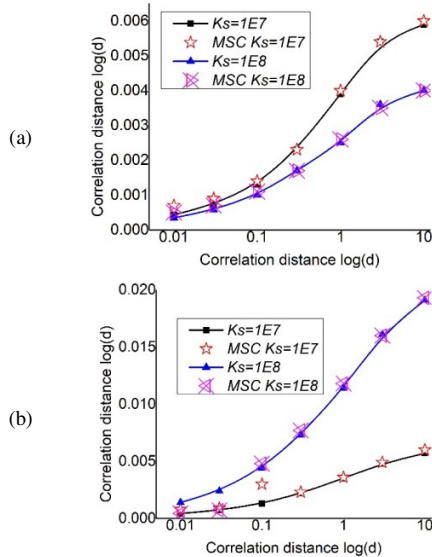


Fig. 4. Coefficient of variation of the eigenvalue. (a) First mode, (b) second mode.

Figure 5 illustrates the statistical probabilistic distribution of the eigenvalues of the first 2 modes obtained from Monte Carlo simulation corresponding with the stiffness of elastic support  $K_s = 10^7 \text{ N/m}$  and two cases of correlation distance  $d=0.01$  and  $d=0.1$ . On Figure 5, it is shown clearly that with smaller correlation distance, the histogram is quite scattered, whereas when the correlation distance is larger, the histogram is close to the normal distribution.

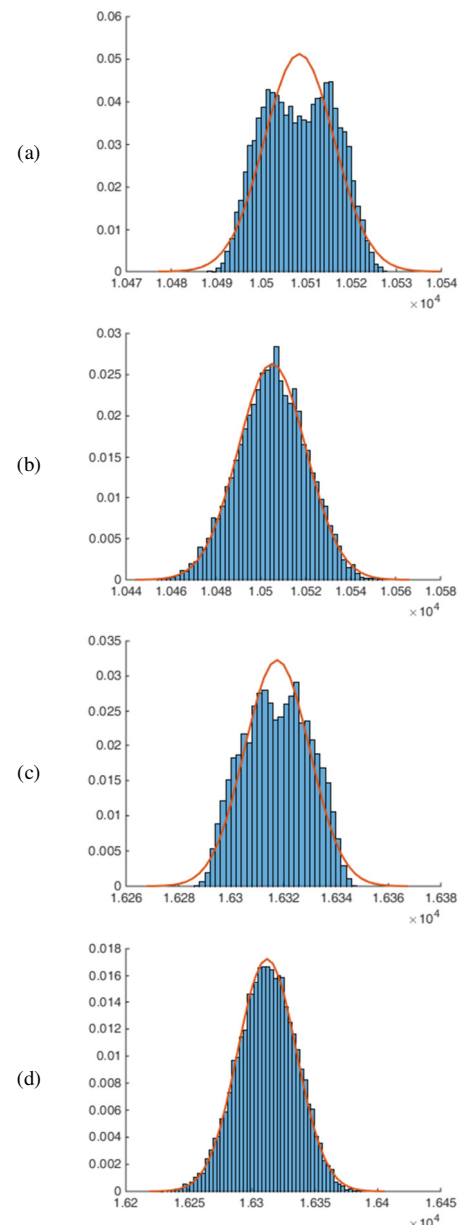


Fig. 5. Probabilistic distribution of the eigenvalue. (a) first mode with  $d=0.01$ , (b) first mode with  $d=0.1$ , (c) second mode with  $d=0.01$ , (d) second mode with  $d=0.1$ .

#### IV. CONCLUSIONS

In this paper, the stochastic finite element method was employed to investigate the response variability of the eigenvalue of the free vibration of a continuous beam on elastic support. The free vibration equation used higher order beam theory to establish finite element formulas. The first-order perturbation solution exhibits a good agreement with Monte Carlo simulation. The response variability of eigenvalues, given in terms of the coefficient of variation, is smaller of the input standard deviation of the random process of the elastic modulus of the beam.



## ACKNOWLEDGMENT

This research is funded by University of Transport and Communications (UTC) under the grant number T2021-CT-007TD.

## REFERENCES

- [1] P. C. Nguyen, "Nonlinear Inelastic Earthquake Analysis of 2D Steel Frames," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6393–6398, Dec. 2020, <https://doi.org/10.48084/etasr.3855>.
- [2] P. C. Nguyen, D. D. Pham, T. T. Tran, and T. Nghia-Nguyen, "Modified Numerical Modeling of Axially Loaded Concrete-Filled Steel Circular-Tube Columns," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7094–7099, Jun. 2021, <https://doi.org/10.48084/etasr.4157>.
- [3] P. C. Nguyen, B. Le-Van, and S. D. T. V. Thanh, "Nonlinear Inelastic Analysis of 2D Steel Frames: An Improvement of the Plastic Hinge Method," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5974–5978, Aug. 2020, <https://doi.org/10.48084/etasr.3600>.
- [4] D. X. Quy and N. V. Thi, "Static analysis of beam resting on elastic foundation by anisotropic beam-foundation element taking into account non-contact between beam and foundation," *Transport and Communications Science Journal*, vol. 72, no. 5, pp. 552–564, Jun. 2021, <https://doi.org/10.47869/tcsj.72.5.4>.
- [5] J. G. R. Iniguez, M. L. Daza-Torres, A. P. Gonzalez, and A. Cros, "Natural frequency of a heavy flexible plate: power law evolution as a function of length," *Latin American Journal of Solids and Structures*, vol. 18, no. 5, Jun. 2021, Art. no. e377, <https://doi.org/10.1590/1679-78256479>.
- [6] A. W. de Q. R. Reis, R. B. Burgos, and M. F. F. de Oliveira, "Nonlinear Dynamic Analysis of Plates Subjected to Explosive Loads," *Latin American Journal of Solids and Structures*, vol. 19, no. 1, Jan. 2022, Art. no. e422, <https://doi.org/10.1590/1679-78256706>.
- [7] V. T. A. Ninh, "Fundamental frequencies of bidirectional functionally graded sandwich beams partially supported by foundation using different beam theories," *Transport and Communications Science Journal*, vol. 72, no. 4, pp. 452–467, 2021.
- [8] P. M. Phuc, "Using phase field and third-order shear deformation theory to study the effect of cracks on free vibration of rectangular plates with varying thickness," *Transport and Communications Science Journal*, vol. 71, no. 7, pp. 853–867, 2020.
- [9] Q.-H. Pham, V. K. Tran, T. T. Tran, P.-C. Nguyen, and P. Malekzadeh, "Dynamic instability of magnetically embedded functionally graded porous nanobeams using the strain gradient theory," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 10025–10044, Dec. 2022, <https://doi.org/10.1016/j.aej.2022.03.007>.
- [10] Q.-H. Pham, T. Thanh Tran, V. Ke Tran, P.-C. Nguyen, and T. Nguyen-Thoi, "Free vibration of functionally graded porous non-uniform thickness annular-nanoplates resting on elastic foundation using ES-MITC3 element," *Alexandria Engineering Journal*, vol. 61, no. 3, pp. 1788–1802, Mar. 2022, <https://doi.org/10.1016/j.aej.2021.06.082>.
- [11] Q.-H. Pham, V. K. Tran, T. T. Tran, T. Nguyen-Thoi, P.-C. Nguyen, and V. D. Pham, "A nonlocal quasi-3D theory for thermal free vibration analysis of functionally graded material nanoplates resting on elastic foundation," *Case Studies in Thermal Engineering*, vol. 26, Aug. 2021, Art. no. 101170, <https://doi.org/10.1016/j.csite.2021.101170>.
- [12] T. T. Tran, P.-C. Nguyen, and Q.-H. Pham, "Vibration analysis of FGM plates in thermal environment resting on elastic foundation using ES-MITC3 element and prediction of ANN," *Case Studies in Thermal Engineering*, vol. 24, Apr. 2021, Art. no. 100852, <https://doi.org/10.1016/j.csite.2021.100852>.
- [13] Q.-H. Pham, P.-C. Nguyen, V.-K. Tran, and T. Nguyen-Thoi, "Finite element analysis for functionally graded porous nano-plates resting on elastic foundation," *Steel and Composite Structures*, vol. 41, no. 2, pp. 149–166, 2021, <https://doi.org/10.12989/scs.2021.41.2.149>.
- [14] P.-C. Nguyen, Q. H. Pham, T. T. Tran, and T. Nguyen-Thoi, "Effects of partially supported elastic foundation on free vibration of FGP plates using ES-MITC3 elements," *Ain Shams Engineering Journal*, vol. 13, no. 3, May 2022, Art. no. 101615, <https://doi.org/10.1016/j.jasej.2021.10.010>.
- [15] Q.-H. Pham, P.-C. Nguyen, and T. Thanh Tran, "Dynamic response of porous functionally graded sandwich nanoplates using nonlocal higher-order isogeometric analysis," *Composite Structures*, vol. 290, p. 115565, Jun. 2022, <https://doi.org/10.1016/j.compstruct.2022.115565>.
- [16] M. V. Shitikova and A. I. Krusser, "Force driven vibrations of nonlinear plates on a viscoelastic winkler foundation under the harmonic moving load," *International Journal for Computational Civil and Structural Engineering*, vol. 17, no. 4, pp. 161–180, Dec. 2021, <https://doi.org/10.22337/2587-9618-2021-17-4-161-180>.
- [17] M. Barabash and P. Bogdan, "Modeling of the subway dynamic influence on the ground structure," *International Journal for Computational Civil and Structural Engineering*, vol. 17, no. 3, pp. 14–23, Sep. 2021, <https://doi.org/10.22337/2587-9618-2021-17-3-14-23>.
- [18] Y. Karmi, Y. Khadri, S. Tekili, A. Daouadi, and E. M. Daya, "Dynamic Analysis of Composite Sandwich Beams with a Frequency-Dependent Viscoelastic Core under the Action of a Moving Load," *Mechanics of Composite Materials*, vol. 56, no. 6, pp. 755–768, Jan. 2021, <https://doi.org/10.1007/s11029-021-09921-w>.
- [19] W.-R. Chen, "Vibration Analysis of Axially Functionally Graded Timoshenko Beams with Non-uniform Cross-section," *Latin American Journal of Solids and Structures*, vol. 18, no. 7, Oct. 2021, Art. no. e397, <https://doi.org/10.1590/1679-78256434>.
- [20] T. D. Hien and P.-C. Nguyen, "Evaluation of Response Variability of Euler-Bernoulli Beam Resting on Foundation Due to Randomness in Elastic Modulus," in *International Conference on Sustainable Civil Engineering and Architecture*, Ho Chi Minh, Vietnam, Oct. 2019, pp. 1087–1092, [https://doi.org/10.1007/978-981-15-5144-4\\_105](https://doi.org/10.1007/978-981-15-5144-4_105).
- [21] H. D. Ta and P.-C. Nguyen, "Perturbation based stochastic isogeometric analysis for bending of functionally graded plates with the randomness of elastic modulus," *Latin American Journal of Solids and Structures*, vol. 17, no. 7, Sep. 2020, Art. no. e306, <https://doi.org/10.1590/1679-78256066>.
- [22] N. V. Thuan and T. D. Hien, "Variability in frequencies of vehicle vibration analysis with multiple random variables," *Transport and Communications Science Journal*, vol. 72, no. 2, pp. 215–226, 2021.
- [23] T.-T. Tran and D. Kim, "Uncertainty quantification for nonlinear seismic analysis of cabinet facility in nuclear power plants," *Nuclear Engineering and Design*, vol. 355, Dec. 2019, Art. no. 110309, <https://doi.org/10.1016/j.nucengdes.2019.110309>.
- [24] T.-T. Tran, K. Salman, S.-R. Han, and D. Kim, "Probabilistic Models for Uncertainty Quantification of Soil Properties on Site Response Analysis," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, vol. 6, no. 3, Sep. 2020, Art. no. 04020030, <https://doi.org/10.1061/AJRUA6.0001079>.
- [25] W. K. Liu, T. Belytschko, and A. Mani, "Random field finite elements," *International Journal for Numerical Methods in Engineering*, vol. 23, no. 10, pp. 1831–1845, 1986, <https://doi.org/10.1002/nme.1620231004>.
- [26] N. V. Thuan and T. D. Hien, "Stochastic Perturbation-Based Finite Element for Free Vibration of Functionally Graded Beams with an Uncertain Elastic Modulus," *Mechanics of Composite Materials*, vol. 56, no. 4, pp. 485–496, Sep. 2020, <https://doi.org/10.1007/s11029-020-09897-z>.
- [27] R. G. Ghanem and P. D. Spanos, "Spectral Stochastic Finite-Element Formulation for Reliability Analysis," *Journal of Engineering Mechanics*, vol. 117, no. 10, pp. 2351–2372, Oct. 1991, [https://doi.org/10.1061/\(ASCE\)0733-9399\(1991\)117:10\(2351\)](https://doi.org/10.1061/(ASCE)0733-9399(1991)117:10(2351)).
- [28] O. H. Galal, W. El-Tahan, M. A. El-Tawil, and A. A. Mahmoud, "Spectral SFEM analysis of structures with stochastic parameters under stochastic excitation," *Structural engineering and mechanics: An international journal*, vol. 28, no. 3, pp. 281–294, 2008.
- [29] T. D. Hien, B. T. Thanh, N. N. Long, N. Van Thuan, and D. T. Hang, "Investigation Into The Response Variability of A Higher-Order Beam Resting on A Foundation Using A Stochastic Finite Element Method," in *5th International Conference on Geotechnics, Civil Engineering Works and Structures*, 2020, pp. 117–122, [https://doi.org/10.1007/978-981-15-0802-8\\_15](https://doi.org/10.1007/978-981-15-0802-8_15).

- 
- [30] C.-K. Choi and H.-C. Noh, "Weighted Integral SFEM Including Higher Order Terms," *Journal of Engineering Mechanics*, vol. 126, no. 8, pp. 859–866, Aug. 2000, [https://doi.org/10.1061/\(ASCE\)0733-9399\(2000\)126:8\(859\)](https://doi.org/10.1061/(ASCE)0733-9399(2000)126:8(859)).
- [31] N. T. Nguyen, H. D. Ta, T. N. Van, and T. N. Dao, "Stochastic finite element analysis of the free vibration of non-uniform beams with uncertain material," *Journal of Materials and Engineering Structures*, vol. 9, no. 1, pp. 29–37, Apr. 2022.
- [32] S. Mohammadzadeh, M. Esmacili, and M. Mehrli, "Dynamic response of double beam rested on stochastic foundation under harmonic moving load," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 38, no. 6, pp. 572–592, 2014, <https://doi.org/10.1002/nag.2227>.
- [33] T.-P. Chang, "Dynamic finite element analysis of a beam on random foundation," *Computers & Structures*, vol. 48, no. 4, pp. 583–589, Aug. 1993, [https://doi.org/10.1016/0045-7949\(93\)90251-8](https://doi.org/10.1016/0045-7949(93)90251-8).
- [34] D. T. Hang, X. T. Nguyen, and D. N. Tien, "Stochastic Buckling Analysis of Non-Uniform Columns Using Stochastic Finite Elements with Discretization Random Field by the Point Method," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8458–8462, Apr. 2022, <https://doi.org/10.48084/etasr.4819>.
- [35] J. Reddy, *An Introduction to the Finite Element Method*, 3rd ed. New York, NY, USA: McGraw-Hill Education, 2005.
- [36] M. Shinozuka and G. Deodatis, "Simulation of Stochastic Processes by Spectral Representation," *Applied Mechanics Reviews*, vol. 44, no. 4, pp. 191–204, Apr. 1991, <https://doi.org/10.1115/1.3119501>.
- [37] P.-C. Nguyen, T. N. Van, and H. T. Duy, "Stochastic Free Vibration Analysis of Beam on Elastic Foundation with the Random Field of Young's Modulus Using Finite Element Method and Monte Carlo Simulation," in *6th International Conference on Geotechnics, Civil Engineering and Structures*, Ha Long, Vietnam, Oct. 2021, pp. 499–506, [https://doi.org/10.1007/978-981-16-7160-9\\_50](https://doi.org/10.1007/978-981-16-7160-9_50).

# Development of a Graphical User Interface for Reflection Loss Calculation in Perovskite-RGO based Microwave Absorbing Composites

Aayushi Arya

Department of Electrical Engineering, Indian Institute of Technology Hyderabad, India  
aayushi.arya@outlook.com  
(corresponding author)

Received: 7 November 2022 | Revised: 24 November 2022 | Accepted: 9 December 2022

## ABSTRACT

In this paper, a novel method is investigated wherein the theoretical and mathematical analysis of the perovskite-Reduced Graphene Oxide (RGO) based composite microwave absorber is used to form a machine learning model using linear regression to predict the reflection loss and the effective dielectric permittivity of a selected perovskite compound in an RGO-based composite. At first, the theoretical derivation is carried out to find a mathematical relationship between the reflection loss and the dielectric permittivity of the composite and the cationic radii of the perovskite structure, which is then used to form the base for the machine learning model to directly calculate the microwave absorption characteristics from the atomic parameters of the given composite structure. Linear regression is used for the machine learning algorithm which is verified with an  $R^2$  of 0.869 with the atomic radii as the input parameters. The model is further used to develop a Graphical User Interface (GUI) to make the prediction more appealing and user-friendly. The current paper provides a new approach to the integration of theoretical knowledge with advanced computing tools to form innovative predictive tools for current microwave-absorbing materials.

**Keywords-**atomic radius; microwave absorbers; machine learning; perovskite; reflection loss; reduced graphene oxide

## I. INTRODUCTION

Microwave absorbing composites have become an important field of research due to their growing applications. This field has seen a variety of research articles dedicated to the experimental studies of the absorbing properties of microwave absorbers [1-3]. There are many theoretical works also, modeling the permittivity and permeability of microwave absorbers [4-7], but they do not analyze their atomic fundamentals. In this work, an effort was conducted to form a machine learning model based on a theoretical framework for a microwave absorber that directly relates the reflection loss to the atomic radii of the elements in the observed compound. Further, a Graphical User Interface (GUI) was developed for predicting the reflection loss of perovskite-Reduced Graphene Oxide (RGO)-based microwave absorbing composites from their ionic radii using the machine learning model. The base equation and the electromagnetic parameters for the model are theoretically evaluated for the given absorber composite. The considered parameters include the combined dielectric permittivity of the filler and matrix material also referred to as the effective dielectric permittivity, the effective magnetic permeability, the radius of the filler particles considered as spherical inclusions, and the effective radius of anion and cation in the perovskite unit cell.

In this work, a perovskite-reduced graphene-based microwave absorbing composite is considered. A combination of perovskite filler and RGO-based matrix was chosen as the microwave absorbing composite. Perovskites are chosen as filler materials due to their good electrical and magnetic properties along with physical and chemical stability [8]. Similarly, RGO has gained popularity in microwave absorbing composites replacing the conventional polymer-based matrix materials [9], due to its high conductivity and excellent physical and chemical properties. Having the advantage of increased point defects and attached functional groups, RGO facilitates the occurrence of interfacial polarization and enhanced polarization sites thereby increasing the overall absorption rate. For an effective microwave absorption in a composite, there has to be good impedance matching for maximum wave absorption followed by high dissipation of the absorbed energy. The given model can be used to form new and efficient perovskite-based microwave absorbing compounds. Perovskite-based absorbers are cost effective, non-toxic, and environmentally friendly. Hence, they can be widely used as microwave absorbers in applications with large human contact like schools, hospitals, etc. They can be applied as microwave absorbing paints, coats, or integrated with absorber foams.

## II. POLARIZATION AND MAGNETISM

Polarizability is the tendency of an atom to allow charge separation when exposed to an electromagnetic wave. There are different polarization kinds that can occur inside a composite system including electronic, ionic, interfacial, and dipole relaxation-based polarizations [10-12]. As depicted in Figure 1(a), for a perovskite and graphene-based composite absorber, all the above phenomena happen at different points resulting in strong EM wave absorption. Dipole polarization occurs across the graphene layers due to the relaxation of dipoles formed between the graphene layers that are loosened with the penetration of perovskite particles between them. Similarly, in ABO<sub>3</sub>-type perovskite structures, ionic polarization occurs at the oxide ions formed by the presence of positively charged A-site cation deficiency. Interfacial polarization is induced by the interaction of the A-site cation with the matrix material that is graphene. This is caused by the interaction across a heterogeneous boundary creating a charge separation on both sides of the boundary. Electronic polarization is attributed to the flow of oxide ions in the 2D and 3D layers across the perovskite particles. The 2D oxide ion conduction occurs across the B-O-B plane, however the introduction of the graphene matrix between the perovskite structure has enabled 3D movement of the oxide ion. Graphene forms interfacial polarization with the A-site cation hence weakens its restricting force over the movement of oxide ions from one-unit cell to another. This phenomenon results in the formation of electronic polarization as well.

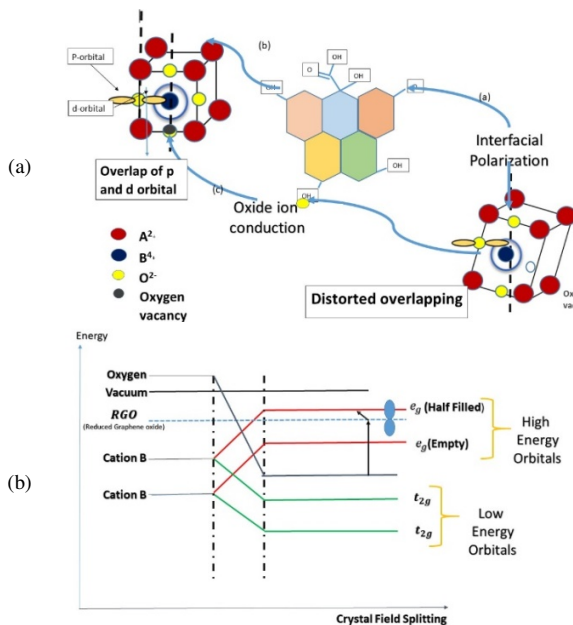


Fig. 1. (a) Polarization mechanism, (b) energy band diagram showing the occurrence of magnetism in perovskite-RGO composite microwave absorber.

The occurrence of magnetism in a perovskite compound occurs via double charge mechanism and involves the influence of RGO which forms a boundary material across the perovskite. The spin coupling in the perovskite filler particles is supported by the RGO layer. As shown in Figure 1(b), cation A

is not disturbed by the bond formation and provides a higher single energy level. On the other hand, cation B when combined with oxide ion and forming a B-O-B bond, splits its energy level into higher ( $e_g$ ) and lower ( $t_{2g}$ ) [13] energy orbitals. In a well ordered perovskite structure, the d-orbital of the B cation overlaps with the p-orbital of the oxide ion in a covalent manner with equal charge distribution. However, due to the presence of an additional energy level of RGO, the octahedron perovskite structure shifts from its center of symmetry. The perovskite oxide can show either ferromagnetic or magnetic properties when the electron charge distribution in the B-O-B bond is such that, one of the B cations is empty while the other is half filled following Goodenough-Kamamori's rule on superexchange interactions [14-15], thus increasing the overall magnetic permeability of the compound [23-24].

## III. THEORETICAL FRAMEWORK

The effective relative permittivity can be calculated from the Maxwell Garnett equation [16]:

$$\epsilon_{eff} = \epsilon_m \frac{2\delta_i(\epsilon_i - \epsilon_m) + \epsilon_i + 2\epsilon_m}{2\epsilon_m + \epsilon_i - \delta_i(\epsilon_i - \epsilon_m)} \quad (1)$$

The dielectric permittivity for the matrix material (RGO) is taken as constant,  $\epsilon_m = 118$ . The maximum value of the dielectric constant is taken to compensate for the high temperature effects on account of the dissipation of the absorbed microwaves [17]. The volume fraction of the filler particles is taken as the critical percolation threshold for the filler particles in an infinite cubic lattice. Its value is 0.1454 in all three dimensions [18]. Similarly, the effective relative magnetic permeability ( $\mu_{eff}$ ) can be calculated from [19]:

$$\mu_{eff} = \mu_d \left( 1 + \frac{\delta_i \omega^2}{10 c^2} (a_{pc})^2 \epsilon_m \right) \quad (2)$$

where  $a_{pc}$  is the atomic radius of the filler inclusions. In this case it is taken as the pseudo cubic lattice constant of the perovskite structure [25].

To determine the pseudocubic lattice constant  $a_{pc}$  of the perovskite filler particles, the effective anion radius, accounting for the effect of oxygen vacancies must be calculated which in turn should be determined from an estimated size of the spherical filler particles from the given dielectric permittivity [20]. The atomic size of the filler particles which are assumed to be spherical can be calculated by the Clausius Mosotti equation:

$$\alpha = \left( \frac{\epsilon_i - 1}{\epsilon_i + 2} \right) a^3 \quad (3)$$

The value of  $\alpha$  is taken as  $2.36 \times 10^{-28}$  from [13] on the theoretical predictions of electric and magnetic parameters of microwave absorbers.

Rearranging (6), it can be rewritten as:

$$a^3 = \alpha \left[ \frac{\epsilon_i}{\epsilon_i - 1} + \frac{2}{\epsilon_i - 1} \right] \quad (4)$$

In (7), the first term accounts for the effect due to the electronic polarizability, while the second term represents the effect due to the dipole polarizability. Since the dipole

polarization is the dominant at microwave frequencies, the first term can be dropped giving the modified equation for the atomic radii of the spherical filler inclusion as:

$$a = \sqrt[3]{\alpha \left[ \frac{2}{\epsilon_i - 1} \right]} \quad (5)$$

This approximated value is used to calculate the effective oxygen anion radius in the perovskite unit cell:

$$r_{x_{eff}} = \frac{a}{2} - r_b \quad (6)$$

The overall pseudo cubic lattice constant of the perovskite structure can be calculated as:

$$a_{pc} = K(V_N)^S r_{av} \quad (7)$$

where  $V_N = V_A V_B V_X$  is the product of the valence electron of cations A and B and O anion respectively. Similarly  $r_{av}$  is the average of the atomic radii of the A, B, and O atoms.  $K$  and  $S$  are constants with values 2.45 and 0.09 for cubic perovskites. Equation (7) is used to find the effective magnetic permeability from (2). Effective permittivity and permeability are further used to calculate the input impedance of the absorber as follows:

$$Z_{in} = \sqrt{\frac{\mu_{eff} \mu_0}{\epsilon_{eff} \epsilon_0}} \quad (8)$$

where  $\mu_0$  and  $\epsilon_0$  are the permeability and permittivity of free space respectively. After finding all the above parameters, the final Reflection Loss (RL) is calculated by:

$$RL = 20 \log \left( \frac{Z_{in} - Z_0}{Z_{in} + Z_0} \right) \quad (9)$$

where  $Z_0$  is the characteristic impedance of the input medium taken as free space.

#### IV. MACHINE LEARNING MODEL-LINEAR REGRESSION

The multiple linear regression model is used to relate the  $RL$  which is the most important absorption characteristic of an absorber to the fundamental value of an atom that is its atomic radius. A plot between the cationic radii and the  $RL$  is presented in Figure 2 suggesting an almost linear relationship between the output and the input parameters. The straight lines connecting the different data points and the shaded area indicate the change in  $RL$  with the change in individual atomic radius. The flowchart in Figure 3 shows the flow of the given work and how the different parameters are ultimately fed to the machine learning algorithm. In order to first apply the multiple linear regression model, the validation of the significance of the input features is conducted by applying the Ordinary Least Square (OLS) method. There are two results in the OLS model summary that are of interest, the coefficient of determination ( $R^2$ ) and the p-value.  $R^2$  measures the amount of variation in the output feature captured by the corresponding variations in the input parameters while the p-value provides a conditional probability of a null output hypothesis.

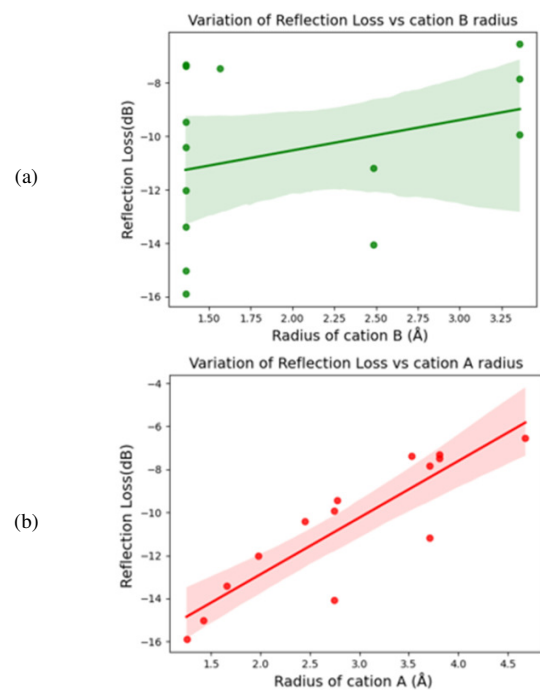


Fig. 2. Linear relationship between the  $RL$  and the cationic radii.

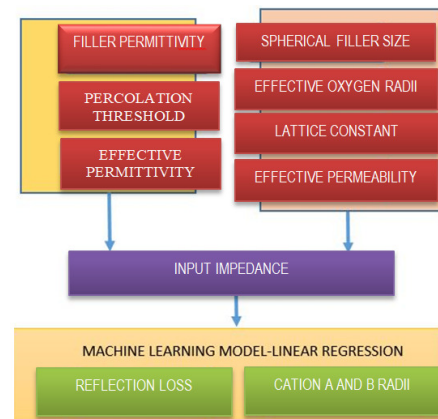


Fig. 3. Step flow and parameter merging of the machine learning algorithm.

If the p-value is less than 0.05, the input features do not fulfil the null hypothesis. In the OLS model summary for the current model, both the above two requirements are met, thus both of our input features have significant effect on the output and can be successfully used in a linear regression relation with the output. After validating the parameters, multiple linear regression is applied. The linear regression forms a linear equation between the input and the output parameters and provides the intercept with the coefficient of the individual input parameters. From the given equation,  $RL$  can be predicted for the given dataset, and can be compared with the actual result. Table I describes the dataset [20] for the various ABO<sub>3</sub> type perovskite materials with their respective permittivity values. This dataset was used for the calculations and the formation of the empirical results.



TABLE I. INPUT AND CALCULATED ELECTROMAGNETIC PARAMETERS OF PEROVSKITE STRUCTURES

Material	$\epsilon_i$	$r_a$	$r_b$	$r_{av}$	$\epsilon_{eff}$	$a$	$V_N$	$a_{pc}$	$\mu_{eff}$	$Z_{in}$	$RL$	$RL_{pred}$
CaTiO <sub>3</sub>	160	10.99	9.94	2.31	1170	1.437	48	8.022	3893.957	687.12	-14.07	-11.33
CaZrO <sub>3</sub>	27	10.99	16.79	2.71	957.3	2.628	48	9.4	5346.693	890.31	-9.93	-11.97
SrZrO <sub>3</sub>	30	18.57	16.79	3.32	964.1	2.534	48	11.524	8035.893	1087.63	-7.84	-9.13
BaZrO <sub>3</sub>	35	28.07	16.79	3.92	975	2.403	48	13.607	11203.44	1277.02	-6.55	-6.28
LaGaO <sub>3</sub>	27	22.86	6.26	3.42	957.3	2.628	54	11.984	8690.231	1135.05	-7.47	-7.53
SrTiO <sub>3</sub>	300	18.57	9.94	2.86	1292	1.164	48	9.941	5979.781	810.4	-11.19	-8.49
NdAlO <sub>3</sub>	22.5	14.68	4.08	2.57	947	2.8	54	8.998	4899.176	856.86	-10.42	-11.59
LaAlO <sub>3</sub>	23	22.86	4.08	3.47	948.2	2.779	54	12.159	8945.896	1157.14	-7.31	-7.38
PrAlO <sub>3</sub>	23	16.67	4.08	2.78	948.2	2.779	54	9.746	5747.551	927.5	-9.45	-10.42
ErAlO <sub>3</sub>	16	7.51	4.08	1.89	931.6	3.157	54	6.62	2651.9	635.6	-15.9	-14.91
DyAlO <sub>3</sub>	18	8.56	4.08	1.96	936.4	3.028	54	6.876	2860.948	658.49	-15.02	-14.4
GdAlO <sub>3</sub>	18	9.94	4.08	2.11	936.4	3.028	54	7.416	3327.942	710.19	-13.39	-13.72
SmAlO <sub>3</sub>	20	11.85	4.08	2.29	941.2	2.918	54	8.034	3905.686	767.42	-12.03	-12.78
YAlO <sub>3</sub>	16	17.64	4.08	3.40	931.6	3.157	54	11.942	8629.466	1146.57	-7.38	-8.21

Reflection Loss vs Atomic Radii Model Visualization

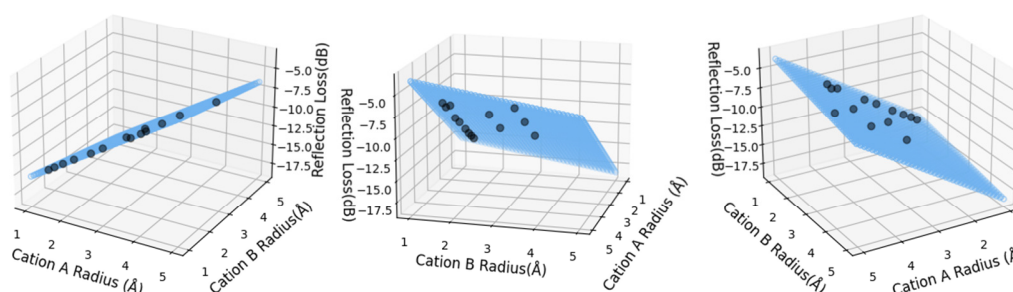


Fig. 4. 3-D model visualization of the given machine learning model.

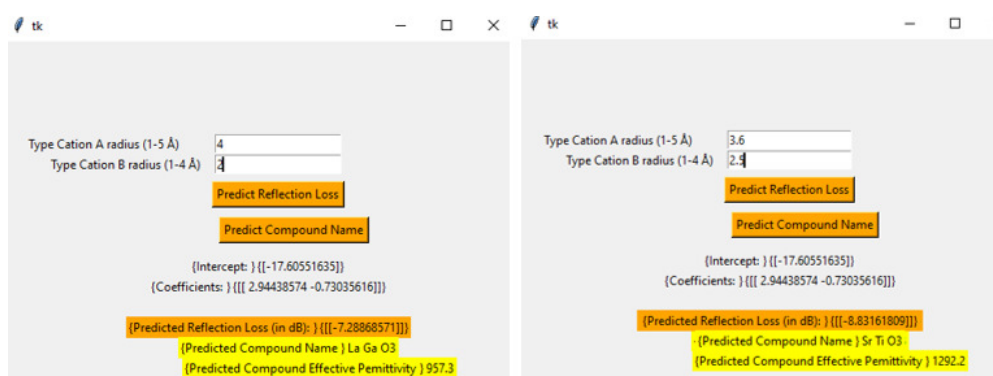


Fig. 5. The developed GUI.

All the values provided in Table I are used to calculate the various electromagnetic parameters derived in the previous theoretical framework which are also tabulated in Table I along with the theoretical and predictive values of the  $RL$ .

3D graphical visualization of the variation in the  $RL$  with respect to cation A and B radii is presented in Figure 4. The inclusion of the two cationic radii as input variables provides a rectangular model area, which covers all the data points provided for model training. The data plots are well contained within the model scope. From the given machine learning model, a GUI is developed, as shown in Figure 5, for the calculation of the  $RL$  of the perovskite compound and its effective permittivity based on the atomic radii provided by the

user. The GUI takes the cations' A and B radii as input and suggests a suitable microwave absorbing perovskite compound. It also shows the output  $RL$  and the effective permittivity based on the machine learning predictive model. The GUI is tested by taking two sets of input atomic radii and the output results are matched with the results of Table I. The results are in good accordance showing the high accuracy of the developed GUI.

TABLE II. MODELING PARAMETERS OF THE PROPOSED LINEAR REGRESSION MODEL

Intercept	Coefficients
-17.605	[-2.944 -0.73]

## V. CONCLUSION

In the present work, a new and innovative method has been suggested of utilizing a theoretical framework to develop a machine learning GUI for estimating the reflection loss in microwave absorbers. This work includes the diverse mechanisms and phenomena working behind the application of microwave absorption in a single equation using machine learning tools. This helps not only the simplification of the accurate calculation of reflection loss without going into huge depths of the physical processes of polarization, magnetism, and dissipation across the absorber, but also opens a new window to develop advanced software and algorithms devoted to the formation of new and complex kinds of microwave absorbers. Using this approach, the microwave absorber can become a part of more general engineering. It is time we take advantage of the true potential of microwave absorbers because such materials can be used for the benefit of society, as in hospitals, schools, libraries, etc. reducing the harmful effects of the EM wave radiation exposure in the ambient surroundings.

## VI. DATA AVAILABILITY STATEMENT

The data that support the findings of this study can be accessed at [22].

## NOMENCLATURE

Parameter	Notation
Effective relative permittivity	$\epsilon_{eff}$
Effective relative permeability	$\mu_{eff}$
Filler permittivity	$\epsilon_i$
Radii of the spherical filler particle	$a$
Effective radius of oxygen anion	$r_{xeff}$
Pseudo cubic Lattice constant	$a_{pc}$
Volume fraction of filler inclusion	$\delta_i$
Product of valence electron of cation A, B, and O	$V_N = V_A V_B V_X$
Average atomic radius of the A, B, and O atoms	$r_{av}$
Input impedance of the absorber	$Z_{in}$
Reflection Loss (dB)	$RL$

## REFERENCES

- [1] M. Green and X. Chen, "Recent progress of nanomaterials for microwave absorption," *Journal of Materiomics*, vol. 5, no. 4, pp. 503–541, Dec. 2019, <https://doi.org/10.1016/j.jmat.2019.07.003>.
- [2] D. Micheli *et al.*, in *Advances in Nanocomposites - Synthesis, Characterization and Industrial Applications*, B. Reddy, Ed. IntechOpen, 2011, pp. 359–384, <https://doi.org/10.5772/15215>.
- [3] A. Arya and I. Srikanth, "Design and Modelling of Carbon Fiber Grid Structure based Carbon/Epoxy Composites for Enhanced Microwave Absorbing Properties," *Advanced Materials Letters*, vol. 11, no. 11, pp. 1–6, Nov. 2020, <https://doi.org/10.5185/amlett.2020.111577>.
- [4] Y. Liu, K. Zhao, M. G. B. Drew, and U. Liu, "A theoretical and practical clarification on the calculation of reflection loss for microwave absorbing materials," *AIP Advances*, vol. 8, no. 1, 2018, <https://doi.org/10.1063/1.4991448>.
- [5] Z. Jia *et al.*, "Progress in low-frequency microwave absorbing materials," *Journal of Materials Science: Materials in Electronics*, vol. 29, no. 20, pp. 17122–17136, Oct. 2018, <https://doi.org/10.1007/s10854-018-9909-z>.
- [6] A. Houbi, Z. A. Aldashevich, Y. Atassi, Z. Bagasharova Telmanovna, M. Saule, and K. Kubanych, "Microwave absorbing properties of ferrites and their composites: A review," *Journal of Magnetism and Magnetic Materials*, vol. 529, Jul. 2021, Art. no. 167839, <https://doi.org/10.1016/j.jmmm.2021.167839>.
- [7] X. Zeng, X. Cheng, R. Yu, and G. D. Stucky, "Electromagnetic microwave absorption theory and recent achievements in microwave absorbers," *Carbon*, vol. 168, pp. 606–623, Oct. 2020, <https://doi.org/10.1016/j.carbon.2020.07.028>.
- [8] M. Tyunina, "Oxygen Vacancies in Perovskite Oxide Piezoelectrics," *Materials*, vol. 13, no. 24, Dec. 2020, Art. no. 5596, <https://doi.org/10.3390/ma13245596>.
- [9] Y. Liu, H. Yu, M. G. B. Drew, and Y. Liu, "A systemized parameter set applicable to microwave absorption for ferrite based materials," *Journal of Materials Science: Materials in Electronics*, vol. 29, no. 2, pp. 1562–1575, Jan. 2018, <https://doi.org/10.1007/s10854-017-8066-0>.
- [10] T. M. N. Abd Aziz, H. Mohamed Kamari, S. Shafinas, and W. Daud, "Polarizability and optical basicity of Er<sup>3+</sup> ions doped tellurite based glasses," *Chalcogenide Letters*, vol. 11, pp. 319–335, Jul. 2014.
- [11] A. Moliton, *Basic Electromagnetism and Materials*. New York, NY, USA: Springer, 2007.
- [12] R. Schmidt, E. Langenberg, J. Ventura, M. Varela, and J. Zhang, "Bi Containing Multiferroic Perovskite Oxide Thin Films," in *Perovskite: Crystallography, Chemistry and Catalytic Performance*, Nova Science Publishers, 2013.
- [13] A. Arya and G. V. V. Sharma, "Prediction of material composition for microwave absorption through mathematical modelling," *Journal of Physics Communications*, vol. 5, no. 12, Sep. 2021, Art. no. 125002, <https://doi.org/10.1088/2399-6528/ac37a7>.
- [14] Y. Wu *et al.*, "Oxygen vacancies regulated microwave absorption properties of reduced graphene oxide/multi-walled carbon nanotubes/cerium oxide ternary nanocomposite," *Journal of Alloys and Compounds*, vol. 819, Apr. 2020, Art. no. 152944, <https://doi.org/10.1016/j.jallcom.2019.152944>.
- [15] A. Arya and G. Sharma, "Effect of Elemental Atomic Radii on the Attenuation Constant of Microwave Absorber: A Theoretical Analysis," in *2022 International Conference for Advancement in Technology (ICONAT)*, Goa, India, Jan. 2022, <https://doi.org/10.1109/ICONAT53423.2022.9726063>.
- [16] O. Levy and D. Stroud, "Maxwell Garnett theory for mixtures of anisotropic inclusions: Application to conducting polymers," *Physical Review B*, vol. 56, no. 13, pp. 8035–8046, Oct. 1997, <https://doi.org/10.1103/PhysRevB.56.8035>.
- [17] C. Yuwen, B. Liu, L. Zhang, S. Guo, and J. Peng, "Synthesis high-quality graphene oxide and temperature-dependent dielectric properties of reduced graphene oxide," *Materials Research Express*, vol. 6, no. 9, Apr. 2019, Art. no. 0950b4, <https://doi.org/10.1088/2053-1591/ab149f>.
- [18] Y. Y. Tarasevich and E. N. Manzhosova, "On site percolation on the correlated simple cubic lattice," *International Journal of Modern Physics C*, vol. 14, no. 10, pp. 1405–1412, Dec. 2003, <https://doi.org/10.1142/S0129183103005480>.
- [19] B. A. Belyaev and V. V. Tyurnev, "Electrodynamic Calculation of Effective Electromagnetic Parameters of a Dielectric Medium with Metallic Nanoparticles of a Given Size," *Journal of Experimental and Theoretical Physics*, vol. 127, no. 4, pp. 608–619, Oct. 2018, <https://doi.org/10.1134/S1063776118100114>.
- [20] M. Sebastian, *Dielectric Materials for Wireless Communication*, 1st ed. Paris, France: Elsevier, 2008.
- [21] H. Ohsato, "Microwave Dielectrics with Perovskite-Type Structure," in *Perovskite Materials - Synthesis, Characterisation, Properties, and Applications*, L. Pan and G. Zhu, Eds. IntechOpen, 2016, <https://doi.org/10.5772/61718>.
- [22] A. Arya, "MicrowaveAbsorberGUI." Accessed: Dec. 13, 2022. [Online]. Available: <https://github.com/aayushiarya25/MicrowaveAbsorberGUI>.
- [23] A. E. Irmak, "Structural and Electrical Properties of Ca<sup>2+</sup> Doped LaFeO<sub>3</sub>: The Effect of A-site Cation Size Mismatch," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5538–5546, Apr. 2020, <https://doi.org/10.48084/etasr.3443>.
- [24] S. C. A. Bikkina and P. V. Y. Jayasree, "Development of a Wire Mesh Composite Material for Aerospace Applications," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9310–9315, Oct. 2022, <https://doi.org/10.48084/etasr.5201>.

- [25] S. Boumous, S. Belkhiat, and F. Kharchouche, "MgO Effect on The Dielectric Properties of BaTiO<sub>3</sub>," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4092–4099, Jun. 2019, <https://doi.org/10.48084/etasr.2705>.

## AUTHORS PROFILE



**Aayushi Arya** is currently pursuing her Phd from the EE Department, Indian Institute of Technology, Hyderabad, India. With keen interest in fundamental and physical concepts, I have done research works in deriving the theoretical and mathematical framework of microwave absorbers followed by the integration of the derived concepts in advanced modeling tools such as machine learning.

# A Brief Review on Cloud Computing Authentication Frameworks

**Abdul Raouf Khan**

Department of Computer Sciences, King Faisal University, Saudi Arabia  
raoufkhan@kfu.edu.sa  
(corresponding author)

**Latifa Khalid Alnwiheh**

Aljaber Finance, Saudi Arabia  
latifaalnwiheh@gmail.com

*Received: 10 November 2022 | Revised: 23 November 2022 | Accepted: 26 November 2022*

## ABSTRACT

Cloud computing is among the most widely used technologies worldwide as it helps individual users and organizations to store and share information conveniently and cost-effectively. However, data security is a major concern in cloud computing. Security involves guaranteed access to the data only to authorized users and protection against various threats and attacks. Security is generally achieved through the appropriate and efficient implementation of access control, authentication, and authorization mechanisms. Various frameworks have been proposed and implemented for authentication and access control in cloud computing. This paper discusses some of the main authentication frameworks in cloud computing, highlighting their methodologies, algorithms, and problems and identifying the issues that should form the focus of research in the area.

*Keywords-cloud security; cloud computing; security framework; cloud authentication*

## I. INTRODUCTION

The reliance on cloud for storage and computing makes it a prominent target for online attacks. Sharing resources over the Internet puts shared data at risk of being compromised or misused. Thus, users are not comfortable sharing sensitive data online and it is important to ensure that they clearly understand the security-related requirements [1]. Regardless of the clear risks, customers use the cloud due to its many convenient features [2]. Therefore, the security of the cloud needs to be improved. Many researchers have considered various security issues related to cloud computing and proposed various ways to improve it [3-4]. The following sections consider and categorize the research on security-related issues concerning the cloud.

## II. AUTHENTICATION IN CLOUD COMPUTING

Authentication is an important aspect of cloud computing in the context of data security. Authentication in an information system can be achieved in multiple ways, for example, based on something the user "has" such as smartcards, something he "knows" such as passwords, and/or something the user "is" such as biometrics (fingerprint, iris scan, etc.).

### A. Biometric Authentication Frameworks

In [5], a biometric-based authentication scheme was proposed to offer secure user identification and authentication

using elliptic curve cryptography for the generation and exchange of secure keys. An authentication scheme based on a contactless smart card was proposed for the user to improve security and avoid problems with key management. In the registration phase, the user must enter his biometric data, which are recorded in the service provider database. The user identifier was based on the username and the data associated with the server and was used to verify his identity with the cloud service provider. The latter collects the user's identifier, biometric data, and hashed password to store in the smart card during the registration phase. Operation is assumed to be supervised by a monitor, and a proxy is issued at the end of each access case. The reader's sensor had a liveness detector for the data to confirm that they were not collected from an earlier scan. Furthermore, the transferred data were encrypted to prevent attackers from obtaining or tracking them, and data transmission occurred only if the server and the customer met the validation criteria. The biometric scheme involved two phases: biometric verification and biometric linking. To avoid the introduction of a replay attack and guarantee the creation of a session, the nonce value is generated during the handshake message. When a user attempts to log in by inserting his biometric data, a Biohash is generated and associated with their ID and known password. To generate a more secure biometric pattern, the Biohash is XORed with the hashed password and then the smartcard compares its data with the hashed password. The authentication server matches the biometric data and the

hashed password with the relevant data in the database and checks the client identifier. Following this, the authentication server generates an encrypted session key that is transferred with a hashed value. By computing the nonce value, the client can decrypt the data sent by the server and verify its origin. The client can then prove their validity by submitting a signed "hello" message with the nonce value. The authentication server then validates the message and submits proxy credentials to access the service. This scheme has low computational complexity, requires scant communication, is scalable over a large range of numbers of clients, and involves little computation on the client side, which renders it suitable for a cloud environment. However, the complexity of the used hash algorithms is high.

In the context of multimodal biometric systems, a framework for biometrics-based mobile cloud computing was proposed in [6], using two-step verification based on fingerprints and iris recognition. If the user cannot be authenticated through these biometric techniques, the system supports a backup authentication code. The fingerprint recognition system takes an image of the user's fingertip through sensors. The iris recognition system scans their iris using a high-definition camera. The system converts the image of the user's fingertip to obtain a ridge structure, and their iris scan is matched with the available database. Images of the fingertip and the iris of the user are recorded in each access case and compared with the saved input to authenticate them. The user's fingerprint is validated before their iris scan. Backup code-based verification serves as an emergency access mode, as the user is given five to ten unique codes at registration, where each is valid for a single use to access the system. This method helps maintain data security in the cloud. The user can generate new unique codes by using the system, which notifies him of each code generation instance. However, the backup codes are at risk of attacks, and code maintenance induces complexity from the user's perspective.

#### B. Password-based Authentication Frameworks

In [7], a model was proposed that unified the mathematical approaches to single- and two-server Password-based Authentication Systems (PAS). The reliability and availability of security services were measured to ensure secure data access through Web applications. A single authentication server was involved in the single-server architecture to authenticate the client, storing a dictionary of all client passwords in its database, which must be available at the time of operation to guarantee successful client authentication. If the server fails, a single point of failure occurs that threatens the entire authentication system. To recover the system in this case, a redundant offline server with an updated database of customer passwords can be used. Two servers, a front- and a back-end authentication server, were used in this architecture. The former acts as a public server, while the latter is responsible for maintaining an updated dictionary of client passwords. The client's authentication requests are sent only to the front-end authentication server. The functional behaviors of the system were represented as Generalized Stochastic Petri Nets (GSPNs) to visualize the impact of the authentication service, considering scaled-up/down versions of the authentication

servers by exploiting a diversity of backup approaches. Hot and cold standby were used to improve the reliability of the system. Platform-Independent Petri Nets (PIPE) were used to generate a system reachability graph to derive the equivalent Continuous-Time Markov Chain (CTMC). The availability and reliability of the system were estimated by applying traditional Markovian state-based analysis, and the performance-related attributes of the authentication framework were monitored. Experimental results showed that the cold standby strategy can be used to improve the availability of the two-server PAS in case of attacks to a greater extent than the single-server PAS, while the reliability of the single-server PAS is higher than that of the two-server.

A multi-level authentication method was proposed in [8] that allowed access by authenticating the user's password at three levels. Organizational authentication was first conducted by verifying that the user has the privileges to access the services provided, where a failure terminated the session. If organizational authentication succeeded, team-level authentication followed, where access was provided to specific cloud services. The third level, user authentication, was then performed, where specific permissions and privileges were granted to the user. The user can access cloud services only if the first two levels of authentication have been satisfied. This method can be used within an organization's cloud storage.

#### C. Image-based Authentication Frameworks

A three-step data security model was proposed in [9] using cryptography and steganography during data storage and sharing in the cloud. The model was composed of three levels. The first level employed cryptography using RSA to prevent unauthorized data access. The second level used steganography, where the data were hidden in an image in the form of ciphertext using StegoTools by creating a symmetric key that was sent to the receiver. In the third authentication level, the data were accessed in the image and decrypted by the RSA. The receiver decrypted the image through StegoTools by specifying the symmetric key created during image steganography. The ciphertext was decrypted using the RSA algorithm to obtain the original plaintext. This model can be applied to audio, video, and text data.

#### D. Digital Certificate-based Authentication Frameworks

Data anonymization was proposed in [10] to prevent the misuse of user data by the cloud provider. This study focused on the trust issue between users and cloud providers based on the anonymity of shared data in the cloud. The idea was to conserve the capability of the cloud provider to charge for the use of reliable cloud services by cutting off the semantic linkage between the data and their owners. The required operations can then be performed by the cloud service provider on an anonymized dataset, following which they can convey the outcomes to the users and charge for the use of services while keeping their identity private. Cloud users can thus safely avail themselves of the services without being identified. This approach guarantees privacy in CC by using ring and group signatures. The study claimed that the anonymity of clients can be preserved using signatures. Typically, this means that each client has a signature that can be verified by the cloud provider by referring to a list of users while preserving their privacy, as



it cannot identify the member of the list to whom the signature in question belongs. This approach requires preparing a set of legitimate users willing to use the cloud resources. However, it is useful only if flat-rate accounting is allowed in the relevant business model. According to it, the user has to pay in advance for access (for a fixed period) to cloud services. When group signatures are used, the group manager is an actor in addition to the verifier and signer. Users can register and obtain credentials to prove that they belong to the set of registered clients. In case of a conflict, the group manager can revoke the anonymity of the group signature.

In [11], a model was proposed based on Elliptical Curve Cryptography (ECC), using a digital signature-based identification protocol for authentication. The study claimed that it does not require modular inversion, which improves its efficiency. However, the system required an algorithm to find the index by using a multiplicative group of finite fields.

#### *E. N-factor-based Authentication Frameworks*

##### *1) Two-factor Authentication*

An Advanced Encryption Standard–Cipher-text-Identity And Attribute-Based Encryption" (AES–CP–IDABE) was proposed in [12]. This system guaranteed the secure sharing of documents for temporally confidential data using two files: the original and a metadata file. The document's time of release and its attribute were stored on the server in a separate file. The secret keys of the metadata and the original document depended on each other. If the file needed to be deleted or updated, the user can do so without having to decrypt it. Furthermore, two-factor authentication was applied to prevent shoulder surfing and keylogger attacks and improve security. The first factor was a unique user ID and password and the second was mobile authentication. Following password authentication, a QR code was generated and sent to the user, which could be read by the built-in application in Android mobile phones. Then, the phone automatically sent the International Mobile Equipment Identity (IMEI) number to the server and an OTP was generated once the IMEI number was verified. The access policy file needed to be decrypted only if the data owner wished to change the data attributes, which saved time.

##### *2) Multi-factor Authentication*

A model of a three-layer authentication mechanism was proposed in [13] to ensure data security, encrypting the data at many levels using several security-related techniques. Unauthorized access was prevented in the first layer through authentication. This layer used various methods, including Fast Identity Online Alliance, multi-factor authentication, one-time password, and the short message service. The second layer was connected to the first to ensure that only legitimate customers could communicate. This layer used different cryptographic encryption techniques, such as homomorphic and proxy-based encryption. The third layer was in charge of the interactions of the second layer and ensured that users who requested access to the data were verified. The study claimed that this model can minimize data security issues in all three layers of cloud service models. This model provided login access to the end user to prevent malicious access to the stored data. To protect the

user's confidential information, it ensured the fast retrieval of data while using advanced data protection security and intelligence. However, multi-factor authentication can cause a computational overhead at the user's end.

In light of the drawbacks of multi-factor authentication, an architecture was proposed in [14] to improve it. This architecture used a combination of explicit and implicit authentication factors to provide users with secure access to cloud services at different levels. Implicit authentication is convenient for the user but cannot deliver the required security because it considers user behavior for authentication and has a high false-positive rate [14]. Explicit authentication is more secure, as the user directly provides the authentication factor, but is not user-friendly. An algorithm was proposed to reduce the perceived authentication difficulty for users at different access levels. The algorithm presented a difficult authentication factor if the user failed implicit authentication. In other words, the user needed to be validated through implicit authentication to be easily explicitly authenticated. The architecture involved sandboxing, implicit and explicit authentication, a metalearner, and component F. User access control was performed by sandboxing, whereby the user's access to different levels of the cloud service was controlled. Implicit and explicit authentication mechanisms were used to authenticate the user and obtain access to different service levels in the cloud. Each explicit authentication factor was assigned an authentication score. The metalearner (machine learning engine) provided the weight of the authentication based on the implicit authentication factor. The F component calculated the user authentication score and then identified a set of the best explicit authentication factors to be used by the user to obtain a higher access level. Each cloud service was assigned an operational sensitivity using sandboxing, which specified its security level. If the service was highly sensitive, only highly trusted users could gain access to it. If the authentication score of the user was high, he was highly trusted. The implicit authentication factor was used to assign weights to the scores of the explicit authentication factors. The weight of an explicit factor decreased (increased) if its corresponding implicit factor was invalidated (validated). The user needed a high authentication score to gain access to sensitive services, equal to or higher than the required operational sensitivity. This method reduced the difficulty of user authentication by 29% compared to other methods and can adapt the difficulty of authentication based on the user's situation. Progressive multi-factor authentication results in user fatigue if it is frequently applied because it requires user involvement, and multi-factor authentication can also excessively burden the user.

In [15], a multifactor authentication framework was proposed, which was essentially an extended version of the one proposed in [12]. This framework had three phases, including registration, multi-factor authentication, and encryption phases. The multi-factor authentication, unlike [12], was implemented through Captcha and OTP (One Time Password), in addition to the traditional password authentication process. The Captcha process verified if the user was a machine or a human, and OTP verified if the user was genuine or not. The problem with this authentication mechanism was the redundant use of Captcha, as the OTP process can prove both.

#### F. Identity-based Authentication Frameworks

In [16], a lightweight mutual authentication system comprising three phases was proposed. In the first phase, the cloud server and the customer generate a random number and an identity request and send it to a third party known as agency. The agency verifies the identity of the customer and the server and responds with a timestamp in the second phase. In the last phase, the user and the server forward identity and timestamp, along with the random number, to each other for mutual authentication. This verification process uses three operations: string concatenation, hashing, and the logical exclusive OR function. The disadvantages of this scheme are that it requires a third party as an agency and transferring computations and identities to the third party makes it vulnerable to attacks.

#### G. Location-based Authentication Frameworks

Message Digest and Location-based Authentication (MDLA) is a mutual authentication scheme proposed in [17] and involves symmetric key encryption. The goal is to authenticate a mobile user and the cloud server. The scheme consists of registration, authentication, and update phases. It starts only if the user is registered to the cloud service provider, where their credentials are stored on a server. The first phase involves the generation of a random number as a secret key and a message digest along with a primary key. The secret key is used to validate the cloud server in the authentication phase. The primary key is obtained using the user's hashed ID, password, and the previous or first location of the mobile device. If the user is registered, he can send an authentication request using an authentication key, which is composed of the user's current location and timestamp, to the cloud server to access resources. When the authentication request is received on the server side, the information on the mobile user is located to obtain the associated expiration period and the primary key. If the expiry period is zero, the authentication request is rejected by the server, which asks the user to re-register or updates the notification to the mobile user. In case of successful registration, the message is decrypted by the cloud server using the primary key to obtain the current location and timestamp that form the authentication key. The encrypted message is decrypted using the authentication key to obtain the digest of the message. The user is considered legitimate if it matches the stored message digest. Following the authentication of the mobile client, an authentication reply is sent to the mobile user containing the encrypted message digest through the previously shared secret key. Sending this response authenticates the cloud server. A random number is selected as the new secret key by the cloud server and is sent to the mobile user for the next authentication request. If the stored message digest matches the received message digest, the cloud server is considered legitimate. There are three types of update phases: key updates, client registration or re-registration if the expiry period is zero, and updates to authentication or re-authentication in case of a connection loss.

#### H. Service Model-based Authentication Frameworks

In [18], a personal CC framework based on a service model was proposed, providing users with a conveniently exposed Open API and a secure connection for accessing the cloud. The environment of "cloud orchestration and single sign-on token"

was considered so that the user could have a smooth experience. The study described various details of its various components based on the service model and then presented an application of the required security technologies.

Optimizing customer security in an Infrastructure as a Service (IaaS)-based cloud and the use of cloud resources were the objectives of [19], where a framework based on Virtual Machines (VMs) was proposed for the secure usage of sensitive data. Usage Management (UM) was applied to control the resources in the cloud and satisfy the performance and security-related requirements. UM involves constantly monitoring the policies related to the resources. The study presented Randomized Algorithms (RAs) and a policy model to specify the security policies. The RAs compute the optimal distribution of cloud resources of the IaaS between users to ensure security. The architecture consists of a central UM framework that manages cloud resources and Amazon S3 data. It ensures that the resources are moved to the VMs only if the security-related requirements are met, otherwise they are rejected. The UM framework also provides cloud VMs. A local Usage Management Module (UMM) was developed, where any VM provided by the central UM framework can be injected to enforce policies during the lifetime of the resource. If the context of the public cloud does not satisfy the policy requirements regarding the resources, the UMM retracts the resources. The architecture starts when the user requests a policy-protected resource. The user submits contextual information (such as credentials) to the UM framework and is authenticated and granted access based on access control policies for the relevant resources. Depending on the level of security and policies, the resources are transferred to the VM. The RA then receives information on the resource utilization of the VMs to calculate the cost function. Following this, the UMM receives the RA results and distributes the resources to multiple VMs based on security parameters and performance.

#### I. Cryptography-based Authentication Frameworks

In [20], security schemes that can guarantee the integrity, confidentiality, and authenticity of data were investigated, proposing a framework that uses symmetric and asymmetric cryptographic techniques to enhance data security, including the AES, RSA, and SHA algorithms. The proposed framework used a two-step authentication process: one for user authentication based on a login password for access to data on the server, and the other used RSA to generate a fingerprint verification mechanism for enhanced authentication. The process is carried out at the sender and receiver ends to make the system invulnerable to man-in-the-middle and hijacking attacks. AES was used to encrypt the key and the message using an asymmetric key algorithm, like the RSA. The simulation results showed that the proposed strategy was scalable, efficient, and cost-effective for simple data sharing/access.

In [21], a framework was proposed for secure data sharing in the cloud using hashed message authentication codes, index building, and data classification. An algorithm was developed to protect the data against several issues, e.g., malicious insiders, traffic hijacking, denial of service, and shared technology vulnerabilities, using public and private key

encryption (RSA, AES) along with a digital signature. An error localization algorithm was also proposed to specify the location of errors during an operation, such as deletion and appending while storing the data. Data are classified into private, public, and restricted access modes, by the data owner, to provide access control according to the attributes of confidentiality, integrity, and availability. Index-building is used to search through the encrypted data as the files are retrieved. A file index was proposed to be generated, and both the index and the files were encrypted. The index was composed of keywords that are encrypted as useless data. The encryption process involves the creation of a private/public key using RSA, and Symantec's encryption desktop tool is then applied to generate the user's authentication certificate and verify his identity. This certificate is a unique identifier that can be used to prevent malicious access by attaching it to the message. The Symantec tool generates the customer's private and public keys, and a request for a passphrase when they enter their username and email. The user is allowed to view the toolbar with the identifier, public key, and subkeys once they have provided the passphrase. A secure shell (SSH) is used to produce the keys, as it can be entered anywhere. A hash function is then used to generate the hash of the message, which is encrypted using the private key of the sender and joining the original message with the encrypted message. The sender and receiver share a symmetric key, which is a session key. Using this shared key, the sender re-encrypts the encrypted message along with the original message. The shared key is then encrypted by the receiver's public key and sent with the re-encrypted message. On the receiver side, the message is decrypted using the private key to obtain the session key, which is checked to determine if it matches the shared key. If they match, the receiver uses the sender's public key and the session key for decryption to obtain the original message. The HMAC is applied as a checksum to detect errors during data transmission. The checksum can determine if errors have occurred in the data but cannot correct them. The user can retrieve the data from the cloud by registering with the relevant organization to obtain his access credentials, and his username is stored in the cloud directory. The username is compared with the one saved in the cloud; if the user requests a segment of the data available for public access, no authentication is needed. However, if he requests access to private or restricted data, authentication is carried out by comparing the provided username with the one stored in the database. Authentication is carried out as the user sends his password to the owner and answers some security questions. Upon the user's request, the owner sends them using a digital signature and a keyword, and then sends the user's digital signature to the cloud along with his ID for subsequent use. When the user applies the received digital signature along with the keyword, the cloud confirms whether their ID is valid. The encrypted keywords help the cloud specify the user's location, which can be decrypted by the owner's sent key. The user then submits a file download request to the cloud, which sends the HMAC along with the encrypted file. This is in turn decrypted using the received key. Fully homomorphic encryption can be performed on the encrypted data if there are multiple sub-customers to whom the data need to be sent. The experimental study using Eclipse ID proved that the proposed algorithm was secure. The hybrid encryption-based scheme was faster than

the RSA but suffered from memory and time limitations such that it cannot be used in the cloud.

A security framework was proposed in [22] combining the processes used in Kerberos and Pretty Good Privacy (PGP). Kerberos provides mutual authentication, confidentiality, and integrity, and prevents eavesdropping and reply attacks. This system used the concept of a trusted third party (TTP). Kerberos has three logical components: the Authentication Server (AS), the Real server (RS), and the Ticket Granting Server (TGS). The authentication server acts as a Key Distribution Center (KDC) and issues a session key after verifying the user credentials. The session key is used for communication between the AS and the TGS. The TGS issues a ticket for the Real Server (RS) and provides a session key for communication between a user and the RS, which provides services to the user. The PGP provides cryptographic and authentication services. In this scheme, the user is registered with the Kerberos KDC, which provides a ticket to communicate with the cloud Service Provider (SP). KDC also sends the user credentials and the ticket to the SP, which stores them and acknowledges their receipt. The user encrypts the data before transmission to the cloud. PGP authenticates the user and sends user-encrypted data to the SP. The SP in turn sends the data, as requested by the user, to PGP, which decrypts them with the user information, and sends decrypted data to the user after authentication. The problem with this framework is that Kerberos does not support non-repudiation.

In [23], a hybrid encryption algorithm was proposed combining RSA and AES cryptography to secure data during cloud upload and download. The algorithm uses three separate keys for encryption and decryption. The first is a public key shared between the parties, the second is a private key only on the customer side, and the third is a secret key for decryption. Three keys are generated as the data are uploaded: an AES secret key, an RSA private key-d, an RSA public key-n, and an RSA public key-e. These keys are generated based on system time. The user is expected to save the AES secret key and the RSA private key, as they are needed to upload the data to the cloud (this key is unknown to the cloud administrator). Upload begins through user authentication (username and password). The data are uploaded in an encrypted form to the cloud and stored in a temporary directory. The RSA and AES algorithms are then called. The user is subsequently required to provide the AES secret key, so the data are permanently kept in the cloud and their temporary directory is deleted. The user has to identify the filename to download for the download process and use the AES secret key and the RSA private key to decrypt the data. This study claimed that this hybrid mode of encryption can prevent unauthorized access to user data, as a secret private key is required for this. This access control method can protect the data even if the private key has been compromised, as the original data are not given (unreadable data are downloaded). This also applies to the cloud administrator if he tries to view the data.

In [24], a mechanism based on the RSA cryptographic system was proposed, however, its key generation process is complex and unreliable for certain processes. In [25], two-layer attribute-based hybrid encryption was proposed. The first layer

used HMAC on sensitive attributes. A substitution mechanism was applied in the second layer to re-encrypt the data, and access control policies were also used. The overall mechanism is inefficient as several copies of the encrypted data are stored using the same key and the computational cost is high. Users are supposed to manage encryption keys, which is a significant drawback of the proposed system.

To simultaneously ensure the security, authenticity, and verification of the data, a three-way mechanism was proposed in [26] using Diffie–Hellman key exchange merged into AES and a digital signature. The Diffie–Hellman algorithm is responsible for key exchange through the generation of keys. Once the encryption algorithm has been executed, the digital signature is applied for user authentication. Two servers are involved in this system: one for storing user data files and the other for the encryption process (it is referred to as the trusted computing platform). The flow starts when Diffie–Hellman key exchange is used to exchange keys once the user attempts to upload data to the cloud server. The user is authenticated through the digital signature. The AES encryption algorithm is used to encrypt user data, and the file is uploaded to the storage server. The same steps are carried out if the user wants to download data.

The method proposed in [27] used the hash function, RSA, and DES to ensure data security in the mobile cloud. The system consists of a data owner, a cloud service provider, and a third-party auditor. On behalf of the data owner, the third-party auditor verifies the integrity of the data stored in the mobile cloud. The data in the cloud are encrypted twice: once by the private key of the data owner and again by the third-party auditor's private key. RSA is applied for message authentication. Each data owner and the third-party auditor uses RSA to generate public and private keys for each key generation phase. Only the third-party auditor's public key is then exchanged with the data owner over a secure channel (key-sharing phase). The owner encrypts the data using their public key and then generates a hash message of the encrypted file. The encrypted file is then re-encrypted using the third-party auditor's public key, as is the hash message. The two sets are joined and sent to the third-party auditor (encryption phase). The latter stores the encrypted hash function to ensure data integrity and generates a random key to encrypt the message (using DES). The received package is decrypted using the third-party auditor's private key. The produced random key is stored for later decryption. To verify the correctness of the data after performing DES, the encrypted package is uploaded to the cloud and submitted to the third-party auditor. A random key generated by the third-party auditor is used to decrypt the message. The third-party auditor then produces a hash of the encrypted file and compares it with the decrypted and stored hash value. The result is used to send the appropriate file to the owner along with the requested file. The latter is encrypted by the data owner using their public key so that only they can decrypt it.

In [28], a framework was proposed based on classifying the data into three segments, private, public, and limited access, and accordingly assigning sensitivity ratings to them. The owner classifies the data into segments. The framework

consists of two phases. The first phase is further divided into three components: index-building, classification, and message authentication. The index builder is used to sort the encrypted data, and the index itself is encrypted to provide security. Encryption is performed using the Secure Socket Layer (SSL) with a single key. A Message Authentication Code (MAC) is generated after encryption to perform an integrity check. Finally, the encrypted data are stored in the cloud based on a sensitivity rating according to the classification. In the second phase, the data retrieval process requires the registration of the user. The user credentials are stored in the cloud, and access to cloud data requires the submission of these credentials along with an access request. The cloud then takes action based on the request and the sensitivity rating of the data. Authentication is not required if the data are classified as public but is needed if the requested segments are private or have limited access. A digital signature mechanism is used for authentication and verification. Finally, an integrity check of the data is performed using MAC. This scheme emphasizes a trade-off of key size, where a large key increases system complexity and leads to throughput problems.

In the context of mobile cloud computing, an authentication scheme called Message Digest Authentication (MDA) was proposed [29]. This scheme does not involve any additional hardware infrastructure, like the IMSI chip, and consists of a registration and an authentication phase. In the former, the user needs to register by creating an account consisting of a user ID and password, as well as some unique identifying information. Following this, the cloud stores the user's ID and the hashed password along with information on the mobile device used. A message digest of the user, which consists of a user certificate and policy, is then generated by the cloud server. A message digest of the cloud, consisting of a cloud certificate and policy, is also generated. The cloud sends the registered device an encrypted message in the first phase that includes the user's message digest, the cloud's message digest, the cloud's public key, and the relevant column reference in the authentication database. A key is generated by XORing the user ID and the hashed password to encrypt the message. The authentication process starts once the user's message digest, the cloud's digest, and the encryption key have been received by the user's device and consists of two operations: the authentication of the cloud by the mobile device and the authentication of the mobile device by the cloud. The key generated by XORing the user ID and hashed password is used when the mobile device sends an authentication request to the cloud. This key is also used to produce an authentication key to encrypt the message digest. The encrypted message along with the column reference is sent to the cloud, which performs decryption after receiving the message. The cloud searches for the specific user ID and the hashed password. If they are found, it generates a key to decrypt and obtain the message digest. The message digest is decrypted using the authentication key. The user is considered legitimate if the stored message digest matches the obtained message digest. The cloud server submits its digital signature, including the encrypted message digest, using its private key when the mobile device has been authenticated. In return, the mobile device decrypts the message by using the public key of the cloud. If the stored message digest matches the obtained

message digest, the cloud server is considered legitimate. This step establishes data transmission. However, the claimed hypothesis fails if the vulnerability of certain parameters is high.

### III. CONCLUSIONS AND FUTURE WORK

This study reviewed several authentication frameworks and their models in cloud computing, reported the characteristics and features of various frameworks, outlined the outstanding threats to them, and proposed solutions to counter these threats. All proposed solutions share a common problem: they are excessively geared toward the client or server side of the cloud architecture. The requirements for developing an efficient cloud computing authentication framework are outlined as follows:

- Client-based solutions are either based on the client's or the cloud provider's features. A common framework should be designed and implemented for both parties. Such solutions should consider the requirements of the clients and the service vendors to avoid interoperability issues within the cloud environment.
- Greater focus is needed on data safety, confidentiality, and privacy, and less on the underlying architecture and processes running within the clients' systems.
- A low computational overhead should be a major desideratum in designing cloud-based services or devices as the scope of application has expanded to low-power systems.
- The vendor-specific solutions in some proposals imply hardware incompatibility in the cloud environment and thus need to be addressed.

### REFERENCES

- [1] M. Alsaif, N. Aljaafari, and A. R. Khan, "Information Security Management in Saudi Arabian Organizations," *Procedia Computer Science*, vol. 56, pp. 213–216, Jan. 2015, <https://doi.org/10.1016/j.procs.2015.07.201>.
- [2] B. Heydari and M. Aajami, "Providing a New Model for Discovering Cloud Services Based on Ontology," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2268–2272, Dec. 2017, <https://doi.org/10.48084/etasr.1577>.
- [3] M. Ramzan, M. S. Farooq, A. Zamir, W. Akhtar, M. Ilyas, and H. U. Khan, "An Analysis of Issues for Adoption of Cloud Computing in Telecom Industries," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3157–3161, Aug. 2018, <https://doi.org/10.48084/etasr.2101>.
- [4] M. F. Hyder, S. Tooba, and Waseemullah, "Performance Evaluation of RSA-based Secure Cloud Storage Protocol using OpenStack," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7321–7325, Aug. 2021, <https://doi.org/10.48084/etasr.4220>.
- [5] G. J. W. Kathrine, "A secure framework for enhancing user authentication in cloud environment using biometrics," in *2017 International Conference on Signal Processing and Communication (ICSPC)*, Coimbatore, India, Jul. 2017, pp. 283–287, <https://doi.org/10.1109/CSPC.2017.8305854>.
- [6] S. K. Khatri, Monica, and V. R. Vadi, "Biometric based authentication and access control techniques to secure mobile cloud computing," in *2017 2nd International Conference on Telecommunication and Networks (TEL-NET)*, Dec. 2017, pp. 1–7, <https://doi.org/10.1109/TEL-NET.2017.8343558>.
- [7] D. Chattaraj and M. Sarma, "Dependability Quantification of Cloud-Centric Authentication Frameworks," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*, San Francisco, CA, USA, Jul. 2018, pp. 840–844, <https://doi.org/10.1109/CLOUD.2018.00117>.
- [8] H. A. Dinesha and V. K. Agrawal, "Multi-level authentication technique for accessing cloud services," in *2012 International Conference on Computing, Communication and Applications*, Dindigul, India, Oct. 2012, pp. 1–4, <https://doi.org/10.1109/ICCCA.2012.6179130>.
- [9] V. K. Pant, J. Prakash, and A. Asthana, "Three step data security model for cloud computing based on RSA and steganography," in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, Greater Noida, India, Jul. 2015, pp. 490–494, <https://doi.org/10.1109/ICGCIoT.2015.7380514>.
- [10] M. Jensen, S. Schäge, and J. Schwenk, "Towards an Anonymous Access Control and Accountability Scheme for Cloud Computing," in *2010 IEEE 3rd International Conference on Cloud Computing*, Miami, FL, USA, Jul. 2010, pp. 540–541, <https://doi.org/10.1109/CLOUD.2010.61>.
- [11] L. Wang and T. Song, "An Improved Digital Signature Algorithm and Authentication Protocols in Cloud Platform," in *2016 IEEE International Conference on Smart Cloud (SmartCloud)*, New York, NY, USA, Aug. 2016, pp. 319–324, <https://doi.org/10.1109/SmartCloud.2016.46>.
- [12] K. S. Gajghate and R. V. Mante, "Secure Document Sharing and Access Control on Cloud for Corporate User," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, India, Apr. 2018, pp. 135–138, <https://doi.org/10.1109/ICICCT.2018.8473095>.
- [13] P. Sirohi and A. Agarwal, "Cloud computing data storage security framework relating to data integrity, privacy and trust," in *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*, Dehradun, India, Sep. 2015, pp. 115–118, <https://doi.org/10.1109/NGCT.2015.7375094>.
- [14] R. Fathi, M. A. Salehi, and E. L. Leiss, "User-Friendly and Secure Architecture (UFSA) for Authentication of Cloud Services," in *2015 IEEE 8th International Conference on Cloud Computing*, New York, NY, USA, Jun. 2015, pp. 516–523, <https://doi.org/10.1109/CLOUD.2015.75>.
- [15] L. K. Alnwhiel and A. R. Khan, "A Novel Cloud Authentication Framework," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, Sep. 2020, pp. 1302–1308, <https://doi.org/10.1109/CSCI51800.2020.00243>.
- [16] J. Shen, D. Liu, S. Chang, J. Shen, and D. He, "A Lightweight Mutual Authentication Scheme for User and Server in Cloud," in *2015 First International Conference on Computational Intelligence Theory, Systems and Applications (CCITSA)*, Ilan, Taiwan, Sep. 2015, pp. 183–186, <https://doi.org/10.1109/CCITSA.2015.47>.
- [17] S. Dey, S. Sampalli, and Q. Ye, "A light-weight authentication scheme based on message digest and location for mobile cloud computing," in *2014 IEEE 33rd International Performance Computing and Communications Conference (IPCCC)*, Austin, TX, USA, Sep. 2014, pp. 1–2, <https://doi.org/10.1109/PPCC.2014.7017041>.
- [18] S. H. Na, J. Y. Park, and E. N. Huh, "Personal Cloud Computing Security Framework," in *2010 IEEE Asia-Pacific Services Computing Conference*, Hangzhou, China, Sep. 2010, pp. 671–675, <https://doi.org/10.1109/APSCC.2010.117>.
- [19] V. Nandina, J. M. Luna, C. C. Lamb, G. L. Heileman, and C. T. Abdallah, "Provisioning Security and Performance Optimization for Dynamic Cloud Environments," in *2014 IEEE 7th International Conference on Cloud Computing*, Anchorage, AK, USA, Jun. 2014, pp. 979–981, <https://doi.org/10.1109/CLOUD.2014.150>.
- [20] S. K. M., "Enhanced Security Framework to Ensure Data Security in Cloud Using Security Blanket Algorithm," *International Journal of Research in Engineering and Technology*, vol. 02, no. 10, pp. 225–229, Oct. 2013, <https://doi.org/10.15623/ijret.2013.0210033>.
- [21] A. Bhandari, A. Gupta, and D. Das, "A framework for data security and storage in Cloud Computing," in *2016 International Conference on Computational Techniques in Information and Communication*

- Technologies (ICCTICT)*, New Delhi, India, Mar. 2016, pp. 1–7, <https://doi.org/10.1109/ICCTICT.2016.7514542>.
- [22] S. C. Patel, R. S. Singh, and S. Jaiswal, "Secure and privacy enhanced authentication framework for cloud computing," in *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*, Coimbatore, India, Oct. 2015, pp. 1631–1634, <https://doi.org/10.1109/ECS.2015.7124863>.
- [23] V. S. Mahalle and A. K. Shahade, "Enhancing the data security in Cloud by implementing hybrid (Rsa & Aes) encryption algorithm," in *2014 International Conference on Power, Automation and Communication (INPAC)*, Amravati, India, Jul. 2014, pp. 146–149, <https://doi.org/10.1109/INPAC.2014.6981152>.
- [24] P. Yellamma, C. Narasimham, and V. Sreenivas, "Data security in cloud using RSA," in *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, Tiruchengode, India, Jul. 2013, pp. 1–6, <https://doi.org/10.1109/ICCCNT.2013.6726471>.
- [25] S. Verma and S. Ahuja, "A hybrid two layer attribute based encryption for privacy preserving in public cloud," in *2016 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, Dec. 2016, vol. 2, pp. 1–5, <https://doi.org/10.1109/INVENTIVE.2016.7824822>.
- [26] P. Rewagad and Y. Pawar, "Use of Digital Signature with Diffie Hellman Key Exchange and AES Encryption Algorithm to Enhance Data Security in Cloud Computing," in *2013 International Conference on Communication Systems and Network Technologies*, Gwalior, India, Apr. 2013, pp. 437–439, <https://doi.org/10.1109/CSNT.2013.97>.
- [27] P. Garg and V. Sharma, "An efficient and secure data storage in Mobile Cloud Computing through RSA and Hash function," in *2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, Ghaziabad, India, Oct. 2014, pp. 334–339, <https://doi.org/10.1109/ICICT.2014.6781303>.
- [28] S. K. Sood, "A combined approach to ensure data security in cloud computing," *Journal of Network and Computer Applications*, vol. 35, no. 6, pp. 1831–1838, Nov. 2012, <https://doi.org/10.1016/j.jnca.2012.07.007>.
- [29] S. Dey, S. Sampalli, and Q. Ye, "MDA: message digest-based authentication for mobile cloud computing," *Journal of Cloud Computing*, vol. 5, no. 1, Nov. 2016, Art. no. 18, <https://doi.org/10.1186/s13677-016-0068-6>.



# Geotechnical Characterization of Phosphate Mining Waste Materials for Use in Pavement Construction

**RachidaMalaoui**

Environmental Laboratory, Civil Engineering Dpt, Echahid Cheikh Larbi Tebessi University, Algeria  
rachida.malaoui@univ-tebessa.dz  
(corresponding author)

**El Haddi Harkati**

Mining Laboratory, Civil Engineering Dpt, Echahid Cheikh Larbi Tebessi University, Algeria  
harkati.elhaddi@gmail.com

**Mohamed Redha Soltani**

Mining Laboratory, Civil Engineering Dpt, Echahid Cheikh Larbi Tebessi University, Algeria  
msoltani.mohamedredha@univ-tebessa.dz

**Adel Djellali**

Environmental Laboratory, Civil Engineering Dpt, Echahid Cheikh Larbi Tebessi University, Algeria  
adel.djellali@univ-tebessa.dz

**Abderraouf Soukeur**

Laboratory of Hydrometallurgy and Inorganic Molecular Chemistry, Faculty of Chemistry, University of Science and Technology Houari Boumediene, Algeria  
abderraoufsoukeur@gmail.com

**Rabah Kechiched**

Underground Reservoirs Laboratory: Oil, Gas, and Aquifers, Kasdi Merbah Ouargla University, Algeria  
rabeh21@yahoo.fr

Received: 16 November 2022 | Revised: 7 December 2022 | Accepted: 9 December 2022

## ABSTRACT

Waste rock materials are becoming widely used in road pavement and building constructions in many countries. In this work, experimental laboratory tests were carried out on the waste rock produced from the extraction of the phosphate in the Kef-Essenoun mine, to study the performance of road pavement foundations built with these types of material. Two types of waste, namely phosphatic limestone (type 1) and limestone (type 2), were initially tested to determine the most suitable one to be used in pavement structures. The characterization tests showed that the presence of carbonate-fluorapatite and carbonate-fluorapatite, and calcite, dolomite, and quartz are predominant in phosphatic limestone and limestone, respectively. The Los Angeles Abrasion (LA) and Micro-Deval (MD) values range from 59.9% to 90.4% and 42.05% to 86.31% for phosphatic limestone and from 43.64% to 95.88% and 38.25% to 75% for limestone. The CBR values of type 1 and type 2 waste were found to be 10.5% and 18.7% respectively. The results show that these materials, classified as B<sub>42s</sub> and B<sub>42s</sub> respectively, could be used cautiously in capping layers and pavement backfilling materials. Furthermore, they must be treated with a hydraulic binder such as cement in order to improve their physical and mechanical properties.

*Keywords-characterization; mine waste; phosphatic limestone; limestone; road construction; Kef-Essenoun*

## I. INTRODUCTION

Road infrastructure projects require large quantities of materials. The growing demand for aggregates in the civil construction sector has become more pronounced in recent years as cities grow more and more. The use of unusual industrial mining products in road construction can contribute to the conservation of non-renewable natural resources and minimize the waste quantities produced by mining industries [1]. Such large amounts of waste, resulting from mining operations, have led to the awareness regarding their impact on environment, ecology, and geotechnical side. The release of toxic components from heavy metals and acid mine drainage are examples of such effects. Underground water contamination, large amounts of space of exploitable natural land, instability of the waste storage zone, and dust cloud generation are some examples of such effects. There are numerous economic and ecological gains in the exploitation of alternative materials for the road construction sector. When waste materials with adequate properties are used, it is possible to reduce energy consumption, transport distances, and extraction costs [2]. For several decades the reuse of waste materials as construction materials has been studied in depth, e.g. mine tailings [1-3], recycled construction and demolition wastes used in road construction, embankments and concrete [4-7], coal mining waste [8-9], plastic and rubber waste in flexible and rigid pavements, ferrocement mortar [10-12], steel slag and glass as secondary aggregates in asphalt mixtures [13-15], tungsten mining waste [16], marine sediments [17], bottom ash and fly ash in geotechnical engineering [18], iron and gold mining waste [19], and red mud and lime [20]. In addition, pozzolanic materials such as cement, lime, fly ash, cement kiln dust, and ordinary Portland cement have been used to stabilize some tailings materials for construction purposes [21-23].

Algeria has a large reserve of phosphate [24]. About 41ha of land are currently under exploitation in Kef-Essenoun. The current removal mining rate is 2.8 million tons of ore with a stripping ratio of approximately 2.2 which produces annually 10 million tons of waste that must be removed and disposed of. The accumulation for several decades of waste rock from open-pit phosphate mines and upgrading processes poses serious environmental, ecological, and health problems in the surrounding areas [25]. These types of waste are unsafe to humans, animals, and vegetation. They pollute the air, the surrounding soil, and contaminate the groundwater.

The main objective of this research consists in suggesting a reusing solution in order to reduce the negative impact of rock waste on the environment. Reusing consists of finding a new use for available waste materials with adequate acceptable properties in their current form in order to avoid disposal on landfills. As reported in [26-31], those materials (phosphate tailings) could be widely used in building and road construction. The proposed reusing approach should be integrated in the Algerian road guide. It is worth noting that studies on phosphate mining waste in Algeria are rare, thus, chemical, mineralogical, physical, and geotechnical characterizations of mining waste from Kef-Essenoun phosphate mine were investigated in this study.

## II. GENERAL SETTING

The Kef-Essenoun phosphate mine (34.726784 E, 7.895978 N) is located on the southern flank of Djebel Onk cretaceous anticline in the Eastern Saharan Atlas [32]. It is situated 7km southeast of Bir El Ater City (Tebessa, northeast Algeria) and about 21km to the Algerian-Tunisian border [33]. The study site extends approximately to 250ha and belongs to the Djebel Onk mining basin which represents the occidental part of Gafsa-Metaloui-Onk basin containing many phosphorite layers deposited from late Cretaceous to early Eocene [32-34]. Although this lateral continuity, the Kef-Essenoun deposit displays substantial differences to the Gafsa Metlaoui basin, especially with regard to the lithology and the succession particularly in the phosphorite formation [35-36]. In the Kef-Essenoun deposit, the sedimentary lithologies consist of ~ 500m thick succession of upper Cretaceous (Maastrichtian) to middle Eocene (Lutetian) age where upper thanetian phosphorite layer is mined for phosphate raw material [32-37].

Mine waste rock from Kef-Essenoun mine (Figure 1) consists of soil and rock excavated during the mining operations after the commercially recoverable part has been recovered. These waste rocks are unloaded and accumulated in huge quantities. Between 2006 and 2017, the estimated quantity of waste rock produced is approximately 5.5 to 18.5Mt/yr. This quantity will increase to more than 18Mt/yr in the future (Figure 2).



Fig. 1. Phosphate mine waste hips nearby the mining site.

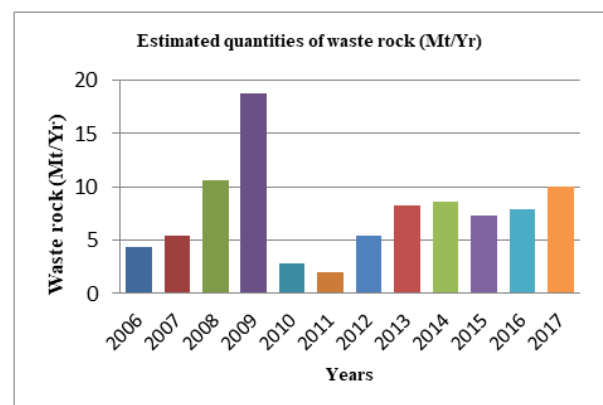


Fig. 2. Estimated quantities of waste rock from Kef-Essenoun [38].



Fig. 3. Tailings of Kef Essenoun mine: (a) phosphatic limestone, (b) limestone.

Those amounts of waste can be utilized in the field of pavement construction. The reuse of mine tailings for subgrade and foundation layers is a relatively new approach in Algeria. The lack of fundamental knowledge about their behavior limits their reuse. Therefore, and in order to examine the various aspects concerning the possibility of using this mining waste in road engineering, complete characterization is necessary. The objective of this experimental study is to make a geotechnical characterization of the two recycled mining materials of Kef-Essenoun. In this research, the studied materials are phosphatic limestone (type 1) and limestone (type 2) (Figure 3).

### III. EXPERIMENTAL PROGRAM AND TEST PROCEDURES

#### A. Initial Testing and Study Material Selection

The mining waste samples used in this research were collected from different locations of stacked tailings located in the surroundings of the Kef-Essenoun mine. The waste consists mainly of phosphatic limestone and limestone. These materials were brined to the laboratory in order to describe their mineralogical, chemical, physical, and mechanical properties.

#### B. Chemical and Mineralogical Characteristics

The analysis of the major elements was conducted by the Algerian phosphate company SOMIPHOS using colorimetric, gravimetric, volumetric, ionometric, and spectrophotometric by Atomic Absorption Spectrometry (AAS) analytical methods. The phosphate ore tailings of crystalline phases were performed by X-Ray Diffraction, (XRD). Diffraction patterns were made with a PANalytical X'Pert Pro diffractometer equipped with a conventional X-ray (Cu K $\alpha$  radiation, running at 40kV and 30mA) with detection in the 2 $\theta$  range of 10-90°.

#### C. Physical and Mechanical Parameters of Waste Materials

The main physical and mechanical characteristics of the studied waste samples are shown in Table I. Using oven drying at 105°C (+/- 5°C), the measured initial water content was about 4.62% and 8.54% for phosphatic limestone and limestone, respectively (NF P94-050). The following parameters were measured:

- Specific gravity was measured with a NF P94-054 Pycnometer.
- Clay fraction activity was measured by methylene blue test (NF P 94-068).

- Liquid and plastic limits (NF P94-051).
- Particle size distribution was measured according to NF P 94-056.
- Modified proctor (NF-P 94-093).
- California bearing capacity (NF-P94-078).
- Los Angeles, micro-deval, and friability (P 18-573, P 18-572, P18-576).
- Unconfined compressive strength procedure NF P98-230-3.

The test particle size consists of a representative sample of soil passing through superposed sieves with openings from 50mm to 0.08mm. Figure 4 shows the particle size distribution curves of the samples. The samples were classified according to the modified LPC classification system. According to the particle size curve, the percentages of cobbles, gravel, sand, silt, and clay in phosphatic limestone and limestone were 51%, 29%, 17.28%, and 2.72%, and 27.5%, 29%, 40.23%, and 3.27%, respectively. The coefficients of uniformity (Cu) and curvature (Cc) obtained for phosphatic limestone and limestone were 69.44, 17.85, and 0.36, 4.11, respectively. Based on the unified soil classification system, both sample types were classified as well-graded gravel (Gm). The particle size distribution and gradation characteristics of the tested samples are shown in Table I.

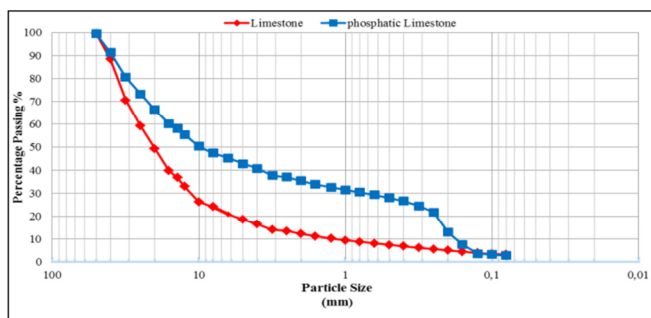


Fig. 4. Particle size distribution curves of phosphatic limestone and limestone waste.

#### D. Soil Classification

Based on the technical guidelines of embankment and capping layer construction (GTR), the tailing materials belong to the B Class of sand and gravel soils with fines. Depending on the nature of these materials, their class is B4. They have a passing percentage less than 12% and less than 70% from the of 80 $\mu$ m and 2mm sieves, respectively. There is no grain-size over 50mm. The methyl blue absorption value is more than 0.2. MD coefficient and LA coefficient are greater than 45. These types of materials are sensitive to water due to the existence of plastic fines and contain a large proportion of coarse particles, so they are generally pervious. The waste materials are classified as: B42ts (very dry) for type 1 and B42s (dry) for type 2 as can be shown in Figure 5.

TABLE I. PHYSICO-MECHANICAL PARAMETERS OF WASTE MATERIALS

Properties of mine railing	Unit	Type1	Type2
Moisture content	%	4.62	8.54
PH		6.8	7.8
<b>Geotechnical properties—Natural parameters</b>			
<b>Grain - size distribution</b>			
Color		Grayish	Beige
Shape		Angular	Angular
<b>Particle size distribution</b>			
Silt and clay		2.78	3.27
Sand		17.28	40.23
Gravel		29.00	29.00
Cobbles		51.00	27.50
Coefficient of uniformity (CU = D <sub>60</sub> /D <sub>10</sub> )	%	69.44	17.85
Coefficient of gradation		0.36	4.11
C <sub>c</sub> = (D <sub>30</sub> ) <sup>2</sup> /(D <sub>10</sub> × D <sub>60</sub> )		Gm	Gm
Group symbol LPC			
Optimum Moisture content (W <sub>opt</sub> )		12.30	14.15
Maximum dry density d <sub>max</sub>	kN/m <sup>3</sup>	1.85	2.00
Unsoaked CBR		24.724	36.675
Soaked CBR		10.548	18.293
Free swelling		0.303	0.308
Liquid limit (LL)		30.88	33.41
Plastic limit (PL)		Not measured	23.805
Plasticity index (PI)		-	9.6
Methylene blue value (MBV)	g/100g	0.83	1
Carbonate content	%	54	88
Organic matter		< 1	< 1
Sand Equivalent (SE)		22.22	12.82
Specific gravity	kN/m <sup>3</sup>	2.737	2.631
Apparent specific gravity		1.434	1.328
<b>Los Angeles abrasion test</b>			
4/6.3 6.3/10 10/14	%	82.9 75.5 72.2	95.88 59.36 43.64
10/25 16/31 25/50	%	68.8 59.9 90.4	43.74 88.3 84.4
<b>Micro Deval and deval test</b>			
4/6.3 6.3/10 10/14		82.6 86.31 81.96	64 73 75
25/50	%	42.05	38.25
Sand friability coefficient		32	31
Material classification		B <sub>42</sub> ts	B <sub>42</sub> s
<b>Unconfined Compressive Strength (UCS) after:</b>			
24 hours		182	520
14 days	kN/m <sup>2</sup>	683	2671
28 days		715	2807

Note: D<sub>10</sub>, D<sub>30</sub> and D<sub>60</sub> represent the percentage of soil particles that are smaller than 10, 30, and 60%, respectively

TABLE II. CHEMICAL COMPOSITION OF THE PHOSPHATIC LIMESTONE (TYPE 1) AND LIMESTONE (TYPE 2) WASTE

Major elements (wt %)														Trace elements (ppm)			
P <sub>2</sub> O <sub>5</sub>		SO <sub>3</sub>	CO <sub>2</sub>	CaO	Al <sub>2</sub> O <sub>3</sub>	MgO	Fe <sub>2</sub> O <sub>3</sub>	SiO <sub>2</sub>	Na <sub>2</sub> O	K <sub>2</sub> O	LOI	F	H <sub>2</sub> O	COrg	Pb	Zn	Cu
Type1	0.40	0.72	0.02	37.84	0.70	4.66	0.35	10.40	0.42	0.04	18.80	2.05	1.54	0.22	30	60	2.50
Type2	0.89	0.03	39.06	49.77	0.25	5.24	0.26	2.30	0.09	0.01	40.70	0.40	0.72	0.02	30	32.5	3.00

#### IV. RESULTS AND DISCUSSION

##### A. Chemical and Mineralogical

##### 1) X-Ray Diffraction (XRD)

The XRD pattern of the two samples is illustrated in Figure 6. It shows the presence of the apatite mineral class with variable substitution rates of CO<sub>3</sub><sup>2-</sup>, F<sup>-</sup> and OH<sup>-</sup>, including carbonate-fluorapatite [Ca<sub>10</sub>(PO<sub>4</sub>)<sub>5</sub>CO<sub>3</sub>F<sub>1.5</sub>(OH)<sub>0.5</sub>], fluorapatite [Ca<sub>5</sub>(PO<sub>4</sub>)<sub>3</sub>F], carbonate-apatite [Ca<sub>10</sub>(PO<sub>4</sub>)<sub>6</sub>], carbonate-hydroxyapatite [Ca<sub>10</sub>(PO<sub>4</sub>)<sub>3</sub>(CO<sub>3</sub>)<sub>3</sub>(OH)<sub>2</sub>], hydroxyapatite [Ca<sub>5</sub>(PO<sub>4</sub>)<sub>3</sub>(OH)], and hydrated phosphate, but the gangue

elements are basically represented by calcite [CaCO<sub>3</sub>], quartz [SiO<sub>2</sub>], dolomite [CaMg(CO<sub>3</sub>)<sub>2</sub>], and gypsum [CaSO<sub>4</sub>] [39-40]. Carbonate-fluorapatite is the dominant phase indexed in a hexagonal symmetry (SG: P6<sub>3</sub>/m) according to the ASTM XRD data reference 98-003-4653 [41]. However, limestone is richer than phosphatic limestone in calcite, quartz, and dolomite.

##### 2) Chemical Analysis

The chemical analysis results of the two phosphate waste types is presented in Table II. The results show a high percentage of major elements CaO, SiO<sub>2</sub>, and MgO for Type 1

and CaO, MgO, and SiO<sub>2</sub> for Type 2. However, Type 1 has higher concentrations of P<sub>2</sub>O<sub>5</sub>. The chemical composition confirms that the mineral composition of phosphatic limestone and limestone contains four main minerals: carbonate-fluorapatite, calcite, dolomite, and quartz. The results are in agreement with the findings of [25, 34-43].

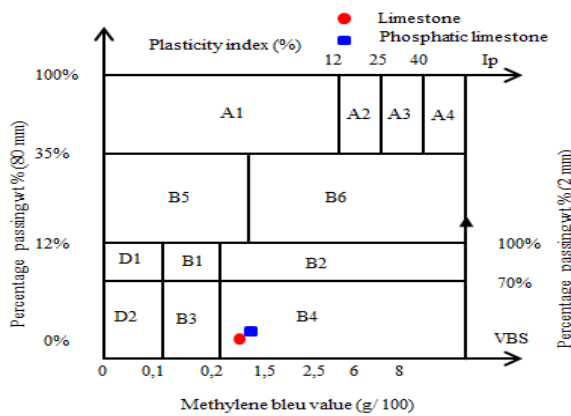


Fig. 5. Classification of studied waste samples: A: fine soil, B: sandy and gravelly soil with fine particles, and D: soil insensitive to water.

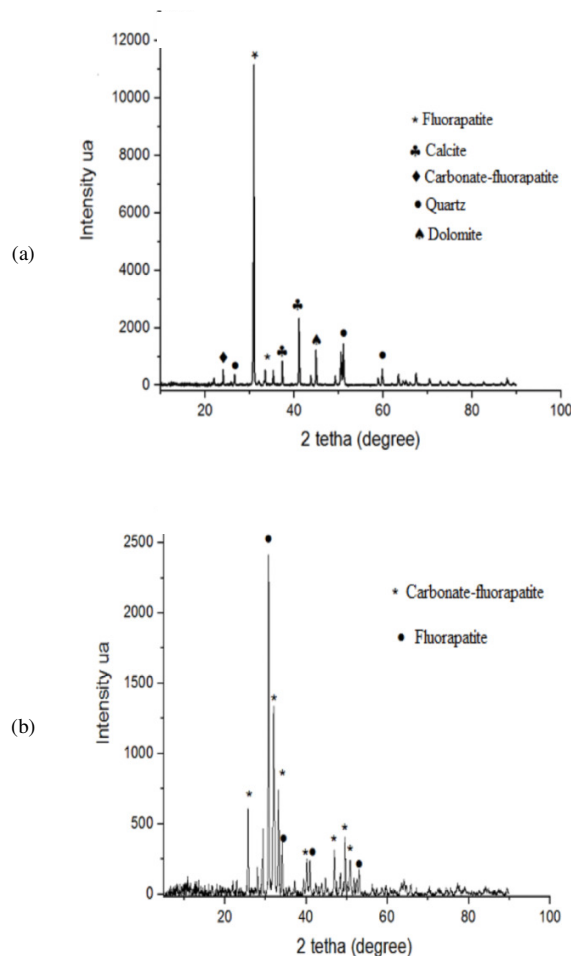


Fig. 6. XRD analysis of Kef-Essenoun mine waste: (a) phosphatic limestone, (b) limestone.

## B. Basic Characteristics

The waste materials used in this study were grayish and beige in color for type 1 and type 2 respectively, with dry moisture content due to the arid climate of the region. The pH values of the tailings are 6.8 and 7.8, indicating basic pH. The specific gravity of limestone is lower than that of phosphatic limestone due to the presence of P<sub>2</sub>O<sub>5</sub> in the latter. The phosphatic limestone sample is non-plastic, and thus, the measurement of the plastic limit and index was avoided. However, the limestone sample has a Plasticity Index (PI) of 9.6% which was validated by the methylene blue test.

## C. Mechanical Behavior of the Materials

### 1) Compaction Characteristics

The results of compaction tests of the waste materials, in terms of dry density and measured water contents, are given in Figure 7.

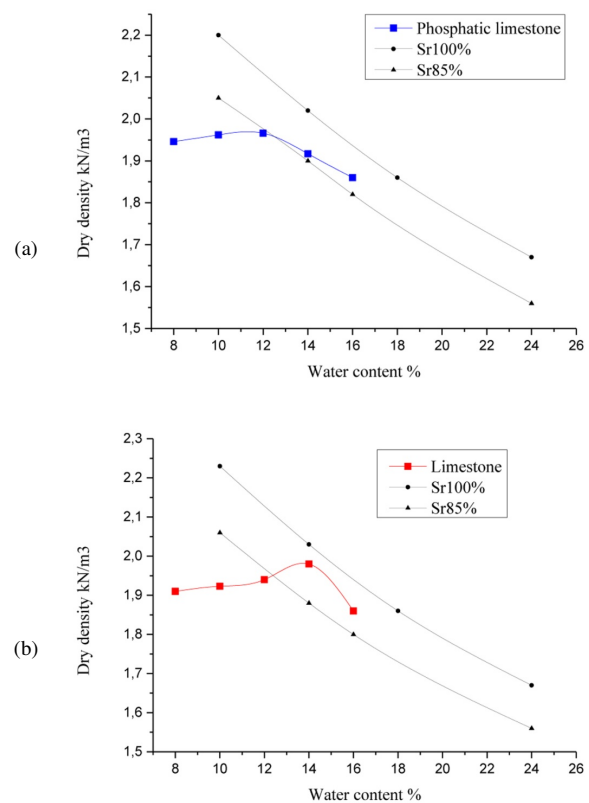


Fig. 7. Corresponding modified proctor compaction curves for mine wastes of (a) phosphatic limestone, (b) limestone.

The 85% and 100% saturation curves are also shown. The results indicate that the limestone waste provides the highest MDD of 2.00g/cm<sup>3</sup> and OMC of 14.15%. Generally, the compaction characteristics depend on both the grain size distribution and the specific gravity of the materials. In this case, the presence of cenospheres (hollow particles of large size) is the main cause for the lower density of the phosphatic limestone samples.



## 2) California Bearing Ratio (CBR)

Figure 8 shows the values of both unsoaked CBR and soaked CBR after 4 days of curing in water related to the MDD. It can be observed that the CBR values of the two types of samples increase as the MDD increases. It is also noticed that the phosphatic limestone waste gives a value of soaked CBR of 10.5%, indicating medium bearing capacity material, while the obtained value of soaked CBR of 18.7% of limestone indicates that it can be used as an appropriate material for the foundation layers of pavements. The FSI of samples obtained for type 1 and type 2 samples are 0.303% and 0.308%, respectively. These results show that these waste materials are not expansive (IS 2720 (Part XL)).

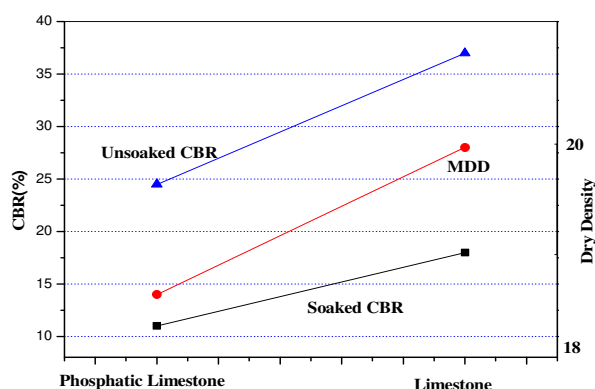


Fig. 8. CBR values related to MDD.

## 3) Los Angeles and Micro-Deval

The variation of LA and MD values is shown in Table I. The LA values range from 59.9% to 90.4% for phosphatic limestone and from 43.64% to 95.88% for limestone for all the considered granular classes. The results indicate that limestone waste gives more suitable values, but these waste materials are unsuitable for pavement constructions. The values of LA and MD tests of materials have to be less than 45%. Sand Friability (FS) coefficient is an indicator to assess the resistance of untreated sandy materials used for pavement subgrade to traffic loading. The FS values of the samples are 32% and 31% for type 1 and 2 materials, respectively. The recommended limit for this use is  $FS \leq 60$ .

## 4) Unconfined Compressive Strength (UCS)

The UCS of the mining waste samples was tested after 1, 14, and 28 days of curing as shown in Figure 9. It is well known that the materials gain in strength overtime and the magnitude of compressive strength increased with increasing MDD of the compacted granular material. It was also shown that the compressive strength of limestone at 28 days is 3.9 times larger than that of the phosphatic limestone. On the other hand, the mechanical characteristics do not deteriorate with time. The limestone closes the interior ports by cementing the grains (lime reacts strongly with water and compaction energy), favored by the plasticity of the compacted material. In this case, limestone and rock behave similarly and the compressive resistance increases over time. Furthermore, the phosphatic limestone does not cement during compaction due

to the absence of plasticity and the presence of a large content of friable phosphate, which explains its lower mechanical resistance.

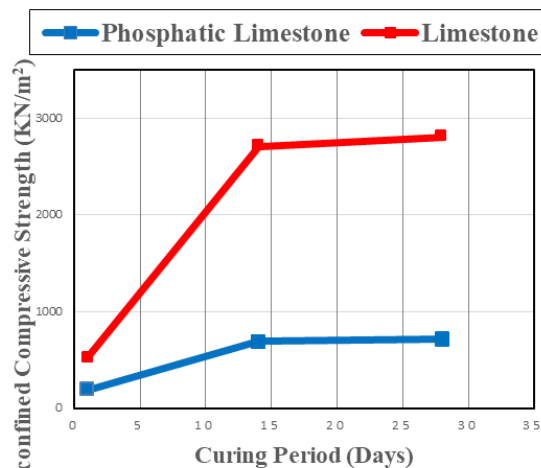


Fig. 9. UCS of samples cured for 1, 14, and 28 days.

## V. POSSIBILITY OF USING WASTE MATERIALS IN ROAD CONSTRUCTION

For the road material field the general rating of the mining wastes as sub-grade materials ranges from fair to poor. The phosphatic limestone waste is considered a water-sensitive material, found in a very dry hydrous state (ts) and is not normally reusable in the embankment construction or capping layer, but in certain cases, its humidification can be considered to bring it to the s (dry) or m (normal) state. The use of limestone waste material could be allowed in road construction with treatment by hydraulic binders. This treatment requires prior measurement of their mechanical strength. For compaction procedure, the very dry soil is considered as being impossible to compact properly by standard methods. In both dry (s) and very dry (ts) states, those soils are not easy to compact to form stable fill structures.

The catalog of the National Agency for Technical Control of Public Works Algerian for new pavement design [44], requires the use of well-graded (GW) materials with PI less than 10% and LA less than 40% in the base layer. The findings indicate that waste rock materials, which are categorized as B42ts and B42s, respectively, should be used with caution as backfill and capping materials for road paving. They need to be treated with a hydraulic binder, such as cement, to increase their mechanical and physical qualities, raising their classes as a result. Increasing the mechanical properties of the tailing materials to be used as pavement sub-base materials will be cost effective and ecologically friendly since it will result in thinner sub-base layers and will maybe reduce the thickness of the pavement layer. In this case, stabilization is the best choice. As the mine is expanded, the waste piles created by its exploitation will ultimately result in great environmental and ecological issues. From the economic and environmental perspectives, the reuses of such waste are beneficial.



## VI. CONCLUSIONS

The management of tailings and waste from mines and quarries is an essential requirement to limit their environmental, ecological, and geotechnical impacts. In the current research work, a characterization of two tailings from the Kef-Essenoun mine (phosphatic limestone and limestone) was investigated in order to allow reusing these materials in pavement foundations. The main findings of this investigation can be summarized as follows:

- The waste materials contain mainly CaO, SiO<sub>2</sub>, and MgO. The phosphatic limestone has a much higher concentration of P<sub>2</sub>O<sub>5</sub> (22.66 wt%) than limestone (0.89 wt%). In terms of trace elements, very low concentrations of Pb, Zn, and Cu were detected. The mineralogical composition of the two phosphate wastes contains four main minerals: carbonate-fluorapatite, calcite, dolomite, and quartz.
- The materials are well-graded gravel containing all grain sizes from cobbles down to clay. The particles were of angular shape with very low plasticity index. They have no grains larger than 50mm, a passing proportion from of 80µm sieve less than 12%, and less than 70% for the 2mm sieve.
- The Los Angeles and Micro-Deval coefficients of the studied materials exceeded 45%, and they are classified as B42ts (phosphatic limestone) and B42s (limestone).
- From the points of mechanical behavior and applicability in road construction, the values of CBR and UCS, increase with MDD and the UCS increases with curing age.
- The mechanical performance of limestone waste is more appropriate than phosphatic limestone's, and is governed essentially by the phosphate content (P<sub>2</sub>O<sub>5</sub> wt%). The mechanical proprieties are improved with higher phosphate content. The phosphatic limestone presents low plasticity and a larger proportion of friable phosphate which lead to less material compaction.
- In terms of environmental impact, the efficient reuse of mining waste in the road sector field helps reduce the production of waste, thereby minimizing the environmental damage and ensuring sustainable construction. Both waste materials considered in this study could be used in pavement structure construction provided that they are treated with hydraulic binders.

Based on this study, the presented methodology could constitute a starting point for the investigation of the reuse of the mining waste of Kef-Essenoun mine in the field of road construction.

## VII. DECLARATIONS

### A. Author Contributions

Rachida Malaoui conducted all the physical and geotechnical characterization tests, investigation, formal analysis, and the writing of the original draft. Abderraouf Soukeur interpreted the chemical and mineralogical characterizations. The geological part was conducted by Rabah Kechiched. The interpretation of the results and the writing of

the paper were done by Rachida Malaoui, Adel Djellali, and Abderraouf Soukeur. El Haddi Harkati and Mohamed Redha Soltani were responsible for conceptualization, supervision, writing, reviewing, editing, and funding acquisition.

### B. Funding

This work is funded by the Environmental Laboratory under the Grant No. A01L05UN120120210001.

## ACKNOWLEDGMENT

The authors are grateful to the National Company of phosphate SOMIPHOS for providing samples and filed assistance. We would also like to thank the staff of the Civil Engineering laboratory of Larbi Tebessi University. We thank warmly Mr. Malik Atout for his help in the X-ray analysis and Mrs. Med El-Nadir Ktir and Med Saleck Ahmed for their aid in the accomplishment of the laboratory tests.

## LIST OF ABBREVIATIONS

CBR	California Bearing Ratio
Corg	Organic carbon content
FSI	Free Swelling Index
LA	Los Angeles abrasion value
LOI	Loss on Ignition
MBV	Methylene Blue Value
MDD	Maximum Dry Density
MD	Micro Deval value
Mt/yr	Million tons per year
OMC	Optimum Moisture Content
PI	Plasticity Index
UCS	Unconfined Compressive Strength
XRD	X-Ray Diffraction

## REFERENCES

- [1] M. Amrani, Y. Taha, A. Kchikach, M. Benzaazoua, and R. Hakkou, "Valorization of Phosphate Mine Waste Rocks as Materials for Road Construction," *Minerals*, vol. 9, no. 4, Apr. 2019, Art. no. 237, <https://doi.org/10.3390/min9040237>.
- [2] S. Lidelow, J. Macsik, I. Carabante, and J. Kumpiene, "Leaching behaviour of copper slag, construction and demolition waste and crushed rock used in a full-scale road construction," *Journal of Environmental Management*, vol. 204, pp. 695–703, Dec. 2017, <https://doi.org/10.1016/j.jenvman.2017.09.032>.
- [3] A. Djellali, M. S. Laouar, B. Saghafi, and A. Houam, "Evaluation of Cement-Stabilized Mine Tailings as Pavement Foundation Materials," *Geotechnical and Geological Engineering*, vol. 37, no. 4, pp. 2811–2822, Aug. 2019, <https://doi.org/10.1007/s10706-018-00796-8>.
- [4] T. H. Nguyen, T. T. T. Nguyen, T. T. H. Nguyen, and V. T. Phan, "The Feasibility of Applying Waste Concrete as Coarse Aggregates in New Concrete," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9192–9195, Oct. 2022, <https://doi.org/10.48084/etasr.5206>.
- [5] A. Arulrajah, J. Piratheepan, M. Ali, and M. Bo, "Geotechnical Properties of Recycled Concrete Aggregate in Pavement Sub-Base Applications," *Geotechnical Testing Journal*, vol. 35, no. 5, pp. 743–751, Aug. 2012, <https://doi.org/10.1520/GTJ103402>.
- [6] A. Arulrajah, J. Piratheepan, M. M. Disfani, and M. W. Bo, "Geotechnical and Geoenvironmental Properties of Recycled Construction and Demolition Materials in Pavement Subbase Applications," *Journal of Materials in Civil Engineering*, vol. 25, no. 8,

- pp. 1077–1088, Aug. 2013, [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0000652](https://doi.org/10.1061/(ASCE)MT.1943-5533.0000652).
- [7] J. Zhang, F. Gu, and Y. Zhang, "Use of building-related construction and demolition wastes in highway embankment: Laboratory and field evaluations," *Journal of Cleaner Production*, vol. 230, pp. 1051–1060, Sep. 2019, <https://doi.org/10.1016/j.jclepro.2019.05.182>.
  - [8] Z. Bian, J. Dong, S. Lei, H. Leng, S. Mu, and H. Wang, "The impact of disposal and treatment of coal mining wastes on environment and farmland," *Environmental Geology*, vol. 58, no. 3, pp. 625–634, Aug. 2009, <https://doi.org/10.1007/s00254-008-1537-0>.
  - [9] L. Haibin and L. Zhenling, "Recycling utilization patterns of coal mining waste in China," *Resources, Conservation and Recycling*, vol. 54, no. 12, pp. 1331–1340, Oct. 2010, <https://doi.org/10.1016/j.resconrec.2010.05.005>.
  - [10] F. A. Al-Fahdawi, A. I. Al-Hadithi, and J. A. Al-Asafi, "The Mechanical Properties of Ferrocement Mortar with Waste Plastic Fibers at Elevated Temperatures," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9347–9350, Oct. 2022, <https://doi.org/10.48084/etasr.5209>.
  - [11] I. M. Khan, S. Kabir, M. A. Alhussain, and F. F. Almansoor, "Asphalt Design Using Recycled Plastic and Crumb-rubber Waste for Sustainable Pavement Construction," *Procedia Engineering*, vol. 145, pp. 1557–1564, Jan. 2016, <https://doi.org/10.1016/j.proeng.2016.04.196>.
  - [12] H. A. Hasan, L. H. A. Mohammed, and L. G. G. Masood, "Effect of rubber tire on behaviour of subgrade expansive Iraqi soils," *IOP Conference Series: Materials Science and Engineering*, vol. 870, no. 1, Mar. 2020, Art. no. 012066, <https://doi.org/10.1088/1757-899X/870/1/012066>.
  - [13] M. M. Disfani, A. Arulrajah, M. W. Bo, and R. Hankour, "Recycled crushed glass in road work applications," *Waste Management*, vol. 31, no. 11, pp. 2341–2351, Nov. 2011, <https://doi.org/10.1016/j.wasman.2011.07.003>.
  - [14] M. M. E. Zumrawi and F. O. A. Khalil, "Experimental Study of Steel Slag Used as Aggregate in Asphalt Mixture," *American Journal of Construction and Building Materials*, vol. 1, no. 1, Feb. 2017, Art. no. 12, <https://doi.org/10.11648/j.ajcbm.20170101.12>.
  - [15] A. Mohajerani, J. Vajna, T. H. H. Cheung, H. Kurmus, A. Arulrajah, and S. Horpibulsuk, "Practical recycling applications of crushed waste glass in construction materials: A review," *Construction and Building Materials*, vol. 156, pp. 443–467, Dec. 2017, <https://doi.org/10.1016/j.conbuildmat.2017.09.005>.
  - [16] J. P. Castro-Gomes, A. P. Silva, R. P. Cano, J. Durán Suarez, and A. Albuquerque, "Potential for reuse of tungsten mining waste-rock in technical-artistic value added products," *Journal of Cleaner Production*, vol. 25, pp. 34–41, Apr. 2012, <https://doi.org/10.1016/j.jclepro.2011.11.064>.
  - [17] V. Dubois, N. E. Abriak, R. Zentar, and G. Ballivy, "The use of marine sediments as a pavement base material," *Waste Management*, vol. 29, no. 2, pp. 774–782, Feb. 2009, <https://doi.org/10.1016/j.wasman.2008.05.004>.
  - [18] C. S. Reddy, S. Mohanty, and R. Shaik, "Physical, chemical and geotechnical characterization of fly ash, bottom ash and municipal solid waste from Telangana State in India," *International Journal of Geo-Engineering*, vol. 9, no. 1, Dec. 2018, Art. no. 23, <https://doi.org/10.1186/s40703-018-0093-z>.
  - [19] T. G. dos Santos, L. F. R. Martins, and E. R. Sosa, "Technological Characterization of Tailings from Iron and Gold Mining with a Geoenvironmental Focus for Reuse in Geotechnical Application," in *8th International Congress on Environmental Geotechnics*, Hangzhou, China, Nov. 2018, pp. 253–260, [https://doi.org/10.1007/978-981-13-2227-3\\_31](https://doi.org/10.1007/978-981-13-2227-3_31).
  - [20] T. Thyagaraj, Ed., *Ground Improvement Techniques and Geosynthetics*. New York, NY, USA: Springer, 2018.
  - [21] N. Bheel, M. A. Jokhio, J. A. Abbasi, H. B. Lashari, M. I. Qureshi, and A. S. Qureshi, "Rice Husk Ash and Fly Ash Effects on the Mechanical Properties of Concrete," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5402–5405, Apr. 2020, <https://doi.org/10.48084/etasr.3363>.
  - [22] S. Ahmari and L. Zhang, "Utilization of cement kiln dust (CKD) to enhance mine tailings-based geopolymer bricks," *Construction and Building Materials*, vol. 40, pp. 1002–1011, Mar. 2013, <https://doi.org/10.1016/j.conbuildmat.2012.11.069>.
  - [23] A. B. Salahudeen, A. O. Eberemu, and K. J. Osinubi, "Assessment of Cement Kiln Dust-Treated Expansive Soil for the Construction of Flexible Pavements," *Geotechnical and Geological Engineering*, vol. 32, no. 4, pp. 923–931, Aug. 2014, <https://doi.org/10.1007/s10706-014-9769-0>.
  - [24] "BioData - Aquatic Bioassessment Data for the Nation," *USGS*. <https://apps.usgs.gov/biodata/>.
  - [25] B. Boumaza, R. Kechiched, and T. V. Chekushina, "Trace metal elements in phosphate rock wastes from the Djebel Onk mining area (Tébessa, eastern Algeria): A geochemical study and environmental implications," *Applied Geochemistry*, vol. 127, Apr. 2021, Art. no. 104910, <https://doi.org/10.1016/j.apgeochem.2021.104910>.
  - [26] J. L. Figueroa, L. Zhou, and W. F. Chang, "Use of Phosphate Mining Waste in Secondary Road Construction," in *Transportation Research Record*, Ithaca, New York, USA, Aug. 1987, vol. 2, pp. 59–64.
  - [27] W. Shen, M. Zhou, and Q. Zhao, "Study on lime-fly ash-phosphogypsum binder," *Construction and Building Materials*, vol. 21, no. 7, pp. 1480–1485, Jul. 2007, <https://doi.org/10.1016/j.conbuildmat.2006.07.010>.
  - [28] W. Shen, M. Zhou, W. Ma, J. Hu, and Z. Cai, "Investigation on the application of steel slag-fly ash-phosphogypsum solidified material as road base material," *Journal of Hazardous Materials*, vol. 164, no. 1, pp. 99–104, May 2009, <https://doi.org/10.1016/j.jhazmat.2008.07.125>.
  - [29] A. A. Ahmed and A. Z. M. Abouzeid, "Potential Use of Phosphate Wastes As Aggregates in Road Construction," *Journal of Engineering Sciences*, vol. 37, no. 2, pp. 413–422, Mar. 2009, <https://doi.org/10.21608/jesaun.2009.125357>.
  - [30] A. N. Ally, M. M. Blanche, U. J. P. Nana, M. M. Grâce, N. François, and C. Pettang, "Recovery of Mining Wastes in Building Materials: A Review," *Open Journal of Civil Engineering*, vol. 11, no. 4, pp. 379–397, Nov. 2021, <https://doi.org/10.4236/ojce.2021.114022>.
  - [31] H. Idrissi *et al.*, "Sustainable use of phosphate waste rocks: From characterization to potential applications," *Materials Chemistry and Physics*, vol. 260, Feb. 2021, Art. no. 124119, <https://doi.org/10.1016/j.matchemphys.2020.124119>.
  - [32] S. Chabou-Mostefai, "Etude de la serie phosphatee tertiaire du Djebel Onk (Algerie): stratigraphie, petrographie, mineralogie et analyse statistique," Ph.D. dissertation, Paul Cezanne University, Aix-en-Provence, France, 1987.
  - [33] L. Gadri *et al.*, "The quarries edges stability in opencast mines: a case study of the Jebel Onk phosphate mine, NE Algeria," *Arabian Journal of Geosciences*, vol. 8, no. 11, pp. 8987–8997, Nov. 2015, <https://doi.org/10.1007/s12517-015-1887-3>.
  - [34] R. Kechiched *et al.*, "Comprehensive REE + Y and sensitive redox trace elements of Algerian phosphorites (Tebessa, eastern Algeria): A geochemical study and depositional environments tracking," *Journal of Geochemical Exploration*, vol. 208, Jan. 2020, Art. no. 106396, <https://doi.org/10.1016/j.gexplo.2019.106396>.
  - [35] R. Kechiched, R. Laouar, O. Bruguier, S. Laouar-Salmi, O. Ameur-Zaimeche, and A. Foufou, "Preliminary Data of REE in Algerian Phosphorites: A Comparative Study and Paleo-redox Insights," *Procedia Engineering*, vol. 138, pp. 19–29, Jan. 2016, <https://doi.org/10.1016/j.proeng.2016.02.048>.
  - [36] R. Kechiched *et al.*, "Glauconite-bearing sedimentary phosphorites from the Tebessa region (eastern Algeria): Evidence of REE enrichment and geochemical constraints on their origin," *Journal of African Earth Sciences*, vol. 145, pp. 190–200, Sep. 2018, <https://doi.org/10.1016/j.jafrearsci.2018.05.018>.
  - [37] Y. Kassatkine, A. Yahyaoui, and S. Chatilov, "The works of prospecting and assessment on phosphate executed in 1976–1978 in the mining district of Djebel Onk," SONAREM (Société Nationale de Recherche et d'Exploration Minière), Internal report 2, 1980.

- [38] "Technical report on the exploitation of the Djebel El Onk quarry preliminary exploitation plan of the Bir el Ater deposit," CERAD, Algeria, 2017.
- [39] D. Nettour, M. Chettibi, G. Bulut, and A. Benselhoub, "Beneficiation of phosphate sludge rejected from djebel onk plant (Algeria)," *Mining of Mineral Deposits*, vol. 13, no. 4, 2019, <https://doi.org/10.33271/mining13.04.084>.
- [40] L. Bounemia and A. Mellah, "Characterization of crude and calcined phosphates of Kef Essennoun (Djebel Onk, Algeria)," *Journal of Thermal Analysis and Calorimetry*, vol. 146, no. 5, pp. 2049–2057, Dec. 2021, <https://doi.org/10.1007/s10973-020-10167-2>.
- [41] A. Soukeur, A. Szymczyk, Y. Berbar, and M. Amara, "Extraction of rare earth elements from waste products of phosphate industry," *Separation and Purification Technology*, vol. 256, Feb. 2021, Art. no. 117857, <https://doi.org/10.1016/j.seppur.2020.117857>.
- [42] R. Buccione, R. Kechiched, G. Mongelli, and R. Sinisi, "REEs in the North Africa P-Bearing Deposits, Paleoenvironments, and Economic Perspectives: A Review," *Minerals*, vol. 11, no. 2, Feb. 2021, Art. no. 214, <https://doi.org/10.3390/min11020214>.
- [43] S. Ferhaoui *et al.*, "Rare earth elements plus yttrium (REY) in phosphorites from the Tébessa region (Eastern Algeria): Abundance, geochemical distribution through grain size fractions, and economic significance," *Journal of Geochemical Exploration*, vol. 241, Oct. 2022, Art. no. 107058, <https://doi.org/10.1016/j.gexplo.2022.107058>.
- [44] CTTP, *Catalogue de Dimensionnement des Chaussees Neuves*. Algiers, Algeria: Organisme National de Controle Technique des Travaux Publics, 2001.

# Investigation of the Penetration Force of Disposable Sterile Needles through Biomedical Textile Surfaces

**Ersin Kayahan**

Laser Technologies Research and Application Center (LATARUM) and Electro-Optics and Systems Engineering, Kocaeli University, Turkey  
kayahan@kocaeli.edu.tr

**Sayit Ozbey**

Laser Technologies Research and Application Center (LATARUM) and Maritime Faculty, Marine Engineering, Kocaeli University, Turkey  
sayit.ozbey@kocaeli.edu.tr

**Ugur Kosa**

Biomedical Engineering, Technology Faculty, Kocaeli University, Turkey  
ugrkosa@gmail.com

**Mehmet Alp Ilgaz**

Faculty of Electrical Engineering, University of Ljubljana, Slovenia  
mehmetalp.ilgaz@gmail.com  
(corresponding author)

**Selma Corovic**

Faculty of Electrical Engineering, University of Ljubljana, Slovenia  
Selma.Corovic@fe.uni-lj.si

*Received: 5 November 2022 | Revised: 3 December 2022 | Accepted: 4 December 2022*

## ABSTRACT

Disposable sterile needles are essential highly consumed medical tools. Medical needles are usually manufactured according to standardized protocols, which currently do not provide the specified minimum tolerance value of the penetration force which strongly depends on needle dimensions, needle cutting edge angle, and the type of the tissue surface to be penetrated. In the present study, experimental measurements were performed according to the ISO 7864 standard to investigate the needle-surface penetration effect via the experimental assessment of the influence of the needle dimensions, cutting edge angle, and three different types of biomedical textiles/artificial tissues (i.e. polyurethane (PU), polypropylene (PP), and artificial leather (AL)) on the penetration force. The results indicate that the smaller the needle's cutting-edge angle, the smaller the penetration force across the target tissue surface. An exponential decaying relationship has been found between the penetration force and the needle diameter/gauge. The results also show that PP provides similar results to other materials that are already included in ISO 7864, and it has a good potential to be accepted as a standardized biomedical textile.

*Keywords-penetration force; sterile needle; biomedical textile; medical device; needle penetration*

## I. INTRODUCTION

Medical needle-based interventions via the insertion of disposable sterile needles into the target tissue are among the most prevalent clinical procedures [1-6]. Disposable sterile

needles are a prerequisite medical tool and an important component of medical devices daily used in medical settings and are produced in different diameters/gauges and lengths. Medical needles are used for various purposes such as spinal applications, dental treatments, veterinary services, etc. [2-7].

Since a medical procedure or therapy may strongly depend on the needle-tissue interaction and the needle tissue procedure they must be manufactured according to strict requirements defined by standards (e.g. ISO 7864, ISO 9626). The ISO 7864 standard specifies the testing procedures for stainless-steel medical needles and the necessary physical and chemical parameters related to the design and production of the needles including the procedures for injecting or withdrawing fluids from the human body via disposable sterile hypodermic needles [8]. This standard also determines the requirements for the diameter of the needles, which ranges from 0.18mm to 1.2mm. The testing procedure of sterile needles to determine the penetration force is also specified in the standard for different model materials/tissues used for testing. Specifically, the following four testing materials/tissues (used as human or animal tissue models) are included in the standard describing their surface properties and thickness through which the needle penetrates: latex rubber, polyurethane (PU), silicone rubber, and polyethylene with respective thickness and their tolerance of  $1\pm0.1\text{mm}$ ,  $0.40\pm0.05\text{mm}$ ,  $0.50\pm0.05\text{mm}$  and  $50\mu\text{m}$ . The ISO 9626 standard classifies the needle types according to the diameter values along with the diameter tolerance values for each needle. For the classification of the needle diameter the gauge (G) expression is used (i.e. the larger the gauge the thinner the needle) [9, 10].

It should be noted that despite the rigorous protocol defined in the standards, several aspects regarding the optimum needle design, needle-tissue interaction, and the adequate penetration force profile remain to be investigated. It is well known that the biggest risk factor of disposable sterile needles is the penetration procedure itself which can lead to injury when not applied optimally. Although the ISO 7864 standard defines protocols for the penetration testing for disposable sterile needles, there is no specified minimum limit or tolerance value for the penetration. Namely, numerous scientific and clinical studies are dealing with the improvement of the medical needle designs which would penetrate the skin or other tissue surfaces easily and safely with minimum influence caused to the tissue (i.e. minimum or no pain and/or other tissue trauma) and to assure an efficient and safe medical treatment or therapy. The critical factors that influence the needle-skin/tissue interaction procedure are the needle geometrical and material properties and the insertion method by carefully taking into account the target tissue properties [11]. Namely, the mechanical properties of the biological tissues vary due to their specific microstructures (i.e. homogeneity, heterogeneity, anisotropy, isotropy, etc.) and the human body itself (i.e. age, gender, BMI, stage of the disease, etc.) and thus may affect the efficacy of the needle's penetration. For example, during the needle insertion through the skin layer, in addition to the type of the needle, deformation of the skin surface may occur due to the skin elasticity, which is an important mechanical property, affecting the efficacy of the needle penetration. Authors in [12] investigated the skin penetration resistance variables such as the needle gauge, the needle insertion velocity, the pre-application pressure, and the penetration depth. Their study demonstrates that adequate pre-application pressure increases the effect of the needle penetration and makes it easier to reach the desired depth. They also demonstrated that even if the

diameter does not show too much effect on skin deterioration, it is directly proportional to the applied strength. Moreover, the penetration depth was confirmed as an essential factor that affects skin deformation during the process of penetration [12]. Authors in [13] determined the penetration force and friction force of hypodermic needles by using PU film and pigskin as models of human skin. In their experiment, 25–75% penetration was achieved successfully by a penetration force ranging from 7N to 10N. Previous studies which experimentally investigated the effect of needle diameter and cutting edge angle on needle insertion showed that beveled needles are more likely to bend, while needles with larger diameter increase friction, requiring more force to be applied to the needle [14, 15]. Furthermore, many recent analytical and numerical studies focused on the modeling and optimization of the needle characteristics (such as needle tip, needle diameter, and cutting edge angle) which became an important prerequisite in adequate needle design [14–16].

Prior to the application of newly designed needles to human or animal tissues/body, they always must be adequately tested. In vivo experimental investigation of the needle penetration through skin is complex and often difficult to be performed. Therefore, different artificial tissues/materials are often used in experimental studies and new materials are being investigated to propose reliable, useful, and cost-effective alternatives [11]. As an alternative to biological tissue, the PU substrate has been successfully used due to advantages such as being homogenous, chemically stable, and easily accessible [17]. Namely, PU and polypropylene (PP) have been efficiently used for various purposes in different applications [18]. PP can also be used as a sustainable material for the production of light weight foam [19]. In addition, PP fibers can be used instead of cotton, due to their admirable chemical stability and hydrophobic nature [20]. Thus, materials such as PU and PP can be used as important alternative biomaterials to efficiently simulate the tissue surface of human/animal skin in medical needle penetration studies.

It should be noted that in addition to the geometrical and material properties of the needles, the needle insertion procedure and the mechanical and geometrical properties of the tissue to be penetrated must be carefully examined. Therefore, the main objective of the present study is to experimentally investigate the key variables that may influence the penetration force profile of the disposable sterile needles such as needle dimensions (i.e. gauges), needle cutting edge angle, and the type of biomedical textile (i.e. human or animal tissue models) to be penetrated. First, our goal was to experimentally identify the relationship between the needle size versus penetration force value. Second, we investigated different needle cutting-edge angles aiming to minimize the penetration force. Finally, three different biomedical textile surfaces were examined, namely PU, PP, and Artificial Leather (AL), to study the needle-surface penetration effect. PP and AL were examined/tested for the first time in this study. They were compared to the reference material PU which has already been standardized by the ISO 7864 standard. We demonstrated that the PP provides similar results to other materials that are already included in ISO 7864 and therefore it has a good

potential to be accepted as a new standardized material for the testing procedure of sterile medical needles.

## II. MATERIALS AND METHODS

Experimental investigations have been conducted to examine the influence of the needle dimensions, needle cutting edge angle, and penetration surface type on the disposable needle penetration force value. All the experimental tests and measurements have been performed according to the ISO 7864 standard. The experiments have been conducted using needles with 6 different diameters (i.e. 18G, 20G, 21G, 25G, 27G, and 31G). These types are among the most frequently used types of needles in medical applications. The results obtained from the experimental measurements were statistically analyzed using the commercial software package Minitab. The influence of the penetration force using different needles (with different diameters) has been examined for 3 different biomedical textile surface types (i.e. PU, PP, and AL). Instron 594312393 measuring device was used to determine the force values by applying 100N load cells as depicted in Figure 1.

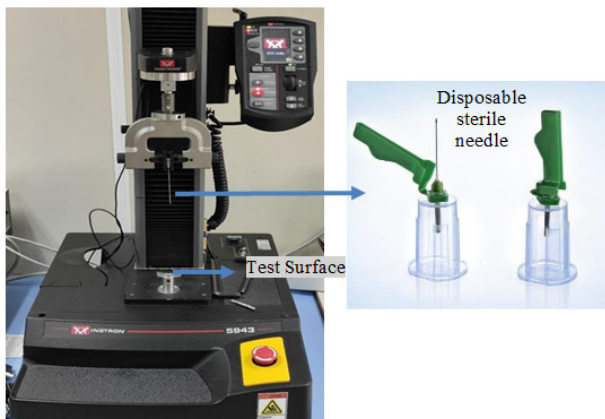


Fig. 1. The experimental measurement setup: the tension testing machine to determine the needle penetration force and the geometry of typical disposable sterile needles.

All analyzed parameters were tested 10 times while keeping the other variables constant. Within the first part of this study, the effect of the needle diameter on the penetration force was investigated using the 6 needles with the above specified diameters on PU, PP, and AL surfaces with a thickness of  $0.4\text{mm} \pm 0.05$  and cutting-edge angle of  $11^\circ \pm 2^\circ$ . To observe the needle cutting edge angle effect on the penetration force, 2 needles, 18G spinal and 18G veterinary steel were used. The cutting-edge angles were measured with the Olympus 3K46661 microscope. 18G veterinary steel needles have a cutting-edge angle of  $11^\circ \pm 2^\circ$ , while 18G spinal needles have a cutting-edge angle of  $17^\circ \pm 2^\circ$ . The physical and mechanical properties of PU and PP are given in Table I. The physical and mechanical properties of AL, which are similar to natural leather's, are given in Table II.

The needles entered into giving surfaces with a constant velocity of 100mm/min. These input values are recorded as a function of the penetration force. A force meter, such as a load cell, was used to measure the force at different stages of the

process. A needle was placed in a way that allowed the friction force to be calculated at the first moment of penetration and the moment of progress. After the forward movement, the needle was removed and the overall force profile was recorded according to the protocol specified in [8].

TABLE I. PHYSICAL AND MECHANICAL PROPERTIES OF PU AND PP [21-23]

Material	Density ( $\text{g/cm}^3$ )	Hardness	Tensile strength (MPa)	Elongation (%)
PU	1.18	80 (Shore A)	20	700
PP	0.95	60 (Shore D)	28.18	45.13

TABLE II. PHYSICAL AND MECHANICAL PROPERTIES OF AL [24, 25]

Density ( $\text{g/cm}^3$ )	Tensile energy ( $\text{gf}\cdot\text{cm/cm}^2$ )	Extension at max. load bending (%)	Shear stiffness ( $\text{gf/cm}\cdot\text{deg}$ )	Geometrical roughness ( $\mu\text{m}$ )
0.9 - 1.6	25.91	13.36	8	3.12

Needles are classified by gauges while they are manufactured according to ISO 9626 standard [9]. The requirements of the standard are given in Table III for needles manufactured between 14G and 32G. Small diameter needles (20-25 gauge) provide sufficient cytological material and often sufficient histological material. They can be used safely when multiple samples are needed and they minimize complications that can occur when reaching the target lesion. However, they are not practical, especially when reaching deep-seated lesions because they tend to deviate from the target. With larger diameter needles (14-19 gauge), it is easier to reach the lesion directly and usually a better sample with fewer entries is provided for cytology and histology, but the risk of bleeding increases with increasing needle diameter [26-29].

TABLE III. OUTER DIAMETER REQUIREMENTS OF NEEDLES ACCORDING TO THE ISO 9626 [9]

Gauge (G)	Minimum outer diameter (mm)	Maximum outer diameter (mm)
18	1.200	1.300
20	0.560	0.920
21	0.800	0.830
25	0.500	0.530
27	0.400	0.420
31	0.254	0.267

## III. RESULTS AND DISCUSSION

The obtained experimental results indicating the penetration force dependency on the needle diameter (the needle gauge) are shown in Figure 2. The obtained results indicate that the penetration force decreases when the gauge size increases, which is in agreement with the results reported in [30]. In our study, the highest penetration force was observed when using the 18G needles with average values of 3.37N, 3.46N, and 3.14N for PU, PP, and AL surfaces, respectively. The results shown in Figure 2 indicate that even though the 20G needles are 28% thinner than 18G needles, the decrease in penetration force was 71% on average for the PU surface.



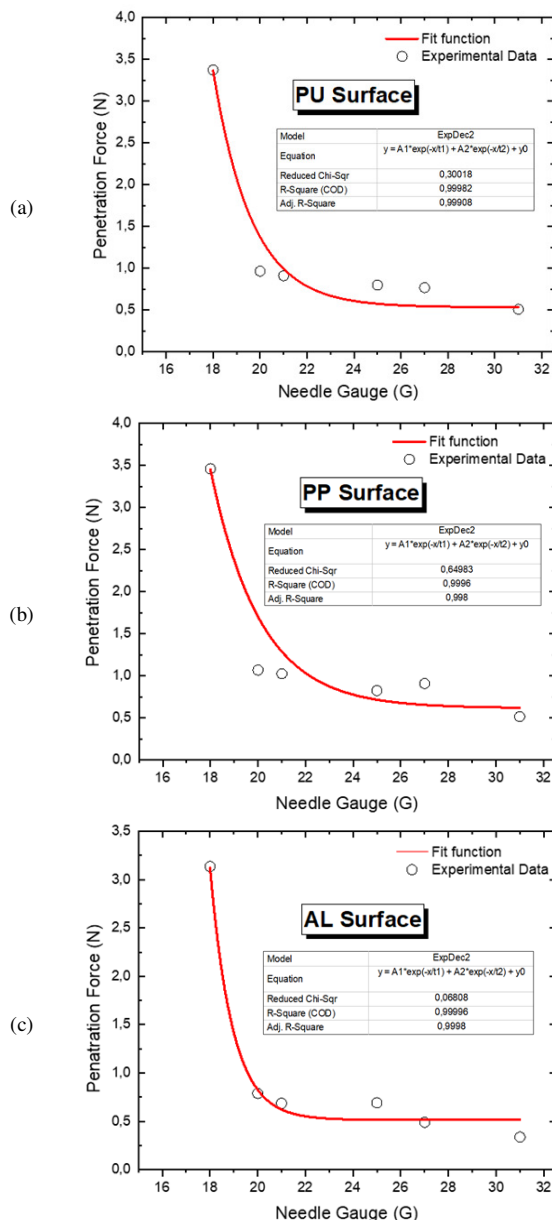


Fig.2. The changes of the needle penetration force with the needle diameter size (gauge) for the (a) PU, (b) PP, and (c) AL surfaces. The solid line shows the second order exponential decay fit function and the circles represent the experimentally measured data.

By using the same comparison method, the results demonstrate that the 21G needles are 9% thinner than 20G and 25G needles are 12% thinner than 21G needles, and the decrease in penetration force was 8% and 12%, respectively for the PU surface. The diameter of 31G needles which is the smallest needle diameter in the sample group is 37% smaller than that of the 27G needles and the penetration force was decreased by 34% for the PU surface. It can be concluded that a decrease in penetration force occurs with a decrease in the needle diameter values. Based on the curve fitting performed in Origin Lab software we demonstrated that the penetration force vs. needle diameter relationship can be described with a

second-order exponential decay function for all material (i.e. PU, PP, and AL) surfaces, as shown in Figure 2. To determine the penetration force changes at different surfaces (i.e. different materials) which should be used for the standard material surface, AL and PP were used and compared with the PU surface which is used as a reference. The obtained results are given in Figure 3 and Table IV. It can be seen that the AL surface exhibits the best performance with the minimum penetration force in comparison with the PP and PU surfaces. The results indicate that the PP surface exhibits a significant similarity with the PU surface for the needle types of 18G, 20G, 21G, 25G, 27G, and 31G. Thus, the PP surface could be a good candidate as a new standardized material surface. It should be also emphasized that PP is cheaper than PU. Therefore, the PP surface could be also used as a reference surface according to the ISO 7864 standard in addition to the PU surface, since it is cost-effective and more accessible. However, the PP must also satisfy other requirements to be fully accepted by ISO 7864 standard.

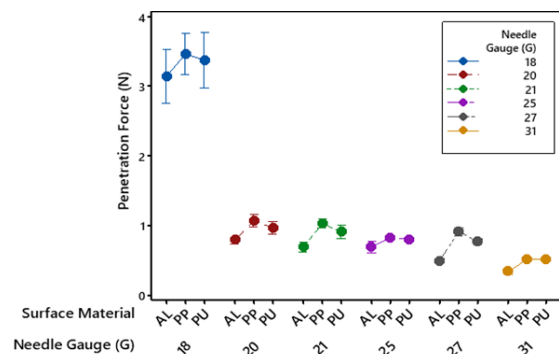


Fig. 3. The variation of penetration force with AL, PP, and PU surfaces.

Data distribution of the samples is found to be statistically significant since it passed the normality test ( $P\text{-value} \leq 0.05$ ), as shown in Table V. Therefore, by using an average penetration force, a significant inference can be obtained. The largest difference in penetration force between PU and PP is only 2% (0.02N) for 18G, 25G, and 31G needles. The PU and PP surfaces have almost the same results in terms of calculated penetration forces. The variation of needle cutting edge angle with the penetration force is given in Figure 4. The average penetration force value was 2.15N for a cutting edge angle of  $11^\circ \pm 2^\circ$  and 3.3 N for  $17^\circ \pm 2^\circ$  for PU. The results in Figure 4 demonstrate that when the needle cutting-edge angles increase, the penetration force also increases for all the examined surfaces, which is also in agreement with the results of [31]. The standard deviation (StDev) of cutting-edge angles of  $11^\circ \pm 2^\circ$  and  $17^\circ \pm 2^\circ$  is 0.1814 and 0.2387, respectively, as shown in Table VI for the PU material. The results were obtained from penetration force measurements performed 10 times. The increase in the penetration force with the cutting-edge angle of  $17^\circ \pm 2^\circ$  was found to be 56%, 16%, and 12% than with  $11^\circ \pm 2^\circ$  for PU, PP, and AL surfaces, respectively. This can be explained by the fact that the increase of the cutting-edge angle increases the contact area between the needle and the surface of the material, which causes a higher penetration force.

TABLE IV. VARIATION OF THE PENETRATION FORCE VALUES WITH DIFFERENT NEEDLE GAUGES AND SURFACE MATERIALS

Gauge	18G			20G			21G			25G			27G			31G		
Sample	PU	PP	AL	PU	PP	AL	PU	PP	AL	PU	PP	AL	PU	PP	AL	PU	PP	AL
1	3.08	3.17	2.92	0.89	1.13	0.77	0.89	1.04	0.81	0.82	0.81	0.67	0.82	0.83	0.43	0.49	0.49	0.32
2	3.51	3.83	3.99	0.93	0.94	0.88	0.99	0.92	0.78	0.75	0.77	0.56	0.84	0.92	0.49	0.48	0.51	0.46
3	2.86	3.81	2.71	0.91	1.06	0.81	0.80	0.98	0.56	0.81	0.82	0.58	0.82	0.98	0.52	0.49	0.50	0.28
4	2.89	3.31	3.93	1.18	0.91	0.72	0.78	1.03	0.61	0.76	0.77	0.87	0.70	0.76	0.56	0.48	0.51	0.30
5	4.30	3.85	3.12	0.89	1.34	0.76	1.09	1.07	0.58	0.85	0.84	0.58	0.80	0.93	0.41	0.55	0.55	0.29
6	3.71	2.96	2.27	0.93	1.12	0.91	0.99	0.91	0.65	0.76	0.85	0.64	0.83	1.02	0.47	0.45	0.56	0.37
7	3.11	3.06	3.17	1.20	1.04	0.68	1.15	1.23	0.83	0.84	0.88	0.73	0.73	0.84	0.46	0.55	0.49	0.34
8	4.28	4.13	2.88	0.80	0.98	0.79	0.80	1.02	0.71	0.80	0.85	0.61	0.63	0.93	0.51	0.48	0.52	0.29
9	3.19	3.09	3.55	0.98	1.02	0.82	0.79	0.99	0.74	0.84	0.84	0.84	0.71	0.98	0.57	0.53	0.50	0.33
10	2.79	3.38	2.82	0.93	1.14	0.74	0.81	1.05	0.59	0.75	0.81	0.83	0.79	0.89	0.48	0.59	0.52	0.40

TABLE V. DATA ANALYSIS OF NEEDLE PENETRATION FORCE VALUES WITH DIFFERENT NEEDLE GAUGES AND SURFACE MATERIALS WITH NORMALITY TEST

Gauge (G)	Surface material	Number of tests	Mean	95% CI of mean	StDev	95% CI of StDev	Min	Median	Max	Normality test	
										P	Decision
18	PU	10	3.37	(2.97, 3.77)	0.56	(0.39, 1.03)	2.79	3.15	4.30	0.111	Pass
18	PP	10	3.46	(3.16, 3.75)	0.41	(0.28, 0.75)	2.96	3.35	4.13	0.190	Pass
18	AL	10	3.14	(2.75, 3.53)	0.55	(0.38, 0.99)	2.27	3.02	3.99	0.482	Pass
20	PU	10	0.95	(0.88, 1.03)	0.11	(0.07, 0.20)	0.80	0.93	1.18	0.054	Pass
20	PP	10	1.07	(0.98, 1.16)	0.12	(0.08, 0.23)	0.91	1.05	1.34	0.496	Pass
20	AL	10	0.79	(0.74, 0.84)	0.07	(0.05, 0.13)	0.68	0.78	0.91	0.925	Pass
21	PU	10	0.91	(0.81, 1.01)	0.14	(0.09, 0.25)	0.78	0.85	1.15	0.056	Pass
21	PP	10	1.02	(0.96, 1.09)	0.09	(0.06, 0.16)	0.91	1.03	1.23	0.207	Pass
21	AL	10	0.69	(0.61, 0.76)	0.10	(0.07, 0.18)	0.56	0.68	0.83	0.391	Pass
25	PU	10	0.80	(0.77, 0.83)	0.04	(0.03, 0.07)	0.75	0.81	0.85	0.136	Pass
25	PP	10	0.82	(0.80, 0.85)	0.04	(0.02, 0.06)	0.77	0.83	0.88	0.443	Pass
25	AL	10	0.69	(0.61, 0.78)	0.12	(0.08, 0.22)	0.56	0.66	0.87	0.142	Pass
27	PU	10	0.77	(0.72, 0.82)	0.07	(0.05, 0.13)	0.63	0.80	0.84	0.141	Pass
27	PP	10	0.91	(0.85, 0.96)	0.08	(0.05, 0.14)	0.76	0.93	1.02	0.702	Pass
27	AL	10	0.49	(0.45, 0.53)	0.05	(0.03, 0.09)	0.41	0.49	0.57	0.941	Pass
31	PU	10	0.51	(0.48, 0.54)	0.04	(0.03, 0.08)	0.45	0.49	0.59	0.135	Pass
31	PP	10	0.52	(0.50, 0.53)	0.03	(0.02, 0.05)	0.49	0.51	0.56	0.054	Pass
31	AL	10	0.34	(0.30, 0.38)	0.06	(0.04, 0.10)	0.28	0.33	0.46	0.203	Pass

TABLE VI. STATISTICAL DATA ANALYSIS OF THE NEEDLE CUTTING-EDGE ANGLE EFFECT ON THE PENETRATION FORCE

Gauge (G)	Surface material	Cutting-edge angle	Test number	Mean	StDev	95% CI
18	PU	11° ± 2°	10	2.15	0.181	(2.01, 2.29)
18	PU	17° ± 2°	10	3.36	0.239	(3.22, 3.50)
18	PP	11° ± 2°	10	2.48	0.091	(2.42, 2.54)
18	PP	17° ± 2°	10	2.89	0.093	(2.83, 2.95)
18	AL	11° ± 2°	10	1.99	0.073	(1.93, 2.04)
18	AL	17° ± 2°	10	1.23	0.086	(2.17, 2.28)

#### IV. CONCLUSIONS

In this study, the effects of needle diameter, cutting edge angle, and material surface type on the penetration force were studied. Based on the experimental investigation the following conclusions were derived:

- The penetration force reduces with needle diameter. The penetration force versus needle diameter relationship can be described with an exponentially decreasing function (a second-order exponential decay function).
- The penetration force increased when the cutting edge angle of the needle increased from 11°±2° to 17°±2° for the 18G needle type. Consequently, the minimum cutting-edge angle reduces the damage to the penetrated skin.
- PP surfaces exhibit similar penetration properties to the standardized PU material surfaces for the 18G-31G medical needles. Therefore, the PP surfaces can be used in experiments to determine the penetration force. Other advantages of the PP are its cost-effectiveness and accessibility. PP has a good potential to be accepted as a new standardized model material for the testing procedures of sterile needles. Further experimental investigations should be conducted in order to characterize PU for other types of needles.
- AL surfaces have less penetration force than the PP and PU surfaces. However, the penetration force of AL does not match the penetration values of the reference surface (PU).

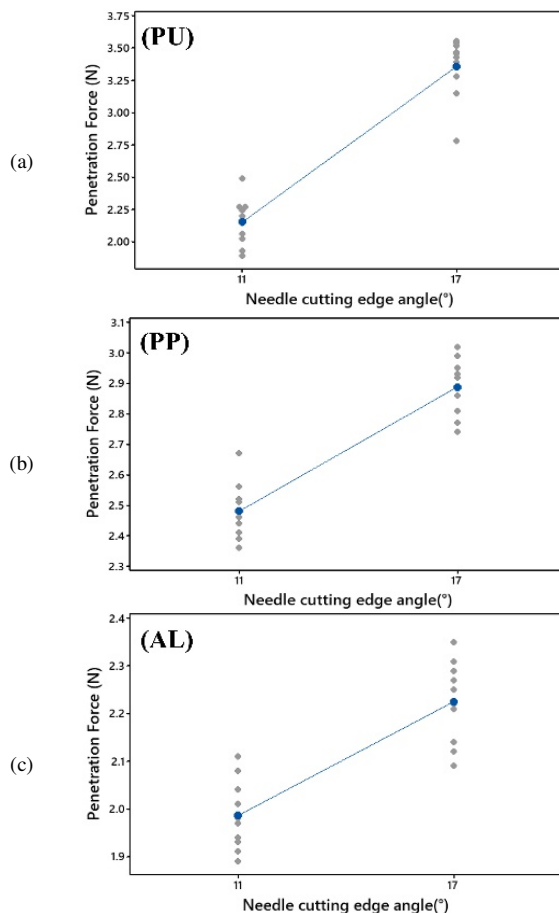


Fig. 4. The variation of penetration force with needle cutting edge angle for (a) PU, (b) PP, and (c) AL surfaces.

In conclusion, the results of our study provide a promising start to specify the limit values for the penetration force which should be considered in when producing standardized needles. Moreover, the limit penetration force value to be determined within the manufacturing standard should be independently calculated for each product, by considering the needle gauge size. It is also recommended to determine the limit value by taking into consideration the cutting-edge angles for needles which are manufactured at different cutting-edge angles with the same gauge size. This study provides a supplementary resource to the manufacturers' tests to evaluate the penetration force during the production stage of the needles. It should be also noted that for the first time alternative materials such as polypropylene and artificial leather were examined in this study to assess the penetration force of disposable sterile needles, and the results are reported with the standardized reference material (polyurethane) used as a reference. The obtained results indicate that polypropylene has a strong potential to be accepted as a new material for the efficient evaluation of the penetration force of sterile medical needles.

#### ACKNOWLEDGMENT

This research was partly supported by Kocaeli University-Laser Technologies Research and Application Center (LATARUM).

#### REFERENCES

- [1] V. D. Sree, A. Ardekani, P. Vlachos, and A. B. Tepole, "The biomechanics of autoinjector-skin interactions during dynamic needle insertion," *Journal of Biomechanics*, vol. 134, Mar. 2022, Art. no. 110995, <https://doi.org/10.1016/j.jbiomech.2022.110995>.
- [2] C. A. Luna, R. Tulcan-Toro, F. Romero, and M. F. Luna, "Measurement of residual volume in spinal needles after spinal anesthesia," *Colombian Journal of Anesthesiology*, vol. 45, pp. 12–15, Dec. 2017, <https://doi.org/10.1016/j.rcae.2017.08.013>.
- [3] L. Yeo *et al.*, "The Development of the Modern Prostate Biopsy," in *Prostate Biopsy*, London, UK: IntechOpen, 2011.
- [4] C. Eraslan, O. F. K. Koseoglu, N. Meydan, N. Culhaci, and A. Oral, "Comparison of the results of ultrasonography-guided percutaneous liver mass biopsy performed with 18 and 20 gauge needles," *Ege Tip Dergisi*, vol. 58, no. 1, pp. 8–12, Mar. 2019, <https://doi.org/10.19161/etd.418122>.
- [5] L. Arendt-Nielsen, H. Egekvist, and P. Bjerring, "Pain following controlled cutaneous insertion of needles with different diameters," *Somatosensory & Motor Research*, vol. 23, no. 1–2, pp. 37–43, Jan. 2006, <https://doi.org/10.1080/08990220600700925>.
- [6] K. L. Reed, S. F. Malamed, and A. M. Fonner, "Local Anesthesia Part 2: Technical Considerations," *Anesthesia Progress*, vol. 59, no. 3, pp. 127–137, Oct. 2012, <https://doi.org/10.2344/0003-3006-59.3.127>.
- [7] S. S. Meschi, A. Farghadan, and A. Arzani, "Flow topology and targeted drug delivery in cardiovascular disease," *Journal of Biomechanics*, vol. 119, Apr. 2021, Art. no. 110307, <https://doi.org/10.1016/j.jbiomech.2021.110307>.
- [8] *ANSI ISO 7864:(2016), Sterile Hypodermic Needles for Single Use*. Washington, DC, USA: American National Standards Institute, 2016.
- [9] *BS EN ISO 9626:(2016), Stainless steel needle tubing for the manufacture of medical devices. Requirements and test methods*. London, UK: British Standards Institution, 2016.
- [10] S. Vogels *et al.*, "Measuring intracompartmental pressures for the chronic exertional compartment syndrome: Challenging commercially available devices and their respective accuracy," *Journal of Biomechanics*, vol. 135, Apr. 2022, Art. no. 111026, <https://doi.org/10.1016/j.jbiomech.2022.111026>.
- [11] C. Yang, Y. Xie, S. Liu, and D. Sun, "Force Modeling, Identification, and Feedback Control of Robot-Assisted Needle Insertion: A Survey of the Literature," *Sensors*, vol. 18, no. 2, Feb. 2018, Art. no. 561, <https://doi.org/10.3390/s18020561>.
- [12] K. D. Butz *et al.*, "Prestress as an optimal biomechanical parameter for needle penetration," *Journal of Biomechanics*, vol. 45, no. 7, pp. 1176–1179, Apr. 2012, <https://doi.org/10.1016/j.jbiomech.2012.01.049>.
- [13] E. Busillo and J. S. Colton, "Characterization of Plastic Hypodermic Needles," *Journal of Medical Devices*, vol. 3, no. 4, Nov. 2009, Art. no. 041004, <https://doi.org/10.1115/1.4000452>.
- [14] M. D. O'Leary, C. Simone, T. Washio, K. Yoshinaka, and A. M. Okamura, "Robotic needle insertion: effects of friction and needle geometry," in *International Conference on Robotics and Automation (Cat. No.03CH37422)*, Taipei, Taiwan, Sep. 2003, vol. 2, pp. 1774–1780 vol.2, <https://doi.org/10.1109/ROBOT.2003.1241851>.
- [15] K. Ehmann and K. Malukhin, "A Generalized Analytical Model of the Cutting Angles of a Biopsy Needle Tip," *Journal of Manufacturing Science and Engineering*, vol. 134, no. 6, Nov. 2012, Art. no. 061001, <https://doi.org/10.1115/1.4007712>.
- [16] A. Wittek, G. Bourantas, B. F. Zwick, G. Joldes, L. Esteban, and K. Miller, "Mathematical modeling and computer simulation of needle insertion into soft tissue," *PLOS ONE*, vol. 15, no. 12, Nov. 2020, Art. no. e0242704, <https://doi.org/10.1371/journal.pone.0242704>.
- [17] L. J. Pavlovich, W. L. McClung, J. G. Thacker, R. F. Edlich, and G. T. Rodeheaver, "A synthetic membrane for testing needle penetration," *Journal of Applied Biomaterials*, vol. 4, no. 2, pp. 157–160, 1993, <https://doi.org/10.1002/jab.770040207>.
- [18] F. A. AlFaraidy and S. Azzam, "Residential Buildings Thermal Performance to Comply With the Energy Conservation Code of Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3949–3954, Apr. 2019, <https://doi.org/10.48084/etasr.2536>.

- [19] A. W. Ali and N. M. Fawzi, "Production of Light Weight Foam Concrete with Sustainable Materials," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7647–7652, Oct. 2021, <https://doi.org/10.48084/etasr.4377>.
- [20] Z. A. Hussain and N. Aljalawi, "Effect of Sustainable Glass Powder on the Properties of Reactive Powder Concrete with Polypropylene Fibers," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8388–8392, Apr. 2022, <https://doi.org/10.48084/etasr.4750>.
- [21] Y. Mobarak and A. Thabet, "Predictable Models and Experimental Measurements for Electric Properties of Polypropylene Nanocomposite Films," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 1, pp. 120–129, Feb. 2016, <https://doi.org/10.11591/ijece.v6i1.9108>.
- [22] H. O. Gulsoy and M. Tasdemir, "Physical and Mechanical Properties of Polypropylene Reinforced with Fe Particles," *International Journal of Polymeric Materials and Polymeric Biomaterials*, vol. 55, no. 8, pp. 619–626, Aug. 2006, <https://doi.org/10.1080/00914030500257664>.
- [23] H. H. Kim, M. Mazumder, S.-J. Lee, and M.-S. Lee, "Laboratory Evaluation of Sustainable PMA Binder Containing Styrene-Isoprene-Styrene (SIS) and Thermoplastic Polyurethane," *Sustainability*, vol. 12, no. 23, Jan. 2020, Art. no. 10057, <https://doi.org/10.3390/su122310057>.
- [24] E. K. Roh, "Mechanical properties and preferences of natural and artificial leathers, and their classification with a focus on leather for bags," *Journal of Engineered Fibers and Fabrics*, vol. 15, no. 1, pp. 1–10, Jan. 2020, <https://doi.org/10.1177/1558925020968825>.
- [25] K. Gerlach, H. J. Pitowski, and K. Schneider, "Synthetic leather product and method of production," US3974320A, Aug. 10, 1976.
- [26] I. N. Akpınar and T. Y. Kuzan, "Perkutan Biyopsi: Igne Secimi ve Goruntuleme Kilavuzlari," *Turk Radyoloji Seminerleri*, vol. 3, pp. 159–168, 2015, <https://doi.org/10.5152/trs.2015.215>.
- [27] J. W. Charboneau, C. C. Reading, and T. J. Welch, "CT and sonographically guided needle biopsy: current techniques and new innovations," *AJR American journal of roentgenology*, vol. 154, no. 1, pp. 1–10, Jan. 1990, <https://doi.org/10.2214/ajr.154.1.2104689>.
- [28] K. D. Hopper, C. S. Abendroth, K. W. Sturtz, Y. L. Matthews, L. A. Stevens, and S. J. Shirk, "Automated biopsy devices: a blinded evaluation," *Radiology*, vol. 187, no. 3, pp. 653–660, Jun. 1993, <https://doi.org/10.1148/radiology.187.3.8497611>.
- [29] C. Reading, J. Charboneau, E. James, and M. Hurt, "Sonographically guided percutaneous biopsy of small (3 cm or less) masses," *American Journal of Roentgenology*, vol. 151, no. 1, pp. 189–192, Jul. 1988, <https://doi.org/10.2214/ajr.151.1.189>.
- [30] A. C. Goncalves, S. Cavassana, F. R. Chavarette, R. Outa, S. J. Casarin, and A. V. Corazza, "Variation of the Penetration Effort in an Artificial Tissue by Hypodermic Needles," *Journal of Healthcare Engineering*, vol. 2020, Sep. 2020, Art. no. 8822686, <https://doi.org/10.1155/2020/8822686>.
- [31] S. Aoyagi, H. Izumi, and M. Fukuda, "Biodegradable polymer needle with various tip angles and consideration on insertion mechanism of mosquito's proboscis," *Sensors and Actuators A: Physical*, vol. 143, no. 1, pp. 20–28, May 2008, <https://doi.org/10.1016/j.sna.2007.06.007>.

# Power Quality Enhancement in a Wind Farm Connected Grid with a Fuzzy-based STATCOM

**Ramchandra Adware**

Department of Electrical Engineering, G H Raisoni College of Engineering, India  
ramchandra.adware@raisoni.net  
(corresponding author)

**Vinod Chandrakar**

Department of Electrical Engineering, G H Raisoni College of Engineering, India  
vinod.chandrakar@raisoni.net

Received: 9 November 2022 | Revised: 30 November 2022 | Accepted: 2 December 2022

## ABSTRACT

Integrating wind farms with electricity transmission networks presents several problems, and power quality plays a vital role among these. This study proposed a novel fuzzy logic controller to reduce the effect of power quality issues in such applications and investigated two different FACTS (Flexible AC Transmission System) devices, the Static Var Compensator (SVC) and the Static Synchronous Compensator (STATCOM). The fuzzy logic controller was designed as a voltage controller to improve power quality. Detailed analysis was carried out to investigate the successful mitigation of voltage sag/swell, active and reactive power improvement, and voltage flickers control by two controllers in a Multi-Terminal Load (MTL) system consisting of wind generators. The results were verified in MATLAB simulations, and a comparison was performed between the proposed shunt controllers. Furthermore, the design of a fuzzy interference system based on  $V_{ref}$  (reference voltage) and the voltage measured at STATCOM location signals was investigated and compared with the SVC in terms of voltage sag, voltage swell, voltage flickers, and Total Harmonic Distortions (THD). The proposed fuzzy system was compared with a PI-based STATCOM and SVC. The power quality issues were exacerbated when using Multi-Terminal Load (MTL) in the transmission network.

**Keywords-**Voltage sag/swell; voltage distortion; THD; SVC; STATCOM; FLC

## I. INTRODUCTION

Wind power's contribution to power generation is rapidly increasing these days. However, the inherent problem of wind power generation is its low quality due to the intermittent power generation that influences the connected grid. Power quality problems, such as harmonic distortions, voltage instability due to voltage sag, voltage swell, and reliability, are concerns that must be addressed [1-2]. Wind Energy Generation Systems (WEGS) with double-fed IG are used in power generation due to their low cost and variable speed constant frequency operation [3]. However, since wind power is typically located far from consumers, it is necessary to provide shunt compensation at Points of Common Coupling (PCC) [4], while Multi Terminal Load (MTL) connected to the wind farm at PCC produces power quality issues such as frequency variations and voltage sags/swells [5]. Voltage flickers are abrupt frequency changes for short durations due to load switching and integration of compensatory devices [6]. As a result, voltage stability suffers from reactive power demand in PCC [7]. If the voltage sag persists for longer durations, it

may lead to voltage collapse and longer-lasting voltage swells or flickers can introduce harmonics into the system [8]. The current produced due to sags/swells or flickers affects the switchgears and the protection system of the PCC. Therefore, various compensators with filters have been adopted to mitigate these power quality problems [9-12]. Many power quality conditioners have been implemented to address power quality issues in distribution networks, such as UPQC, DVR, and DSTATCOM [4, 9, 13, 14]. Among the different FACTS controllers, SVC and STATCOM play an important role in reactive power compensation [12, 16]. In [8], two types of FACTS controllers, SVC and STATCOM, were considered to mitigate power quality issues. However, when selecting a FACTS device, the performance of different types of shunt compensators should be examined based on fast response, reactive power compensator range, switching losses, resonance problems, energy source requirements, and DC link capacitor value [13]. Various control techniques have been proposed for the operating performance of various shunt compensators, such as nonlinear control, Lyapunov function-based control [17], instantaneous p-q theory [18], negative and zero sequence

control [6, 19-20], instantaneous d-q theory [13], backpropagation control [4], and the instantaneous id-iq method [7, 16]. Hybrid structures were introduced in [7, 21] to improve STATCOM operation with lower current rating APFs and PPFs. A hybrid structure fails during capacitive load and gives better performance only during inductive loading [13]. To improve the compensation range with lower current characteristics of the APF, one more hybrid device with SVC in parallel with an active power filter in the phase distribution system was evaluated. The APF was used primarily to eliminate harmonics and compensate for reactive power on the load bus left by SVC [8]. However, it does not work on high-voltage transmission lines because the cost of APF is higher with multilevel structures. Some studies investigated the use of PI controllers [12]. Furthermore, controllers such as PI and PID require an intensive mathematical model of the system, making them more sensitive to changes in the system parameters [20]. As a result, they failed to meet the requirements of robust system performance.

Many studies investigated the reactive power compensation for the stability of a power system, however, very few tested power quality issues. Shunt FACTS devices such as SVC and STATCOM are popular for voltage regulation and voltage stability improvements. PI-based SVC and STATCOM are mainly used in industry. The PI gains  $k_p$  and  $k_i$  are fixed and unchanged under changing system operating conditions, therefore STATCOM and SVC performance suffers. Fuzzy-based SVC and STATCOM are more suitable to replace PI. Very few studies addressed fuzzy-based FACTS shunt FACTS devices for power quality issues.

## II. PROPOSED TEST SYSTEM

The test system consists of two generator systems of 500kV, 3000MVA, and 500kV, 2500MVA, connected with buses B1, B2, and B3, as shown in Figure 1. The multi-terminal load system of 6000MW load with a wind generator of 9MW was connected to bus B3. Power quality issues were considered while the wind farm was connected to PCC or at load bus B3. Power quality analysis was performed while B3 was connected to the MTL load with the wind generator. The proposed model was simulated in MATLAB/Simulink to investigate the performance of the shunt compensator with SVC, STATCOM, and STATCOM with FLC3.

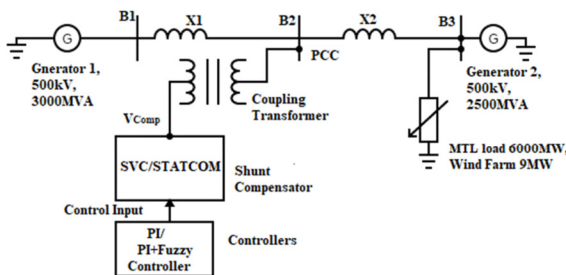


Fig. 1. Proposed system model.

### A. Proposed System with SVC

As shown in Figure 2, SVC dynamically generates static reactive power to adjust the characteristics of a transmission

line by altering the reactive impedance in its network. SVC is mostly used to maintain reactive power using different static power electronics controllers, such as a Thyristor Controlled Reactor (TCR) and a Thyristor Switched Capacitor (TSC). The transmission line requires reactive power, which is supplied and absorbed by the SVC. The SVC keeps the system voltage constant at its endpoints. When the system voltage is low, it produces capacitive reactive power. When the system voltage is high, inductive reactive power can be produced by simply changing the impedance using a thyristor switching circuit. Due to their ability to quickly compensate for power loss with minimal maintenance, shunt compensators are increasingly widely used in power systems.

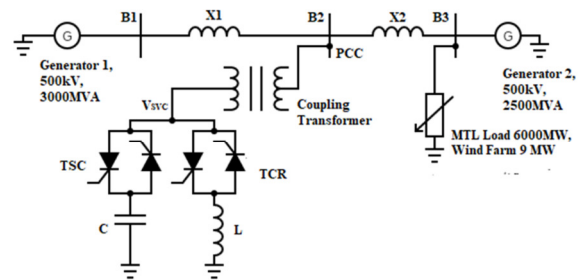


Fig. 2. Test system with SVC.

### B. Proposed System with STATCOM

As illustrated in Figure 3, the converter output is changed to achieve reactive power transfer between the AC transmission network and the converter. If the output voltage exceeds the system bus voltage, the converter supplies capacitive reactive power and the converter's reactive current will go to the utility. The converter absorbs inductive reactive power from the utility when the output voltage is lower than the system voltage. As a result, there is no reactive current between the utility and STATCOM, as it serves as a floating element in the system when the utility and the converter output voltages are equal.

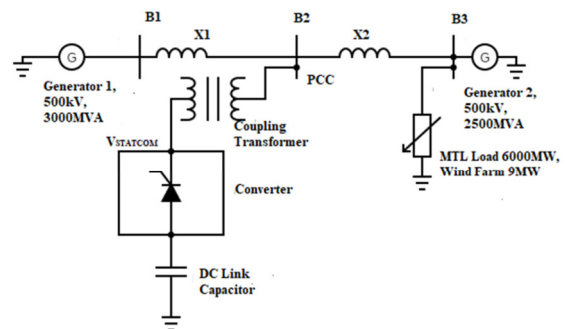


Fig. 3. Test system with STATCOM.

### C. Implementation of the Proposed FLC

The hybrid fuzzy controller was designed using the FIS GUI editor tools in the fuzzy logic toolbox. The suggested fuzzy controller was modeled on Mamdani's controller, which employs an "if-then" logic as part of its inference engine. Each controller input is taken into account as a fuzzy variable using



membership functions. A triangular membership function is used for input and output variables. Figure 4 shows the function of the fuzzy controller, which consists of three functional blocks: fuzzification, rule inference, and defuzzification. Error  $e$  and the change in error  $de$  are the input variables in the proposed system, and the FLC output is a reactive current.

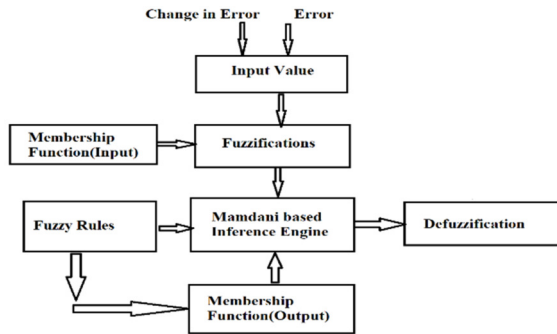


Fig. 4. Fuzzy logic flow diagram.

#### D. Hybrid FLC for STATCOM

Figure 5 shows the FLC scheme and the hybrid controller being the hybridization of a PI controller with a fuzzy logic controller. To obtain precise output from the proposed controller, the initial error signal generated by computing  $V_{ref}$  and measured values from the PI Controller is then tuned, and an  $I_{qref}$  value is obtained to compensate for the reactive current. The error signal is once again sent to the FLC in the proposed hybrid controller, along with the error and change in error as a second variable to obtain the accurate estimation of  $I_{qref}$ . Table I shows a list of the proposed FLC governing rules. The error and the change in error are provided for the fuzzy sets. To eliminate inaccuracy and improve the dynamic response of the controller, conditions for 9 sets of rules were applied.

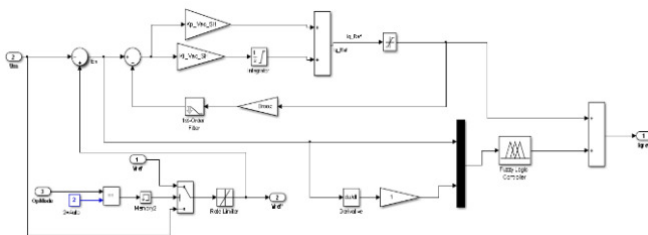


Fig. 5. Proposed fuzzy hybrid controller.

TABLE I. FUZZY RULES

de/e	N	Z	P
N	N	N	Z
Z	N	Z	P
P	Z	P	P

### III. POWER QUALITY CASE STUDIES

This section investigates the performance of the proposed hybrid fuzzy controller in a system with SVC, STATCOM, and STATCOM with FLC, under various power quality issues, when connected with an MTL and a 9MW wind farm.

#### A. Voltage Sag Analysis

Voltage sag is a frequent and severe issue caused by system faults, rapid increases in load, or the start of powerful motors. Figure 6 shows a line-to-line ground fault that occurred between 0.2 and 0.3s with an  $80\Omega$  fault resistance and a  $0.001\Omega$  ground resistance close to bus 1. The voltage sag was developed in the given system and computed with SVC and STATCOM.

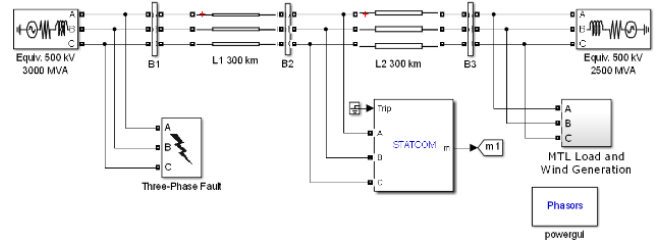


Fig. 6. Proposed system with Voltage sag.

At PCC, a STATCOM and an SVC are placed independently. Four parameters were used to investigate the dynamic behavior of the system. The first step was to investigate the effect of voltage sag on transient voltage stability by measuring the  $V_{rms}$  voltage at the PCC with and without a compensator. Similarly, active, reactive power, and harmonics were measured during fault, pre-fault, and post-fault phases. When the reactive power absorbed by the load bus increases, the voltage decreases, requiring FACTS controllers to compensate for the reactive power consumed in the system. A short-circuit line-to-line fault was simulated at generator bus B1 to investigate the effects of voltage sag. Figure 7 shows the system's condition during a short circuit without a compensator. The  $V_{rms}$  voltage drops to 0.64pu after a defect, is maintained at 0.77pu with the SVC and PI controller, becomes 0.84pu with the STATCOM PI controller, and increases to 0.85pu with the suggested fuzzy controller. Similarly, STATCOM with the fuzzy controller effectively maintains active power during faults compared to other compensators. Reactive power during faults demands approximately 35VAR power without a compensator, whereas with SVC it delivers 312VAR reactive power, -470 VAR with STATCOM and PI controller, and with fuzzy-based STATCOM it delivers approximately 485VAR reactive power. Figures 7-9 show how STATCOM executes superior compensation to restore the system voltage to normal. In this instance, voltage recovery cannot be achieved with the installed capacity of the SVC.

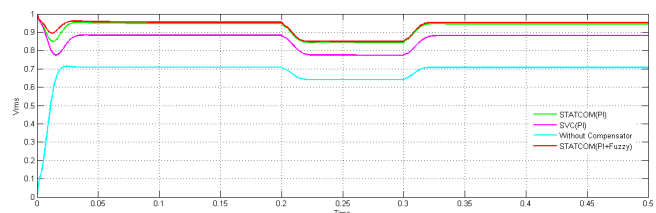


Fig. 7. Vrms during voltage sag.

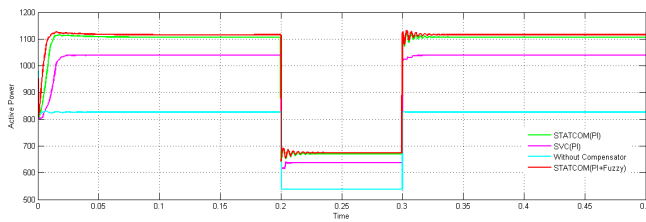


Fig. 8. Active power during voltage sag.

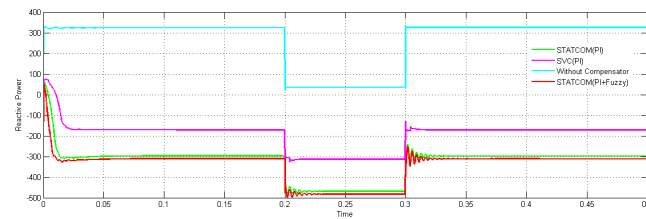


Fig. 9. Reactive power during voltage sag.

TABLE II. PERFORMANCE PARAMETERS (VOLTAGE SAG)

Performance parameter		Without comp.	SVC	STATCOM (PI)	STATCOM (PI+fuzzy)
Vrms (PU)	I	0.71	0.88	0.945	0.955
	II	0.64	0.775	0.84	0.85
	III	0.72	0.88	0.945	0.955
Active power (W)	I	825	1040	1100	1116
	II	538	637	670	675
	III	825	1040	1100	1116
Reactive power (VAR)	I	325	-170	-300	-310
	II	35	-312	-470	-485
	III	325	170	-300	-310
% THD		485.89	417.52	329.02	320.28

(I-prefault, II-during fault, III-post fault)

### B. Analysis of Voltage Swell

The MTL load system consisted of a parallel load with 500MW active power, 100VAR inductive reactive power, and 100VAR capacitive reactive power to produce voltage swell. Figure 10 shows the three-phase circuit breaker to produce the Voltage swell condition. Table III shows the proposed system including FLC-based STATCOM, STATCOM, and SVC with PI controller. Figures 11-13 show the analysis of the test system with and without compensation. The impact of various compensators was examined taking into account a momentary voltage swell lasting 0.2-0.3s.

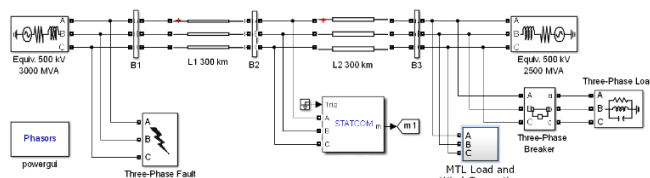


Fig. 10. Proposed system with Voltage swell.

Table III shows that when there is a sudden loss of load at load bus B3, the electrical mode's stability drops and becomes unstable in the absence of STATCOM. Furthermore, the RMS voltage under fault conditions without a compensator is

0.708pu and improves with SVC to 0.785, whereas it is 0.945 with STATCOM and reaches 0.95 with the suggested STATCOM with fuzzy, as it gradually increases towards the rated value. The active and reactive power during the same period without a compensator is 826W and 325VAR, but with SVC and STATCOM, the compensator delivers approximately 1039W and -171VAR, and 1106W and -297VAR, respectively. Using fuzzy-based STATCOM produced around -230VAR of reactive power and 1050W of active power. Similar results for the THD percentage demonstrated improved results compared to the results with and without compensation.

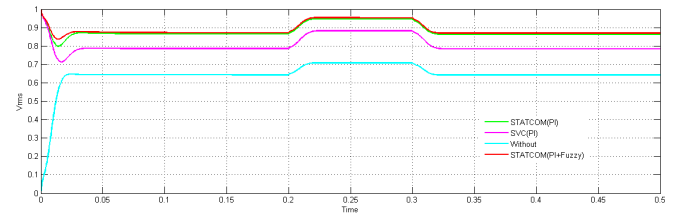


Fig. 11. Vrms during voltage swell.

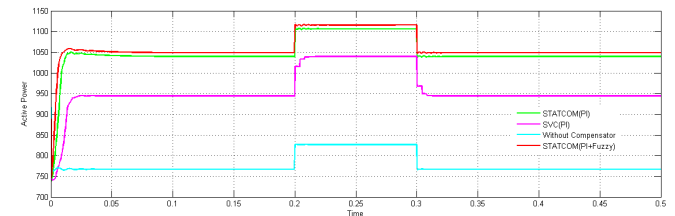


Fig. 12. Active power during voltage swell.

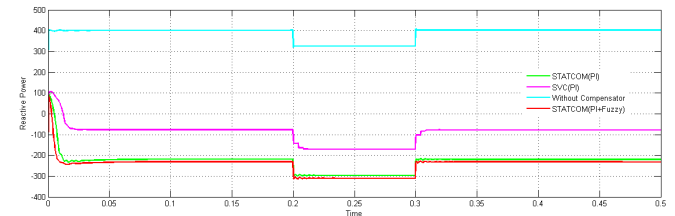


Fig. 13. Reactive power during voltage swell.

TABLE III. VOLTAGE SWELL PERFORMANCE PARAMETERS

Performance parameter		Without comp.	SVC	STATCOM (PI)	STATCOM (PI+fuzzy)
Vrms (PU)	I	0.643	0.785	0.865	0.87
	II	0.708	0.882	0.945	0.95
	III	0.644	0.784	0.864	0.87
Active power (W)	I	767	944	1040	1050
	II	826	1039	1106	1116
	III	766	944	1040	1050
Reactive power (VAR)	I	400	-77	-219	-231
	II	325	-171	-297	-310
	III	400	-78	-220	-230
% THD		485.72	449.14	407.95	401.96

(I-prefault, II-during fault, III-post fault)

### C. Analysis of Capacitor Bank Switching at Load Bus

Load voltage distortion and voltage flickers are the most frequent problems in a power system. Capacitor banks were

connected to B3 through a switch to imitate voltage distortion or flickers, as shown in Figure 14. Capacitor banks were used as shunt compensators with the load bus to compensate for reactive var power on the load bus. Load voltage distortion occurs at the load bus whenever capacitor banks are switched on or off. The simulations in this section ran in transient mode, with the capacitor switched at load bus B3. When a capacitor is turned on during steady-state operation or when load compensation is required, sudden voltage flickers or voltage distortion are introduced at the bus, which is analyzed at PCC with and without a compensator, as well as with the proposed fuzzy-based STATCOM.

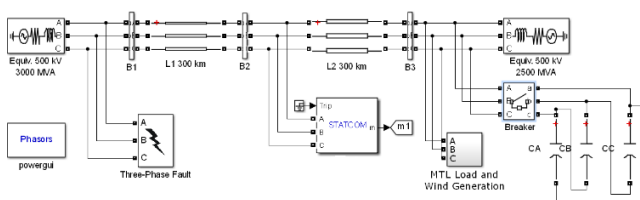


Fig. 14. Proposed system with capacitor switching.

Table IV shows the performance of various compensators under different transient conditions for all transmission line performance parameters. Figures 15-17 show that the Vrms voltage at the PCC increased up to 0.975, closer to the rated value, using the proposed fuzzy-based hybrid STATCOM. Similarly, simulations of other performance metrics, such as active and reactive power, demonstrated that the proposed fuzzy-based STATCOM was capable of producing better active and reactive power than the PI-based SVC and STATCOM. The results reveal that the proposed controller improved THD by 10 to 15%.

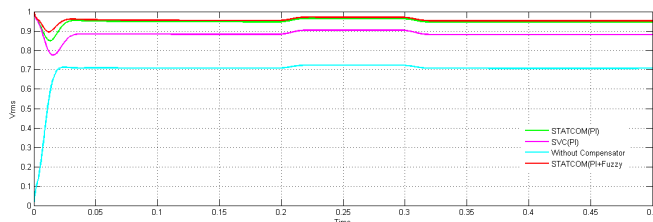


Fig. 15. Vrms with capacitor switching.

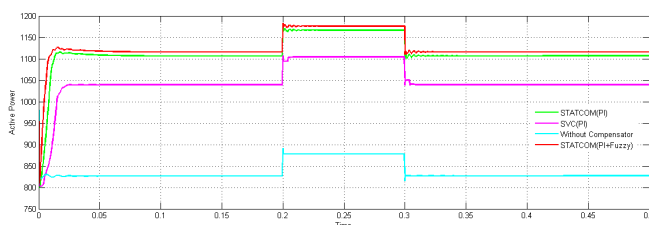


Fig. 16. Active power with capacitor switching.

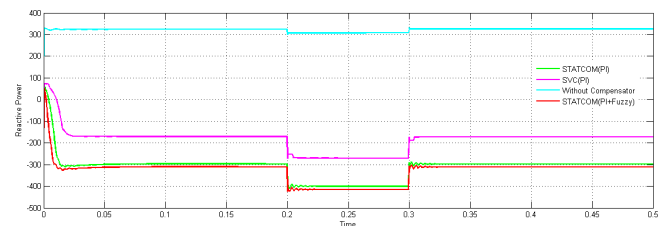


Fig. 17. Reactive power with capacitor switching.

TABLE IV. PERFORMANCE PARAMETERS (CAP. SWITCHING)

Performance parameter		Without comp.	SVC	STATCOM (PI)	STATCOM PI+fuzzy
Vrms (PU)	I	0.645	0.884	0.945	0.955
	II	0.71	0.903	0.964	0.975
	III	0.644	0.883	0.944	0.954
Active power (W)	I	767	1040	1106	1116
	II	846	1104	1166	1176
	III	766	1139	1107	1117
Reactive power (VAR)	I	400	-170	-300	-310
	II	325	-270	-400	-415
	III	401	-171	-301	-310
% THD		485.86	420.56	326.86	310.16

(I-prefault, II-during fault, III-post fault)

#### IV. DISCUSSION

This study investigated a novel hybrid fuzzy controller for AC voltage control of STATCOM. Many studies used either PI controllers or replaced PI with fuzzy controllers. This study used both PI and fuzzy controllers, introducing a novel hybrid fuzzy controller. Other studies showed that the proposed hybrid controller can be easily implemented with simple fuzzy rules and improve system voltage stability and power flow. Similarly, other studies presented techniques based on a SMIB system with an ideal nonlinear load. This study used two dynamic sources connected on either side with a multiterminal load considering the connected wind farm. Different power quality issues were presented, showing that the proposed hybrid controller improves RMS voltages in every power quality issue near the rated values. Furthermore, there is a 10% to 15% improvement in active and reactive power. On other hand, results in Tables II, III, and IV showed 10% to 20 % improvements in THD.

#### V. CONCLUSION

This study investigated the modeling of a system with and without a compensator in several transient situations, revealing power quality concerns in MATLAB Simulink. Simulation results showed that STATCOM and SVC with a PI controller both efficiently alleviated voltage sag/swell, but their efficacy varied. The results obtained by simulating the system with a fuzzy-based hybrid controller with STATCOM were superior. In comparison to SVC, STATCOM with PI, and the system without a compensator, the Vrms during voltage sag, swell, and capacitor bank switching in the suggested STATCOM was substantially more stable. Moreover, the suggested STATCOM enhanced the system voltage at the point of common coupling. STATCOM quickly restores the system voltage, which is

disrupted during voltage sag and swell. Moreover, harmonics are produced as a result of resonance, needing a longer compensation time than in the case of STATCOM. When a quick change occurs, such as the loss of a high load or the switching of capacitor banks, the proposed STATCOM produced a remarkably stable voltage profile with minimum transient overshoots. Furthermore, the recommended STATCOM exhibited lower distortions in terms of load voltage, line active power, and reactive power than earlier compensators.

## REFERENCES

- [1] O. P. Bharti, R. K. Saket, and S. K. Nagar, "Controller Design of DFIG Based Wind Turbine by Using Evolutionary Soft Computational Techniques," *Engineering, Technology & Applied Science Research*, vol. 7, no. 3, pp. 1732–1736, Jun. 2017, <https://doi.org/10.48084/etasr.1231>.
- [2] A. N. Deshmukh and V. K. Chandrakar, "Power quality issues and their mitigation techniques in grid tied Solar Photovoltaic Systems-A review," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, Jan. 2021, pp. 1–6, <https://doi.org/10.1109/ICCCI50826.2021.9402581>.
- [3] A. B. Lajimi, S. A. Gholamian, and M. Shahabi, "Modeling and Control of a DFIG-Based Wind Turbine During a Grid Voltage Drop," *Engineering, Technology & Applied Science Research*, vol. 1, no. 5, pp. 121–125, Oct. 2011, <https://doi.org/10.48084/etasr.60>.
- [4] S. K. Rani and S. Prabakaran, "ANN Based DC Link Control of STATCOM in Wind Integrated Distribution System for Power Quality Conditioning," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5896–5902, Aug. 2020, <https://doi.org/10.48084/etasr.3630>.
- [5] J. Kaniewski, Z. Fedyczak, and G. Benysek, "AC Voltage Sag/Swell Compensator Based on Three-Phase Hybrid Transformer With Buck-Boost Matrix-Reactance Chopper," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 8, pp. 3835–3846, Dec. 2014, <https://doi.org/10.1109/TIE.2013.2288202>.
- [6] M. Moghbel, M. A. S. Masoum, A. Fereidouni, and S. Deilami, "Optimal Sizing, Siting and Operation of Custom Power Devices With STATCOM and APLC Functions for Real-Time Reactive Power and Network Voltage Quality Control of Smart Grid," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5564–5575, Aug. 2018, <https://doi.org/10.1109/TSG.2017.2690681>.
- [7] R. Okuyama, N. Morishima, Y. Ashizaki, and Y. Itaya, "Static synchronous compensator to stabilize grid voltage for wind and photovoltaic power plant," in *2018 International Power Electronics Conference (IPEC-Niigata 2018 -ECCE Asia)*, Feb. 2018, pp. 3450–3454, <https://doi.org/10.23919/IPEC.2018.8507648>.
- [8] M. M. N. Akhtar and V. K. Chandrakar, "Security Enhancement of Hybrid Source in the Grid with STATCOM," in *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, Jun. 2019, pp. 49–54, <https://doi.org/10.1109/ICECA.2019.8821798>.
- [9] V. T. Kullarkar and V. K. Chandrakar, "Power quality improvement in power system by using static synchronous series compensator," in *2017 2nd International Conference for Convergence in Technology (I2CT)*, Mumbai, India, Apr. 2017, pp. 1035–1040, <https://doi.org/10.1109/I2CT.2017.8226286>.
- [10] S. Rahmani, A. Hamadi, N. Mendalek, and K. Al-Haddad, "A New Control Technique for Three-Phase Shunt Hybrid Power Filter," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 8, pp. 2904–2915, Dec. 2009, <https://doi.org/10.1109/TIE.2008.2010829>.
- [11] 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>.
- [12] D. M. Holey and V. K. Chandrakar, "Harmonic Domain Modelling of Space Vector Based STATCOM," *Energy and Power Engineering*, vol. 8, no. 4, 2016, Art. no. 195, <https://doi.org/10.4236/epe.2016.84018>.
- [13] L. Wang, C.-S. Lam, and M.-C. Wong, "A Hybrid-STATCOM With Wide Compensation Range and Low DC-Link Voltage," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 6, pp. 3333–3343, Jun. 2016, <https://doi.org/10.1109/TIE.2016.2523922>.
- [14] R. Jadeja, S. Patel, and S. Chauhan, "STATCOM – A Preface to Power Quality in Power Systems Performance," *Engineering, Technology & Applied Science Research*, vol. 6, no. 1, pp. 895–905, Feb. 2016, <https://doi.org/10.48084/etasr.603>.
- [15] L. Ji *et al.*, "A Multi-Objective Control Strategy for Three Phase Grid-Connected Inverter During Unbalanced Voltage Sag," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2490–2500, Dec. 2021, <https://doi.org/10.1109/TPWRD.2020.3025158>.
- [16] A. Kumar and P. Kumar, "Power Quality Improvement for Grid-connected PV System Based on Distribution Static Compensator with Fuzzy Logic Controller and UVT/ADALINE-based Least Mean Square Controller," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 6, pp. 1289–1299, Nov. 2021, <https://doi.org/10.35833/MPCE.2021.000285>.
- [17] L. K. Haw, M. S. A. Dahidah, and H. A. F. Almurib, "A New Reactive Current Reference Algorithm for the STATCOM System Based on Cascaded Multilevel Inverters," *IEEE Transactions on Power Electronics*, vol. 30, no. 7, pp. 3577–3588, Jul. 2015, <https://doi.org/10.1109/TPEL.2014.2341318>.
- [18] J. He, Y. W. Li, and F. Blaabjerg, "Flexible Microgrid Power Quality Enhancement Using Adaptive Hybrid Voltage and Current Controller," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 6, pp. 2784–2794, Jun. 2014, <https://doi.org/10.1109/TIE.2013.2276774>.
- [19] G. Shahgholian, E. Mardani, and A. Fattollahi, "Impact of PSS and STATCOM Devices to the Dynamic Performance of a Multi-Machine Power System," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2113–2117, Dec. 2017, <https://doi.org/10.48084/etasr.1381>.
- [20] S. Belwanshi, V. Chandrakar, and S. Dhurvey, "Performance Evaluation of IPFC by Using Fuzzy Logic Based Controller for Damping of Power System Oscillations," in *2011 Fourth International Conference on Emerging Trends in Engineering & Technology*, Port Louis, Mauritius, Aug. 2011, pp. 168–173, <https://doi.org/10.1109/ICETET.2011.43>.
- [21] P. Jagtap and V. Chandrakar, "Comparative Study of UPFC Controllers to Improve Transient and Dynamic Stability of Power System," in *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*, Kuala Lumpur, Malaysia, Sep. 2021, pp. 1–7, <https://doi.org/10.1109/GUCON50781.2021.9573804>.

# Design Coordination of a Fuzzy-based Unified Power Flow Controller with Hybrid Energy Storage for Enriching Power System Dynamics

**Vedashree P. Rajderkar**

Department of Electrical Engineering, G H Raisoni College of Engineering, India  
vedashree.rajderkar@raisoni.net  
(corresponding author)

**Vinod K. Chandrakar**

Department of Electrical Engineering, G H Raisoni College of Engineering, India  
vinod.chandrakar@raisoni.net

*Received: 21 November 2022 | Revised: 5 December 2022 | Accepted: 7 December 2022*

## ABSTRACT

In power networks, operation planning has become necessary due to the continuously increased load demand. The Unified Power Flow Controller (UPFC) controls the power flow through the network and improves the dynamics of the power system. In this paper, the design application of the coordination of a fuzzy-based UPFC with hybrid energy storage has been proposed to enrich the system's dynamic performance during abnormal conditions and improve the hybrid energy storage performance during transients when the DC link voltage is insufficient to meet the converter voltage demand. The hybrid energy storage consists of a supercapacitor and a battery connected across the UPFC Vdc shunt capacitor. The proposed system to improve the UPFC performance during large disturbances on a multi-machine system was designed and tested in MATLAB/SIMULINK.

**Keywords:** fuzzy logic; UPFC; hybrid energy storage; super capacitor; battery; dynamic stability

## I. INTRODUCTION

In modern power networks, the interconnected transmissions have to limit the challenges on stability, security, and control operation of the power system in order to operate the transmission network below its thermal capability. When transferring power at a high voltage level, a power system's transient stability refers to its capacity to keep equipment running synchronously in the face of significant disturbances [1]. Flexible AC Transmission System (FACTS) controllers are utilized in power transmission systems to increase the stability of the electrical network. The most adaptable FACTS device is the UPFC. The UPFC's main duty is to regulate the flow of both reactive and actual power by providing a voltage to the line. Voltage control, improved transient stability and damping oscillation are the secondary functions of the UPFC [2]. In the UPFC design, the shunt converter's primary job is to generate or absorb active power from the line, much like a shunt compensator. The DC link capacitor can be charged by the shunt converter, which also uses it to power the series converter. Any losses and useful power utilized or supplied by the series branch must be compensated by a shunt branch. If the power balance is not maintained, voltage cannot remain constant. The active power can have a closed channel via the converter but the reactive power cannot due to the associated

DC link capacitor between the two converters. During small transients, the DC bus compensates for the converter losses in the UPFC by charging or discharging operations [3]. The energy in the DC bus cannot be suppressed during large transient oscillations without degrading the DC voltage. In order to reduce oscillations without impacting the effect of DC link voltage, it is desirable to interchange real power with the storage device when using a fully discharged hybrid energy storage system with a back to back DC-DC converter [4-5]. To offer a significant amount of short-term actual power exchange, a shunt with a DC bus capacitor and a hybrid energy storage device are used. The components of a UPFC with hybrid energy storage are coordinated by using fuzzy logic to improve the transient stability of the system. Authors in [1] introduce the pole-shifting controller-based CSC-STATCOM, which is intended to improve the transient stability of a two-area network. In order to boost oscillation damping capabilities, a damping stabilizer based on PSS is being used. However, they are unable to confirm how well the fuzzy-based UPFC with hybrid energy storage will operate in a dynamic environment. The way a coordinated UPFC based on fuzzy logic affects the flow of actual power was evaluated in both normal and abnormal circumstances in [2]. The performance of the hybrid energy storage is not a metric in this case. The STATCOM can be used along with battery energy storage with the performance

indices technique for improved power situation when using various power controllers [3]. Due to the additional control freedom, the FACTS/BESS has greater flexibility and damping capabilities than the standard FACTS. However, the use of fuzzy-based hybrid energy storage in conjunction with the UPFC for dynamic performance was not tried. Authors in [4] offer a new approach of decreasing the error by the coupling of a neuro-fuzzy scheme with the PI controller for boosting transient stability. However, the damping capability of the hybrid energy storage was not tested. The first swing stability problem is dampened by the UPFC utilizing a local measurement approach. The use of hybrid energy storage in conjunction with the fuzzy-based UPFC to maintain the necessary DC bus voltage to transfer real power across converters during an interruption was not demonstrated [5]. Authors in [6] present an innovative, best-practice fuzzy PID controller based on UPFC, its gains are increased when the time-weighted absolute error integral is taken into consideration by the PSO-GWO method. Low frequency oscillations in the power system have been reduced with the proposed controller, but the impact of storage with UPFC was not covered. In [2], a fuzzy logic technique was used to increase a power network's transient stability. The unified power flow controller has three synchronized control inputs. However, how well the fuzzy-based UPFC works with storage hasn't been tested. Authors in [7] examined quadrature voltage control and in-phase voltage control as fuzzy logic-based UPFC solutions for enhancing power system transient stability utilizing the energy function method. However, the compatibility of fuzzy-based UPFC with storage was not examined. In order to improve the transient stability for UPFC, the authors in [8] looked towards designing a neuro-fuzzy controller, although the fuzzy-based UPFC with storage was not investigated. For improving transient stability, RBFN-based UPFC performance was reported in [9], but UPFC with storage was not considered. For improving transient stability, RBFN Based UPFC Performance was reported in [10]. Many researchers presented combinations of fuzzy-based UPFCs [11-13]. However, they did not focus on the effects of the hybrid energy storage during abnormal conditions [14-16]. The above mentioned authors did not covered in depth the hybrid energy storage, which consists of a supercapacitor with lithium-ion batteries connected across the Vdc of the UPFC, during large sudden disturbances in a power network. Hybrid energy enhances the UPFC primary function of power flow and the secondary function of oscillation damping. The PI-based UPFC has limitations under various system conditions in large nonlinear systems. The fuzzy-based UPFC replaced the PI with significant improvement of performance.

In this study, we present the design of a fuzzy inference for series and shunt Voltage Source Converter (VSC)-based UPFC controller to improve the transient performance of UPFC, which is comparable with the PI-based UPFC. The transient performance of the UPFC was enhanced during sudden large disturbances in the power network by connecting hybrid storage across Vdc. According to the simulation results, the coordination of the fuzzy-based UPFC combined with the hybrid energy storage will function better in enhancing the dynamics under significant disturbances in the network. The

improved UPFC performance has led to an improvement in system performance. MATLAB was used to test the proposed fuzzy-based UPFC design with hybrid energy storage in a multi-machine scenario.

## II. THE PROPOSED MODEL

The power network used in this paper is a multi-machine system with UPFC in combination with hybrid energy storage during abnormal condition as shown in Figure 1. Generator G1 represents dynamic sources, while Generators G2 and G3 provide static sources. Two six-pulse converters linked to a common capacitor make up the power injection model of the UPFC. The UPFC shunt converter regulates the useful power insertion from the shunt converter to the power network by controlling the DC bus voltage across the capacitor. In contrast, a VSC that is connected in series can inject a voltage at the line frequency that has a controlled magnitude and phase angle. As a result, the VSC connected in series receives a very less amount of useful power from the transmission system to change the injected series voltage. A back-to-back dc-dc converter is used to create a shunt connection between the DC bus of the UPFC device and the hybrid energy storage, as shown in Figure 1. In order to enhance the dynamic stability of the power network, the UPFC device often generates or absorbs reactive power. When the DC bus capacitor is unable to support the UPFC during a transient circumstance, the hybrid energy storage offers its energy storage to the UPFC [17, 18]. The converters' ability to exchange real power is required anytime. When there are disturbances, it can maintain a balance in the power transmission between generation and demand without violating the stability limit. While the UPFC-hybrid energy storage is connected between bus B1 and B2, a 3-phase short circuit fault occurs near bus B1 for durations of 0.3s and 0.2s.

The following equation describes the generator dynamics:

$$M \frac{d^2 \delta}{dt^2} = P_m - P_e \quad (1)$$

where  $P_m$  is turbine power,  $P_e$  is output electrical power,  $M$  is the inertia constant, and  $\delta$  is the torque angle.

The fundamental  $P$  and  $Q$  equations are:

$$P = \frac{V_t E_f}{x_d} \sin \delta + \frac{V_t^2}{2} \left[ \frac{1}{x_q} - \frac{1}{x_d} \right] \sin 2\delta \quad (2)$$

$$Q = \frac{E_f V_t}{x_d} \cos \delta + V_t^2 \left[ \frac{1}{x_q} - \frac{1}{x_d} \right] \sin^2 \delta - \frac{V_t^2}{x_d} \quad (3)$$

The UPFC equations are:

$$P = \frac{V_i V_j}{x} \sin \delta + \frac{V_{se} V_j}{x} \sin(\delta + \rho) \quad (4)$$

$$Q = \frac{V_i V_j}{x} (\cos \delta - 1) + \frac{V_{se} V_j}{x} \cos(\delta + \rho) \quad (5)$$

when  $\rho = 90^\circ - \delta$ .

The hybrid energy storage consisting of a super capacitor and a battery is demonstrated in Figure 2.



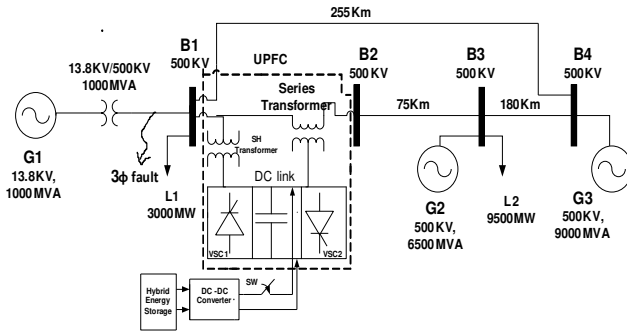


Fig. 1. System model with combination of UPFC and hybrid energy storage.

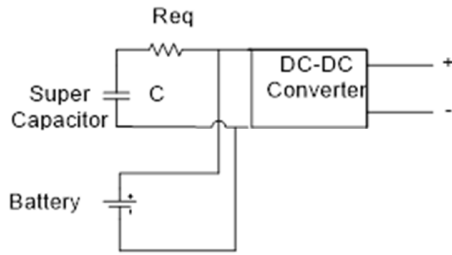


Fig. 2. Hybrid energy storage.

The supercapacitor output voltage is expressed as:

$$V_{SC} = \frac{N_s Q_T d}{N_p N_e \epsilon \epsilon_0 A_i} + \frac{2 N_e N_s R T}{F} \sinh^{-1} \left( \frac{Q_T}{N_p N_e^2 A_i \sqrt{8 R T \epsilon \epsilon_0 c}} \right) - R_{SC} \cdot i_{SC} \quad (6)$$

where  $Q_T$  is the electric charge,  $N_s$  is the number of series supercapacitors,  $N_p$  the number of parallel supercapacitors,  $R$  the ideal gas constant,  $T$  the operating temperature,  $d$  the molecular radius,  $N_e$  the number of layers of electrodes,  $A_i$  the interfacial area between the electrodes and the electrolyte,  $c$  the molar concentration, and  $F$  the Faraday's constant.

The super capacitor energy is directly proportional to the DC voltage and its capacitance is [17]:

$$E_{SC} = \frac{1}{2} C V_{SC}^2 \quad (7)$$

The maximum power of the super capacitor is:

$$P_{max} = \frac{V_{SC}^2}{4 R_{eq}} \quad (8)$$

where  $R_{eq}$  is the equivalent resistance.

The stored energy in a battery depends on the stored charge and its voltage. The operating power of a battery is represented by:

$$P = VI = (E - IR) I = EI - I^2 R \quad (9)$$

The following equations are used in the lithium-ion battery model:

a) Model for Discharge ( $i^* > 0$ )

$$f_1(it, i^*, i) = E_0 - K \cdot \frac{Q}{Q - it} \cdot i^* - K \cdot \frac{Q}{Q - it} \cdot it + A \cdot \exp(-B \cdot it) \quad (10)$$

b) Model for Charge ( $i^* < 0$ )

$$f_2(it, i^*, i) = E_0 - K \cdot \frac{Q}{Q - it} \cdot i^* - K \cdot \frac{Q}{Q - it} \cdot it + A \cdot \exp(-B \cdot it) \quad (11)$$

where  $A$  is the exponential voltage,  $B$  is the exponential capacity,  $E_0$  is the constant voltage,  $K$  is the polarization constant,  $i^*$  is the low-frequency current dynamics,  $i$  is the battery current, and  $Q$  is the maximum battery capacity.

The hybrid energy storage voltage and the steady state DC bus voltage are related through the duty cycle  $D$ , which can be modified in such a way that it will help damp power oscillations as:

$$V_{dc} = \frac{D}{1-D} V_{HES} \quad (12)$$

#### A. Fuzzy Logic-based UPFC

In comparison to the traditional control strategy, the advantage of the fuzzy logic technique is that it does not need exact quantitative values for the control inputs and system variables. Fuzzy logic controllers are more adaptable and work without the need for a mathematical description of the power network. The fuzzy logic approach is a good and effective method for regulating a UPFC.

#### B. Fuzzy-based Shunt Voltage Regulator of the VSC

The input to the fuzzy logic controller consists of  $V_{ref}$  and  $V_{mes}$ . The process of fuzzification entails mapping the input variables onto linguistically fuzzy variables. There is a distinct membership function for each fuzzy variable. The inputs are fuzzy-set using 3 trapezoid-shaped membership functions. The voltage regulator is achieved by adjusting the difference of  $V_{ref}$  and  $V_{mes}$  which is the input to the fuzzy logic as shown in Figure 3. On the basis of fuzzed linguistic features, control judgments are made. Inference entails guidelines for selecting which outputs to employ. The input variables of the fuzzy coordination controller of the shunt voltage regulator consist of 3 fuzzy variables and 9 fuzzy rules to control the shunt current.

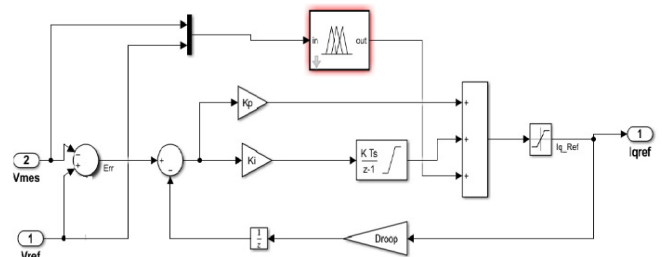


Fig. 3. Fuzzy-based shunt voltage regulator of VSC.

Table I shows the input output mapping of the fuzzy rule-based system. The Mamdani inference system produces linguistic variables as its output. They need to be quantitatively translated into the output. The fuzzy controller employs the centroid approach.

#### C. Fuzzy-based Series Current Regulator of the VSC

Using a current regulator, the line currents are forced to match the corresponding reference values in order to control the useful and reactive power flow in the line. Each fuzzy variable has a different membership function. For input and

output, trapezoidal membership functions were chosen. As shown in Figure 4, the fuzzy control produces the series injected voltage in collaboration with 2 separate PI controllers. A collection of basic membership functions is produced using the current. Nine concepts are used by the fuzzy coordination controller to control series voltage. A time-based pulse in a model of all-systems has a defuzzifier as its final state.

TABLE I. INPUT-OUTPUT MAPPING OF THE SHUNT VSC

$V_{ref}$	$S_{Vm}$	$M_{Vm}$	$L_{Vm}$
$S_{Vref}$	$S_{O/P}$	$M_{O/P}$	$L_{O/P}$
$M_{Vref}$	$M_{O/P}$	$M_{O/P}$	$L_{O/P}$
$L_{Vref}$	$L_{O/P}$	$L_{O/P}$	$L_{O/P}$

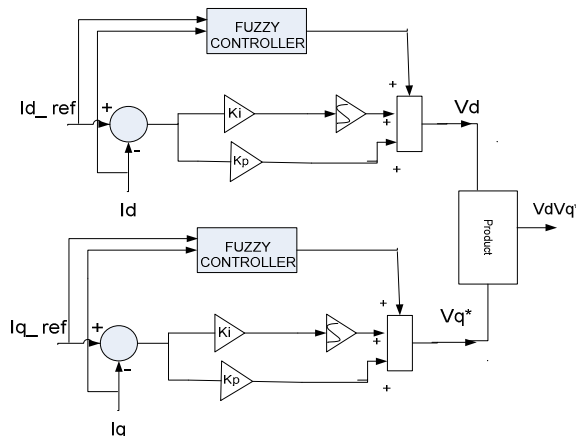


Fig. 4. Series current regulator of VSC with fuzzy logic.

The design of the series current regulator for the fuzzy inference system of the VSC is shown in Tables II and III.

TABLE II. SERIES CURRENT REGULATOR SIGNAL ID INFERENCE SYSTEM COORDINATION FOR VSC

$I_d$	$S_{Id}$	$M_{Id}$	$L_{Id}$
$S_{Idref}$	$S_{O/P}$	$M_{O/P}$	$L_{O/P}$
$M_{Idref}$	$M_{O/P}$	$M_{O/P}$	$L_{O/P}$
$L_{Idref}$	$L_{O/P}$	$L_{O/P}$	$L_{O/P}$

TABLE III. COORDINATED SERIES CURRENT REGULATOR SIGNAL IQ OF THE VSC INFERENCE SYSTEM

$I_q$	$S_{Iq}$	$M_{Iq}$	$L_{Iq}$
$S_{Iqref}$	$S_{O/P}$	$M_{O/P}$	$L_{O/P}$
$M_{Iqref}$	$M_{O/P}$	$M_{O/P}$	$L_{O/P}$
$L_{Iqref}$	$L_{O/P}$	$L_{O/P}$	$L_{O/P}$

### III. SIMULATION RESULTS

The digital simulation of the hybrid energy storage with fuzzy logic-based UPFC was carried out in a multi-machine system in MATLAB when a 3-phase fault occurs at the sending end of the transmission lines near bus B1. The aim is to enhance the power system dynamic performance as well as to check the secondary function of UPFC to damp out the oscillations during abnormal conditions. The design coordinated action of the fuzzy-based UPFC with hybrid

energy storage improves the damping performance of the system and reduces the dynamics.

Figure 5 demonstrates how the fuzzy-based UPFC with hybrid energy storage reduces the power fluctuations in the rotor speed of the generator G1 in all the considered scenarios during fault durations. The system with the fuzzy-based UPFC and the hybrid energy storage is superior than without UPFC, with PI-UPFC, and with fuzzy logic-based UPFC. It significantly improves the dynamic performance of the system. The proposed fuzzy-based UPFC with hybrid energy storage was tested in a multi-machine system for fault durations of 0.2 and 0.3s. The terminal voltage of the generator G1 is shown in Figure 6. The results indicate that the proposed fuzzy controller has satisfactory operation over the PI-based controller and gets better enrichment in the dynamic performance of the system. The fuzzy-based UPFC with hybrid energy storage stabilizes the generator G1 terminal voltage waveform in minimum number of cycles.

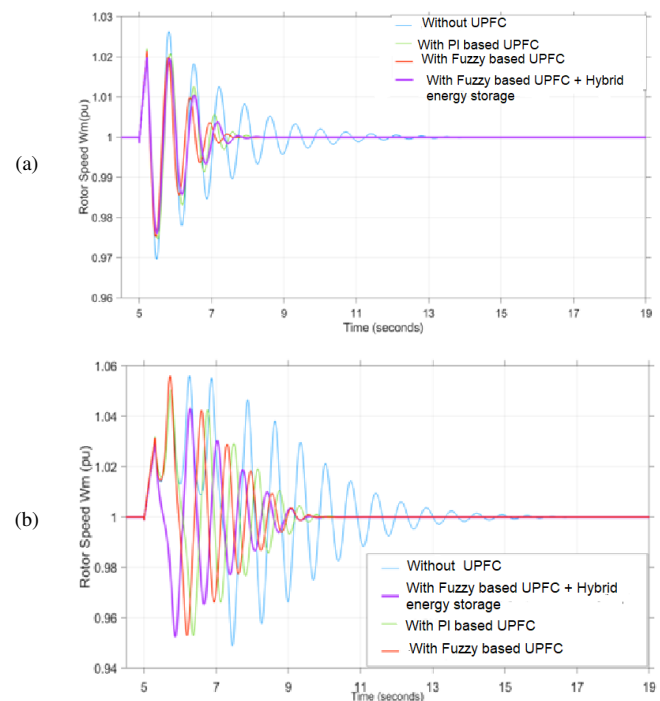


Fig. 5. Speed of rotor at G1 without UPFC, with UPFC based on PI, with fuzzy logic-based UPFC, and with fuzzy logic-based UPFC and hybrid energy storage during (a) a 0.2s fault and (b) a 0.3s fault.

Figure 7 shows that the initial swing is suppressed under post fault conditions. The flow of real power with the fuzzy-based UPFC with hybrid energy storage indicates that it increases the power handling capability of the UPFC during transient conditions when the DC link capacitor is inadequate to exchange the real power between the converters.

Figure 8 indicates that the fuzzy-based UPFC with hybrid energy storage provides reactive power support during large disturbances to maintain the DC link voltage between the converters and also helps maintain voltage profile after clearing the fault.

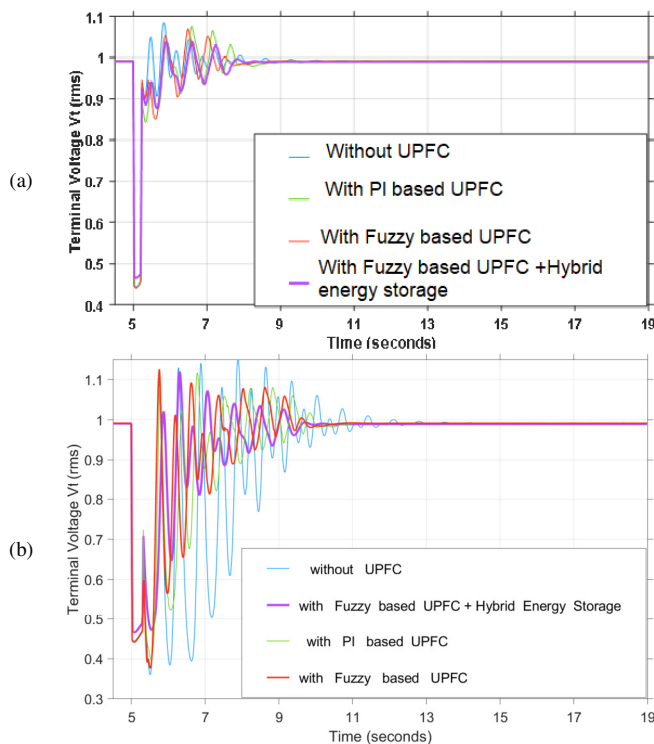


Fig. 6. G1 terminal voltage without UPFC, with UPFC based on PI, with fuzzy logic-based UPFC, and with fuzzy logic-based UPFC and hybrid energy storage during (a) a 0.2s fault and (b) a 0.3s fault.

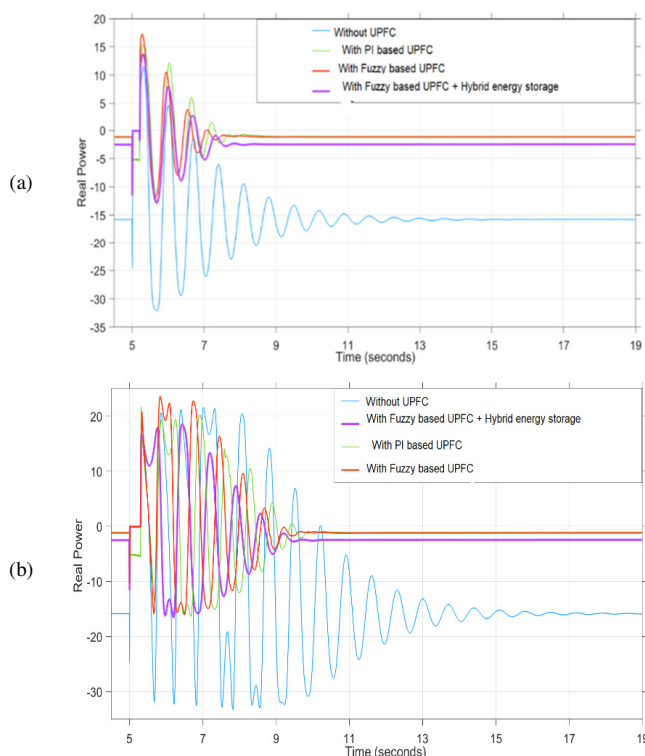


Fig. 7. Bus B2 real power voltage without UPFC, with UPFC based on PI, with fuzzy logic-based UPFC, and with fuzzy logic-based UPFC and hybrid energy storage during (a) a 0.2s fault and (b) a 0.3s fault

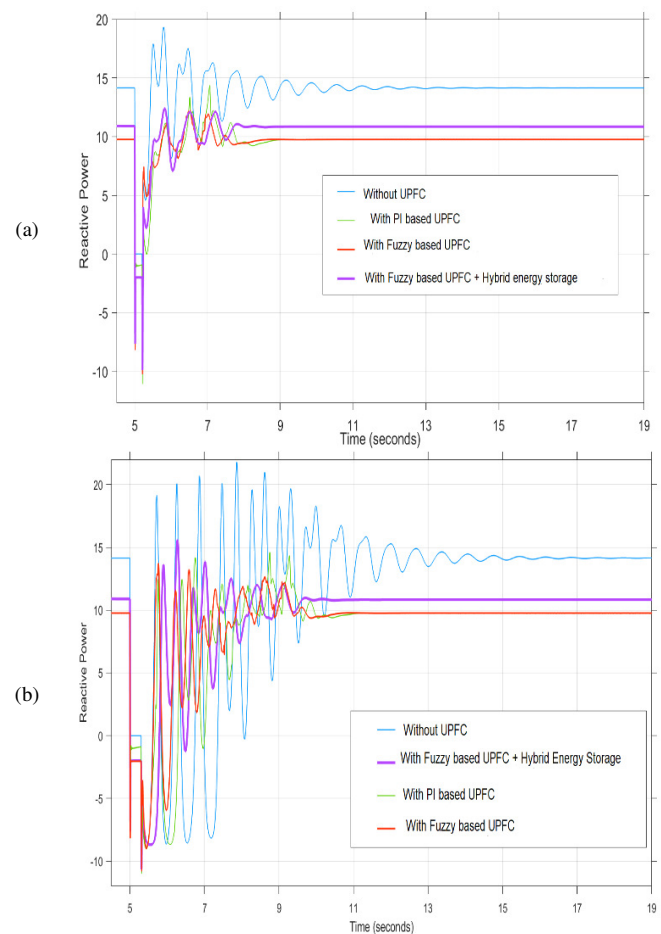


Fig. 8. Bus B2 reactive power voltage without UPFC, with UPFC based on PI, with fuzzy logic-based UPFC, and with fuzzy logic-based UPFC and hybrid energy storage during (a) a 0.2s fault and (b) a 0.3s fault

#### IV. DISCUSSION

Rotor speed damping varies from 13s without UPFC to 7.5s with the fuzzy-based UPFC with hybrid energy storage for the fault duration of 0.2s. Therefore, settling time reduces by 42.3%. Similarly, for 0.3s fault duration, the settling time of the rotor speed without UPFC is 9.5s which reduces to 5.5s with the fuzzy-based UPFC with hybrid energy storage (38.7% reduction). These values indicate the significant dynamic performance of the fuzzy-based UPFC with hybrid energy storage for improving the dynamic stability of the system. The primary function of the UPFC is to regulate the terminal voltage which is enhanced by the fuzzy-based UPFC with hybrid energy storage. The basic function of the UPFC is to control real and reactive power under abnormal condition, something that has been significantly achieved with the proposed controller. The proposed hybrid energy storage with UPFC improves the fundamental performance of the UPFC during sudden large disturbances hence the system dynamic stability is improved.

#### V. CONCLUSION

In this paper, a fuzzy-based UPFC with hybrid energy storage is proposed to enrich the power system dynamics. The

design objectives are to maximize the power flow and to minimize the oscillation damping during large disturbances in a multi-machine system. The simulation results illustrate that the combination of the fuzzy-based UPFC with the hybrid energy storage meets the design objectives and demonstrate the robustness of the fuzzy logic controller over the PI controller. Due to the hybrid energy storage application during large disturbances with duration of 0.2s and 0.3s, the real power support from  $V_{dc}$  to the system increases and helps oscillation damping. The fuzzy inference is designed to coordinate the PI controller with the storage device. The overall performance of the UPFC during disturbance conditions has been significantly improved.

## REFERENCES

- [1] S. Gupta and R. Tripathi, "Two-area power system stability improvement using a robust controller-based CSC-STATCOM," *Acta Polytechnica Hungarica*, vol. 11, no. 7, pp. 135–155, Jan. 2014.
- [2] V. K. Chandrakar and A. G. Kothari, "Fuzzy Logic Based Unified Power Flow Controllers for Improving Transient Stability," *International Journal of Power and Energy Systems* 2008, vol. 28, no. 3, Jan. 2008, <https://doi.org/10.2316/Journal.203.2008.2.203-3746>.
- [3] L. Zhang, C. Shen, M. L. Crow, L. Dong, S. Pekarek, and S. Atcitty, "Performance Indices for the Dynamic Performance of FACTS and FACTS with Energy Storage," *Electric Power Components and Systems*, vol. 33, no. 3, pp. 299–314, Dec. 2004, <https://doi.org/10.1080/15325000590474438>.
- [4] S. Mishra, "Hybrid-Neuro-Fuzzy UPFC for Improving Transient Stability Performance of Power System," *Electric Power Components and Systems*, vol. 33, no. 1, pp. 73–84, Jan. 2005, <https://doi.org/10.1080/15325000590454548>.
- [5] E. Gholipour and S. Saadate, "Improving of transient stability of power systems using UPFC," *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 1677–1682, Apr. 2005, <https://doi.org/10.1109/TPWRD.2005.846354>.
- [6] N. Nahak and R. K. Mallick, "Enhancement of small signal stability of power system using UPFC based damping controller with novel optimized fuzzy PID controller," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 1, pp. 501–512, Jan. 2018, <https://doi.org/10.3233/JIFS-169606>.
- [7] S. Limyingcharoen, U. D. Annakkage, and N. C. Pahalawaththa, "Fuzzy logic based unified power flow controllers for transient stability improvement," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 145, no. 3, pp. 225–232, May 1998, <https://doi.org/10.1049/ip-gtd:19982000>.
- [8] S. Latha and C. K. Rani, "Transient Stability Improvement with Unified Power Flow Controller Using Fuzzy Logic and ANFIS Approach," *Advanced Materials Research*, vol. 768, pp. 378–387, 2013, <https://doi.org/10.4028/www.scientific.net/AMR.768.378>.
- [9] A. N. Shirazi, B. Mozaffari, and S. Soleymani, "Transient stability improvement with neuro-fuzzy control of GUPFC in multi machine system," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 1, pp. 611–623, Jan. 2019, <https://doi.org/10.3233/JIFS-171488>.
- [10] V. K. Chandrakar and A. G. Kothari, "RBFN Based UPFC for Improving Transient Stability Performance," *WSEAS Transactions on Power Systems*, vol. 2, no. 1, 2007.
- [11] K. D. Joshi and V. Chandrakar, "Transient Stability Improvement Using UPFC-SMES in A Multi Machine Power System," *International Journal of Applied Power Engineering (IJAPE)*, vol. 5, no. 1, pp. 14–21, Apr. 2016, <https://doi.org/10.11591/ijape.v5.i1.pp14-21>.
- [12] D. M. Holey and V. K. Chandrakar, "Dynamic Harmonic Domain Modelling of Space Vector Based UPFC," *American Journal of Electrical Power and Energy Systems*, vol. 5, no. 1, Apr. 2016, Art. no. 1, <https://doi.org/10.11648/j.epe.20160501.11>.
- [13] S. N. Dhurvey and V. K. Chandrakar, "Performance comparison of UPFC in co-ordination with optimized POD and PSS on damping of power system oscillations," *WSEAS Transactions on Power Systems*, vol. 3, no. 5, pp. 287–299, 2008.
- [14] G. Shahgholian, E. Mardani, and A. Fattollahi, "Impact of PSS and STATCOM Devices to the Dynamic Performance of a Multi-Machine Power System," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2113–2117, Dec. 2017, <https://doi.org/10.48084/etasr.1381>.
- [15] M. Y. A. Khan, U. Khalil, H. Khan, A. Uddin, and S. Ahmed, "Power Flow Control by Unified Power Flow Controller," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3900–3904, Apr. 2019, <https://doi.org/10.48084/etasr.2587>.
- [16] O. Kahouli, B. Ashammari, K. Sebaa, M. Djebali, and H. H. Abdallah, "Type-2 Fuzzy Logic Controller Based PSS for Large Scale Power Systems Stability," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3380–3386, Oct. 2018, <https://doi.org/10.48084/etasr.2234>.
- [17] S. Kandasamy and A. Rajapalan, "Dynamic control modeling and simulation of a UPFC-SMES compensator in power systems," *Ain Shams Engineering Journal*, vol. 6, no. 4, pp. 1179–1186, Dec. 2015, <https://doi.org/10.1016/j.asej.2015.04.011>.
- [18] V. P. Rajderkar and V. K. Chandrakar, "Enhancement of Power System Security by Fuzzy based Unified Power Flow Controller," in *2022 2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, Jun. 2022, <https://doi.org/10.1109/CONIT55038.2022.9847699>.

# Study of Gamma-ray Shielding of Two Different Heavy Metals and their Combination for Cs-137 and Co-60 Sources

**Mohamed E. M. Eisa**

Department of Physics, Northern Border University, Saudi Arabia | Department of Physics, Sudan University of Science and Technology, Sudan  
memeisa@yahoo.com  
(corresponding author)

**M. D. M. Ali**

Department of Physics, Sudan University of Science and Technology, Sudan  
mohameddaffallah@yahoo.com

**Mustafa J. Abuualreish**

Department of Chemistry, Northern Border University, Saudi Arabia  
mustjeed\_2008@hotmail.com

Received: 23 November 2022 | Revised: 7 December 2022 | Accepted: 12 December 2022

## ABSTRACT

This article presents data collected by measurements of lead (Pb) and iron (Fe) and their combination as heavy shielding materials. Measurements were performed using gamma photon energies of 662, 1173, and 1332keV for the Cs-137 and Co-60 sources. The theoretical data part was calculated using WinXCom, Phy-X, and Py-MLUBF software packages. Tables and graphs of the photon Mass Attenuation Coefficient (MAC), Linear Attenuation Coefficient (LAC), Half Value Layer (HVL), Tenth Value Layer (TVL) and Mean Free Path (MFP) are presented for both heavy metals and their combination to study the shielding properties experimentally and theoretically. The results will contribute to the ongoing research as a database for future use.

**Keywords-linear attenuation coefficient; photon mass attenuation coefficient; radiation resistance data; heavy metals; gamma photon energies; gamma radiation properties**

## I. INTRODUCTION

Radiation protective shields play many functions, the most important of which is to reduce radiation exposure [1]. Radioactivity is common in the rocks and soil, in the water and oceans, and even in our building and home materials [2-4]. With the increasing use of gamma radiation in various applications such as industry, medicine, agriculture, nuclear reactors, and particle accelerators, radiation exposure for longer duration can cause very harmful effects on human health. Therefore, the use of shielding becomes paramount [5]. The interaction with gamma radiation photons takes place in various processes like the photoelectric effect, incoherent scattering, coherent scattering, and pair production and depends upon the energy of the photons [6-7]. Possible exposure to radiation emitted from radionuclides can occur in houses, offices, and other working places. In order to be able to assess radiological hazard, it is important to study the levels of radiation emitted from these building materials [8]. Different

radiation shielding materials have been developed in the past. Ionizing photons such as X-rays and gamma rays can change the chemical structures of molecules, which may result in biological damages, cellular level mutations, or deterioration of materials [9-10]. Protection of biological entities from the harmful effects of radiation exposure is a fundamental requirement in the application of nuclear technology. Radiation exposure can be avoided by methods involving time, distance, and shielding, with shielding being the most important method [11-14].

When a gamma-ray beam passes through a sample of thickness  $x$  (cm) under narrow beam geometry, the photons are transmitted according to Beer-Lambert's equation [2, 15, 16]:

$$I = I_0 e^{-\mu x} \quad (1)$$

where  $I_0$  and  $I$  represent the gamma-ray intensity before and after passing through the sample, respectively, and  $\mu$  ( $\text{cm}^{-1}$ ) is the Linear Attenuation Coefficient (LAC) of the sample.

The LAC can be described in terms of the Mass Attenuation Coefficient (MAC) as follows [10, 17-19]:

$$\mu = \left(\frac{\mu}{\rho}\right) \mu_s \quad (2)$$

where  $\mu_s \left(\text{cm}^2 \text{g}^{-1}\right)$  is the MAC and  $\rho \left(\text{g cm}^{-3}\right)$  is the density of the sample.

The Mean Free Path (MFP) (cm) is:

$$MFP = \frac{1}{\mu} \quad (3)$$

The effectiveness of the shielding capability of a material to a photon can be described by the Half Value Layer (HVL) (cm) [20-21]:

$$HVL = \frac{\ln(2)}{\mu} \quad (4)$$

Similarly, the Tenth Value Layer (TVL) (cm) is given by:

$$TVL = \frac{\ln(10)}{\mu} \quad (5)$$

## II. MATERIALS AND METHODS

Measurements were conducted at the irradiation room of the Secondary Standard Dosimetry Laboratory of the Sudan Atomic Energy Commission (SAEC) on model OB-85 gamma calibrator manufactured by the Buchler GmbH and by using Cs-137 and Co-60 radioactive sources. Experimental measurement Tables II-IV were constructed using a secondary standard ionization chamber which has a volume of  $1000\text{cm}^3$  and displays the radiation dose of the sources and the thicknesses. The ionization chamber was chosen to be the center of a sphere relative to the reference source in a 2m distance from the source to the reference point. The holder of shielding samples was placed close to the OB-85 gamma calibrator. The shielding materials were placed inside the holder one by one. This procedure was followed for all shielding materials [22, 23]. Through the measurements the intensity of radiation was recorded, before and after placing the shielding material in the holder. The experimental coefficients were evaluated using gamma energies of 662keV for Cs-137 and 1173 and 1332KeV for Co-60 and the gamma transmission parameters such as the MAC, HVL, MFP, and transmission factor were determined and calculated.

## III. RESULTS

Density, chemical composition, and concentration of the iron and lead samples and their combination can be seen in Table I. Table II-III present the experimental results of the

attenuation coefficients and HVLs of Fe and Pb slabs using different thicknesses and doses of Cs-137 and Co-60. Table IV presents the experimental results of the combination of Fe and Pb using Cs-137 gamma ray. The theoretical coefficient of the two materials and their combination were determined using Py-MLUBF, Phy-X, and XCOM [24, 26, 28-31] software packages which gave very close results. The results were compared against the experimental findings and gave very good agreement. Tables V-VII present the simulation results of Fe and Pb shielding using Py-MLUBF [32-33], Phy-X and WinXCOM software. The combination of the two shielding materials, Pb and Fe, is shown in Tables VIII and IX. Figures 1 and 2 show the relation between HVL and MAC vs Energy by using Py-MLUBF for Fe. Figures 3 and 4 show the relation between HVL and MAC vs Energy by using Py-MLUBF for Fe and Pb calculated by Py-MLUBF for Cs-137. Figures 8 and 9 show the relation between the gamma ray shielding factor for Fe and Pb calculated by Py-MLUBF for energies of 1173 and 1332KeV, respectively for the Co-60 source [33].

TABLE I. DENSITY AND CHEMICAL CONCENTRATION OF LEAD (PB) AND IRON (FE) SAMPLES

Material	Lead	Iron
Density(g/cm <sup>3</sup> )	11.34	7.87
Concentration(%)	Sn	0.0023
	Sb	0.00051
	Bi	0.0025
	As	0.0019
	Ag	0.00064
	Cd	0.00006
	Zn	0.00022
	Te	0.00026
	Au	0.00096
	In	0.00015
	Na	0.00028
	Ca	0.001
	Pb	99.99
	C	0.062
	Si	0.035
	Mn	0.345
	Cr	0
	Mo	0.022
	Ni	0.00029
	Al	0.00039
	Cu	0
	Ti	0
	V	0
	Nb	0
	Co	0.046
	W	0
	Fe	99.474

TABLE II. EXPERIMENTAL RESULTS FOR FE USING CS-137 AND CO-60

Thickness (cm)	Attenuation coefficients and HVL for Fe slabs using Cs-137 gamma rays with initial dose of 474.76μGy.				Attenuation coefficients and HVL for Fe slabs using Co-60 gamma rays with initial dose of 1.906 μGy.			
	Dose (μGy)	LAC (cm <sup>-1</sup> )	MAC (cm <sup>2</sup> /g)	HVL (cm)	Dose (μGy)	LAC (cm <sup>-1</sup> )	MAC (cm <sup>2</sup> /g)	HVL (cm)
0.202	434.58	0.437	0.055	1.583	1.797	0.291	0.037	2.377
0.522	373.03	0.462	0.058	1.500	1.602	0.332	0.042	2.081
1.036	289.54	0.477	0.060	1.451	1.346	0.335	0.042	2.063
1.350	249.21	0.477	0.060	1.451	1.202	0.341	0.043	2.029



TABLE III. EXPERIMENTAL RESULTS FOR PB USING CS-137 AND CO-60

Attenuation coefficients and HVL for lead slabs using Cs-137 gamma rays with initial dose of 474.76μGy.					Attenuation coefficients and HVL for lead slabs using Co-60 gamma rays with initial dose of 1.906 μGy.			
Thickness (cm)	Dose (μGy)	LAC (cm <sup>-1</sup> )	MAC (cm <sup>2</sup> /g)	HVL (cm)	Dose (μGy)	LAC (cm <sup>-1</sup> )	MAC (cm <sup>2</sup> /g)	HVL (cm)
0.154	400.46	1.105	0.097	0.627	1.744	0.576	0.050	1.211
0.472	274.69	1.159	0.102	0.597	1.428	0.611	0.053	1.132
0.812	190.42	1.125	0.099	0.615	1.169	0.602	0.053	1.151
0.966	164.63	1.096	0.096	0.632	1.079	0.589	0.051	1.176

TABLE IV. EXPERIMENTAL RESULTS FOR FE AND PB USING CS-137 LAC AND HVL FOR LEAD AND IRON SLABS USING CS-137 GAMMA RAYS WITH INITIAL DOSE OF 237.74μGy

Thickness (cm)	Dose (μGy)	Experimental results	
		LAC μ(cm <sup>-1</sup> )	HVL (cm)
0.356	182.72	0.7394	0.937
0.676	157.52	0.608	1.138
0.994	110.09	0.774	0.894
1.014	87.74	0.983	0.705
1.334	76.17	0.853	0.812
1.822	74.86	0.634	1.092
1.848	59.86	0.746	0.928
2.162	51.84	0.704	0.983

TABLE V. RESULTS OF FE BY PHY-X AND PY-MLBUF

Phy-X(Fe)							Py-MLBUF (Fe)				
Energy (MeV)	Source	MAC (cm <sup>2</sup> /g)	LAC (1/cm)	HVL (cm)	TVL (cm)	MFP (cm)	MAC (cm <sup>2</sup> /g)	LAC (1/cm)	HVL (cm)	TVL (cm)	MFP (cm)
6.62E-01	Cs (137)	0.073	0.578	1.199	3.982	1.729	7.35E-02	5.78E-01	1.1983	3.9806	1.730
8.00E-01		0.067	0.527	1.315	4.368	1.897	6.70E-02	5.27E-01	1.3147	4.3675	1.897
8.26E-01	Co (60)	0.066	0.519	1.336	4.437	1.927					
1.00E+00		0.060	0.472	1.469	4.881	2.120	6.00E-02	4.72E-01	1.4691	4.8804	2.118
1.17E+00	Co (60)	0.055	0.435	1.594	5.295	2.300	5.53E-02	4.35E-01	1.5927	5.2907	2.298
1.33E+00	Co (60)	0.052	0.408	1.701	5.650	2.454	5.18E-02	4.08E-01	1.7003	5.6482	2.450
1.50E+00		0.049	0.384	1.804	5.992	2.602	4.88E-02	3.84E-01	1.8037	5.9918	2.604
2.00E+00		0.043	0.336	2.065	6.861	2.980	4.27E-02	3.36E-01	2.0651	6.86	2.976

TABLE VI. RESULTS OF PB BY PHY-X AND PY-MLBUF

Phy-X(Pb)							Py-MLBUF (Pb)				
Energy (MeV)	Source	MAC (cm <sup>2</sup> /g)	LAC (1/cm)	HVL (cm)	TVL (cm)	MFP (cm)	MAC (cm <sup>2</sup> /g)	LAC (1/cm)	HVL (cm)	TVL (cm)	MFP (cm)
6.62E-01	Cs 137	1.10E-01	1.25E+00	0.5557	1.8459	0.800	1.10E-01	0.110	1.250	0.555	1.843
8.00E-01		8.87E-02	1.01E+00	0.6891	2.2892	0.990	8.87E-02	0.089	1.006	0.689	2.289
1.00E+00		7.10E-02	8.05E-01	0.8607	2.8591	1.2422	7.10E-02	0.071	0.805	0.861	2.859
1.17E+00	Co (60)	6.18E-02	7.01E-01	0.9891	3.2856	1.426	6.18E-02	0.062	0.700	0.990	3.288
1.33E+00	Co (60)	5.61E-02	6.36E-01	1.0896	3.6194	1.572	5.61E-02	0.056	0.636	1.089	3.618
1.50E+00		5.22E-02	5.92E-01	1.1705	3.8884	1.689	5.22E-02	0.052	0.592	1.170	3.888
2.00E+00		4.61E-02	5.22E-01	1.3268	4.4074	1.915	4.61E-02	0.046	0.522	1.327	4.408

TABLE VII. RESULTS OF FE BY WINXCOM

WinXCOM (Fe)					
Photon energy (MeV)	Scattering		Photoelectric absorption (cm <sup>2</sup> /g)	Total attenuation	
	Coherent (cm <sup>2</sup> /g)	Incoherent (cm <sup>2</sup> /g)		With coherent scattering (cm <sup>2</sup> /g)	Without coherent scattering (cm <sup>2</sup> /g)
6.620E-01	9.955E-04	7.159E-02	8.713E-04	7.346E-02	7.246E-02
8.000E-01	6.833E-04	6.575E-02	5.650E-04	6.699E-02	6.631E-02
1.000E+00	4.381E-04	5.916E-02	3.514E-04	5.995E-02	5.951E-02
1.022E+00	4.196E-04	5.853E-02	3.335E-04	5.928E-02	5.887E-02
1.173E+00	3.188E-04	5.466E-02	2.524E-04	5.526E-02	5.494E-02
1.250E+00	2.808E-04	5.292E-02	2.256E-04	5.350E-02	5.322E-02
1.332E+00	2.473E-04	5.122E-02	2.014E-04	5.181E-02	5.156E-02
1.500E+00	1.951E-04	4.811E-02	1.627E-04	4.883E-02	4.864E-02
2.000E+00	1.099E-04	4.107E-02	1.003E-04	4.265E-02	4.254E-02
2.000E+00	1.099E-04	4.107E-02	1.003E-04	4.265E-02	4.254E-02

TABLE VIII. RESULTS OF THE COMBINATION OF THE TWO SHIELDING MATERIALS

Fe+Pb combination, Py-MLUBF results							Fe+Pb combination, Phy-X results				
Energy (MeV)	Source	MAC (cm <sup>2</sup> /g)	LAC (1/cm)	HVL (cm)	TVL (cm)	MFP (cm)	MAC (cm <sup>2</sup> /g)	LAC (1/cm)	HVL (cm)	TVL (cm)	MFP (cm)
6.62E-01	Cs(137)	1.10E-01	1.25E+00	0.555	1.845	0.800	0.102	0.984	0.705	2.341	1.017
8.00E-01		8.87E-02	1.01E+00	0.689	2.289	0.990	0.084	0.808	0.858	2.851	1.238
1.00E+00		7.10E-02	8.05E-01	0.860	2.859	1.242	0.069	0.660	1.051	3.491	1.516
1.17E+00	Co (60)	6.18E-02	7.01E-01	0.989	3.285	1.426	0.060	0.580	1.195	3.971	1.724
1.33E+00	Co (60)	5.61E-02	6.36E-01	1.089	3.619	1.572	0.055	0.530	1.307	4.342	1.886
1.50E+00		5.22E-02	5.92E-01	1.170	3.8884	1.689	0.052	0.495	1.401	4.655	2.022
2.00E+00		4.61E-02	5.22E-01	1.326	4.407	1.915	0.045	0.435	1.592	5.287	2.296

TABLE IX. RESULTS OF THE COMBINATION OF THE TWO SHIELDING MATERIALS BY WINXCOM

WinXCOM (Pb+Fe)					
Photon energy (MeV)	Scattering		Photoelectric absorption (cm <sup>2</sup> /g)	Total attenuation	
	Coherent (cm <sup>2</sup> /g)	Incoherent (cm <sup>2</sup> /g)		With coherent scattering (cm <sup>2</sup> /g)	Without coherent scattering (cm <sup>2</sup> /g)
6.620E-01	3.831E-03	6.586E-02	2.210E-02	9.180E-02	8.796E-02
8.000E-01	2.652E-03	6.056E-02	1.464E-02	7.785E-02	7.519E-02
1.000E+00	1.714E-03	5.454E-02	9.223E-03	6.548E-02	6.377E-02
1.022E+00	1.643E-03	5.399E-02	8.826E-03	6.445E-02	6.281E-02
1.173E+00	1.253E-03	5.044E-02	6.727E-03	5.851E-02	5.725E-02
1.250E+00	1.106E-03	4.884E-02	5.955E-03	5.613E-02	5.502E-02
1.332E+00	9.754E-04	4.728E-02	5.282E-03	5.398E-02	5.301E-02
1.500E+00	7.712E-04	4.443E-02	4.242E-03	5.053E-02	4.976E-02
2.000E+00	4.363E-04	3.795E-02	2.567E-03	4.436E-02	4.392E-02

Metals are commonly used in the design of radiation protection systems. Lead is dense and can be used against various high-energy applications of radiation. Lead possesses specific characteristics, e.g. its HVL increases with increase in energy [34]. To assess the shielding ability of a material, HVL is inversely related to the shielding effectiveness. With the increase in energy, HVL increases, due to the reducing intensity of incident gamma radiation to one half, more target thickness will be required. In iron, the HVL increases slightly. This is attributed to the small decrease in LAC with the decrease in iron content (Figures 1 and 2). We can see the significant differences in the HVL of Fe and Pb in Figures 1 and 3. The MAC of Fe and Pb can be seen in Figures 2 and 4. However, the combination is slightly changed in HVL and MAC. Gamma ray factor is larger through Cs-137 source and it is decreased in Co-60. The theoretical and experimental values of MAC, HVL, and TVL of gamma energy at 662, 1173 and 1333keV for lead and iron show that there is a good agreement between the theoretical and the experimental values [35].

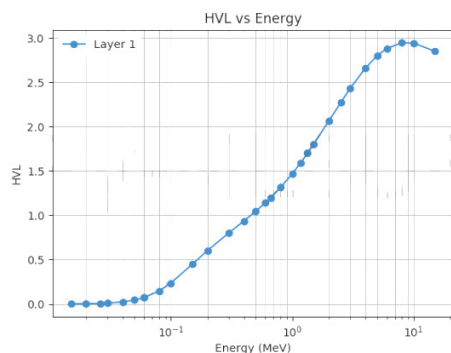


Fig. 1. Py-MLUBF (Fe) HVL vs energy.

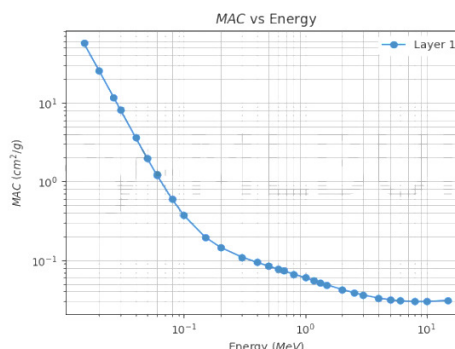


Fig. 2. Py-MLUBF (Fe) MAC vs energy.

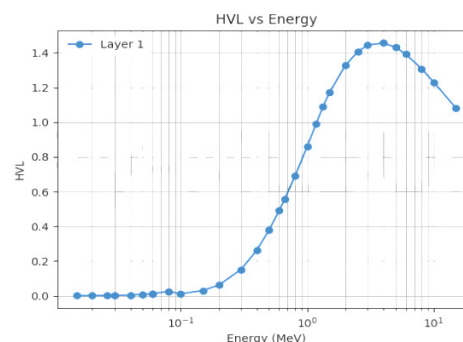


Fig. 3. Py-MLUBF (Pb) HVL vs energy.

#### IV. CONCLUSION

In this paper, mass attenuation coefficients and half-value layer values have been studied using COM, Phy-X, and Py-MLUBF. The codes provide quick calculations of gamma-ray

interaction parameters of the sample for the selected energies. The MAC values are found to reduce exponentially with increasing energy, whereas the HVL values are found to increase exponentially and radiation shielding values increase with increasing energy. The theoretical results are in good agreement with the results from the experimental work.

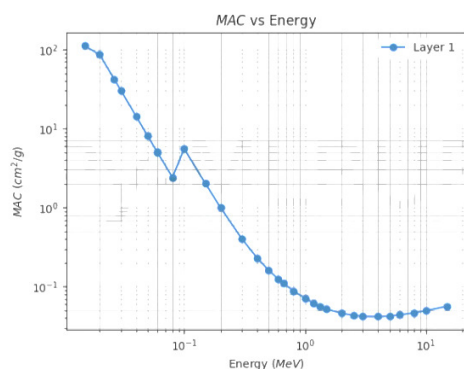


Fig. 4. Py-MLUBF (Pb) MAC vs energy.

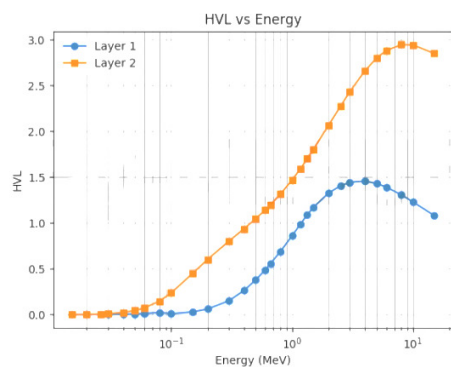


Fig. 5. Fe+Pb combination HVL vs energy.

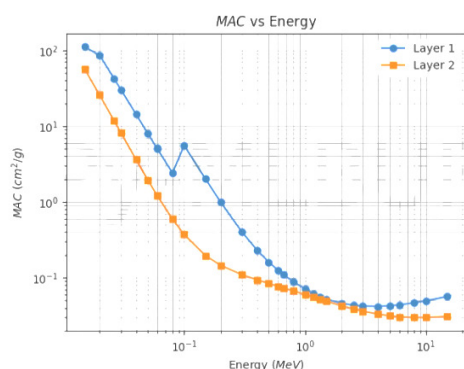


Fig. 6. Fe+Pb combination MAC vs energy.

The efficiency of the shielding material depends on its thickness and interaction energy. Knowing MAC, HVL, TVL, and MFP values may help determine which material reduces the radiation intensity more effectively. The results of this study may encourage the authorities to disseminate the radioprotection culture to the public. The heat and radiation resistance of materials such as Fe and Pb should be studied quantitatively in experiments and the results can help as a data

base for future use. With this information at hand, future investigations will allow us to enhance radiation control.

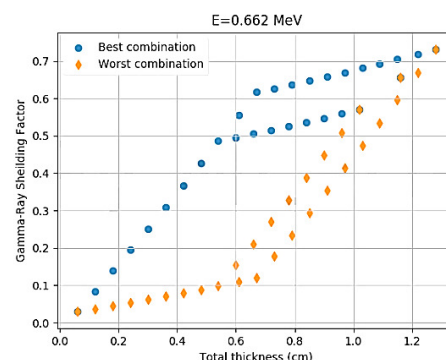


Fig. 7. Gamma Ray shielding factor for Fe+Pb calculated by Py-MLUBF for Cs-137.

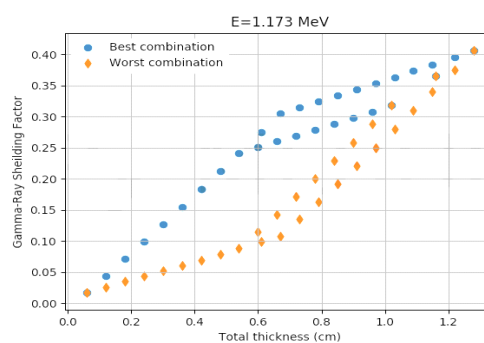


Fig. 8. Gamma Ray shielding factor for Fe+Pb calculated by Py-MLUBF for Co-60.

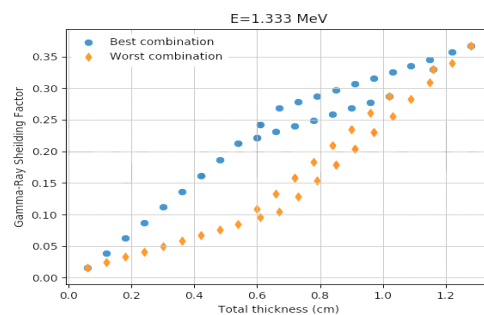


Fig. 9. Gamma Ray shielding factor for Fe+Pb calculated by Py-MLUBF for Co-60.

## REFERENCES

- [1] M. J. R. AL-Dhuhaibat, "Study of the shielding properties for some composite materials manufactured from polymer epoxy supported by cement, aluminum, iron and lead against gamma rays of the cobalt radioactive source (Co-60)," *International Journal of Application or Innovation in Engineering & Management*, vol. 4, no. 6, pp. 90–98, 2015.
- [2] S. F. Olukotun *et al.*, "Investigation of gamma radiation shielding capability of two clay materials," *Nuclear Engineering and Technology*, vol. 50, no. 6, pp. 957–962, Aug. 2018, <https://doi.org/10.1016/j.net.2018.05.003>.
- [3] S. Harb, A. H. El-Kamel, A. Abbady, A. M. Zahran, and F. A. Ahmed, "Natural Radioactivity Measurements of Basalt Rocks in Aden

- governorate, South of Yemen on Gulf of Aden," *Journal of Applied Physics*, vol. 5, no. 6, pp. 39–48, Jan. 2014, <https://doi.org/10.9790/4861-0563948>.
- [4] A. M. Abdelmonem, "Gamma rays and thermal neutron attenuation studies of special composite mixes for using in different applications," *Radiation Physics and Chemistry*, vol. 186, Sep. 2021, Art. no. 109541, <https://doi.org/10.1016/j.radphyschem.2021.109541>.
- [5] Y. Elmahroug, B. Tellili, and C. Souga, "Determination of total mass attenuation coefficients, effective atomic numbers and electron densities for different shielding materials," *Annals of Nuclear Energy*, vol. 75, pp. 268–274, Jan. 2015, <https://doi.org/10.1016/j.anucene.2014.08.015>.
- [6] U. S. Rajurkar and P. P. Pawar, "Measurement of attenuation coefficient and mean free path of some vitamins in the energy range 0.122-1.330 MeV," *Journal of Chemical and Pharmaceutical Research*, vol. 8, no. 5, pp. 852–856, 2016.
- [7] N. Nakao *et al.*, "Attenuation length of high energy neutrons through a thick concrete shield measured by activation detectors at CHARM," *Journal of Nuclear Science and Technology*, vol. 57, no. 9, pp. 1022–1034, Sep. 2020, <https://doi.org/10.1080/00223131.2020.1751740>.
- [8] N. Ibrahim, "Natural activities of  $^{238}\text{U}$ ,  $^{232}\text{Th}$  and  $^{40}\text{K}$  in building materials," *Journal of Environmental Radioactivity*, vol. 43, no. 3, pp. 255–258, May 1999, [https://doi.org/10.1016/S0265-931X\(98\)00033-2](https://doi.org/10.1016/S0265-931X(98)00033-2).
- [9] E. Kavaz, N. Ekin, H. O. Tekin, M. I. Sayyed, B. Aygun, and U. Perisanoglu, "Estimation of gamma radiation shielding qualification of newly developed glasses by using WinXCOM and MCNPX code," *Progress in Nuclear Energy*, vol. 115, pp. 12–20, Aug. 2019, <https://doi.org/10.1016/j.pnucene.2019.03.029>.
- [10] M. B. Gili and F. Hila, "Investigation of Gamma-ray Shielding Features of Several Clay Materials Using the EPICS2017 Library," *Philippine Journal of Science*, vol. 150, pp. 1017–1026, Jul. 2021, <https://doi.org/10.56899/150.05.13>.
- [11] A. M. Madbouly and A. El-Sawy, "Calculation of Gamma and Neutron Parameters for Some Concrete Materials as Radiation Shields for Nuclear Facilities," *International Journal of Emerging Trends in Engineering and Development*, vol. 8, no. 3, pp. 7–17, Aug. 2018, <https://doi.org/10.26808/rs.ed.i8v4.02>.
- [12] H. Q. Vu, V. H. Tran, P. T. Nguyen, N. T. H. Le, and M. T. Le, "Radiation Shielding Properties Prediction of Barite used as Small Aggregate in Mortar," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6469–6475, Dec. 2020, <https://doi.org/10.48084/etasr.3880>.
- [13] F. A. Al-Mufadi, A. El-Taher, and G. A. Gamal, "Influence of  $\gamma$ -Irradiation on the Structural Properties of Iodine Monoselenide Crystals," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1264–1268, Dec. 2016, <https://doi.org/10.48084/etasr.703>.
- [14] M. R. Abdullah, O. K. Alghazawi, and M. Al-Ayyad, "Non-uniform Heat Source and Radiation Effect on a Transient MHD Flow Past a Vertical Moving Plate with Inclined Magnetic Field and Periodic Heat Flux," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4361–4366, Aug. 2019, <https://doi.org/10.48084/etasr.2779>.
- [15] J. Räisänen, "Experimental arrangements for the simultaneous use of PIXE and complementary accelerator based techniques," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 49, no. 1, pp. 39–45, Apr. 1990, [https://doi.org/10.1016/0168-583X\(90\)90213-E](https://doi.org/10.1016/0168-583X(90)90213-E).
- [16] M. I. Sayyed *et al.*, "Experimental and Theoretical Study of Radiation Shielding Features of  $\text{CaO-K}_2\text{O-Na}_2\text{O-P}_2\text{O}_5$  Glass Systems," *Materials*, vol. 14, no. 14, Jan. 2021, Art. no. 3772, <https://doi.org/10.3390/ma14143772>.
- [17] N. Singh, K. J. Singh, K. Singh, and H. Singh, "Gamma-ray attenuation studies of  $\text{PbO-BaO-B}_2\text{O}_3$  glass system," *Radiation Measurements*, vol. 41, no. 1, pp. 84–88, Jan. 2006, <https://doi.org/10.1016/j.radmeas.2004.09.009>.
- [18] E. O. Echeweozo, A. D. Asiegbu, and E. L. Efurumibe, "Investigation of kaolin - Granite composite bricks for gamma radiation shielding," *International Journal of Advanced Nuclear Reactor Design and Technology*, vol. 3, pp. 194–199, Jan. 2021, <https://doi.org/10.1016/j.jandrt.2021.09.007>.
- [19] K. Won-In, N. Sirikulrat, and P. Dararutana, "Radiation Shielding Lead-Free Glass Based on Barium-Bearing Glass Using Thailand Quartz Sands," *Advanced Materials Research*, vol. 214, pp. 207–211, 2011, <https://doi.org/10.4028/www.scientific.net/AMR.214.207>.
- [20] O. Agar, "Investigation on Gamma Radiation Shielding Behaviour of  $\text{CdO-WO}_3\text{-TeO}_2$  Glasses from 0.015 to 10 MeV," *Cumhuriyet Science Journal*, vol. 39, no. 33, pp. 983–990, 2018, <https://doi.org/10.17776/csj.451770>.
- [21] N. R. Abd Elwahab, N. Helal, T. Mohamed, F. Shahin, and F. M. Ali, "New shielding composite paste for mixed fields of fast neutrons and gamma rays," *Materials Chemistry and Physics*, vol. 233, pp. 249–253, May 2019, <https://doi.org/10.1016/j.matchemphys.2019.05.059>.
- [22] A. Martin, S. Harbison, K. Beach, and P. Cole, *An Introduction to Radiation Protection*, 7th Ed. Boca Raton, FL, USA: CRC Press, 2018.
- [23] K. S. Mann, "Investigation of gamma-ray shielding by double layered enclosures," *Radiation Physics and Chemistry*, vol. 159, pp. 207–221, Jun. 2019, <https://doi.org/10.1016/j.radphyschem.2019.03.007>.
- [24] K. S. Mann and S. S. Mann, "Py-MLBUF: Development of an online-platform for gamma-ray shielding calculations and investigations," *Annals of Nuclear Energy*, vol. 150, Jan. 2021, Art. no. 107845, <https://doi.org/10.1016/j.anucene.2020.107845>.
- [25] L. Gerward, N. Guilbert, K. B. Jensen, and H. Levring, "WinXCom—a program for calculating X-ray attenuation coefficients," *Radiation Physics and Chemistry*, vol. 71, pp. 653–654, Oct. 2004, <https://doi.org/10.1016/j.radphyschem.2004.04.040>.
- [26] M. I. Sayyed, M. Elsafi, A. H. Almuqrin, K. Cornish, and A. M. Elkhatib, "Novel Shielding Mortars for Radiation Source Transportation and Storage," *Sustainability*, vol. 14, no. 3, Jan. 2022, Art. no. 1248, <https://doi.org/10.3390/su14031248>.
- [27] E. Sakar, O. F. Ozpolat, B. Alim, M. I. Sayyed, and M. Kurudirek, "Phy-X / PSD: Development of a user friendly online software for calculation of parameters relevant to radiation shielding and dosimetry," *Radiation Physics and Chemistry*, vol. 166, Jan. 2020, Art. no. 108496, <https://doi.org/10.1016/j.radphyschem.2019.108496>.
- [28] M. S. Eid *et al.*, "Implementation of waste silicate glass into composition of ordinary cement for radiation shielding applications," *Nuclear Engineering and Technology*, vol. 54, no. 4, pp. 1456–1463, Apr. 2022, <https://doi.org/10.1016/j.net.2021.10.007>.
- [29] A. El-Sayed Abdo, "Calculation of the cross-sections for fast neutrons and gamma-rays in concrete shields," *Annals of Nuclear Energy*, vol. 29, no. 16, pp. 1977–1988, Nov. 2002, [https://doi.org/10.1016/S0306-4549\(02\)00019-1](https://doi.org/10.1016/S0306-4549(02)00019-1).
- [30] G. AlMisned *et al.*, "Gamma-Ray Protection Properties of Bismuth-Silicate Glasses against Some Diagnostic Nuclear Medicine Radioisotopes: A Comprehensive Study," *Materials*, vol. 14, no. 21, Jan. 2021, Art. no. 6668, <https://doi.org/10.3390/ma14216668>.
- [31] A. H. Almuqrin and M. I. Sayyed, "Gamma Ray Shielding Properties of  $\text{Yb}_3\text{+Doped Calcium Borotellurite Glasses}$ ," *Applied Sciences*, vol. 11, no. 12, Jan. 2021, Art. no. 5697, <https://doi.org/10.3390/app11125697>.
- [32] M. I. Sayyed, "Half value layer, mean free path and exposure buildup factor for tellurite glasses with different oxide compositions," *Journal of Alloys and Compounds*, vol. 695, pp. 3191–3197, Feb. 2017, <https://doi.org/10.1016/j.jallcom.2016.11.318>.
- [33] F. Akman, O. Agar, M. R. Kacal, and M. I. Sayyed, "Comparison of experimental and theoretical radiation shielding parameters of several environmentally friendly materials," *Nuclear Science and Techniques*, vol. 30, no. 7, Jun. 2019, Art. no. 110, <https://doi.org/10.1007/s41365-019-0631-1>.
- [34] R. Singh, S. Singh, G. Singh, and K. S. Thind, "Gamma Radiation Shielding Properties of Steel and Iron Slags," *New Journal of Glass and Ceramics*, vol. 7, no. 1, 2017, Art. no. 72939, <https://doi.org/10.4236/njgc.2017.71001>.
- [35] A. B. Aziz, Md. F. Rahman, and M. M. Prodhan, "Comparison of Lead, Copper and Aluminium as Gamma Radiation Shielding Material through Experimental Measurements and Simulation Using MCNP Version 4c," *International Journal of Contemporary Research and Review*, vol. 9, no. 8, pp. 20193–20206, Aug. 2018, <https://doi.org/10.15520/ijcrr/2018/9/08/584>.

# Behavior of Strip Footing/s above Void in Sandy Soil

**Assma Benbouza**

LGC-ROI Civil Engineering Laboratory, Department of Civil Engineering, Faculty of Technology,  
University of Batna 2, Algeria  
as.benbouza@univ-batna2.dz

**Tarek Mansouri**

LGC-ROI Civil Engineering Laboratory, Department of Civil Engineering, Faculty of Technology,  
University of Batna 2, Algeria  
t.mansouri@univ-batna2.dz

**Khelifa Abbeche**

LGC-ROI Civil Engineering Laboratory, Department of Civil Engineering, Faculty of Technology,  
University of Batna 2, Algeria  
k.abbeche@univ-batna2.dz

Received: 17 November 2022 | Revised: 4 December 2022 | Accepted: 12 December 2022

## ABSTRACT

This paper presents a numerical study that utilizes finite element analysis under the plain strain condition performed on sand with isolated strip footing and two closely spaced strip footings above a continuous void. The Bearing Capacity Ratio (BCR) and the efficiency factor  $\zeta\gamma$  were introduced to determine the effect of the void on the ultimate bearing capacity of footing/s. The influence of various parameters including spacing (S/B) (i.e. edge to edge) between the two interfering footings along with the location and the shape of the void were studied. In general, the results indicate that the presence of a void reduces the bearing capacity and affects the performance of footing/s and there is a critical value of S/B beyond which the effect of the void on the bearing capacity of the interfering footings becomes negligible.

**Keywords-***bearing capacity; strip footing; finite element method; underground voids; granular soil; interfering footings*

## I. INTRODUCTION

During the recent years, urban development has been one of the Algerian government's top priorities, which has greatly reduced the amount of land available for the construction of projects, thus, it has become impossible to avoid constructing buildings close to each other in order to make the most of the land and reduce cost, but this type of construction can have an influence on the bearing capacity of the foundations. The calculation of the bearing capacity of shallow footings is usually done using a method similar to that developed in [1] for isolated footings. In reality, footings are rarely isolated and interfere with each other to some extent. Shallow footings interference has been studied theoretically in [2] using the limit equilibrium method and the stress characteristic method. Authors in [3] presented laboratory test results for two closely spaced footings on sand. They found that the interference factors are generally lower than those predicted by theory. Authors in [4] studied the interaction between two closely spaced rough and smooth strip footings on dense and loose

sand. Authors in [5] used the method of characteristics for two different failure mechanisms. Authors in [6] used the upper bound limit analysis for finding the interference effect of two nearby strip footings on sand. Authors in [7] examined the closely spaced footings on geogrid-reinforced sand, and authors in [8] studied the interference effect of strip and square footings on sand reinforced with geosynthetics. Authors in [9] studied numerically the bearing capacity of two interfering strip footings on sands. Authors in [10] studied the failure surface in granular soil under two closely spaced strip footings and concluded that the failure patterns observed for granular soil conform to those proposed by the theory in [2]. Authors in [11] experimentally studied the interference effect of two closely spaced strip footings constructed on bi-layer soil. Authors in [12] studied experimentally and numerically the interference of closely spaced square footings on sandy soil, and concluded that the maximum interference effect is observed when the spacing between the two footings is  $0.5B$ , and was approximately negligible when the spacing between the footings was equal to  $2B$ , where  $B$  is the footing's width.

The existence of underground voids (e.g. natural or artificial caves) may be attributed to two reasons: dissolution of soluble materials (e.g. salt, dolomite, and limestone) and artificial underground activities such as tunneling, mining, subway excavations, etc. These voids could cause serious engineering problems leading to poor performance of shallow foundations, structure collapses, road settlements, etc. which need special attention in engineering practice. In many cases, footings are positioned on soil containing voids that are either not visible before construction or are formed after the construction. Several studies on the subject are available and some of them deal with cavity-footing interaction, e.g. [13], while others studied experimentally, theoretically, and numerically the effect of a void on the stability of shallow footings [14-18, 19]. Other studies investigated the yielding pressure of strip footings above multiple-shaped cavities using FEM [20, 21]. Authors in [22] adopted FE analysis to estimate the undrained bearing capacity of surface strip footings above one and two voids. Authors in [23] studied with FEM the behavior of shallow strip footing on twin voids and clarified the failure mechanism. Authors in [24] investigated the ultimate bearing capacity and failure mechanism of strip footings subjected to vertical load placed on c- $\phi$  soil with square voids. The critical and adverse locations of voids were analyzed. Authors in [25] investigated numerically the undrained stability of strip footings above voids in two-layered clays. Using finite element limit analysis, the undrained bearing capacity factor  $N_s$  of strip footings has been calculated and the effect of the thickness of the top layer and the effect of undrained shear stress ratio on  $N_s$  were studied. In case of slope, a numerical study was conducted to analyze the bearing capacity behavior of strip footings on a reinforced sand slope with a single circular void [26]. Further investigations have been carried out on the effect of artificial cavities on deep foundations [27, 28].

The influence of voids on the performance of shallow interfering footings has not been well covered or is not available in the literature. For this reason, this study aims to investigate numerically the effect of interference of two closely spaced strip footings above single continuous void in sandy soil. The main purpose of this study, in order to evaluate the bearing capacity of strip footing/s above void, was to reveal the effect of various parameters, such as the spacing ( $S/B$ ) (i.e. edge to edge) between two footings, the shape and the location of the void, on the ultimate bearing capacity.

## II. PROBLEM DEFINITION

The geometry and the key parameters of the problem analyzed in this paper are illustrated in Figure 1. The geometry consists of an isolated rigid strip footing and two closely spaced rigid strip footings with  $B$  and  $S$  representing the footing width and the spacing between the two footings respectively. They are placed on the horizontal surface of an isotropic homogenous soil of friction angle  $\phi$  and unit weight  $\gamma$ . A static vertical load is imposed. By considering the case of zero surcharge and zero soil cohesion, the bearing capacity formula reduces to:

$$q_u = \frac{1}{2} \gamma B N_\gamma \zeta_\gamma \quad (1)$$

where  $\zeta_\gamma$  is the efficiency factor which is function of  $S/B$  and  $\phi$ .

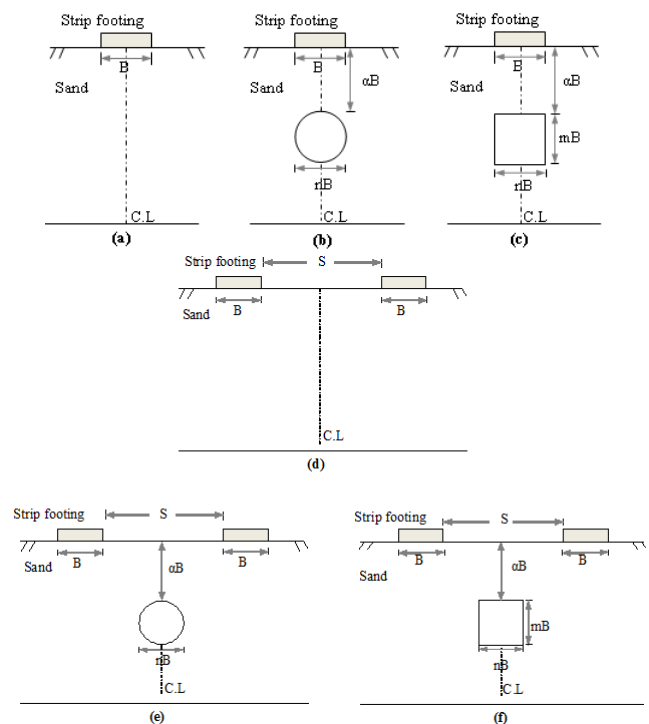


Fig. 1. Schematic view of footing/s and void.

In case of voids below the footings, one square and one circular voids were adopted in this study. Shape and location of the void and the spacing between voids are quantified in terms of dimensionless parameters, i.e.  $m$ ,  $n$ ,  $\alpha$ . Parameters  $m$  and  $n$  represent the void height and width or diameter normalized by the footing width  $B$  and they are equal to 1 in this study. Parameter  $\alpha$  designates the relative vertical distance from the centerline of the footing to the void normalized by  $B$ . The configuration of a single square or circular void is shown in Figure 1. According to [20], the void continuously extends horizontally. Furthermore, the ratio of void width to the width of the strip footing is equal to 1, which corresponds to the presupposition of [22].

## III. FINITE ELEMENT MODEL

The commercially available finite element program PLAXIS was used to model the footings-voids system. The behavior of soil was numerically simulated as an elastic-perfectly plastic material considering Mohr-Coulomb failure criterion in conjunction with a non-associated flow rule (i.e.  $\phi \neq \psi$ ). The values of the geotechnical properties are: Young's modulus  $E = 30\text{MPa}$ , Poisson's ratio  $\nu = 0.2$ , bulk unit weight  $\gamma = 20\text{KN/m}^3$ , and friction angle  $\phi = 35^\circ$ . The soil was modeled with 15-node triangular elements. Furthermore, Figure 1 shows two shapes of void: the continuous circular void which is considered as a tunnel without lining and the continuous square void which is introduced by excavation of the soil at the desired depth, to model a natural cavity. Thus the vertical and horizontal limits of this model are chosen appropriately far to avoid any influence on the results [17]. The dimensions of the area for this analysis are  $15B$  in vertical and  $30B$  in horizontal taking that  $B=1\text{m}$ . Figure 2 shows a



schematic of the boundary conditions and the finite element mesh used in the current study. The vertical borders have been fixed in the horizontal direction and full fixities have taken place at the bottom of the model. The mesh is refined in the area adjacent to the footings and around voids to enhance the accuracy of the numerical results. The footings are considered very stiff and rough. In this study, instead of modeling the footing itself, the settlement of the footing is imposed by means of a uniform indentation at the top of the sandy layer until the ground reaches the failure state. The uniformly displacement is automatically decided with trial calculations in the program. Because of the symmetrical nature of the problem and in order to reduce the time required for each run, only half of the model was taken in the numerical simulation.

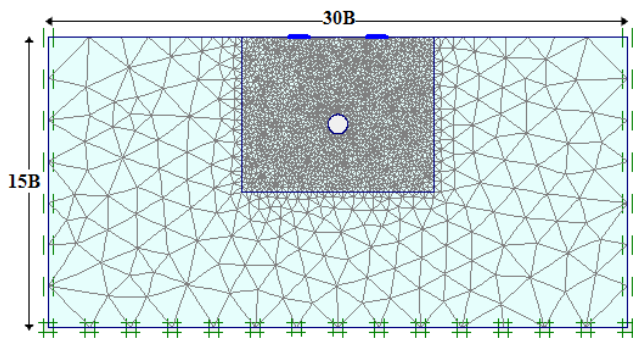


Fig. 2. Finite element mesh with boundary conditions.

#### IV. TEST PROGRAM

A total of 98 parametric tests were performed on the load-bearing capacity of rigorous strip foundations resting on one cavity of circular or square shape, as shown in Table I.

TABLE I. DETAILS OF NUMERICALLY MODEL TESTS

Test series	Footings number	Shape of voids	Variable parameters		Fixed parameters
			$\alpha$	S/B	
A	Isolated footing	Without void	/	/	$n=m=B=1\text{m}$ $\phi=35^\circ$ $L/B=1$
		With a circular void	1 - 2 - 3 - 4 - 5 - 6 - 7 and 8	/	
		With a square void	1 - 2 - 3 - 4 - 5 - 6 - 7 and 8	/	
B	Two interfering footings	Without void	/	0 - 0,5 - 1 - 1,5 - 2 - 3 - 4 - 5 and 6	
		With a circular void	1 - 3 - 5 and 7	0 - 0,5 - 1 - 1,5 - 2 - 3 - 4 - 5 and 6	
		With a square void	1 - 3 - 5 and 7	0 - 0,5 - 1 - 1,5 - 2 - 3 - 4 - 5 and 6	

The efficiency factor  $\zeta_\gamma$  was introduced to determine the influence of the void on the ultimate bearing capacity.  $\zeta_\gamma$  is expressed as:

$$\zeta_\gamma = \frac{q_{u \text{ int v}}}{q_{u \text{ isol}}} \quad (2)$$

where  $q_{u \text{ int v}}$  is the ultimate bearing capacity of two strip footings with or without void, and  $q_{u \text{ isol}}$  is the ultimate bearing capacity of an isolated strip footing without void.

#### V. VALIDATION

To validate the numerical model, the bearing capacity factor  $N_\gamma$  of a single strip footing on sand is calculated. The obtained  $N_\gamma$  value is compared with the results from the literature while the friction angle is equal to  $35^\circ$ . As shown in Table II, the result of this study is remarkably close to that given in the literature. This good accordance can be taken as a validation of the present numerical model.

TABLE II. VALUES OF BEARING CAPACITY FACTOR  $N_\gamma$

Reference	[1]	[29]	[30]	[31]	Current study
$N_\gamma$	45.41	37.15	48.03	33.92	45.72

#### VI. RESULTS AND DESCUTION

##### A. Single Strip Footing in Soil with and without Void

The bearing pressure-displacement curves for the strip footing in soil without void and with a circular or square void are shown in Figure 3. For all cases, the footing reaches a clear limit load, which was taken as the ultimate bearing capacity. The presence of the void has a big influence on the bearing pressure. The magnitude of bearing pressure is higher for the soil without a void, therefore the presence of voids in the soil reduces bearing capacity. In addition, the magnitude of bearing pressure for a soil with circular void is slightly higher than for a square void.

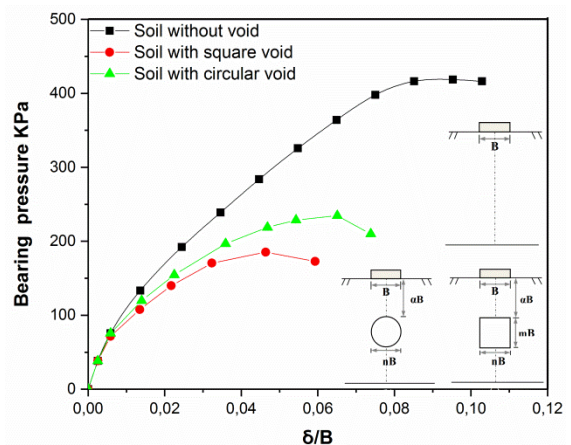


Fig. 3. Bearing pressure-displacement ( $\alpha=3$ ).

##### 1) Effect of Void Depth from the Single Footing Base

The influence of  $\alpha$  on the bearing capacity is presented in Figure 4 which shows the relationship between the Bearing

Capacity Ratio (BCR) as a function of the void depth  $\alpha$ . BCR is defined as:

$$BCR = \frac{q_{u \text{ isol } v}}{q_{u \text{ isol}}} \quad (3)$$

where  $q_{u \text{ isol } v}$  is the ultimate bearing capacity of an isolated strip footing with void and  $q_{u \text{ isol}}$  is the ultimate bearing capacity of an isolated strip footing without void.

As can be noticed in Figure 4 the bearing capacity ratio increases linearly when the value of  $\alpha$  increases from 1 to 4 for circular void and from 1 to 5 for square void and remains almost constant thereafter for either circular or square void. The BCR recorded is 0.14 to 1 for  $\alpha$  ranging from 1 to 8. This indicates that the existence of circular or square void under the foundation at a depth equal to or greater than 6 times the foundation width does not affect the bearing capacity of the footing and the void impact is eliminated. These results are in accordance with the findings of [24]. Also, it is observed from Figure 4 that the values of the BCR for the circular void case are larger than those of the BCR for the square void case.

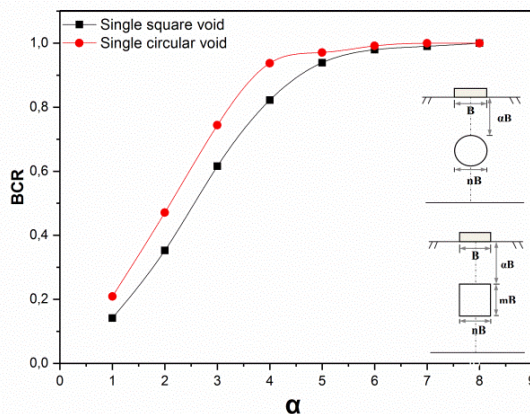


Fig. 4. Variation of BCR as a function of  $\alpha$ .

## B. Interfering Effect of Two Strip Footings with and without Void

### 1) Two Adjacent Strip Footings without Void

The results of numerical analysis obtained by the Plaxis code shown in Figure 5 prove that for  $0 \leq S/B \leq 1$ , the efficiency factor ( $\zeta_\gamma$ ) magnitude increased to its maximum value, which means that the ultimate bearing capacity of each strip footing increases almost by 60% and for  $1 \leq S/B \leq 4$ ,  $\zeta_\gamma$  decreases with an increase in spacing ratio. Finally, for  $S/B \geq 4$ ,  $\zeta_\gamma$  remains constant. This means that for a spacing ratio greater than 4B, no interference effect was observed and each footing acted as an isolated footing.

To verify the accuracy of the results obtained by the present study, they were compared with the obtained numerical analysis results from [8, 9], theoretical analysis [2], and experimental test [3], which are presented in Figure 5.

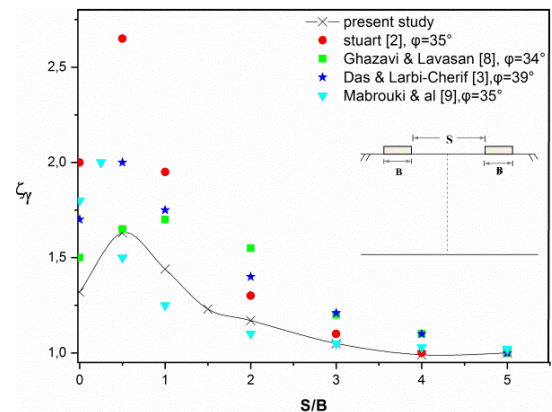


Fig. 5. Comparison between the literature values and present study.

It appears that the general trend of interference factor variations found in this study is similar to those predicted by other studies, but there is a large variation in amplitudes between theory, the experimental, and the numerical results. From this Figure, the numerical results agree very well with the experimental test results [10].

### 2) Two Adjacent Strip Footings with a Single, Square or Circular Void

Figure 6 indicates the variation of  $\zeta_\gamma$  with different spacing ratios  $S/B$  for two interfering strip footings above a single circular or square void for a variation of void depths  $\alpha=1, 3, 5$ , and 7 at 2m intervals, for a sandy soil with a void located in the centerline of the model. It can be noted that for  $S/B$  varying from 0 to 1.5 for circular void and from 0 to 2 for square void, the value of  $\zeta_\gamma$  increases with increasing  $\alpha$  until the void is present within a certain critical depth  $\alpha$  which is about 7 for a circular void and a little higher for a square void. The influence of the void gradually becomes insignificant so only the interference effect of the two footings remains. When the footings are located away from the centerline of the model (i.e. when the  $S/B$  value increases), the efficiency factor magnitude increases with an increase in  $\alpha$  and reaches the maximum value when  $S/B = 1$  for  $\alpha = 1, 5$ , or 7 and  $S/B = 1.5$  for  $\alpha = 3$ . In most cases  $\zeta_\gamma$  decreases until the ultimate bearing capacity remains at 100% for  $S/B$  almost equal to 4, indicating no effect of the void on footing stability while there is not much interference effect. The only exception to this observation occurs when  $\alpha = 1$  where  $\zeta_\gamma$  values are significantly lower and keep invariant approximately for  $S/B$  varying between 1.5 and 2.5 for a circular void and between 1 and 1.5 for a square void, particularly due to the effect of the stability of the soil above the void which is more dominant than that of the effect of footings' interference.  $\zeta_\gamma$  continues its increase to reach 1 at  $S/B=6$  for a circular void and almost 1 at  $S/B=6$  for a square void, so the void effect is neglected and there is not much interference effect. On the other hand, for the case of a circular void, the  $\zeta_\gamma$  for all values of  $S/B$  and  $\alpha$  is more than that for the case of the square void shown in Figure 6. It clearly can be observed that same remark in Figures 3, 5, and 6, in which the values of bearing capacity for the circular void are larger than those of the bearing capacity for the square void, due to the section of the square void which is greater than that of the

circular void and not to the shape of the void, therefore the effect of the void's shape can be neglected [32].

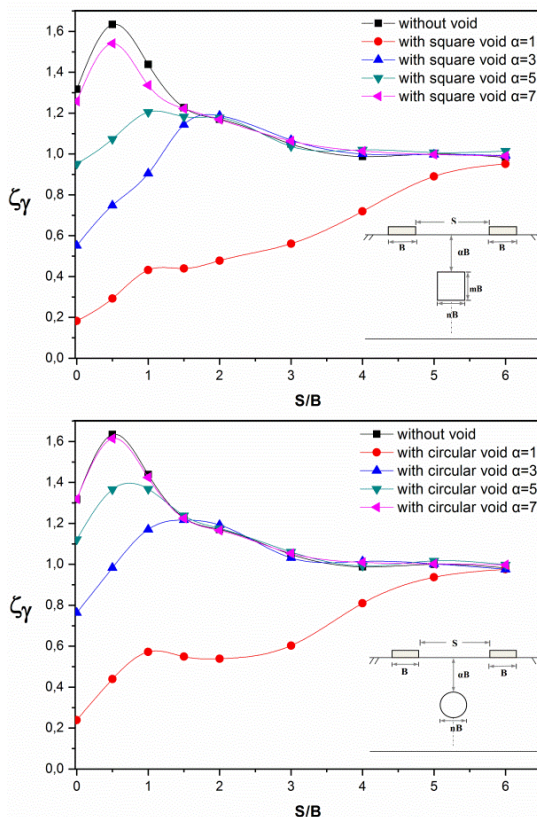


Fig. 6. Variation of the efficiency factor  $\zeta\gamma$  versus spacing ratio  $S/B$  for two interfering strip footings above a single circular or square void.

## VII. CONCLUSION

In this paper, the behavior of strip footings above voids in sandy soil has been investigated numerically, and the BCR and  $\zeta\gamma$  factors of footings have been calculated. In addition, the critical location of the void and the spacing between footings have also been discussed. The main conclusions of the current study are:

- The presence of void has a big influence on bearing pressure. The magnitude of the bearing pressure is higher for a soil without voids, therefore the presence of voids in the soil reduces bearing capacity.
- With single strip footing, the existence of a circular or square void under the foundation at a depth equal to or greater than 6 times the foundation width does not affect the bearing capacity of the footing and the void impact is eliminated.
- With two interfering strip footings without void the results of numerical analysis proved that for a spacing ratio greater than  $4B$ , no interference effect was observed and each footing acted as an isolated footing.
- With two interfering strip footings with a void located in the centerline of the model, it can be noted that for values of

$S/B$  varying from 0 to 1.5 for circular void and from 0 to 2 for square void, the value of  $\zeta\gamma$  increases with increasing  $\alpha$  until the void is present within a certain critical depth  $\alpha$  which is about 7 for a circular void and a little higher for a square void. The influence of the void becomes insignificant so only the interference effect of the two footings remains.

- There is a critical value of  $S/B$  beyond which the effect of the void on the bearing capacity of the interfering footings becomes negligible.
- The values of bearing capacity for the circular void case are larger than those for the square void case, due to the section of the square void which is greater than that of the circular void and not to the shape of void, therefore the effect of the void shape can be neglected.

## REFERENCES

- [1] K. Terzaghi, *Theoretical Soil Mechanics*. New York, NY, USA: Wiley, 1943.
- [2] J. G. Stuart, "Interference Between Foundations, with Special Reference to Surface Footings in Sand," *Géotechnique*, vol. 12, no. 1, pp. 15–22, Mar. 1962, <https://doi.org/10.1680/geot.1962.12.1.15>.
- [3] B. M. Das and S. Larbi-Cherif, "Bearing Capacity of Two Closely-Spaced Shallow Foundations on Sand," *Soils and Foundations*, vol. 23, no. 1, pp. 1–7, Mar. 1983, <https://doi.org/10.3208/sandf1972.23.1>.
- [4] E. C. J. Hazell, "Interaction of closely spaced strip footings," Department of Engineering Science, University of Oxford, Final year project report, 2004.
- [5] J. Kumar and P. Ghosh, "Ultimate Bearing Capacity of Two Interfering Rough Strip Footings," *International Journal of Geomechanics*, vol. 7, no. 1, pp. 53–62, Jan. 2007, [https://doi.org/10.1061/\(asce\)1532-3641\(2007\)7:1\(53\)](https://doi.org/10.1061/(asce)1532-3641(2007)7:1(53)).
- [6] J. Kumar and P. Ghosh, "Upper bound limit analysis for finding interference effect of two nearby strip footings on sand," *Geotechnical and Geological Engineering*, vol. 25, no. 5, pp. 499–507, May 2007, <https://doi.org/10.1007/s10706-007-9124-9>.
- [7] A. Kumar and S. Saran, "Closely Spaced Footings on Geogrid-Reinforced Sand," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 129, no. 7, pp. 660–664, Jul. 2003, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2003\)129:7\(660\)](https://doi.org/10.1061/(ASCE)1090-0241(2003)129:7(660)).
- [8] M. Ghazavi and A. A. Lavasan, "Interference effect of shallow foundations constructed on sand reinforced with geosynthetics," *Geotextiles and Geomembranes*, vol. 26, no. 5, pp. 404–415, Oct. 2008, <https://doi.org/10.1016/j.geotexmem.2008.02.003>.
- [9] A. Mabrouki, D. Benmeddour, R. Frank, and M. Mellas, "Numerical study of the bearing capacity for two interfering strip footings on sands," *Computers and Geotechnics*, vol. 37, no. 4, pp. 431–439, Jun. 2010, <https://doi.org/10.1016/j.compgeo.2009.12.007>.
- [10] A. Benbouza, L. Arabet, and K. Abbeche, "Numerical Study of the Failure Surface in Granular Soil Under Two Closely Spaced Strip Footings," in *Sustainable Civil Infrastructures*, Egypt, Jul. 2017, pp. 165–172, [https://doi.org/10.1007/978-3-319-61905-7\\_14](https://doi.org/10.1007/978-3-319-61905-7_14).
- [11] R. Boufarh, K. Abbeche, and A. Abdi, "Experimental Investigation of Interference Between Adjacent Footings on Layered Cohesionless Soil," *Soil Mechanics and Foundation Engineering*, vol. 56, no. 2, pp. 128–135, May 2019, <https://doi.org/10.1007/s11204-019-09580-z>.
- [12] A. Gupta and T. G. Sitharam, "Experimental and numerical investigations on interference of closely spaced square footings on sand," *International Journal of Geotechnical Engineering*, vol. 14, no. 2, pp. 142–150, Apr. 2018, <https://doi.org/10.1080/19386362.2018.1454386>.
- [13] A. Badie and M. C. Wang, "Interaction Between Strip Footing and Soft Ground Tunnel," in *13th International Conference on Soil Mechanics and Foundation Engineering*, New Delhi, India, 1994, pp. 571–574.



- [14] R. L. Baus and M. C. Wang, "Bearing Capacity of Strip Footing above Void," *Journal of Geotechnical Engineering*, vol. 109, no. 1, pp. 1–14, Jan. 1983, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1983\)109:1\(1\)](https://doi.org/10.1061/(ASCE)0733-9410(1983)109:1(1)).
- [15] L. A. Wood and W. J. Carnach, "The Behaviour of Footings Located Above Voids," in *Proceedings of the Eleventh International Conference on Soil Mechanics and Foundation Engineering*, San Francisco, CA, USA, Aug. 1985.
- [16] A. Badie and M. C. Wang, "Stability of Spread Footing Above Void in Clay," *Journal of Geotechnical Engineering*, vol. 110, no. 11, pp. 1591–1605, Nov. 1984, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1984\)110:11\(1591\)](https://doi.org/10.1061/(ASCE)0733-9410(1984)110:11(1591)).
- [17] M. C. Wang and A. Badie, "Effect of Underground Void on Foundation Stability," *Journal of Geotechnical Engineering*, vol. 111, no. 8, pp. 1008–1019, Aug. 1985, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1985\)111:8\(1008\)](https://doi.org/10.1061/(ASCE)0733-9410(1985)111:8(1008)).
- [18] M. C. Wang and C. W. Hsieh, "Collapse load of strip footing above circular void," *Journal of Geotechnical Engineering*, vol. 113, no. 5, May 1987, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1987\)113:5\(511\)](https://doi.org/10.1061/(ASCE)0733-9410(1987)113:5(511)).
- [19] G. Azam, C. W. Hsieh, and M. C. Wang, "Performance of Strip Footing on Stratified Soil Deposit with Void," *Journal of Geotechnical Engineering*, vol. 117, no. 5, pp. 753–772, May 1991, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1991\)117:5\(753\)](https://doi.org/10.1061/(ASCE)0733-9410(1991)117:5(753)).
- [20] M. Kiyosumi, O. Kusakabe, M. Ohuchi, and F. Le Peng, "Yielding Pressure of Spread Footing above Multiple Voids," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 133, no. 12, pp. 1522–1531, Dec. 2007, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2007\)133:12\(1522\)](https://doi.org/10.1061/(ASCE)1090-0241(2007)133:12(1522)).
- [21] M. Kiyosumi, O. Kusakabe, and M. Ohuchi, "Model Tests and Analyses of Bearing Capacity of Strip Footing on Stiff Ground with Voids," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 137, no. 4, pp. 363–375, Apr. 2011, [https://doi.org/10.1061/\(asce\)gt.1943-5606.0000440](https://doi.org/10.1061/(asce)gt.1943-5606.0000440).
- [22] J. K. Lee, S. Jeong, and J. Ko, "Undrained stability of surface strip footings above voids," *Computers and Geotechnics*, vol. 62, pp. 128–135, Oct. 2014, <https://doi.org/10.1016/j.compgeo.2014.07.009>.
- [23] A. A. Lavasan, A. Talsaz, M. Ghazavi, and T. Schanz, "Behavior of Shallow Strip Footing on Twin Voids," *Geotechnical and Geological Engineering*, vol. 34, no. 6, pp. 1791–1805, Dec. 2016, <https://doi.org/10.1007/s10706-016-9989-6>.
- [24] H. Zhou, G. Zheng, X. He, X. Xu, T. Zhang, and X. Yang, "Bearing capacity of strip footings on c-φ soils with square voids," *Acta Geotechnica*, vol. 13, no. 3, pp. 747–755, Feb. 2018, <https://doi.org/10.1007/s11440-018-0630-0>.
- [25] Y. Xiao, M. Zhao, and H. Zhao, "Undrained stability of strip footing above voids in two-layered clays by finite element limit analysis," *Computers and Geotechnics*, vol. 97, pp. 124–133, May 2018, <https://doi.org/10.1016/j.compgeo.2018.01.005>.
- [26] B. Mazouz, T. Mansouri, M. Baazouzi, and K. Abbeche, "Assessing the Effect of Underground Void on Strip Footing Sitting on a Reinforced Sand Slope with Numerical Modeling," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9005–9011, Aug. 2022, <https://doi.org/10.48084/etasr.5131>.
- [27] M. A. Soomro, M. A. Keerio, M. A. Soomro, and D. K. Bangwar, "3D Centrifuge Modeling of the Effect of Twin Tunneling to an Existing Pile Group," *Engineering, Technology & Applied Science Research*, vol. 7, no. 5, pp. 2030–2040, Oct. 2017, <https://doi.org/10.48084/etasr.1393>.
- [28] M. A. Soomro, D. K. Bangwar, M. A. Soomro, and M. A. Keerio, "3D Numerical Analysis of the Effects of an Advancing Tunnel on an Existing Loaded Pile Group," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2520–2525, Feb. 2018, <https://doi.org/10.48084/etasr.1693>.
- [29] G. G. Meyerhof, "Some Recent Research on the Bearing Capacity of Foundations," *Canadian Geotechnical Journal*, vol. 1, no. 1, pp. 16–26, Sep. 1963, <https://doi.org/10.1139/t63-003>.
- [30] A. S. Vesic, "Analysis of Ultimate Loads of Shallow Foundations," *Journal of the Soil Mechanics and Foundations Division*, vol. 99, no. 1, pp. 45–73, Jan. 1973, <https://doi.org/10.1061/JSEFAQ.0001846>.
- [31] J. B. Hansen, "A General Formula for Bearing Capacity," *Ingeniøren - International Edition*, 1961.
- [32] A. Al-Tabbaa, "Model tests of footings above shallow cavities," *Ground Engineering*, vol. 22, no. 7, pp. 39–42, Oct. 1989.

# Aquifer Vulnerability Assessment by DRASTIC and SI methods: The case study of Guelma Plain, Northeast Algeria

**Houria Baazi**

Natural Hazards and Territory Planning Laboratory (LRNAT), Institute of Earth Sciences and Universe, Geography and Territory Planning Department, University of Batna 2, Algeria

[h.baazi@univ-batna2.dz](mailto:h.baazi@univ-batna2.dz)

(corresponding author)

Received: 27 November 2022 | Revised: 9 December 2022 | Accepted: 18 December 2022

## ABSTRACT

Guelma plain, located in northern Algeria, contains an alluvial nappe that provides economic opportunities such as industrial and agricultural employment. In Guelma, the pressure of anthropogenic origin is combined with unfavorable climatological conditions that expose it to environmental and social risks. The current study aims to assess the level of groundwater vulnerability to pollution using two methods, DRASTIC and SI. These are coupled with the Arc Gis software in order to carry out the mapping. From south to north, several levels of vulnerability have been identified, from low to high. The result comparison allows us to conclude that the SI method is the most appropriate for assessing the vulnerability to pollution of the study region's wetlands. The result of this study can be used as a decision-support tool for the management and protection of the quality of water resources in the Guelma plain.

**Keywords:** Guelma; groundwater; nappe; vulnerability; pollution; DRASTIC; SI

## I. INTRODUCTION

Water is one of the most important ecosystem, climate, and social components [1, 2]. Water in sufficient quantity and of acceptable quality is a fundamental necessity, but maintaining its quality is challenging [3]. Groundwater is severely overexploited because it is the main source of water for drinking, domestic, industrial purposes, and livestock operations, and it is in a critical state based on quantitative and qualitative evaluations [4-6]. Surface and ground water are the main resources of drinkable water, since the 97.5% of the total water on the globe is saline. Civilizations have long promoted an understanding of the need of protecting water resources. Technical methods are crucial for conserving these resources. These tools are based on evaluations of the sensitivity and susceptibility of water to pollution, as well as a compilation of maps that clearly delineate zones of sensitivity and vulnerability at various levels. The drop in groundwater level and ensuing rise in susceptibility are caused by human and agricultural activities. This study sought to ascertain Guelma alluvial plain's sensitivity to groundwater pollution using the following vulnerability methods: the general DRASTIC [7] approach and the SI [8] method combined with GIS. In order to establish measures to conserve these resources and prevent the loss of this wealth, the created maps of groundwater pollution sensitivity using various methodologies help us better pinpoint the location of polluted areas.

## II. STUDY AREA

The wilaya of Guelma is located in the North-East of the country. It is limited by the wilaya of Annaba in the North, El Tarf in the North East, Skikda in the North-West, Souk Ahras and Oum El-Bouaghi in the South East, and finally Constantine in the West. It is a part of the Maghreb chain of eastern Algeria [9] and extends between 7° and 7°30' East and between 36° and 36°30' North, having between 180 and 250m altitude [10].

This area is characterized by subtropical climate. The rainy season lasts from October until May, with an annual precipitation that doesn't surpass 550mm and with sparse vegetation. In addition, the annual average value of temperature is 17.75°C, and the monthly temperatures variate between a minimum of 9.72°C and a maximum 26.76°C. It is surrounded by various thrust sheets, giving rise to the current mountains [9, 11-13]: Mahouna (1411m) to the south, Haouara (1292m) to the north, and Débar (1060m) to the North-West [10]. The reliefs are made up of allochthonous lands belonging mainly to the Tellian domain made up of Meso-Cenozoic marls and carbonates [14].

## III. GEOLOGICAL AND HYDROGEOLOGICAL SETTINGS

The geology of Guelma region consists of three major sets: pre-aquifer, a Mio-Pliocene set (continental, from Guelma basin), and a recent set (Pliocene and Quaternary). Guelma

region includes Djebel Debagh Neritic domain, Heliopolis, and Guelma South. This Jurassic-Cretaceous carbonate facies unit, which is more or less karstified, is overlain by several thrust layers and subjected to significant tectonic events. Along its course between Nador and Medjez Amar, the Seybouse wadi has deposited alluvium [9]. The center of the plain represents a zone of collapse filled with Miocene (clay and gypsum marl) and Quaternary (terraces of heterogeneous alluvium) deposits (Figure 2) [10]. These permeable alluviums are nourished by the infiltration waters of the rains and the lateral contributions of the Seybouse wadi's watershed. There are three distinct terrace levels (low, medium, and high) [10].

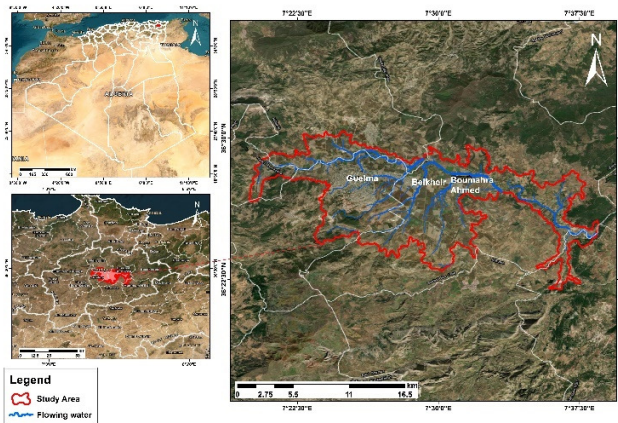


Fig. 1. Study area.

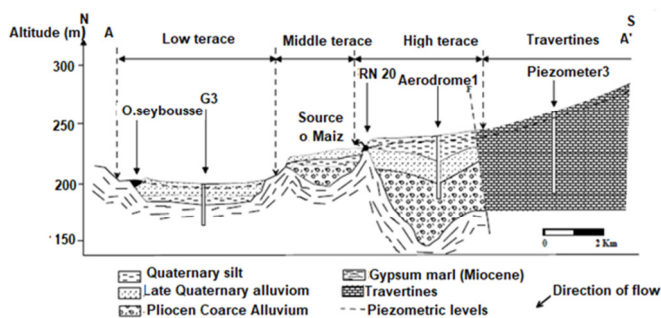


Fig. 2. Hydrogeological section in the Guelma north-south plain.

#### IV. MATERIALS AND METHODS

##### A. Application of the DRASTIC and SI Methods

The DRASTIC index was developed by the US Environmental Protection Agency as a standardized approach for measuring how sensitive groundwater is to pollution [6]. DRASTIC can be utilized by people with a wide range of backgrounds and degrees of competence due to its simplicity [6]. DRASTIC evaluates groundwater vulnerability by incorporating 7 physical and hydrogeological parameters: Depth of water (D), net recharge (R), aquifer media (A), soil media (S), topography (T), vadose zone impact (I), and hydraulic conductivity are the (C) parameters. For convenience of use, they have been organized to form the Abbreviation DRASTIC. The DRASTIC factors are measurable parameters for which data is generally available from a number of sources without extensive investigation. The Susceptibility Index (SI) is

a modification of the DRASTIC approach for determining aquifer sensitivity to disperse agricultural pollution. Depth to water, net recharge, aquifer medium, topography, and land use are the 5 parameters involved. It differs from DRASTIC in that it includes the land-use factor while excluding the soil media, vadose zone, and aquifer. Land use is a critical and influential factor in groundwater contamination caused by anthropogenic activities such as agricultural land use which can negatively impact groundwater through the application of fertilizers, pesticides, and industrial land use, which can negatively impact groundwater through the disposal of chemical compounds into water bodies and landfills [11].

TABLE I. RATINGS AND WEIGHTS OF THE DRASTIC METHOD PARAMETERS [7]

D (m)	Weight	Evaluation
0-1.5	5	10
1.5-4.5		9
4.5-9		7
9-15		5
15-22.5		3
22.5-30		2
> 30		1
A	Weight	Evaluation
Karst limestone	3	10
Sand and gravel		8
Solid sandstone		6
Altered metamorphic		4
Metamorphic		3
Metamorphic		1
Sand and fine gravel		5
T (slope %)	Weight	Evaluation
0-2	1	10
2-6		9
6-12		5
12-18		3
> 18		1
C (m/s)	Weight	Evaluation
$> 9.4 \times 10^{-4}$	3	10
$4.9 \times 10^{-4} - 9.4 \times 10^{-4}$		8
$32.9 \times 10^{-5} - 4.7 \times 10^{-4}$		6
$14.7 \times 10^{-5} - 32.9 \times 10^{-5}$		4
$4.7 \times 10^{-5} - 14.7 \times 10^{-5}$		2
$4.7 \times 10^{-7} - 4.7 \times 10^{-5}$		1
R (mm)	Weight	Rating
>25.5	4	9
25.5-17.5		8
17.5-10		6
10-5		3
5-0		1
S	Weight	Rating
Thin or absent	2	10
Sands		9
Sandy loam		6
Silt		4
Silty loam		3
Clays		1
I	Weight	Rating
Karstic limestone	5	10
Sand and gravel		9
Sand and gravel with silt and clay		8
Sandstone		6
Limestone		6
Silt and clay		1



For each used models we have applied the following formulas:

- DRASTIC:

$$ID = D_r \times D_w + R_r \times R_w + A_r \times A_w + S_r \times S_w + T_r \times T_w + I_r \times I_w + C \times C_p \quad (1)$$

where D: depth, R: net recharge; A: aquifer media, S: soil media, T: Topography, I: unsaturated zone; C: hydraulic conductivity, and the subscripts r and w are the corresponding rating and weights, respectively:

- SI:

$$I_s = D_c \times D_p + R_c \times R_p + A_c \times A_p + S_c \times S_p + T_c \times T_p + lu_c \times lu_p \quad (2)$$

where D, R, A, S, T, and Lu are the parameters of the SI method, p the weight of the parameter, and c the associated score.

The land cover parameter (Lu) was obtained by processing the sentinel satellite image taken in July 2020. This image was used to establish the land cover map of the study area. Tables I-III are used to elaborate the maps of the parameters of the DRASTIC and SI methods.

TABLE II. WEIGHTS ASSIGNED TO THE SI METHOD PARAMETERS [8]

N°	Parameters	Weight
1	D: Depth	0.186
2	R: Net recharge	0.212
3	A: Aquifer media	0.259
4	T: Topography	0.121
5	LU: Land Use	0.222

TABLE III. VULNERABILITY ASSESSMENT CRITERIA OF THE DRASTIC AND SI METHODS [7,12,13,8]

Vulnerability index	DRASTIC	SI
Very low	< 80	---
low	180-120	<45
Moderate	121-160	45-64
High	161-200	65-85
Very high	>200	>85

### B. Elaboration of Maps of the Specific Parameters of the Vulnerability of the Aquifer

In order to use the DRASTIC and the SI methods, the following steps must be followed:

- Data collection
- Data digitization
- Calculation of scores (Score  $\times$  Weight)
- Assignment of scores
- Classification and reclassification of the 7 specific parameters of the groundwater
- Map development for each specific parameter of the tablecloth on Raster mode

- Calculation of the global vulnerability index by the sum of the 7 raster maps of the specific parameters of the vulnerability of the tablecloth
- Classification and development of the final vulnerability map

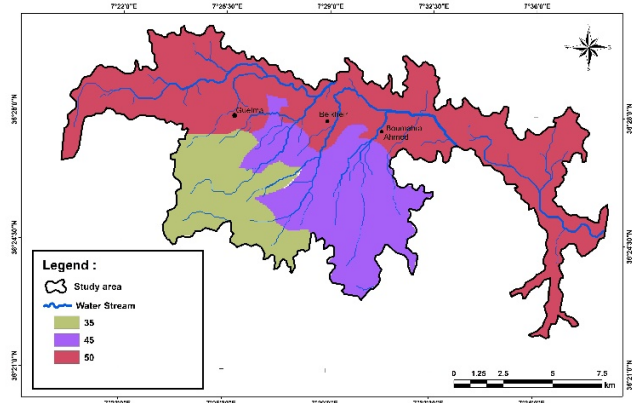


Fig. 3. Map spatial distribution of depth to water.

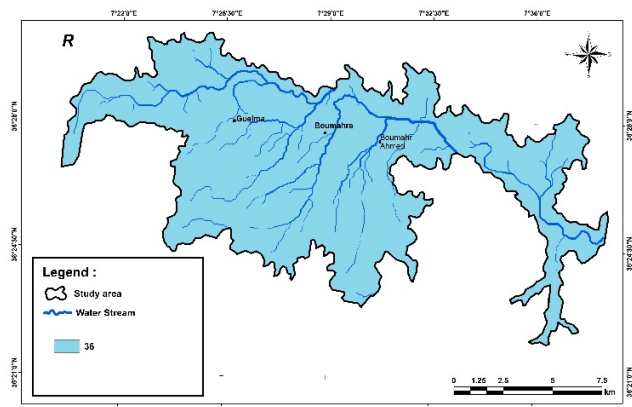


Fig. 4. Spatial distribution of net recharge.

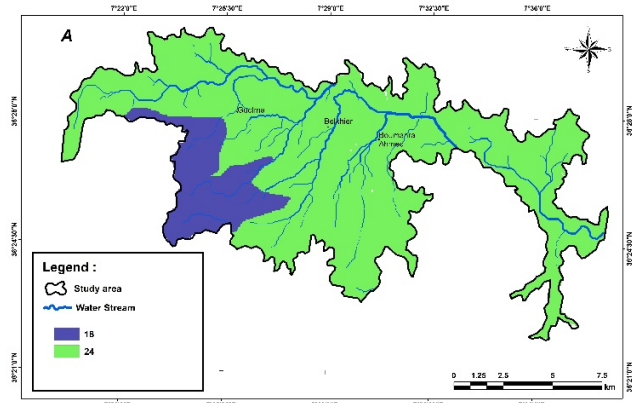


Fig. 5. Spatial distribution of aquifer materials.

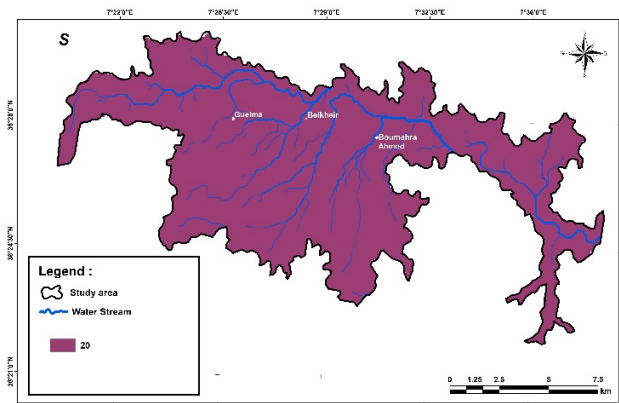


Fig. 6. Slope spatial repair.

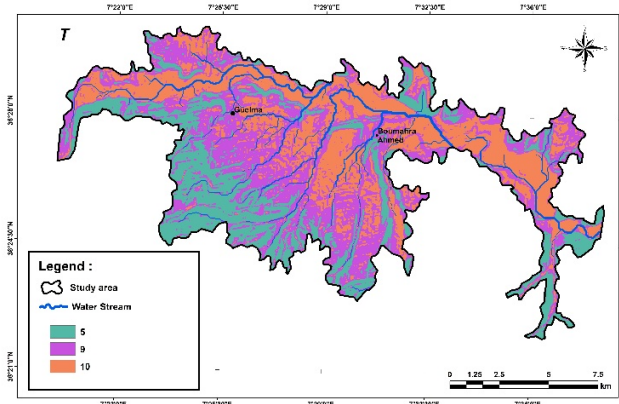


Fig. 7. Spatial distribution of soil type in the region.

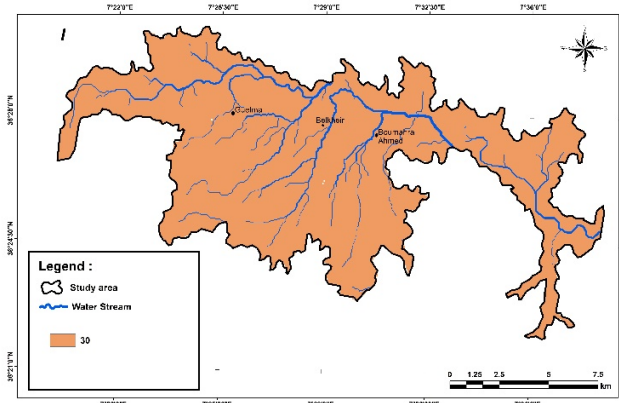


Fig. 8. Distribution of values of the nature of the vadose zone index.

## V. RESULTS AND DISCUSSION

### A. Synthesis Map by the DRASTIC Method

The application of the DRASTIC method at the Guelma plain has allowed the realization of the vulnerability map of Figure 11, from which we can distinguish three classes of different degrees of vulnerability. We notice that the areas of low vulnerability are encountered in the southeast of the plain. This can be explained by the nature of the soil that is characterized by a fine texture and the slope that varies from

medium to strong. Similarly, the areas of medium vulnerability are located in the center of the plain to the south, more precisely in the southeast or at the wastewater discharges (drainage of wastewater by the various tributaries of Oued Seybouse) and at the level of settlements which contain a large number of boreholes for agricultural development.

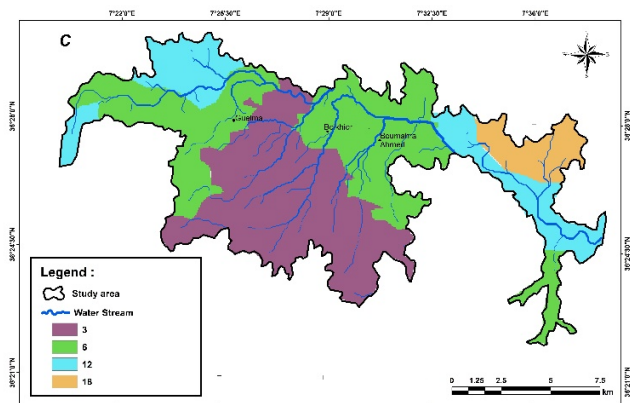


Fig. 9. Distribution of hydraulic conductivity values.

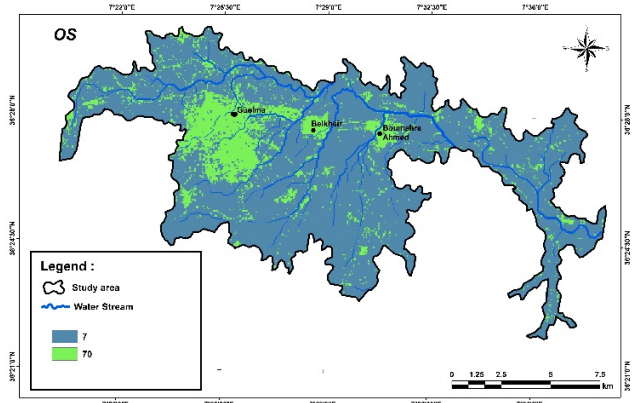


Fig. 10. Land cover map.

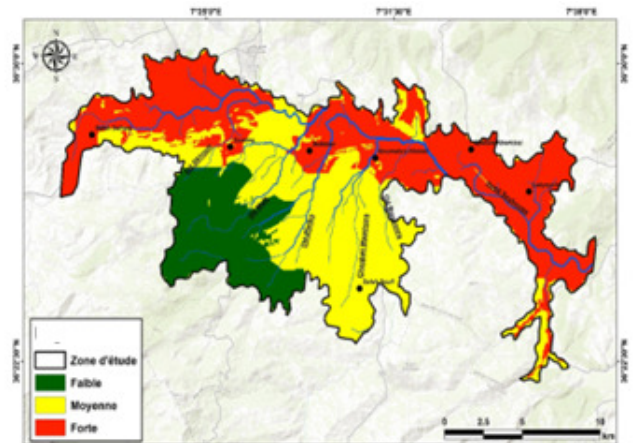


Fig. 11. Aquifer vulnerability map (DRASTIC method).

The area of high vulnerability is noticed at the North-East and North-South edges, which can be explained by the drainage of wastewater by the main wadi (Seybouse) interrupted by a site of medium vulnerability. This elevation is explained by the gravelly nature of the soil of the vadose zone, the materials of the aquifer which are very permeable, and the urban discharges.

### B. Synthesis Map by the SI method

The application of the SI method at the Guelma plain has allowed the realization of the vulnerability map of Figure 12, from which we can again distinguish three different classes of degrees of vulnerability. The areas of low vulnerability are distributed throughout the plain but the large area of this class is located in the south, affirming the result found in the case of the DRASTIC method, which is explained by the nature of the soil at an important depth that can reach 25m. A sector of medium vulnerability is distributed all over the plain, where the wastewater discharges are located and at the level of the agglomerations. The area of high vulnerability is mainly noticed all along the wadi Seybouse (drainage of wastewater by Oued Seybouse) and even in the center going south, where there is the presence of an unauthorized dump site and in the industrial area between Oued Skhoun and Oued Maiz.

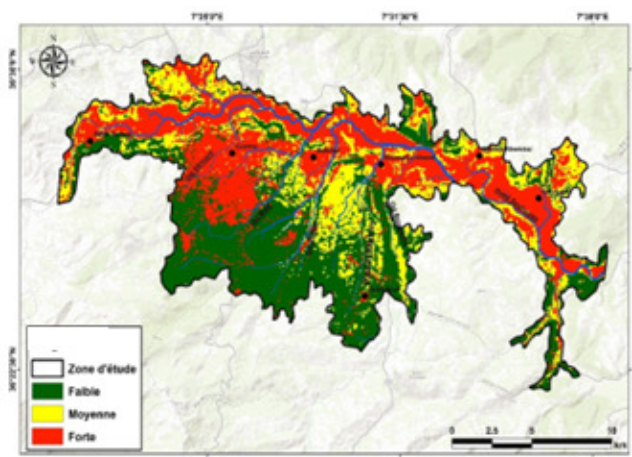


Fig. 12. Aquifer vulnerability map (SI method).

## VI. CONCLUSION

Surface and ground water are the main resources of drinkable water [14]. With the ever-increasing demand for water and the increasing lack and deterioration of water quality, the need for vulnerability mapping methods is necessary. Due to the large number of methods in use, it is essential to recognize which ones are the most reliable and effective.

The plain of Guelma is made up of agricultural land that is irrigated from the Seybouse wadi, which is the receiving medium for all the urban and industrial wastewater from the city of Guelma and its agglomerations. The contamination of the plain's groundwater has been brought to light by this study. In order to achieve this goal, the application of the DRASTIC and SI methods allowed the vulnerability assessment of the pollution of the Guelma alluvial groundwater. This

vulnerability increases from the center of the plain to the northern edges and especially along the Oued Seybouse and its tributaries (Oued Maiz, Oued Zimba, Chaabet Maamora). The distribution of pollution sources has induced this growth. The analysis of the two vulnerability maps resulting from the application of these methods revealed three distinct classes of low, medium, and high vulnerability. Although the two vulnerability maps are relatively close, the vulnerability map made from the SI method reflects more the reality of the groundwater pollution of the Guelma plain, which can be explained by the high vulnerability area located in the southern part (presence of an industrial area and an unauthorized dump site). As the pollution of groundwater in the study area is confirmed, consequently the most vulnerable perimeters from now on must be subjected to a rigorous control in order to define protection measures to avoid and stop the degradation of this natural source.

## REFERENCES

- [1] M. A. Keerio, N. Bhatti, S. R. Samo, A. Saand, and A. A. Bhuriro, "Ground Water Quality Assessment of Daur Taluka, Shaheed Benazir Abad," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2785–2789, Apr. 2018, <https://doi.org/10.48084/etasr.1925>.
- [2] N. V. Duc, "Experimental Water Quality Analysis from the Use of High Sulfuric Fly Ash as Base Course Material for Road Building," *Engineering, Technology & Applied Science Research*, vol. 9, no. 5, pp. 4627–4630, Oct. 2019, <https://doi.org/10.48084/etasr.2991>.
- [3] K. Praveen and L. B. Roy, "Assessment of Groundwater Quality Using Water Quality Indices: A Case Study of Paliganj Distributary, Bihar, India," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8199–8203, Feb. 2022, <https://doi.org/10.48084/etasr.4696>.
- [4] S. S. Singha, S. Pasupuleti, S. Singha, R. Singh, and A. S. Venkatesh, "A GIS-based modified DRASTIC approach for geospatial modeling of groundwater vulnerability and pollution risk mapping in Korba district, Central India," *Environmental Earth Sciences*, vol. 78, no. 21, Oct. 2019, Art. no. 628, <https://doi.org/10.1007/s12665-019-8640-2>.
- [5] UNESCO World Water Assessment Programme, *The United Nations World Water Development Report 2022: groundwater: making the invisible visible*. UNESCO, 2022.
- [6] C. Denizman, "Land use changes and groundwater quality in Florida," *Applied Water Science*, vol. 8, no. 5, Aug. 2018, Art. no. 134, <https://doi.org/10.1007/s13201-018-0776-9>.
- [7] L. Aller, T. Bennet, J. H. Lehr, and R. J. Petty, *DRASTIC: A standardized system for evaluating ground water pollution potential using hydrogeologic settings*. Ada, OK, USA: EPA, 1985.
- [8] L. Ribeiro, *SI: a new index of aquifer susceptibility to agricultural pollution*. Lisboa, Portugal: ERS/HA/CVRM, Instituto Superior Técnico, 2000.
- [9] J.-M. Vila, "La chaîne alpine d'Algérie orientale et des confins algéro-tunisien," Ph.D. dissertation, Université Pierre et Marie Curie, Paris, France, 1980.
- [10] B. K., C. A., and B. L., "Variations spatiales des caractéristiques physico-chimiques des eaux souterraines de la plaine alluvionnaire de Guelma, (Nord-est algérien)," *Bulletin des sciences géographiques de l'INCT*, vol. 18, no. 1, pp. 52–62, Sep. 2014.
- [11] J.-M. Vila, R. Busnardo, A. Devries, J. Magne, and J. Sigal, "Données Stratigraphiques Sur La Série Renversée Et Charriée Du Djebel Bou Sba Et Étude De Son Cadre Structural (région De Guelma, Constantinois, Algérie)," 1968.
- [12] J. F. Raoult, "Chevauchement d'âge éocène dans la dorsale du Djebel Bou Aded (Est de la chaîne numidique, Algérie)," *Comptes Rendus de l'Académie des Sciences*, vol. 266, pp. 861–864, 1975.

- [13] J.-C. Lahondère, "Les séries ultratelliennes d'Algérie nord-orientale et les formations environnantes dans leur cadre structural," Ph.D. dissertation, Université Paul Sabatier, Toulouse, France, 1987.
- [14] *Etude géophysique de la plaine de Guelma. Rapport interne.* Algérie Geophysique, 2004.
- [15] T. Nouar, "Impact de la pollution sur les ressources en eau superficielles et souterraines de la région de Guelma," Ph.D. dissertation, Université Badji Mokhtar de Annaba, Annaba, Algeria, 2007.
- [16] N. Brahmi, "Evaluation et gestion des ressources en eau dans le bassin versant de la moyenne Seybouse," Ph.D. dissertation, Université Badji Mokhtar de Annaba, Annaba, Algeria, 2009.
- [17] M. Anane, B. Abidi, F. Lachaal, A. Limam, and S. Jellali, "GIS-based DRASTIC, Pesticide DRASTIC and the Susceptibility Index (SI): comparative study for evaluation of pollution potential in the Nabeul-Hammamet shallow aquifer, Tunisia," *Hydrogeology Journal*, vol. 21, no. 3, pp. 715–731, May 2013, <https://doi.org/10.1007/s10040-013-0952-9>.
- [18] A. Corniello, D. Ducci, and P. Napolitano, "Comparison between parametric methods to evaluate aquifer pollution vulnerability using a GIS: An example in the Piana Campana, Southern Italy," presented at the International Symposium on Engineering Geology and the Environment, Athens, Greece, 1997, pp. 1721–1726.
- [19] M. J. Focazio, T. E. Reilly, M. G. Rupert, and D. R. Helsel, "Assessing ground-water vulnerability to contamination: Providing scientifically defensible information for decision makers," U.S. Geological Survey, USGS Numbered Series 1224, 2002, <https://doi.org/10.3133/cir1224>.
- [20] N. Kumar, A. A. Mahessar, S. A. Memon, K. Ansari, and A. L. Qureshi, "Impact Assessment of Groundwater Quality using WQI and Geospatial tools: A Case Study of Islamkot, Tharparkar, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5288–5294, Feb. 2020, <https://doi.org/10.48084/etasr.3289>.



# Substation Danger Sign Detection and Recognition using Convolutional Neural Networks

**Wajid Ali**

School of Automation and Electrical Engineering, Lanzhou Jiaotong University, China  
wajidaliustb@gmail.com

**Guo Wang**

School of Automation and Electrical Engineering, Lanzhou Jiaotong University, China  
wangguo@lztu.edu.cn  
(corresponding author)

**Kalim Ullah**

University of Science and Technology Bannu, Pakistan  
engr\_kalim125@yahoo.com

**Muhammad Salman**

School of Automation and Electrical Engineering, Lanzhou Jiaotong University, China  
salmankhan5780@gmail.com

**Sajad Ali**

School of Automation and Electrical Engineering, Lanzhou Jiaotong University, China  
sajad8371@gmail.com

*Received: 14 November 2022 | Revised: 7 December 2022 | Accepted: 15 December 2022*

## ABSTRACT

This paper focuses on the training of a deep neural network regarding danger sign detection and recognition in a substation. It involved applying the concepts of neural networks and computer vision to achieve results similar to traffic sign and number plate detection systems. The input data were captured in three distinct formats, i.e. grayscale, RGB, and YCbCr, which have been used as a base for comparison in this paper. The efficiency of the neural network was tested on a unique data set involving danger signs present in industrial and processing facilities. The data set was unique, consisting of four distinct symbols. The trained data were selected so that they would not facilitate overfitting and also would not be under fitted. The accuracy of the model varied with the input type and was tested with two distinct classifiers, CNN and SVM, and the results were compared. The model was designed to be fast and accurate, and it can be implemented on mobile devices.

**Keywords-**CNN; neural networks; danger sign detection

## I. INTRODUCTION

Signs play an important role in everyday life and come in a huge variety of color schemes [1], symbols, and words, originating from different regions, countries, and working environments. The main objective of this study was to recognize four different and distinct types of signs used in the industrial process worldwide. Automatic danger sign detection can help prevent accidents that could endanger employees and

civilians. The presence of an application on a smartphone that can recognize and warn about the meaning and danger associated with a sign can save lives. It is an ideal device for new and less experienced workers. Nowadays, almost every person has a mobile phone, and about 80% of these are smartphones, able to integrate with the internet and run complex applications [2-4]. A neural network-based application can easily run on this type of phone. The target audience of this research is huge. It can benefit seasonal

workers, new industrial recruits, engineers, etc. The study focuses on using Neural Networks (NNs) to distinguish several danger signals. The application will be entirely web-based, having a mobile application to integrate with the user and even a web page for a direct online interface. The world has become heavily reliant on IT, and the COVID19 pandemic has shown that the potential of mobile applications and software is at its peak. A danger sign detection technique can be used in any application. The impacting factors are the sample size and training parameters, which have been tuned to give promising results and can be generalized to other symbols with slight modification and testing. The application shows it can accurately identify and classify data and warn people by recommending the prescribed precautions. This information is easy and can be quickly accessed on a phone.

The database used in this study was compiled from scratch, as no dataset of this kind existed. This involved extracting images from online repositories and personal sight visits. Image modification techniques were also used to add diversity and unpredictability to the dataset. The data set can be consistently updated and improved by adding more images. The more the images in the data set, the more accurate the training and testing will be. Although this paper is focused on the danger signs found explicitly in the substation, the same techniques could be used with other data sets to recognize danger signs in different environments. The concept involves computer vision and NN integration. The main objective is to replicate an accurate and similar response to that of traffic sign detection and number plate identification. They both use similar tools and techniques. While extensive literature exists in the field of traffic sign detection, studies related to the identification of signs in other environments are limited at best and non-existent to the best of our knowledge. This study aims to fill this scientific gap.

## II. BASIC ABOUT CNNs

A typical Convolutional Neural Network (CNN) contains pooling, activation, normalization function, convolutional and Multinomial Logical Regression (MLR) layers. Usually, the last layer of the network is the MLR. This layer has a subpart known as the softmax function, which determines the final classification of data. The structure of the CNN model is chosen according to the difficulty and the problem that needs to be solved. The CNN model is designed with specific layers and neurons in mind. We can aim for higher accuracy for our results, but this will require more training time, a larger data set, and a more complex network structure [5].

## III. RELATED WORK

The ongoing progress in the field of NNs includes improvements in the processing power required, reduction of the complexity of the algorithm, and advancements in computer vision. The closest mainstream topic related to computer vision, CNN, and image identification, is traffic sign or number plate detection. Many papers have been published regarding traffic sign detection with various techniques and methods [1, 6-8]. Most of these studies have existing data sets. These sets are classified into groups and are ready to be fed into the NN algorithm. In the current study, we had to make

and classify the data and set and test the identifying markers and check their accuracy for our algorithm from scratch.

## IV. CONVOLUTIONAL NEURAL NETWORKS

NNs are often used to train and identify objects. These can be images, trends, anything that can be quantized in the form of data. Though CNNs are a powerful tool, they require high computational power and resources to function properly [9]. As microelectronics gets smaller, it makes nano electronics more prevalent and viable in the electronic landscape. Accurate NN models require several layers, increasing the algorithm's complexity, the computational power required, and the energy needed to execute and process it. Fortunately, the semiconductor industry has made this possible, providing handheld devices with the power of a modest computer. Researchers have tried bringing deep learning models implemented in CNNs to mobile devices. These belong to one of the three following categories:

- Network pruning [10]
- Network quantization [11]
- Training small networks directly.

Finding the ideal number of layers and neurons required for the accurate and efficient working of the network is necessary. Overfitting, excess complexity, and several hidden layers can make the model accurate regarding the corresponding training data but sensitive to change, lacking robustness. The effects of overfitting and its correlation with robustness, validation, and training data were studied in [12].

There have been improvements in different techniques and methodologies to advance the techniques used in NNs, such as [13, 14]. CNNs have strong discrimination powers in terms of classifying objects. They can also accurately localize a portion of an image or data cluster using tools such as regression. In our efforts to deploy a model of CNN compatible with image extraction, we took inspiration from literature such as [14-16]. The main methodology applied in recent literature for detection networks are:

- Direct detection [14, 15]
- Regional detection [13, 17]

### A. Direct Detection

The direct detection and prediction algorithms are based on the position or class of the target object. This information is provided with either regression or classification, depending on the problem and type of learning involved [18]. This can be extracted from the convolutional layer of the algorithm, making runtime faster in comparison.

### B. Regional Detection

In regional detection, several regions are generated for the data in question, then data classes, and then the prediction operation is performed on object position and class for every region. In this method, regression and classification are conducted twice in different stages. This makes the regional detection method longer and more processing intensive, but more accurate in the computation of direct detection.



For our purpose of danger sign identification and recognition, the regional detection method is preferred, as accurate identification of hazardous symbols can help prevent accidents and provide precautionary safety information for workers. There is no processing and time restraint, thus, the delay is an acceptable trade-off for higher accuracy.

## V. BOUNDARY DETECTION TECHNIQUES

Many CNN object identifiers use boundary boxes and class labels to identify and extract the exact image for learning [19, 20]. This can be useful in predicting the exact boundary boxes for the classified images. A danger sign varies depending on the type of hazard. Some are triangular, while others are square or circular. These distinct and different bound conditions for the sign make image extraction significantly more accurate. We can identify an image using its boundary estimation, segmentation of the image itself, or sign and symbol category classification. The pixel segmentation mask would require high processing power and would be slower in identifying the letters and words shown on the symbol. However, this is possible and has many potential applications, such as universal danger sign detection software [20]. This could be compatible with all languages and work environments. Such ideas have been forgone, as failure to identify a single letter in the sign can cause false or inconclusive results. Thus the main techniques used were boundary detection and symbol recognition.

## VI. DEEP NEURAL NETWORKS

Deep Neural Networks (DNNs) have become an indispensable tool in image classification used in many applications, such as facial recognition, voice analysis, word or text analysis, etc. Using DNN models helps in achieving high accuracy because of their ability to train and recognize complex input data [21]. The definition of a DNN is defined as a network with a level which is represented by the presence of three or more hidden layers.

## VII. CONVOLUTIONAL NEURAL NETWORKS-SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a supervised model in machine learning [22] that needs to be given labelled input data and a classification algorithm is used to train itself for each given label. Its main advantage lies in its speed and higher accuracy with a lower input sample count. This can make this algorithm ideal for a task requiring robust classification [23-25]. The SVM technique is used mostly as a classifier in data evaluation. The ability of SVM in data analysis can help find the ideal margin between two distinct classifications. SVM is used to determine a hyperplane for each set of distinct outcomes from a given input. If the hyper plane is situated ideally between the trained data sets, it can provide a degree of robustness to the system. SVM is used in traffic signal recognition and training. The basic CNN is used for convolution and training, while the Support Vector Classifier (SVC) is used for classification. The hybrid combination of SCV and CNN is combined to give accurate results for a small number of training sets [26].

## VIII. METHODOLOGY

The training of the system involved five basic steps:

1. Making a data set appropriate for integration.
2. Deciding the most appropriate and accurate type of NN to choose.
3. Training strategy used and data optimization in order to acquire the best results.
4. Model selection involves the layers and neurons in the algorithm, basically determining the complexity of the algorithm and its robustness.
5. The model in question used for the training of our model has 3 convolutional layers followed up by a pooling layer each. The end results of the layers are fed into a flattened layer to convert a 3-dimensional image into 1-dimensional weights. These weights are used in the deep layer to find the weight coefficients to recognize and categorize the image into one of 4 different categories.

The first step was taking in the respective input format. The images captured for processing could be used in one of three ways:

- Grayscale [27]
- RGB [28]
- YCbCr [29]

## IX. IMAGE ANALYSIS

### A. RGB

The RGB data set is the default spectrum used as input for most image recognition applications. Though it is thoroughly used, our results indicated that a modification of RGB, known as YCbCr, gives better results. The luminance information, known as Y, is scattered in all the dimensions of RGB. This information can be used to get better image parameters for training data. Below is a simple conversion from RGB to YCbCr [31].

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

$$Cb = 0.564(B - Y) \quad (2)$$

$$Cr = 0.713(R - Y) \quad (3)$$

### B. YCbCr

The respective training error was the least in YCbCr results for the limited training data set. The YCbCr is inherently superior to grayscale and RGB, as in grayscale many colors and contact information are lost and therefore are not taken into account while training. Any colored image can be converted into a grayscale by multiplying the colored coefficients (usually RGB) with constants. This converts 3-dimensional data into a 1-dimensional data set. Grayscale results were faster to train but had a training error twice that of the YCbCr data set [30]. The model was trained with the same image data but with different feature extraction.

### C. Grayscale

The following converts an RGB image to grayscale:

$$\text{Grayscale} = 0.299R + 0.587G + 0.114B \quad (4)$$

The grayscale image is less complex as the data can only vary in one parameter. This can help make the data set more robust and less prone to overfitting, though much of the image information is lost when converting an image into grayscale data. The decrease in complexity did not adversely affect the results. Quite the contrary, the grayscale results proved better than the RGB.

## X. TRAINING INFORMATION

The training process involved using four distinct image types for categorization. These were:

- Electric hazard
- Flammable risk
- Radiation danger
- Toxic materials



Fig. 1. Image categories.

We chose similar image shapes as already much work has been done on boundary identification markers. Instead of focusing on the shiny shape, we trained the CNN to identify the color pattern and the symbols present [32]. Each image was similar to the others to an extent, e.g. similar sign shape and base color contrast (yellow and black). Thus, the main objective was convolving the image and using the DNN nodes to recognize the distinct shapes.

## XI. IMAGE STANDARDIZATION

We used bounding box regression to standardize the extracted image from the input. It can be taken from an angle and still correctly extract the image [14]. The regression vector can compress or expand the image to the ideal size for data input. We used shape classification for images with different shapes. Most images are of similar shape, including circular, triangular, rectangular, square, hexagonal, and octagonal. For the ease of data usage, we normalized the coordinates of both the x and y-axis. We can use this to find the exact extraction of all shapes except circular images and in our data set, there were no circular image boundaries, so that was not a major concern.

Circular images are detected with small sheer errors, though the CNN-trained system was still able to recognize the symbols and give accurate results.

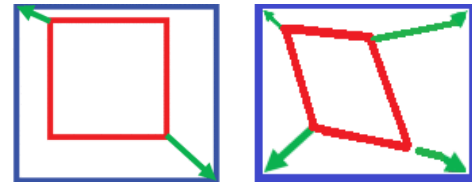


Fig. 2. Image fitting using regression vector.

## XII. MODEL OVERVIEW

The model was trained on the gathered images that made a completely new and authentic data set. The network parameters were tuned during training and validation. ReLu, Softmax Deep CNN in a combination with SVM were used. Figure 3 shows some randomly selected images from the data set.



Fig. 3. Samples of the training set.

The data used for training consisted of 425 images, and the testing set included twice that number. The convolutional layers were used to segment the image into 64 blocks, each for a different color, for both RGB and YCbCr sets. The network architecture is shown in Figure 4. We used 3 convolutional layers for quick response. More can be incorporated to increase the accuracy of the results. The training images were transformed into YCbCr-colored data.

$$\text{Feature map size} = N - F + 1 \quad (5)$$

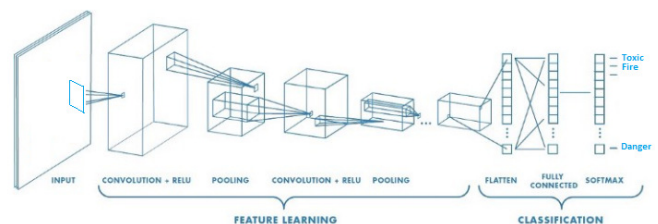


Fig. 4. Network architecture.

### XIII. DETAILED WORKING OF THE MODEL

In our training, we passed 3 convolutional layers followed by max pooling. The input image was sized at 64×64 pixels. After passing through the convolution and ReLu layer, our resulting image was a 62×62-pixel image and our kernel was 3×3. We extracted 16 feature maps in these two stages. We have monitored and identified 448 parameters in this step. Then, max-pooling by a 2×2 pool size and 2 unit stride took place, and the processed image reduced in size to 31×31 pixels. This input image was fed to the second convolution layer, resulting in a 29×29-pixel image, granting 32 feature maps and 4046 parameters. The next max pooling layer resulted in a 14×14-pixel image. Pool size, kernel, and stride were kept constant for all convolution and pooling layers. By the last convolution layer, we had 64 feature maps of a 6×6-pixel image. After, a flatten operator was used to convert all the data into a linear array which passed as weights to the neurons. Our model has a total of 388 neurons, 256 in the 1<sup>st</sup> layer, 128 in the 2<sup>nd</sup> layer, and 4 in the last layer for the respective outputs. Figures 5-6 show the features of the modeled system.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 62, 62, 16)	448
max_pooling2d (MaxPooling2D)	(None, 31, 31, 16)	0
conv2d_1 (Conv2D)	(None, 29, 29, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 14, 14, 32)	0
conv2d_2 (Conv2D)	(None, 12, 12, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 6, 6, 64)	0
flatten (Flatten)	(None, 2304)	0
dense (Dense)	(None, 256)	590880
dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 4)	516

Fig. 5. CNN model data.

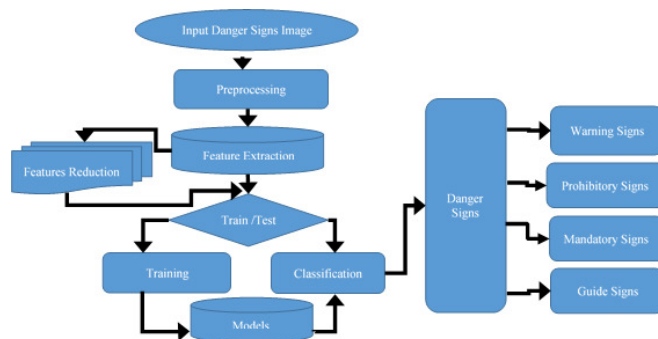


Fig. 6. Flow diagram.

### XIV. SVM

The system was tested using both SVM and CNN, and both linear and non-linear SVM types were tested. Linear SVM uses a dot product between two features, and its product is the linear product combination of the input features.

$$f(ax_n, bx_m) = f(ax_n) + f(bx_m) \quad (6)$$

In non-linear SVM, this rule is not valid. Instead, it uses Kernel. This reduces the dimensionality of the input into various classifiers, replacing the dot product of linear SVM. This SVM can be used to learn shapes in higher dimensional data when using the Kernel.

### XV. NEURAL NETWORK

An NN with more than one layer is non-linear, and there are 3 main activation functions used in such cases, which are:

$$\text{Logistic function } z = \sigma y \quad (7)$$

$$\text{Hyperbolic tangent } z = \tanh y \quad (8)$$

$$\text{Softmax } \phi(y) = \frac{e^{y_i}}{\sum_{i=1}^k e^{y_i}} \quad (9)$$

### XVI. COMPARISON

For our study, we used the Softmax function. A basic understanding of NNs is necessary for the classification of data and while comparing SVM and NN. In contrast, SVM was a multipurpose non-linear NN designed for classifying the problem. Thus we based our classification approach on the NN. Another reason for choosing SVM was the lack of a large data set. SVM is simpler than the complex NN and has given better results on smaller data sets. Thus our CNN model used the ReLu classifier and our SVM model used the Softmax activation function to clearly identify the effect on filter training and its effect.

### XVII. CNN

In our CNN, we have used ReLu (Rectified Linear Unit) as activation function and the Dropout Layer. The NN is assumed to have:

- Linear independence in the input features.
- Low dimensionality in the input space.

By implementing various convolution layers and pooling, we lowered the dimensionality of the image. This helps us resolve the problem of dimensionality by reducing the data size and making learning easier. The purpose of an activation function is to provide non-linearity, but we don't want any negative activation function because that would not assume that the data are independent. We used ReLu as our activation function because it is positive and very easy and simple to model [33]. It is easy to calculate and involve the only comparison between the input and 0.

$$\text{ReLU: } f(x) = \max(0, x) \quad (10)$$

This is the last part of feature extraction before it is fed into the hidden layer for training and finding the neuron coefficient weights. The weights are necessary for classification as they

assign the neuron their values which recognize features and result in the identification and classification of an image. For a probabilistic result, we used the Softmax activation function [34]. This will give us as output one of the four considered classes each time.

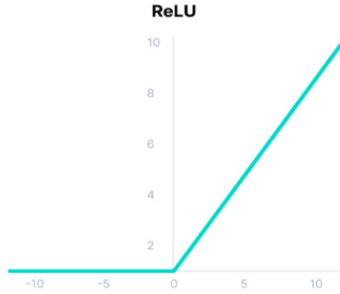


Fig. 7. ReLU function.

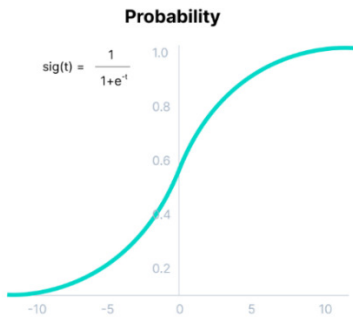


Fig. 8. Softmax function.

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_i \exp(z_i)} \quad (11)$$

For loss calculation, we used Categorical Cross-Entropy loss [35]. This gives us the error caused due to training the data set.

$$\text{CE} = -\sum_i^c t_i \log(f(s)_i) \quad (12)$$

$$\text{CE} = -\sum_{i=1}^{c'-2} t_i \log(f(s)_i) = -t_1 \log(f(s)_1) - (1 - t_1) \log(1 - f(s)_1) \quad (13)$$

## XVIII. EXPERIMENTAL RESULTS

Figure 9 shows the training accuracy and loss of the respective data. The model was run for 100 epochs, approaching a maximum accuracy of 91%. The loss is approaching a value close to 8%, signified by the value of 0.3. The grayscale and the YCbCr sets used the CNN classifier, but the RGB data set were modified to use the SVM classifier. The results showed that the RGB classifier did not increase by any significant value after 100 epochs, maybe due to the data limitation. Still, the validation accuracy was very dynamic and fluctuated consistently. In contrast, both grayscale and YCbCr dataset verification accuracy increased and showed a much smoother increase with a few disturbances and small dynamic behaviors, providing concrete proof that the CNN classifier is better and more accurate than SVM for smaller data sets.

The results show that the model's training and validation accuracy are consistently increasing with a slight offset. This is the ideal result required. It is trained on the modal parameters while still having a significant variation to be considered robust. Thus we conclude that the YCbCr model is well-trained and has a high validation accuracy to be considered robust and accurate. The loss graph shows the training and validation losses. It is based on (12) and (13). The cross-entropy loss is used to distinguish two distinct probability distributions. We can see this as proof of the model's ability to recognize and identify the images. The validation loss function traces a similar path in a logarithmic pattern. This is all consistent with accurate and reliable results.

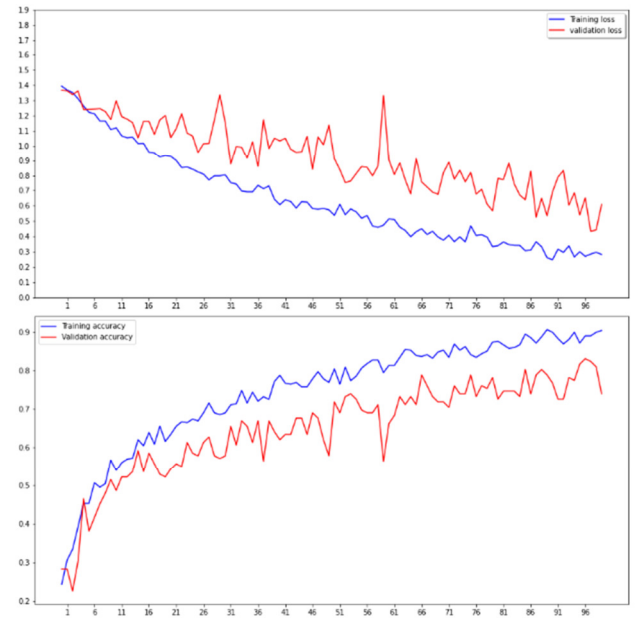


Fig. 9. Training results and loss comparison (YCbCr 100 epochs). Up: loss, down: accuracy.

Similarly, we trained the model with RGB and grayscale images. The results are apparent in Figure 10. It is clear that the RGB data are under-fitted. In the literature, it has been observed that using SVM as a classifier with a small or limited data pool will result in a less accurate and under-fitted plot. The exact pattern is being observed here. Regarding the grayscale results, they are accurate, robust, and ideal. The data are accurately trained and validated, and the loss function is within range. The reason for comparing grayscale and RGB images was to test how the classifier's variation would affect the training and validation results. The RGB dataset used an SVM classifier, while the grayscale dataset used a NN classifier, specifically a CNN classifier. The grayscale data were simpler and less detailed but had better results. As mentioned above, a small dataset could be the reason for low accuracy while using SVM, but the CNN classifier gave ideal results, both with the grayscale data set and YCbCr data. So in our study, the CNN classifier was found to be superior.

Figure 11 shows how increasing training time and duration affected the training accuracy. The YCbCr model validation



data closely followed the training results with a slight set. We increased to training duration to observe the highest accuracy provided by this dataset. We kept all the other parameters, such as the learning rate and the dataset, constant. The training accuracy peak values were achieved at 200 epochs, after which its increase in accuracy was minuscule, while it even started to decrease at 400 epochs. The observed results are consistent with the literature and show that our model's ideal training time is 200 epochs for our specified learning rate. Over-training can result in a loss of accuracy, and this concept is demonstrated here. We also identified a peak accuracy of 99.6%

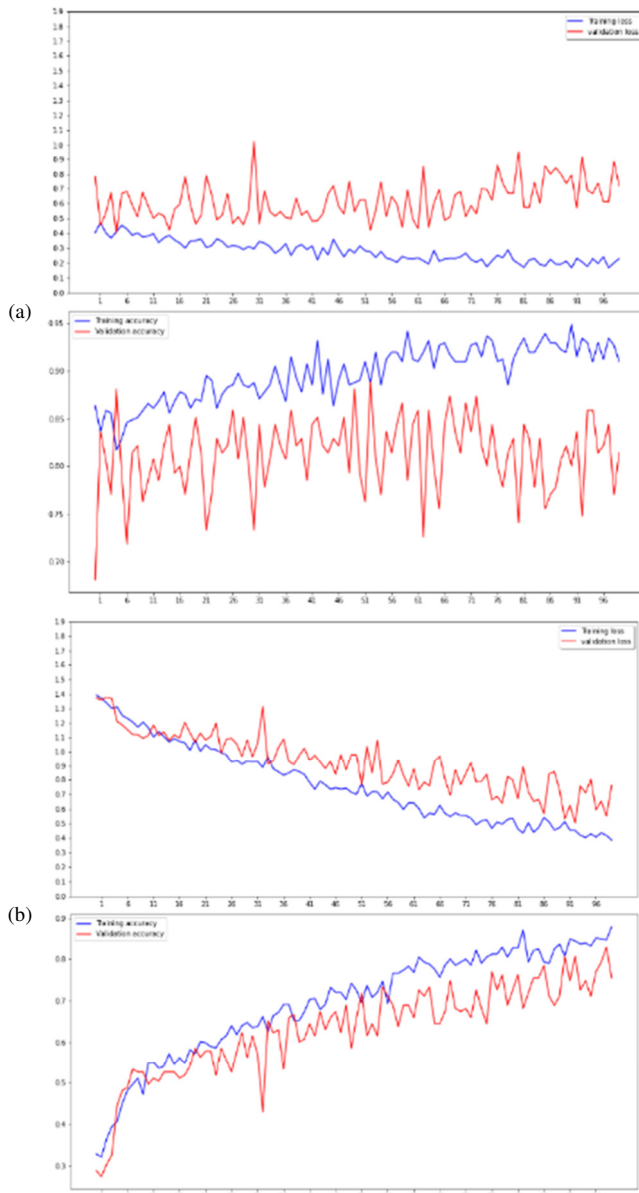


Fig. 10. (a) RGB image training, (b) grayscale image training (100 epochs). Up: loss, down: accuracy.

The effect and duration varied for the number of epochs the network was trained. The result for training and validation

accuracy achieved at 500 epochs for the image dataset are shown in Table I. Only the training duration was varied, and readings at 10, 100, and 500 were plotted. The table values were gathered using the same above-mentioned plots extended to 500 epochs, but a smoothing filter was used in order to make the results more consistent and trends easier to identify.

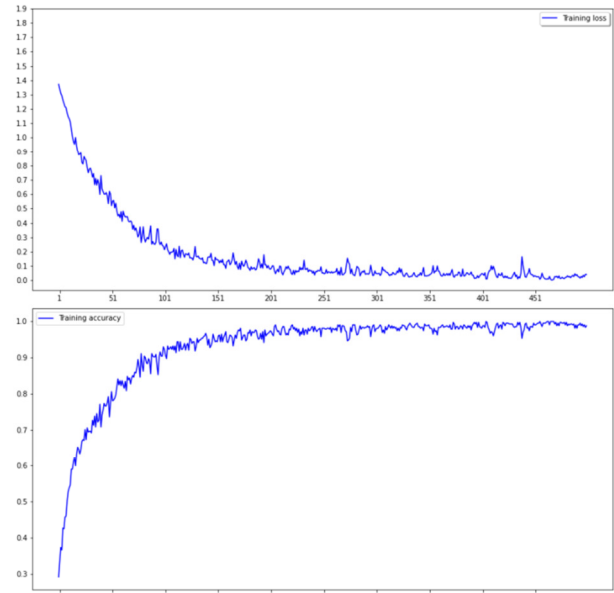


Fig. 11. YCbCr image training for 500 epochs. Up: loss, down: accuracy

TABLE I. RESULTS OF VARIATION OF IMAGE INPUT TYPE AND TRAINING

Time (s)	Epochs	Training accuracy	Validation accuracy	Type of input image
19.13	10	70.7%	66.3%	Grayscale
19.79	10	73%	68%	RGB
20.10	10	86.0%	71.1%	YCbCr
174.11	100	85.3%	80.1%	Grayscale
175.91	100	89.4%	79.9%	RGB
178.35	100	91.4%	86.7%	YCbCr
952.58	500	93.5%	83.7%	Grayscale
972.32	500	94.4%	80.3%	RGB
982.11	500	97.1%	92.5%	YCbCr

Table I compares the grayscale and YCbCr accuracies. The training model is identical. The only difference is that the image set input was extracted in YCbCr format, which is a 3-dimensional input in matrix form, while grayscale is a 1-dimensional matrix input. This is also evident in the fact that the grayscale program runs in a shorter duration. Also, the YCbCr data show the highest accuracy, followed by the grayscale, and RGB models.

## XIX. CONCLUSION

In this paper, we observed the effects on CNN with some slight modifications on a totally new dataset. We used a dataset with 4 different types of images. We varied several factors and compared their effects on the data training and validation accuracy. The image inputs varied and grayscale, RGB, and YCbCr images were utilized. The most accurate results came from the YCbCr model, as it was able to incorporate the effects

of grayscale and normal colors in the mathematical model. The difference was not significant but noticeable nevertheless. Thus the YCbCr image input proved to be a better input type for data extraction. CNN and SVC were used for image classification. The results were recorded, and it was observed that CNN provided better validation accuracy. The most possible reasons for this could be the smaller data set or the lower training duration and the average learning rate. We observed the accuracy variation due to under and over-training from 5 to 500 epochs and their results. The best results are between 100 and 200 epochs regarding time and accuracy. We varied the classification method, the kernel used, and the number of convolution layers. Other factors affecting the accuracy are the optimizer and the loss function calculation, but we did not observe any effect of these factors. We also confirmed the superior behavior of the CNN classifier for this dataset in comparison to the SVM. For future recommendations we should use a model with more layers and a bigger dataset, in order to have a more robust model with better classification ability.

The limitations of this study include the effects of optimizer calculation of the loss function for various values in order to find the optimal position. These steps were not monitored and can be considered as a project for future work and possible improvements.

This study was helpful in identifying that the best method to classify input image is CNN, at least for smaller datasets, and finally, it shows that CNN provides better classification results than SVM.

## REFERENCES

- [1] H. S. Lee and K. Kim, "Simultaneous Traffic Sign Detection and Boundary Estimation Using Convolutional Neural Network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1652–1663, Feb. 2018, <https://doi.org/10.1109/TITS.2018.2801560>.
- [2] "US smartphone ownership 2021," *Statista*. <https://www.statista.com/statistics/219865/percentage-of-us-adults-who-own-a-smartphone/>.
- [3] "Mobile Fact Sheet," *Pew Research Center: Internet, Science & Tech*. <https://www.pewresearch.org/internet/fact-sheet/mobile/>.
- [4] D. G. Gibson, A. Pereira, B. A. Farrenkopf, A. B. Labrique, G. W. Pariyo, and A. A. Hyder, "Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review," *Journal of Medical Internet Research*, vol. 19, no. 5, May 2017, Art. no. e7428, <https://doi.org/10.2196/jmir.7428>.
- [5] G. Montavon, W. Samek, and K.-R. Muller, "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1–15, Feb. 2018, <https://doi.org/10.1016/j.dsp.2017.10.011>.
- [6] S. Song, Z. Que, J. Hou, S. Du, and Y. Song, "An efficient convolutional neural network for small traffic sign detection," *Journal of Systems Architecture*, vol. 97, pp. 269–277, Aug. 2019, <https://doi.org/10.1016/j.sysarc.2019.01.012>.
- [7] A. Bouti, M. A. Mahraz, J. Riffi, and H. Tairi, "A robust system for road sign detection and classification using LeNet architecture based on convolutional neural network," *Soft Computing*, vol. 24, no. 9, pp. 6721–6733, May 2020, <https://doi.org/10.1007/s00500-019-04307-6>.
- [8] R. Antar, S. Alghamdi, J. Alotaibi, and M. Alghamdi, "Automatic Number Plate Recognition of Saudi License Car Plates," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8266–8272, Apr. 2022, <https://doi.org/10.48084/etasr.4727>.
- [9] Y. Hatolkar, P. Agrawal, and S. Patil, "A Survey on Road Traffic Sign Recognition System using Convolution Neural Network," *International Journal of Current Engineering and Technology*, vol. 8, no. 1, pp. 104–108, Jan. 2018, <https://doi.org/10.14741/ijcet/v.8.1.21>.
- [10] C. Qi *et al.*, "An efficient pruning scheme of deep neural networks for Internet of Things applications," *EURASIP Journal on Advances in Signal Processing*, vol. 2021, no. 1, Jun. 2021, Art. no. 31, <https://doi.org/10.1186/s13634-021-00744-4>.
- [11] M. Nagel, M. Fournarakis, R. A. Amjad, Y. Bondarenko, M. van Baalen, and T. Blankevoort, "A White Paper on Neural Network Quantization," *arXiv*, Jun. 15, 2021, <https://doi.org/10.48550/arXiv.2106.08295>.
- [12] S. Lawrence, C. L. Giles, and A. C. Tsoi, "Lessons in Neural Network Training: Overfitting May be Harder than Expected," in *Fourteenth National Conference on Artificial Intelligence*, Menlo Park, CA, USA, Jan. 1997, pp. 540–545.
- [13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, <https://doi.org/10.1109/tpami.2016.2577031>.
- [14] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," in *14th European Conference on Computer Vision*, Amsterdam, Netherlands, Oct. 2016, pp. 21–37, [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788.
- [16] M. Bhalekar and M. Bedekar, "D-CNN: A New model for Generating Image Captions with Text Extraction Using Deep Learning for Visually Challenged Individuals," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8366–8373, Apr. 2022, <https://doi.org/10.48084/etasr.4772>.
- [17] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: object detection via region-based fully convolutional networks," in *30th International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, Dec. 2016, pp. 379–387.
- [18] M. Njah and R. E. Hamdi, "A Constrained Multi-Objective Learning Algorithm for Feed-Forward Neural Network Classifiers," *Engineering, Technology & Applied Science Research*, vol. 7, no. 3, pp. 1685–1693, Jun. 2017, <https://doi.org/10.48084/etasr.968>.
- [19] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, Mar. 2021, Art. no. 53, <https://doi.org/10.1186/s40537-021-00444-8>.
- [20] 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>.
- [21] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Computer Science*, vol. 2, no. 6, Aug. 2021, Art. no. 420, <https://doi.org/10.1007/s42979-021-00815-1>.
- [22] A. Shmilovici, "Support Vector Machines," in *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach, Eds. Boston, MA, USA: Springer US, 2005, pp. 257–276.
- [23] H. Basly, W. Ouada, F. E. Sayadi, B. Ouni, and A. M. Alimi, "CNN-SVM Learning Approach Based Human Activity Recognition," in *9th International Conference on Image and Signal Processing*, Marrakesh, Morocco, Jun. 2020, pp. 271–281, [https://doi.org/10.1007/978-3-030-51935-3\\_29](https://doi.org/10.1007/978-3-030-51935-3_29).
- [24] S. Ahlawat and A. Choudhary, "Hybrid CNN-SVM Classifier for Handwritten Digit Recognition," *Procedia Computer Science*, vol. 167, pp. 2554–2560, Jan. 2020, <https://doi.org/10.1016/j.procs.2020.03.309>.
- [25] M. Elleuch, R. Maalej, and M. Kherallah, "A New Design Based-SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition," *Procedia Computer Science*, vol. 80, pp. 1712–1723, Jan. 2016, <https://doi.org/10.1016/j.procs.2016.05.512>.
- [26] Y. Lai, N. Wang, Y. Yang, and L. Lin, "Traffic Signs Recognition and Classification based on Deep Feature Learning," in *7th International Conference on Pattern Recognition Applications and Methods*, Funchal,



- Portugal, Jan. 2018, pp. 622–629, <https://doi.org/10.5220/0006718806220629>.
- [27] C. Kanan and G. W. Cottrell, "Color-to-Grayscale: Does the Method Matter in Image Recognition?," *PLOS ONE*, vol. 7, no. 1, Jan. 2012, Art. no. e29740, <https://doi.org/10.1371/journal.pone.0029740>.
- [28] "R. Zhu, D. Yu, S. Ji, and M. Lu, "Matching RGB and Infrared Remote Sensing Images with Densely-Connected Convolutional Neural Networks," *Remote Sensing*, vol. 11, no. 23, Jan. 2019, Art. no. 2836, <https://doi.org/10.3390/rs11232836>.
- [29] K. B. Shaik, P. Ganesan, V. Kalist, B. S. Sathish, and J. M. M. Jenitha, "Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space," *Procedia Computer Science*, vol. 57, pp. 41–48, Jan. 2015, <https://doi.org/10.1016/j.procs.2015.07.362>.
- [30] Z. S. G. Tari, J. Shah, and H. Pien, "Extraction of Shape Skeletons from Grayscale Images," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 133–146, May 1997, <https://doi.org/10.1006/cviu.1997.0612>.
- [31] J. Basilio, G. Torres, G. Sanchez-Perez, L. K. T. Medina, H. Meana, and G. S. Perez, "Explicit Image Detection using YCbCr Space Color Model as Skin Detection," in *Applications of Mathematics and Computer Engineering*, 2011, pp. 123–128.
- [32] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," *Procedia Computer Science*, vol. 132, pp. 679–688, Jan. 2018, <https://doi.org/10.1016/j.procs.2018.05.069>.
- [33] K. Eckle and J. Schmidt-Hieber, "A comparison of deep networks with ReLU activation function and linear spline-type methods," *Neural Networks*, vol. 110, pp. 232–242, Feb. 2019, <https://doi.org/10.1016/j.neunet.2018.11.005>.
- [34] S. Bruch, X. Wang, M. Bendersky, and M. Najork, "An Analysis of the Softmax Cross Entropy Loss for Learning-to-Rank with Binary Relevance," in *ACM SIGIR International Conference on Theory of Information Retrieval*, New York, NY, USA, Oct. 2015, pp. 75–78, <https://doi.org/10.1145/3341981.3344221>.
- [35] Z. Zhang and M. R. Sabuncu, "Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels," May 2018. Available: <https://ui.adsabs.harvard.edu/abs/2018arXiv180507836Z>.

# Technoeconomic Analysis of a Hybrid Energy System for an Academic Building

**Suchismita Roy**

School of Electrical Engineering, KIIT Deemed to be University, India  
roysuchismita@yahoo.co.in  
(corresponding author)

**Pradeep Kumar Sahu**

School of Electrical Engineering, KIIT Deemed to be University, India  
pksahu.nitrkl@gmail.com

Received: 14 October 2022 | Revised: 28 November 2022 | Accepted: 30 November 2022

## ABSTRACT

This work is mainly based on the optimal design of a standalone Hybrid Renewable Energy System (HRES) consisting of PV/diesel/battery systems, implemented in an academic building. Different hybrid system configurations such as PV-diesel generator-battery, diesel generator-battery, and PV-diesel generator are compared based on Net Present Cost (NPC) and Cost Of Energy (COE) to find out the best economically viable and environmentally friendly solution. Li-ion and lead-acid batteries were taken into consideration, and the optimization was done in HOMER PRO software. The PV-DG-Li-ion battery configuration emits approximately 2825387kg/year CO<sub>2</sub> whereas the conventional DG system emits 4565074kg/year. It is concluded that the PV-DG-Li-ion battery configuration provides the cleanest and most environment-friendly and techno-economically feasible solution.

**Keywords-**hybrid system; HOMER; NPC; COE; technoeconomical analysis

## I. INTRODUCTION

Nowadays, clean Renewable Energy Sources (RESs) play an important role in fulfilling the increasing energy demands and environmental pollution issues, since most are not only abundant in nature, but also environment friendly. The PV system is considered as the most suitable among RESs due to several advantages like low cost, environment-friendly nature, and higher efficiency. But due to its intermittency nature, it is often used as sole distributed generation in residential buildings, offices, and academic buildings commonly integrated with storage systems to enhance the stability of the system and power quality issues [1-2]. Energy storage techniques such as Li-ion and lead-acid batteries were commonly providing services from small to large power applications [3]. Li-ion batteries have several advantages over the lead-acid batteries, such as fast charging operation, higher density in power and energy, and lower maintenance cost [3]. Different studies on battery energy storage systems reveal that the plays a major key role in energy generation power system when incorporated with RESs. Two key parameters, NPC and COE, are used for evaluating the performance of the battery storage system. The NPC of a Li-ion battery is relatively better than its counterpart lead-acid battery. Moreover, the life cycle of the Li-ion battery is more than five years longer than the lead-acid battery when it is interfaced with a solar PV system. The environmental effect along with the maintenance of Li-ion

battery is also less [3]. Though the Li-ion battery wins over the lead-acid battery, but still for stationary applications, lead-acid batteries provide better efficiency. The energy storage system not only counter effects the unpredictability and variation in the output power produced by RESs but also increases the reliability and efficiency of the system [4-6].

A PV system can also be integrated with wind energy and a Diesel Generator (DG), in an installment also known as a Hybrid Energy System (HES) for efficient energy generation and enhanced reliability. If only the DG is operated for the generation of energy, then the overall maintenance and fuel cost will be much higher. HES has also several other disadvantages, such as a higher level of noise and emission of greenhouse gases. When the sources of the HES are integrated with the battery storage system, then the operational cost is drastically reduced. Moreover, the daily and seasonal variation demand affects the operational cost. During summer and winter seasons the fuel costs for weekends are higher compared to weekdays. Due to lower solar radiation, the use of conventional energy sources increases during winter. So, the variation of daily and seasonal energy demands changes is an important aspect as these affect the operational cost [7-8]. The reliability and compatibility of the PV solar system increases when it is integrated with other energy sources like DGs, wind farms, and battery. Such a system is called a HES. Presently, the HES powered by RESs is the most cost-effective and reliable.

Past works have mainly concentrated on the optimal sizing and the economic aspect of the designed hybrid system configuration either with solar PV- wind system, solar PV-wind-diesel generator system coupled with battery storage system [9-13]. Only a few published papers are available that consider the wind resource not adequately available in a particular location. Anyhow, the techno economic analysis and the optimal sizing of the hybrid system consisting of PV panels, DG, and a battery system consisting of Li-ion and Lead acid battery is not available in previous works. Since the wind energy resource is not abundantly available for the location under study, it is not considered for the present work.

In this paper, the technical and economic analysis of solar/DG/battery storage HES implemented in an academic building is proposed. The proposed hybrid system consists of PVs, a DG, and a battery storage system consisting of Li-ion and/or lead-acid batteries. The techno-economic analysis of the hybrid system is carried out in HOMER (Hybrid Optimization Model for Electric Renewable) software based on NPC and COE and the results are compared regarding cost-effectiveness and efficiency. The optimal sizing of the system is estimated based on Total Energy Deficit (TED), NPC, and COE to meet the load demand of the system. The DG and battery bank are designed to compensate for the load demand during night time when the solar energy is not available. According to the developed methodology, the configurations suggested should ensure 0% TED. The optimal system is decided based on the minimum NPC and COE. The optimal hybrid system is designed and compared with other energy sources by using the solar radiation data and ambient temperature obtained from HOMER for the particular location under study [14-15].

The proposed configuration consists of a solar PV system, a DG along with a battery storage system, a converter, and an inverter system. The converter converts the ac voltage produced by the DG to dc voltage and the inverter converts the dc voltage produced by the hybrid system into ac to meet the household load requirements.

## II. SOURCE AND LOAD ASSESMENT

To develop an efficient hybrid system in a particular location, source and load assessment are essential for its techno economical analysis.

### A. Site Location

The site under consideration is the academic institution KIIT Deemed to be University, located in the city of Bhubaneswar, Odisha, India. It consists of 5 academic blocks A, B, C, D, and E (Figure 1). Block A consists of 10 classrooms, 1 conference hall, 12 laboratories, and 5 washrooms. Block B contains the dean's office and 5 administrative offices. Block C consists of 12 classrooms, 5 laboratories, and 5 washrooms. Block D consists of 6 classrooms, 12 faculty cabins, and 6 washrooms. Block E consists of 20 classrooms, 4 laboratories, and 6 washrooms.

### B. Source Assessment

The institute is located at 20.3555°N, 85.8161°E and solar radiation is plentifully available. During the summer season, the electrical load is quite high, so there is a necessity for grid

independent hybrid RESs. For this location, solar power is abundantly available. Figure 2 shows the monthly variation of solar irradiance and Figure 3 shows the monthly temperature variation.



Fig. 1. Building blocks.



Fig. 2. Monthly solar irradiance.

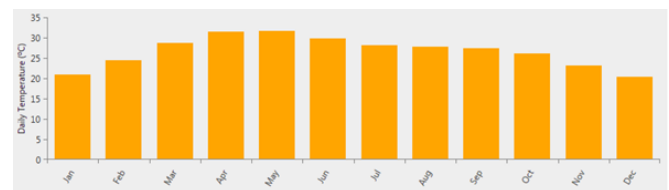


Fig. 3. Monthly temperature variation.

### C. Load Assessment

Maximum load occurs during the daytime hours, from 8 am to 8 pm when the academic activities are carried out. During night time, the load is minimal. During spring months the load is at its peak. During vacations the load is less. The total load demand of the institute is 10427.5kwh/day and the peak load is 1038.1kw. The proposed system is designed to satisfy all the demands economically. The system is designed to mitigate the peak load demand during summer and provide a reliable solution for all the existing energy problems. The hourly load demand and the monthly average solar global horizontal irradiance are plotted with variation of clearness index. With seasonal changes the temperature also varies, so the variation of daily temperature over one year is also plotted. Table I provides detailed load calculation of the institution under consideration. The hourly load profile is plotted in Figure 4 and the monthly load profile is plotted in Figure 5. The total load is maximum at 2pm, obtained from the daily load curve and the monthly peak is during May, i.e. during summer time.

## III. RESOURCE ASSESSMENT

In this paper, PVs, DG, and batteries were considered as the resources for HOMER optimization. The location of the institute under study receives plenty radiation all over the year, so it is suitable for harnessing solar energy.

TABLE I. LOAD ASSESSMENT

Campus 3, Block A	Load (KW)	h/day	KWh/day
Light load	12.072	8	96.576
Fan load	6.08	12	72.96
AC and laboratory load	320.234	5	3202.34
Total	338.386		3371.876
Campus3, Block B	Load (KW)		
Light load	2.357	8	18.856
Fan load	2.4	10	24
AC and laboratory load	40.16	5	481.92
Total	44.917		524.776
Campus 3, Block C	Load (KW)		
Light load	9.395	8	75.16
Fan load	4.56	12	54.72
AC and laboratory load	281.145	5	2811.45
Total	295.1		2941.33
Campus 3, Block D	Load (KW)		
Light load	7.087	8	56.696
Fan load	5.12	12	61.44
AC and laboratory load	120.639	5	1206.39
Total	132.846		1324.52
Campus 3, Block E	Load (KW)		
Light load	26.517	8	212.136
Fan load	24.56	12	294.72
AC and laboratory load	175.816	5	1758.16
Total	226.893		2265.01

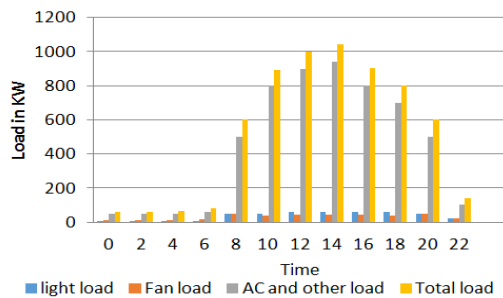


Fig. 4. Daily load curve.

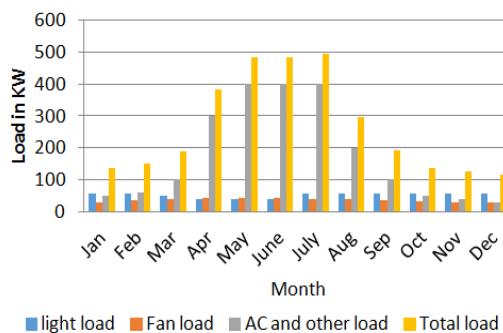


Fig. 5. Monthly load curve.

During summer the temperature can go as high as 45°C. Solar radiation is available for approximately 10-12 hours with longer sunny days during summer whereas during winter the radiation is available for almost 8-10 hours with shorter sunny days. During the rainy season, it is hardly available for almost two months. The solar radiation data for this particular location are obtained from the site of NASA using the solar energy database from HOMER software. The annual average solar

radiation is 4.81kWh/m<sup>2</sup>/day with an average clearness index of 0.519. Figure 2 shows the monthly variation in daily radiation which indicates the ample solar radiation available. The energy generated by the PV panels can be expressed as:

$$E_{PV} = A \times \eta \times H \times PR \quad (1)$$

where  $E_{PV}$  is the total energy output of the PV panel,  $\eta$  is solar panel efficiency,  $A$  is the total panel area (m<sup>2</sup>),  $H$  the average solar irradiance on a tilted panel annually (kWh/m<sup>2</sup>/day), and  $PR$  the performance ratio of the panel

#### IV. SYSTEM DESCRIPTION AND COMPONENTS

The HRES can be designed consisting of a PV module, a DG, a power converter, and a battery bank consisting of lead-acid or lithium-ion batteries. In this configuration, the PV panels and the battery bank are connected to the DC bus system while the DG is connected to the ac bus system. The electrical load is connected to the ac bus system. The solar PV system generates the power to compensate for the electrical load demands and the excess generated power will charge the battery bank which can be used along with the DG during adverse weather conditions or during night time. The size of all the components used here is chosen such that there is no undue load on the system. The system configuration is given in Figure 6. If the energy generated by the hybrid system is more than what is required by the load, the excess energy is stored in the battery. If the energy generated is less than the load demand, then the battery bank will be used. In that scenario, the DG will operate and meet the deficit energy if the state of charge of the battery drops to the minimum value.

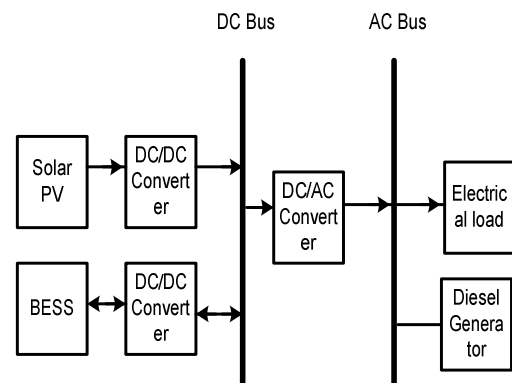


Fig. 6. System configuration.

##### A. PV Module

In India, there is ample scope for using solar energy which is a major RES. The generated power is DC in nature from which the maximum power can be extracted by using the MPPT technique. By changing the tilt angle of the solar panel the maximum solar radiation can be received according to the changing position of the sun. The PV power generated can be given by [16]:

$$P_{PV} = F_{PV} Y_{PV} \frac{I_T}{I_S} \quad (2)$$

where  $P_{PV}$  is the PV power (W),  $F_{PV}$  is the derating factor of PV which includes the effect of dust accumulation and temperature rise,  $Y_{PV}$  is the rated capacity of the PV panel,  $I_T$  is the incident solar radiation under operating conditions, and  $I_S$  is the incident solar radiation under standard operating condition ( $W/m^2$ ). The details of the PV module are given in Table II.

TABLE II. DETAILS OF THE PV MODULE

VarioTrack VT -65	Value
Rated capacity	1000kW
Panel Type	Flat plate
Operating temperature	45°C
Efficiency	17.3%
Capital cost	602\$/kW
Replacement cost	602\$/kW
O&M cost	100\$/kW

### B. Diesel Generator

The DG gives the best possible combination for HESs incorporating solar RES with batteries. In the case of grid integrated systems, DGs provide an efficient backup during power shutdown periods. The capital cost of the DG is minimal but the operation and management cost is quite significant due to the fuel cost. The DG is incorporated into the system to evaluate its environmental impact. The details of the DG are given in Table III [17].

TABLE III. DETAILS OF THE DIESEL GENERATOR

CAT 1750	Value
Capital cost	500\$/kw
Replacement cost	450\$/kw
O&M cost	0.03\$/hour
Lifetime	15000h
Efficiency	45%

### C. Battery Bank

The energy produced by the solar PV system depends upon the intermittent solar irradiance. During night time or in bad weather conditions, it is not available. The main constraint of using RESs is that they are fluctuating in nature. So, in order to reduce the dependency on RES availability, battery energy sources are used in this study. In this work, both lithium-ion and lead-acid batteries were taken into consideration. The details of the battery bank systems are given in Tables IV-V.

Lead-acid batteries are used for residential applications for storing the energy produced by the solar PV system and to provide energy during adverse weather conditions and night time. These batteries have good reliability, low cost, and reduced discharge capacity. Despite all their advantages, they have reduced lifetime due to continuous loading and discharges. The lead-acid battery of the Surrette S-260 model with a nominal voltage of 12V and string size 2 is considered in this work. The initial state of charge of the battery is 100% and the minimum state of charge is 30%. The throughput of the battery is 1704kWh and its lifetime is 10 years. Its capital cost is 205\$ and the operation and management cost is 50\$/year [6].

TABLE IV. DETAILS OF LEAD-ACID BATTERY

Surrette S-260	Value
Nominal voltage	12V
Nominal capacity	3.12kWh
Maximum capacity	260Ah
Round the trip efficiency	80%
Capital cost	205\$
Replacement cost	205\$
O&M cost	50\$
Lifetime	10years

TABLE V. DETAILS OF LITHIUM-ION BATTERY

Generic 1kWh Li-ion battery	Value
Nominal voltage	6V
Nominal capacity	1kWh
Maximum capacity	167Ah
Round the trip efficiency	90%
Capital cost	132\$
Replacement cost	100\$
O&M cost	10\$
Lifetime	10years

Lithium-ion batteries can be used as alternatives because they have high efficiency and reliability and prolonged lifetime. Their only drawback is that they are quite expensive compared to the lead-acid batteries. The generic 1kWh Li-ion battery with a nominal voltage of 6V and string size 4 is considered in this paper. The initial and minimum state of charge is of 100% and 30%, the throughput is 3000kWh, and its lifetime is 10 years. Its capital cost is 132\$ and the operation and management cost is 10\$ [6].

The battery rating should be designed such that it can meet the load demand of the HES. The Wh capacity of the battery system is given by (3):

$$C_{Wh} = E_L \times AD \times \eta_{Bat} \times \eta_{Inv} \times DoD \quad (3)$$

where  $C_{Wh}$  represents the WH capacity of the load,  $E_L$  is the daily average load in kWh/day,  $AD$  is the autonomy of the battery,  $\eta_{Bat}$  and  $\eta_{Inv}$  are the efficiency of the battery and the inverter, respectively, and  $DoD$  is the battery Depth of Discharge.

### D. Power Converter

The system consists of both AC and DC systems, so the power converters act as bridges that transfer power between these two systems. The solar PV produces DC power and the DG and the load are connected to the AC. So, power converters act as an interface between these two. The rating of the converter under consideration is 1725kW. The efficiency is 95% and the lifetime of the converter is 10 years. Its capital cost is 32189\$ and its operation and maintenance cost is 8000\$/year [18].

## V. HOMER OPTIMIZATION

HOMER software is used to optimize the HES based on certain input parameters to provide the best possible combination in terms of techno-economical feasibility. The optimization aims to minimize the NPC and COE as objective functions. It also computes the size of the PV array, the number of batteries, and the size of the DG and the converter [19-20].

The NPC of the system includes all the installation, operational, and management costs minus the revenue generated over the period. The NPC is calculated by:

$$NPC = \frac{C_{Tot,annual}}{CRF(i,T)} \quad (4)$$

where  $C_{Tot,annual}$  is the annual total cost of the system (\$/year),  $i$  is the real interest rate annually (%),  $T$  is the time period of the project work (years), and  $CRF(i,T)$  is the capital recovery factor.

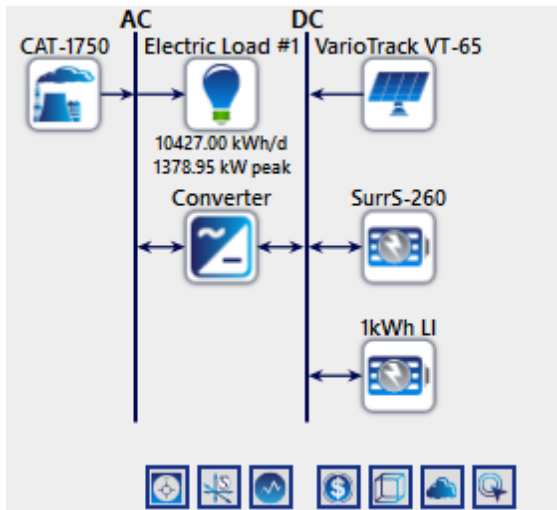


Fig. 7. Hybrid system configuration in HOMER.

COE represents the cost of the production of electricity by the HES under consideration on an annual basis. It is calculated by:

$$COE = \frac{C_{Tot,annual}}{E_s} \quad (5)$$

where  $C_{Tot,annual}$  is the annual total cost of the system and  $E_s$  is the annual energy supplied.

The payback period (PBP) of the optimal system can also be calculated to ensure that the time period required for the recovery of capital or investment cost should be less:

$$PBP = \frac{\text{Total Capital cost}}{COE \times \text{Annual Energy Consumption}} \quad (6)$$

The proposed HES model developed in HOMER software is shown in Figure 7. The optimization is done based on the input data given in Tables II-V. The project lifetime is 25 years with an expected inflation rate of 2% and a nominal discount rate of 8%.

## VI. RESULT ANALYSIS AND DISCUSSION

The simulation is carried out with the following configurations:

- PV-DG-battery (Li-ion and lead-acid)
- DG-battery
- PV-DG

### • DG

The optimization is done on the above-mentioned configurations and the NPC and COE of these architectures are determined accordingly in HOMER. The first configuration is considered for two types of batteries, Li-ion and lead-acid. The result analysis shows the list of hybrid energy configurations based on the lowest NPC and COE. Table VI gives the optimal size of different possible hybrid energy configurations.

TABLE VI. OPTIMAL SIZE OF DIFFERENT CONFIGURATIONS

System configuration	PV (kW)	DG (kW)	No. of Li-ion battery strings	No. of lead-acid battery strings	Converter (kW)
PV-DG-Li-ion battery-converter	584	1400	274		670
DG- Li-ion battery-converter		1400	279		690
PV-DG-lead-acid battery-converter	2921	1400		304	604
DG-lead-acid battery-converter		1400		306	592
DG		1400			
PV-DG-converter	869	1400			172

The total NPC, the levelized COE, and the operating cost of all configurations are given in Table VII.

TABLE VII. COST ANALYSIS OF ALL CONFIGURATIONS

System configuration	NPC(\$)	COE(\$)	Operating cost (\$)
PV-DG-Li-ion battery-converter	21.5M\$	0.437\$	1.65M\$
DG- Li-ion battery-converter	21.58M\$	0.438\$	1.65M\$
PV-DG-lead-acid battery-converter	22.91M\$	0.465\$	1.76M\$
DG-lead-acid battery-converter	22.98M\$	0.467\$	1.76M\$
DG	33.6M\$	0.684\$	2.603M\$
PV-DG-converter	33.6M\$	0.684\$	2.603M\$

From the above configurations, the PV-DG-Li-ion battery-converter system gives the most cost-effective solution. From Table VII, it can be observed that the above mentioned system has an NPC of 21.5M\$, COE of 0.437\$, and overall operating cost of 1.65M\$. Out of the 6 system configurations, the system containing only DG has the highest fuel cost. Among the PV-DG-Li-ion battery-converter and the PV-DG-lead-acid battery-converter, the first one gives a better economical solution. Although the initial cost of a Li-ion battery is more, the life cycle cost is less compared to the lead-acid battery. So, the overall NPC, COE, and operating cost of the configuration containing a Li-ion battery is less compared to the total cost of a system with a lead-acid battery. The optimal system consists of a PV cell of 584kW capacity, a 1400kW DG, a 670kW converter, and 274 strings of Li-ion batteries. The highest amount of electricity is produced in April and May as the solar radiation is the highest during these two months. Until around 18:00hr, the PV and the DG of the HES meet the load demand of the system and the surplus energy is used to charge the



battery system. But after that, the solar radiation decreases and it eventually becomes zero by 19:00hrs. During that time the DG alone can't meet the load demand so the battery will be utilized. The load demand decreases after 20:00hrs. Therefore, excess power is generated by the DG and the surplus energy is used to charge the battery. Figure 8 depicts the cost analysis of all 6 possible configurations in terms of NPC, COE, and operating cost. The PBP of all possible configurations are estimated by assuming a project life period of 25 years with inflation rate and nominal discount rate of 2% and 8%, respectively. The PBP of the proposed optimal system consisting of PV-DG-Li-ion battery is calculated as 5.6 years which is sufficient to recover all the cost invested in the system.

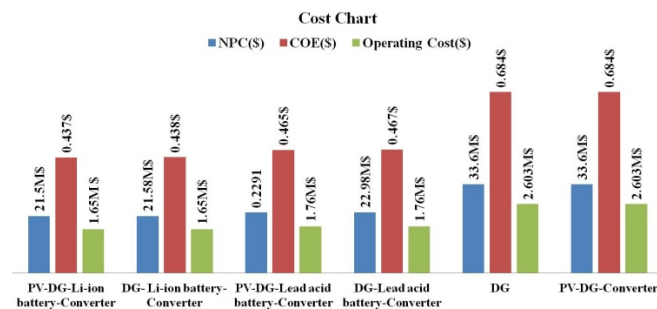


Fig. 8. Cost chart of various configurations of the hybrid system.

## VII. ENVIRONMENTAL EFFECT

RESs produce clean and sustainable energy. Table VIII shows the emission of pollutant gases like carbon dioxide (CO<sub>2</sub>), carbon monoxide (CO), unburned hydrocarbons, sulphur dioxide (SO<sub>2</sub>), and nitrogen oxide (NO) from the different configurations obtained from HOMER software. Compared to the conventional DG system which emits 4565 tons of CO<sub>2</sub>, the optimal hybrid energy system emits 2825 tons which are quite less. This CO<sub>2</sub> emission reduction is due to the penetration of the RES in the system. The other pollutants are also reduced significantly. Figure 9 shows the graph of the emission of harmful pollutants by the 6 possible configurations under consideration and it can be observed that the PV-DG-Li-ion battery-converter system emits less pollutants than the other systems.

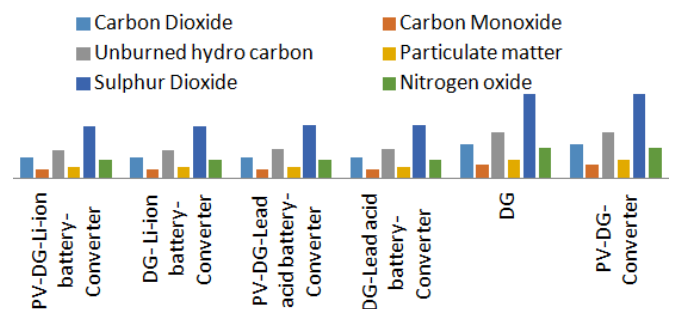


Fig. 9. Pollutant emissions.

TABLE VIII. POLLUTANT EMISSIONS

Emissions (kg/year)	System configurations					
	PV-DG-Li-ion battery-converter	DG-Li-ion battery-converter	PV-DG-lead-acid battery-converter	DG-lead-acid battery-converter	DG	PV-DG-converter
CO <sub>2</sub>	2825387	2836377	2887232	2898694	4565074	4561485
CO	1187	1191	1213	1217	1917	1916
Unburned hydro carbons	385	386	393	395	622	621
Particulate matter	160	161	164	165	259	259
SO <sub>2</sub>	7010	7037	7163	7192	11326	11317
NO	25389	25488	25945	26048	41022	40990

## VIII. CONCLUSION

In this paper, the techno-economic analysis of a hybrid energy system in HOMER software for an educational institute is briefly discussed. The annual average solar radiation of the site under consideration is 4.81kWh/m<sup>2</sup>/day with an average clearness index of 0.519. This is the potential location for the installation of PV/DG/battery system which gives the best economically viable solution by optimizing through HOMER software. The optimal system consists of a PV cell of 584kW capacity, a Diesel generator of 1400kW capacity, a 670kW converter, and 274 strings of Li-ion batteries with a nominal voltage of 6V and nominal capacity of 167Ah. The conventional Diesel generator system is not an economically viable solution due to its high fuel cost. It has also a hazardous effect on the environment due to its high level of emissions. In the present system, the hybrid energy system model is combined with the conventional DG system and the resulting system model gives the best possible result in terms of economy. The system not only reduces the number of operating

hours of the Diesel generator and hence the fuel cost, but also reduces the emission of harmful gases. It is found that PV-DG-Li-ion battery system emits approximately 2825387kg CO<sub>2</sub> per year, when the conventional Diesel generator system emits 4565074kg/year. The PBP is calculated for the optimal system consisting of PV-DG-Li ion battery and it is found to be 5.6 years which is sufficient to recover all the costs invested in the system. Finally, the results obtained from this study can be successfully implemented in similar geographical location areas and the economy of these areas can be improved significantly along with the standard of living.

## REFERENCES

- [1] E. A. Al-Ammar, N. H. Malik, and M. Usman, "Application of using Hybrid Renewable Energy in Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 1, no. 4, pp. 84–89, Aug. 2011, <https://doi.org/10.48084/etasr.33>.
- [2] T. V. Krishna, M. K. Maharana, and C. K. Panigrahi, "Integrated Design and Control of Renewable Energy Sources for Energy Management," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5857–5863, Jun. 2020, <https://doi.org/10.48084/etasr.3613>.

- [3] A. A. Kebede *et al.*, "Techno-economic analysis of lithium-ion and lead-acid batteries in stationary energy storage application," *Journal of Energy Storage*, vol. 40, Aug. 2021, Art. no. 102748, <https://doi.org/10.1016/j.est.2021.102748>.
- [4] A. Mayyas, A. Chadly, S. T. Amer, and E. Azar, "Economics of the Li-ion batteries and reversible fuel cells as energy storage systems when coupled with dynamic electricity pricing schemes," *Energy*, vol. 239, Jan. 2022, Art. no. 121941, <https://doi.org/10.1016/j.energy.2021.121941>.
- [5] L. Olatomiwa, S. Mekhilef, A. S. N. Huda, and K. Sanusi, "Techno-economic analysis of hybrid PV–diesel–battery and PV–wind–diesel–battery power systems for mobile BTS: the way forward for rural development," *Energy Science & Engineering*, vol. 3, no. 4, pp. 271–285, 2015, <https://doi.org/10.1002/ese3.71>.
- [6] M. M. Symeonidou, C. Zioga, and A. M. Papadopoulos, "Life cycle cost optimization analysis of battery storage system for residential photovoltaic panels," *Journal of Cleaner Production*, vol. 309, Aug. 2021, Art. no. 127234, <https://doi.org/10.1016/j.jclepro.2021.127234>.
- [7] O. Krishan and S. Suhag, "Techno-economic analysis of a hybrid renewable energy system for an energy poor rural community," *Journal of Energy Storage*, vol. 23, pp. 305–319, Jun. 2019, <https://doi.org/10.1016/j.est.2019.04.002>.
- [8] A. Kaabeche and R. Ibtouen, "Techno-economic optimization of hybrid photovoltaic/wind/diesel/battery generation in a stand-alone power system," *Solar Energy*, vol. 103, pp. 171–182, May 2014, <https://doi.org/10.1016/j.solener.2014.02.017>.
- [9] B. K. Das, M. A. Alotaibi, P. Das, M. S. Islam, S. K. Das, and M. A. Hossain, "Feasibility and techno-economic analysis of stand-alone and grid-connected PV/Wind/Diesel/Batt hybrid energy system: A case study," *Energy Strategy Reviews*, vol. 37, Sep. 2021, Art. no. 100673, <https://doi.org/10.1016/j.esr.2021.100673>.
- [10] S. Gabra, J. Miles, and S. A. Scott, "Techno-economic analysis of stand-alone wind micro-grids, compared with PV and diesel in Africa," *Renewable Energy*, vol. 143, pp. 1928–1938, Dec. 2019, <https://doi.org/10.1016/j.renene.2019.05.119>.
- [11] R. H. T. Djiela, P. T. Kapen, and G. Tchuen, "Techno-economic design and performance evaluation of Photovoltaic/Diesel/Batteries system through simulation of the energy flow using generated solar radiation data," *Energy Conversion and Management*, vol. 248, Nov. 2021, Art. no. 114772, <https://doi.org/10.1016/j.enconman.2021.114772>.
- [12] R. Dufo-López and J. L. Bernal-Agustín, "Techno-economic analysis of grid-connected battery storage," *Energy Conversion and Management*, vol. 91, pp. 394–404, Feb. 2015, <https://doi.org/10.1016/j.enconman.2014.12.038>.
- [13] K. R. Khalilpour and A. Vassallo, "Technoeconomic parametric analysis of PV-battery systems," *Renewable Energy*, vol. 97, pp. 757–768, Nov. 2016, <https://doi.org/10.1016/j.renene.2016.06.010>.
- [14] J. Wesly, A. C. P. Brasil, C. A. Frate, and R. K. Badibanga, "Techno-economic analysis of a PV-wind-battery for a remote community in Haiti," *Case Studies in Chemical and Environmental Engineering*, vol. 2, Sep. 2020, Art. no. 100044, <https://doi.org/10.1016/j.cscee.2020.100044>.
- [15] M. B. Hamida, W. Ahmed, M. Asif, and F. A. Almaziad, "Techno-Economic Assessment of Energy Retrofitting Educational Buildings: A Case Study in Saudi Arabia," *Sustainability*, vol. 13, no. 1, Jan. 2021, Art. no. 179, <https://doi.org/10.3390/su13010179>.
- [16] R. Kannan, K. C. Leong, R. Osman, H. K. Ho, and C. P. Tso, "Life cycle assessment study of solar PV systems: An example of a 2.7kWp distributed solar PV system in Singapore," *Solar Energy*, vol. 80, no. 5, pp. 555–563, May 2006, <https://doi.org/10.1016/j.solener.2005.04.008>.
- [17] M. L. Tuballa and M. L. S. Abundo, "Operational Impact of RES Penetration on a Remote Diesel-Powered System in West Papua, Indonesia," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2963–2968, Jun. 2018, <https://doi.org/10.48084/etasr.1984>.
- [18] R. F. Bastos, G. H. Fuzato, C. R. Aguiar, R. V. A. Neves, and R. Q. Machado, "Model, design and implementation of a low-cost HIL for power converter and microgrid emulation using DSP," *IET Power Electronics*, vol. 12, no. 14, pp. 3833–3841, 2019, <https://doi.org/10.1049/iet-pel.2019.0302>.
- [19] D. I. Papaioannou, C. N. Papadimitriou, A. L. Dimeas, E. I. Zountouridou, G. C. Kiokos, and N. D. Hatziaargyriou, "Optimization & sensitivity analysis of microgrids using HOMER software- A case study," in *MedPower 2014*, Athens, Greece, Aug. 2014, pp. 1–7, <https://doi.org/10.1049/cp.2014.1668>.
- [20] B. U. Kansara and B. R. Parekh, "Modelling and simulation of distributed generation system using HOMER software," in *2011 International Conference on Recent Advancements in Electrical, Electronics and Control Engineering*, Sivakasi, India, Sep. 2011, pp. 328–332, <https://doi.org/10.1109/ICONRAEECE.2011.6129804>.

# Assessment of the Flexural Strength of No-Fines Recycled Aggregate Concrete Prisms

**Ihsan Ali Rahu**

Department of Civil Engineering, Quaid-e Awam University of Engineering, Science & Technology, Pakistan  
Ihsanali0999@gmail.com

**Bashir Ahmed Memon**

Department of Civil Engineering, Quaid-e-Awam University of Engineering, Science & Technology, Pakistan  
bashir\_m@hotmail.com

**Mahboob Oad**

Department of Civil Engineering, Quaid-e-Awam University of Engineering, Science & Technology, Pakistan  
engrmahboob04@gmail.com  
(corresponding author)

**Shakeel Ahmed Dahri**

Department of Civil Engineering, Quaid-e-Awam University of Engineering, Science & Technology, Pakistan  
enr.shakeelahmed9@gmail.com

**Abdul Raqeeb Memon**

Department of Structural Engineering, Faculty of Engineering, Coventry University, UK  
raqeeba2@uni.coventry.ac.uk

**Amjad Hussain Bhutto**

Department of Civil Engineering, Quaid-e-Awam University of Engineering, Science and Technology, Pakistan  
amjad.bhutto62@gmail.com

*Received: 2 November 2022 | Revised: 25 November 2022 | Accepted: 28 November 2022*

## ABSTRACT

This study investigated experimentally the flexural strength of no-fines recycled aggregate concrete, produced using 20-70% replacement of conventional coarse aggregates with coarse aggregates from demolished waste. Six prisms in each dosage with a 1:4 mix and 0.5 water-binder ratios were prepared. A batch of prisms with conventional aggregates was also cast to compare them with the proposed concrete. An equal number of samples were cured for 7 and 28 days and tested under a gradually increasing central point load to examine failure load, central deflection, and flexural strength. The comparison of results showed an increasing trend in deflection with an increase in the dosage of recycled aggregates. The 7-day cured samples had approximately 2.6 times the deflection of conventional concrete. However, the deflection at all replacement levels remained less than that allowed by ACI-318. The results showed a decreasing trend in flexural strength with an increase in the dosage of recycled aggregates. The 40% replacement sample had a less than 20% strength reduction and is recommended as the optimum level of replacement of conventional aggregates for the production of no-fines recycled aggregate concrete.

*Keywords-flexural strength; no-fines recycled aggregate concrete; demolishing waste*

## I. INTRODUCTION

The construction industry has used a variety of materials to build safe and reliable infrastructures and meet the needs of inhabitants and trends of society. As concrete is currently the most widely used material in the construction industry due to its versatility and durability, the rapid construction development brings a burden on its natural resources. In parallel, changes in infrastructure lead to a drastic increase in demolitions and, consequently, waste generation. In previous decades, landfills were commonly available to accommodate this waste, but nowadays the lack of space around city centers, aesthetics, and environmental policies complicates the management of such waste. Furthermore, this waste cannot be dumped in or around cultivated lands, as it would cause serious impacts on the agricultural sector and the economy of the area in general. Most demolishing waste is used in floors, plinth protection, or similar filling sections but still, the residual waste raises worries. A possible solution could be the reuse of concrete waste material in new concrete to produce an indigenous green product. However, there is a need to investigate better the behavior of waste materials and their effects on the fresh and hardened properties of concrete, although several studies attempted to examine the use of demolishing waste in new concrete [1].

Concrete is used in a variety of forms in modern infrastructures. Low-strength concrete, such as no-fines concrete, is commonly used in non-structural members, such as in partition walls in place of bricks or pavements for better drainage. However, it may also be used for other types of structures in non-structural members or low-load areas. This study investigated the use of demolition waste as coarse aggregates to produce no-fines recycled aggregate concrete and examined its flexural strength.

## II. LITERATURE REVIEW

The use of recycled aggregates is an globally active research field. A summary of the most recent developments was presented in [1], discussing processes, issues, and effects on the hardened product, and proposing a few recommendations for its smooth and confident use in the industry. In [2], demolition waste was used as low-grade coarse aggregates, while authors in [3] concluded that better concrete properties can be achieved if the source of the recycled aggregates is high-strength concrete. In [4], an investigation on the use of high-strength recycled concrete did not show an effect of the replacement percentage on the slump flow of the resulting concrete but about a 10% reduction in compressive, split tensile, and flexural strengths. In [5], a 5% reduction in split tensile strength was observed at equal replacement of conventional coarse with recycled aggregates. In [6], the properties of concrete from demolished waste were investigated in fresh and hardened states, along with its long-term effects. It is obvious that there is a wide scatter of results of the relevant studies.

In parallel to modifications in the normal concrete matrix for better performance, several studies investigated the development of no-fines concrete. No-fines concrete was initially developed for road pavement in areas of low traffic

volumes, as its main features are good drainage of surface water or recharge of the groundwater table. Its initial development and progress were reviewed in [7]. Several studies examined the mechanical properties of no-fines concrete for road pavement, highlighting its efficiency in low-volume traffic areas [8-12]. The compressive strength of no-fines concrete for road pavement was evaluated in [13], finding that although its strength was lower compared to conventional concrete, it was suitable for low-traffic pavements. Normally, single-sized coarse aggregates are used in the development of no-fines concrete. In contrast, graded aggregates were used in [14] to develop no-fines concrete and study its compressive strength, finding that 10-20mm was the most suitable grading as it showed a compressive strength equal to 15.7MPa. Different additives such as silica fume [15], limestone dust [16], fly ash [17-18], ground granulated blast furnace slag [18], etc., have also been used to improve the performance of no-fines concrete. The use of additives improved the performance of the no-fines concrete but it is still not as that of the conventional concrete. Therefore, additional research is required to improve its properties and performance. Improvement in the performance of no-fines concrete has also been attempted using very small quantities of fines in concrete, showing improvement in strength but a reduction in its permeability characteristics, which, in some cases, is the main requirement [19]. Other studies attempted to improve no-fines concrete using uncrushed aggregates [20], fly ash with admixtures [21], different mixtures and water-binder ratios [22], GGBS [23], etc.

Not only compressive strength is important, but also flexural strength. Therefore, a variation of fly ash and plasticizer was used in [24] to study the flexural strength of no-fines concrete, designing 5 different mixes, and concluding that 60% fly ash with 0.5% plasticizer and 0.17 water-binder ratio was optimum for producing equal or more flexural strength than conventional concrete. Furthermore, it was noted that extreme care should be taken in dose decisions, as higher doses of plasticizer can increase slump and reduce flexural strength. In [25], 1:2 and 1:3 mixes were used to develop no-fines concrete and study its compressive, tensile, and flexural strengths, concluding that 12.5mm was the most suitable size for coarse aggregates, the 1:2 mix was better than the 1:3, and observing a 23% reduction in flexural strength in comparison to the conventional. In [26], 30% less flexural strength was observed in no-fines concrete. The development of green concrete was also attempted using plastic waste and nano-silica [27], recycled binder concrete [28], and sugarcane bagasse ash and limestone fines [29]. Green concrete made with RCA was also studied in [30] for strain and displacement under sustained loading of 12 months, showing only a 10.34% deviation from the control concrete.

Despite the scattering in the results, it is evident that many research efforts focus on improving no-fines concrete, but it is also evident that few studies focus on the use of recycled aggregates in the production of no-fines concrete. This study aimed to investigate varying percentages of recycled aggregates in the production of no-fines concrete, study its flexural strength, and provide a good insight into its use.

### III. MATERIALS AND METHODS

Demolition waste was collected from slab debris from a two-story building to prepare no-fines concrete. Large pieces of concrete, as shown in Figure 1, were hammered manually to produce coarse aggregates of 25mm maximum size. Conventional coarse aggregates from Nooriabad hills and fine aggregates from Bolhari hills were obtained from the local market. Conventional and demolished coarse aggregates were cleaned and washed, as shown in Figure 2, and sieved to ensure their quality. Figure 3 shows the gradation curve of both aggregate types, and it can be observed that the trend of both curves and the ranges on various sieves fall in the specified ranges of ASTM C-136 [31].



Fig. 1. Demolishing waste.

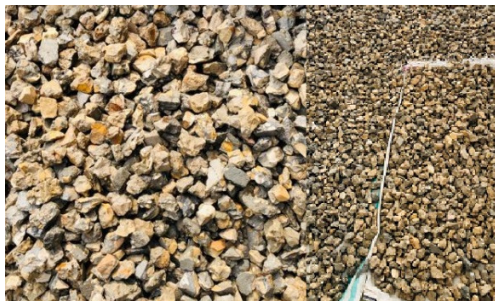


Fig. 2. Aggregates.

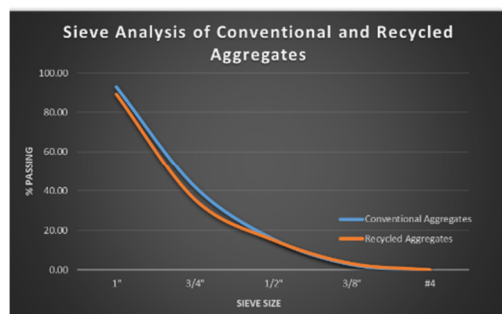


Fig. 3. Sieve analysis of coarse aggregates.

Table I shows the basic properties of both conventional and recycled aggregates.

TABLE I. PROPERTIES OF AGGREGATES

Description	CA	RA
Water absorption	1.14	2.51
Specific gravity	2.61	2.65
Fineness modulus	6.2	6.43
Surface texture	Fairly smooth	Rough
Appearance	Angular	Angular

### IV. RESULTS AND DISCUSSION

After curing, the samples were tested on a Universal Testing Machine (UTM) under a central point load following the provisions of ASTM C-293 [32], as shown in Figure 5. The samples were carefully monitored during testing, and the crushing load and central deflection were recorded at failure. The average of the observations in each batch and curing age were calculated. Flexural strength was evaluated using the standard formula:

$$\sigma_f = \frac{3PL}{2bd^2} \quad (1)$$

where the  $\sigma_f$  is flexural strength,  $P$  is failure load,  $L$  is length, and  $b_d$  is the cross-section of the prism. Tables II and III show the average values of load, deflection, and flexural strength for 7- and 28-days cured specimens, respectively.



Fig. 4. Prism specimens.



Fig. 5. Testing of specimens.

TABLE II. AVERAGE LOAD, DEFLECTION, AND FLEXURAL STRENGTH OF 7-DAY CURED SPECIMENS

Concrete	Load (kN)	Deflection (mm)	Flexural Strength (MPa)
Conventional Concrete	3.87	0.7	2.90
Concrete with 20% RCA	3.51	0.9	2.63
Concrete with 30% RCA	3.11	1.1	2.34
Concrete with 40% RCA	2.98	1.3	2.23
Concrete with 50% RCA	2.65	1.5	1.99
Concrete with 60% RCA	2.07	1.6	1.55
Concrete with 70% RCA	1.95	1.8	1.46



TABLE III. AVERAGE LOAD, DEFLECTION, AND FLEXURAL STRENGTH OF 28-DAY CURED SPECIMENS

Concrete	Load (kN)	Deflection (mm)	Flexural Strength (MPa)
Conventional Concrete	4.42	0.6	3.31
Concrete with 20% RCA	4.12	0.7	3.09
Concrete with 30% RCA	3.81	0.9	2.85
Concrete with 40% RCA	3.66	1.1	2.74
Concrete with 50% RCA	3.15	1.2	2.36
Concrete with 60% RCA	2.60	1.4	1.95
Concrete with 70% RCA	2.36	1.5	1.77

Figures 6 and 7 show a comparison of the average flexural strength for the 7- and 28-days cured samples of the proposed with the conventional no-fines concrete. The average flexural strength of the no-fines conventional concrete was recorded at 2.9MPa for 7- and 3.31MPa for 28-day cured samples, 71% and 62% higher than in [21]. The recycled concrete samples achieved 2.63MPa to 1.46MPa from 20% to 70% replacement of conventional coarse aggregates, which were better for up to 40% replacement than in [21]. On the other hand, for 28-day cured samples up to 50% replacement, the results were higher than in [21]. Similarly to the 7-day cured samples, they showed a decreasing trend in flexural strength with increasing recycled aggregates. The maximum strength (3.31MPa) was observed at 20% replacement, while the minimum (1.77MPa) was recorded for the maximum replacement. Table IV shows the percentile difference in flexural strength of no-fines recycled aggregate concrete concerning the conventional. The reduction in strength is mainly attributed to the attached mortar, the age, and the exposure of the old concrete.

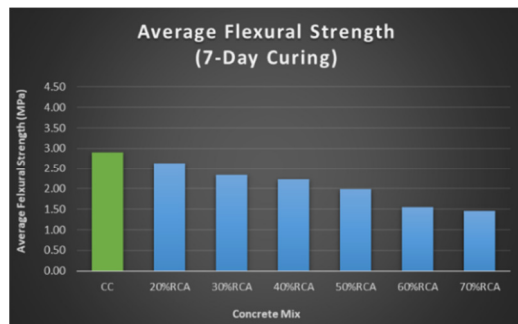


Fig. 6. Average flexural strength (7-day curing).

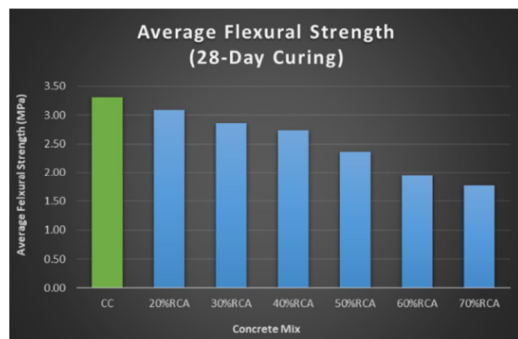


Fig. 7. Average flexural strength (28-day curing).

Figures 8 and 9 show a comparison of the average central point deflection at failure load in all prisms to no-fines conventional concrete. An increasing pattern in deflection with increasing the replacement of coarse aggregates is recorded. The maximum deflection was 1.8mm and 1.5mm for the 7- and 28-day cured specimens, respectively. Table IV also lists the percentile difference in deflection for both curing ages of the proposed concrete specimens. The maximum deflection was about 2.5 times the deflection of the conventional 7-day cured no-fines concrete samples. It is also noted that all deflection values are within the allowable limits of ACI-318 [33] for simply supported beams.

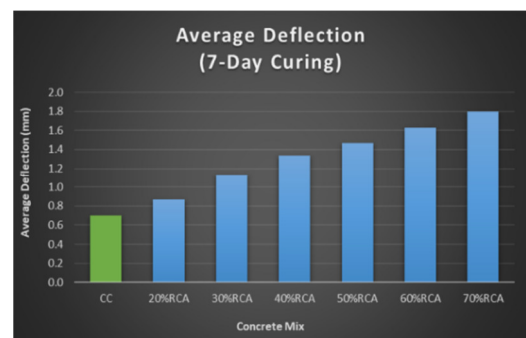


Fig. 8. Average deflection (7-day curing).

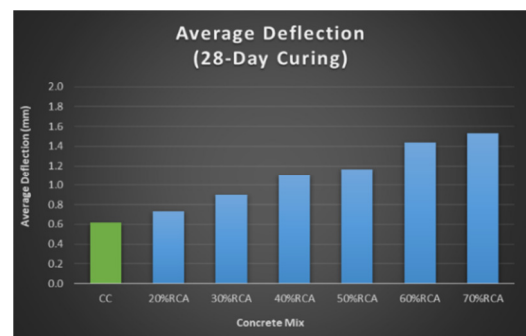


Fig. 9. Average deflection (28-day curing).

TABLE IV. PERCENTILE DIFFERENCE OF FLEXURAL STRENGTH AND DEFLECTION

Concrete Mix	7-Day Curing		28-Day Curing	
	Deflection	FS	Deflection	FS
Conventional Concrete	--	--	--	--
Concrete with 20% RCA	23.81	-9.22	18.92	-6.81
Concrete with 30% RCA	61.90	-19.45	45.95	-13.81
Concrete with 40% RCA	90.48	-22.99	78.38	-17.18
Concrete with 50% RCA	109.52	-31.42	89.19	-28.74
Concrete with 60% RCA	133.33	-46.58	132.43	-41.11
Concrete with 70% RCA	157.14	-49.61	148.65	-46.55

The observed failure mode of the beams was flexural failure, as shown in Figure 10. As plain concrete prisms were used in this study, flexural failure confirmed the concrete failure mode. Hence, it was concluded that recycled aggregates from demolished waste are suitable for producing no-fines concrete. However, considering the flexural strength and its loss (about 7%) compared to conventional no-fines concrete, the 20% replacement was considered as optimal.





Fig. 10. Failure mode.

## V. CONCLUSION

Based on the experimental investigation on the flexural strength of no-fines recycled aggregate concrete, the following can be concluded:

- The flexural strength of no-fines recycled aggregate concrete decreases with increasing replacement level of coarse aggregates.
- Central point deflection increases with increasing replacement level of coarse aggregates.
- The 20% replacement specimens, cured for 28 days, had maximum residual strength.
- All recorded deflections for both 7- and 28-day cured samples were less than the allowable deflection for simply supported beams as per ACI-318.
- All prisms failed in flexure mode at the center.

Therefore, it can be concluded that demolished waste may be used as recycled aggregates in no-fines concrete with 20% as the optimum replacement level for conventional aggregates.

## REFERENCES

- [1] 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, 2016.
- [2] S. C. Kou, C. S. Poon, and H. W. Wan, "Properties of concrete prepared with low-grade recycled aggregates," *Construction and Building Materials*, vol. 36, pp. 881–889, Nov. 2012, <https://doi.org/10.1016/j.conbuildmat.2012.06.060>.
- [3] M. Oad, A. H. Buller, B. A. Memon, and N. A. Memon, "Flexural Stress-Strain Behavior of RC Beams Made with Partial Replacement of Coarse Aggregates with Coarse Aggregates from Old Concrete: Part-1: 1:2:4 Ratio," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 3048–3053, Jun. 2018, <https://doi.org/10.48084/etasr.1982>.
- [4] B. S. Hamad and A. H. Dawi, "Sustainable normal and high strength recycled aggregate concretes using crushed tested cylinders as coarse aggregates," *Case Studies in Construction Materials*, vol. 7, pp. 228–239, Dec. 2017, <https://doi.org/10.1016/j.cscm.2017.08.006>.
- [5] S. Tanwani and B. A. Memon, "Tensile Strength Of Concrete Cylinders Made By Partial Replacement Of Natural Coarse Aggregates With Coarse Aggregates From Old Concrete," *International Journal of Emerging Technology and Innovative Engineering*, vol. 1, no. 4, 2015, <https://doi.org/10.13140/RG.2.2.26278.37443>.
- [6] A. Gholampour and T. Ozbakkaloglu, "Time-dependent and long-term mechanical properties of concretes incorporating different grades of coarse recycled concrete aggregates," *Engineering Structures*, vol. 157, pp. 224–234, Feb. 2018, <https://doi.org/10.1016/j.engstruct.2017.12.015>.
- [7] A. Patel, D. Jobanputra, and P. Patel, "A Review on No-Fines Concrete Application to Pavements," *International Journal of Innovative Science and Research Technology*, vol. 4, no. 1, Jan. 2019.
- [8] 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, Apr. 2018.
- [9] A. Abinayaa and V. Jeyasudha, "Mechanical Properties of No Fines - Pervious Concrete," *International Research Journal of Engineering and Technology*, vol. 5, no. 2, pp. 828–831, Feb. 2018.
- [10] 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>.
- [11] 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.
- [12] K. B. Thombre, A. B. More, and S. R. Bhagat, "Investigation of Strength and Workability in No-Fines Concrete," *International Journal of Engineering Research and Technology*, vol. 5, no. 9, pp. 390–393, Sep. 2016.
- [13] 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.
- [14] 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>.
- [15] K. Rubini, E. Sampathkumar, D. C. Santhosh, and M. Ramprasath, "An Experimental Study on No Fines Concrete with Partially Replacement of Cement by Silica Fume," *International Journal of Engineering Research & Technology*, vol. 6, no. 3, pp. 390–394, Mar. 2017, <https://doi.org/10.17577/IJERTV6IS030353>.
- [16] M. S. Ayubi and M. T. Ahmed, "Influence of Limestone Dust on No-Fines Concrete," *IOSR Journal of Mechanical and Civil Engineering*, vol. 14, no. 2, pp. 12–16, May 2017, <https://doi.org/10.9790/1684-1402081216>.
- [17] U. M. Muthaiyan and S. Thirumalai, "Studies on the properties of pervious fly ash-cement concrete as a pavement material," *Cogent Engineering*, vol. 4, no. 1, Jan. 2017, Art. no. 1318802, <https://doi.org/10.1080/23311916.2017.1318802>.
- [18] 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>.
- [19] M. Hashemi, P. Shafigh, M. Abbasi, and I. Asadi, "The effect of using low fines content sand on the fresh and hardened properties of roller-compacted concrete pavement," *Case Studies in Construction Materials*, vol. 11, Dec. 2019, Art. no. e00230, <https://doi.org/10.1016/j.cscm.2019.e00230>.
- [20] 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>.
- [21] O. G. Kumar and M. P. Kumar, "Effect on Mechanical Properties of No-Fine Concrete by Partial Replacement of Fly Ash in Cement with Polypropylene Fiber and Hydrogen Peroxide," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 6, no. 12, pp. 23031–23038, 2017, <https://doi.org/10.15680/IJIRSET.2017.0612014>.

- [22] A. Malik, "An Experimental Study on Properties of No-Fines Concrete," *Imperial Journal of Interdisciplinary Research*, vol. 2, no. 10, Sep. 2016.
- [23] Panimalar Engineering College, M. Mageswari, M. P. Karthikeyan, S. Pavithran, M. R. kumar, and R. G. rajan, "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>.
- [24] F. Monika, "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.
- [25] D. G. N. Devi, "Development of High Quality No Fine Concrete," *Journal of Advances in Natural Sciences*, vol. 3, no. 2, pp. 277–282, Oct. 2016, <https://doi.org/10.24297/jns.v3i2.3973>.
- [26] S. S. K and S. B. T, "Study and Comparison of Mechanical Properties, Durability and Permeability of M15, M20, M25 Grades of Pervious Concrete with Conventional Concrete," *International Journal of Applied Research*, vol. 1, no. 10, pp. 676–681, Oct. 2015.
- [27] M. F. Qasim, Z. K. Abbas, and S. K. Abed, "Producing Green Concrete with Plastic Waste and Nano Silica Sand," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7932–7937, Dec. 2021, <https://doi.org/10.48084/etasr.4593>.
- [28] D. D. Nguyen and V. T. Phan, "Compressive Strength Studies on Recycled Binder Concrete," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7332–7335, Aug. 2021, <https://doi.org/10.48084/etasr.4230>.
- [29] 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>.
- [30] A. H. Buller, N. M. Husain, M. Oad, B. A. Memon, and I. N. Sodhar, "Investigating the Deflection and Strain of Reinforced Green Concrete Beams Made With Partial Replacement of RCA under Sustained Loading," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9203–9207, Oct. 2022, <https://doi.org/10.48084/etasr.5170>.
- [31] "Standard Test Method for Sieve Analysis of Fine and Coarse Aggregates," American Society for Testing and Materials, West Conshohocken, PA, USA, Standard ASTM C136/C136M-08, 2008.
- [32] "Standard Test Method for Flexural Strength of Concrete (Using Simple Beam With Center-Point Loading)," American Society for Testing and Materials, West Conshohocken, PA, USA, Standard ASTM C293/C293M-16, 2016.
- [33] "Building Code Requirements for Structural Concrete (ACI 318-05) and Commentary (ACI 318R-05)," American Concrete Institute, Farmington Hills, MI, USA, ACI 318-05, 2005.

# A Comparative Numerical Study between Minorant Functions and Line Search Methods in Penalty Methods for Linear Optimization

Assma Leulmi

Department of Mathematics, Ferhat Abbas University Setif-1, Algeria  
as\_smaleulmi@yahoo.fr  
(corresponding author)

Received: 16 November 2022 | Revised: 27 November 2022 | Accepted: 2 December 2022

## ABSTRACT

The aim of this paper is to present a comparative numerical study between the minorant functions and line search methods in computing the step size in the penalty method for linear optimization. The minorant functions were confirmed by many interesting numerical experimentations to be more beneficial than the classical line search methods.

**Keywords**-linear optimization; penalty methods; line search; minorant function

## I. INTRODUCTION

Optimization is an especially interesting topic in data science. Linear Optimization (LO) or Linear Programming (LP) problems form an important class of optimization problems, aiming to find the feasible region and optimize the solution in order to have the highest or lowest value of a function. It is now well established as an important and very active branch of applied mathematics. The wide applicability of LO models and its richness as a mathematical theory that underlines these models and the methods developed to solve them have been the driving forces behind the rapid and continuing evolution of the subject. The template is general enough to express many different problems in engineering, industry, commerce, economics, business administration, physical sciences, and mathematics, or in any other area where decisions (in a board sense) must be taken in some complex (or conflicting) situation that can be represented by a mathematical model. There are two classes of methods for the resolution of LO problems, simplex method and interior point methods. In this study, we are interested in interior point methods. These are efficient methods developed to solve LO and NP problems. Several algorithms have been proposed to solve LO problems. Some fundamental classes of interior point methods are the affine method, the projective method with the potential reduction of Karmarkar and its alternatives [5, 9], central trajectory methods, and penalty/barrier methods [6, 13]. Our work is based on the latter type of interior point methods for solving LO problems.

In this paper, we propose a logarithmic barrier interior-point method for solving LO problems. In fact, the main difficulty to be anticipated in establishing iterations in such a method will come from the determination and computation of the step-size. The aim of this paper is to present a comparative

numerical study between the line search methods and the minorant function to compute the step-size along the direction in barrier logarithmic methods.

## II. PROBLEM FORMULATION

We consider the following LO problem

$$\begin{cases} \min_x b^T x \\ A^T x \geq c, x \in R^m. \end{cases} \quad (1)$$

where  $A \in R^{m \times n}$ , such that  $\text{rang} A = m < n$ ,  $c \in R^n$  and  $b \in R^m$ . The problem (1) is the dual of the following linear program:

$$\begin{cases} \max_y c^T y \\ A^T y = b \\ y \in R^n, y \geq 0. \end{cases} \quad (2)$$

The problem (1) can be written in the following standard form:

$$\begin{cases} \min_x b^T x \\ A^T x - c = s \\ x \in R^m, s \in R^n, s \geq 0. \end{cases} \quad (3)$$

One of the advantages of problem (1) with respect to its dual problem (2) is that the variable of the objective function is a vector instead of a matrix. Furthermore, under certain convenient hypothesis, the resolution of problem (1) is equivalent to (2) in the sense that the optimal solution of one of the two problems can be reduced directly from the other through the application of the theorem of the sladeness complementary [15]. In the rest of the paper, the following are denoted:

- $X = \{x \in R^m : A^T x - c \geq 0\}$ , the set of feasible solutions of (1).
- $\hat{X} = \{x \in R^m : A^T x - c > 0\}$ , the set of strictly feasible solutions of (1).
- $F = \{y \in R^n : Ay = b, y \geq 0\}$ , the set of feasible solutions of (2).
- $\hat{F} = \{y \in R^n : Ay = b, y > 0\}$ , the set of strictly feasible solutions of (2).

Let  $u, v \in R^n$ . Their scalar product is defined by:

$$\langle u, v \rangle = u^T v = \sum_{i=1}^n u_i v_i$$

We suppose that the sets  $\hat{X}$  and  $\hat{F}$  are not empty.

#### A. The Perturbed Problem of (1)

Problem (1) is approximated by the following perturbed problem:

$$\begin{cases} \min_x f_\eta(x) \\ x \in R^m. \end{cases} \quad (4)$$

where  $\eta > 0$  is the the penalty parameter and  $f_\eta$  is the barrier function defined by:

$$f_\eta(x) = \begin{cases} b^T x + n\eta \ln \eta - \eta \ln \sum_{i=1}^n \langle e_i, A^T x - c \rangle & \text{if } A^T x - c > 0 \\ +\infty & \text{if not} \end{cases}$$

where  $(e_1, e_2, \dots, e_n)$  is the canonical base in  $R^n$ . We are interested in solving the problem (4).

The idea of this new approach consists to introduce one original process to calculate the step-size based on minorant functions. The main advantage of (4) resides in the strict convexity of its objective function and its feasible domain. Consequently, the conditions of optimality are necessary and sufficient. This fosters theoretical and numerical studies of the problem. In the next section, the existence and uniqueness of the optimal solution of (4) is proved and we show its convergence to (1), in particular the behavior of its optimal value and its primal solutions when  $\eta \rightarrow 0$ , then  $\lim_{\eta \rightarrow 0} x_\eta = x^*$  is an optimal solution of (1). In Section III, we propose an interior point algorithm based on the Newton's approach which allows us to solve the nonlinear system resulting from the optimality conditions. The iteration of this algorithm is of descent type, defined by:  $x_{k+1} = x_k + \alpha_k d_k$ , where  $d_k$  is the descent direction and  $\alpha_k$  is the step-size. Also, we present different steps-sizes by minimizing a minorant function which approximates the one-dimensional function  $\theta(\alpha_k) = \min_{\alpha > 0} f(x + \alpha d)$ . The last Section is dedicated to the presentation of comparative numerical tests to illustrate the effectiveness of our approach and to determine the most efficient algorithm.

The main advantage of (4) resides in the strict convexity of its objective function and its feasible domain. Consequently, the conditions of optimality are necessary and sufficient. This fosters theoretical and numerical studies of the problem. Before this, it is necessary to show that (4) has at least an optimal solution.

#### B. Convergence of the Perturbed Problem

Firstly, we give the following definition:

**Definition:** Let  $f$  be a function defined from  $R^m$  to  $R \cup \{\infty\}$ .  $f$  is called inf-compact if for all  $\eta > 0$ , the set  $X_\eta(f) = \{x \in R^m : f(x) \leq \eta\}$  is compact, which comes in particular to say that its cone of recession is reduced to zero.

To prove that (4) has an optimal solution, we show that  $f_\eta$  is inf-compact. For that, it is enough to prove that the cone of recession  $\hat{X}((f_\eta)_\infty) = \{d \in R^n : (f_\eta)_\infty(d) \leq 0\}$ , is reduced to the origin, i.e.  $((f_\eta)_\infty(d) \leq 0) \Rightarrow (d = 0)$ , where  $(f_\eta)_\infty$  is defined by:

$$(f_\eta)_\infty(d) = \lim_{\alpha \rightarrow +\infty} \frac{f_\eta(x + \alpha d) - f_\eta(x)}{\alpha} = b^T d.$$

This needs the following proposition:

**Proposition:**  $d = 0$  whenever  $b^T d \leq 0$  and  $A^T d \in \hat{X}$ .

Then, the problem (4) has an optimal solution. We know that the Hessian matrix  $H = \nabla^2 f_\eta(x)$  is positive definite, then the problem (4) is strictly convex, and if it has an optimal solution, then it is unique. We have:

$$f_\eta(x) = b^T x + n\eta \ln \eta - \eta \ln \sum_{i=1}^n \langle e_i, A^T x - c \rangle$$

Then:

$$\nabla f_\eta(x) = b - \eta \sum_{i=1}^n \frac{Ae_i}{\langle e_i, A^T x - c \rangle}$$

and:

$$\nabla^2 f_\eta(x) = \eta \sum_{i=1}^n \frac{Ae_i(Ae_i)^T}{\langle e_i, A^T x - c \rangle^2}$$

As  $f_\eta$  is inf-compact and strictly convex, therefore the problem (4) admits a unique optimal solution. We denote by  $x(\eta)$  or  $x_\eta$  the unique optimal solution of (4).

#### C. Convergence of the Perturbed Problem to the Initial Problem

For  $x \in \hat{X}$ , let's introduce the symmetrical definite positive matrix  $B_i$  of rank  $m, i = 1, \dots, n$  and the lower triangular matrix  $L$ , such that  $B_i = Ae_i(Ae_i)^T = LL^T$ , which implies that  $H$  is a positive definite matrix. In what follows, we will be interested in the behavior of the optimal value and the optimal solution  $x(\eta)$  of problem (4).

**Proposition:** For  $\eta > 0$ , let  $x_\eta$  an optimal solution of the problem (4) then there is an optimal solution of (1)  $x \in X$ , such that,  $\lim_{\eta \rightarrow 0} x_\eta = x$ .

**Remark:** We know that if one of the problems (1) and (2) has an optimal solution, and the values of their objective functions are equal and finite, the other problem has an optimal solution.

### III. RESOLUTION OF THE PERTURBED PROBLEM

In this part, we are interested in the numerical solution of problem (4). We use a logarithmic barrier interior point method. This method types are based on the optimality conditions which are necessary and sufficient, and consist of constructing a sequence of iterate  $x_{k+1} = x_k + \alpha_k d_k$ , where  $x_\eta$  is an optimal solution of (4) if it satisfies the following condition:

$$\nabla f_\eta(x_\eta) = 0 \quad (5)$$

To solve (5) we use the Newton's approach which means to find in each iteration a vector  $x_{\eta k} + d_k$  checking the following linear system:

$$H_k d_k = -\nabla f_\eta(x_{\eta k}). \quad (6)$$

As  $H_k = \nabla^2 f_\eta(x_{\eta k})$  is a symmetric positive definite matrix, the Cholesky methods and the conjugate gradient methods are the best convenient for solving (6). To ensure the convergence of the algorithm towards an optimal solution  $x^*$  of (4), it should be made sure that all the iterations  $x_{\eta k} + d_k$  remain strictly feasible. For that, we introduce a step-size  $\alpha_k$  checking the condition:

$$A^T(x_{\eta k} + \alpha_k d_k) - c > 0$$

#### A. Effectual Computation of the Step Size

There are two main techniques used for computing the displacement step  $\alpha_k$ .

##### 1) Line Search

These methods try out a sequence of candidate values of  $\alpha_k$ , stopping to accept one of these values when some conditions are satisfied. An ideal choice of the step-size is the global minimization of the one-dimensional function  $\theta(\cdot)$  defined by:

$$\theta(\alpha_k) = \min_{\alpha > 0} f(x + \alpha d)$$

The most used line search methods are Wolfe, Goldstein—Armijo, Fibonacci, etc. Unfortunately, these methods have big computational cost.

##### 2) Principle of Approximate Function

A minorant function  $\tilde{\varphi}$  must be close to:

$$\varphi(\alpha) = \frac{1}{\eta} [f_\eta(x + \alpha d) - f_\eta(x)]$$

which must give the  $\min_\alpha \tilde{\varphi}(\alpha)$  in  $[0, \hat{\alpha}]$  by a simple and easy manner, which permits the computation of the step-size at each iteration in a relatively short time and with a smaller number of instructions in contrast to line search technique.

Authors in [12] gave a simple form for the function, which is presented in the following proposition:

**Proposition [12]:** Let  $\hat{\alpha} = \sup\{\alpha: 1 + z_i \alpha\}$  with  $z_i = \frac{\langle e_i, A^T d \rangle}{\langle e_i, A^T x - c \rangle}$ ,  $\forall i = 1, \dots, n$ . For all  $\alpha \in [0, \hat{\alpha}]$ , the following function  $\varphi(\alpha)$  is well defined:

$$\varphi(\alpha) = n(\sum_{i=1}^n z_i) \alpha - \|z\|^2 \alpha - \sum_{i=1}^n \ln(1 + z_i \alpha)$$

such that,  $\varphi(\alpha)$  verifies the following properties:

$$\|z\|^2 = n(\bar{z}^2 + \sigma_z^2) = \varphi''(0) = -\varphi'(0), \varphi(0) = 0$$

#### B. Minorant Function

Authors in [10] proposed 3 minorant functions in 2019. In this paper, we are interested in their best minorant function defined as:

$$\tilde{\varphi}_1(\alpha) = \delta \alpha - (n-1) \ln(1 + \beta \alpha) - \ln(1 + \gamma \alpha)$$

with:

$$\delta = n\bar{z} - \|z\|^2$$

$$\beta = \bar{z} - \frac{\sigma_z}{\sqrt{n-1}}$$

$$\gamma = \bar{z} + \sigma_z \sqrt{n-1}.$$

In addition, they proved that the minorant function  $\tilde{\varphi}_1$  is defined and convex on  $[0, \hat{\alpha}]$ ,  $\varphi(\alpha) > \tilde{\varphi}_1(\alpha)$  ( $\tilde{\varphi}_1$  minorant function on  $[0, \hat{\alpha}]$ ), and the function  $\tilde{\varphi}_1$  verifies the following properties:

$$\|z\|^2 = n(\bar{z}^2 + \sigma_z^2) = \tilde{\varphi}''(0) = -\tilde{\varphi}'(0), \tilde{\varphi}(0) = 0$$

The minimum of  $\tilde{\varphi}_1$  is obtained in  $\bar{\alpha}_i = \alpha_{opt}$ , such that,  $\tilde{\varphi}_1'(0) = 0$ . We are then coming back to solve the second order following equation:  $\alpha^2 - 2b\alpha + c = 0$ , with:

$$b = \frac{1}{2} \left( \frac{n}{\delta} - \frac{1}{\beta} - \frac{1}{\gamma} \right) \text{ and } c = \frac{-\|z\|^2}{\beta\gamma\delta}$$

The roots of this equation are of the type  $\bar{\alpha} = b \pm \sqrt{b^2 - c}$ . Let's take one root of the two that belong to  $[0, \hat{\alpha}]$ . Thus, the  $\bar{\alpha}$  is explicitly computed, then, we consider it belongs to the interval  $(0, \hat{\alpha} - \varepsilon)$  and  $\varphi'(\alpha) < 0$ , with  $\varepsilon > 0$  being a fixed precision.

**Remark:** The calculation of  $\bar{\alpha}$  is performed by a dichotomous procedure, in the cases where  $\bar{\alpha}_i \notin (0, \hat{\alpha} - \varepsilon)$ , and  $\varphi'(\alpha) > 0$ , as follows:

Put  $\alpha = 0$  and  $b = \hat{\alpha} - \varepsilon$  while  $|b - a| > 10^{-4}$ . If  $\varphi\left(\frac{a+b}{2}\right) < 0$ , then  $b = \frac{a+b}{2}$ . Else  $a = \frac{a+b}{2}$ , so  $\bar{\alpha} = b$ . This calculation guarantees a better approximation of the minimizer of  $\tilde{\varphi}'(\alpha)$  while remaining in the domain of  $\varphi$ .

**Proposition [10]:** Let  $x_{k+1}$  and  $x_k$  be two strictly feasible solutions of (4) obtained respectively at the  $k+1$  and  $k$  iterations, so we have  $f_\eta(x_{k+1}) < f_\eta(x_k)$ .

### IV. ALGORITHM DESCRIPTION AND NUMERICAL RESULTS

In this section, we present the algorithm of our approach to obtain an optimal solution  $\bar{x}$  to problem (1) and some numerical tests.

#### A. The Algorithm

In this section, we present the algorithm of our approach to obtain an optimal solution  $\bar{x}$  to the problem (1).

For simplicity, we consider  $x_k$  instead of  $x_{\eta k}$  and  $x$  instead of  $x_\eta$ .

Begin algorithm

Initialization

$x_0$  is a strictly feasible solution of  
(1),  $d_0 \in R^m$  and  $\varepsilon > 0$  is a given precision.

Iteration

While  $|b^T d_k| > \varepsilon$  do

Solve the system  $H_k d_k = -\nabla f_\eta(x_{\eta k})$ .

Compute the step-size using the strategy  
of minorant function or line search  
method.

Take the new iterate  $x_{k+1} = x_k + \alpha_k d_k$ .

Take  $k = k + 1$ .

End while

End algorithm

This approach tries to reduce the number of iterations and the time of calculation. Some examples are presented below.

### B. Numerical Results

To measure the numerical performance of the proposed methods, we present a numerical comparison of the results obtained by the proposed algorithm using the minorant function given in [10] to compute the step-size and those obtained by using the line search Wolfe's method. We use examples with fixed and variable sizes to carry out the numerical tests.

The following examples are taken from the literature [1, 4, 8] and were implemented in MATLAB. We took  $\varepsilon = 1.0e - 006$ . In the result table:

- (size) represents the size of the example.
- (itrat) represents the number of iterations necessary to obtain an optimal solution.
- (time) represents the time of computation in seconds (s).
- (mf st) represents the strategy that uses minorant function.
- (lr st) represents the strategy that uses Wolfe's line search.

Recall that the considered problem (1) is:

$$\begin{cases} \min_x b^T x \\ A^T x \geq c, x \in R^n \end{cases}$$

We note that the matrices used in the numerical tests are full matrices.

#### 1) Examples with Fixed Size

Example 01:

$$A = \begin{bmatrix} 2 & 3 & 1 & 2 \\ 3 & 0 & -2 & 1 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \text{ and } c = [4 \quad 1 \quad 2 \quad 0]^T$$

- The initial strictly feasible point is  $x^0 = [1 \quad 1.5 \quad 1 \quad 1]^T$ .
- The optimal solution is  $x^0 = [0 \quad 0.67 \quad 0 \quad 0]^T$ .

Example 02:

$$A = \begin{bmatrix} 2 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{and } c = [3 \quad -1 \quad 1 \quad 0 \quad 0 \quad 0]^T$$

- The initial strictly feasible point is  $x^0 = [1 \quad 1 \quad 2]^T$ .
- The optimal solution is  $x^* = [0.5 \quad 0.0713 \quad 0.5]^T$ .

Example 03:

$$A = \begin{bmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 6 \\ 2 \\ 4 \end{bmatrix}$$

$$\text{and } c = [4 \quad -2 \quad -2 \quad 0 \quad 0 \quad 0]^T$$

- The initial strictly feasible point is  $x^0 = [0.5 \quad 1 \quad 1]^T$ .
- The optimal solution is  $x^0 = [0 \quad 0 \quad 0]^T$ .

Example 04:

$$A = \begin{bmatrix} 1 & 0 & -4 & 3 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 5 & 3 & 1 & 0 & -1 & 3 & 0 & 1 & 0 & 0 & 0 & 0 \\ 4 & 5 & -3 & 3 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2 & 1 & -5 & 0 & 0 & 0 & 1 & 0 & 0 \\ -2 & 1 & 1 & 1 & 2 & 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & -3 & 2 & -1 & 4 & 5 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b = [1 \quad 4 \quad 4 \quad 5 \quad 7 \quad 5]^T$$

$$c = [-4 \quad -5 \quad -1 \quad -3 \quad 5 \quad -8 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]^T$$

- The initial strictly feasible point is:  
 $x^0 = [-2 \quad 4 \quad 1 \quad 1 \quad 1 \quad 1]^T$
- The optimal solution is  
 $x^* = [0.5 \quad 1.5 \quad 0 \quad 0 \quad 1.5 \quad 0]^T$

Example 05:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 5 & 4 & 3 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 6 & 7 & 8 & 9 & 10 & 5 & 2 & 8 & 3 & 1 & 0 & 1 & 0 & 0 & 0 \\ 11 & 12 & 13 & 14 & 15 & 6 & 7 & 80 & 90 & 10 & 0 & 0 & 1 & 0 & 0 \\ 1 & 10 & 20 & 30 & 40 & 50 & 60 & 80 & 90 & 10 & 0 & 0 & 0 & 1 & 0 \\ 3 & 9 & 27 & 60 & 45 & 60 & 75 & 8 & 9 & 46 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b_i = 10^4, i = 1, \dots, 5,$$

$$c = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]^T$$

- The initial strictly feasible point is  
 $x^0 = [1 \quad 1 \quad 1 \quad 1 \quad 1]^T$
- The optimal solution is  
 $x^0 = [0 \quad 0 \quad 0.0888 \quad 0 \quad 0.0078]^T$

The results of the last examples are given in Table I.

TABLE I. EXAMPLES WITH FIXED SIZE

Size	mf st		ls st	
	itrat	time	itrat	time
2 × 4	5	0.032	13	0.130
3 × 6	6	0.044	33	0.128
3 × 6	7	0.051	34	0.132
6 × 12	9	0.055	34	0.306
5 × 15	8	0.048	22	0.170



## 2) Example with Variable Size

$n = 2m$ , For  $i, j = 1, \dots, m$ ,

$$A[i, j] = 0 \text{ if } i \neq j \text{ or } (i + 1) \neq j$$

$$A[i, j] = A[i, i + m] = 1, \quad b[i] = 2.$$

- The initial strictly feasible point is  $x^0 = [1 \quad 1 \quad \dots \quad 1]^T$ .
- The optimal solution is  $x^* = [0 \quad 0 \quad \dots \quad 0]^T$ .

Table II resumes the obtained results.

TABLE II. EXAMPLES WITH FIXED SIZE

Size	mf st		ls st	
	itrat	time	itrat	time
$50 \times 100$	1	0.031	49	8.5512
$100 \times 200$	1	0.053	50	32.6145
$200 \times 400$	2	0.088	51	91.6524
$400 \times 800$	3	0.096	52	161.4374
$500 \times 1000$	3	0.12	52	411.8901

## V. CONCLUSION

The conducted numerical study clearly shows that the strategy of the minorant function seems to be more efficient than that of the line search in time and number of iterations. Our future work will consider further improving the computational time of the logarithmic barrier algorithm by proposing another, better, approximate function. But the extensions would be envisaged to the nonlinear, and not necessarily relevant to the LO problem.

## REFERENCES

- [1] M. Achache, "A polynomial-time weighted path-following interior-point algorithm for linear optimization," *Asian-European Journal of Mathematics*, vol. 13, no. 2, Mar. 2020, Art. no. 2050038, <https://doi.org/10.1142/S1793557120500382>.
- [2] P. Armaos, "A Study of Joint Cost Inclusion in Linear Programming Optimization," *Engineering, Technology & Applied Science Research*, vol. 3, no. 4, pp. 473–478, Aug. 2013, <https://doi.org/10.48084/etasr.327>.
- [3] B. Badri-Koochi, R. Tavakkoli-Moghaddam, and M. Asghari, "Optimizing Number and Locations of Alternative-Fuel Stations Using a Multi-Criteria Approach," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3715–3720, Feb. 2019, <https://doi.org/10.48084/etasr.2474>.
- [4] M. Bouafia, D. Benterki, and A. Yassine, "A new efficient short-step projective interior point method for linear programming," *Operations Research Letters*, vol. 46, no. 3, pp. 291–294, May 2018, <https://doi.org/10.1016/j.orl.2018.02.004>.
- [5] J.-P. Crouzeix and B. Merikhi, "A logarithm barrier method for semi-definite programming," *RAIRO - Operations Research*, vol. 42, no. 2, pp. 123–139, Apr. 2008, <https://doi.org/10.1051/ro:2008005>.
- [6] J.-P. Chehab and M. Raydan, "Geometrical properties of the Frobenius condition number for positive definite matrices," *Linear Algebra and its Applications*, vol. 429, no. 8, pp. 2089–2097, Oct. 2008, <https://doi.org/10.1016/j.laa.2008.06.006>.
- [7] R. M. Freund and S. Mizuno, "Interior Point Methods: Current Status and Future Directions," in *High Performance Optimization*, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, Eds. Boston, MA, USA: Springer US, 2000, pp. 441–466, [https://doi.org/10.1007/978-1-4757-3216-0\\_18](https://doi.org/10.1007/978-1-4757-3216-0_18).
- [8] N. Karmarkar, "A new polynomial-time algorithm for linear programming," in *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, New York, NY, USA, Sep. 1984, pp. 302–311, <https://doi.org/10.1145/800057.808695>.
- [9] A. Leulmi and S. Leulmi, "Logarithmic Barrier Method Via Minorant Function for Linear Programming | Journal of Siberian Federal University," *Journal of Siberian Federal University. Mathematics & Physics*, vol. 12, no. 2, pp. 191–201, 2019, <https://doi.org/10.17516/1997-1397-2019-12-2-191-201>.
- [10] A. Leulmi, B. Merikhi, and D. Benterki, "Study of a Logarithmic Barrier Approach for Linear Semidefinite Programming," *Journal of Siberian Federal University. Mathematics & Physics*, vol. 11, no. 3, pp. 1–13, 2018, <https://doi.org/10.17516/1997-1397-2018-11-3-300-312>.
- [11] N. Karmarkar, "A new polynomial-time algorithm for linear programming," in *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, New York, NY, USA, Sep. 1984, pp. 302–311, <https://doi.org/10.1145/800057.808695>.
- [12] H. Mansouri and M. Zangiabadi, "An adaptive infeasible interior-point algorithm with full-Newton step for linear optimization," *Optimization*, vol. 62, no. 2, pp. 285–297, Feb. 2013, <https://doi.org/10.1080/02331934.2011.611881>.
- [13] M. R. Rezoug, M. Benaouadj, D. Taibi, and R. Chenni, "A New Optimization Approach for a Solar Tracker Based on an Inertial Measurement Unit," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7542–7550, Oct. 2021, <https://doi.org/10.48084/etasr.4330>.
- [14] H. Wolkowicz and G. P. H. Styan, "Bounds for eigenvalues using traces," *Linear Algebra and its Applications*, vol. 29, pp. 471–506, Feb. 1980, [https://doi.org/10.1016/0024-3795\(80\)90258-X](https://doi.org/10.1016/0024-3795(80)90258-X).
- [15] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects (Universitext)*. Berlin, Heidelberg, Germany: Springer-Verlag, 2006.
- [16] R. T. Rockafellar, *Convex Analysis: (PMS-28)*, vol. 30. Princeton, NJ, USA: Princeton University Press, 1970.

# Load Flow Analysis and the Impact of a Solar PV Generator in a Radial Distribution Network

**Mohamed Ali Zdiri**

CEM Laboratory, Engineering School of Sfax, Tunisia  
mohamed-ali.zdiri@enis.tn  
(corresponding author)

**Bilel Dhouib**

CEM Laboratory, Engineering School of Sfax, Tunisia  
dhouib.bilel@gmail.com

**Zuhair Alaas**

Department of Electrical Engineering, Faculty of Engineering, Jazan University, Saudi Arabia  
zalaas@jazanu.edu.sa

**Fatma Ben Salem**

CEM Laboratory, Engineering School of Sfax, Tunisia  
fatma\_bs@yahoo.fr

**Hsan Hadj Abdallah**

CEM Laboratory, Engineering School of Sfax, Tunisia  
hsan.hajabdallah@enis.tn

*Received: 17 November 2022 | Revised: 10 December 2022 | Accepted: 18 December 2022*

## ABSTRACT

The distribution system acts as a conduit between the consumer and the bulk power grid. Due to characteristics such as a high resistance/reactance ratio, distribution networks cannot be solved using conventional methods, such as the Gauss-Seidel and Newton-Raphson. This research proposes a method for the calculation of the power flow in radial networks that considers their wide range of resistance and reactance values, PV generator characteristics, and radial structure. An iterative methodology is employed, with each iteration beginning with the branch that has the highest accurate power flow solution. The procedure is reliable and effective over various workloads and network configurations. To confirm the effectiveness of the suggested technique on the simple and IEEE 33-bus radial distribution system, simulations were carried out in MATLAB. The implications of including a renewable energy source, such as a PV generator, in the network under consideration are investigated by simulation result comparison. The optimal location of the PV generator was also determined using an Artificial Neural Network (ANN) controller. The results of the identification process improve the already exceptional efficacy and performance of the ANN controller.

**Keywords-radial network; PV generator; power flow; identification; ANN controller; MATLAB**

## I. INTRODUCTION

Distribution systems operate as conduits between the bulk power grid and the customers. Primary or main feeders and lateral distributors are frequently used in the distribution circuits. The main feeder travels across the main load centers after leaving the substation. Radial distribution systems are referred to as lateral distributors because they join various load

sites to the main feeder. The ease of use and low cost of radial systems result in their increasing popularity.

Power flow is an essential tool in power system analysis. The maximum currents carried by distribution feeders and the associated voltage dips, annual energy loss, and dependability of meeting consumer needs must all be considered in order to design properly the expansion and operation characteristics of distribution networks. Given that each optimization study

requires several power flow runs, the effectiveness of such an algorithm is crucial. However, the numerous inputs used as the foundation for these studies, including load forecast, load model coefficients, network parameters, and formation of bus shunts, are frequently estimated with a certain degree of error. Owing to practical challenges in data collection in large and complicated distribution systems, this uncertainty is typically non-statistical in nature.

Electric distribution networks have certain inherent characteristics, including:

- Radial or weakly mesh structures
- Unbalanced operation and unbalanced distributed loads
- High number of buses and branches
- Wide range of resistance and reactance values
- Operating in multiple phases

During the past two decades, the Newton–Raphson and rapidly decoupled power flow solution approaches and their variations, have successfully been utilized to solve "well-behaved" power systems [1, 2]. However, flaws have been noted when these algorithms have been extensively applied to faulty and improperly initialized power systems. The Gauss-Seidel power flow approach, another traditional power flow method, has proven to be incredibly ineffective when fixing large power systems while being quite durable. Distribution networks fall within the category of ill-conditioned power systems for the generic Newton–Raphson and rapid decoupled power flow algorithms due to their wide range of resistance and reactance values and radial structure. Furthermore, due to the development of renewable energy sources, the resolution of distribution systems should consider them [3]. To solve the distribution system problem, the load flow method must be modified.

Power flow is often separated into two categories. The first category, which suggests recursive solutions without recurring derivatives in their formulations, comprises power flow strategies that incorporate the distribution network graph simply by rearranging the power flow equations to obtain a recursive formula [4–9]. Iterative solutions based on linear approximations belong to the second class of approaches that use Taylor's series expansion of power flow equations in real and complex domains. Additionally, the use of Distributed Generation (DG) systems, which heavily rely on renewable energy, has become increasingly popular [10], and their development has rapidly advanced [11]. The installation of DG units at the distribution level has numerous advantages, such as reduced peak loads, enhanced system security, dependability, voltage stability, grid strengthening, cutting on-peak operating costs, and reduced network loss [12, 13].

To calculate power flow, fundamental radial network assumptions are considered in this study. This approach is comparable to others where an iterative procedure is used along the branches. Because it is based on a precise power flow solution for a single branch that provides a comprehensive solution, the essential concept is offered in the literature [14, 15]. Even for extremely large power flow loads, this approach

consistently yields an accurate computation. The consequences of integrating a solar generator into the current network are the subject of the following phase of our inquiry.

This study describes a useful technique for determining power flow in radial networks. To test the accuracy and efficacy of the suggested technique on the simple and IEEE 33-bus radial distribution system, MATLAB simulations were carried out. A wide range of simulation results were examined to understand the effects of the addition of a renewable energy source such as the PV generator to the radial network. An ANN controller was used to determine the location of the PV generator that had already been integrated to the radial network. The outcomes of the identification process enhance the ANN controller's already outstanding efficacy and performance.

## II. PV-RADIAL DISTRIBUTION NETWORK

Bus injection into the branch current matrix (BIBC) and analogous current injections are used to control the load flow of a radial distribution network [16]. This section explains the development process in detail. The corresponding current injection-based model is more useful for distribution networks. The complex apparent power load  $S_i$  for the buses is expressed as follows:

$$S_i = P_i + jQ_i \quad (1)$$

where  $P_i$  and  $Q_i$  are the active and reactive load powers, respectively; and  $i$  is equal to 1, 2, 3, ...,  $n$ .

The analogous current injection that corresponds to the  $ii$ -th solution iteration is:

$$I_i^{ii} = I_i^r(V_i^{ii}) + sI_i^i(V_i^{ii}) = ((P_i + jQ_i)/V_i^{ii})^* \quad (2)$$

where  $V_i^{ii}$  and  $I_i^{ii}$  represent the bus voltage and the equivalent current injection at the  $ii$ -th iteration and  $s^2 = -1$ . The corresponding current injection of the bus at the  $ii$ -th iteration is represented by the real and imaginary components  $I_i^r$  and  $I_i^i$  respectively.

### A. Construction of Relationship Matrices

Figure 1 depicts a simple radial distribution network.

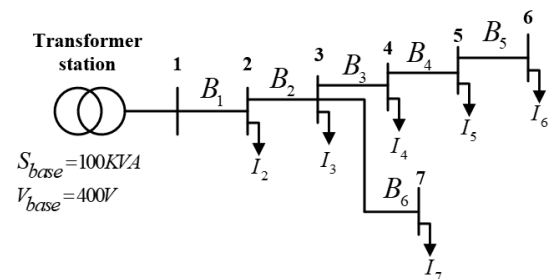


Fig. 1. Simple radial distribution network.

The injected currents are determined using (2). Kirchhoff's current law is used in the radial distribution network to determine the branch current. Thus, branch currents can be

formulated as functions of equivalent current injections. The branch currents  $B_1, B_2, B_3, B_4, B_5$ , and  $B_6$  can be represented as:

$$\begin{cases} B_1 = I_2 + I_3 + I_4 + I_5 + I_6 + I_7 \\ B_2 = I_3 + I_4 + I_5 + I_6 + I_7 \\ B_3 = I_4 + I_5 + I_6 + I_7 \\ B_4 = I_5 + I_6 \\ B_5 = I_6 \\ B_6 = I_7 \end{cases} \quad (3)$$

Consequently, the connection between the branch currents and bus current injections can be described as:

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \\ I_6 \end{bmatrix} \quad (4)$$

Equation (4) can be re-written as:

$$[B] = [BIBC] \cdot [I] \quad (5)$$

The branch injection to the branch current matrix is referred to as *BIBC*. The *BIBC* matrix only has the values 0 and 1, and it is an upper triangular matrix.

#### B. BIBC Matrices Development

The procedure for constructing the *BIBC* matrix presented in (4) is [15]:

- **Step 1:** The dimensions of the *BIBC* matrix for a distribution system with  $m$  branch sections and  $n$  buses are  $(m, n-1)$ .
- **Step 2:** The column of the *BIBC* matrix's  $i$ -th bus should be copied to the column of the  $j$ -th bus if the line section  $B_{ii}$  lies between buses  $i$  and  $j$ . Subsequently, the locations of the  $ii$ -th row and  $j$ -th bus column should be filled with +1.
- **Step 3:** Repeat Step 2 until the *BIBC* matrix contains each line portion.

#### C. Bus Voltage Development

Using the generalized equations shown below, a forward sweep across the ladder network is used to determine the receiving-end bus voltages:

$$V(m2) = V(m1) - B(jj) Z(jj) \quad (6)$$

where  $m1$  and  $m2$  are the transmitting and receiving ends, respectively and  $jj$  and  $B$  represent the branch number and current, respectively.

Equation (7) can be used to determine the relationship between the branch currents and bus voltages.

$$\begin{cases} V_2 = V_1 - B_1 Z_{12} \\ V_3 = V_2 - B_2 Z_{23} \\ V_4 = V_3 - B_3 Z_{34} \\ V_5 = V_4 - B_4 Z_{45} \\ V_6 = V_5 - B_5 Z_{56} \\ V_7 = V_6 - B_6 Z_{67} \end{cases} \quad (7)$$

The following is the matrix representation for the aforementioned relations:

$$[\Delta V] = [BCBV] \cdot [B] \quad (8)$$

with:

$$[BCBV] = \begin{bmatrix} Z_{12} & 0 & 0 & 0 & 0 & 0 \\ Z_{12} & Z_{23} & 0 & 0 & 0 & 0 \\ Z_{12} & Z_{23} & Z_{34} & 0 & 0 & 0 \\ Z_{12} & Z_{23} & Z_{34} & Z_{45} & 0 & 0 \\ Z_{12} & Z_{23} & Z_{34} & Z_{45} & Z_{56} & 0 \\ Z_{12} & Z_{23} & 0 & 0 & 0 & Z_{37} \end{bmatrix}$$

Branch current variations cause equivalent bus voltage changes, which can be calculated using the *BCBV* matrix. The relationship between bus current injections and bus voltages can be expressed as follows.

$$[V] = [V_1] - [BCBV] \cdot [BIBC] \cdot [I] \quad (9)$$

The updated voltage values are used to perform a new top-down iteration. This process is repeated until the voltage levels in a row are within the permissible range.

Furthermore, branch  $jj$ 's actual and reactive power loss is provided by:

$$\begin{cases} p_{loss} = |B(jj)|^2 \cdot |R(jj)| \\ q_{loss} = |B(jj)|^2 \cdot |X(jj)| \end{cases} \quad (10)$$

#### D. PV Bus Development

Evidently, the ideal location for a PV source on a radial distribution line that is used to obtain the least amount of loss is at the end of the line. Equation (11) states the maximum power that a solar plant can supply [17]:

$$P_{GPV} = P_1 \cdot E_c \cdot [1 + P_2 \cdot (E_c - E_{cref}) + P_3 \cdot (T_{jc} - T_{jcref})] \quad (11)$$

where  $E_c$  corresponds to panel insolation ( $\text{W/m}^2$ ),  $E_{cref}$  represents the reference insolation of  $1000 \text{W/m}^2$ ,  $T_{jcref}$  denotes the reference panel temperature of  $25^\circ\text{C}$ , and parameters  $P_1$ ,  $P_2$ , and  $P_3$  have fixed values. With only 3 constant variables and a straightforward equation, this streamlined model allows the computation of the maximum power generated by a collection of panels for a specific panel illuminance and temperature.

Bus  $i$ 's active power changes with respect to:

$$P_{ci} = P_{ci0} - P_{GPV} \quad (12)$$

where  $P_{ci0}$  is the first power consumed before the PV generator is injected.

### III. SIMULATION RESULTS AND DISCUSSION

#### A. Application on a Radial Distribution Test Network

We chose a 7-node radial distribution test network with 6 loads, 1 transformer station, and 6 lines for our study. Figure 2 shows the radial test network utilized. Note that the  $K_1$  and  $K_2$  switches are complementary.

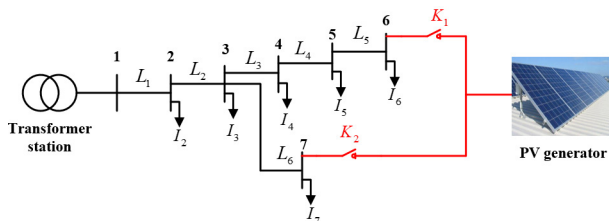


Fig. 2. Structure of the seven-node PV-radial distribution test network.

The line data and the data relating to the nodes are illustrated in Tables I and II, respectively.

TABLE I. LINE DATA

Number of the line	Liaison	Impedance (p.u.)
1	1-2	$0.015 + j0.050$
2	2-3	$0.010 + j0.070$
3	3-4	$0.030 + j0.050$
4	3-7	$0.030 + j0.075$
5	4-5	$0.025 + j0.030$
6	5-6	$0.035 + j0.015$

TABLE II. NODE DATA

Bus No	Types	Active power consumed (p.u.)	Reactive power consumed (p.u.)
1	P-V	0	0
2	P-Q	0.3	0.25
3	P-Q	0.4	0.15
4	P-Q	0.3	0.4
5	P-Q	0.7	0.45
6	P-Q	0.6	0.5
7	P-Q	0.5	0.7

Amerisolar-Worldwide Energy and Manufacturing USA Co., Ltd AS-6M-350W;  
1 series modules; 1 parallel strings

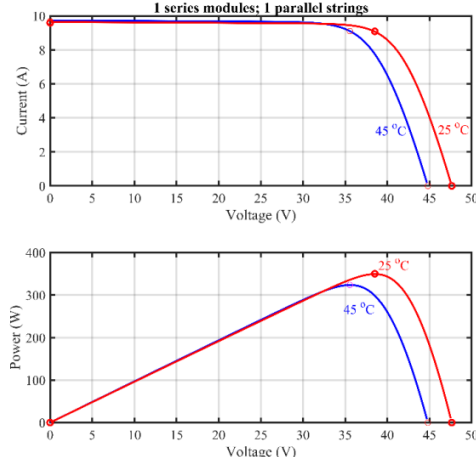


Fig. 3. PV module characteristics.

The Amerisolar-Worldwide Energy and Manufacturing USA Co., Ltd AS-6M-350W PV array type, shown in Figure 3, is suitable to perform the current requirements. The PV array parameters are reported in Table III.

TABLE III. PV MODULE PARAMETERS

Maximum power (W)	350
Voltage at MPP (V)	38.51
Current at MPP (A)	9.09
Open-circuit voltage (V)	38.51
Short-circuit current (A)	9.59

#### B. Simulation Results

In this study, two cases, with and without the PV generator (in node 6 or node 7) in the simple radial distribution network, were simulated and discussed. The *BIBC* matrix and equivalent current injection were used to determine the load flow in the radial distribution network. The load flow solution of a 7-bus distribution system for the radial network without the PV generator is summarized in Table V.

TABLE IV. LOAD FLOW SOLUTION OF THE RADIAL NETWORK

Bus N°	Voltage (p.u.)	Angle (rad.)
1	1.15	0
2	0.9769	-0.2306
3	0.9255	-0.6257
4	0.8986	-0.8477
5	0.8963	-0.9672
6	0.8748	-1.0186
7	0.8781	-0.7325

The radial distribution load flow had a total active power loss of 1.3924 p.u. and a total reactive power loss of 4.4914 p.u. Considering the case of an integrated PV generator in node 6 (case 01) or node 7 (case 02), Table IV illustrates the voltage solution of each bus with and without PV generator integration. Note that the PV generator power was set to 3.9 p.u.

TABLE V. LOAD FLOW SOLUTION WITH AND WITHOUT PV GENERATOR

Bus N°	Voltage (p.u.)		
	Without	Case 01	Case 02
1	1.15	1.15	1.15
2	0.9769	1.0584	1.0172
3	0.9255	0.9491	0.9087
4	0.8986	0.9877	0.8016
5	0.8963	1.0513	0.7598
6	0.8748	<b>1.1635</b>	0.7618
7	0.8781	0.8876	<b>1.2134</b>

Evidently from Table VI, the PV generator integration ensured a voltage increase in the node where the PV generator was integrated. In conclusion, after integrating a PV generator, the voltage modules were improved from different nodes. The most important voltages are those at which the PV generator is connected.

The branch currents for the two cases and the case without a PV generator are listed in Table VII. It is evident that the

injection of the PV generator increased the current magnitude of the branch where the PV generator was integrated. Moreover, the branch currents decreased, except for the nodes of the branches connected to the nodes where the PV generator was integrated. Figure 4 shows the active and reactive power losses for the case without a PV generator and for cases 01 and 02. Comparing these two cases with that without the PV generator, it is seen that the active and reactive power losses decreased after PV generator integration. Indeed, the power losses of case 01 were lower than those of case 02, indicating that the best integration node was node 6. In addition, the voltage plan of the case 01 was more reliable than that of case 02. The results were better for the most distant PV locations.

TABLE VI. BRANCH CURRENT MAGNITUDE SOLUTION OF THE PV-RADIAL NETWORK

N°	Branch current magnitude (p.u.)		
	Without PV generator	Case 01	Case 02
1	5.7284	2.4254	5.1432
2	5.3283	2.3567	5.1174
3	3.4965	2.3854	2.5879
4	2.7432	2.4689	1.9892
5	<b>1.3240</b>	<b>2.9720</b>	0.9786
6	<b>1.3280</b>	0.7605	<b>4.9831</b>

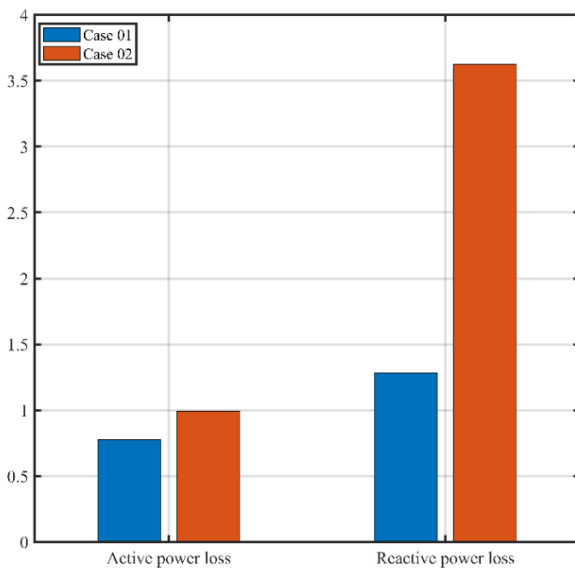


Fig. 4. Active and reactive power loss with and without PV generator.

TABLE VII. POWER LOSS SOLUTION OF THE PV-RADIAL NETWORK

Actual PV location nodes	Active power loss (p.u.)	Reactive power loss (p.u.)
02	1.3418	4.1688
03	0.9390	2.9652
04	3.3720	9.6416
05	0.9827	2.6175
06	0.7760	1.2828
07	0.9920	3.6237

To detect PV generator localization, an ANN control technique has been proposed [18, 19]. Table VIII lists the active and reactive power losses of the radial network system with respect to the node where the PV generator was injected. Figure 5 shows the ANN controller modeling design based on the data reported in Table VIII.

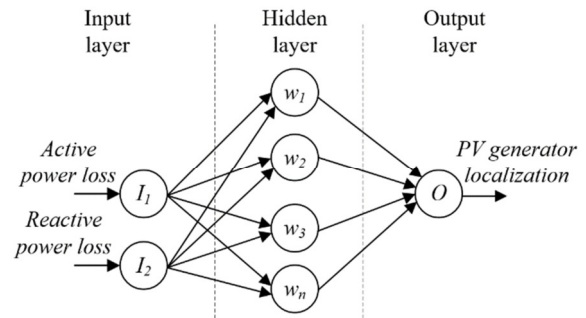


Fig. 5. ANN controller modeling design.

The MATLAB/Simulink environment was used to create the ANN model. Furthermore, the employed feedforward NN model had a hidden layer of 10 neurons, a hidden layer of 2 neurons, and an output layer of 1 neuron. A collection of localized PV generator data points based on simulations were trained using the Levenberg–Marquardt algorithm. The training performance curve demonstrates that the MSE error achieved at epoch 5 was 0.24085, as presented in Figure 6.

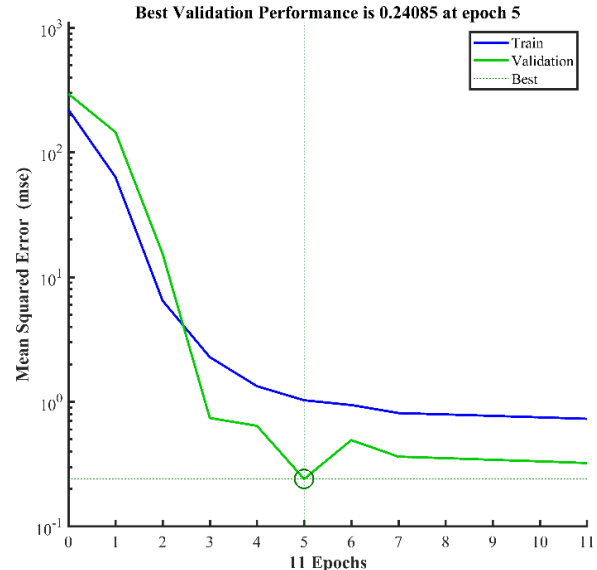


Fig. 6. ANN performance.

The best placement for the PV generator was determined by testing and verifying the 3 random values of the active and reactive power losses. The results of the identification procedure, which improves the efficacy and performance of the ANN controller, are displayed in Table IX.



TABLE VIII. ANN LOCALIZATION SOLUTION OF THE PV-RADIAL NETWORK

Active power loss (p.u.)	Reactive power loss (p.u.)	ANN localization
0.951	2.971	<b>04</b>
0.933	2.633	<b>05</b>
0.997	3.65	<b>07</b>

The proposed algorithm has also been implemented on the IEEE 33-bus radial distribution system with and without a PV Generator (PVG) as presented in Figure 7, to demonstrate its effectiveness. Thirty-three buses and 32 lines make up the system [20].

TABLE IX. LOAD FLOW SOLUTION OF THE IEEE 33-BUS RADIAL NETWORK WITHOUT PVG

Bus N°	Voltage (p.u.)	Angle (rad.)
1	1.0000000	0.0000000
2	0.9970252	0.0002549
3	0.9828930	0.0016899
4	0.9753834	0.0028438
5	0.9679571	0.0040162
6	0.9494792	0.0023728
7	0.9459544	-0.0016709
8	0.9322985	-0.0043500
9	0.9259658	-0.0056504
10	0.9200917	-0.0067634
11	0.9192229	-0.0066342
12	0.9177081	-0.0064295
13	0.9115324	-0.0080554
14	0.9092423	-0.0094567
15	0.9078155	-0.0101279
16	0.9064335	-0.0105421
17	0.9043854	-0.0119185
18	0.9037720	-0.0120894
19	0.9964968	0.0000658
20	0.9929192	-0.0011032
21	0.9922147	-0.0014411
22	0.9915772	-0.0017962
23	0.9793071	0.0011495
24	0.9726356	-0.0003994
25	0.9693105	-0.0011622
26	0.9475495	0.0030617
27	0.9449852	0.0040422
28	0.9335434	0.0054904
29	0.9253236	0.0068506
30	0.9217655	0.0086887
31	0.9176035	0.0072149
32	0.9166879	0.0068126
33	0.9164042	0.0066776

The load flow solution of an IEEE 33-bus distribution system for the radial network without PVG is illustrated in Table IX. Considering the case of an integrated PVG in nodes 1, 2...33, Table X illustrates the active and reactive power loss of each case.

Table XI represents the power loss solution of branches 15-16, 16-17, and 17-18. The GPV at node 16 supplies a significant portion of the energy from the GPV source installed at node 16. Additionally, nodes 17 and 18 in branches 15-16,

16-17, and 17-18 use the most energy. As you move away from node 16, the current in the branches decreases, which in turn causes the branches' losses to decrease. Thus, node 16, which has the lowest active and reactive losses, is the ideal position for the integration of the GPV.

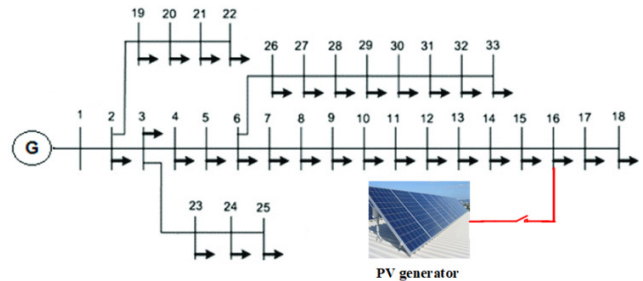


Fig. 7. Structure of the IEEE 33-bus radial distribution network with PVG.

TABLE X. POWER LOSS SOLUTION OF THE PVG- IEEE 33 BUS RADIAL NETWORK

Actual PVG location nodes	Active power loss (p.u.)	Reactive power loss (p.u.)
2	0.2093	0.1422
3	0.2014	0.1380
4	0.1973	0.1359
5	0.1932	0.1337
6	0.1843	0.1263
7	0.1832	0.1233
8	0.1762	0.1182
9	0.1731	0.1160
10	0.1703	0.1141
11	0.1699	0.1139
12	0.1692	0.1136
13	0.1667	0.1117
14	0.166	0.1109
15	0.1657	0.1106
<b>16</b>	<b>0.1655</b>	<b>0.1105</b>
17	0.1658	0.1108
18	0.1662	0.1111
19	0.2092	0.1420
20	0.20861	0.14151
21	0.20860	0.14150
22	0.2089	0.1419
23	0.1997	0.1369
24	0.1968	0.1346
25	0.1957	0.1338
26	0.1834	0.1258
27	0.1823	0.1252
28	0.1783	0.1218
29	0.1756	0.1195
30	0.1743	0.1188
31	0.1728	0.1175
32	0.1727	0.1173
33	0.1729	0.1177

Figure 8 illustrates the active and reactive power losses for the IEEE 33-bus network with PVG (01: injected in node 16) and without PVG (02). The reduction of active and reactive

losses in this study, when compared to previous works [21-21], is equal to 21.09% and 22.73%, respectively, as shown in Figure 8. This demonstrates the strong performance at this reduction level and even in the absence of shunt capacity, which lowers the cost of system installation. Additionally, compared to the other cases, the voltage plan for the case that integrated the PVG in node 16 was more reliable.

TABLE XI. BRANCH POWER LOSS SOLUTION OF THE PVG-IEEE 33 BUS RADIAL NETWORK

Branch	Without PVG		With PVG (in node 16)	
	Active power loss (p.u.)	Reactive power loss (p.u.)	Active power loss (p.u.)	Reactive power loss (p.u.)
15-16	0.287	0.210	<b>0.155</b>	<b>0.113</b>
16-17	0.257	0.343	<b>0.243</b>	<b>0.324</b>
17-18	0.054	0.043	<b>0.051</b>	<b>0.040</b>

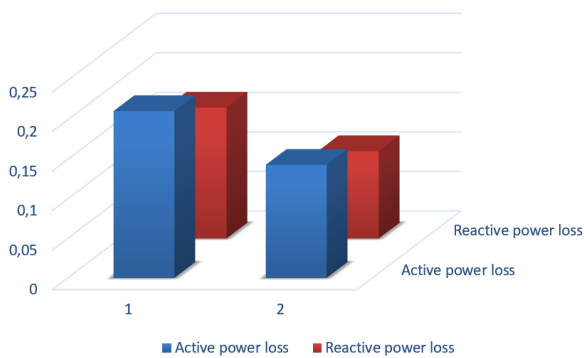


Fig. 8. Active and reactive power loss (1) without and (2) with PVG of the IEEE 33 bus network.

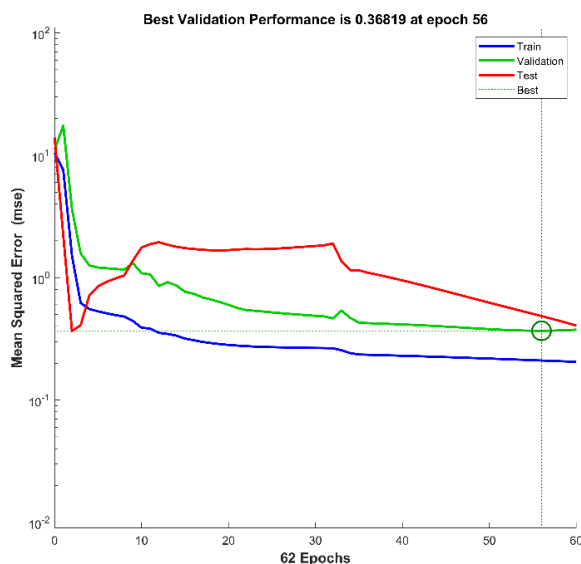


Fig. 9. IEEE 33-bus network ANN performance.

The Levenberg-Marquardt approach was used to train a set of localized PV generator data points based on simulations (Table XII). The representation of the training performance curve in Figure 9 shows that the obtained MSE at epoch 62 was

0.36819. The active and reactive power losses of 3 random variables were tested and verified to find the ideal location of the PV generator. The outcomes of the identification process, which enhances the ANN controller's effectiveness and performance, are shown in Table XIII.

TABLE XII. ANN LOCALIZATION SOLUTION OF THE PVG-IEEE 33-BUS RADIAL NETWORK

Active power loss (p.u.)	Reactive power loss (p.u.)	ANN localization
0.1168	0.1118	<b>13</b>
0.1996	0.1368	<b>23</b>
0.1759	0.1198	<b>30</b>

#### IV. CONCLUSION

In this study, an efficient technique for resolving the load flow radial distribution network problem was proposed. An ideal solution to the load flow issue was produced using two matrices generated from the topological properties of the distribution networks. The *BIBC* matrix was calculated based on Kirchhoff's current law. The *BCBV* matrix, which depicts the relationship between the bus voltages and branch currents, was used to resolve the load flow. These two matrices offer an easy methodology for handling load-flow problems when coupled. Consequently, the proposed method performed well in actual practice. Simulation results indicate that the proposed technique performs well for a widespread radial distribution network.

Owing to the developments in renewable energy sources, PV generators have been integrated into the radial system under study. To obtain the best location for a PV generator, the simulation results regarding the voltage amplitudes, current branches, and active and reactive power losses were presented and discussed in this paper. Based on these results, the best location to integrate a PV generator was determined as the farthest node, which provides reliable voltage and less power loss. In addition, to locate a PV generator that was already integrated into the radial network, an ANN controller was applied. The results of the identification procedure improved the efficacy and performance of the ANN controller.

#### ACKNOWLEDGMENT

The authors would like to thank Editage for English language editing.

#### REFERENCES

- [1] 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>.
- [2] M. A. Zdiri, A. S. Alshammari, A. A. Alzamil, M. B. Ammar, and H. H. Abdallah, "Optimal Shedding Against Voltage Collapse Based on Genetic Algorithm," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7695–7701, Oct. 2021, <https://doi.org/10.48084/etasr.4448>.
- [3] C. Seneviratne and C. Ozansoy, "Frequency response due to a large generator loss with the increasing penetration of wind/PV generation – A literature review," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 659–668, May 2016, <https://doi.org/10.1016/j.rser.2015.12.051>.

- [4] M. C. Herrera-Bríñez, O. D. Montoya, L. Alvarado-Barrios, and H. R. Chamorro, "The Equivalence between Successive Approximations and Matricial Load Flow Formulations," *Applied Sciences*, vol. 11, no. 7, Jan. 2021, Art. no. 2905, <https://doi.org/10.3390/app11072905>.
- [5] T. Shen, Y. Li, and J. Xiang, "A Graph-Based Power Flow Method for Balanced Distribution Systems," *Energies*, vol. 11, no. 3, Mar. 2018, Art. no. 511, <https://doi.org/10.3390/en11030511>.
- [6] O. D. Montoya and W. Gil-González, "On the numerical analysis based on successive approximations for power flow problems in AC distribution systems," *Electric Power Systems Research*, vol. 187, Art. no. 106454, Oct. 2020, <https://doi.org/10.1016/j.epsr.2020.106454>.
- [7] O. D. Montoya, W. Gil-González, and D. A. Giral, "On the Matricial Formulation of Iterative Sweep Power Flow for Radial and Meshed Distribution Networks with Guarantee of Convergence," *Applied Sciences*, vol. 10, no. 17, Art. no. 5802, Jan. 2020, <https://doi.org/10.3390/app10175802>.
- [8] O. D. Montoya, L. E. Rueda, W. Gil-González, A. Molina-Cabrera, H. R. Chamorro, and M. Soleimani, "On the Power Flow Solution in AC Distribution Networks Using the Laurent's Series Expansion," in *2021 IEEE Texas Power and Energy Conference (TPEC)*, Oct. 2021, pp. 1–5, <https://doi.org/10.1109/TPEC51183.2021.9384936>.
- [9] B. Sereeter, A. S. Markensteijn, M. E. Kootte, and C. Vuik, "A novel linearized power flow approach for transmission and distribution networks," *Journal of Computational and Applied Mathematics*, vol. 394, Art. no. 113572, Oct. 2021, <https://doi.org/10.1016/j.cam.2021.113572>.
- [10] K. Anoune, M. Bouya, A. Astito, and A. B. Abdellah, "Sizing methods and optimization techniques for PV-wind based hybrid renewable energy system: A review," *Renewable and Sustainable Energy Reviews*, vol. 93, pp. 652–673, Oct. 2018, <https://doi.org/10.1016/j.rser.2018.05.032>.
- [11] M. Zangiabadi, R. Feuillet, and H. Lesani, "An approach to deterministic and stochastic evaluation of the uncertainties in distributed generation systems," in *CIREN 2009 - 20th International Conference and Exhibition on Electricity Distribution - Part 1*, Prague, Czech Republic, Jun. 2009, pp. 1–4, <https://doi.org/10.1049/cp.2009.1098>.
- [12] D. R. H. Al-Rubayi and A. M. Alrawi, "Optimal Size and Location of Distributed Generators using Intelligent Techniques," *Engineering And Technology Journal*, vol. 28, no. 23, pp. 6623–6633, 2010.
- [13] M. Akorede, H. Hizam, I. Aris, and Z. Kadir, "A Review of Strategies for Optimal Placement of Distributed Generation in Power Distribution Systems," *Research Journal of Applied Sciences*, vol. 5, no. 2, pp. 137–145, Feb. 2010, <https://doi.org/10.3923/rjasci.2010.137.145>.
- [14] M. A. Mates, "A new power flow method for radial networks," in *2003 IEEE Bologna Power Tech Conference Proceedings*, Bologna, Italy, Jun. 2003, vol. 2, <https://doi.org/10.1109/PTC.2003.1304335>.
- [15] J.-H. Teng, "A direct approach for distribution system load flow solutions," *IEEE Transactions on Power Delivery*, vol. 18, no. 3, pp. 882–887, Jul. 2003, <https://doi.org/10.1109/TPWRD.2003.813818>.
- [16] S. Sivanagaraju, J. Viswanatha Rao, and M. Giridhar, "A loop based load flow method for weakly meshed distribution network," *ARPJ Journal of Engineering and Applied Sciences*, vol. 3, no. 4, pp. 55–59, Jan. 2008.
- [17] D. Rekioua and S. Aissou, "Photovoltaic Panels Characteristics Methods," in *Proceedings Engineering & Technology*, Jan. 2013, vol. 1, pp. 68–174.
- [18] M. A. Zdiri *et al.*, "Design and Analysis of Sliding-Mode Artificial Neural Network Control Strategy for Hybrid PV-Battery-Supercapacitor System," *Energies*, vol. 15, no. 11, p. 4099, Jan. 2022, <https://doi.org/10.3390/en15114099>.
- [19] J. Chakravorty, S. Shah, and H. N. Nagraja, "ANN and ANFIS for Short Term Load Forecasting," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2818–2820, Apr. 2018, <https://doi.org/10.48084/etasr.1968>.
- [20] V. Vita, "Development of a Decision-Making Algorithm for the Optimum Size and Placement of Distributed Generation Units in Distribution Networks," *Energies*, vol. 10, no. 9, Sep. 2017, Art. no. 1433, <https://doi.org/10.3390/en10091433>.
- [21] M. Dixit, P. Kundu, and H. R. Jariwala, "Incorporation of distributed generation and shunt capacitor in radial distribution system for techno-economic benefits," *Engineering Science and Technology, an International Journal*, vol. 20, no. 2, pp. 482–493, Apr. 2017, <https://doi.org/10.1016/j.jestch.2017.01.003>.

# An Adaptive Sliding Mode Controller for a PAM-based Actuator

**Quy-Thinh Dao**

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam  
thinh.daoquy@hust.edu.vn

**Vuong Van Dinh**

Hanoi University of Science and Technology, Vietnam | Hanoi Vocational College of High Technology, Vietnam  
vuongdv@hht.edu.vn

**Chien Tuan Vu**

Hanoi University of Science and Technology, Vietnam  
chien.vt181355@sis.hust.edu.vn

**Thuyet Quang Pham**

Hanoi University of Science and Technology, Vietnam  
thuyet.pq181778@sis.hust.edu.vn

**Duc Minh Duong**

Hanoi University of Science and Technology, Vietnam  
duc.duongminh@hust.edu.vn  
(corresponding author)

*Received: 5 December 2022 | Revised: 20 December 2022 | Accepted: 23 December 2022*

## ABSTRACT

The Pneumatic Artificial Muscle (PAM) is a promising actuator for developing the human-robot interaction system. However, modeling and controlling PAM-based actuators are a significant difficulty due to the inherent uncertainty and hysteresis of PAM. Besides, the control approach of a PAM-based system also deals with unknown disturbances that always exist in any system. This study developed a sliding mode controller that employs an adaptive law to deal with issues and improve control performance. Furthermore, the stability of the proposed controller is proven based on the Lyapunov stability criterion. Finally, through a series of tests, the effectiveness of the proposed control approach is verified.

*Keywords-pneumatic artificial muscle; sliding mode control; adaptive law; antagonistic configuration*

## I. INTRODUCTION

Robotics industry has grown rapidly during the recent years, especially regarding robots that move in the same way as humans do. PAM is a potential actuator mechanism that responds to this tendency, because it can stretch in the same way that human muscles can. Besides its flexibility in movement, it possesses benefits such as high power-to-weight, ease of maintenance, inherent safety, and low cost [1-5]. The above characteristics have made PAM a special research object, with common applications including manipulators, robotic orthosis, etc. [6-13]. However, controlling PAM faces many challenges. The deviation in mathematical model

estimation and the sensitivity to external disturbances are obstacles that need to be overcome if the PAM control is to achieve optimal performance. Many algorithms have been proposed to solve the control problem for this actuator, such as the traditional PID controller and its enhanced variations [14-19]. Nevertheless, the above controllers show limitations for the hysteresis and nonlinear characteristics of PAM. Nonlinear controllers to address the limitations of PID controllers, such as the fuzzy controller [20, 21], the neural network based controller [22], and the nonlinear ADRC controller [23], have been employed in several recent studies. In addition, the sliding mode controller and its various versions are always the leading candidates for the tasks of controlling nonlinear objects [17,

24-27]. Nonetheless, PAM is a nonlinear object, and the parameter estimation for the model of the PAM-based system is approximate and if the system disturbance is very high, it definitely affects the control quality of the aforementioned methods. As a result, creating a disturbance observer will aid in overcoming PAM's restrictions. In this study, a disturbance observer based on the adaptive law is developed to not only estimate the system disturbance component accurately but also to help the controller adapt to sudden external influences affecting the PAM system. As a result, control performance is improved. The effectiveness of the employed control approach is verified by experiments under different conditions. In addition, the stability of the overall system is proven by Lyapunov theory.

## II. SYSTEM DESCRIPTION

The system structure consists of two artificial muscles in antagonistic setup as shown in Figure 1. Two PAMs are supplied from an air compressor through two proportional valves (SMC, ITV-2030-212S-X26). One muscle bundle contracts while the other relaxes when the artificial muscle system is inflated and deflated through two proportional valves, causing the pulley to revolve about its axis of rotation. A potentiometer (WDD35D8T) measures the rotational angle produced. For software, the proposed controller is programmed on an education controller type from National Instrument (myRIO-1900). The controller calculates the potentiometer's feedback angle value and sends the control output to the electric control valves. LabVIEW software is used to monitor the entire procedure.

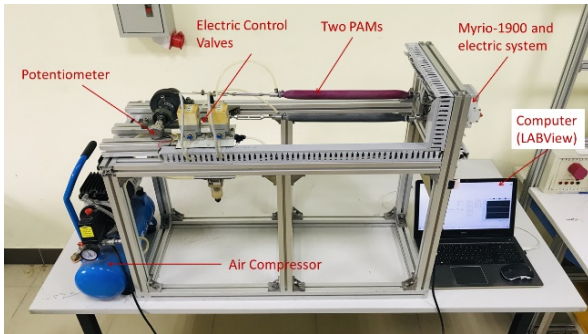


Fig. 1. The PAM-based experiment platform.

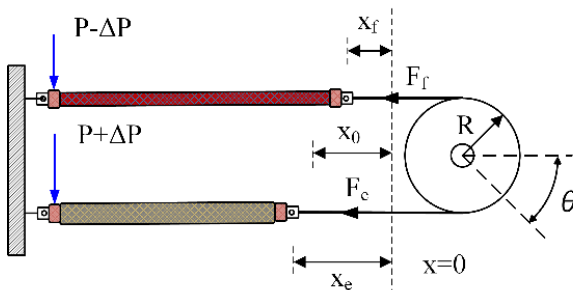


Fig. 2. The working principle of PAM-based antagonistic configuration.

Figure 2 depicts the operating principle of the PAMs. The two muscles are simultaneously initialized with the same

amount of gas at the same pressure  $P_0 = 0.2\text{MPa}$ , causing them to contract to the same position  $X_0$ . When the lower muscle bundle is injected with a pressurized amount of air  $\Delta P$  and at the same time the upper muscle is reduced by a corresponding amount of air, the lengths of the two muscle bundles change as  $X_0 - X_e$  and  $X_0 - X_f$  respectively. As a result, the pulley rotates at an angle of  $\theta$ . Therefore, controlling the change in air pressure inside the two muscles will result in a change in the pulley rotation angle. The internal pressure of both PAMs is obtained as follows:

$$\begin{cases} P_1 = P_0 + \Delta P \\ P_2 = P_0 - \Delta P \end{cases} \quad (1)$$

The control voltage of the proportional regulator valve PAM is determined using (1) as follows:

$$\begin{cases} u_1 = u_0 + u = k_0(P_0 + \Delta P) \\ u_2 = u_0 - u = k_0(P_0 - \Delta P) \end{cases} \quad (2)$$

where  $u_1$  and  $u_2$  are two voltages applied to two proportional regulator valves corresponding to the two PAMs,  $k$  is a constant indicating the voltage-pressure relationship on the valve, and  $u$  is the control variable that regulates the amount of pressure change  $\Delta P$  because the rotation angle of the pulley is changed by altering the control output  $u$  of the two proportional valves. As a result, the control output  $u$  would be the only input variable, with the observed joint angle  $\theta$  being the output. After introducing the disturbance, the discrete-time SISO model of the system can be expressed by:

$$y_k = -\sum_{i=0}^n a_i y_{k-i} + \sum_{j=0}^m b_j u_{k-j} + p_k \quad (3)$$

where  $u_k$  is the control output at instant  $k$ ,  $y_k$  is the pulley's angle,  $p_k$  represents the system disturbances. The model parameters are  $a_i$  and  $b_j$ . In [17], the authors chose a second-order discrete-time model (with  $m = n = 2$ ) and the least-squares method to determine the mathematical model for PAM. The coefficients of the model are shown in Table I.

TABLE I. MATHEMATICAL MODEL PARAMETERS

Parameter	Value
$a_1$	$-1.9567 \pm 0.0092$
$a_2$	$-0.9576 \pm 0.0128$
$b_1$	$0.0126 \pm 0.0013$
$b_2$	$0.0124 \pm 0.0049$

## III. CONTROLLER DESIGN

The design of the Adaptive Sliding Mode Controller (ASMC) is described in this section. Firstly, the SMC is designed based on the sliding surface theory. The system's uncertainty disturbance component is then estimated using an adaptive law-based noise observer. As a result, the control quality of the SMC is improved. Finally, the controller is verified for stability based on the Lyapunov stability condition. Figure 3 illustrates the block diagram of the overall system. The recommended SMC is described in the next subsection.

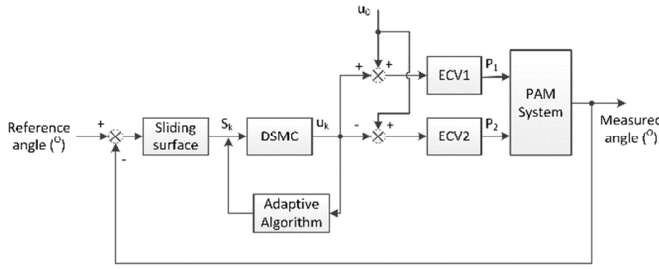


Fig. 3. Block diagram of the proposed control approach.

#### A. Sliding Mode Controller

Tracking deviation is determined by:

$$e_k = y_k^* - y_k \quad (4)$$

where  $y_k^*$  is the reference signal.

Consider a sliding surface as follows:

$$s_k = e_k + \alpha e_{k-1} \quad (5)$$

where  $\alpha$  is the tuning parameter  $0 < \alpha < 1$ . Equation (5) is rewritten as:

$$s_k = y_k^* - y_k - \alpha e_{k-1} \quad (6)$$

Replacing  $y_k$  from (3) and (6) can be expressed as:

$$s_k = y_k^* + \sum_{i=0}^n a_i y_{k-i} - \sum_{j=0}^m b_j u_{k-j} - p_k + \alpha e_{k-1} \quad (7)$$

To ensure the sliding variable is driven on the sliding surface, the following reaching law is taken into consideration.

$$\Delta s_k = s_k - s_{k-1} = -K_{SW} \text{sign}(s_k) \quad (8)$$

or:

$$s_k = s_{k-1} - K_{SW} \text{sign}(s_k) \quad (9)$$

where  $K_{SW} > 0$  is the control gain. By replacing  $s_k$  from (9) into (7), the control signal  $u_k$  becomes:

$$u_k = \frac{1}{b_1} \left[ y_k^* + \sum_{i=1}^n a_i y_{k-i} - \sum_{j=0}^m b_j u_{k-j} - p_k + \alpha e_{k-1} - s_{k-1} + K_{SW} \text{sign}(s_k) \right] \quad (10)$$

Component  $p_k$  in  $u_k$  is a variable of the unknown system disturbance. As a result, we propose in this paper an observer that uses adaptive law to estimate that component, and thereby improving the efficiency of the SMC. Finally, with the estimated value  $\hat{p}_k$  of  $p_k$ , the control signal  $u_k$  is calculated as:

$$u_k = \frac{1}{b_1} \left[ y_k^* + \sum_{i=1}^n a_i y_{k-i} - \sum_{j=0}^m b_j u_{k-j} - \hat{p}_k + \alpha e_{k-1} - s_{k-1} + K_{SW} \text{sign}(s_k) \right] \quad (11)$$

The architecture of an adaptive law-based disturbance observer is examined in greater depth below.

#### B. Adaptive Algorithm

As previously stated, the unknown disturbance  $p_k$  has a significant impact on the control signal  $u_k$ . The proposed adaptive algorithm for estimating the  $p_k$  component improves the control signal's accuracy. The following is a presentation of the adaptive law:

$$\hat{p}_k = \hat{p}_{k-1} - \varphi s_k \quad (12)$$

where  $\varphi$  is a strictly positive constant related to the adaptation rate. It is clear that:

$$\Delta \hat{p}_k = \hat{p}_k - \hat{p}_{k-1} - \varphi s_k \quad (13)$$

Equation (13) also shows that there is no adaptation when states are on the sliding surface:

$$\Delta \hat{p}_k = 0 \quad \text{for } s_k = 0 \quad (14)$$

#### C. Stable Analysis of Adaptive Sliding Mode Controller

This subsection presents the stability analysis of the proposed controller using the Lyapunov stability condition. It allows the ASMC's parameter to be determined. Consider the error value of the actual system disturbance component and the estimated component of the observer.

$$\tilde{p}_k = p_k - \hat{p}_k \quad (15)$$

It is obvious that:

$$\Delta \tilde{p}_k = \Delta p_k - \Delta \hat{p}_k \quad (16)$$

When the sliding surface variable  $s$  is 0, (14) indicates that there won't be any adaptation, resulting in the following calculation of  $\Delta \tilde{p}_k$ :

$$\Delta \tilde{p}_k = -\Delta \hat{p}_k = \varphi s_k \quad (17)$$

The following Lyapunov function is used as the base for the stability proof of the proposed controller:

$$V_k = \frac{1}{2} s_k^2 + \frac{1}{2\varphi} \tilde{p}_k^2 \quad (18)$$

Hence, we have:

$$\Delta V_k = V_k - V_{k-1} = \frac{1}{2} s_k^2 - \frac{1}{2} s_{k-1}^2 + \frac{1}{2\varphi} \tilde{p}_k^2 - \frac{1}{2\varphi} \tilde{p}_{k-1}^2 \quad (19)$$

Consider the following expression:

$$\Delta_s = \frac{1}{2} s_k^2 - \frac{1}{2} s_{k-1}^2 = \Delta s_k s_k - \frac{1}{2} (\Delta s_k)^2 \quad (20)$$

On the other hand,  $\Delta s_k$  is represented as:

$$\Delta s_k = s_k - s_{k-1} = y_k^* + \sum_{i=1}^n a_i y_{k-i} - \sum_{j=0}^m b_j u_{k-j} - p_k + \alpha e_{k-1} - s_{k-1} \quad (21)$$

Substituting  $u_k$  from (11) into (21) we get:



$$\Delta s_k = -\tilde{p}_k - K_{sw} \text{sign}(s_k) \quad (22)$$

Then, (20) becomes:

$$\begin{aligned} \Delta_s &= [-\tilde{p}_k - K_{sw} \text{sign}(s_k)] s_k - \frac{1}{2} (\Delta s_k)^2 \\ &= -\tilde{p}_k s_k - K_{sw} \text{sign}(s_k) s_k - \frac{1}{2} (\Delta s_k)^2 \end{aligned} \quad (23)$$

Next, we consider the following expression:

$$\Delta_p = \frac{1}{2\varphi} \tilde{p}_k^2 - \frac{1}{2\varphi} \tilde{p}_{k-1}^2 = \frac{1}{\varphi} \Delta \tilde{p}_k \cdot \tilde{p}_k - \frac{1}{2\varphi} (\Delta \tilde{p}_k)^2 \quad (24)$$

Replacing  $\Delta \tilde{p}_k = \varphi s_k$  into (24), we obtain:

$$\Delta_p = \frac{1}{\varphi} \Delta \tilde{p}_k \cdot \tilde{p}_k - \frac{1}{2\varphi} (\Delta \tilde{p}_k)^2 = \tilde{p}_k s_k - \frac{1}{2\varphi} (\Delta \tilde{p}_k)^2 \quad (25)$$

From (19), (23), and (25) we have:

$$\begin{aligned} \Delta V_k &= \Delta_s + \Delta_p \\ &= -\tilde{p}_k s_k - K_{sw} \text{sign}(s_k) s_k - \frac{1}{2} (\Delta s_k)^2 + \tilde{p}_k s_k - \frac{1}{2\varphi} (\Delta \tilde{p}_k)^2 \\ &= -K_{sw} \text{sign}(s_k) s_k - \frac{1}{2} (\Delta s_k)^2 - \frac{1}{2\varphi} (\Delta \tilde{p}_k)^2 \leq 0 \end{aligned}$$

Thus, the system is asymptotically stable.

#### IV. EXPERIMENTAL RESULTS

A series of different circumstances is conducted to verify the effectiveness of the proposed ASMC controller. The experiment uses two types of input signals: pure-sine and mixed-sine trajectories, as follows:

$$f(t) = 20 \sin 2\pi f \quad (24)$$

$$f(t) = 20 \sin 2\pi f + 12.8 \sin \pi f \quad (25)$$

These types of signals change frequency, from 0.1Hz to 0.8Hz, during the experiments. Furthermore, the system is evaluated with two scenarios, with load and no-load, for each input signal at a certain frequency. With the sample time  $T_s$  of 5ms, the control algorithm will be built using the LabVIEW/MyRIO toolbox and then incorporated into the MyRIO-1900 controller. On the LabVIEW interface, the whole motion trajectory of the object will be measured and presented. The conventional SMC is used as a counterpart to demonstrate the proposed controller's usefulness. Table II shows the parameters of the ASMC and SMC after fine-tuning.

Below the system is tested to track pure-sine orbitals with frequency varying between 0.1 and 0.8Hz. The analytical evaluation will show the advantages of tracking the performance of ASMC than of SMC. A load of 5.0kg also is added to the system to evaluate the robustness of the system when there are external disturbances. In subsection B, the same test scenarios and evaluations as subsection A are repeated. However, the input signal will be mixed-sine orbitals to test the capabilities of ASMC compared to SMC.

TABLE II. ADSMC AND DSMC PARAMETERS

Parameters	$k_{sw}$	$\alpha$	$\varphi$
ASMC	0.5	0.1	0.04
SMC	0.75	0.1	

##### A. Experiment with Pure-Sinusoidal Reference Signal

The system is evaluated when tracking pure-sine wave signals with frequencies ranging from 0.1 to 0.8Hz. The system operates with both load and no-load situations for each input signal. For example, Figure 4 demonstrates the experimental results when the system tracks 0.1 and 0.5Hz desired signals without a load. It can be seen that, at zero load, the tracking performance of ASMC is better than SMC's, especially in the steady state. Particularly, the maximum deviation values of ASMC and SMC at 0.1Hz are about  $1.0^\circ$  and  $3.0^\circ$ , respectively. Meanwhile, the performance of two controllers at 0.5Hz are  $3.0^\circ$  and  $4.0^\circ$ , respectively. There is no significant difference between the two controllers at the transient state.

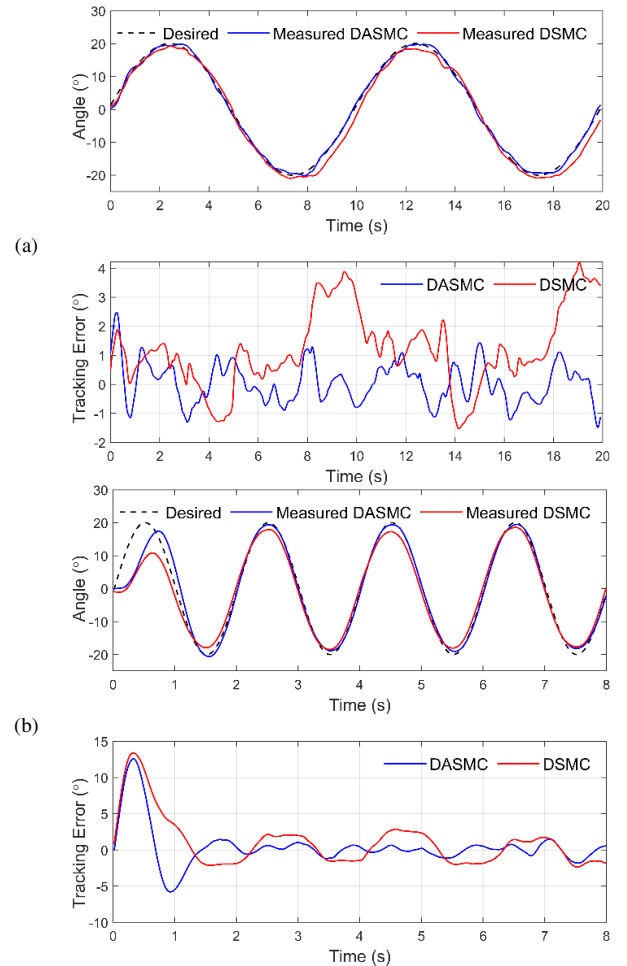


Fig. 4. Experiment results for tracking a pure-sinusoidal signal without a load: (a) 0.1Hz, (b) 0.5Hz.

With the system under loaded scenario in Figure 5, we observe that both controllers are affected. However, ASMC still outperforms SMC in terms of tracking performance. In

more detail, the highest deviation values for SMC and SMC at 0.1 Hz are 2.0 and 4.0°, respectively. The tracking performance of both controllers is lowered at 0.5Hz, with the highest errors of 4.5 and 5.0°, respectively. Table III shows the RMSE of the two controllers under the first scenarios experiment.

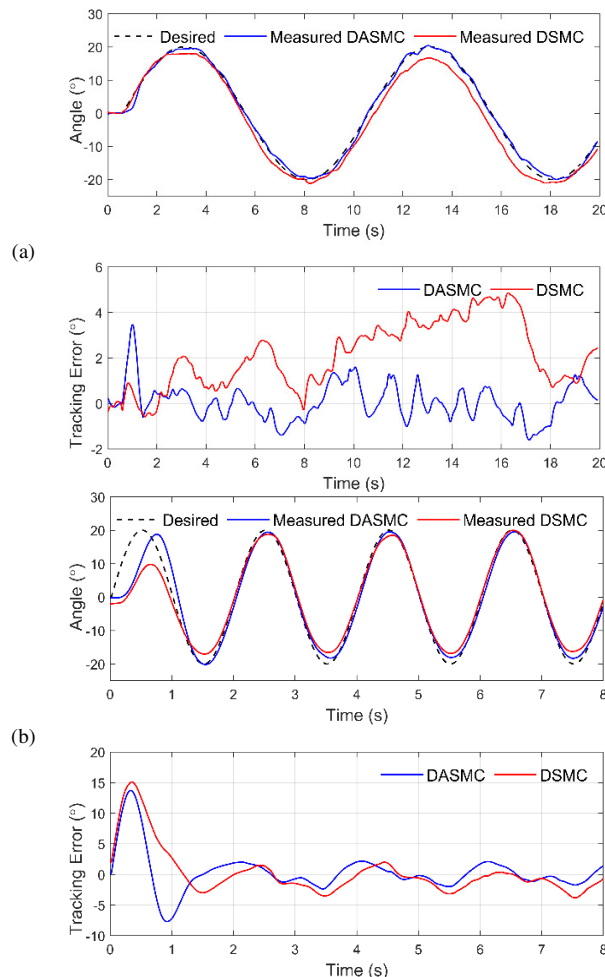


Fig. 5. Experiment results for tracking a pure sinusoidal trajectory with a load ( $m = 5.0\text{kg}$ ): (a) 0.1Hz, (b) 0.5Hz.

### B. Experiment with Mixed-Sine Reference Signal

Finally, ASMC and SMC are tested with mixed-sine orbital input. Figures 6 and 7 depict the tracking performance results of the two controllers when the system is tested under no-load and loaded conditions, respectively. In the case of no-load scenario, although the trajectory is more complicated, SAMC still performs well in terms of tracking, in contrast to SMC. It is observed that the response rate of SMC is slower than the change of the orbital reference signal. For example, at 0.25Hz, ASMC's highest deviation value is 3.0°, while SMC's is 6.0°. At 0.4Hz, the maximum error value of the two controllers is 4.0 and 7.0°, respectively. When the system had a load, the general trend of the two controllers is still to show reduced tracking performance, but ASMC still shows its ability in orbit tracking.

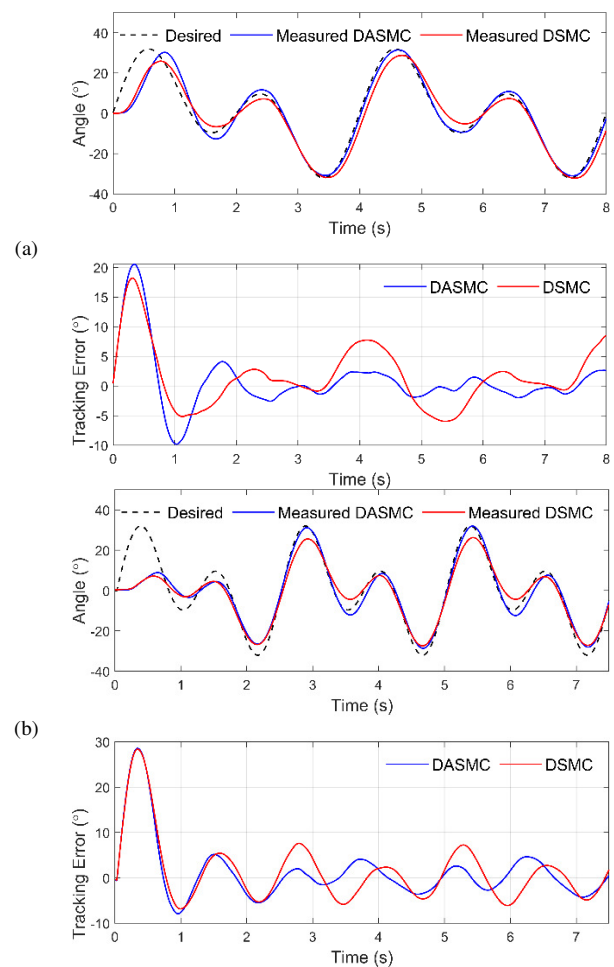


Fig. 6. The results of the experiments for tracking a mixed sinusoidal trajectory without a load:(a) 0.25Hz, (b) 0.4Hz.

TABLE III. RMSE OF THE TWO CONTROLLERS WITH SINUSOIDAL ORBITAL SIGNAL

	Frequency (Hz)	Without load		Load $m = 5\text{kg}$	
		ASMC	SMC	ASMC	SMC
Pure-sine signal	0.1	0.7	1.8	0.8	2.5
	0.5	2.8	3.5	3.3	3.9
	0.8	3.9	5.0	4.1	5.2
Mixed-sine signal	0.05	0.9	1.9	0.9	2.0
	0.25	5.0	5.9	5.5	6.4
	0.4	6.6	7.1	6.7	7.3

In both pure-sine and mixed-sine trajectories, with or without load, the ASMC always shows advantages over the conventional SMC. The proposed adaptive algorithm has observed the disturbance successfully. Thus, the disturbance is rejected, and the tracking performance is increased significantly. In the case of sudden load change, the driving trajectory of the PAM under the control of ASMC almost immediately returns to the steady-state as before the load changed. Moreover, in a load scenario with a pure sinusoidal input signal at low frequency (0.1Hz), the RMSE value of ASMC is only 1.0°, while that of traditional SMC is larger (2.0°). With mixed-sine trajectories, the results are the same. In

terms of 0.4Hz frequency trajectories, the ASMC has a RMSE value of  $6.0^\circ$ , which is lower than the RMSE value of the traditional SMC ( $7.0^\circ$ ). However, the ASMC still has the limitation that the transient time is larger than that of the traditional SMC. This is a common disadvantage of adaptive controllers because it takes a certain amount of time to update the value of the parameter it estimates. The ASMC still shows outstanding performance in the steady state in all experimental scenarios.

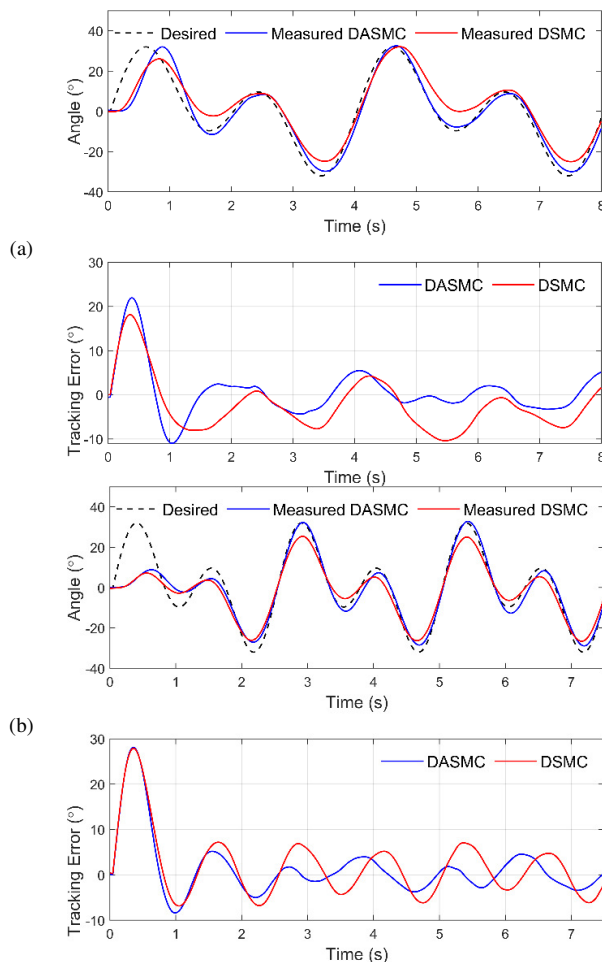


Fig. 7. Experiment results for tracking a mixed sinusoidal trajectory with a load ( $m = 5.0$  kg): (a) 0.25Hz, (b) 0.4Hz.

## V. CONCLUSION

This paper proposes a sliding mode controller with an adaptive disturbance observer for a PAM-based actuator. The proposed adaptive sliding mode control shows accuracy quality in orbit tracking, and good adaptability to external disturbances. Those qualities are all verified through different test scenarios with pure-sine and mixed-sine input signals at the frequency range from 0.05 to 0.8Hz. The adaptive algorithm helps to accurately estimate the system noise, thereby improving the control quality of conventional sliding mode controllers. Because of the disturbance observation process, the transient time of ASMC is larger than that of the traditional

SMC. The ASMC still shows outstanding performance in the steady state in all experimental scenarios. In following research, solutions to speed up the disturbance estimation will be studied in order to shorten the transient process for the proposed controller, thereby helping the controller to show better quality in the two aspects of transient and steady-state times for PAM.

## ACKNOWLEDGMENT

This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2022-PC-002.

## REFERENCES

- [1] F. Daerden and D. Lefeber, "Pneumatic artificial muscles: Actuators for robotics and automation," *European Journal of Mechanical and Environmental Engineering*, vol. 47, no. 1, pp. 11–21, 2002.
- [2] C.-P. Chou and B. Hannaford, "Measurement and modeling of McKibben pneumatic artificial muscles," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 1, pp. 90–102, Oct. 1996, <https://doi.org/10.1109/70.481753>.
- [3] M. Sekine, K. Shiota, K. Kita, A. Namiki, and W. Yu, "A lightweight shoulder prosthesis with antagonistic impact-absorbing hybrid actuation for bimanual activities of daily living," *Advances in Mechanical Engineering*, vol. 8, no. 4, pp. 1–17, Apr. 2016, <https://doi.org/10.1177/1687814016645982>.
- [4] K. K. Aun and H. P. H. Anh, "System Modeling Identification and Control of the Two-Link Pneumatic Artificial Muscle Manipulator Optimized with Genetic Algorithms," in *IEEE International Conference on Control and Automation*, Guangzhou, China, Jun. 2007, pp. 501–506, <https://doi.org/10.1109/ICCA.2007.4376407>.
- [5] C.-J. Chiang and Y.-C. Chen, "Neural network fuzzy sliding mode control of pneumatic muscle actuators," *Engineering Applications of Artificial Intelligence*, vol. 65, pp. 68–86, Oct. 2017, <https://doi.org/10.1016/j.engappai.2017.06.021>.
- [6] I. D. Walker, "Continuous Backbone 'Continuum' Robot Manipulators," *International Scholarly Research Notices*, vol. 2013, Jul. 2013, Art. no. e726506, <https://doi.org/10.5402/2013/726506>.
- [7] T.-Y. Choi and J.-J. Lee, "Control of Manipulator Using Pneumatic Muscles for Enhanced Safety," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 8, pp. 2815–2825, Dec. 2010, <https://doi.org/10.1109/TIE.2009.2036632>.
- [8] G. Kim, S. Kang, H. Cho, J. Ryu, M. Mun, and K. Kim, "Modeling and simulation of powered hip orthosis by pneumatic actuators," *International Journal of Control, Automation and Systems*, vol. 8, no. 1, pp. 59–66, Feb. 2010, <https://doi.org/10.1007/s12555-010-0108-9>.
- [9] G. S. Sawicki and D. P. Ferris, "A pneumatically powered knee-ankle-foot orthosis (KAFO) with myoelectric activation and inhibition," *Journal of NeuroEngineering and Rehabilitation*, vol. 6, no. 1, Jun. 2009, Art. no. 23, <https://doi.org/10.1186/1743-0003-6-23>.
- [10] S. Z. Ying, N. K. Al-Shammari, A. A. Faudzi, and Y. Sabzehmeidani, "Continuous Progressive Actuator Robot for Hand Rehabilitation," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5276–5280, Feb. 2020, <https://doi.org/10.48084/etasr.3212>.
- [11] P. Beyl *et al.*, "Safe and Compliant Guidance by a Powered Knee Exoskeleton for Robot-Assisted Rehabilitation of Gait," *Advanced Robotics*, vol. 25, no. 5, pp. 513–535, Jan. 2011, <https://doi.org/10.1163/016918611X558225>.
- [12] S. Hussain, S. Q. Xie, and P. K. Jamwal, "Control of a robotic orthosis for gait rehabilitation," *Robotics and Autonomous Systems*, vol. 61, no. 9, pp. 911–919, Sep. 2013, <https://doi.org/10.1016/j.robot.2013.01.007>.
- [13] D. Kim, Y.-P. Hong, and K.-S. Kim, "Bipedal Walking and Impact Reduction Algorithm for a Robot with Pneumatically Driven Knees," *International Journal of Control, Automation and Systems*, vol. 19, no. 12, pp. 3937–3946, Dec. 2021, <https://doi.org/10.1007/s12555-020-0613-4>.

- [14] G. Andrikopoulos, G. Nikolakopoulos, and S. Manesis, "Non-linear control of Pneumatic Artificial Muscles," in *21st Mediterranean Conference on Control and Automation*, Patanias, Greece, Jun. 2013, pp. 729–734, <https://doi.org/10.1109/MED.2013.6608804>.
- [15] K. K. Ahn and T. D. C. Thanh, "Nonlinear PID control to improve the control performance of the pneumatic artificial muscle manipulator using neural network," *Journal of Mechanical Science and Technology*, vol. 19, no. 1, pp. 106–115, Jan. 2005, <https://doi.org/10.1007/BF02916109>.
- [16] D. Zhang, X. Zhao, and J. Han, "Active Model-Based Control for Pneumatic Artificial Muscle," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 2, pp. 1686–1695, Oct. 2017, <https://doi.org/10.1109/TIE.2016.2606080>.
- [17] Q.-T. Dao, D.-H. Mai, D.-K. Nguyen, and N.-T. Ly, "Adaptive Parameter Integral Sliding Mode Control of Pneumatic Artificial Muscles in Antagonistic Configuration," *Journal of Control, Automation and Electrical Systems*, vol. 33, no. 4, pp. 1116–1124, Aug. 2022, <https://doi.org/10.1007/s40313-022-00902-5>.
- [18] T. Nuchkrua and T. Leephakpreeda, "Fuzzy Self-Tuning PID Control of Hydrogen-Driven Pneumatic Artificial Muscle Actuator," *Journal of Bionic Engineering*, vol. 10, no. 3, pp. 329–340, Sep. 2013, [https://doi.org/10.1016/S1672-6529\(13\)60228-0](https://doi.org/10.1016/S1672-6529(13)60228-0).
- [19] H. Wang and J. Lu, "Research on Fractional Order Fuzzy PID Control of the Pneumatic-hydraulic Upper Limb Rehabilitation Training System Based on PSO," *International Journal of Control, Automation and Systems*, vol. 20, no. 1, pp. 310–320, Jan. 2022, <https://doi.org/10.1007/s12555-020-0847-1>.
- [20] A. Rezoug, F. Hamerlain, and M. Hamerlain, "Application of Fuzzy Sliding Mode to control of Manipulator Robot actuated by Pneumatic artificial Muscles," *IFAC Proceedings Volumes*, vol. 42, no. 19, pp. 580–585, Jan. 2009, <https://doi.org/10.3182/20090921-3-TR-3005.00099>.
- [21] Z. Nadjat, B. Behih, Z. Bouchama, and K. Zehar, "Robust Adaptive Fuzzy Control of Nonlinear Systems," *Engineering, Technology & Applied Science Research*, vol. 12, pp. 8328–8334, Apr. 2022, <https://doi.org/10.48084/etasr.4781>.
- [22] S. Tian, G. Ding, D. Yan, L. Lin, and M. Shi, "Nonlinear Controlling of Artificial Muscle System with Neural Networks," in *IEEE International Conference on Robotics and Biomimetics*, Shenyang, China, Aug. 2004, pp. 56–59, <https://doi.org/10.1109/ROBIO.2004.1521751>.
- [23] K. Xing, J. Huang, Y. Wang, J. Wu, Q. Xu, and J. He, "Tracking control of pneumatic artificial muscle actuators based on sliding mode and non-linear disturbance observer," *IET Control Theory & Applications*, vol. 4, no. 10, pp. 2058–2070, Oct. 2010, <https://doi.org/10.1049/iet-cta.2009.0555>.
- [24] J. H. Lilly and P. M. Quesada, "A two-input sliding-mode controller for a planar arm actuated by four pneumatic muscle groups," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 3, pp. 349–359, Sep. 2004, <https://doi.org/10.1109/TNSRE.2004.831490>.
- [25] D. M. Duc, T. X. Tuy, and P. D. Phuoc, "A Study on the Response of the Rehabilitation Lower Device using Sliding Mode Controller," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7446–7451, Aug. 2021, <https://doi.org/10.48084/etasr.4312>.
- [26] W. M. Bessa and R. S. S. Barreto, "Adaptive fuzzy sliding mode control of uncertain nonlinear systems," *Sba: Controle & Automação Sociedade Brasileira de Automatica*, vol. 21, pp. 117–126, Apr. 2010, <https://doi.org/10.1590/S0103-17592010000200002>.
- [27] C. P. Vo, X. D. To, and K. K. Ahn, "A Novel Adaptive Gain Integral Terminal Sliding Mode Control Scheme of a Pneumatic Artificial Muscle System With Time-Delay Estimation," *IEEE Access*, vol. 7, pp. 141133–141143, 2019, <https://doi.org/10.1109/ACCESS.2019.2944197>.

# A Novel Algorithm for Optimal Harmonic Load Flow including Harmonic Compensation

**Prachi Godbole**

Department of Electrical Engineering, Fr. C. R. Institute of Technology, University of Mumbai, India  
prachigodbole@gmail.com  
(corresponding author)

**Sincy George**

St. Francis Institute of Technology, University of Mumbai, India  
principal@sfit.ac.in

*Received: 10 November 2022 | Revised: 18 December 2022 | Accepted: 21 December 2022*

## ABSTRACT

Load flow analysis is widely used for finding voltage at various parts of a large, interconnected power system. The exponential increase in the use of power electronic devices and the noticeable percentage of integration of renewable energy sources in modern power systems result in the occurrence of non-sinusoidal voltages at various parts of the power system. By reformation of the Newton-Raphson load flow method, a few harmonic load flow algorithms have been developed for modern transmission systems. However, for larger transmission systems, complexity and heavy computational burden are often encountered in the load flow solution process due to the increased size of the Jacobean matrix that further increases with multiple non-linear load buses and compensation requirements. This leads to convergence problems and increased execution time. In this paper, an optimal harmonic load flow algorithm is proposed, that employs a modified PSO technique to select control variables and the harmonic load flow method to find solutions for load flow analysis with a reduced size of the Jacobean matrix. The formulated method uses a simple technique to take into account nonlinear loads and results in faster convergence. The novelty of the algorithm is that an optimal load flow solution with the desired amount of reactive and harmonic compensation currents is obtained. The algorithm can help in getting load flow solutions in future power systems with distorted voltage and currents with compensation. The performance of the proposed algorithm is tested on a modified IEEE 30-bus system with multiple non-linear load buses and is validated with the Simulink model of the system.

**Keywords-***compensation; harmonic load flow; optimization; Particle Swarm Optimization (PSO); power system; transmission losses*

## I. INTRODUCTION

Load flow analysis plays an important role in power system operation, control, and planning. In modern power systems, the proliferation of power electronic-based nonlinear loads for industrial and commercial use results in the occurrence of non-sinusoidal bus voltages at various parts of the power system. As per [1], the Indian transmission system voltage and current were found to be rich in 5<sup>th</sup> and 7<sup>th</sup> harmonics due to large industrial loads such as rolling mills driven by variable speed drives and arc furnaces. Also, due to the integration of renewable energy sources in power grids [2], there is a noticeable decrease in voltage quality in modern transmission systems. The standard regulations and recommendations for harmonic studies [3] provide harmonic distortion limits for system voltage. As distributed generation and non-linear loads are increasing day by day, it is needed to find the desired reactive and harmonic compensation under non-sinusoidal voltage conditions while conducting load flow analysis in order

to maintain stable operation in future transmission systems. Newton-Raphson Load Flow (NRLF) is widely used for power flow analysis under sinusoidal conditions of a power system. Under non-sinusoidal voltage conditions and in order to find harmonic load flow solution, authors in [4] have modified the conventional NRLF method with detail nonlinear load modeling and satisfactory results were obtained with reasonable accuracy. However, it is implemented for an 8-bus system. Different harmonic load flow methods have been developed in [5-7]. Most use nonlinear load modeling and gave satisfactory results. Authors in [5] used a unified Newton's approach to simplify the solution, while authors in [6] proposed a decoupled approach suitable for transmission systems. Recently, the harmonic monitoring method [7] has been developed using PMUs in the transmission system. By understanding the importance of the Optimal Power Flow (OPF) solution for economic and reliable operation under linear load conditions, traditionally, classical or deterministic

optimization techniques have been used. Recently, comparatively easier to implement heuristic methods have been preferred to undertake global search [8-10].

The future transmission systems will be more and more at the risk of non-sinusoidal bus voltage conditions and under the adverse effects of harmonics. Under such conditions, with the optimal placement of capacitors and with strategic placement and sizing of passive or active filters for harmonic compensation using deterministic [11] and heuristic [12-14] optimization techniques, optimal load flow solutions can be obtained. However, in most of the reviewed studies, distribution systems have been considered for harmonic analysis and associated optimization using the decoupled harmonic load flow approach, as such systems are more sensitive to harmonic distortions. Looking at today's situation of modern transmission systems with noticeable percentage of harmonics, the utilities are redirecting their focus on transmission system harmonics and their mitigation. To the best of our knowledge, there is no sufficient solution available for harmonic load flow in transmission systems with compensation. In this paper, an optimal harmonic load flow algorithm is proposed for transmission systems with reduced size of Jacobean matrix which results in faster convergence. The desired amount of reactive and harmonic compensation is also addressed.

## II. THE PROPOSED OPTIMAL HARMONIC LOAD FLOW ALGORITHM

A simple and efficient optimal harmonic load flow algorithm is developed by modifying the NRLF method and employing the PSO technique that helps in getting faster a global optimal solution under non-sinusoidal voltage condition.

Consider a  $n$  bus system with 1 to  $m$  linear buses including the slack bus and  $(m+1)$  to  $n$  nonlinear load buses out of which 2 to  $l$  buses as PV buses and  $(l+1)$  to  $m$  buses as linear PQ buses, as shown in Figure 1. The specified variables and unknowns are given in Table I.

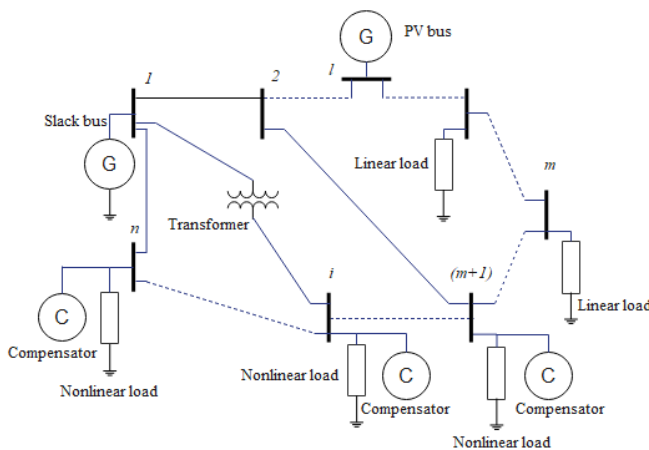


Fig. 1. Layout of a typical  $n$ -bus power system with compensators.

TABLE I. SPECIFIED AND UNKNOWN DATA OF VARIOUS BUSES

Bus No. (i)	Bus type	Specified variables	Unknown variables
1	Slack	$ V_{1,1} , \delta_{1,1}$	$P_{1,1}, Q_{1,1}, \delta_{1,h},  V_{1,h} $
2 to l	PV	$P_{i,1},  V_{i,1} $	$\delta_{i,1}, Q_{i,1}, \delta_{i,h},  V_{i,h} $
(l+1) to m	Linear PQ	$P_{i,1}, Q_{i,1}$	$\delta_{i,1},  V_{i,1} , \delta_{i,h},  V_{i,h} $
(m+1) to n	Non-linear PQ	$P_{i,1}, S_{i,t}$	$Q_{i,1}, \delta_{i,1},  V_{i,1} , \delta_{i,h},  V_{i,h} $

In Table I, the first subscript of a variable represents the bus number and the second represents the order of harmonics. Under non-sinusoidal condition, the NRLF equation [15] can be modified as:

$$[\Delta \bar{M}_{(mod)}] = [\bar{J}_{(mod)}] \cdot [\Delta \bar{U}_{(mod)}] \quad (1)$$

where  $\Delta$  represents change in the specified and calculated value of a parameter, while  $[\Delta \bar{M}_{(mod)}]$  is the modified mismatch vector and  $[\Delta \bar{U}_{(mod)}]$  is the modified correction vector. Equation (1) can be rewritten as:

$$\begin{bmatrix} \Delta \bar{W}_{i(mod)} \\ \Delta \bar{I}_{i,h} \end{bmatrix} = [\bar{J}_{mod}] \cdot \begin{bmatrix} \Delta \bar{\delta}_{i,1} \\ \Delta |\bar{V}_{i,1}| \\ \Delta \bar{\delta}_{i,h} \\ \Delta |\bar{V}_{i,h}| \end{bmatrix} \quad (2)$$

$[\Delta \bar{M}_{(mod)}]$  includes the power mismatch vector  $[\Delta \bar{W}_{i(mod)}]$  and the harmonic current mismatch vector  $[\Delta \bar{I}_{i,h}]$ , which are explained below.

### 1) Power Mismatch Vector

$[\Delta \bar{W}_{i(mod)}]$  is given by:

$$[\Delta \bar{W}_{i(mod)}] = \begin{bmatrix} \Delta P_{2,1}, \dots, \Delta P_{m,1}, \Delta Q_{2,1}, \dots, \Delta Q_{m,1}, \\ \Delta P_{m+1,t}, \dots, \Delta P_{n,t}, \Delta Q_{m+1,t}, \dots, \Delta Q_{n,t} \end{bmatrix}^T \quad (3)$$

where subscript  $t$  represents the total power including fundamental and harmonic powers calculated at the  $i^{th}$  bus as:

$$P_{i,t} = \sum_{h=1}^{h_{max}} (|V_{i,h}| \cdot \sum_{k=1}^n |V_{k,h}| \cdot |Y_{ik,h}| \cdot \cos(\theta_{ik,h} - \delta_{i,h} + \delta_{k,h})) \quad (4)$$

$$Q_{i,t} = \sum_{h=1}^{h_{max}} (|V_{i,h}| \cdot \sum_{k=1}^n |V_{k,h}| \cdot |Y_{ik,h}| \cdot \sin(\theta_{ik,h} - \delta_{i,h} + \delta_{k,h})) \quad (5)$$

where  $|Y_{ik,h}|$  is an element of bus admittance matrix  $[Y_{bus,h}]$  at the  $h^{th}$  harmonic component with magnitude  $|Y_{ik,h}|$  and phase angle  $\theta_{ik,h}$ . The size of  $[\Delta \bar{W}_{i(mod)}]$  is  $[2(n-1)] \times 1$ . However, if we consider 2 to  $l$  buses as PV buses, the corresponding  $\Delta \bar{Q}_{i,1} = 0$  reduces the size of the matrix.

### 2) Harmonic Current Mismatch Vector

$[\Delta \bar{I}_{i,h}]$  is formed for  $i = 1$  to  $n$  and  $h = 2$  to  $h_{max}$  given by:

$$[\Delta \bar{I}_{i,h}] = [\Delta I_{1,2}, \dots, \Delta I_{n,2}, \dots, \Delta I_{1,h_{max}}, \dots, \Delta I_{n,h_{max}}]^T \quad (6)$$

where  $\Delta I_{i,h} = I_{i,h(spec)} - I_{i,h(cal)}$  while  $spec$  and  $cal$  are representing the specified and calculated values of the parameters.  $I_{i,h(spec)}$  at the  $i^{th}$  bus is obtained from the



fundamental load current  $I_{i,1(spec)}$  obtained using the specified powers as follows:

$$I_{i,1(spec)} = \left( \frac{P_{i,1(spec)} + jQ_{i,1(spec)}}{V_{i,1}} \right)^* \quad (7)$$

where  $P_{i,1(spec)}$  and  $Q_{i,1(spec)}$  are the specified fundamental active and reactive powers at nonlinear load buses. Equation (7) calculates  $I_{i,1(spec)}$  for  $i = (m + 1)$  to  $n$  non-linear load buses.  $Q_{i,1(spec)}$  is iteratively calculated using the  $S_{i,t(spec)}$  equation that includes voltage and current distortion powers in addition to  $P_{i,1(spec)}$  and  $Q_{i,1(spec)}$ :

$$S_{i,t}^2 = P_{i,t}^2 + Q_{i,t}^2 + D_{i,t}^2 \quad (8)$$

where  $D_{i,t}^2 = (D_{Ii,t}^2 + D_{Vi,t}^2)$  with  $D_{Ii,t}^2$  and  $D_{Vi,t}^2$  are the squared values of total current and total voltage distortion powers respectively. The specified harmonic currents  $I_{i,h(spec)}$  for  $h = 2$  to  $h_{max}$  are obtained using a typical harmonic spectrum for non-linear load as follows [16]:

$$|I_{i,h(spec)}| = |I_{i,1(spec)}| \cdot \alpha_h \quad (9)$$

$$\theta_{i,h(spec)} = \theta_{h-spectr} + h \cdot (\theta_{i,1(spec)} - \theta_{1-spectr}) \quad (10)$$

where  $|I_{i,1(spec)}|$  and  $\theta_{i,1(spec)}$  are the magnitude and the phase angle for  $I_{i,1(spec)}$ ,  $\alpha_h$  is the ratio of the  $h$ -order harmonic component to the fundamental component of the current, and  $\theta_{h-spectr}$  is the phase angle for the  $h$ -order harmonic component.

$I_{i,h(cal)}$  for  $i = 1$  to  $n$  to form the mismatch current vectors is obtained as:

$$I_{i,h(cal)} = \sum_{k=1}^n |Y_{ik,h}| \cdot |V_{k,h}| \cdot (\cos(\theta_{ik,h} + \delta_{k,h}) + j \sin(\theta_{ik,h} + \delta_{k,h})) \quad (11)$$

The elements of  $[\Delta \bar{I}_{i,h}]$  are expressed in real and imaginary terms as:

$$[\Delta \bar{I}_{i,h}] = [\Delta \bar{I}_{ir,h}, \Delta \bar{I}_{ii,h}]^T \quad (12)$$

The harmonic current mismatch vector is updated with the desired harmonic compensation current,  $[I_{i,h(comp)}]$  as:

$$\Delta I_{ri,h(mod)} = (I_{ri,h(spec)} - I_{ri,h(comp)}) - I_{ri,h(cal)} \quad (13)$$

$$\Delta I_{ii,h(mod)} = (I_{ii,h(spec)} - I_{ii,h(comp)}) - I_{ii,h(cal)} \quad (14)$$

where:

$$I_{ri,h(comp)} = I_{ri,h(cal)} \cdot a_{i,h} \quad (15)$$

$$I_{ii,h(comp)} = I_{ii,h(cal)} \cdot a_{i,h} \quad (16)$$

where  $a_{i,h}$  is the fractional multiplier for the harmonic current to be compensated.

The desired reactive compensation current,  $I_{ii,1(comp)}$  is obtained as:

$$I_{ii,1(comp)} = I_{ii,1} \cdot a_{i,1} \quad (17)$$

where  $a_{i,1}$  is the fractional multiplier for the reactive current to be compensated.

$[\Delta \bar{U}_{(mod)}]$  includes fundamental voltage correction vectors,  $\Delta \bar{\delta}_{i,1}$  and  $\Delta |\bar{V}_{i,1}|$  for  $i = 2$  to  $n$  and harmonic voltage correction vectors,  $\Delta \bar{\delta}_{i,h}$  and  $\Delta |\bar{V}_{i,h}|$  for  $i = 1$  to  $n$  as given below:

$$[\Delta \bar{U}_{(mod)}] = \begin{bmatrix} \Delta \delta_{2,1}, \dots, \Delta \delta_{n,1}, \Delta |V_{2,1}|, \dots, \Delta |V_{n,1}|, \Delta \delta_{1,h}, \dots, \Delta \delta_{n,h}, \\ \Delta |V_{1,h}|, \dots, \Delta |V_{n,h}| \end{bmatrix}^T \quad (18)$$

Thus, the complete harmonic load flow equation with compensation can be written as:

$$\begin{bmatrix} \Delta \bar{W}_{i,1(mod)} \\ \Delta \bar{W}_{i,t} \\ \Delta \bar{I}_{ri,h(mod)} \\ \Delta \bar{I}_{ii,h(mod)} \end{bmatrix} = \begin{bmatrix} \bar{J}_{\bar{W}_{i,1}, \bar{\delta}_{i,1}} & \bar{J}_{\bar{W}_{i,1}, |\bar{V}_{i,1}|} & \bar{J}_{\bar{W}_{i,t}, \bar{\delta}_{i,h}} & \bar{J}_{\bar{W}_{i,t}, |\bar{V}_{i,h}|} \\ \bar{J}_{\bar{W}_{i,t}, \bar{\delta}_{i,1}} & \bar{J}_{\bar{W}_{i,t}, |\bar{V}_{i,1}|} & \bar{J}_{\bar{W}_{i,t}, \bar{\delta}_{i,h}} & \bar{J}_{\bar{W}_{i,t}, |\bar{V}_{i,h}|} \\ \bar{J}_{\bar{I}_{ri,h}, \bar{\delta}_{i,1}} & \bar{J}_{\bar{I}_{ri,h}, |\bar{V}_{i,1}|} & \bar{J}_{\bar{I}_{ri,h}, \bar{\delta}_{i,h}} & \bar{J}_{\bar{I}_{ri,h}, |\bar{V}_{i,h}|} \\ \bar{J}_{\bar{I}_{ii,h}, \bar{\delta}_{i,1}} & \bar{J}_{\bar{I}_{ii,h}, |\bar{V}_{i,1}|} & \bar{J}_{\bar{I}_{ii,h}, \bar{\delta}_{i,h}} & \bar{J}_{\bar{I}_{ii,h}, |\bar{V}_{i,h}|} \end{bmatrix} \times \begin{bmatrix} \Delta \bar{\delta}_{i,1} \\ \Delta |\bar{V}_{i,1}| \\ \Delta \bar{\delta}_{i,h} \\ \Delta |\bar{V}_{i,h}| \end{bmatrix} \quad (19)$$

For  $K$  order of harmonics under consideration, the size of  $[\Delta \bar{U}_{mod}]$  is  $[2(n-1) + 2Kn] \times 1$ . However, for  $i = 2$  to  $l$  as PV buses,  $\Delta |\bar{V}_{i,1}| = 0$ . In the proposed method, due to the use of the harmonic current spectrum, the nonlinear load parameters are not required to be included in the unknown vectors, resulting in reduction in the size of  $[\Delta \bar{U}_{mod}]$  by  $2(n-m)$ . Thus, the need of fundamental mismatch vector,  $[\Delta \bar{I}_{i,1}]$  in  $[\Delta \bar{M}_{(mod)}]$  has been eliminated and reduced by  $2(n-m)$ . This ultimately reduces the size of the modified Jacobean matrix  $[\bar{J}_{(mod)}]$ .  $[\bar{J}_{(mod)}]$  is formed by sub-matrices represented with the notation as a partial derivative of the  $[\Delta \bar{M}_{(mod)}]$  element given by the first subscript with respect to the  $[\Delta \bar{U}_{(mod)}]$  element. The size of  $[\bar{J}_{(mod)}]$  is  $[2(n-1) + 2Kn] \times [2(n-1) + 2Kn]$ . However, for the system with  $i = 2$  to  $l$  as PV buses, the size of  $[\bar{J}_{(mod)}]$  is  $[(2(n-1) - (l-1)) + 2Kn] \times [(2(n-1) - (l-1)) + 2Kn]$ . In the solution process,  $I_{i,1(spec)}$  and  $I_{i,h(spec)}$  are iteratively calculated using the updated bus voltages  $V_{i,1}$ , the specified powers  $P_{i,1(spec)}$ , and  $Q_{i,1(spec)}$  until the convergence set by tolerance  $\epsilon$ . This eliminates the decoupled effect to keep the solution accuracy unaffected.

In the proposed algorithm, the PSO technique is employed to find optimal harmonic load flow solution to achieve certain objectives keeping all inequalities within limits. Two types of inequalities, i.e. control variables and functional inequalities are considered. The optimization problem is formed as follows.

#### A. Selection of Control Variables and Formation of Penalty Function to Handle Functional Inequalities

The limits considered for inequalities are given in Table II. The control variables  $x_c$  are system parameters like generator voltage  $V_{g,i}$ , transformer tap ratio  $T_i$ , and current multipliers  $a_{i,h}$  and  $a_{i,1}$  for optimal compensation. The boundaries of the search space are set by control variable limits. In modified PSO, Michalewicz's non-uniform variable mutation operator,  $mp$  [17] is used to accelerate particles towards the

global best solution by adding variability into the population in each iteration. Rather than generating particles with the random variables, a small differential variable  $dx$  is defined to generate new non-repeated particle positions in each iteration for enhanced search:

$$dx = mp \cdot (x_{c(max)} - x_{c(min)}) \quad (20)$$

where  $x_{c(max)}$  and  $x_{c(min)}$  are the maximum and minimum limits of the control variables. All the selected control variables are applied to the system to get the optimal power flow solution.

The functional inequalities  $f_p$  considered are the generator reactive power  $Q_{gi}$ , the bus voltage magnitudes for PQ buses  $|V_i|$ ,  $\%V_{thdi}$ , and the total MVA capacity of lines  $S_{tli}$  as mentioned in Table II that are handled by selecting optimum static penalty factor  $\rho_{fp}$ . The corresponding penalty function,  $\Omega_{fp}$  for  $f_p$  inequality is defined for  $i = 1$  to  $N_{fp}$  parameters with maximum and minimum limits. The total penalty function,  $\Omega_{fp(total)}$  for  $N_p$  inequalities is:

$$\Omega_{fp(total)} = \sum_{t=1}^{N_p} \Omega_{f_{pt}} \quad (21)$$

TABLE II. LIMITS FOR CONTROL VARIABLES AND FUNCTIONAL INEQUALITIES

System component / bus	Numbers	Inequality limits
<b>Control variables <math>x_c</math></b>		
Generator bus	$i = 1$ to $n_g$	$ V_g _{(min)} \leq  V_{gi}  \leq  V_g _{(max)}$
Transformer	$i = 1$ to $n_T$	$T_{(min)} \leq T_i \leq T_{(max)}$
Compensator bus	$i = 1$ to $n_{comp}$	$a_{1(min)} \leq a_{i,1} \leq a_{1(max)}$ $a_{h(min)} \leq a_{i,h} \leq a_{h(max)}$
<b>Functional inequalities <math>f_{pi}</math></b>		
Generator bus	$i = 1$ to $n_g$	$Q_{g(min)} \leq Q_{gi} \leq Q_{g(max)}$
PQ bus	$i = 1$ to $n_{PQ}$	$ V _{(min)} \leq  V_i  \leq  V _{(max)}$ $\%V_{thdi} \leq \%V_{thd(max)}$
Transmission line	$i = 1$ to $n_{tl}$	$S_{tli} \leq S_{tl(max)}$

$n_g$ =no. of generators,  $n_T$ =no. of transformers,  $n_{comp}$ = no. of compensator buses,  $n_{PQ}$ = no. of PQ buses,  $n_{tl}$ =no. of transmission lines, All minimum and maximum limits are expressed with a subscript *min* and *max* respectively

### B. Formation of the Objective Function

As active power loss  $P_{loss}$  is of great concern in modern deregulated transmission systems with huge power flows over a long distance, the main objective of the proposed algorithm is considered as the minimization of  $P_{loss}$  under non-sinusoidal voltage conditions.

$$F = P_{loss} = \sum_{h=1}^{h_{max}} \left( \sum_{i=1, i \neq k}^n \left( \sum_{k=1, k \neq i}^n P_{ik,h} \right) \right) \quad (22)$$

where  $P_{ik,h}$  is the active harmonic power flow through transmission line between the  $i^{th}$  and the  $k^{th}$  bus.

The objective function is modified using the total penalty function,  $\Omega_{fp(total)}$  as:

$$F_t = F + \Omega_{fp(total)} \quad (23)$$

This helps in getting a global solution to the power flow problem which makes the system parameters within the specified margin. The optimal harmonic power flow solution is

obtained by selecting optimal control variables within the set PSO iterations  $Iter_{max}$ . The specialty of the algorithm is that voltage THD is allowed within the limit using harmonic compensation and voltage magnitude is controlled using fundamental reactive compensation with optimal multipliers for compensators. The flowchart for the optimal harmonic power flow is given in Figure2.

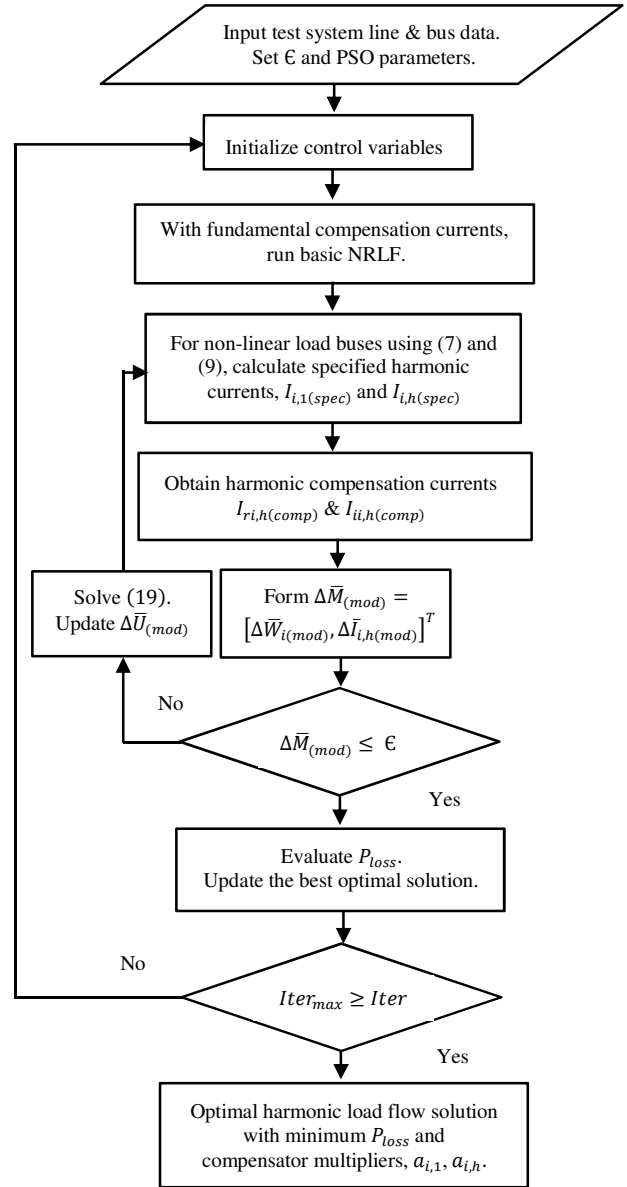


Fig. 2. Flowchart of the proposed optimal harmonic load flow algorithm.

### III. SIMULATION RESULTS

A generalized program for the proposed optimization algorithm was developed in the MATLAB m-file platform. To verify the performance of the proposed harmonic optimization algorithm, the 30-bus IEEE system [18] is modified by connecting non-linear loads in a few buses. A 6-pulse rectifier

is taken as the non-linear load connected at buses 7, 15, 21, and 30 with load demands and total apparent power demand  $S_{i,t}$  that includes distortion powers in addition to active and reactive powers as shown in Table III. The load at these buses is assumed such to get distorted voltage at various buses.

TABLE III. NON-LINEAR LOAD BUS DATA FOR THE MODIFIED IEEE 30-BUS SYSTEM

Bus no.	$P_d$ (MW)	$Q_d$ (MVar)	$S_{total}$ (MVA)
7	22.8	0.9	25
15	15	2.5	17
21	22	0.2	24
30	10.6	1.9	12

The initial assumptions for the unknown variables are taken as  $|V_{i,1}| = 1$  p.u.,  $|V_{i,h}| = 0.1$  p.u.,  $\delta_{i,1}$  and  $\delta_{i,h} = 0$ deg, and  $Q_{i,1} = 0$  p.u. for non-linear load bus and the harmonic load flow solution is obtained without optimal control. It is observed that the rms voltages  $|V_i|$  and  $\%V_{i,thd}$  are exceeding the limits specified by the statutory standards of 5% at some buses. It was also found that the total active power loss including harmonic loss  $P_{loss}$  is 6.83MW with fundamental active power loss,  $P_{loss,1} = 6.76$  MW obtained from the power flows. The proposed optimization algorithm is implemented with the control variables and functional inequality limits shown in Table IV.

TABLE IV. CONTROL AND FUNCTIONAL INEQUALITY LIMITS

Control variable set limits <sup>a</sup>							
	$T$	$ V_g $	$a_5$	$a_7$	$a_1$		
min. limit	0.95	0.95	0.5	0.5	0.8		
max. limit	1.05	1.1	1.0	1.0	1.0		
Functional variable set limits <sup>b</sup>							
	$Q_{g1}$	$Q_{g2}$	$Q_{g5}$	$Q_{g8}$	$Q_{g11}$	$Q_{g13}$	$ V $
min. limit	-29.8	-24	-30	-26.5	-7.5	-7.8	0.95
max. limit	59.6	48	60	53	15	15.5	1.05

<sup>a</sup> $|V_g|$  limits are specified in p.u. while  $T$  and multipliers  $a_5$ ,  $a_7$  and  $a_1$  are unitless. <sup>b</sup>All  $Q_g$  limits are specified in MVar while those for  $|V|$  are in p.u.

For rigid margins of functional inequalities like  $|V|$  and voltage Total Harmonic Distortion (THD), the high value of static penalty factor of 200 is selected. For the soft margins of  $Q_g$  and  $S_{tl}$ , the comparatively small static penalty factors of 5 and 10 are selected. The mutation ratios  $\mu$  and the acceleration coefficients  $c_1$ ,  $c_2$  are selected as 0.1, 2.05, and 2.05, respectively. The inertia coefficient  $\omega$  used to control the convergence behavior of PSO is taken as 0.5 and set to gradually decrease using a damping factor of 0.7296 in each iteration. The population  $n_p$  and the maximum number of PSO iterations  $Iter_{max}$  are set as 30 and 20. The optimal obtained values of the control variables are given in Table V. It is observed that the obtained optimal values of  $a_{i,1}$  and  $a_{i,5}$ ,  $a_{i,7}$  are such that the required compensation will be provided depending on the fundamental reactive currents and the 5<sup>th</sup> and 7<sup>th</sup> harmonic components of the current, respectively, at that bus. PSO with careful selection of parameters and modified with mutation operator exhibits better convergence characteristics. The optimal obtained harmonic power flow solution is given in Table VI.

TABLE V. OPTIMAL CONTROL VARIABLES OBTAINED BY THE MODIFIED PSO

System control variables				
$T_{6-9}$	$T_{6-10}$	$T_{4-12}$	$T_{27-28}$	$V_{g1}$
1.0011	1.05	0.9916	0.9798	1.0615
$V_{g2}$	$V_{g5}$	$V_{g8}$	$V_{g11}$	$V_{g13}$
1.036	0.9904	1.0166	1.0384	1.0004
Compensator control variables				
$a_{3,1}$	$a_{10,1}$	$a_{24,1}$	$a_{7,1}$	$a_{7,5}$
0.942	0.846	0.8148	0.8151	1.0
$a_{7,7}$	$a_{15,1}$	$a_{15,5}$	$a_{15,7}$	$a_{21,1}$
0.7633	0.9774	0.6428	0.5413	0.9209
$a_{21,5}$	$a_{21,7}$	$a_{30,1}$	$a_{30,5}$	$a_{30,7}$
0.6030	0.8830	0.9343	0.5	0.784

It is observed from Table VI that all bus voltage magnitudes are within the limits of 0.95p.u. and 1.05p.u., showing effective reactive power compensation with optimal values of multipliers  $a_{i,1}$  for the selected buses and system parameters. The voltage harmonic magnitudes at the buses are observed to be considerably reduced. As an effect, the values of  $\%V_{thd}$  at all buses are reduced below 5%. This demonstrates the effectiveness of harmonic compensation with optimal values of multipliers  $a_{i,5}$ ,  $a_{i,7}$ .

TABLE VI. OPTIMAL HARMONIC LOAD FLOW SOLUTION OBTAINED BY THE PROPOSED ALGORITHM

Bus $i$	$ V_{i,1} $ (p.u.)	$ V_{i,5} $ (p.u.)	$ V_{i,7} $ (p.u.)	$ V_i $ (p.u.)	$\%V_{i,thd}$
7	0.9999	0.0002	0.0015	0.9999	0.1550
10	0.9769	0.0009	0.0072	0.9769	0.7398
15	0.9692	0.0049	0.0109	0.9693	1.2375
21	0.9643	0.0071	0.0099	0.9644	1.2668
22	0.97	0.0025	0.0078	0.9700	0.8481
23	0.9638	0.0044	0.0093	0.9639	1.0665
24	0.96	0.0041	0.0088	0.9601	1.0082
25	0.9774	0.0049	0.0059	0.9774	0.7832
26	0.9599	0.0081	0.0038	0.9600	0.9286
27	0.9963	0.0003	0.0125	0.9964	1.2590
29	0.979	0.0067	0.0113	0.9791	1.3374
30	0.9695	0.0008	0.0445	0.9705	4.5859

To validate the above results obtained by the proposed optimization algorithm, a Simulink model was developed by the Simpower block set of MATLAB as shown in Figure 3. The complete system model consists of subsystems for group of buses, compensators, and nonlinear loads connected. The obtained results by the Simulink model and by the proposed algorithm are matching.

TABLE VII. RESULT COMPARISON BEFORE AND AFTER OPTIMIZATION

Parameter	Before compensation	After compensation
$P_{loss}$	6.83 MW	5.72 MW
$\%V_{thd}$	11.4935	4.5859
$ V_{30} $	0.8783	0.9705

From Table VII, it is observed that  $P_{loss}$  is considerably reduced to 5.72MW from 6.83MW with 5<sup>th</sup> and 7<sup>th</sup> harmonic active power losses as  $5.4037e^{-3}$  MW and  $7.1817e^{-3}$  MW, respectively. The voltage magnitude at the highest affected bus 30 has been improved from 0.8783p.u. to 0.9705p.u.

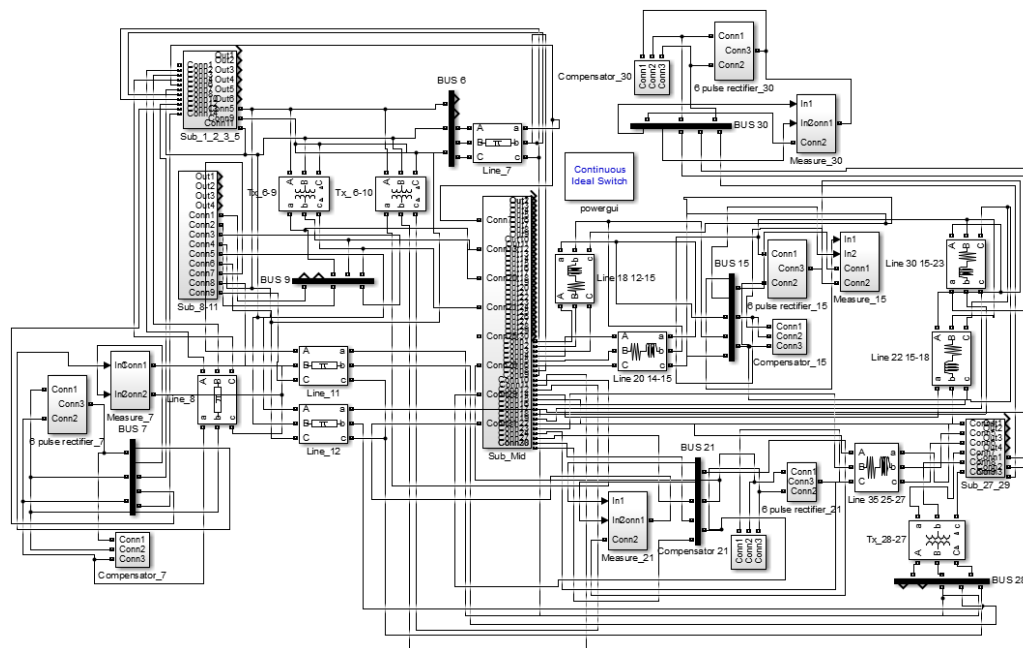


Fig. 3. Simulation model developed for IEEE 30 bus system with compensators.

#### IV. CONCLUSIONS

The increased use of power electronic-based devices in various applications along with the incorporation of renewable energy sources in modern power systems result in frequent occurrence of non-sinusoidal voltages at system buses. The proposed optimal harmonic load flow algorithm gives a global solution for the transmission systems under non-sinusoidal operation conditions with optimum amount of reactive and harmonic current compensation. The objective of the algorithm is to minimize the transmission losses while maintaining the fundamental voltage magnitude and  $\%V_{thd}$  within limits, while the size of the Jacobean matrix is reduced, resulting in faster convergence. The specialty of the proposed method is that with the addition of a greater number of nonlinear loads, the size of the Jacobean matrix remains the same, being independent of the number of nonlinear loads. Another advantage of the algorithm is that the multipliers  $a_{i,1}$ ,  $a_{i,5}$ , and  $a_{i,7}$  in each bus are controlled for the desired current compensation. By optimally compensating reactive and distortion powers, the transmission power loss can be reduced to 5.72MW from 6.83MW.

#### REFERENCES

- [1] *Report on Power Quality of Electricity Supply to the Consumers*. New Delhi, India: Central Electricity Regulatory Commission, 2018.
- [2] X. Liang, "Emerging Power Quality Challenges Due to Integration of Renewable Energy Sources," *IEEE Transactions on Industry Applications*, vol. 53, no. 2, pp. 855–866, Mar. 2017, <https://doi.org/10.1109/TIA.2016.2626253>.
- [3] "IEEE Recommended Practice for Conducting Harmonic Studies and Analysis of Industrial and Commercial Power Systems," *IEEE Std 3002.8-2018*, pp. 1–79, Jul. 2018, <https://doi.org/10.1109/IEEESTD.2018.8479357>.
- [4] D. Xia and G. T. Heydt, "Harmonic Power Flow Studies - Part II Implementation and Practical Application," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 6, pp. 1266–1270, Jun. 1982, <https://doi.org/10.1109/TPAS.1982.317172>.
- [5] C. N. Bathurst, B. C. Smith, N. R. Watson, and J. Arrillaga, "A modular approach to the solution of the three-phase harmonic power-flow," *IEEE Transactions on Power Delivery*, vol. 15, no. 3, pp. 984–989, Jul. 2000, <https://doi.org/10.1109/61.871363>.
- [6] W.-M. Lin, T.-S. Zhan, and M.-T. Tsay, "Multiple-frequency three-phase load flow for harmonic analysis," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 897–904, Feb. 2004, <https://doi.org/10.1109/TPWRS.2004.825906>.
- [7] B. Vedik, C. K. Shiva, and P. Harish, "Reverse harmonic load flow analysis using an evolutionary technique," *SN Applied Sciences*, vol. 2, no. 9, p. 1584, Aug. 2020, <https://doi.org/10.1007/s42452-020-03408-4>.
- [8] Z. Sahli, A. Hamouda, S. Sayah, D. Trentesaux, and A. Bekrar, "Efficient Hybrid Algorithm Solution for Optimal Reactive Power Flow Using the Sensitive Bus Approach," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8210–8216, Feb. 2022, <https://doi.org/10.48084/etasr.4680>.
- [9] M. Salman, S. Hongsheng, M. A. Aman, and Y. Khan, "Enhancing Voltage Profile and Power Loss Reduction Considering Distributed Generation (DG) Resources," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8864–8871, Aug. 2022, <https://doi.org/10.48084/etasr.5046>.
- [10] Y. A. Mobarak, A. M. Hemeida, A. El-Bahnasawy, and M. M. Hamada, "Reactive Power Compensation on Egypt Electricity Network for Optimal Power Saving," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3699–3704, Feb. 2019, <https://doi.org/10.48084/etasr.2451>.
- [11] Y.-Y. Hong, "Optimal harmonic power flow," *IEEE Transactions on Power Delivery*, vol. 12, no. 3, pp. 1267–1274, Jul. 1997, <https://doi.org/10.1109/61.637003>.
- [12] G. W. Chang, H.-L. Wang, and S.-Y. Chu, "Strategic placement and sizing of passive filters in a power system for controlling voltage distortion," *IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1204–1211, Jul. 2004, <https://doi.org/10.1109/TPWRD.2003.822954>.
- [13] M. Mohammadi, "Particle swarm optimization algorithm for simultaneous optimal placement and sizing of shunt active power conditioner (APC) and shunt capacitor in harmonic distorted distribution system," *Journal of Central South University*, vol. 24, no. 9, pp. 2035–2048, Sep. 2017, <https://doi.org/10.1007/s11771-017-3613-7>.

- [14] A. Ulinuha, M. A. S. Masoum, and S. Islam, "Hybrid genetic-fuzzy algorithm for volt/var/total harmonic distortion control of distribution systems with high penetration of non-linear loads," *IET Generation, Transmission & Distribution*, vol. 5, no. 4, pp. 425–439, Apr. 2011, <https://doi.org/10.1049/iet-gtd.2010.0168>.
- [15] W. Tinney and C. Hart, "Power Flow Solution by Newton's Method," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, no. 11, pp. 1449–1460, 1967.
- [16] "Modeling and simulation of the propagation of harmonics in electric power networks. I. Concepts, models, and simulation techniques," *IEEE Transactions on Power Delivery*, vol. 11, no. 1, pp. 452–465, Jan. 1996, <https://doi.org/10.1109/61.484130>.
- [17] P. S. Andrews, "An Investigation into Mutation Operators for Particle Swarm Optimization," in *2006 IEEE International Conference on Evolutionary Computation*, Jul. 2006, pp. 1044–1051, <https://doi.org/10.1109/CEC.2006.1688424>.
- [18] B. Zhao, C. X. Guo, and Y. J. Cao, "A multiagent-based particle swarm optimization approach for optimal reactive power dispatch," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1070–1078, Feb. 2005, <https://doi.org/10.1109/TPWRS.2005.846064>.

## AUTHORS PROFILE



**Prachi Godbole** received the BE degree in Electrical Engineering and the ME degree in Electrical Engineering (Power Systems) from the Walchand College of Engineering, Sangli, India in 1999 and 2002, respectively. She is currently pursuing the Ph.D. degree in Electrical Engineering at the University of Mumbai, India. Her research interests include power quality and renewable energy technology.



**Sincy George** received the BTech and MTech degrees in Electrical Engineering from the Government College of Engineering, Trivandrum-Kerala in 1985 and 1989 respectively and a Ph.D. degree in Electrical Engineering from the Indian Institute of Technology, Bombay. She worked as a Professor and the Head of Electrical Engineering Department at Fr. C. R. Institute of Technology, India for 20 years. Since 2016, she is the Principal at St. Francis Institute of Technology, India. Her research interests include power quality and power electronics. She is a senior member, IEEE, a fellow of the Institution of Engineers India (IEI), a Member of the Indian Society for Technical Education (ISTE), and Computer Society of India (CSI).

# Building an Application that reads Secure Information Stored on the Chip of the Citizen Identity Card in Vietnam

**Van-Hoan Le**

Weapon Institute, Vietnam  
hoanle.aiot@gmail.com

**Nhu-Quynh Luc**

Academy of Cryptography Techniques, Vietnam  
quynhln@actvn.edu.vn  
(corresponding author)

**Toan Thanh Dao**

University of Transport and Communications, Vietnam  
daotoan@utc.edu.vn

**Quang-Trung Do**

Academy of Cryptography Techniques, Vietnam  
trungdq1980@actvn.edu.vn

*Received: 29 November 2022 | Revised: 24 December 2022 | Accepted: 31 December 2022*

## ABSTRACT

Reading the information on the CIC/passport is very meaningful in serving the life activities of citizens in Vietnam and of foreign citizens visiting Vietnam. This research is based on the operating modes, such as BAC, FACE, and EAC to read the data contained in the chip put on the Citizen Identity Card (CIC) out securely. Specifically, the authors used the BAC mode to perform safe data reading from the CIC's chip. BAC mode uses 3DES and SHA1 algorithms to encrypt data to ensure security, so when the data are transmitted from the chip they are encrypted and decryption is performed by the application. In this paper, a complete application has been built for reading personal information stored securely on CIC. This application is built based on the BAC reading mode corresponding to CIC in Vietnam and meets the requirements of the ICAO 9303 standard, so it is able to read passports or CICs of other countries that use this standard. The performance of this application when reading data stored on fields DG1, DG2, and DG13 is improved with a speed of about 1.2s - 1.3s for all 3 data fields.

**Keywords-***BAC; EAC; PACE; 3DES; MRZ; SHA; digital signature*

## I. INTRODUCTION

Smart card technology was first proposed in 1977 [1]. During the '90s, there was an explosion in smart card technology in Europe with the appearance of SIMs (in GSM devices) and the international payment systems MasterCard, Visa, and Euro-pay [2-4]. In 1994, the EMV system (Euro-pay, MasterCard, Visa) was widely used in many countries [5, 6]. Currently, smart cards are applied in many fields [7]. Physically, the smart card is compact in size and embedded with an integrated circuit to control access to resources [4, 8]. Communication between the smart card and the card reader is direct physical contact (Contact Smartcard) or through short-

range wireless connection such as radio frequency identification (Contactless Smartcard) [8-10]. Operating systems designed for smart cards can provide personal identification, authentication, data storage, and application processing [9, 11, 12]. Smart cards using chips have very high reliability and safety for practical applications because some security algorithms and biometric features are installed [13-15]. Smart card technology with security is applied to digital ID cards (e-passport) as legal proof for different services in many countries [13, 16-18].

In Vietnam, the Government promulgated the Decree 577 on ID cards and regulations on issuance of ID cards in 1957, to



help citizens facilitate daily transactions [19]. After a while, this ID card showed some disadvantages: it can be faked, many people share the same number of ID cards, and the information on the ID card is lost after a period of use. Currently, with the trend of global digital transformation, the Government of Vietnam has launched a chip-based CIC to store and secure user information [20], in which, if the data retrieved in the chip are exactly the same as the data recorded on the card, the user information is valid [21]. Moreover, CIC with chips have many other advantages such as making international, online, and digital transactions [22].

CICs with chip in Vietnam are manufactured based on smart card technology in compliance with ICAO 9303 [22, 23]. In particular, the chip used in CIC has built-in cryptographic solutions to secure personal data, ensuring that information is safe and reliably authenticated. According to ICAO 9303, the access process to read security information stored on the CIC's chip operates in three modes: BAC (Basic Access Control), PACE (Password Authenticated Connection Establishment), and EAC (Extended Access Control) [17, 22-23]. The command architecture used is the APDU (Application Protocol Data Unit) protocol [7, 21, 24]. Data are stored and organized on the chip according to ICAO 9303 [22, 25], ISO/IEC 14443-4 [26], and ISO/IEC 7816-4:2013 [27] standards.

In Vietnam, a new chip-based CIC is going to be deployed in the community in 2022. There are almost no studies related to the reading of the information contained in the CIC/passport chip. Almost every citizen's transactions related to finance (such as withdrawing money from an ATM, buying and selling online, inter-bank financial transactions, etc.) and activities in life are related to the CIC/passport. Therefore, reading the information on the CIC/passport will make the transaction process between citizens, government agencies, and service providers much more convenient. In this article, an application that can read security information stored on the DG1, DG2, and DG13 partitions of the CIC's chip has been built and is discussed.

## II. RELATED WORK

### A. The Physical Surface of the CIC

Figure 1 shows the physical surface (front and back) design of the CIC. The information recorded on CIC includes full name, date of birth, expiration date, gender, hometown, portrait, and country. The back side has the information on the relevant characteristics of the citizen. The MRZ information area recorded on the back of the CIC is an area used to be read by a machine, facilitating inspection and reducing execution time in administrative procedures [25]. In addition, MRZ also provides the ability to verify information in the VIZ. MRZ is formatted according to ICAO 9303 standard to ensure machine readability in different countries. Therefore, OCR-B font is used to store data in MRZ [25].

The hardware design of the chip ID card is organized into two main parts: Power Reserve and Main Card. In particular, the Power Reserve helps CIC collect and reserve energy to ensure that the chip on CIC works when performing transactions related to the CIC. The Main Card part consists of the RFID chip, the CIC chip, antenna, and integrated circuit. The RFID

chip is in charge of receiving power and making the connection with the antenna. The data will be transmitted through the antenna and if encryption or decryption is needed, the RFID chip will make the connection and perform encryption and decryption with the CIC chip. The CIC chip performs the main job of executing cryptographic algorithms and communicating with the RFID chip to transmit information securely and store encrypted personal information. The CIC chip is connected to the power source to ensure the energy used to store the password key and ID of the card when the power is not directly supplied.



Fig. 1. (a) Front side and (b) back side (b) of the CIC.

### B. File Organization on the Operating System in the Chip of the CIC

According to the ICAO standard [19, 20], the data in the chip of CIC are in a hierarchical form. First, there is the root directory (Master File -MF) and then the dedicated directory (Dedicated File - DF), and finally the Elementary Files (EF). The root directory of the CIC is divided into 2 DFs and 3 EFs, in which Future Application is a dedicated directory for future development, EF.CardAccess is the elementary file used for PACE mode access, EF.CardSecurity is the elementary file used for PACE mode access. EMRTD is a dedicated directory that stores the basic data files of each citizen. This is the most important folder of the CIC because all user information will be located in there. The EMRTD stores basic files, also known as data partitions and is described according to the international standards ICAO 9303 for CIC/e-passport. Citizens' data are stored in basic files in the EMRTD Application directory including EF.COM, EF.SOD, and DG1-16. EF.COM stores Unicode version information and a list of data groups available in the CIC, DG1 stores individual basic details (MRZ), DG2 stores citizens' facial images, DG3 stores the fingerprint, DG4 stores an iris picture, DG5 stores features of the face, DG6 is expanded to store information that will be added later, DG7 stores citizen's signature, DG8, DG9, DG10, DG11 are currently not used, DG12 stores supporting information, DG13, DG16 store citizen-related information, DG14 is used for the EAC access mechanism, DG15 stores the public key, and EF.SOD stores the hash value of each field.

### C. Communication between CIC and the Computer

The computer is connected to the card reader via a serial port, USB, or Bluetooth to read the card data through either contact or contactless communication. After the CIC is inserted into the slot of the card reader, the computer connected to the card reader will detect the device through the ATR message sent from the CIC. After receiving the card, the computer and the card establish a connection and select the protocol for data transmission (TAPDU). When the data transfer protocol is

selected, the process of sending and receiving data between the computer and the card is done through APDU commands. When the computer sends the APDU command to the card, the card will receive the command and send a response back to the computer. Finally, after the successful data exchange, the CIC is removed from the slot. The computer will detect the discarded device through the card reader.

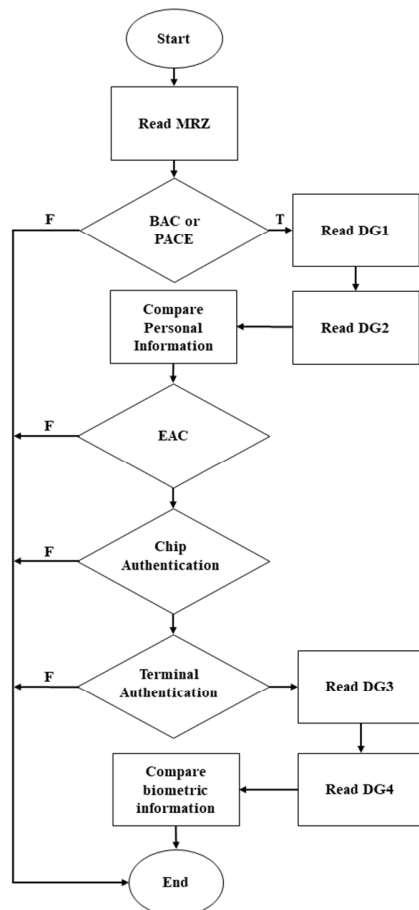


Fig. 2. Reading CIC data with BAC, PACE, and EAC modes.

#### D. Data Reading Modes

The read access process of the CIC chip includes three secure access modes, each of them determining the data partitions allowed to be accessed. The CIC access modes are BAC, PACE, and EAC [22, 24, 25]. The BAC mode ensures that only authorized parties have access to read the data. The PACE mode implements password-based access and Diffie-Hellman key agreement. The password used for PACE depends on the digital certificate of the CIC, using AES or 3DES algorithms to encrypt and authenticate the data [28, 29]. This mode allows reading, similar to BAC mode, the user's information data including name, date of birth, card number, portrait, relative information, and other personal characteristics. EAC mode is an access mode that combines advanced security features between the terminal and the CIC to protect and restrict access to sensitive personal data contained on the chip. It enables mutual authentication and establishes a secure

communication channel between the CIC and a terminal. In contrast to ordinary personal data that can be protected by basic mechanisms, more sensitive data (such as fingerprints or iris images) must have enhanced protection to prevent unauthorized access. To access this mode, users need to establish a secure connection based on either BAC or PACE mode. EAC access mode is implemented through two parts: Chip Authentication that uses the Diffie-Hellman key exchange algorithm to exchange new session keys and Terminal Authentication that uses RSA digital signature algorithm or digital signature on an elliptic curve. Figure 2 describes in detail the process of reading data when combining the three modes. To retrieve the data in CIC, we first need to get the MRZ information contained on the CIC that can be entered or scanned. Then, we proceed to access with two modes, BAC or PACE. When either of these two modes is successfully set up, the individual can access the partitions DG1 and DG2 to get information, portraits and some information in the related data partition (if permitted). That information is then compared with the personal information contained on the CIC to check whether the user information is valid. After successful data retrieval, EAC advanced access mode can be selected to read data in partitions DG3 and DG4 to get information about fingerprint and iris images and compare them with other personal information. In this study, we aim to use BAC mode to read information stored on the CIC chip. The data read include: information of citizens themselves stored on the DG1, portrait information stored on the DG2, and citizen-related information such as parents, special identities, etc. that are stored on DG13.

### III. DESIGNING AND BUILDING AN APPLICATION TO READ INFORMATION SAFELY FROM CIC WITH CHIP IN VIETNAM

#### A. Setting Up the BAC Mode to Read Data

Figure 3 shows the details of the four stages needed to be performed when the application connects to the CIC with chip, including: First, get information from the MRZ on the CIC as a key seed. Second, establish session key pair (KENC, KMAC) from the key seed. Third, authenticate and exchange new session keys with CIC. Finally, proceed to read the data on the chip using the new session key pair to encrypt and authenticate each other. BAC mode is the access mode to the CIC using 3DES symmetric encryption in CBC mode to encrypt as well as generate the token authentication code. Figure 3 presents details of the test execution procedure for the CIC with chip to connect to the card reader with BAC mode.

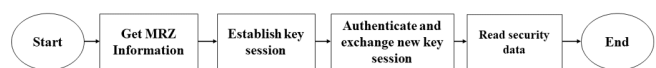


Fig. 3. Software read stages with the BAC mode.

First, the PC will get the MRZ information contained on the CIC either manually or by scanning. The PC receiving the MRZ will generate a key encryption code and a key message authentication code (KENC, KMAC). Next, the PC requests a notification from the CIC via the APDU command sent to the CIC. CIC generates a response message M sending a response

back to the PC. The PC receives the message  $M$  from the CIC and then generates a message  $M'$  and the derived key  $K_{ifd}$ . It calculates  $S = M \parallel M' \parallel K_{ifd}$ , encrypts  $S$  using a  $K_{enc}$  key, and generates message authentication code with ciphertext using the  $K_{mac}$  key. Next, the PC constructs the APDU instruction with the data  $E \parallel MAC$  sent to the CIC. CIC's chip will receive decryption and authentication data with  $K_{enc}$  and  $K_{mac}$  keys and check if the message  $M$  sent before is correct. If it is correct, the CIC will generate the derived key  $K_{ic}$  and a new key seed by xor-ing  $K_{ic}$  with the key derived from the incoming PC code ( $K_{ifd}$ ). Finally, the CIC generates a new session key pair, i.e.  $K_{enc}(new)$  and  $K_{mac}(new)$ , generates a new message  $M_2 = M \parallel M' \parallel K_{ic}$ , encrypts message  $M_2$  with the  $K_{enc}$  key, generates the token authentication code, constructs the APDU instruction with the data  $E \parallel MAC$ , and sends a response back to the PC. At this point, the PC receives, decrypts, and authenticates the data to check if it is the message  $M'$  sent before. If it is correct, the PC will generate the key seed by xor-ing  $K_{ifd}$  and the incoming CIC derived key ( $K_{ic}$ ), and will generate a new session key pair  $K_{enc}(new)$ ,  $K_{mac}(new)$ .

When both the chip and the card reader complete the connection tests successfully, the two sides will communicate using the new session key pair to encrypt and authenticate data, with the following process: the PC sends the APDU command read data to CIC, CIC encrypts data and sends the APDU commands back to the PC. Then, the PC receives the decoded data, validates the data, and checks the received information. This is demonstrated in the application by the following steps: First, the application will proceed to get the MRZ information data as the key seed, then will calculate and set up a session key pair including  $K_{enc}$  and  $K_{mac}$ . The application then sends a connection to the CIC and the CIC authenticates and initiates a new session key exchange. When the new session key exchange is successful, the application will proceed to read the specified data partitions using the session key pair to encrypt and authenticate the data.

### 1) Session Key Establishment

Session Key Establishment is the process by which a user generates a session key pair based on the MRZ information provided by the CIC. The MRZ information serves as the key seed to generate the session key pair [25]. According to the ICAO 9303 standard, the process of establishing session key [22, 25] consists of four main steps:

- **Step 1:** Build MRZ information (organized with 24 numbers according to the following structure: *Document Number* - Consists of 9 digits of CIC from the end, *Birth* is 6 digits on the CIC, *Expiry* is 6 digits of the expiration date on the CIC, *Check* is 1 byte calculated by the formula  $Check = (Document\ Number + 7 + 3 + 1 + 7 + 3 + 1 + 7 + 3 + 1) \bmod 10$ ).
- **Step 2:** Generate the key (here using SHA1 hash).
- **Step 3:** Generate  $K_{enc}$  encryption key (using true random source in chip).
- **Step 4:** Generate key message authentication key (using SHA1 and  $K_{mac}$ ). This process is detailed in the ICAO standard [19].

After successful session key setup, a key pair consisting of an encryption key ( $K_{enc}$ ) and a key message authentication key ( $K_{mac}$ ) will be generated for secure exchange with the CIC.

### 2) Authenticating and Exchanging the New Session Key

This process is done after the session key pair has been previously established. The application will request a message from the CIC, then proceed to use the previously set session key pair to encrypt and generate an authentication code for that message, and then send it to the CIC. The CIC decodes and authenticates the incoming data. If successful, the card will proceed to generate a new session key pair consisting of  $K_{enc}(new)$  and  $K_{mac}(new)$  and then send it back to the application. The two sides of the application and the CIC will be accessed through this new session key pair [25]. First, during the process of authentication and exchange, a new session key is performed, an RND.IC message (8 bytes) is randomly requested from the CIC. The application performs the following operations: Generate random RND.IFD (8 bytes), generate a random K.IFD (16 bytes) key, Join  $S = RND.IFD \parallel RND.IC \parallel K.IFD$  (RND.IC sends request to CIC), encode  $E_{IFD} = E(K_{enc}, S)$ , calculate the message authentication code  $M_{IFD} = MAC(K_{mac}, E_{IFD})$ , build command APDU( $E_{IFD} \parallel M_{IFD}$ ), and send the APDU command to the CIC. Then, CIC receives the data sent to the APDU( $E_{IFD} \parallel M_{IFD}$ ). E<sub>IFD</sub> conducts decryption and authentication with application via MAC, extracts RND.IC from  $S$  and checks if it returns the correct value, generates K.IC key and generates  $K_{seed}(new) = K.IC \text{ XOR } K.IFD$ , generates a new session key pair  $K_{enc}(new)$ ,  $K_{mac}(new)$ , joins  $R = RND.IC \parallel RND.IFD \parallel K.IC$ , encrypts  $E_{IC} = E(K_{enc}, R)$  and computes the message authentication code  $M_{IC} = MAC(K_{mac}, E_{IC})$ , constructs the PDU( $E_{IC} \parallel M_{IC}$ ) command, and sends the APDU command back to the application. The application receives the response data from the APDU tag( $E_{IC} \parallel M_{IC}$ ), conducts E<sub>IC</sub> decryption and authentication with the CIC via MAC, extracts the RND.IFD from  $R$  and checks if the tag returns the correct value, generates  $K_{seed}(new) = K.IC \text{ XOR } K.IFD$ , and generates a new session key pair ( $K_{enc}(new)$ ,  $K_{mac}(new)$ ). Finally, the application and the CIC will generate a new session key pair that is used to encrypt, decrypt, and authenticate the data when reading the data file used for secure data reading.

### 3) The Process of Reading Secure Data

It is done after the software and the CIC have authenticated and set up the session key. During the reading process, the application and the CIC encrypt, decrypt, and authenticate the data based on a new session key pair to ensure that they are confidential and authenticated. Card reading is done through properly constructed APDU commands sent to the chip [DO'85' or DO'87'] [DO'97'] DO'8E'. The response to the APDU command is [DO'85' or DO'87'] [DO'99'] DO'8E'. In this module, the data on each partition are different, so the structure of the APDU command sent to the card is also built differently and must follow the rules of sending access commands to that partition. After sending a command to that partition, there will be response data with status bytes. If the returned data are correct and the status bytes return "9000", then the command sent is correct. In the CICs in Vietnam, the data areas are: DG1

(MRZ information), DG2 (face image encoding), DG3 (fingerprint encryption), DG4...DG12 (backup), DG13 (additional personal information), DG14 (read condition of EAC mode), DG15 (passive authentication), DG16 (backup).

### B. Analysis, Evaluation, and Design of the Application

Figure 4 presents the details of the system and the operation of the reading data application. The application is built with the following devices: Scanner (gets MRZ information from the card), card reader (communication device that transmits APDU commands from the computer to the CIC and vice versa), and software (retrieve and process data in CIC). The user will provide a CIC chip and the scanner connected to the computer will take the MRZ information from the card as input. At the same time, a card reader is connected to the computer through the built-in data reader software. After obtaining the MRZ information as input to the data reader software, CIC will communicate with the software through the previously connected card reader. The application then processes and retrieves data from the card.

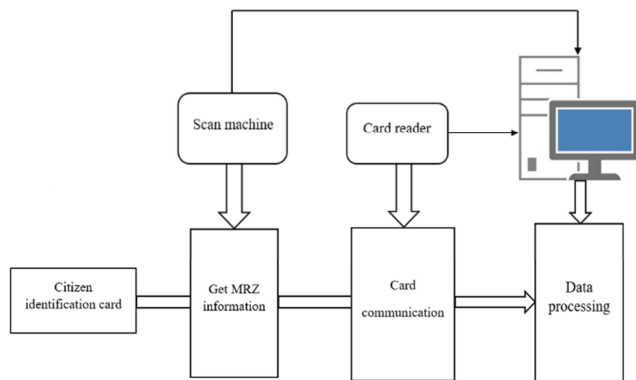


Fig. 4. Built-in system and program that reads the security information of the CIC.

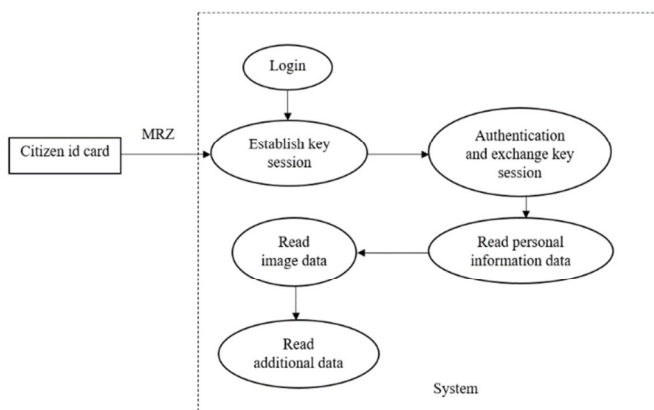


Fig. 5. ID card security information reader program flow.

Figure 5 shows that the software is built into many different modules, each with different functions. The software is built based on the BAC mode. The software is divided into the following modules: Login module, module for setting session key, authentication and session key exchange module, personal data reading module, image data reading module, and

supplementary information reading module. First, users need a password account to log in to the main module of the software. The MRZ of CIC is then put into the session key setup module to generate a session key pair for the next process. The session key exchange and authentication module receives the session key pair, establishes a secure connection to the CIC, and exchanges a new session key pair. The personal information data read module retrieves the data from the DG1 partition. The image data reading module will get the data from the DG2 partition. The additional information reading module retrieves data from partition DG13.

## IV. RESULTS AND DISCUSSION

### A. Building and Designing Modules in CIC Reader Software

Figure 6(a) describes in detail the operation flow of the login module. At first, on the login interface of the software, the user needs to enter the account/password. Next, when the software's module works, if the information is valid, it will take the user to the main interface. In case of incorrect input, the software will not allow login. Figure 6(b) details the operation flow of the authentication module and session key establishment. This module will process information according to the following steps:

- Step 1: The user enters the MRZ on the CIC into the software.
- Step 2: The entered MRZ information is processed in accordance with the provisions of ICAO 9303 standard to get the necessary information.
- Step 3: After obtaining the necessary information from the MRZ, the module will proceed to generate the key.
- Step 4: After having the key seed, the processing software generates a pair of session keys KENC (encryption) and KMAC (key message authentication code).

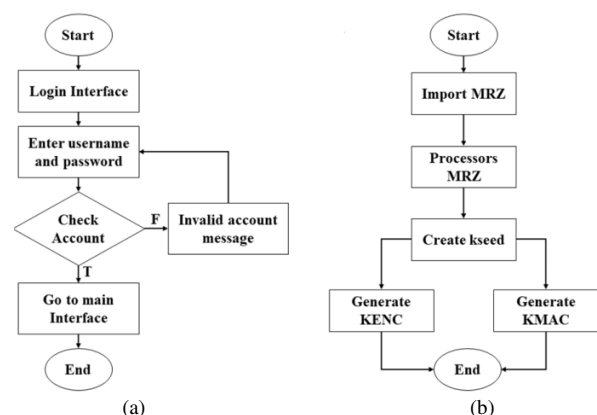


Fig. 6. (a) Login module activity flow, (b) session key exchange and authentication module activity flow.

The new session key exchange and authentication module is operational:

- Step 1: User generates a random plaintext block.

- Step 2: Encrypt and generate token authentication code from plaintext block from KENC AND KMAC set up before.
- Step 3: Connect the cipher-text and the generated message authentication code.
- Step 4: Send S to the card for personal authentication.
- Step 5: If it is legitimate, the card will respond and receive a resend notice.
- Step 6: Decrypt the message and authenticate with the card.
- Step 7: Get the key seed from the card's response message.
- Step 8: Set up a new session key for encryption and token authentication. The details of the operation flow of this module are shown in Figure 7(a).

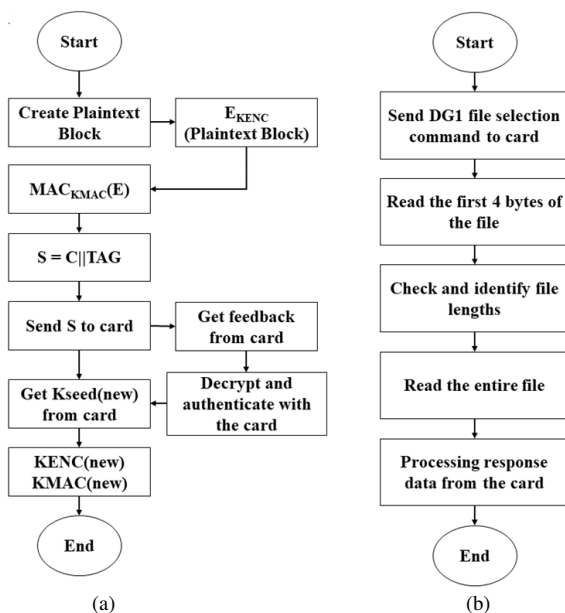


Fig. 7. (a) New session key exchange and authentication module activity flow, (b) operation flow of the data reading module (personal information).

Figure 7(b) shows the details of the operation flow of the personal data reading module. When this module works, the information stored on the CIC read out is processed through the following steps:

- Step 1: Send the file selection command to DG1 to read the card's personal information.
- Step 2: Send the command to read the first 4 bytes of the DG1 file.
- Step 3: Get a response from a valid check card and determine the length of data contained in the file.
- Step 4: Proceed to read the entire file.
- Step 5: Get response from the data processing card from decoding, validating, and extracting valid information. In addition, there is a module to read image data and

additional information. The module that reads image data and reads additional information is basically the same as the module that reads personal data, except for the method of file selection and the processing of the received data.

### B. Building and Perfecting the Functions of the CIC Reader Software

In this study, the results of the readings from the CIC will be obscured, due to concerns regarding confidentiality and privacy. To build an application that reads security information on a CIC, we used a number of devices and supporting tools: a Duali DE-620 smart card reader, a PC with Visual Studio 2019 installed, programming language: C/C++, OpenSSL library. Figure 8(a) shows the designed login module interface. On this interface, citizens only need to enter the correct account and password and then they can start working with the CIC reader software with the main interface of Figure 8(b). On this main interface, if a citizen wants to read the information stored on the CIC, it is only needed to enter the MRZ information string (written on the back of the CIC).

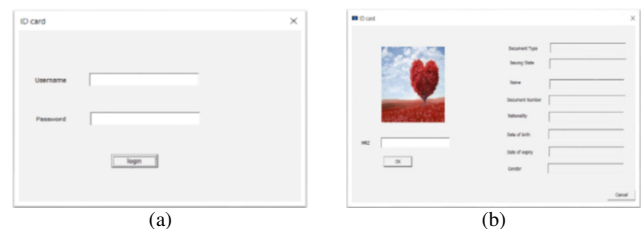


Fig. 8. (a) Login interface, (b) main interface.

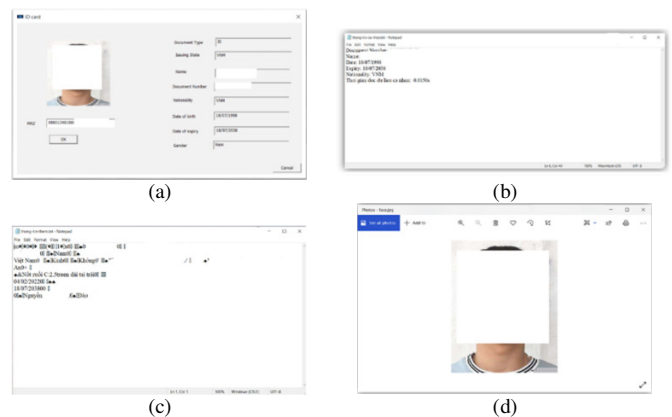


Fig. 9. (a) Results of reading personal information on the CCCD, (b) results of reading data on DG1, (c) data reading results on DG13, (d) data reading results on DG2.

Figure 9 shows the results by the software when reading security information stored on the CIC's chip. Figure 9(a) is the citizen's output displayed on the CIC reader software. Figures 9(b),(c) show the detailed results of personal information read on DG1 and DG13. The read out results are saved on 2 separate corresponding txt files. Figure 9(d) shows the result of a detailed portrait of an individual read on the DG2 partition that is saved on a txt file. The results obtained when reading individual data on the chip CIC include: Document Type, Issuing State, Full Name, CIC Number, Nationality, Date of



Birth, Expiration Date, Gender, Portrait Image, and other personal data.

TABLE I. READING TIME RESULT OF SECURED DATA ON THE CIC CHIP

No	End-to-end authentication time (s)	Reading time of DG1 (s)	Reading time of DG2 (s)	Reading time of DG13 (s)	Sum (s)
1	0.1950	0.0260	0.9630	0.0360	1.22
2	0.1960	0.0260	0.9640	0.0380	1.224
3	0.1960	0.0260	0.9570	0.0370	1.216
4	0.1950	0.0260	0.9430	0.0380	1.202
5	0.1950	0.0250	0.9640	0.0390	1.223
Average	0.1954	0.0258	0.9582	0.0376	1.217

Finally, in order to evaluate the performance of the CIC reader software, the time to read citizen information stored on each partition DG1, DG2, DG13 and end-to-end authentication of chip CIC was determined. Table I gives the results of the software's execution speed performance when reading data. It is seen that the total time to read the entire data (end-to-end authentication, DG1, DG2 and DG13) is approximately 1.217s. So, the software reading data on the CIC has an average time of about 1.2 to 1.3s, which is a relatively fast read speed, which is suitable for real time applications.

### C. Analyzing, Evaluating, and Testing the Source Code of the CIC Reader Software

The Fortify Static Code Analyzer toolkit (Version 22.1.0.0166) was used to analyze and evaluate the CIC reader software source code. Table II gives the detailed results. There are error issues related to Buffer Overflow, Variable Never Used, Integer to Character, Signed to Unsigned, Unchecked Return Value, but these problems are not of a serious nature and don't affect the security of the software. Specifically, these issues are analyzed below.

- There is a problem related to buffer overflow (Buffer Overflow: 1 issue), which is warned in the memcpy function (void \*memcpy(void \*str1, const void \*str2, size\_t n) of the reader, occurring when the size of the copied array is smaller than the data array size, which does not guarantee program optimization according to [30-33]. The Fortify Static Toolkit Code Analyzer showed how to fix the source code (check the correct copy array length and number of bytes to copy) according to [34].

TABLE II. RESULTS OF SOURCE CODE OF CIC SOFTWARE TESTING WITH THE FORTIFY STATIC CODE ANALYZER

Category	Fortify Priority (audited/total)				Total Issues
	Critical	High	Medium	Low	
Buffer Overflow	0	0/1	0	0	0/1
Poor Style: Variable Never Used	0	0	0	0/2	0/2
Type Mismatch: Integer to Character	0	0/1	0	0	0/1
Type Mismatch: Signed to Unsigned	0	0/1	0	0	0/1
Unchecked Return Value	0	0	0	0/1	0/1
Weak Cryptographic Hash	0	0	0	0/5	0/5

- There are two problems related to unused variables (Variable Never Used: 2 issues). The error occurs because two variables, num and lRetVal have been declared but not used. The fix is to delete these two variables.
- There is an issue related to integer to character conversion (Integer to Character: 1 issues). The error is that the function returns an unsigned char to an int type, but the return value is assigned to a char type. The fix is to resynchronize the return value with the same type.
- There is an issue with the signed to unsigned conversion (Signed to Unsigned: 1 issues). The error occurs because the arrays ch[] (type char) and mrz\_input[] (type unsigned char) are not of the same data type. The fix is to re-declare one of the two arrays with the same data type char or unsigned char.

```

Sink: Variable: num
Enclosing Method: OnBnClickedOk()
File: readercccdDlg.cpp:336
Taint Flags:

333 fclose(fp);
334 //////////////////////////////////////////////////
335
336 int num;
337 char c;
338 //FILE* fp;
339 fopen_s(&fp, "test.txt", "r");

Sink: Variable: lRetVal
Enclosing Method: OnBnClickedOk()
File: readercccdDlg.cpp:204
Taint Flags:

201
202 BYTE atr[40];
203 INT atrLength;
204 LONG lRetVal;
205
206
207

```

Fig. 10. Unused variables.

```

Sink: AssignmentStatement
Enclosing Method: OnBnClickedOk()
File: readercccdDlg.cpp:343
Taint Flags:

340 char mang[200];
341 //Đọc từng ký tự từ file cho tới khi gặp EOF
342 int i = 0;
343 while ((c = fgetc(fp)) != EOF)
344 {
345 //Xuất từng ký tự ra màn hình
346 mang[i] = c;

```

Fig. 11. Integer to character conversion issue.

```

Sink: AssignmentStatement
Enclosing Method: OnBnClickedOk()
File: readercccdDlg.cpp:250
Taint Flags:

247
248 for (int i = 0; i < 23; i++)
249 {
250 mrz_input[i] = ch[i];
251 mrz_input[i] = mrz_input[i] - 0x30;
252 }
253 memcpy(mrz, mrz_input, 9);

Sink: InitInstance()
Enclosing Method: InitInstance()
File: readercccd.cpp:53
Taint Flags:

50 InitCtrls.dwICC = ICC_WIN95_CLASSES;
51 InitCommonControlsEx(&InitCtrls);
52
53 CWinApp::InitInstance();
54
55
56 AfxEnableControlContainer();

```

Fig. 12. Unchecked return value issue.

- There is an issue with the unchecked return value (Unchecked Return Value: 1 issue). This error occurs when the InitInstance() function returns no value. Since InitInstance() is an interface function of the MFC available when creating the interface, it cannot be dropped.

Analyzing, evaluating, and testing the problems in CIC reader software with Fortify Static Code Analyzer toolkit (Version 22.1.0.0166), shows that the developed and built software is guaranteed source code-safe. This is enough to confirm that the CIC reader software will be secure against



some types of attacks when deployed in real-world commercial applications.

## V. CONCLUSION

Reading citizen information stored on a CIC with chip is practical for daily services, such as electronic payments, online money transfers, online public services, etc. In this paper, we used the BAC mode to set up a secure channel in the process of reading citizen information stored on CIC. We used cryptographic algorithms such as 3DES, SHA1, and MAC, to ensure that during the authentication process between the chip and the terminal the communication in the process of reading citizen information from the chip on CIC is secure. As a result, the software reads the information stored on DG1, DG2, and DG3 fields with safety and security. The performance of the software has a reading and data processing speed of approximately 1.2 s to 1.3 s. The CIC reader software source code has been evaluated and tested with the Fortify Static Code Analyzer tool (Version 22.1.0.0166) to confirm the software its robustness and safety.

## ACKNOWLEDGMENT

The authors thank the Academy of Cryptography Techniques and the Minister of Education and Training (MOET) for supporting this work under grant number B2022-GHA-10.

## REFERENCES

- [1] L. C. Guillou and M. Ugon, "Smart Card a Highly Reliable and Portable Security Device," in *Lecture Notes in Computer Science*, New York, NY, USA: Springer, 1987, pp. 464–479.
- [2] K. Vedder, "GSM: Security, Services, and the SIM," in *Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, 1998, pp. 224–240.
- [3] V. Guyot, "Smart card, the stealth leaker," *Journal in Computer Virology*, vol. 8, no. 1, pp. 29–36, May 2012, <https://doi.org/10.1007/s11416-012-0159-y>.
- [4] W. Rankl and W. Effing, *Smart Card Handbook*, Fourth Edition. New York, NY, USA: John Wiley & Sons, 2010.
- [5] D. Basin, R. Sasse, and J. Toro-Pozo, "The EMV Standard: Break, Fix, Verify," in *IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, Dec. 2021, pp. 1766–1781, <https://doi.org/10.1109/SP40001.2021.00037>.
- [6] M. E. Haykin and R. B. J. Warnar, *Smart Card Technology: New Methods for Computer Access Control*. Gaithersburg, MD, USA: National Institute of Standards and Technology, 1988.
- [7] K. Markantonakis, "Multi Application Smart Card Platforms and Operating Systems," in *Smart Cards, Tokens, Security and Applications*, K. E. Mayes and K. Markantonakis, Eds. Boston, MA, USA: Springer, 2008, pp. 51–83.
- [8] W. Rankl and W. Effing, *Smart Card Handbook*, 1st Edition. Chichester, WS, England: Wiley, 2010.
- [9] H. Guo, *Smart Cards and their Operating Systems*. Helsinki, Finland: Helsinki University of Technology, 2002.
- [10] 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>.
- [11] K. R. Wilcox, "Multi-application smart cards: Card operating systems and application security," presented at the 21st Computer Science Seminar, 2003.
- [12] E. Barker and N. Mouha, *Recommendation for Triple Data Encryption Algorithm (TDEA) Block Cipher*. Gaithersburg, MD, USA: National Institute of Standards and Technology, 2017.
- [13] B. Robisson *et al.*, "Smart security management in secure devices," *Journal of Cryptographic Engineering*, vol. 7, no. 1, pp. 47–61, Apr. 2017, <https://doi.org/10.1007/s13389-016-0143-4>.
- [14] T. Rosteck, *Connected Secure Systems (CSS) Division Call hosted by Deutsche Bank*. infineon, 2021.
- [15] A. H. Al-Omari, "Lightweight Dynamic Crypto Algorithm for Next Internet Generation," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4203–4208, Jun. 2019, <https://doi.org/10.48084/etasr.2743>.
- [16] P. A. Karger, D. C. Toll, E. R. Palmer, S. K. McIntosh, S. Weber, and J. W. Edwards, "Implementing a High-Assurance Smart-Card OS," in *Financial Cryptography and Data Security*, Tenerife, Canary Islands, Jan. 2010, pp. 51–65, [https://doi.org/10.1007/978-3-642-14577-3\\_7](https://doi.org/10.1007/978-3-642-14577-3_7).
- [17] O. Dagdelen, "The Cryptographic Security of the German Electronic Identity Card," Ph.D. dissertation, Technical University of Berlin, Berlin, Germany, 2013.
- [18] U. Iftikhar, K. Asrar, M. Waqas, and S. A. Ali, "Evaluating the Performance Parameters of Cryptographic Algorithms for IOT-based Devices," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7867–7874, Dec. 2021, <https://doi.org/10.48084/etasr.4263>.
- [19] *Order of Identification and Regulations for Issue Issuance*. 1957.
- [20] *Circular No. 59/2021/TT-BCA detailing the implementation of the Law on Citizen Identification*. 2021.
- [21] *ISO/IEC 18013-3:2017, Information technology — Personal identification — ISO-compliant driving licence — Part 3: Access control, authentication and integrity validation*. ISO, 2017.
- [22] *Doc 9303: Machine Readable Travel Documents: Part 3: Specifications Common to all MRTDs*, 8th ed. ICAO, 2021.
- [23] "Radio Frequency Protocol and Application Test Standard for eMRTD – Part 3," ICAO, Technical Report, Mar. 2018.
- [24] D. Cooper, H. Ferraiolo, K. Mehta, S. Francomacaro, R. Chandramouli, and J. Mohler, *NIST Special Publication 800-73-4: Interfaces for Personal Identity Verification – Part 1: PIV Card Application Namespace, Data Model and Representation*. Gaithersburg, MD, USA: NIST, US Department of Commerce, 2015.
- [25] *Advanced Security Mechanisms for Machine Readable Travel Documents and eIDAS Token – Part 1*. Germany: Federal Office for Information Security, 2015.
- [26] *ISO/IEC 14443-4:2008, Identification cards — Contactless integrated circuit cards — Proximity cards — Part 4: Transmission protocol*. ISO, 2008.
- [27] *ISO/IEC 7816-4:2020(en), Identification cards — Integrated circuit cards — Part 4: Organization, security and commands for interchange*. ISO, 2020.
- [28] H. Mestiri, I. Barraj, and M. Machhout, "AES High-Level SystemC Modeling using Aspect Oriented Programming Approach," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6719–6723, Feb. 2021, <https://doi.org/10.48084/etasr.3971>.
- [29] A. S. Alshammari, "Comparison of a Chaotic Cryptosystem with Other Cryptography Systems," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6187–6190, Oct. 2020, <https://doi.org/10.48084/etasr.3745>.
- [30] G. McGraw, "Software Security: Building Security In," in *17th International Symposium on Software Reliability Engineering*, Raleigh, NC, USA, Nov. 2006, <https://doi.org/10.1109/ISSRE.2006.43>.
- [31] A. Apvrille and M. Pourzandi, "Secure software development by example," *IEEE Security & Privacy*, vol. 3, no. 4, pp. 10–17, Jul. 2005, <https://doi.org/10.1109/MSP.2005.103>.
- [32] J. Koziol *et al.*, *The Shellcoder's Handbook: Discovering and Exploiting Security Holes*. New York, NY, USA: Wiley, 2004.
- [33] M. Howard and D. LeBlanc, *Writing Secure Code, Second Edition*, 2nd ed. Redmond, WA, USA: Microsoft Press, 2003.
- [34] "About Strsafe.h - Win32 apps," *Microsoft*. <https://learn.microsoft.com/en-us/windows/win32/menurc/strsafe-ovw>.

# A Simplified Deformation Estimation Method for Anchor Piles of Sheet Pile Quay Walls under Kinematic Forces during Earthquakes

**Kenichiro Miyashita**

Pacific Consultants Co, Ltd, Japan  
kenichirou.miyashita@os.pacific.co.jp  
(corresponding author)

**Takashi Nagao**

Kobe University, Japan  
nagao@people.kobe-u.ac.jp

Received: 6 November 2022 | Revised: 26 December 2022 | Accepted: 29 December 2022

## ABSTRACT

In the seismic design of quay walls, it is necessary to evaluate the deformation of the walls during earthquakes as well as the safety of structural members. However, conventional seismic design methods for sheet pile quay walls cannot accurately determine the degree of deformation. One reason for this is that conventional methods do not consider kinematic forces acting on an anchor pile due to the deformation of the ground. This study proposes a simplified estimation method for anchor pile deformation under the influence of kinematic forces. The results of two-dimensional finite element analysis reveal that anchor pile deformation involves rotational and translational components caused by the kinematic forces, which the conventional methods do not consider. The deformation of the anchor pile caused by kinematic forces was 30%–40% of the total deformation at the pile head. It was clarified that unlike horizontally stratified ground, shear stress is generated in the ground before an earthquake resulting in the kinematic force acting on the anchor pile during the earthquake. Furthermore, a simplified method for estimating the deformation of the anchor pile under kinematic forces that uses one-dimensional seismic response analysis considering the predicted shear stress based on a theoretical equation is proposed. It was demonstrated that the proposed method accurately reproduces the anchor pile deformation.

*Keywords-sheet pile quay wall; seismic design; finite element analysis; anchor pile; kinematic force*

## I. INTRODUCTION

In the seismic design of structures, the seismic safety of structural members is the main focus of the assessment [1, 2]. Quay walls are a major part of port facilities, and are constructed on soft ground in coastal areas. In the event of an earthquake, seaward deformation of the quay walls can occur, which may interfere with the use of berths, even if the structural members are safe [3–6]. Therefore, in seismic design of quay walls, it is necessary to evaluate the degree of deformation of the quay wall with high accuracy. A sheet pile quay wall is a structure that is often constructed in quite soft ground. In areas where significant ground motions can occur, an anchor pile is often driven into the ground at a position away from the sheet pile to resist seismic forces. Conventional seismic design of anchor piles generally only considers tie rod tension and subgrade reaction as loads acting on the anchor piles [1]. Although this conventional method has been effective for designing safe cross-sections of piles to resist seismic forces, it predicts the deformation degree with low accuracy [7,

8]. Various studies have shown that two-dimensional (2D) Finite Element Analysis (FEA) can accurately reproduce the seismic deformation of quay walls [9–18]. However, 2D FEA is not a standard technique in design practice, due to its large computational burden. One reason for the low deformation calculation accuracy by the conventional method is because it does not consider the force generated by the deformation of the ground (kinematic force). Both inertial and kinematic forces act against a sheet pile quay wall during earthquakes [19–21]. However, the conventional method only considers inertial forces, such as hydrostatic and seismic earth pressure. Although several studies have been conducted on applying kinematic force to seismic design of piles [22–24], they focused on piles driven into horizontally stratified ground. To date, no studies have been conducted on piles under kinematic forces installed in uneven ground, such as that behind a sheet pile wall. This study proposes a simplified estimation method for the deformation of the anchor pile during earthquakes by kinematic forces. First, deformation of the anchor pile of a sheet pile quay wall was evaluated using 2D FEA. The bending

deformation of the anchor pile was calculated based on its curvature. By subtracting the bending deformation from the deformation of the anchor pile, its deformation due to kinematic forces was obtained. Because the ground in front of the anchor pile is not horizontally stratified, but irregular, shear stress is generated in this area before an earthquake and residual deformation occurs in the quay wall during the earthquake. We clarified that this is responsible for the kinematic force acting on the anchor pile during an earthquake. This paper proposes a simplified estimation method for the deformation of an anchor pile by kinematic forces. The method applies one-dimensional (1D) seismic response analysis considering an initial shear stress distribution in the ground predicted by a theoretical equation.

## II. CONVENTIONAL SEISMIC DESIGN METHOD FOR ANCHOR PILES

The conventional method [1] assumes that an active failure region is generated behind the sheet pile and a passive failure region in front of the anchor pile. These regions are independent and do not affect each other (Figure 1). Here, each region is determined based on the seismic earth pressure by Mononobe-Okabe theory [25]. The anchor pile is located at the point where the active failure line is generated behind the sheet pile at the seabed level, which intersects the passive failure line generated in front of the anchor pile at  $l_{m1}/3$  below the tie rod mounting height, where  $l_{m1}$  is the depth at which the bending moment of the anchor pile first becomes zero below the tie rod mounting height according to the Winkler foundation model. By installing the pile at this position, the conventional method assumes that the pile is embedded in horizontally stratified ground. Figure 2 shows the application method of the Winkler foundation model to the anchor pile. By applying the tie rod tension as a concentrated load and the subgrade reaction as a distributed load, the deformation magnitude and the cross-sectional force of the pile are calculated by solving the differential equation (1). Since the elongation of the tie rod is negligible, the deformation of the anchor pile is assumed to be equal to that of the quay wall.

$$EI \frac{d^4 u}{dy^4} = -Bp \quad (1)$$

where  $EI$  is the flexural rigidity of the pile,  $y$  is the depth,  $u$  is the lateral displacement of the pile at depth  $y$ ,  $p$  is the subgrade reaction, and  $B$  is the pile width.

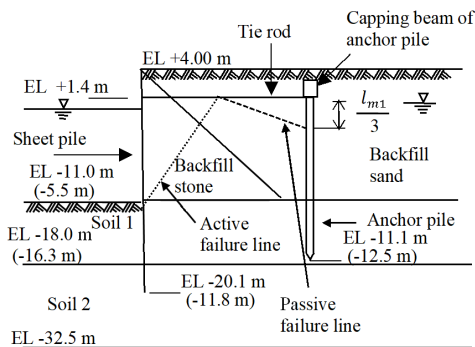


Fig. 1. Profile of the studied sheet pile quay wall. Note: values in parentheses are the ones for water depth of 5.5m, EL: elevation.

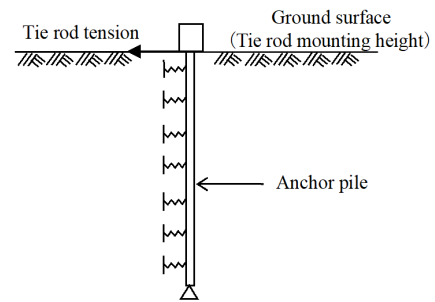


Fig. 2. Forces applied to the pile.

The nonlinear load–displacement relationship of a pile is often modeled considering the nonlinear characteristics of the soil [26–28]. A nonlinear relationship between  $p$  and  $u$  is considered based on experimental test results as follows [1]:

$$p = k_c u^{0.5} \quad (2)$$

$$p = k_s y u^{0.5} \quad (3)$$

where  $k_s$  and  $k_c$  are the Subgrade Reaction Modulus (SRM) values for S-type and C-type ground, respectively. The ground is classified as S-type when the N-values obtained from the standard penetration test increase with increasing depth and as C-type when the N-value is independent of the depth. The current study used (2), assuming C-type of ground. The SRM is often considered to be dependent on the foundation width [29, 30], however, in the design of anchor piles, this dependence is ignored [1].

## III. METHODS

Two sheet pile quay wall profiles designed according to [1] were used. As shown in Figure 1, the water depths were –5.5 and –11.0m, and the seismic coefficients were 0.10 and 0.15, respectively. The elevation of the ground surface was 4.0m. These values were set considering typical values used in design practice. Tables I and II list the ground conditions and dimensions of the structural members. The natural period of the ground was set to 0.8s for both ground conditions based on the values typically observed in the field. This study applied the FEA code FLIP [31], which is widely used for evaluating the seismic response of port facilities. Accurate reproduction of the seismic responses of various quay walls and grounds has been demonstrated using this code [32–34]. The code includes the multi-spring model for the ground [35] which considers the ground response under principal stress axis rotation. The nonlinear characteristics of the soil were modeled using the hyperbolic model [36] and (4) and (5). The hyperbolic model shows a substantial damping coefficient compared with the experimental result in the large shear strain range. Thus, the code modifies the hysteresis curve to reduce the area of the hysteresis loop compared to that given by the Masing rule [37] to prevent the damping coefficient from exceeding the maximum value, as shown in Table I [38]. The shear modulus of the soil was assumed to be dependent on the effective confining stress, as shown in (6).

$$\tau = \frac{G\gamma}{(1 + |\frac{\gamma}{\gamma_r}|)} \quad (4)$$

$$\gamma_r = \frac{\sigma'_m \sin \phi}{G} \quad (5)$$

$$G = G_{ma} \left( \frac{\sigma'_m}{\sigma'_{ma}} \right)^{m_g} \quad (6)$$

where  $G$  is the shear modulus ( $\text{kN/m}^2$ ),  $\tau$  is the shear stress ( $\text{kN/m}^2$ ),  $\gamma$  is the shear strain,  $\gamma_r$  is the reference shear strain,  $\sigma'_m$  is the effective confining stress ( $\text{kN/m}^2$ ),  $m_g$  is a parameter indicating the dependency on the confining pressure and  $m_g = 0.5$  in based on [39]. The standard method for setting the parameters of the soil properties in [40] was followed in the present analysis. The effects of liquefaction were neglected.

TABLE I. GROUND CONDITION

Soil	$\rho$ ( $\text{t/m}^3$ )	$G_{ma}$ ( $\text{kN/m}^2$ )	$\sigma'_{ma}$ ( $\text{kN/m}^2$ )	$\phi$ ( $^\circ$ )
Backfill sand	2.0	58300	89.8	38
Soil 1	2.0	72200	198.5	39
Soil 2	2.0	125000	279.2	39
Backfill stone	2.0	101250	98	40
Wall friction angle ( $^\circ$ )	In front of the wall: $\delta = 15$ Behind the wall: $\delta = 0$			
Common physical properties	$h_{max} = 0.24$ , $K_w = 2200000 \text{ kN/m}^2$ $\nu = 0.33$			

$\rho$  is the saturated unit weight,  $G_{ma}$  is the reference shear modulus,  $\sigma'_{ma}$  is the reference average effective confining stress,  $\phi$  is the shear resistance angle,  $h_{max}$  is the maximum damping coefficient,  $K_w$  is the bulk modulus of pore water,  $\nu$  is the Poisson's ratio.

TABLE II. DIMENSIONS OF THE STRUCTURAL MEMBERS

Water depth (m)	Sheet pile		Tie rod	
	Depth of embedment (m)	Moment of inertia of area ( $\text{m}^4/\text{m}$ )	Area ( $\text{m}^2/\text{m}$ )	Length (m)
-5.5	-11.8	0.000104	0.00063	13.5
-11.0	-20.1	0.000791	0.00128	18.9

Water depth (m)	Anchor pile	
	Pile length (m)	Are moment of inertia ( $\text{m}^4/\text{m}$ )
-5.5	14.2	0.000304
-11.0	12.5	0.000516

Elastic modulus of all structural members is  $200 \text{ kN/mm}^2$

Sheet pile, anchor pile, and tie rod were modeled as linear beam elements. A thin steel-pipe pile was used as the anchor pile. Viscos boundaries were applied to both bottom and side boundaries. A bi-linear joint element was applied to the boundary between the wall (i.e. sheet pile and capping beam of the anchor pile) and the soil to consider wall friction. The friction between the tie rod and the soil was neglected. As the current study used 2D analysis, special care was required in setting the boundary conditions between the anchor pile and the surrounding soil. The 3D effect at the boundary between the soil and the pile was considered using a soil spring element [41]. The soil spring force  $k_{hp}$  calculated using (7) is proportional to the ratio of the increments of shear stress to shear strain of the soil ( $d\tau/d\gamma$ ), where the coefficients  $\alpha_p$  and  $\beta_p$  were selected according to the diameter and installation interval of the pile:

$$k_{hp} = \frac{\alpha_p d\tau}{\beta_p d\gamma} \quad (7)$$

Figure 3 shows the FE mesh around the quay wall that was used in the calculations of water depth  $-5.5 \text{ m}$  as an example. The FE model had a horizontal length of  $310 \text{ m}$ . The mesh height was set to transmit seismic waves up to  $15 \text{ Hz}$ . Two ground motions were used in the calculations: the Hachinohe wave and the Iwakuni wave. These ground motions were calculated using seismic hazard analysis considering the source, path, and site amplification characteristics at the Hachinohe Port and Iwakuni Port, Japan, with a return period of 75 years [42]. The site amplification characteristics considered here are the ones due to shallow and deep subsurfaces [43–45]. Figure 4 shows the time history waveforms for the two ground motions. The maximum acceleration values were  $1.0$  and  $1.5 \text{ m/s}^2$  for the Hachinohe wave and the Iwakuni wave, respectively. The predominant frequency ranges are different between the Hachinohe and Iwakuni waves. The Hachinohe wave is dominated by low frequencies in the range of  $0.4$ – $1.5 \text{ Hz}$ , whereas the Iwakuni wave is dominated by high frequencies around  $4.0 \text{ Hz}$ .

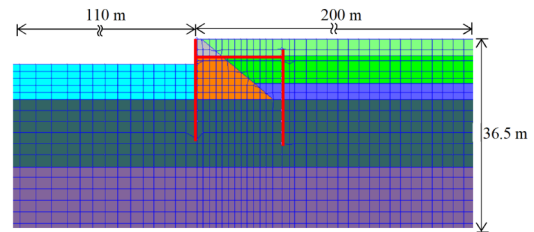


Fig. 3. FE mesh.

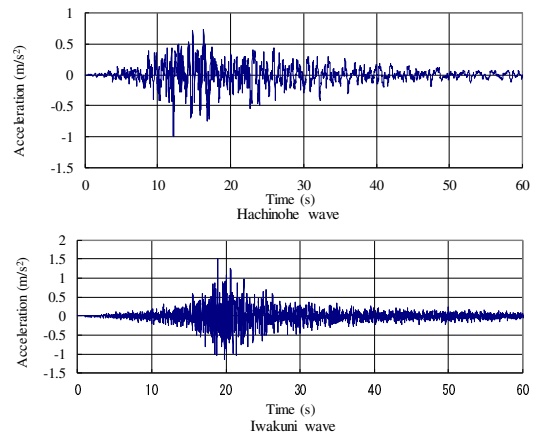


Fig. 4. Ground motion.

#### IV. RESULTS AND DISCUSSION

##### A. Effect of the Kinematic Force on the Deformation of Anchor Piles

Figure 5 presents the residual displacement of the anchor pile calculated using FEA and the conventional method. The displacement of the anchor pile using the conventional method assumed the abovementioned nonlinear relationship between the pile displacement and subgrade reaction. The tie rod tensile force was set to the residual tensile force of the tie rod calculated by FEA, as shown in Table III.

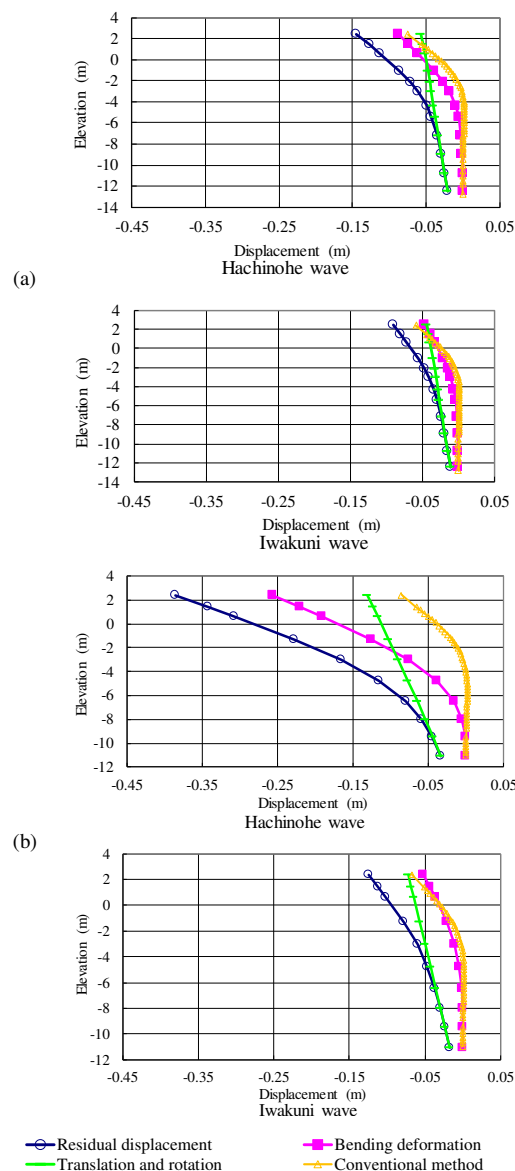


Fig. 5. Displacement of an anchor pile. Water depth (a) -5.5m, (b) -11.0m.

TABLE III. RESIDUAL TENSILE FORCE OF A TIE ROD

Water depth (m)	Hachinohe wave (kN)	Iwakuni wave (kN)
-5.5	216	185
-11.0	349	300

It is seen that the Hachinohe wave left larger displacements than the Iwakuni wave. The explanation is that the Hachinohe wave is dominated by the low frequency range, which has a large influence on ground deformation, while the Iwakuni wave is dominated by the high frequency range, which has a little influence on the ground deformation. The bending deformation was calculated using FEA by double integration of the residual curvature of the anchor pile when both the translation and rotation angles were zero at the pile tip. Translation and rotation were obtained by subtracting the bending deformation from the residual displacement. Translation and rotation at the

pile head were 30%–40% of the residual displacement, which cannot be neglected. The translational and rotational displacements are due to the kinematic force. No study has shown explicitly that the displacement of anchor pile includes three components, i.e. translation, rotation, and bending deformation, however, it can be confirmed by model experimental results [13] and analytical results [8] that the displacement of anchor pile includes components other than bending deformation.

In contrast, the conventional method assumes that the anchor pile is restrained by the ground at its base and cannot deform, so its deformation is only the bending deformation. The kinematic force is not considered in the conventional method, which contributes to its low estimation accuracy for the deformation degree of the anchor piles. Several studies estimated the amount of the anchor pile displacement assuming the occurrence of a slip line in the ground passing through the lower end of the anchor pile [15, 46]. However, it has been pointed out that the technique underestimates the displacement amount of the anchor pile [15]. To accurately calculate the displacement amount of the anchor pile, it is required to evaluate the kinematic force appropriately. When the degrees of bending deformation calculated by FEA and the conventional method are compared, the depth at which the bending deformation begins to increase greatly differs. Therefore, in the case of a Hachinohe wave applied to a quay wall with a water depth of -11m, the deformation at the pile head obtained by the conventional method is much smaller than that calculated by FEA. This difference is attributed to the differences in the resistive properties of the ground in front of the anchor pile assumed by the two methods.

### B. Origins of the Kinematic Force

Sheet pile quay walls have a heterogeneous geometry, resulting in shear stresses being generated over a wide area of the nearby ground. Figure 6 shows a contour map of the initial shear stress. The initial shear stress above the seabed level is small in the active area behind the sheet pile and large in other areas between the sheet pile and the anchor pile. The initial shear stress below the seabed level is generated over a wide area around the sheet pile quay wall. For this reason, the shear stress of the ground fluctuates around the initial shear stress value, rather than zero, at the time of the earthquake. This leads to the generation of considerable shear strain during an earthquake due to the nonlinearity of the soil, which can result in significant residual ground deformation. To examine the influence of the initial shear stress on the residual deformation of the soil, 1D nonlinear seismic response analysis of the ground was performed by applying the initial shear stress extracted from the 2D FEA, where the initial shear stress refers to that at the positions shown in the red column in Figure 6. This is because the 1D analysis should use the initial shear stress free from the effects of the displacement of the anchor pile, and the large initial shear stress that causes soil deformation. As a reference, 1D analysis without the initial shear stress was also performed. Figure 7 compares the residual displacements of 1D analysis and 2D FEA results. In the case of the 2D FEA results, the translational and rotational displacements shown in Figure 5 are illustrated. As expected,



1D analysis without the initial shear stress resulted in almost no residual displacement, while 1D analysis with the initial shear stress resulted in a larger seaward residual displacement. Hence, the initial stress was found to act as implicit lateral force acting on the ground. The displacements determined by 1D analysis with the initial shear stress agreed well with those calculated by the 2D FEA.

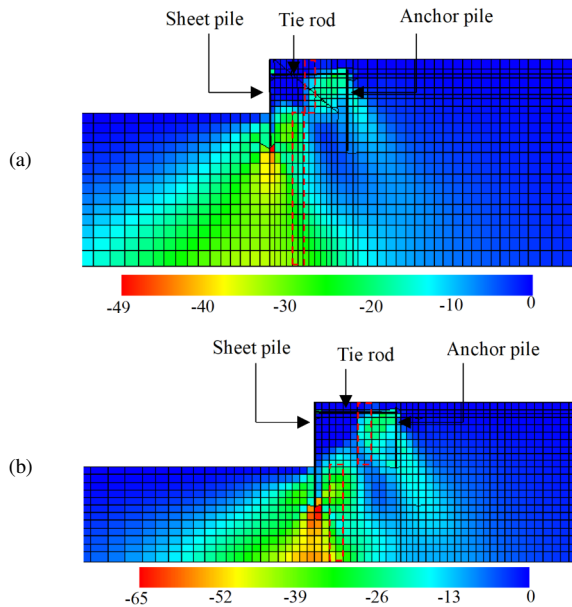


Fig. 6. Contour map of the initial stress, water depth:(a) -5.5m, (b) -11m.

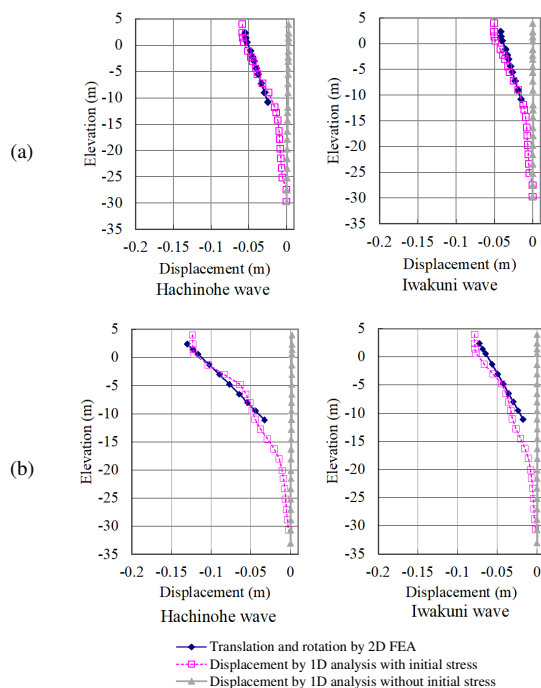


Fig. 7. Comparison of displacements. Water depth:(a) -5.5m, (b) -11m.

### C. A Simplified Estimation Method of Anchor Pile Deformation under Kinematic Force

A simplified method for estimating the deformation of anchor piles under kinematic force is proposed in this paper, which does not require reference to the results of 2D FEA. Since a wharf constitutes irregular ground, the effective weight at the same depth is different in the ground at the front and the back of the sheet pile wall. Thus, shear stress is generated in the ground. An equation for calculating the ground shear stress when a distributed load is applied, based on Boussinesq's equation [47], is presented here. The simplified estimation method calculates the shear stress in the ground using the effective weight of the ground as a distributed load based on (8)-(10) and Figure 8.

$$\tau_{xy}(y) = \int_0^y \frac{\gamma_s}{2\pi} \cos 2\alpha_1(z) dz - \int_0^{y-H} \frac{\gamma_s}{2\pi} \cos 2\alpha_2 dz \quad (8)$$

$$\alpha_1(z) = -\tan^{-1} \frac{L_R}{y-z} \quad (9)$$

$$\alpha_2(z) = \tan^{-1} \frac{L_L}{y-H-z} \quad (10)$$

where  $\gamma_s$  is the unit weight of the ground ( $\text{kN/m}^3$ ),  $z$  is the depth (m),  $\alpha_1$  and  $\alpha_2$  are the angles between the vertical line and the line connecting the estimation position and the edge of the distributed load ( $^\circ$ ),  $H$  is the wall height (m),  $L_R$  is the distance between the sheet pile wall and the anchor pile (m), and  $L_L$  is the distance from the sheet pile wall to the intersection of the passive failure line generated at the bottom of the sheet pile meets the seabed (m). The passive failure angle is  $17^\circ$  based on the Coulomb passive earth pressure.

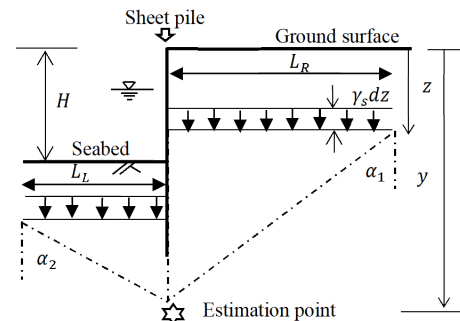


Fig. 8. Schematic diagram of the simplified estimation method for shear stress.

The first and second terms on the right side of (8) are the shear stress due to the effective weight of the ground behind and in front of the sheet pile wall, respectively. The range of the distributed load behind the sheet pile wall was determined from the wall to the anchor pile as the range where large shear stresses are generated (based on 2D FEA results). The range of the distributed load in front of the sheet pile wall is defined by the width of the passive failure region generated at the bottom of the sheet pile wall. This is assumed to be the region that mainly affects the shear stress. Since the simplified estimation method does not consider active failure regions behind the wall, the shear stress is maximum at the wall position, which was thus used to calculate the shear stress in the proposed estimation method.



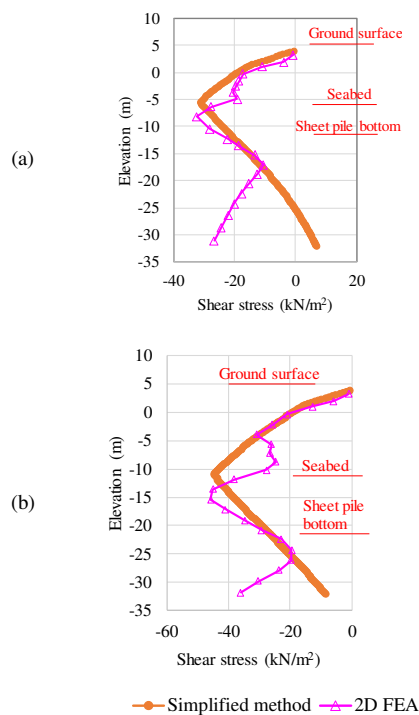


Fig. 9. Comparison of the shear stress. Water depth:(a) -5.5m, (b) -11m.

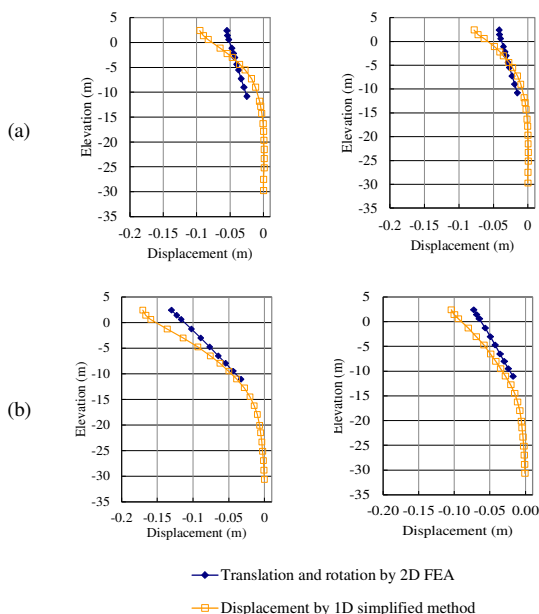


Fig. 10. Comparison of the displacements under various conditions. Water depth:(a) -5.5m, (b) -11m. Left: Hachinohe wave, right: Iwakuni wave.

Figure 9 compares the shear stress estimated by the simplified method and that shown in Figure 6. Although the method somewhat overestimates the shear stress near the seabed, it generally reproduced the shear stress distribution obtained by the 2D FEA from the ground surface to the seabed. At depths below the seabed, the shear stress determined by 2D

FEA decreased from the seabed to a position slightly deeper than the sheet pile bottom and increased at deeper depths, whereas the shear stress determined by the simplified estimation method decreased monotonically below the seabed. However, it is assumed that the shear stress below the seabed has no significant effect on the degree of deformation of the anchor pile. The deformation obtained using the 1D simplified estimation method and the translational and rotational deformation of the anchor pile obtained by 2D FEA are compared in Figure 10. Although the deformation obtained by the simplified estimation method is a little larger than that obtained by the 2D FEA at the pile head, the results are in accordance in general. Therefore, the simplified estimation method is considered an effective method for estimating the degree of deformation of the anchor piles under kinematic forces.

## V. CONCLUSIONS

This study investigated the effect of kinematic forces on the deformation of the anchor pile of a sheet pile quay wall. The main conclusions drawn from the current study are:

- The deformation of the anchor pile due to the kinematic forces is 30%–40% of the total deformation at the pile head, which is a non-negligible amount. Therefore, the low accuracy of the deformation obtained by the conventional method is attributed to the fact that it only considers inertial and not kinematic forces.
- Shear stress is generated in the ground between the sheet pile and the anchor pile before an earthquake due to the irregular shape of the quay wall. This initial shear stress results in large shear deformation of the ground during an earthquake, resulting in kinematic forces acting on the anchor pile.
- A simplified method was proposed for estimating the shear stresses generated in the ground around the sheet pile quay wall and the deformation of the anchor pile under kinematic forces, without reference to the 2D FEA results. It was demonstrated that the simplified estimation method effectively reproduces the 2D FEA results, but with a much lower computational effort. Therefore, this strategy is useful for evaluating the deformation of anchor piles under kinematic force.

## REFERENCES

- [1] *Technical standards and commentaries for port and harbour facilities in Japan*. Tokyo, Japan: The Overseas Coastal Area Development Institute of Japan, 2009.
- [2] *BS EN 1998-1(2004), Eurocode 8: Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings*. London, UK: British Standards Institution, 2004.
- [3] E. Galal, D. E. Mohamed, and E. Tolba, "A study of sheet pile quay wall rehabilitation methods," *Port-Said Engineering Research Journal*, vol. 26, no. 3, pp. 37–45, Sep. 2022, <https://doi.org/10.21608/psrj.2022.131369.1174>.
- [4] G. A. Athanasopoulos *et al.*, "Lateral spreading of ports in the 2014 Cephalonia, Greece, earthquakes," *Soil Dynamics and Earthquake Engineering*, vol. 128, Jan. 2020, Art. no. 105874, <https://doi.org/10.1016/j.soildyn.2019.105874>.
- [5] S. Werner *et al.*, "Seismic Performance of Port de Port-au-Prince during the Haiti Earthquake and Post-Earthquake Restoration of Cargo

- Throughput," *Earthquake Spectra*, vol. 27, no. 1, pp. 387–410, Oct. 2011, <https://doi.org/10.1193/1.3638716>.
- [6] T. Sugano, A. Nozu, E. Kohama, K. Shimosako, and Y. Kikuchi, "Damage to coastal structures," *Soils and Foundations*, vol. 54, no. 4, pp. 883–901, Aug. 2014, <https://doi.org/10.1016/j.sandf.2014.06.018>.
- [7] L. Wang, Z. Yajun, W. Gong, J. Li, and A. Ishimwe, "Calculation Method of Seismic Residual Displacement of Sheet Pile Quay Walls," *Advances in Civil Engineering*, vol. 2019, Aug. 2019, Art. no. e1251295, <https://doi.org/10.1155/2019/1251295>.
- [8] G. Gazetas, E. Garini, and A. Zafeirakos, "Seismic analysis of tall anchored sheet-pile walls," *Soil Dynamics and Earthquake Engineering*, vol. 91, pp. 209–221, Dec. 2016, <https://doi.org/10.1016/j.soildyn.2016.09.031>.
- [9] M. Mohajeri, Y. Kobayashi, K. Kawaguchi, and M. Sato, "Numerical Study on Lateral Spreading of Liquefied Ground Behind a Sheet Pile Model in a Large Scale Shake Table Test," in *13th World Conference on Earthquake Engineering*, Vancouver, BC, Canada, Aug. 2004, pp. 1–10.
- [10] U. Cilingir, S. K. Haigh, S. P. G. Madabhushi, and X. Zeng, "Seismic behaviour of anchored quay walls with dry backfill," *Geomechanics and Geoengeering*, vol. 6, no. 3, pp. 227–235, Sep. 2011, <https://doi.org/10.1080/17486025.2011.578670>.
- [11] A. Zekri, M. H. Aminfar, A. Ghalandarzadeh, and P. Ghasemi, "Effects of Liquefiable Layer on Dynamic Response of Anchored Flexible Quay Wall," in *10th International Congress on Advances in Civil Engineering*, Ankara, Turkey, Oct. 2012, pp. 1–10.
- [12] K. Ichii, S. Iai, Y. Sato, and H. Liu, "Seismic Performance Evaluation Charts for Gravity Type Quay Walls," *Structural Engineering / Earthquake Engineering*, vol. 19, no. 1, pp. 21s–31s, 2002, <https://doi.org/10.2208/jscseee.19.21s>.
- [13] P. Ghasemi, A. Ghalandarzadeh, A. Zekri, and M. Aminfar, "Mitigation of Seismic Deformation of Anchored Quay Wall by Compacting," in *7th International Conference on Case Histories in Geotechnical Engineering*, Chicago, IL, USA, Dec. 2013, vol. 7th International Conference on Case Histories in Geotechnical Engineering, pp. 1–5.
- [14] H. Tan, Z. Jiao, and J. Chen, "Field testing and numerical analysis on performance of anchored sheet pile quay wall with separate pile-supported platform," *Marine Structures*, vol. 58, pp. 382–398, Mar. 2018, <https://doi.org/10.1016/j.marstruc.2017.12.006>.
- [15] C. J. W. Habets, D. J. Peters, J. G. de Gijt, A. Metrikine, and S. N. Jonkman, "Model Solutions for Performance-Based Seismic Analysis of an Anchored Sheet Pile Quay Wall," *International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering*, vol. 10, no. 3, pp. 248–260, 2016.
- [16] M. F. R. Khazi and M. Vazeer, "FEM analysis of anchored sheet pile quay wall: A case study on the failure of WQ-7 berth of Visakhapatnam port," *International Journal of Research in Engineering and Technology*, vol. 5, no. 14, pp. 25–31, Sep. 2016, <https://doi.org/10.15623/ijret.2016.0526006>.
- [17] Y.-Y. Ko and H.-H. Yang, "Deriving seismic fragility curves for sheet-pile wharves using finite element analysis," *Soil Dynamics and Earthquake Engineering*, vol. 123, pp. 265–277, Aug. 2019, <https://doi.org/10.1016/j.soildyn.2019.05.014>.
- [18] T. Hamamoto, T. Nagano, and T. Yamao, "Analytical study of the sheet pile quay damaged by 2005 Fukuoka earthquake," in *Proceedings of 8th Annual IIIR Conference on Disaster Management*, Kumamoto, Japan, 2012, pp. 134–141.
- [19] K. Tokimatsu and Y. Asaka, "Effects of Liquefaction-Induced Ground Displacements on Pile Performance in the 1995 Hyogoken-Nambu Earthquake," *Soils and Foundations*, vol. 38, pp. 163–177, Sep. 1998, [https://doi.org/10.3208/sandf.38.Special\\_163](https://doi.org/10.3208/sandf.38.Special_163).
- [20] S. Nikolaou, G. Mylonakis, G. Gazetas, and T. Tazoh, "Kinematic pile bending during earthquakes: analysis and field measurements," *Geotechnique*, vol. 51, no. 5, pp. 425–440, Jun. 2001, <https://doi.org/10.1680/geot.2001.51.5.425>.
- [21] M. N. Hussien, T. Tobita, S. Iai, and M. Karay, "Soil-pile-structure kinematic and inertial interaction observed in geotechnical centrifuge experiments," *Soil Dynamics and Earthquake Engineering*, vol. 89, pp. 75–84, Oct. 2016, <https://doi.org/10.1016/j.soildyn.2016.08.002>.
- [22] K. Tokimatsu, H. Suzuki, and M. Sato, "Effects of inertial and kinematic interaction on seismic behavior of pile with embedded foundation," *Soil Dynamics and Earthquake Engineering*, vol. 25, no. 7, pp. 753–762, Aug. 2005, <https://doi.org/10.1016/j.soildyn.2004.11.018>.
- [23] R. W. Boulanger, C. J. Curras, B. L. Kutter, D. W. Wilson, and A. Abghari, "Seismic Soil-Pile-Structure Interaction Experiments and Analyses," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 125, no. 9, pp. 750–759, Sep. 1999, [https://doi.org/10.1061/\(asce\)1090-0241\(1999\)125:9\(750\)](https://doi.org/10.1061/(asce)1090-0241(1999)125:9(750)).
- [24] J.-S. Chiou, W.-Y. Hung, Y.-T. Lee, and Z.-H. Young, "Combined dynamic structure-pile-soil interaction analysis considering inertial and kinematic effects," *Computers and Geotechnics*, vol. 125, Sep. 2020, Art. no. 103671, <https://doi.org/10.1016/j.compgeo.2020.103671>.
- [25] N. Mononobe, "On determination of earth pressure during earthquake," in *World Engineering Congress*, Tokyo, Japan, 1929, vol. 9, pp. 177–185.
- [26] H. Matlock, "Correlation for Design of Laterally Loaded Piles in Soft Clay," in *2nd Annual Offshore Technology Conference*, Houston, TX, USA, Apr. 1970, <https://doi.org/10.4043/1204-MS>.
- [27] M. Georgiadis, C. Anagnostopoulos, and S. Safflekou, "Cyclic Lateral Loading of Piles in Soft Clay," *Journal of Geotechnical Engineering*, vol. 23, no. 1, pp. 47–60, Jun. 1992.
- [28] M. D. Dewaikar and S. D. Patil, "Behavior of Laterally Loaded Piles in Cohesionless Soil under One-Way Cyclic Loading," in *The New Millennium Conference*, 2001, pp. 97–100.
- [29] T. Nagao, "Effect of Foundation Width on Subgrade Reaction Modulus," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6253–6258, Oct. 2020, <https://doi.org/10.48084/etasr.3668>.
- [30] T. Nagao and R. Tsutaba, "Evaluation Methods of Vertical Subgrade Reaction Modulus and Rotational Resistance Moment for Seismic Design of Embedded Foundations," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7386–7392, Aug. 2021, <https://doi.org/10.48084/etasr.4269>.
- [31] S. Iai, Y. Matsunaga, and T. Kameoka, "Strain Space Plasticity Model for Cyclic Mobility," *Soils and Foundations*, vol. 32, no. 2, pp. 1–15, Jun. 1992, [https://doi.org/10.3208/sandf1972.32.2\\_1](https://doi.org/10.3208/sandf1972.32.2_1).
- [32] S. Iai and T. Kameoka, "Finite Element Analysis of Earthquake Induced Damage to Anchored Sheet Pile Quay Walls," *Soils and Foundations*, vol. 33, no. 1, pp. 71–91, Mar. 1993, <https://doi.org/10.3208/sandf1972.33.71>.
- [33] S. Higuchi *et al.*, "Evaluation of the seismic performance of dual anchored sheet pile wall," in *15th World Conference on Earthquake Engineering*, Lisbon, Portuga, Sep. 2012, pp. 1–10.
- [34] T. Nagao and Y. Kurachi, "An Experimental and Analytical Study on the Seismic Performance of Piers with Different Foundation Bottom Widths," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9142–9148, Oct. 2022, <https://doi.org/10.48084/etasr.5088>.
- [35] I. Towhata and K. Ishihara, "Modelling soil behavior under principal stress axes rotation," in *International Conference on Numerical Methods in Geomechanics*, Nagoya, Japan, 1985, pp. 523–530.
- [36] B. O. Hardin and V. P. Drnevich, "Shear Modulus and Damping in Soils: Design Equations and Curves," *Journal of the Soil Mechanics and Foundations Division*, vol. 98, no. 7, pp. 667–692, Jul. 1972, <https://doi.org/10.1061/JSFEAQ.0001760>.
- [37] G. Masing, "Eigenspannungen und Verfestigung beim Messing," in *Second International Congress of Applied Mechanics*, Zurich, Switzerland, 1926, pp. 332–335.
- [38] O. Ozutsumi and S. Iai, "Adjustment Method of the Hysteresis Damping for Multiple Shear Spring Model," in *4th International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*, San Diego, CA, USA, Mar. 2001.
- [39] I. Suetomi and N. Yoshida, "Nonlinear Behavior of Surface Deposit during the 1995 Hyogoken-Nambu Earthquake," *Soils and Foundations*, vol. 38, pp. 11–22, Sep. 1998, [https://doi.org/10.3208/sandf.38.Special\\_11](https://doi.org/10.3208/sandf.38.Special_11).
- [40] T. Morita, "Simplified method to determine parameter of FLIP," *Technical note of the Port and Harbour Research Institute*, vol. 869, pp. 1–36, 1997.

- [41] O. Ozutsumi, Y. Tamari, Y. Oka, K. Ichii, S. Iai, and Y. Umeki, "Modeling of soil-pile interaction subjected to soil liquefaction in plane strain analysis," in *53rd Japan National Conference on Geotechnical Engineering*, Takamatsu, Japan, 2003, pp. 1899-1900.
- [42] T. Nagao, M. Yamada, and A. Nozu, "Probabilistic Seismic Hazard Analysis with Focus on Fourier Amplitude and Group Delay Time," in *15th World Conference on Earthquake Engineering*, Lisbon, Portugal, Sep. 2012, pp. 1–10.
- [43] 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>.
- [44] 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>.
- [45] 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>.
- [46] C.-Y. Ku, J.-J. Jang, J.-Y. Lai, and M.-J. Hsieh, "Modeling of dynamic behavior for port structures using the performance-based seismic design," *Journal of Marine Science and Technology*, vol. 25, no. 6, pp. 732–741, Dec. 2017, <https://doi.org/10.6119/JMST-017-1226-14>.
- [47] J. H. Michell, "On the Direct Determination of Stress in an Elastic Solid, with application to the Theory of Plates," *Proceedings of the London Mathematical Society*, vol. s1-31, no. 1, pp. 100–124, 1899, <https://doi.org/10.1112/plms/s1-31.1.100>.

# Bio-Based Jet Fuel Production by Transesterification of Nettle Seeds

**Sinem Gurkan Aydin**

Department of Faculty of Applied Sciences, Aircraft Maintenance and Repair, Istanbul Gelisim University, Turkey  
sgurkan@gelisim.edu.tr  
(corresponding author)

**Arzu Ozgen**

Department of Medical Services and Techniques, Istanbul Gelisim University, Turkey  
aozgen@gelisim.edu.tr

*Received: 12 December 2022 | Revised: 23 December 2022 | Accepted: 5 January 2023*

## ABSTRACT

The use of petroleum-based fuels in air transport and the increase in oil prices over the years have increased fuel costs. Due to this increase, fuel manufacturers and airline companies have started to search for alternative fuels. Since aviation has an important place in the transportation sector, biomass has the greatest potential in the search for renewable energy sources. Biological substances of plant and animal origin and containing carbon compounds are energy sources, and the fuels produced from them are called biofuels. Biofuels are an important source of sustainable energy, which greatly reduces the greenhouse gas effect, improves weather conditions, reduces dependence on oil produced from fossil fuels, and is important for new markets. The nettle seed oil used in the current study was purchased from the local market and was obtained using the cold-pressing method at low temperatures. After the completion of the transesterification process, a two-phase mixture consisting of biofuel-glycerin was obtained, and the upper phase containing fatty acids was taken and transferred to a clean tube. After the final washing processes, bio jet fuel was obtained by adding chemicals at certain rates. The analysis of the obtained fuel was conducted at the Tubitak Marmara Research Centre. When the report was evaluated and compared with international standards, consistent results were obtained. It can be predicted that sustainable fuels can replace fossil fuels in the future.

**Keywords-**sustainable energy; bio-based jet fuel; transesterification; nettle seeds

## I. INTRODUCTION

Sustainable and renewable fuel production ways have an important place in terms of environmental, social, and economic aspects. Jet fuels constitute an important part of the relevant economy with a growing rapidly share [1]. When we look at the transportation sector in particular, it needs 2627.02 million tons of oil equivalent, which represents 27.9% of the total energy produced in the world [2]. In order to change the aircraft design and reduce fuel consumption, several attempts have been made to reduce CO<sub>2</sub> emissions, but these measures have yielded lower results than CO<sub>2</sub> emissions from the increase in flights [3, 4]. As a result, airline operators have sought different ways to solve this problem. Biofuels are a likely solution due to their renewable nature and low CO<sub>2</sub> emissions [4-6]. In addition, these fuels bring a new approach to the sector, increase safety, and reduce harmful emission particles [7-11]. Sustainable biofuels are important alternative fuels that require a large amount of raw material resources. These fuels can be obtained by means such as Fischer Tropsch synthesis [12-14] and alcohol synthesis [15-17], which have

become known ways of sustainable fuel production using intermediates such as alcohol and synthesis gas. The goal of minimizing the CO<sub>2</sub> emissions in the carrying sector has become a major driver for the advance of renewable fuels [18]. According to the estimations of the International Energy Agency, it is predicted that by 2050, sustainable biofuels will supply 27% of all energy in the field of transportation. For this reason, sustainable fuels have been the focus of attention [19, 20]. Traditional fossil-based jet fuels consist of approximately 20% paraffins, 40% isoparaffins, 20% naphthenes, and 20% aromatics [19]. This composition gives physical properties such as a freezing point of -47°C and 43.28MJ/kg energy [20].

Renewable bio jet energy sources consist of hydrocarbons in the boiling range of traditional fossil-based energy sources, showing that they have a structure close to the traditional fossil-based jet fuel. In the absence of aromatic compounds, harmful particles emitted from renewable bio jet fuels are lower than those emitted by fossil fuels [22].

The most accepted jet fuel standard is ASTM D1655-17 [23]. The limits of this standard allow biodiesel blending up to 50ppm, however, their use in air transport is still limited despite their advantages. Several hydrotreating routes have been proposed to manufacture renewable aviation fuel [24-30]. In addition, there have been new developments in the methods of obtaining rich aromatic and cycloparaffin hydrocarbons during the production of renewable jet fuels [31-33]. After that, the fuel standard ASTM D1655 can be limited as the high aromatic and/or cycloparaffin content in the fuel meets the density requirements, low heating value, and especially smoke point limits. To date, 5 different fuels that can be mixed with fossil-based jet fuel and used in gas turbine engines have been approved by the American Society for Testing Materials (ASTM). ASTM D7566-19 describes the detailed requirements and compositions [34].

The genus *Urtica*, a member of the Urticaceae family, is in the main group of Angiosperms (flowering plants) [35, 36]. The main varieties of the genus *Urtica*, known for their painful hairs on their leaves and stems, are *Urtica dioica* L., *Urtica urens* L., *Urtica pilulifera* L., *Urtica cannabina* L., *Urtica membranacea* Poir., and *Urtica kiovensis* Rogoff [38]. *Urtica* spp. is more commonly known as nettle. It grows all over the world, especially in regions with temperate climates. It prefers open or partially shaded habitats with plenty of moisture and is often found in forests, rivers or streams, and along roadsides. It is common in Europe and North America, North Africa, and parts of the Asia [35]. Nettle seeds are rich in oil, and nettle seed oil contains monounsaturated fatty acids, which is one of the reasons why that plant is preferred in biodiesel studies. The fatty acid content of nettle seed is palmitic acid, stearic acid, oleic acid, linoleic acid [37, 39].

## II. MATERIALS AND METHODS

Pure oil obtained from the nettle plant seeds was obtained. Methanol, ethanol, NaOH, fuel system anti-icing, kerosene, isooctane, nettle seed oil, and purified water were used. A single-stage basic reaction method, which is generally preferred for refined and crude vegetable oils, was used for transesterification. The flow chart of the procedure is given in Figure 3. 200ml methanol and 10g NaOH were used in appropriate proportions for every 1000ml of nettle seed oil. They were mixed with 200ml methanol in 10g NaOH for 30 minutes in a magnetic stirrer heater (Figure 1). The sodium methoxide ( $\text{CH}_3\text{ONa}$ ) obtained as a result of this process was kept in an oven at  $60^\circ\text{C}$  in order to maintain its temperature.

Then, the nettle seed oil was heated to  $70^\circ\text{C}$  using a magnetic heater and  $\text{CH}_3\text{ONa}$  was added to it, and the mixture was stirred at  $60\text{--}70^\circ\text{C}$  for 4 hours. The mixture was left at room temperature for 120 hours for the transesterification process to be completed. At the end of this process, a two-phase mixture consisting of biofuel-glycerin was obtained, and the upper phase containing fatty acids was carefully removed with the help of a pipettor and transferred to a clean tube (Figure 2). The fatty acid-glycerin boundary part was centrifuged at 1000 rpm for 5 minutes and the fatty acid part was recovered.

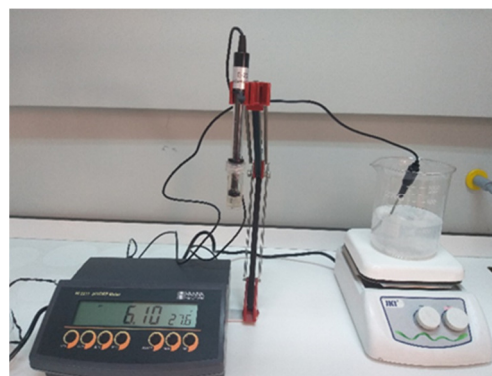


Fig. 1. Formation of  $\text{CH}_3\text{ONa}$  by a single-stage basic reaction method.

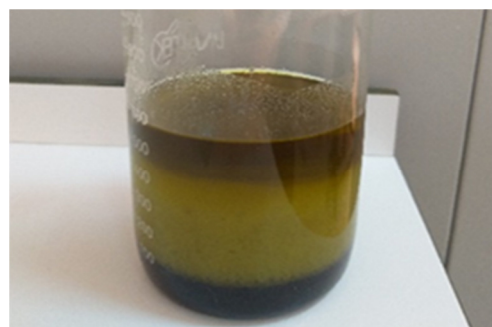


Fig. 2. Bi-phase product resulting from the transesterification process.

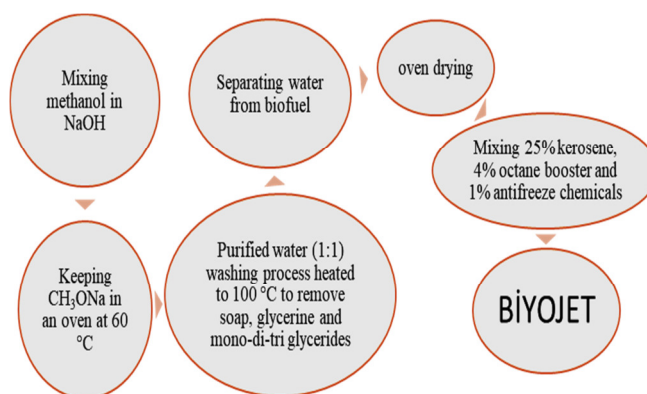


Fig. 3. Flowchart of the procedure.

In order to remove soap, glycerin and mono-di-tri glycerides that may be present in the biofuel, washing process was carried out 10 times by spraying on the biofuel using pure water (1:1) heated up to  $100^\circ\text{C}$ . After the last washing, the washing water was separated from the biofuel and settled to the bottom. The biofuel in the upper phase was carefully removed and dried in an oven at  $110\text{--}120^\circ\text{C}$ , and the remaining pure water and alcohol were completely removed. At the last stage, 3.5L 70% bio-based bio jet fuel were obtained by mixing 25% kerosene, 4% octane increaser, and 1% antifreeze chemicals and was sent to Tubitak Marmara Research Centre in order to determine its properties (Figure 4).





Fig. 4. Bio-based jet fuel obtained by transesterification of nettle seed

### III. RESULTS AND DISCUSSION

Generally, triglycerides, lignocelluloses, and syngas are used to produce renewable aviation fuels. Biofuels can be obtained from plants, animal fats, and biomass. Especially biofuels produced from plant and organic waste reduce CO<sub>2</sub> emissions, in addition to reducing the dependence on fossil fuels. There are also bacteria, yeast, and algae that have the capacity to produce fuel molecules and are used in the production of biofuels. The assortment can be enhanced by modifying the existing methods or by engineering applications and synthetic means. In such applications, genetic optimization is necessary, although the selection of organisms takes up a lot of space to increase efficiency. HEFA technology is a production method of sustainable fuels. It is conducted by hydroprocessing vegetable oils and animal fats. Approximately 1.2 tons of vegetable oil is required to produce 1 ton of HEFA fuel. One of the main advantages of this technology is to integrate this process into an oil refinery (with an additional step) and eliminate the need to develop a dedicated production facility. The HEFA production process is proven and approved for mixing ratios of up to 50%. In addition, current investments in infrastructure show that the process has an economically viable scope in the near future. In our study, a 70% biological-based mixture was tried and successful results were obtained.

Pure vegetable oil cannot be used as fuel in aircraft gas turbines. For this reason, its combustion characteristics should be approximated to diesel. Four techniques can be used for fuel regulation, aiming to reduce viscosity and eliminate atomization problems. These techniques are heating, dilution/mixing, microemulsion, and transesterification. The jet fuel sample produced with nettle seed oil was analyzed in the laboratory in accordance with the international standards. The characteristics of the fuel produced in our study and its comparison with the standards are given in Table I. As a result, it has been concluded that the characteristics of the fuel are compatible with the characteristics determined by international authorities. The distillation profile is made in accordance with the ASTM D 86 standard test. By evaporation and recondensation of 100ml sample, the temperatures at which 5, 10, 20, and 30 volume fractions were collected and determined, and a temperature-volume curve, i.e. a distillation profile, was obtained. A highly volatile fuel makes the engine easier to start,

but increases problems such as vapor plugging, icing, and fuel-scalding. In order to minimize these problems, it is necessary to balance the volatility characteristics of avgas. The distillation graph of the produced jet fuel is given in Figure 5.

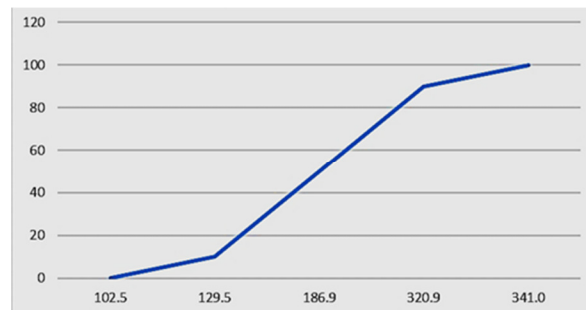


Fig. 5. Distillation graph of the bio-based jet fuel obtained with transesterification of nettle seeds.

Fuel instability is a process of multi-step oxidation reactions that occur between some of the compounds it contains. The initial reaction products are hydroperoxides and peroxides. These products are dissolved in the fuel, but they affect the elastomeric materials of the fuel system and shorten its useful life. In addition, soluble gum and insoluble particles are formed with the ongoing reactions. These substances clog fuel filters and cause residue to build up on fuel system walls, making it difficult for fuel to flow. The existing gum and potential gum values in our product are in the range of values that are quite suitable for use. The bio jet fuels studied before include 50% sustainability, while the product in our study has 70% sustainability [38]. We have previously discussed the fuel additives used with a detailed unit analysis [39]. Factors that act as drivers for switching from one fuel to another to improve energy efficiency have been identified in the literature, including various performance parameters that support environmental protection [40]. Biofuel has been previously produced using cheap waste cooking oil collected from Pakistan's Nawabshah local market [41]. The fuel produced in our study will provide a very good approach in the aviation industry.

### IV. CONCLUSION

Air transport is a feature of our modern, globalized world that connects people, commerce, and most importantly, businesses across continents. The benefits and importance of air travel are undisputed, but there are also important aspects to consider. Traditionally, environmental problems involving aviation have focused on noise and air pollution and the recent issue of global climate change has focused the attention on the CO<sub>2</sub> emission volumes of airplanes. When greenhouse gases produced from fuels burned in flights are emitted into the atmosphere, they have a significant negative impact on the environment. Aviation is likely to need around 450-500 million tons of sustainable aviation fuel per year by 2050. Jet fuels produced using renewable bio-based are important for the aviation industry and reduce its dependency on fossil fuels. It is an important breakthrough, as it will contribute to the targets of reducing emissions against global problems.



TABLE I. ANALYSIS RESULTS OF BIO-BASED JET FUEL OBTAINED BY TRANSESTERIFICATION OF NETTLE SEEDS

Parameter	Jet a	Jet b	Analysis result	Analysis method
Density (15°C)	775-840	751-802	803.5 kg/m <sup>3</sup>	ASTMD 4052
Distillation				ASTMD 86
Initial boiling point	151.5	152.7	102.5 °C	
%10	205		129.5 °C	
%50	200	190	186.9 °C	
%90	246.7	245	320.9 °C	
Ultimate boiling point	300		341.0 °C	
Distillation residue	1.5	1.5	0.8%v/v	
Distillation loss	1.5	1.5	0.2%v/v	ASTMD 3228
Flash point	38	-	16.5 °C	
Color (saybond)	-	-	-15	ASTMD 6045
Electrical conductivity	500-600	-	504 pS/m	ASTMD 2624
Copperstrip corrosion 2h100 °C	1b	-	1b	ASTMD 130
Mercaptan sulfur	0.003	--	0.0004% m/m	ASTMD 3227
Acid number	0.1		<0.1 mgKOH/g	ASTMD 664
Current gum	2000	2000	2254mg/100mL	ASTMD 381
Potential gum	-	--	3454 mg/100mL	ASTMD873
Sulfur	-	-	0.064% m/m	ASTMD 4294
Kinematic viscosity	8	-	3.948 mm <sup>2</sup> /s	ASTMD 445
Doctor test	-	-	NEGATIF	TS 2884
Freezing point	-40. -47	-50	-31 °C	ASTMD 2386
Smoke point	25	25	26 mm	ASTMD 1322
Heat of combustion	42.8	42.8	43.3MJ/kg	ASTMD 3338
FIA aromatic	15-25	25	17.7%v/v	ASTMD 1319
FIA olefin	%0.2- %35	-	0.9%v/v	
FIA saturated	-	-	81.4%v/v	
Thermal stability				
Control temperature 260 °C min	450	450	450mL	ASTMD 3241
Pressure difference	-	-	252mmHg	
Heating pipe sediment class	1	1	<3	
Water reaction	3	3	4 mL	ASTMD 1094
Aniline point			54	ASTMD 611-A
Octane number RON	0-100	0-100	50.1	ASTMD 2699
Octane number MON	0-100	0-100	48.0	ASTMD 2700

The current project deals with aspects such as global warming, sustainability, and reduction of carbon emissions. Unlike the previously produced planetary fuels, the proposed fuel is not separated due to its high efficiency and 70% biological nature. In conclusion, the biojet fuel under study is produced using a sustainable resource, nettle seed oil, and is an excellent candidate product, with the highest biological rate.

#### ACKNOWLEDGMENT

The authors wish to thank the Istanbul Gelisim University Scientific Research Projects Application Center for funding this work. Project Number: DUP-100920-SGA, Year 2021. The authors wish to thank the Istanbul Gelisim University Scientific Research Projects Application Center.

#### REFERENCES

- [1] "Jet Fuel Price Monitor," IATA. <https://www.iata.org/en/publications/economics/fuel-monitor>.
- [2] IEA, *Key World Energy Statistics*. Paris, France: International Energy Agency, 2016.
- [3] "Environmental Trends in Aviation to 2050," in *ICAO Environmental Report, Aviation and Climate Change*, ICAO, 2016, pp. 16–22.
- [4] *Biofuels for Aviation: Technology Brief*. IRENA, 2017.
- [5] *IATA Economic Briefing: Airline Fuel and Labour Cost Share*. IATA, 2010.
- [6] *IATA 2015 Report on Alternative Fuels*. IATA, 2015.
- [7] C. D. Klingshirn *et al.*, "Hydroprocessed Renewable Jet Fuel Evaluation, Performance, and Emissions in a T63 Turbine Engine," *Journal of Engineering for Gas Turbines and Power*, vol. 134, no. 5, Mar. 2012, Art. no. 051506, <https://doi.org/10.1115/1.4004841>.
- [8] T. Rahmes *et al.*, "Sustainable Bio-Derived Synthetic Paraffinic Kerosene (Bio-SPK) Jet Fuel Flights and Engine Tests Program Results," in *9th AIAA Aviation Technology, Integration, and Operations Conference*, Hilton Head, SC, USA, Sep. 2009, <https://doi.org/10.2514/6.2009-7002>.
- [9] S. Blakey, L. Rye, and C. W. Wilson, "Aviation gas turbine alternative fuels: A review," *Proceedings of the Combustion Institute*, vol. 33, no. 2, pp. 2863–2885, Jan. 2011, <https://doi.org/10.1016/j.proci.2010.09.011>.
- [10] R. L. Speth, C. Rojo, R. Malina, and S. R. H. Barrett, "Black carbon emissions reductions from combustion of alternative jet fuels," *Atmospheric Environment*, vol. 105, pp. 37–42, Mar. 2015, <https://doi.org/10.1016/j.atmosenv.2015.01.040>.
- [11] M. Badami, P. Nuccio, D. Pastrone, and A. Signoretto, "Performance of a small-scale turbojet engine fed with traditional and alternative fuels," *Energy Conversion and Management*, vol. 82, pp. 219–228, Jun. 2014, <https://doi.org/10.1016/j.enconman.2014.03.026>.
- [12] J. Li, G. Yang, Y. Yoneyama, T. Vitidsant, and N. Tsubaki, "Jet fuel synthesis via Fischer–Tropsch synthesis with varied 1-olefins as additives using Co/ZrO<sub>2</sub>–SiO<sub>2</sub> bimodal catalyst," *Fuel*, vol. 171, pp. 159–166, May 2016, <https://doi.org/10.1016/j.fuel.2015.12.062>.
- [13] T. Hanaoka, T. Miyazawa, K. Shimura, and S. Hirata, "Jet fuel synthesis in hydrocracking of Fischer–Tropsch product over Pt-loaded zeolite catalysts prepared using microemulsions," *Fuel Processing Technology*, vol. 129, pp. 139–146, Jan. 2015, <https://doi.org/10.1016/j.fuproc.2014.09.011>.

- [14] T. Hanaoka, T. Miyazawa, K. Shimura, and S. Hirata, "Jet fuel synthesis from Fischer–Tropsch product under mild hydrocracking conditions using Pt-loaded catalysts," *Chemical Engineering Journal*, vol. 263, pp. 178–185, Mar. 2015, <https://doi.org/10.1016/j.cej.2014.11.042>.
- [15] M. He, M. Wang, G. Tang, Y. Fang, and T. Tan, "From medium chain fatty alcohol to jet fuel: Rational integration of selective dehydration and hydro-processing," *Applied Catalysis A: General*, vol. 550, pp. 160–167, Jan. 2018, <https://doi.org/10.1016/j.apcata.2017.11.009>.
- [16] G. Nie, X. Zhang, L. Pan, M. Wang, and J.-J. Zou, "One-pot production of branched decalins as high-density jet fuel from monocyclic alkanes and alcohols," *Chemical Engineering Science*, vol. 180, pp. 64–69, Apr. 2018, <https://doi.org/10.1016/j.ces.2018.01.024>.
- [17] K. P. Brooks *et al.*, "Chapter 6 - Low-Carbon Aviation Fuel Through the Alcohol to Jet Pathway," in *Biofuels for Aviation*, C. J. Chuck, Ed. New York, NY, USA: Academic Press, 2016, pp. 109–150.
- [18] "Technology Roadmap - Biofuels for Transport – Analysis," IEA. <https://www.iea.org/reports/technology-roadmap-biofuels-for-transport>.
- [19] M. Bernabei, R. Reda, R. Galiero, and G. Bocchinfuso, "Determination of total and polycyclic aromatic hydrocarbons in aviation jet fuel," *Journal of Chromatography A*, vol. 985, no. 1, pp. 197–203, Jan. 2003, [https://doi.org/10.1016/S0021-9673\(02\)01826-5](https://doi.org/10.1016/S0021-9673(02)01826-5).
- [20] A. Agosta, "Development of a chemical surrogate for JP-8 aviation fuel using a pressurized flow reactor," M.S. thesis, Drexel University, Philadelphia, PA, USA, 2002.
- [21] *Aviation Fuels: Technical Review*. San Ramon, CA, USA: Chevron, 2007.
- [22] J. Holmgren, "Biofuels: Unlocking the Potential," in *The International Conference on Biorefinery*, Syracuse, NY, USA, Oct. 2009, pp. 1–44.
- [23] *ASTM D1655-20(2020), Standard Specification for Aviation Turbine Fuels*. West Conshohocken, PA, USA: ASTM International, 2020.
- [24] K. K. Gupta, A. Rehman, and R. M. Sarviya, "Bio-fuels for the gas turbine: A review," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 9, pp. 2946–2955, Dec. 2010, <https://doi.org/10.1016/j.rser.2010.07.025>.
- [25] G. Liu, B. Yan, and G. Chen, "Technical review on jet fuel production," *Renewable and Sustainable Energy Reviews*, vol. 25, pp. 59–70, Sep. 2013, <https://doi.org/10.1016/j.rser.2013.03.025>.
- [26] D. Chiamonti, M. Prussi, M. Buffi, and D. Tacconi, "Sustainable bio kerosene: Process routes and industrial demonstration activities in aviation biofuels," *Applied Energy*, vol. 136, pp. 767–774, Dec. 2014, <https://doi.org/10.1016/j.apenergy.2014.08.065>.
- [27] U. Neuling and M. Kaltschmitt, "Conversion routes for production of biokerosene—status and assessment," *Biomass Conversion and Biorefinery*, vol. 5, no. 4, pp. 367–385, Dec. 2015, <https://doi.org/10.1007/s13399-014-0154-2>.
- [28] M. Mohammad, T. Kandaramath Hari, Z. Yaakob, Y. Chandra Sharma, and K. Sopian, "Overview on the production of paraffin based-biofuels via catalytic hydrodeoxygenation," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 121–132, Jun. 2013, <https://doi.org/10.1016/j.rser.2013.01.026>.
- [29] W.-C. Wang and L. Tao, "Bio-jet fuel conversion technologies," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 801–822, Jan. 2016, <https://doi.org/10.1016/j.rser.2015.09.016>.
- [30] J. Yang, Z. Xin, Q. (Sophia) He, K. Corscadden, and H. Niu, "An overview on performance characteristics of bio-jet fuels," *Fuel*, vol. 237, pp. 916–936, Feb. 2019, <https://doi.org/10.1016/j.fuel.2018.10.079>.
- [31] T. Wang *et al.*, "Aviation fuel synthesis by catalytic conversion of biomass hydrolysate in aqueous phase," *Applied Energy*, vol. 136, pp. 775–780, Dec. 2014, <https://doi.org/10.1016/j.apenergy.2014.06.035>.
- [32] J. Fu, C. Yang, J. Wu, J. Zhuang, Z. Hou, and X. Lu, "Direct production of aviation fuels from microalgae lipids in water," *Fuel*, vol. 139, pp. 678–683, Jan. 2015, <https://doi.org/10.1016/j.fuel.2014.09.025>.
- [33] Y. Zhang *et al.*, "Production of jet and diesel biofuels from renewable lignocellulosic biomass," *Applied Energy*, vol. 150, pp. 128–137, Jul. 2015, <https://doi.org/10.1016/j.apenergy.2015.04.023>.
- [34] Q. Liu, C. Zhang, N. Shi, X. Zhang, C. Wang, and L. Ma, "Production of renewable long-chained cycloalkanes from biomass-derived furfurals and cyclic ketones," *RSC Advances*, vol. 8, no. 25, pp. 13686–13696, 2018, <https://doi.org/10.1039/C8RA01723A>.
- [35] "The Plant List: A Working List of All Plant Species." <http://www.theplantlist.org>.
- [36] R. Upton, "Stinging nettles leaf (*Urtica dioica* L.): Extraordinary vegetable medicine," *Journal of Herbal Medicine*, vol. 3, no. 1, pp. 9–38, Mar. 2013, <https://doi.org/10.1016/j.hermed.2012.11.001>.
- [37] Z. Jafari, S. A. Samani, and M. Jafari, "Insights into the bioactive compounds and physico-chemical characteristics of the extracted oils from *Urtica dioica* and *Urtica pilulifera*," *SN Applied Sciences*, vol. 2, no. 3, Feb. 2020, Art. no. 416, <https://doi.org/10.1007/s42452-020-2219-0>.
- [38] S. Gurkan Aydin and A. Ozgen, "Sustainable Jet Fuel Production: Using Pumpkin Seed Oil," *TEM Journal*, vol. 10, no. 2, pp. 879–882, 2021.
- [39] S. G. Aydin, O. Polat, A. Ozgen, and E. Turali, "Calculated Optimized Structure and Geometric Analysis of Oxygenated Fuel Additives: Alcohols and Ethers," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5632–5636, Jun. 2020, <https://doi.org/10.48084/etasr.3491>.
- [40] 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>.
- [41] I. Naim and T. Mahara, "Fuel Substitution for Energy Saving: A Case Study of Foundry Plant," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3439–3444, Oct. 2018, <https://doi.org/10.48084/etasr.2298>.

# The Improved CURLI Method for Multi-Criteria Decision Making

Anh-Tu Nguyen

Faculty of Mechanical Engineering, Hanoi University of Industry, Vietnam

tuna@hau.edu.vn

(corresponding author)

Received: 5 December 2022 | Revised: 26 December 2022 | Accepted: 31 December 2022

## ABSTRACT

Multi-Criteria Decision Making (MCDM) investigates the best available choice in the presence of multiple conflicting criteria, whereas the Collaborative Unbiased Rank List Integration (CURLI) method has been proposed recently and has been applied in various fields of daily life. However, most previous works concentrated on analyzing cases in which the factor of a criterion is a specific quantity. The present paper proposes an approach developed from the original CURLI method, named Improved CURLI. This improvement helps solve a problem when the factors of the criteria can be linguistic variables or a data set. The proposed method is applied to rank the alternatives for two case studies: choosing the best grinding wheel and the best service suppliers. The ranking results are compared to those obtained using other methods. Furthermore, sensitivity analysis is also conducted to examine the stability and reliability of the ranking results in various scenarios. The results demonstrate the validity of the Improved CURLI method and prove that it is applicable for making decisions in various fields.

**Keywords-***MCDM; CURLI method; improved CURLI method; data set*

## I. INTRODUCTION

The concept of MCDM is increasingly being used in various fields [1, 2]. MCDM problems are mentioned as MCDA or MADM. The essence of MCDM is evaluating and ranking the available options in order to select the best and avoid the worst. Researchers have suggested different MCDM approaches, and many of them have been applied in a variety of contexts [1, 3]. However, the original versions of the MCDM approaches are deemed inadequate for circumstances in which the criteria are expressed as a set of factors or linguistic variables [4, 5]. The main reason for this phenomenon is that the uncertainty about the object depends on a lot of factors, like experts' opinions, time, location, the way of data collection, etc. [6]. To deal with this issue, approaches that combine the fuzzy method with MCDM have been proposed. That combination is referred to as fuzzy - MCDM. Many studies are progressively employing the fuzzy - MCDM methods. For instance, TOPSIS integrated with fuzzy is used in many applications such as supplier selection [7], project selection [8], healthcare software [9], supply chain management [10], manager selection [11], etc. In [12], a fuzzy VIKOR - MCDM is proposed to analyze and evaluate the quality of information security policies as well as the content of press agencies in Gulf countries. Realizing this method's exploitation potential, numerous investigations have been launched in different tasks like warehouse location selection [13], mobile services [14], and sustainable development in Islamic countries [15]. In other approaches, the fuzzy MARCOS method is utilized for road traffic risk analysis [16],

assessment of drone-based city logistics [17], e-service quality in airline industry [18], and so on.

When applying fuzzy MCDM, it is necessary to determine the weights of the criteria which play an important role to make the final decision [19-21]. In this manner, calculating the weights for the criteria is relatively complex, time-consuming, and difficult in situations requiring prompt decision-making. Furthermore, the ambiguity of the decision makers (or surveyed experts) regarding the research object influences the weights determined by these methods. Hence, the accuracy of the weights after the calculation is not guaranteed. As a result, no longer will the ranking of the options be certain [22-24]. In addition, most fuzzy MCDM methods still inherit some stages from the original MCDM methods, one of which is data normalization. But the data normalization processes of the MCDM methods are also not the same, leading to the ranking results of the alternatives also being highly dependent on the data normalization and sensitive to change [25-27].

Generally, despite the widespread application of fuzzy MCDM approaches, they still have limitations to some extent. If a suggested MCDM method can solve two major problems, it will make significant contributions to this field. Due to the mentioned limitations associated with the right determination for the criteria and the normalization of data, it is necessary to propose an MCDM approach that does not require the determination of weights for the criteria and the metric normalization. To overcome these issues, the CURLI method was first introduced in 2016. It is used to rank applicants for medical programs [28]. Despite the fact that it has existed for

six years, only a handful of research projects about this technique have been carried out. The applying CURLI method is proposed to inspect the quality of the X12 steel grinding process [29]. A turning test progress based on the CURLI approach was designed in [30]. The result reveals that the proposed method is as precise as the PEG method and better than the PSI method. In addition, recent works have shown that the CURLI method is just as accurate as the R and CODAS approaches to ranking robots, just as accurate as the R, SAW, WASPAS, TOPSIS, VIKOR, MOORA, COPRAS, and PIV approach to rating the turning process, and just as accurate as the R and MABAC approaches to ranking bridge construction [31]. However, all the above mentioned studies are only considered when the criteria are clearly defined (not fuzzy set). In this paper, an approach based on the original CURLI method, which allows overcoming the mentioned drawbacks is proposed. The key contributions of the current paper are:

- This paper proposes an improved CURLI method to allow solving the MCDM problems, when the factors of the criteria can be linguistic variables or a data set.
- The proposed method does not need the input data to be normalized or the weights of the criteria to be evaluated.
- The proposed approach is the first step towards enhancing the CURLI method to optimize MCDM problems as well as applying it to various fields of daily life.

## II. THE IMPROVED CURLI METHOD

The original CURLI method involves four steps as follows [28-31]:

**Step 1:** Establish a decision-making matrix with  $m$  options and  $n$  criteria as in Table I.  $C_{ij}$  is the factor of criterion  $j^{th}$  in option  $i^{th}$ , where  $i = 1 \div m, j = 1 \div n$ .

TABLE I. DECISION-SCORE MATRIX

Alternatives	Criteria			
	$C_1$	$C_2$	$C_1$	$C_n$
$A_1$	$C_{11}$	$C_{12}$	$C_{1j}$	$C_{1n}$
$A_2$	$C_{21}$	$C_{22}$	$C_{2j}$	$C_{2n}$
$A_i$	$C_{i1}$	$C_{i2}$	$C_{ij}$	$C_{in}$
$A_m$	$C_{m1}$	$C_{m2}$	$C_{mj}$	$C_{mn}$

**Step 2:** Create score-point matrices of level  $m$ . In criterion  $j$ , the entry corresponding to row  $t$  and column  $v$  ( $1 \leq t, v \leq m$ ) is determined based on the following regulations:

- If the factor of criterion  $j$  in alternative  $A_t$  is worse than that of alternative  $A_v$ , the entry in the corresponding row and column will be scored as -1.
- If the factor of criterion  $j$  in alternative  $A_t$  is better than that of alternative  $A_v$ , the entry in the corresponding row and column will be scored as 1.
- If the factor of criterion  $j$  in alternative  $A_t$  is equal to that of alternative  $A_v$ , the entry in the corresponding row and column will be scored as 0.

All the entries on the main diagonal will be blank. After this step,  $n$  square score-point matrices are established in total.

**Step 3:** Create the process score-point matrix. The entries of this matrix are calculated by summing all the corresponding entries of the score-point matrix in step 2.

**Step 4:** Rearrange the process score-point matrix. The arrangement will be performed by moving the rows and columns of the process score-point matrices by the following rules: the order of columns from the left to the right corresponding to the order of the rows from the top to the bottom. The number of the negative and zero entries above the main diagonal is maximal. After sorting, the solution in the first row will be the best choice. Priority of selection reduces from the first to the last row.

The CURLI approach has been used extensively to solve optimization problems in several disciplines, in which the factors of each criterion at an alternative are a specific quantity. However, in reality, a criterion may be represented by linguistic variables or a data set. This paper proposes an approach based on the original CURLI method for handling these issues as follows:

**Step 1:** Establish a decision-making matrix as in Table II.

**Step 2:** Create score-point matrices. This step is implemented similarly to the original CURLI. However, the score-point matrices will be determined with every factor of the criteria. The score point of each criterion is denoted as  $p^f$  ( $1 \leq f \leq k$ ), where  $k$  is the number of factors of a criterion.

TABLE II. DECISION-MAKING MATRIX WITH MULTI-CHOICE OF CRITERIA IN EACH SOLUTION

Alternatives	Criteria			
	$C_1$	$C_2$	$C_1$	$C_n$
$A_1$	$(C_{11}^1, C_{11}^2, \dots, C_{11}^k)$	$(C_{12}^1, C_{12}^2, \dots, C_{12}^k)$	$(C_{1j}^1, C_{1j}^2, \dots, C_{1j}^k)$	$(C_{1n}^1, C_{1n}^2, \dots, C_{1n}^k)$
$A_2$	$(C_{21}^1, C_{21}^2, \dots, C_{21}^k)$	$(C_{22}^1, C_{22}^2, \dots, C_{22}^k)$	$(C_{2j}^1, C_{2j}^2, \dots, C_{2j}^k)$	$(C_{2n}^1, C_{2n}^2, \dots, C_{2n}^k)$
$A_i$	$(C_{i1}^1, C_{i1}^2, \dots, C_{i1}^k)$	$(C_{i2}^1, C_{i2}^2, \dots, C_{i2}^k)$	$(C_{ij}^1, C_{ij}^2, \dots, C_{ij}^k)$	$(C_{in}^1, C_{in}^2, \dots, C_{in}^k)$
$A_m$	$(C_{m1}^1, C_{m1}^2, \dots, C_{m1}^k)$	$(C_{m2}^1, C_{m2}^2, \dots, C_{m2}^k)$	$(C_{mj}^1, C_{mj}^2, \dots, C_{mj}^k)$	$(C_{mn}^1, C_{mn}^2, \dots, C_{mn}^k)$

**Step 3:** Build the process score-point matrix. The entries of this matrix are calculated based on the sum of all points in corresponding marking point matrices. The detailed formulation is expressed as:

$$p_j = \frac{\sum_{f=1}^k p_j^f}{k} \quad (1)$$

**Step 4:** Rearrange the process score-point matrix. This step is the same as in the original CURLI method.

## III. VERIFICATION STUDY AND DISCUSSION

### A. Case Study 1

In this sub-section, the proposed method is applied to determine which grinding wheel to choose in an MCDM problem [32, 33]. There are 8 different types of grinding wheels ( $A_1 \div A_8$ ), and each wheel is described by 7 criteria ( $C_1 \div C_7$ ). Each criterion has 3 alternatives. The decision is

made to satisfy the conditions: (a) for C7 the smallest is the best and (b) for the other criteria, the biggest is the best. The objective of MCDM is to identify the alternative that simultaneously assures that C7 is the smallest and the remaining criteria (from C1 to C6) are maximal. This work has also been accomplished by employing the Fuzzy TOPSIS method [33] and 6 variations of the VIKOR method [32]. The results of rating the alternatives using these two ways will be compared to the results of ranking the alternatives using the proposed method in this study.

In the first step, the decision-making matrix is composed as in Table III. In the second step, scoring for the criteria will be performed. In this case, the value of each criterion at each alternative has three levels of values. The number of

alternatives that need to be ranked is 8. Thus, the score for each alternative (for each criterion) is a set of numbers as shown in (2):

$$P_j^f = \{P_j^1, P_j^2, P_j^3\} \text{ with } \{j \in \mathbf{N} | 1 \leq j \leq 8\} \quad (2)$$

After the scoring process, the decision-score matrixes are illustrated from Table IV to Table X. In the next step, the process score-point matrix is determined based on (1) and is presented in Table XI. The process score-point matrix is rearranged based on the mentioned rules in Section 2. Following the arranging procedure, the alternative in the first row could be seen as the best choice. The final arrangement is shown in Table XII.

TABLE III. DECISION-SCORE MATRIX FOR CHOOSING A GRINDING WHEEL [32, 33]

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
A1	(2700, 3200, 3700)	(391, 451, 511)	(2925, 3475, 4025)	(581, 756, 931)	(12, 17, 22)	(2.65, 4.15, 5.65)	(12, 18, 24)
A2	(2000, 2400, 2800)	(590, 690, 790)	(4275, 4975, 5675)	(1099, 1324, 1549)	(68, 98, 128)	(2.2, 3, 3.8)	(45, 60, 75)
A3	(4400, 5000, 5600)	(725, 850, 975)	(6000, 6900, 7800)	(1282, 1532, 1782)	(9, 13, 17)	(3.55, 4.5, 5.45)	(714, 864, 1014)
A4	(2600, 3000, 3400)	(350, 400, 450)	(3200, 3800, 4400)	(729, 879, 1029)	(21, 30, 39)	(3.15, 4, 4.85)	(107, 152, 197)
A5	(7300, 8000, 8700)	(818, 953, 1088)	(5900, 6700, 7500)	(4188, 4688, 5188)	(950, 1200, 1450)	(6.45, 8.6, 10.8)	(1050, 1300, 1550)
A6	(2150, 2550, 2950)	(370, 440, 510)	(3950, 4600, 5250)	(400, 480, 560)	(150, 200, 250)	(2.6, 3.1, 3.6)	(6.5, 10, 13.5)
A7	(2400, 2800, 3200)	(385, 460, 535)	(1421, 1721, 2021)	(425, 600, 775)	(55, 90, 125)	(1.95, 2.5, 3.05)	(36, 50, 64)
A8	(900, 1200, 1500)	(115, 160, 205)	(1350, 1750, 2150)	(495, 620, 745)	(1.4, 2.2, 3)	(5.75, 8.2, 10.7)	(33, 45, 57)

TABLE IV. SCORE-POINT MATRIX FOR CRITERION C1

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		-1, -1, -1	1, 1, 1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A2	1, 1, 1		1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1
A3	-1, -1, -1	-1, -1, -1		-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A4	1, 1, 1	-1, -1, -1	1, 1, 1		1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A5	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1
A6	1, 1, 1	-1, -1, -1	1, 1, 1	1, 1, 1	1, 1, 1		1, 1, 1	-1, -1, -1
A7	1, 1, 1	-1, -1, -1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1		-1, -1, -1
A8	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	

TABLE V. SCORE-POINT MATRIX FOR CRITERION C2

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		1, 1, 1	1, 1, 1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A2	-1, -1, -1		1, 1, 1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A3	-1, -1, -1	-1, -1, -1		-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A4	1, 1, 1	1, 1, 1	1, 1, 1		1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1
A5	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1
A6	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1	1, 1, 1		1, 1, 1	-1, -1, -1
A7	1, -1, -1	1, 1, 1	1, 1, 1	-1, -1, -1	1, 1, 1	-1, -1, -1		-1, -1, -1
A8	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	

TABLE VI. SCORE-POINT MATRIX FOR CRITERION C3

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1	-1, -1, -1
A2	-1, -1, -1		1, 1, 1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A3	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A4	-1, -1, -1	1, 1, 1	1, 1, 1		1, 1, 1	1, 1, 1	-1, -1, -1	-1, -1, -1
A5	-1, -1, -1	-1, -1, -1	1, 1, 1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1
A6	-1, -1, -1	1, 1, 1	1, 1, 1	-1, -1, -1	1, 1, 1		-1, -1, -1	-1, -1, -1
A7	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1		-1, 1, 1
A8	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, -1, -1	

TABLE VII. SCORE-POINT MATRIX FOR CRITERION C4

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A2	-1, -1, -1		1, 1, 1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A3	-1, -1, -1	-1, -1, -1		-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A4	-1, -1, -1	1, 1, 1	1, 1, 1		1, 1, 1	-1, -1, -1	-1, -1, -1	-1, -1, -1
A5	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1
A6	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1		1, 1, 1	1, 1, 1
A7	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1		1, 1, -1
A8	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1	-1, -1, 1	

TABLE VIII. SCORE-POINT MATRIX FOR CRITERION C5

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		1, 1, 1	-1, -1, -1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1
A2	-1, -1, -1		-1, -1, -1	-1, -1, -1	1, 1, 1	1, 1, 1	-1, -1, -1	-1, -1, -1
A3	1, 1, 1	1, 1, 1		1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1
A4	-1, -1, -1	1, 1, 1	-1, -1, -1		1, 1, 1	1, 1, 1	1, 1, 1	-1, -1, -1
A5	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1
A6	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1	1, 1, 1		-1, -1, -1	-1, -1, -1
A7	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	1, 1, 1	1, 1, 1		-1, -1, -1
A8	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	

TABLE IX. SCORE-POINT MATRIX FOR CRITERION C6

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		-1, -1, -1	1, 1, -1	1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	1, 1, 1
A2	1, 1,		1, 1, 1	1, 1, 1	1, 1, 1	1, 1, -1	-1, -1, -1	1, 1, 1
A3	-1, -1, 1	-1, -1, -1		-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	1, 1, 1
A4	-1, 1, 1	-1, -1, -1	1, 1, 1		1, 1, 1	-1, -1, -1	-1, -1, -1	1, 1, 1
A5	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1	-1, -1, -1
A6	1, 1, 1	-1, -1, 1	1, 1, 1	1, 1, 1	1, 1, 1		-1, -1, -1	1, 1, 1
A7	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1		1, 1, 1
A8	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1	

TABLE X. SCORE-POINT MATRIX FOR CRITERION C7

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1	1, 1, 1	-1, -1, -1	-1, -1, -1
A2	1, 1, 1		-1, -1, -1	-1, -1, -1	-1, -1, -1	1, 1, 1	1, 1, 1	1, 1, 1
A3	1, 1, 1	1, 1, 1		1, 1, 1	-1, -1, -1	1, 1, 1	1, 1, 1	1, 1, 1
A4	1, 1, 1	1, 1, 1	-1, -1, -1		-1, -1, -1	1, 1, 1	1, 1, 1	1, 1, 1
A5	1, 1, 1	1, 1, 1	1, 1, 1	1, 1, 1		1, 1, 1	1, 1, 1	1, 1, 1
A6	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1	-1, -1, -1		-1, -1, -1	-1, -1, -1
A7	1, 1, 1	-1, -1, -1	-1	-1, -1, -1	-1, -1, -1	1, 1, 1		1, 1, 1
A8	1, 1, 1	-1, -1, -1	-1	-1, -1, -1	-1, -1, -1	1, 1, 1	-1, -1, -1	

TABLE XI. SCORE-POINT MATRIX FOR THE EVALUATION PROCESS

Alternatives	Points							
	P1	P2	P3	P4	P5	P6	P7	P8
A1		1	2.3333	-0.333	5	-1	-3.6667	-5
A2	-1		3	-3	5	0.3333	-3	-3
A3	-2.3333	-3		-3	3	-3	-3	-3
A4	0.3333	3	3		5	1	-1	-3
A5	-5	-5	-3	-5		-5	-5	-5
A6	1	-0.3333	3	-1	5		-1	-3
A7	3.6667	3	3	1	5	1		-0.3333
A8	5	3	3	3	5	3	0.3333	

Table XII indicates that all the entries above the principal diagonal are negative, therefore A5 is the best choice and A8 is the worst one. The rank of alternatives is as follows: A5 > A3 >

A6 > A2 > A1 > A4 > A7 > A8. In Table XIII, the ranking results of the present study are compared to Fuzzy TOPSIS method [33] and 6 variants of the VIKOR method.



TABLE XII. THE PROCESS SCORE-POINT MATRIX AFTER REARRANGEMENT

Alternatives	Points							
	P5	P3	P6	P2	P1	P4	P7	P8
A5		-3	-5	-5	-5	-5	-5	-5
A3	3		-3	-3	-2.3333	-3	-3	-3
A6	5	3		-0.3333	-1	-1	-1	-3
A2	5	3	0.3333		-1	-3	-3	-3
A1	5	2.3333	1	1		-0.3333	-3.6667	-5
A4	5	3	1	3	0.3333		-1	-3
A7	5	3	1	3	3.6667	1		-0.3333
A8	5	3	3	3	5	3	0.3333	

TABLE XIII. THE COMPARISON AMONG DIFFERENT METHODS FOR CHOOSING THE GRINDING WHEEL

Alternatives	Methods							
	Original VIKOR	Comprehensive VIKOR	Fuzzy VIKOR	Regret VIKOR	Modified VIKOR	Interval VIKOR	Fuzzy TOPSIS	Improved CURLI
A1	6	6	6	6	6	5	6	5
A2	3	3	3	5	3	4	3	4
A3	2	2	2	2	2	2	2	2
A4	4	5	4	3	5	3	4	6
A5	1	1	1	1	1	1	1	1
A6	5	4	5	4	4	6	5	3
A7	7	7	7	7	7	8	7	7
A8	8	8	8	8	8	7	8	8

The comparison results reveal a high degree of correlation between the methodologies. All approaches offer the same optimal and secondary options. In addition, the proposed method produces 7<sup>th</sup> and 8<sup>th</sup> alternatives similar to the other methods. There is a variation in the arrangement of alternatives between the 3<sup>rd</sup> and 6<sup>th</sup>. However, the difference is small, and this does not significantly impact the overall conclusion. It is evident that analyzing the sensitivity of the method plays an important role in solving the MCDM problem [34]. In this article, we inspect the sensitivity of the proposed approach by ranking the alternatives in the case of withdrawing at least one alternative out of the group randomly. The rank of the alternatives then is compared to the ideal order. The term ideal means that there is no reverse occurring if an alternative is eliminated. In the first scenario, the worst choice A8 is withdrawn from the calculation (Figure 1).

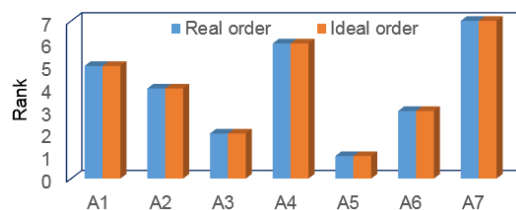


Fig. 1. The order of the solution after removing A8.

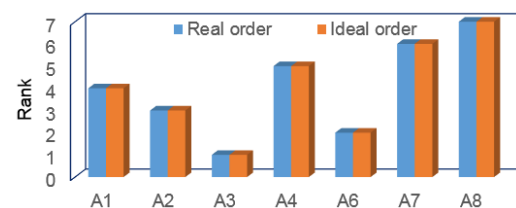


Fig. 2. The order of the solution after removing A5.

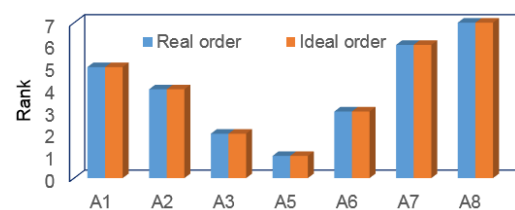


Fig. 3. The order of the solution after removing A4.

After that, the best alternative is eliminated from the list (Figure 2), and finally, the alternative in the middle of the rank is eliminated (Figure 3). It is observed that in all the considered cases, the order of the alternatives is exactly the same as the ideal order. This proves that the proposed method is reliable and applicable in solving MCDM problems.

### B. Case Study 2

In the second case study, the proposed method is applied to find out the best service suppliers. Each criterion is demonstrated by linguistic variables and the decision is made based on 5 criteria: product quality (C1), price (C2), delivery process (C3), service quality (C4), and past efficiency (C5). Each criterion in an alternative row has 3 sub-options and the meaning of linguistic variables is illustrated in Table XIV [35]. The smallest criterion C2 is the best, and for the others, the largest is the best.

The Fuzzy TOPSIS method is applied to rank the alternatives [35]. The statistics of this alternate ranking approach will be compared to those of the proposed method. Applying the improved CURLI method to estimate the alternatives in Table XIV is the same as in the case study 1, and the ranking results are compared to the Fuzzy TOPSIS [35] in Table XV.

TABLE XIV. DECISION-SCORE MATRIX FOR CHOOSING SERVICE SUPPLIER [35]

Alternatives	Criteria				
	C1	C2	C3	C4	C5
A1	(M, G, G)	(VG, M, G)	(G, G, G)	(SB, G, G)	(G, SG, VG)
A2	(G, G, VG)	(SG, G, G)	(VG, VG, VG)	(M, G, G)	(G, G, VG)
A3	(G, G, M)	(SG, VG, SG)	(G, SG, SB)	(G, G, G)	(G, VG, SG)

M – medium; G – good; VG – very good; SG – small good; SB – small bad

TABLE XV. COMPARISON BETWEEN THE IMPROVED CURLI AND THE FUZZY TOPSIS FOR CHOOSING THE SERVICE SUPPLIERS

Alternatives	Methods	
	Fuzzy TOPSIS	Improved CURLI
A1	2	2
A2	1	1
A3	3	3

It can be seen that the ranks of the alternatives of the two methods are precisely the same: A2 is the best and A3 is the worst alternative. This once again confirms the reliability of the proposed method. The sensitivity analysis of the alternative ranking is also performed and evaluated in detail. Figures 4-6 indicate the chart of the solution ranking after eliminating A1, A2, and A3.

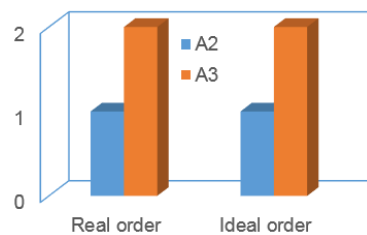


Fig. 4. The order of the solution after removing A1.

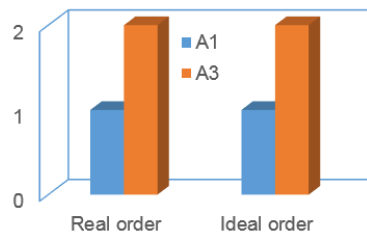


Fig. 5. The order of the solution after removing A2.

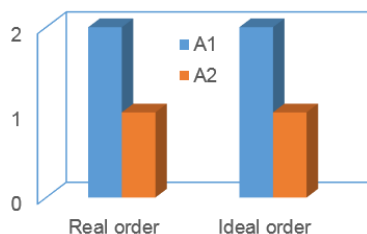


Fig. 6. The order of the solution after removing A3.

The comparison shows that there is no reverse phenomenon appearing in all considered scenarios. This once again confirms the success of the proposed improved CURLI method in the determination of the priority order.

#### IV. CONCLUSION

This paper proposes the improved CURLI approach based on the original CURLI method. Two case studies were implemented to evaluate the methodology efficiency with different types of variables. The ranking results are compared to those of other methods to verify the reliability of the proposed method. The following conclusions are drawn:

- It is notable that there was no major difference between the top and bottom positions. Additionally, the sensitivity of the improved CURLI was examined in the case of removing some random alternatives. The results indicate that there is no reverse of rank in any of the considered case studies. This is strong evidence that supports the idea that the improved CURLI method is better than the old MCDM when there are problems with data and uncertainty.
- It is also interesting that the proposed method can rank problems in which multiple factors influence each criterion at an alternative level without the requirement that the input data should be normalized or the criteria weights be evaluated.
- The improved CURLI method allows solving MCDM problems when the factors of the criteria can be linguistic variables or a data set. This study is the first step towards enhancing the understanding of the CURLI method to optimize MCDM problems.

#### LIST OF ACRONYMS

MCDM	Multi-Criteria Decision-Making
CURLI	Collaborative Unbiased Rank List Integration
MCDA	Multiple-Criteria Decision Analysis
MADM	Multi-Attribute Decision-Making
TOPSIS	Technique for Order Performance by Similarity to Ideal Solution
VIKOR	Vlsekriterijumska optimizacija i Kompromisno Resenje (in Serbian)
MARCOS	Measurement Alternatives and Ranking according to Compromise Solution
PEG	Pareto-Edgeworth Grierson
CODAS	COmbinative DIstance-based ASsessment
R	Ranking of the attributes and alternatives
SAW	Simple Additive Weighting
WASPAS	Weighted Aggregates Sum Product Assessment
MOORA	Multi-Objective Optimization on the basis of Ratio Analysis
COPRAS	Complex PRoportional ASsessment
PIV	Proximity Indexed Value
MABAC	Multi-Attributive Border Approximation area Comparison

#### REFERENCES

- [1] C. Zopounidis and M. Doumpos, Eds., *Multiple Criteria Decision Making: Applications in Management and Engineering*, 1st ed. New York, NY, USA: Springer, 2017.

- [2] S. Alshehri, "Multicriteria Decision Making (MCDM) Methods for Ranking Estimation Techniques in Extreme Programming," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 3073–3078, Jun. 2018, <https://doi.org/10.48084/etasr.2104>.
- [3] R. Umar, Sunardi, and Y. B. Fitriana, "Taxonomy of Fuzzy Multi-Attribute Decision Making Systems in Terms of Model, Inventor and Data Type," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2568–2571, Feb. 2018, <https://doi.org/10.48084/etasr.1747>.
- [4] I. Kaya, M. Colak, and F. Terzi, "A comprehensive review of fuzzy multi criteria decision making methodologies for energy policy making," *Energy Strategy Reviews*, vol. 24, pp. 207–228, Apr. 2019, <https://doi.org/10.1016/j.esr.2019.03.003>.
- [5] C. Kahraman, B. Oztaysi, I. Ucal Sari, and E. Turanoglu, "Fuzzy analytic hierarchy process with interval type-2 fuzzy sets," *Knowledge-Based Systems*, vol. 59, pp. 48–57, Mar. 2014, <https://doi.org/10.1016/j.knsys.2014.02.001>.
- [6] I. Dzitac, F. G. Filip, and M.-J. Manolescu, "Fuzzy Logic Is Not Fuzzy: World-renowned Computer Scientist Lotfi A. Zadeh," *International Journal of Computers Communications & Control*, vol. 12, no. 6, pp. 748–789, Dec. 2017.
- [7] U. M. Modibbo, M. Hassan, A. Ahmed, and I. Ali, "Multi-criteria decision analysis for pharmaceutical supplier selection problem using fuzzy TOPSIS," *Management Decision*, vol. 60, no. 3, pp. 806–836, Jan. 2022, <https://doi.org/10.1108/MD-10-2020-1335>.
- [8] A. Shamsuzzoha, S. Piya, and M. Shamsuzzaman, "Application of fuzzy TOPSIS framework for selecting complex project in a case company," *Journal of Global Operations and Strategic Sourcing*, vol. 14, no. 3, pp. 528–566, Jan. 2021, <https://doi.org/10.1108/JGOSS-07-2020-0040>.
- [9] M. T. J. Ansari, F. A. Al-Zahrani, D. Pandey, and A. Agrawal, "A fuzzy TOPSIS based analysis toward selection of effective security requirements engineering approach for trustworthy healthcare software development," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, Sep. 2020, Art. no. 236, <https://doi.org/10.1186/s12911-020-01209-8>.
- [10] K. Palczewski and W. Salabun, "The fuzzy TOPSIS applications in the last decade," *Procedia Computer Science*, vol. 159, pp. 2294–2303, Jan. 2019, <https://doi.org/10.1016/j.procs.2019.09.404>.
- [11] N. H. Baharin, N. F. Rashidi, and N. F. Mahad, "Manager selection using Fuzzy TOPSIS method," *Journal of Physics: Conference Series*, vol. 1988, no. 1, Apr. 2021, Art. no. 012057, <https://doi.org/10.1088/1742-6596/1988/1/012057>.
- [12] A. M. Talib, "Fuzzy VIKOR Approach to Evaluate the Information Security Policies and Analyze the Content of Press Agencies in Gulf Countries," *Journal of Information Security*, vol. 11, no. 4, pp. 189–200, Aug. 2020, <https://doi.org/10.4236/jis.2020.114013>.
- [13] S. Emec and G. Akkaya, "Stochastic AHP and fuzzy VIKOR approach for warehouse location selection problem," *Journal of Enterprise Information Management*, vol. 31, no. 6, pp. 950–962, Jan. 2018, <https://doi.org/10.1108/JEIM-12-2016-0195>.
- [14] Y. Suh, Y. Park, and D. Kang, "Evaluating mobile services using integrated weighting approach and fuzzy VIKOR," *PLOS ONE*, vol. 14, no. 6, May 2019, Art. no. e0217786, <https://doi.org/10.1371/journal.pone.0217786>.
- [15] M. S. Ismail and A. Felix, "Integrated fuzzy VIKOR and TOPSIS system for the sustainable development in Islam," *AIP Conference Proceedings*, vol. 2385, no. 1, Jan. 2022, Art. no. 130027, <https://doi.org/10.1063/5.0070740>.
- [16] M. Stankovic, Z. Stevic, D. K. Das, M. Subotic, and D. Pamucar, "A New Fuzzy MARCOS Method for Road Traffic Risk Analysis," *Mathematics*, vol. 8, no. 3, Mar. 2020, Art. no. 457, <https://doi.org/10.3390/math8030457>.
- [17] M. Kovac, S. Tadic, M. Krstic, and M. B. Bouraima, "Novel Spherical Fuzzy MARCOS Method for Assessment of Drone-Based City Logistics Concepts," *Complexity*, vol. 2021, Dec. 2021, Art. no. e2374955, <https://doi.org/10.1155/2021/2374955>.
- [18] M. Bakır and O. Atalık, "Application of Fuzzy AHP and Fuzzy MARCOS Approach for the Evaluation of E-Service Quality in the Airline Industry," *Decision Making: Applications in Management and Engineering*, vol. 4, no. 1, pp. 127–152, Mar. 2021, <https://doi.org/10.31181/dmame2104127b>.
- [19] S. Miao, R. J. Hammell, T. Hanratty, and Z. Tang, "Comparison of Fuzzy Membership Functions for Value of Information Determination," in *25th Midwest Artificial Intelligence and Cognitive Science Conference*, Washington, DC, USA, Apr. 2014, pp. 1–8.
- [20] S. Princy and S. S. Dhenakaran, "Comparison of triangular and trapezoidal fuzzy membership function," *Journal of Computer Science and Engineering*, vol. 2, no. 8, pp. 46–51, 2016.
- [21] H. K. Le, "Multi-Criteria Decision Making in the Milling Process Using the PARIS Method," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9208–9216, Oct. 2022, <https://doi.org/10.48084/etasr.5187>.
- [22] D. Duc Trung, "Multi-criteria decision making under the MARCOS method and the weighting methods: applied to milling, grinding and turning processes," *Manufacturing Review*, vol. 9, Jan. 2022, Art. no. 3, <https://doi.org/10.1051/mfreview/2022003>.
- [23] R. M. Dawes and B. Corrigan, "Linear models in decision making," *Psychological Bulletin*, vol. 81, pp. 95–106, 1974, <https://doi.org/10.1037/h0037613>.
- [24] H. J. Einhorn and W. McCoach, "A simple multiattribute utility procedure for evaluation," *Behavioral Science*, vol. 22, no. 4, pp. 270–282, 1977, <https://doi.org/10.1002/bs.3830220405>.
- [25] D. D. Trung, "Development of data normalization methods for multi-criteria decision making: applying for MARCOS method," *Manufacturing Review*, vol. 9, 2022, Art. no. 22, <https://doi.org/10.1051/mfreview/2022019>.
- [26] N. Vafaei, R. A. Ribeiro, and L. M. Camarinha-Matos, "Normalization Techniques for Multi-Criteria Decision Making: Analytical Hierarchy Process Case Study," in *Doctoral Conference on Computing, Electrical and Industrial Systems*, Costa de Caparica, Portugal, Apr. 2016, pp. 261–269, [https://doi.org/10.1007/978-3-319-31165-4\\_26](https://doi.org/10.1007/978-3-319-31165-4_26).
- [27] A. Jahan and K. L. Edwards, "A state-of-the-art survey on the influence of normalization techniques in ranking: Improving the materials selection process in engineering design," *Materials & Design*, vol. 65, pp. 335–342, Jan. 2015, <https://doi.org/10.1016/j.matdes.2014.09.022>.
- [28] J. R. Kiger and D. J. Annibale, "A new method for group decision making and its application in medical trainee selection," *Medical Education*, vol. 50, no. 10, pp. 1045–1053, 2016, <https://doi.org/10.1111/medu.13112>.
- [29] D. D. Trung, N. N. Ba, and D. H. Tien, "Application of the Curli method for multi-critical decision of grinding process," *Journal of Applied Engineering Science*, vol. 20, no. 3, pp. 634–643, 2022, <https://doi.org/10.5937/jaes0-35088>.
- [30] D. D. Trung, "Multi-criteria decision making of turning operation based on PEG, PSI and CURLI methods," *Manufacturing Review*, vol. 9, 2022, Art. no. 9, <https://doi.org/10.1051/mfreview/2022007>.
- [31] D. D. Trung, "Comparison R and CURLI methods for multi-criteria decision making," *Advanced Engineering Letters*, vol. 1, no. 2, pp. 46–56, Jul. 2022, <https://doi.org/10.46793/adeletters.2022.1.2.3>.
- [32] P. Chatterjee and S. Chakraborty, "A comparative analysis of VIKOR method and its variants," *Decision Science Letters*, vol. 5, no. 4, pp. 469–486, 2016.
- [33] S. R. Maity and S. Chakraborty, "Grinding Wheel Abrasive Material Selection Using Fuzzy TOPSIS Method," *Materials and Manufacturing Processes*, vol. 28, no. 4, pp. 408–417, Apr. 2013, <https://doi.org/10.1080/10426914.2012.700159>.
- [34] D. S. Pamucar, D. Bozanic, and A. Randelovic, "Multi-criteria decision making: An example of sensitivity analysis," *Serbian Journal of Management*, vol. 12, no. 1, pp. 1–27, 2017, <https://doi.org/10.5937/sjml2-9464>.
- [35] A. Ozbek, "Supplier Selection with Fuzzy TOPSIS," *Journal of Economics and Sustainable Development*, vol. 6, no. 18, pp. 114–125, Jan. 2015.

# Design and Implementation of a Second Order Continuous-Time $\Sigma\Delta$ Modulator for ECG Signal Acquisition

**M. Kavitha**

Department of Electronics and Communication Engineering, JSS Academy of Technical Education, India  
mkavitha@jssateb.ac.in  
(corresponding author)

**S. Akhila**

Department of Electronics and Communication Engineering, B.M.S College of Engineering, India  
akhilas.ece@bmsce.ac.in

**Anand Kannan**

Texas Instruments India Private Ltd, India  
anandk@ti.com

*Received: 15 December 2022 | Revised: 31 December 2022 | Accepted: 3 January 2023*

## ABSTRACT

The recent developments in biosignal acquisition devices for continuous supervision of cardiovascular signs of high-risk patients require a high-precision and low-power Analog Front End (AFE) circuit. The proposed design adopts Continuous-Time (CT) Sigma-Delta Modulator ( $\Sigma\Delta$ ) architecture to achieve high resolution and Signal-to-Noise And Distortion ratio (SINAD) requirements. The proposed modulator is a second-order CT- $\Sigma\Delta$  with Cascade of Integrators Feed-Forward (CIFF) architecture that consists of a CT loop filter, a single-bit quantizer, and a Digital-to-Analog Converter (DAC). The use of single-bit quantization in the design reduces circuit complexity and power consumption. To use the designed  $\Sigma\Delta$  for measuring ECG signals, a bandwidth ( $B_w$ ) of 150 Hz is considered with a sampling frequency ( $f_s$ ) of 153.6kHz to achieve an oversampling ratio of 512. The design is simulated in a standard Cadence Virtuoso EDA tool at 180nm CMOS technology, operating at 1.8V supply voltage at the block level. The simulation results for the designed modulator show that SINAD is 104.5dB, the Effective Number Of Bits (ENOB) is 17.06bits, with power consumption of 24 $\mu$ W, and achieves Schreier's Figure-Of-Merit (FOM) equal to 172.45dB.

**Keywords-***sigma delta modulator; continuous-time; quantizer; DAC*

## I. INTRODUCTION

Innovations in communications and advances in CMOS technology along with low power design techniques have given impetus to the research in devices intended for the acquisition of biopotential signals. These signals are very weak, having amplitude and bandwidth ranging from  $\mu$ V to mV and from DC to a few kHz [1]. Typical biopotential acquisition systems for acquiring signals such as Electroencephalogram (EEG), Electromyogram (EMG), Electro-cardiogram (ECG), Electrooculogram (EOG), as shown in Figure 1, consist of an Instrumentation Amplifier (IA) to provide adequate gain with low noise, a Low-Pass Filter (LPF) to confine noise bandwidth and an Analog-to-Digital Converter (ADC) that interfaces the AFE and the digital processing unit. Several architectures to design the IA for the AFE of a biopotential signal acquisition

have been reported using techniques such as three operational amplifiers [2], differential difference amplifier [3], capacitively-coupled chopper [4], and current-balancing IA [5]. All these techniques have their own advantages and limitations. The LPF circuit designed using OTA-C continuous-time filtering [6] operates in the sub-threshold region to save power for portable ECG signal detection and a second order LPF [7] is used with low noise and low power for programmable AFEs. ADCs also play a significant role in biosignal acquisition and health monitoring integrated circuits [8].

ADCs that are most commonly used for the ECG signal acquisition are Successive Approximation Register (SAR) and oversampled ADCs. SAR ADCs have moderate accuracy and excellent power to performance ratio, but low resolution of 8 to 12 bits [9, 10]. Higher resolution SAR ADCs need high

resolution feedback DACs, resulting in increased area and power consumption.  $\Sigma\Delta$  ADCs are oversampled converters, comprising of either a Switched-Capacitor (SC) or a Continuous-Time (CT)  $\Sigma\Delta$  Modulator. A state-of-the-art SC implementation provides high precision and accuracy but due to switched integrators it suffers from stringent settling and slew requirements that lead to more power consumption [11, 12]. Compared to SC, CT architecture consumes less power due to inherent anti-alias filtering that obviates the need for a dedicated analog filter required in the signal acquisition system. In order to attain high resolution, wide dynamic range and low power consumption CT $\Sigma\Delta$ Ms are used with active integrators [13-15]. Active RC and Gm-C based are the two types of CT $\Sigma\Delta$  structures. This paper discusses the implementation of CT $\Sigma\Delta$ M with active RC integrators, summer, quantizer, and resistive DAC. The designed modulator exhibits SINAD of 104.5dB and ENOB of 17.06bits. The experimental results are aligned with mathematical model and simulations.

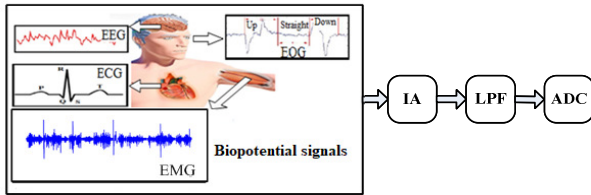


Fig. 1. Block diagram of a biopotential acquisition system.

## II. DESIGN METHODOLOGY

The proposed work uses the top-down hierarchical design of  $\Sigma\Delta$  synthesis methodology shown in Figure 2.

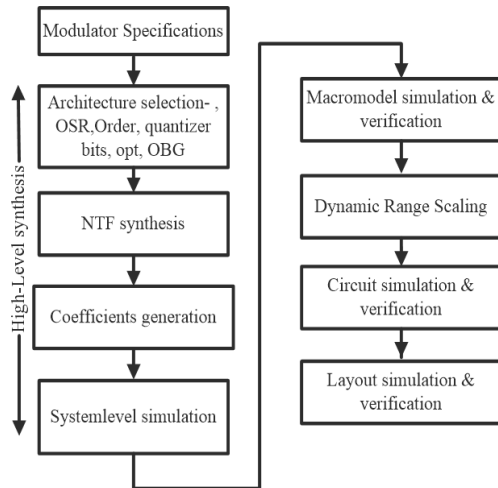


Fig. 2. Design methodology flow graph.

### A. Modulator Specifications

For acquiring the ECG signal, the system-level specifications of  $\Sigma\Delta$ M are set to achieve ENOB of 16-18 bits, SNR > 100dB with sampling rate  $f_s$  of around 150kHz and signal bandwidth  $B_w$  of 150Hz.

### B. Architecture Selection

The design starts with the selection of the Noise Transfer Function (NTF) and the modulator architecture with parameters such as the Oversampling Ratio (OSR) and the order of the modulator (L) and N-bit quantizer. The value of OSR or clock rate is constrained by technology node and power consumption, which varies between 8 and 512 [16]. Higher order modulators having  $L > 2$  improve the SNR significantly, but with an increase in circuit complexity and deterioration of stability. Hence, the proposed design aims for a lower order modulator with higher OSR. The single-bit quantizer is intrinsically linear without mismatch in the quantization step, whereas the multibit quantizer exhibits mismatched quantization steps but has low in-band quantization noise and allows a more aggressive NTF with higher Out-of-Band Gain (OBG), reduced sensitivity to the clock jitter, and lower slew rate requirements from the loop filter [17]. According to the rule-of-thumb [18], for a single-bit quantizer to ensure stability the value of OBG must be about 1.5 and for a multibit quantizer between 1.5 to 3.5. Higher values of OBG cause overloading of the quantizer resulting in modulator instability.

### C. NTF Synthesis

$\Sigma\Delta$ M employs a NTF implemented using the high-pass Butterworth or inverse Chebyshev transfer function, having a cut-off frequency located outside the signal band. Considering the parameters OSR, L, N, OBG, optimization value and type of filter, the NTF is determined using Schreier's MATLAB Delta-Sigma toolbox [19]. For an ideal  $L^{\text{th}}$  order system, the NTF is given by  $(1 - Z^{-1})^L$  with all poles located at  $Z = 0$  and zeros at  $Z = 1$ . Spreading these zeros within the signal band and moving the poles within the unit circle, significantly improves SQNR and stability. The modulator becomes unstable if system poles move out of the unit circle.

### D. Coefficient Generation

NTF is realized to obtain CT $\Sigma\Delta$ M coefficients for feedback (FB) and feedforward (FF) form topologies. Each of them is further distinguished by two different forms to generate four topologies - Cascade of Resonators FB (CRFB), Cascade of Integrators FB (CIFB), Cascade of Resonators FF (CRFF), and Cascade of Integrators FF (CIFF). The FB structure consists of several DACs that fed back to all the integrator output and does not require a large summer before the quantizer. In FF structure, a single DAC is required in the FB path without excess loop delay compensation, which is more area efficient [20]. The other advantage of the FF structure is the reduced output swing of the first integrator. Given a certain output range, the first integrator allows a bigger loop gain and hence lower performance requirements on the following stages. However, the FF architecture requires a multi-input adder to sum all the FF branches before the quantizer. In the current work, a CT $\Sigma\Delta$ M with CIFF architecture is implemented with its state-space representation as shown in Figure 3. The CTABCD matrix is given by (1) along with its description in Table I.

$$[ABCDc] = \left[ \begin{array}{ccc|ccc} 0 & A & 0 & 1 & B & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0.6667 & 0.2288 & C & 0 & D & 0 \end{array} \right] \quad (1)$$

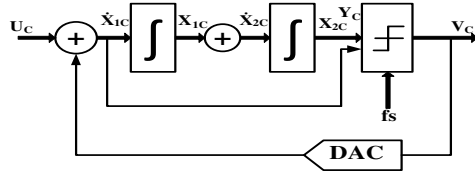


Fig. 3. State-space representation of the CTΣΔM.

TABLE I. DESCRIPTION OF THE ABCDc MATRIX

Matrix & order	Function	Representation
A L×L	Loop filter interconnects	0: No FB connection from $X_{1c}$ to $\dot{X}_{1c}$ 0: No FB connection from $X_{2c}$ to $\dot{X}_{1c}$ 1: $X_{1c}$ directly feeds $\dot{X}_{2c}$ 0: No FB connection from $X_{2c}$ to $\dot{X}_{2c}$
B L×L	Modulator i/p & the f/b DAC o/p to the loop filter.	1: FB connection from $U_c$ to $\dot{X}_{1c}$ -1: FB connection from $V_c$ to $\dot{X}_{1c}$ 0: No connection from $U_c$ to $\dot{X}_{2c}$ 0: No FB connection from $V_c$ to $\dot{X}_{2c}$
C 1×L	K coefficients, i/p from loop filter to the summer.	$K_1$ : Coefficient value from $X_{1c}$ to $Y_c$ $K_2$ : Coefficient value from $X_{2c}$ to $Y_c$ $K_1 = 0.6667$ $K_2 = 0.2288$
D 1×L	Excess loop delay	0: No connection from $U_c$ to $Y_c$ 0: No connection from $V_c$ to $Y_c$

Based on (1), the obtained block diagram of the 2<sup>nd</sup> order CTΣΔM with CIFF is as illustrated in Figure 4.

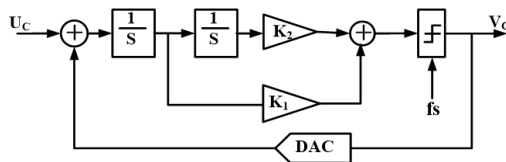


Fig. 4. Block diagram of the second order CTΣΔM with CIFF.

### E. System Level Simulation

The behavioral simulation is performed using MATLAB excluding circuit noise to determine the key parameters and individual block specifications. To use the designed ΣΔM for measuring ECG signals, bandwidth of 150Hz is considered with a sampling frequency of 153kHz. For OSR=512, L=2, N=1, OBG=1.5, opt =1, the pole-zero plot by which design stability is verified and the frequency response of NTF is illustrated in Figure 5. The pole-zero map of the 2<sup>nd</sup> order system has two complex conjugate poles (marked with X) and two zeros (marked with O) at dc that suppress the in-band quantization noise. Frequency response shows that the OBG is 3.438dB (1.5). The resulting NTF is given by:

$$\text{NTF}(z) = \frac{Z^2 - 2Z + 1}{Z^2 - 1.225Z + 0.4415} \quad (2)$$

The time domain response of ΣΔM with a two-level quantizer is shown in Figure 6. For a sinusoidal input, the output is pulse-density modulated. The density of the pulse output waveform varies with input amplitude. The frequency domain response of the designed CTΣΔM is shown in Figure 7

in which the power is maximum at the input frequency of the signal and it clearly exhibits 2<sup>nd</sup> order noise-shaping with a slope of 40dB/decade. The ΣΔM achieves high resolution by the combination of oversampling and noise shaping. For the designed modulator, the MATLAB simulation depicts in-band SNR of 103.85dB at input amplitude of 0.5V with ENOB of 16.96 bits as shown in Figure 7.

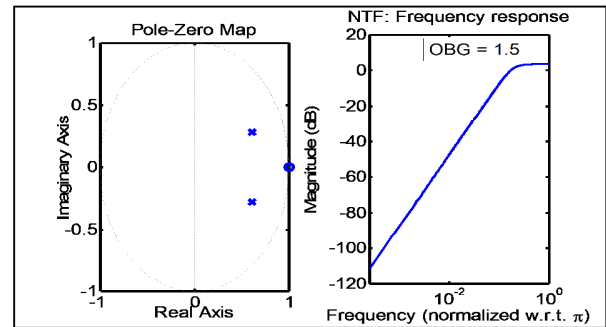


Fig. 5. Left: pole-zero plot, right: NTF frequency response.

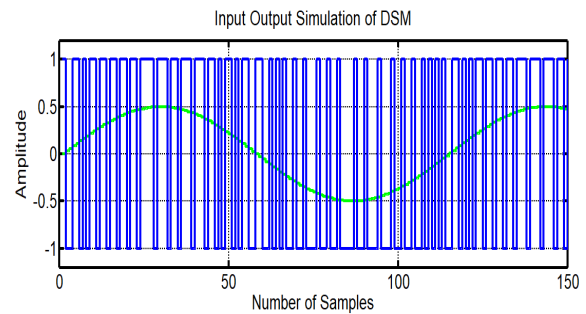


Fig. 6. Time domain plot of input and output with 2-level quantizer.

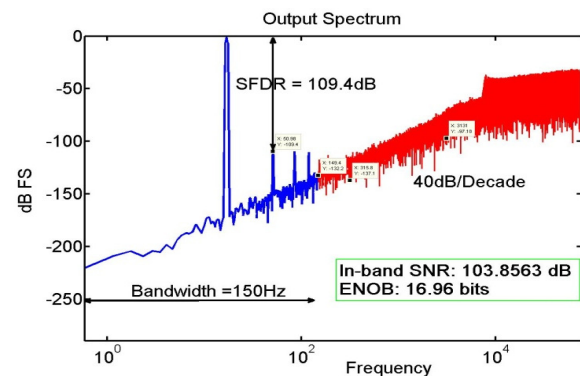


Fig. 7. Measured output spectrum.

The obtained value of peak SNR of 122.7dB at -1.7dBFS input signal with dynamic range of 120dB is shown in Figure 8. The values of in-band SNR and ENOB for different values of OSR of 2<sup>nd</sup> order CTΣΔM is shown in Figure 9. It's observed that SNR increases with OSR at the rate of approximately 13dB/octave.



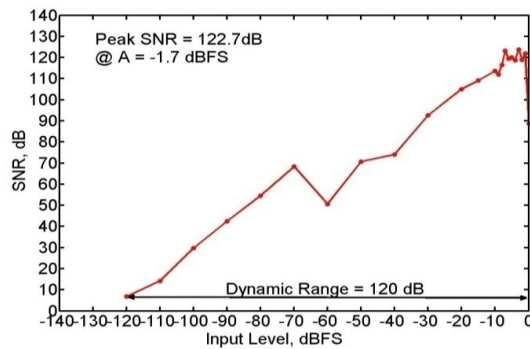


Fig. 8. Measured SNR versus input level.

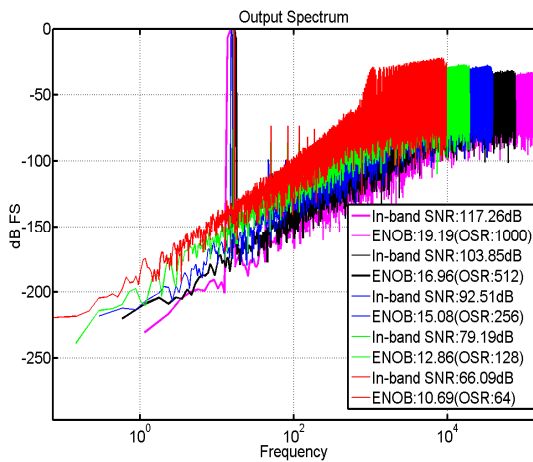
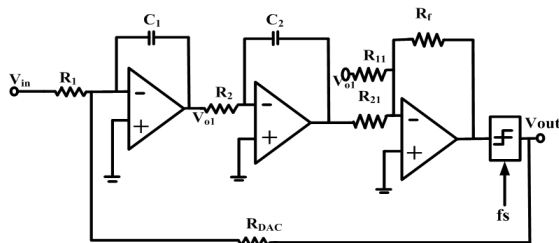


Fig. 9. PSD of the modulator output with values of SNR and ENOB for different OSR values.

#### F. Macro Model Simulation and Verification

The obtained system-level behavioral model of the  $\Sigma\Delta$  architecture that satisfies the required specifications was translated into a macro model consisting of active and passive components in standard CMOS technology using cadence spectre. Ideal building blocks (clock generator, integrators, quantizer, and DAC) [21] were used for validating the design architecture. The schematic of the single-ended 2<sup>nd</sup> order CT active RC type  $\Sigma\Delta$  is shown in Figure 10.

Fig. 10. Schematic of single ended 2<sup>nd</sup> order CTΣΔM.

The schematic is composed of a single-loop, CIFF structure, and single-bit feedback DAC. Active integrators are used to obtain the low power 2<sup>nd</sup> order CT system, VerilogA model for a single-bit quantizer, and resistive DAC. Ideal

switch and RESET signal are used across the integrator capacitors at the beginning of the simulation. When the RESET is high, the switch is shorted and sets the DC voltage on the integrator output.

The values of the active RC integrators are calculated by using the sampling frequency  $f_s$ . The summer circuit is designed using the K-coefficients. The designed values are  $R_1 = R_2 = 100k\Omega$ ,  $C_1 = C_2 = 66.67pF$ ,  $R_f = 100k\Omega$ ,  $R_{11} = 150k\Omega$ ,  $R_{21} = 437.06k\Omega$ . The simulation of the designed modulator with macro models is carried out using the components from analogLib and ahdlLib of cadence for a signal with input amplitude of 0.25V, 150Hz bandwidth, and 153.6kHz sampling frequency. The swing of integrators, summer, quantizer, DAC before scaling is shown in Table II. Before scaling, the integrator outputs were saturated having values more than the supply voltage, hence through scaling the output voltage levels were reduced.

#### G. Dynamic Range Scaling

An integrator's output doesn't respond to changes in its input if it saturates. In order to avoid integrator saturation, the loop filter coefficients were scaled without modifying the loop filter transfer function. This is referred to as dynamic range scaling. The values of active RC integrators are calculated using (3) where  $f_s$  is the sampling frequency and  $\alpha$  and  $\beta$  are scaling coefficients.

$$f_s = \frac{1}{\alpha R_1 C_1} \quad \text{and} \quad f_s = \frac{1}{\beta R_2 C_2} \quad (3)$$

The weighted addition of the integrator outputs is carried out using a summing amplifier. The design of the summing amplifier is calculated using the K-coefficients obtained from Table I.

$$K_1 = \frac{K_1}{\alpha} ; \quad R_{11} = \frac{R_f}{K_1} \quad (4)$$

$$K_2 = \frac{K_2}{\alpha\beta} ; \quad R_{21} = \frac{R_f}{K_2} \quad (5)$$

The lower and upper limits for R and C are set by the matching consideration and thermal noise level that is fixed by the overall dynamic range requirements. Table II shows the value of the output swing across different blocks before and after scaling. It is clear that after scaling the voltage levels are reduced, maintaining the desired SNR. The obtained parameter values before and after scaling are shown in Table II.

TABLE II. OUTPUT VALUES BEFORE AND AFTER SCALING.

Parameters	Before scaling	After scaling
Integrator-1	-0.375 to 2.19V	0.46 to 1.34V
Integrator-2	-0.94 to 2.69V	0.32 to 1.42V
Summer	0.023 to 1.69V	0.20 to 1.56V
Quantizer	0 to 1V	0 to 1V
DAC	0 to 1.8V	0 to 1.8V
Signal power input	-12.04	-12.04
Signal power output	-17.15	-17.15
Sinad input	234.8	234.8
Sinad output	-13.97	-13.97
SINAD	102.02	104.5
ENOB	16.65	17.06

### III. SECOND-ORDER CTΣΔM

The proposed modulator is a 2<sup>nd</sup>-order CTΣΔM with CIFF architecture that consists of a CT loop filter, single-bit quantizer, and resistive DAC as shown in Figure 11. Fully differential architecture implementation minimizes even-order harmonics and common-mode noise [22]. The circuits implemented with active RC integrators are highly linear, insensitive to parasitics, and simple to design. The first stage processes in-band and DAC signal. The designed values of the first stage integrator input resistance  $R_1$  and capacitance  $C_1$  are 100kΩ and 161.8pF respectively. In order to reduce power and to enhance the linearity of the modulator,  $R_1$  is maximized up to the input-referred thermal noise,  $P_N$  given by:

$$P_N = 8kTB_w \left[ R_1 + R_{DAC} \frac{R_1^2}{R_{DAC}^2} + \frac{2}{3g_{m,OTA}} \left( 1 + \frac{R_1}{R_{DAC}} \right)^2 \right] \quad (6)$$

where  $k$ ,  $T$ ,  $B$ ,  $R_{DAC}$ ,  $g_{m,OTA}$  represent the Boltzmann constant, the absolute temperature, signal bandwidth, feedback DAC resistors, and amplifier input transconductance, respectively. The values of the 2<sup>nd</sup> stage integrator  $R_2$  and  $C_2$  are 200kΩ, 41.89pF, respectively.

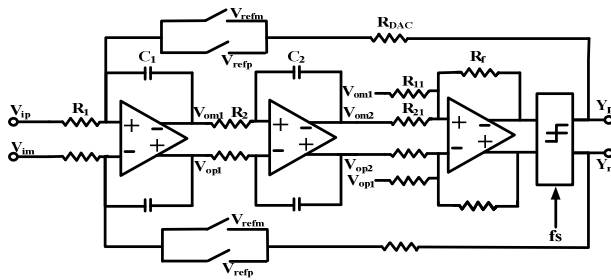


Fig. 11. Schematic of the fully differential 2<sup>nd</sup>-order CTΣΔM designed with active-RC type.

The summing amplifier uses the resistors  $R_f = 80k\Omega$ ,  $R_{11} = 61.8k\Omega$ ,  $R_{21} = 143.26k\Omega$  designed using (4) and (5) FF coefficients. Switched-resistor feedback DAC,  $R_{DAC}$ , is chosen to be equal to  $R_1$  for better matching performance that provides less noise [23] than the current-steering DAC.

### IV. MEASUREMENT RESULTS

The presented modulator has been implemented in 180nm CMOS technology at 1.8V supply voltage, for applied input of a sinusoidal signal of 45Hz with input amplitude of 0.25V and sampling frequency of 153KHz achieving OSR of 512. For Hanning window of 16384-points, the block level simulation depicts SINAD of 104.5dB with ENOB of 17.06 with a power consumption of 24μW and achieves Schreier's Figure-of-Merit of 172.45dB. Table III shows the performance comparison of the proposed work with other known methods, showing an increase in the values of SINAD and ENOB.

### V. CONCLUSION

This paper proposes the CTΣΔM design for the acquisition of ECG signals. The cascade of integrators with FF is opted for the implementation of loop filter, with active-RC integrators and resistive feedback DAC that reduces noise and power

dissipation. The set target of ENOB is attained by the systematic coefficient generation using MATLAB and the loop filter coefficients scaling using the cadence EDA tool. The designed modulator implemented in 180nm CMOS technology with signal bandwidth of 150Hz, high OSR of 512 has SINAD of 104.5dB and ENOB of 17.06 achieving around 10% increase in comparison with existing techniques. The modulator consumes 24μW power from 1.8V supply and achieves Schreier's FOM of 172.45dB.

TABLE III. CTΣΔM PERFORMANCE COMPARISON

Reference	[15] <sup>b</sup>	[24] <sup>a</sup>	[21] <sup>a</sup>	Proposed
Year of publication	2021	2018	2012	
Technology [nm]	180	180	180	180
Bandwidth	2KHz	200Hz	20KHz	150Hz
OSR	256	256	512	512
SINAD in dB	94.28	96	87.3	104.5
ENOB *	15.37	16	10	17.06
Power	32.34 μW	-	-	24 μW
FoMs **	177.9dB	-	-	172.45dB

\*\*FoMs in dB = SINAD + 10log<sub>10</sub>  $\left( \frac{B_w}{P_{power}} \right)$ , \*ENOB =  $\frac{SINAD-1.76}{6.02}$ ,  
<sup>a</sup> block-level simulation, <sup>b</sup> transistor-level simulation

### REFERENCES

- [1] F. Bagheri, N. Ghafarnia, and F. Bahrami, "Electrocardiogram (ECG) Signal Modeling and Noise Reduction Using Hopfield Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 3, no. 1, pp. 345–348, Feb. 2013, <https://doi.org/10.48084/etasr.243>.
- [2] E. M. Spinelli, R. Pallas-Areny, and M. A. Mayosky, "AC-coupled front-end for biopotential measurements," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 391–395, Mar. 2003, <https://doi.org/10.1109/TBME.2003.808826>.
- [3] W.-S. Wang, Z.-C. Wu, H.-Y. Huang, and C.-H. Luo, "Low-Power Instrumental Amplifier for Portable ECG," in *2009 IEEE Circuits and Systems International Conference on Testing and Diagnosis*, Chengdu, China, Apr. 2009, <https://doi.org/10.1109/CAS-ICTD.2009.4960836>.
- [4] Q. Fan, F. Sebastiano, J. H. Huijsing, and K. A. A. Makinwa, "A 1.8 μW 60 nV/√Hz Capacitively-Coupled Chopper Instrumentation Amplifier in 65 nm CMOS for Wireless Sensor Nodes," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 7, pp. 1534–1543, Jul. 2011, <https://doi.org/10.1109/JSSC.2011.2143610>.
- [5] P. Sutha and V. E. Jayanthi, "Implementation of a Biopotential Amplifier with a Conventional and Current-Balancing Approach for Foetal ECG Monitoring," *Circuits, Systems, and Signal Processing*, vol. 39, no. 6, pp. 2860–2879, Jun. 2020, <https://doi.org/10.1007/s00034-019-01311-x>.
- [6] S.-Y. Lee and C.-J. Cheng, "Systematic Design and Modeling of a OTA-C Filter for Portable ECG Detection," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 1, pp. 53–64, Oct. 2009, <https://doi.org/10.1109/TBCAS.2008.2007423>.
- [7] M. K. Adimulam, A. Divya, K. Tejaswi, and M. B. Srinivas, "A low power, low noise Programmable Analog Front End (PAFE) for biopotential measurements," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju, Korea (South), Jul. 2017, pp. 3844–3847, <https://doi.org/10.1109/EMBC.2017.8037695>.
- [8] M. Hamdani, M. Youcefi, A. Rabehi, B. Nail, and A. Douara, "Design and Implementation of a Medical TeleMonitoring System based on IoT," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8949–8953, Aug. 2022, <https://doi.org/10.48084/etasr.5040>.
- [9] W. Mao, Y. Li, C.-H. Heng, and Y. Lian, "A Low Power 12-bit 1-kS/s SAR ADC for Biomedical Signal Processing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 2, pp. 477–488, Oct. 2019, <https://doi.org/10.1109/TCSI.2018.2859837>.

- [10] J. Zheng, W.-H. Ki, and C.-Y. Tsui, "A Fully Integrated Analog Front End for Biopotential Signal Sensing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 11, pp. 3800–3809, Aug. 2018, <https://doi.org/10.1109/TCSI.2018.2854741>.
- [11] V. Sharma, N. Kumar Y.B., and V. M.h., "67 dB SNDR 20 kHz BW SC third-order  $\Sigma\Delta$  modulator with single Op-Amp and 20  $\mu$ W power consumption for bio-medical applications," *IET Circuits, Devices & Systems*, vol. 14, no. 6, pp. 881–891, 2020, <https://doi.org/10.1049/iet-cds.2019.0414>.
- [12] S.-Y. Lee, P.-H. Su, K.-L. Huang, Y.-W. Hung, and J.-Y. Chen, "High-Pass Sigma-Delta Modulator With Techniques of Operational Amplifier Sharing and Programmable Feedforward Coefficients for ECG Signal Acquisition," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 3, pp. 443–453, Jun. 2021, <https://doi.org/10.1109/TBCAS.2021.3082545>.
- [13] J. M. de la Rosa, "Sigma-Delta Modulators: Tutorial Overview, Design Guide, and State-of-the-Art Survey," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 1, pp. 1–21, Jan. 2011, <https://doi.org/10.1109/TCSI.2010.2097652>.
- [14] M. Kavitha, S. Akhila, and A. Kannan, "Flexible Analog Front End Architecture for Biomedical Applications," in *2021 IEEE 3rd PhD Colloquium on Ethically Driven Innovation and Technology for Society (PhD EDITS)*, Bangalore, India, Aug. 2021, <https://doi.org/10.1109/PhDEDITS53295.2021.9649471>.
- [15] F. Hemmati and E. Najafi Aghdam, "A low-power CT 2nd order Delta Sigma modulator using a new design methodology for biomedical applications," *AEU - International Journal of Electronics and Communications*, vol. 137, Jul. 2021, Art. no. 153779, <https://doi.org/10.1016/j.aeue.2021.153779>.
- [16] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*, 2nd ed. Hoboken, NJ, USA: Wiley-IEEE Press, 2017.
- [17] S. Pavan, N. Krishnapura, R. Pandarinathan, and P. Sankar, "A Power Optimized Continuous-Time  $\Delta\Sigma$  ADC for Audio Applications," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 351–360, Oct. 2008, <https://doi.org/10.1109/JSSC.2007.914263>.
- [18] W. L. Lee, "A Novel Higher Order Interpolative Modulator Topology for High Resolution Oversampling A/D Converters," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 1987.
- [19] "Delta Sigma Toolbox," *Mathworks*. <https://www.mathworks.com/matlabcentral/fileexchange/19-delta-sigma-toolbox>.
- [20] A. Anand, N. Koirala, R. K. Pokharel, H. Kanaya, and K. Yoshida, "Systematic Design Methodology of a Wideband Multibit Continuous-Time Delta-Sigma Modulator," *International Journal of Microwave Science and Technology*, vol. 2013, Mar. 2013, Art. no. e275289, <https://doi.org/10.1155/2013/275289>.
- [21] G. Zheng, S. P. Mohanty, and E. Kougianos, "Design and modeling of a continuous-time delta-sigma modulator for biopotential signal acquisition: Simulink vs. Verilog-AMS perspective," in *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*, Jul. 2012, pp. 1–6, <https://doi.org/10.1109/ICCCNT.2012.6396103>.
- [22] V. Kledrowetz, L. Fucik, R. Prokop, and J. Háze, "A 1 V 92 dB SNDR 10 kHz Bandwidth Second-Order Asynchronous Delta-Sigma Modulator for Biomedical Signal Processing," *Sensors*, vol. 20, no. 15, Jan. 2020, Art. no. 4137, <https://doi.org/10.3390/s20154137>.
- [23] W. Mohguen and S. Bouguezel, "Denoising the ECG Signal Using Ensemble Empirical Mode Decomposition," *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7536–7541, Oct. 2021, <https://doi.org/10.48084/etasr.4302>.
- [24] N. Beigh, A. Hamid, F. Beigh, and F. Ahmad, "2nd Order Sigma Delta Modulator Design using Delta Sigma Toolbox," *Asian Journal of Electrical Sciences*, vol. 7, no. 2, pp. 41–45, Sep. 2018.

# Improvement of Absorption and Emission Phenomena of 1.55 $\mu$ m Quantum Dot Laser using Indium Nitride

**Md. Abdullah Al Humayun**

Department of EEE, Eastern University, Bangladesh  
humayun0403063@gmail.com

**Mohd Abdur Rashid**

Department of EEE, Noakhali Science and Technology University, Bangladesh | Division of Electronics and Informatics, Gunma University, Japan  
marashid.eee@nstu.edu.bd  
(corresponding author)

**Anna Kuwana**

Division of Electronics and Informatics, Gunma University, Japan  
kuwana.anna@gunma-u.ac.jp

**Haruo Kobayashi**

Division of Electronics and Informatics, Gunma University, Japan  
koba@gunma-u.ac.jp

*Received: 23 November 2022 | Revised: 19 December 2022 | Accepted: 23 December 2022*

## ABSTRACT

This paper presents the improvement of emission and absorption phenomena of 1.55 $\mu$ m quantum dot laser by enhancing the stability of oscillation frequency and minimizing absorption loss. Among the key parameters related to these absorption and emission phenomena, the effective density of state, rate of change of carrier mobility, and optical feedback level were extensively investigated to reduce the fluctuation of the oscillation frequency of the emitted light as well as the absorption loss. These phenomena were investigated through mathematical analysis and numerical simulation using AlN, GaN, and InN quantum dots as the active material of the laser structure. The numerical results were compared, confirming that the InN quantum dot in the active layer significantly improved the absorption and emission phenomena of the quantum dot laser. At the same time, InN is capable of emitting light at a wavelength of 1.55 $\mu$ m due to its band gap of 0.7eV. Consequently, InN has a superior potentiality to other existing materials to fabricate the quantum dot laser operating at the 1.55 $\mu$ m wavelength, which is the window of an optical fiber communication system that offers the lowest attenuation. Therefore, the InN quantum dot laser is expected to be a promising candidate not only in the field of semiconductor technology but also in the field of optical communication in the imminent future.

*Keywords-density of state; feedback level; frequency fluctuation; momentum relaxation time; laser*

## I. INTRODUCTION

Nowadays researchers focus on the absorption and emission phenomena of Quantum Dot (QD) as a top priority, which has led to its use as monolithic coherent light source such as lasers, electroabsorption modulators, and saturable absorber mirrors [1]. Lasers fabricated from such an active layer exhibit higher device performance compared to the traditional semiconductor lasers based on bulk or Quantum Well (QW), or Quantum Wire

(QWR) active medium. Improvements in Quantum Dot Laser (QDL) performance have appeared in wide modulation bandwidth, relative noise intensity, minimum lasing threshold, output power temperature sensitivity, reduced linewidth enhancement factors, reduced crosstalk between amplified signals at low power level, and improved four-wave mixing at high power levels and longer device lifetime [2-5]. Currently, the absorption properties of QDL such as effective density of state, carrier density at the threshold, level of optical feedback,

rate of change of mobility of carriers, and emission properties have become the focus of numerous theoretical and experimental investigations for further improvement [3, 6]. The stability of laser characteristics depends on the structural parameters and alterations of the density of the carriers and photons along the length of the cavity structure and its dependence on the current above-threshold conditions [7]. Among the absorption properties, effective density of state and mirror reflectivity are the most important parameters assisting optical feedback, which plays an important role in ensuring more uniformity in the mobility (reduction of the rate of change of mobility) of the carrier and finally in improving the stability of the semiconductor laser frequency. The active region of the laser can be designed in the form of three-dimensional bulk materials, two-dimensional QWLs, single-dimensional QWRs, or zero-dimensional QDs [8]. Recently, it has been discussed that enhancement of device characteristics is expected for lasers with lower dimensionality of the active region. This reduction in the dimensionality was achieved by introducing QWR and QD as the active layer material of the laser structure. In QD, the motion of the carriers is restricted in all three directions, which helps overcome the problems of QWL and other available categories of material [8, 9].

The existing materials used to design QDL are ternary  $\text{Ga}_{47}\text{In}_{53}\text{As}$  [10] and quaternary  $\text{InGaAsN}$  [11],  $\text{InGaAsN}$  [9],  $\text{GaN}$ , and  $\text{AlN}$  [10]. As these materials are difficult to grow and not compatible with  $1.55\mu\text{m}$  emission, researchers are looking for alternative materials. Recent findings showed that  $\text{InN}$  has a band gap of around  $0.7\text{eV}$ , which is compatible with  $1.55\mu\text{m}$  emission [12]. The emission energy depends on many parameters, both on the physical properties of QD, such as shape, dimensions, strain, and chemical compositions. Additionally, the emission energy depends on the composition and the residual strain of the confinement layers as well as the ground-quantized energy levels of carriers. These parameters depend not only on the design of the structure but also on growth conditions [11]. For long-wavelength operations at  $1.55\mu\text{m}$ ,  $\text{InN}$  QDL is predicted to be suitable for long-distance communication with high-performance and low-power consumption light sources for optical fiber systems [12]. The growth technique is very vital as the operating wavelength of the laser varies with the active layer thickness. The wavelength variations and tenability of the device are estimated to be about  $0.15\text{nm}$  [5].

This study investigates the effective density of state, carrier density at threshold, feedback level of the gain medium, uniformity of carrier mobility, the stability of laser frequency, and the minimization of absorption loss. These major absorption and emission properties were investigated using  $\text{GaN}$ ,  $\text{AlN}$ , and  $\text{InN}$  QD in the active layer of the laser structure. The numerical results were compared to analyze the improvement of these phenomena of QDL using  $\text{InN}$ . The result comparison revealed that the effective density of state, carrier density at threshold, and the feedback level of QDL are higher by using  $\text{InN}$  QD than  $\text{GaN}$ 's and  $\text{AlN}$ 's in the active layer of the laser structure. Furthermore, the rate of change of the mobility of the carrier and fluctuation of the laser frequency along with the absorption loss were also minimized

significantly by employing  $\text{InN}$  QD in the active layer of the laser structure.

## II. MATERIALS AND METHODS

### A. Device Structure

A laser structure composed of an  $\text{InN}$ -based quantum dot in the active layer was considered to investigate the improvement of the cavity length dependence on the laser characteristics operating at the  $1.55\mu\text{m}$  wavelength. The details of the laser structure with  $\text{InN}$ -based quantum dot are shown in Table I [4].

TABLE I. LAYER STRUCTURE OF INN-BASED QDL

p GaN contact layer (77 nm)
P $\text{Al}_{0.13}\text{Ga}_{0.87}\text{N}$ upper cladding layer (1000 nm)
p $\text{In}_{0.82}\text{Ga}_{0.18}\text{N}$ guiding layer (117 nm)
<b>InN undoped QDs active layer (2.7 nm)</b>
n $\text{In}_{0.82}\text{Ga}_{0.18}\text{N}$ guiding layer (117 nm)
n $\text{Al}_{0.13}\text{Ga}_{0.87}\text{N}$ Lower cladding layer (1000 nm)
n- GaN contact layer (77 nm)
C sapphire(0001) substrate

The structure of the  $\text{InN}$ -based QDL consists of an  $\text{InN}$ -plane sapphire wafer (oriented along 0001 direction) as the substrate along with a  $77\text{nm}$  thick  $\text{n+GaN}$  contact layer, a  $1000\text{nm}$  thick  $\text{n Al}_{0.13}\text{Ga}_{0.87}\text{N}$  lower cladding layer, a  $117\text{nm}$  thick  $\text{n In}_{0.82}\text{Ga}_{0.18}\text{N}$  guiding layer, a  $2.7\text{nm}$  thick  $\text{InN}$  active region with single layer undoped quantum dots, a  $117\text{nm}$  thick  $\text{p In}_{0.82}\text{Ga}_{0.18}\text{N}$  guiding layer, a  $1000\text{nm}$  thick  $\text{p Al}_{0.13}\text{Ga}_{0.87}\text{N}$  upper cladding layer, and a  $77\text{nm}$  thick  $\text{p+GaN}$  contact layer. Thus, the mean size of the quantum dot was  $2.7\text{nm}$  and the optical confinement layer thickness was  $236.7\text{nm}$ . The FIB etching technique was used to form Fabry-Perot cavity mirrors by etching facets, providing high-quality etched facets. A Fabry-Perot resonator is composed of two highly reflective mirrors that allow small portions of light to pass through while reflecting most of the light back through the active region, where it can be further replicated through stimulated emission. A pair of parallel planes (or facets) was etched. Under the appropriate biasing conditions, the laser light would be emitted from these planes. The two remaining sides were roughened to eliminate lasing in any direction other than the main.

### B. Numerical Analysis

This section presents the numerical analysis of some of the major characteristics of the QDL. These QDL characteristics were analyzed using  $\text{InN}$ ,  $\text{GaN}$ , and  $\text{AlN}$  QD as the active layer material of the laser structure. The characteristics of the laser analyzed in this study were the effective density of state of the carriers, momentum relaxation, absorption loss, optical feedback level, and oscillation frequency fluctuation.

#### 1) Effective Density of State

The QD semiconductor materials have discrete energy levels. Due to these discrete energy levels, symmetric emission of light is obtained, which is predicted to be one of the subjects of great current interest to enhance the sensitivity of QDL to those optical feedback levels. QDL has gained greater importance after significant progress in nanostructure growth

through the self-assembling technique [13]. The effective density of state per unit volume in the QD heterostructure is given by [14]:

$$\rho = 2 \left( \frac{m_e \pi K_B T}{h^2} \right)^{\frac{3}{2}} \quad (1)$$

where  $m_e$  is the electron's effective mass,  $K_B$  is Boltzmann's constant,  $h$  is Planck's constant, and  $T$  is the temperature.

## 2) Momentum Relaxation and Absorption Loss

The mobility of the cladding layer is an important parameter, which affects the minority carrier leakage current. The mobility of the carrier is related to the momentum relaxation time by [5]:

$$\mu(T) = \mu_0 \left( \frac{\tau_0}{\tau(T)} \right) \quad (2)$$

The free carrier absorption loss is given by [14]:

$$\alpha_{fc} = \alpha_{fc_0} \left( \frac{\tau(T)}{\tau_0} \right) \quad (3)$$

where  $\alpha_{fc_0}$  is the free carrier absorption loss at room temperature,  $\tau$  is the momentum relaxation time, and  $T$  is the temperature. The temperature dependence of momentum relaxation time is given by [13]:

$$\tau(T) = \frac{\tau_0}{1 + \beta \left( \frac{T - T_0}{T_1 - T_0} \right)} \quad (4)$$

where  $\tau_0$  is the momentum relaxation time at room temperature,  $T_0$  is the lower limit of the operating temperature, and  $T_1$  is the upper limit of the operating temperature range. To investigate the effect of temperature variation on the mobility of the carrier and the carrier absorption loss, using (4) in (2) and (3) gives:

$$\mu(T) = \mu_0 \left[ 1 + \beta \left( \frac{T - T_0}{T_1 - T_0} \right) \right] \quad (5)$$

$$\alpha_{fc} = \frac{\alpha_{fc_0}}{1 + \beta \left( \frac{T - T_0}{T_1 - T_0} \right)} \quad (6)$$

## 3) Carrier Density at Threshold

This section investigates the phenomenon of carrier density at threshold. The gain coefficient and the photon lifetime have a significant role in the carrier densities at threshold [15]. The relationship between the photon lifetime and the carrier density threshold is given by:

$$\rho_{th} = \rho_0 + \frac{1}{\tau_p G} \quad (7)$$

where  $\rho_0$  is the carrier density at transparency,  $\tau_p$  is the photon lifetime, and  $G$  is the modal gain coefficient.

## 4) Optical Feedback Level

The optical feedback level of a laser depends on several parameters, has further effects on the operating characteristics of a diode laser, and is given by:

$$\gamma = \frac{\sqrt{R_{eff}}}{\tau_{active}} \quad (8)$$

where  $\tau_{active}$  is the round-trip time in the active region, given by:

$$\tau_{active} = \frac{2L}{v_g} \quad (9)$$

where  $L$  is the length of the cavity and  $v_g$  is the group velocity of the light wave in the gain medium of the laser.

## 5) Fluctuation of Laser Frequency

The fluctuation of laser frequency depends strongly on the refractive index and is given by [16]:

$$\delta\omega = - \frac{E_g}{h n_g} \Delta n_{eff} \quad (10)$$

where  $E_g$  is the band gap energy of the active layer material,  $h$  is Planck's constant,  $\Delta n_{eff}$  is the total effective refractive index variation due to plasma and the effect of QD, and  $n_o$  is the refractive index of the active layer material. In this equation, the minus sign indicates the decrement in the laser frequency. Now, taking the modulus on both sides of (5) gives:

$$|\delta\omega| = \frac{E_g}{h n_g} \Delta n_{eff} \quad (11)$$

## III. RESULTS AND ANALYSIS

This section presents the results obtained using the numerical analysis mentioned above. The improvement of QDL's characteristics using InN was achieved in terms of higher effective density of state of the carriers, enhanced optical feedback level, reduced frequency fluctuation, and reduced rate of change momentum relaxation and mobility as well as minimization of absorption loss.

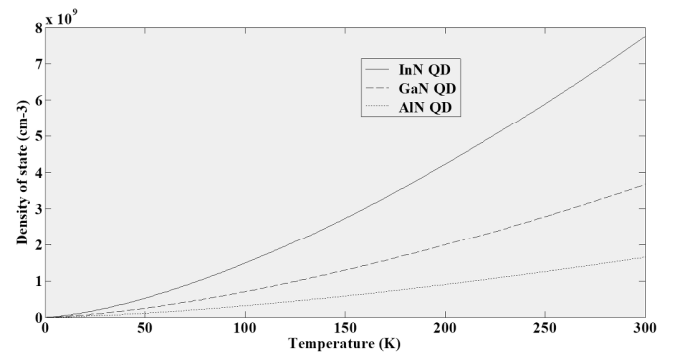


Fig. 1. Effect of temperature on the effective density of state of QDL using InN, GaN, and AlN QD as active layer material of the laser.

Figure 1 shows the effect of temperature on the effective density of state of the carriers of QDL using InN, GaN, and AlN QD as the active layer material of the laser structure. The effective density of state increased with a temperature increase for any material. InN QDL experienced an increase of  $2 \times 10^9 \text{ cm}^{-3}$  in the density of state, from  $5.9 \times 10^9$  to  $7.9 \times 10^9 \text{ cm}^{-3}$  over the temperature range of 250-300K. GaN QDL experienced an increase of  $7 \times 10^8 \text{ cm}^{-3}$ , from  $2.9 \times 10^9$  to  $3.6 \times 10^9 \text{ cm}^{-3}$ , and AlN experienced an increase of  $7 \times 10^8 \text{ cm}^{-3}$ , from  $1.2 \times 10^9$  to  $1.5 \times 10^9$  within the same temperature range. Remarkably, InN QDL experienced the highest upward trend in the density of state, showing that a prominent enhancement of the effective density of state was achieved using InN, and therefore, it more promising to provide higher effective density of state than other existing materials.



Figure 2 shows the temperature dependence of the momentum relaxation time using InN, GaN, and AlN QD as the active material of the laser. InN QDL experienced a decrease of 0.0125nsec in the momentum relaxation time, from 0.755 to 0.5625nsec due to a change in temperature from 250 to 300K. GaN QDL experienced a decrease of 0.025nsec, from 0.5 to 0.475nsec, and AlN experienced a decrease of momentum relaxation time of 0.025nsec, from 0.4 to 0.375nsec within the same temperatures. Therefore, InN showed the highest momentum relaxation time at any temperature among the three materials. Thus, using InN QD as active layer material offers a longer time for momentum relaxation, which helps achieving uniformity of carrier mobility.

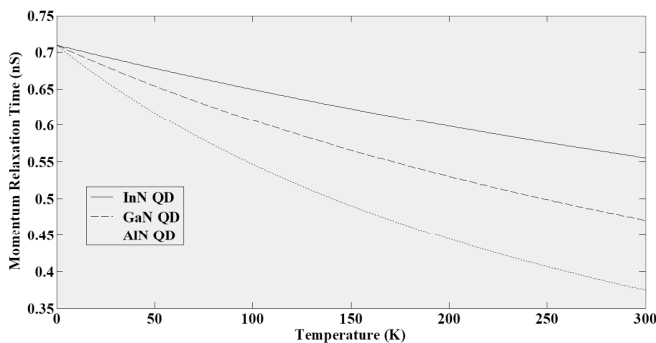


Fig. 2. Temperature dependence of momentum relaxation time using InN, GaN, and AlN QD as active layer material of the laser structure.

Figure 3 demonstrates the temperature dependence of the mobility of the carrier using InN, GaN, and AlN QD as the active layer material of the laser structure. AlN QDL experienced an increase of  $59\text{cm}^2\text{V}^{-1}\text{s}^{-1}$  in mobility, from 1066 to  $1125\text{cm}^2\text{V}^{-1}\text{s}^{-1}$  over the temperature range of 250-300K. GaN QDL experienced an increase of  $33.34\text{cm}^2\text{V}^{-1}\text{s}^{-1}$  in mobility, from 866.66 to  $900\text{cm}^2\text{V}^{-1}\text{s}^{-1}$  and InN experienced an increase of  $25\text{cm}^2\text{V}^{-1}\text{s}^{-1}$  in mobility, from 750 to  $775\text{cm}^2\text{V}^{-1}\text{s}^{-1}$  within the same temperature range. InN showed the lowest rate of change of mobility concerning temperature among the three materials used in the simulation, providing strong evidence that it is the most promising to provide the lowest rate of change in mobility with an increase in temperature.

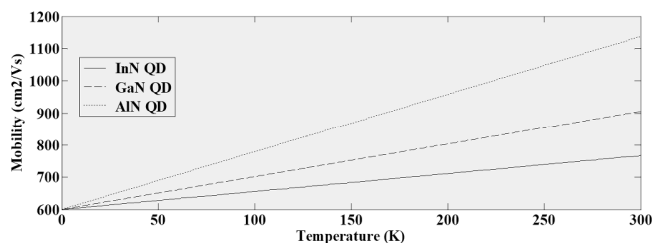


Fig. 3. Temperature dependence of the mobility of carrier using InN, GaN, and AlN QD as active layer material of the laser structure.

Figure 4 shows the temperature dependence of the absorption loss using InN, GaN, and AlN QD as the laser's active material. InN had the lowest absorption loss at any temperature, as it experienced a decrease of  $0.135\text{cm}^{-1}$  in

absorption loss, from  $0.925$  to  $0.79\text{cm}^{-1}$  during a temperature change from 250 to 300K. GaN QDL experienced a decrease of  $0.095\text{cm}^{-1}$ , from  $0.9$  to  $0.805\text{cm}^{-1}$ , and AlN experienced a decrease in absorption loss of  $0.0075\text{cm}^{-1}$ , from  $0.81$  to  $0.8025\text{cm}^{-1}$  within the same range. Therefore, InN QDL had the lowest absorption loss and the slope of the solid curve is minimal, indicating that the absorption loss had the least temperature dependence.

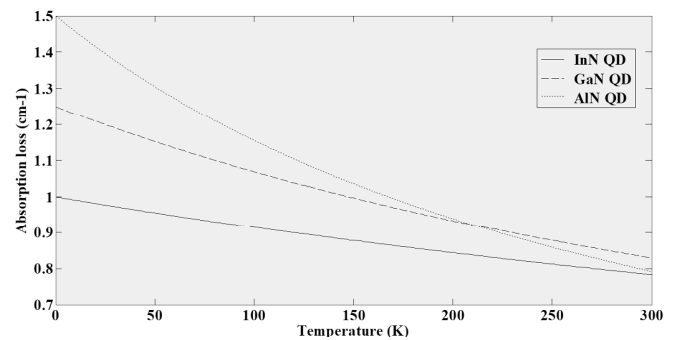


Fig. 4. Temperature dependence of absorption loss using InN, GaN, and AlN QD as active layer material of the laser structure.

Figure 5 represents the carrier density at threshold of QDL as a function of photon lifetime, which decreases exponentially with the increase of photon lifetime. InN QDL experienced a decrease of  $1.525 \times 10^8\text{cm}^{-3}$  in carrier density at threshold, from  $5.215 \times 10^9$  to  $5.0625 \times 10^9\text{cm}^{-3}$  due to the change in photon lifetime from 0.5 to 1s. GaN QDL experienced a decrease of  $10^8\text{cm}^{-3}$ , from  $5.175 \times 10^9$  to  $5.075 \times 10^9\text{cm}^{-3}$ , and AlN experienced a decrease of  $3.75 \times 10^7\text{cm}^{-3}$ , from  $5.125 \times 10^9$  to  $5.0875 \times 10^9\text{cm}^{-3}$  within the same temperature range. Among the three materials used, InN showed the highest density of carrier at any cavity length and is the most promising in providing a higher effective density of state as well as the highest carrier density at threshold, which enhances the amplification of light.

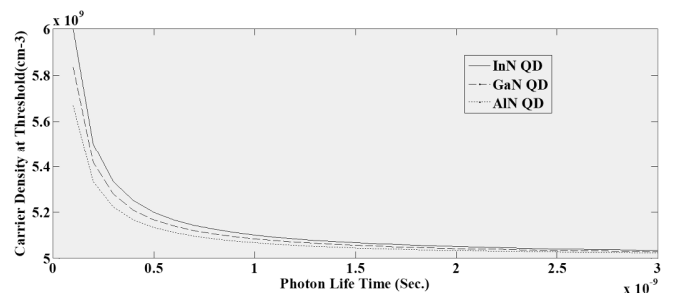


Fig. 5. Dependence of carrier density at the threshold on photon lifetime using InN, GaN, and AlN QD as active layer material of the laser structure.

Figure 6 describes the cavity length dependence of the optical feedback level of QDL, which reduces as the cavity length increases for GaN, AlN, and InN as the active layer material. InN QDL experienced a decrease of 0.2 (20%) in the feedback level, from 0.3727 to 0.1727 due to the change in cavity length from 40 to  $100\mu\text{m}$ . GaN QDL experienced a decrease of 0.164 (16.4%) from 0.309 to 0.145, and AlN experienced a decrease of 0.16 (16%) from 0.28 to 0.12 within

the same cavity range. As InN QD in the active layer of the laser provided the highest feedback level for any cavity length, it is the most promising in providing an enhanced level of feedback along with the highest carrier concentration and effective density of state.

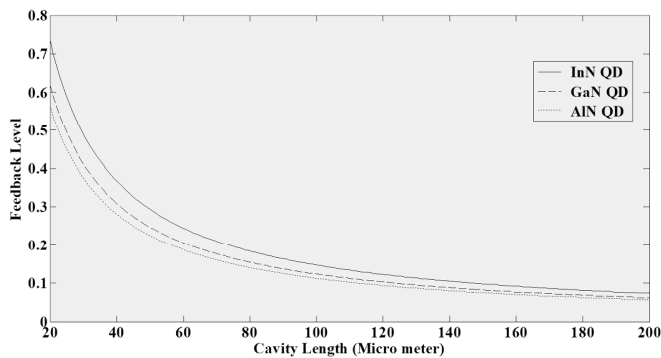


Fig. 6. Dependence of feedback level of laser on cavity length using InN, GaN, and AlN QD as active layer material of the laser structure.

Figure 7 illustrates the effect of the refractive index on the fluctuation of laser oscillation frequency using InN, GaN, and AlN QD as the active layer material of the laser. InN, GaN, and AlN QDL experienced a frequency fluctuation of  $1.8 \times 10^7$  Hz,  $6 \times 10^7$  Hz, and  $8.9 \times 10^7$  Hz, for a 1.5 refractive index. The refractive index value was considered for the most widely used silica fibers around the globe. The fluctuation of the oscillation frequency of QDL is higher for AlN and GaN compared to InN. Additionally, the slope of the curve for InN QD was flatter than the other materials, indicating that the stability of the frequency of the laser output was improved significantly.

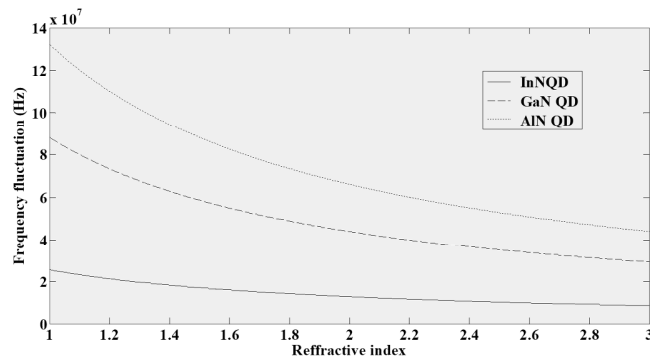


Fig. 7. Effect of refractive index on the fluctuation of laser frequency using InN, GaN, and AlN QD as active layer material of the laser structure.

#### IV. DISCUSSION

Among the three communication windows of optical fiber communication systems (0.89  $\mu$ m, 1.3  $\mu$ m, and 1.55  $\mu$ m), 1.55  $\mu$ m offers the lowest attenuation, greater repeater spacing, and higher bit rate. These phenomena made it possible to use coherent optical sources compatible with the standard silicon fibers used in optical fiber communication [17]. Semiconductor sources and detectors with Group III-V compounds in active layers have been studied extensively and used almost

exclusively for the present light-wave communication systems in these three wavelength regions [19-21]. Among the Group III nitride trios, InN has a unique band gap of 0.7 eV, which is compatible with 1.55  $\mu$ m. In addition, InN QDL achieved the highest effective density of state of the carriers and reduced frequency fluctuation, rate of change momentum relaxation, and mobility as well as minimized absorption loss.

#### V. CONCLUSION

This study investigated extensively the major absorption and emission phenomena of QDL, focusing mainly on the absorption loss and stability of laser frequency and their key parameters. The effective density of state, carrier density at threshold, feedback level, and rate of change of mobility are the most important phenomena to enhance the stability of the oscillation frequency of QDL and minimize absorption loss. Considering the results of the numerical analysis, it can be concluded that InN provides higher effective density of state, enhanced feedback level, and higher uniformity in the mobility of the carrier. The improvement of these key parameters significantly minimized the fluctuation of laser frequency and absorption loss. Therefore, InN is a promising material for enhancing the absorption and emission properties of QDL, like higher stability of laser oscillation frequency and reduction of absorption loss.

#### ACKNOWLEDGMENT

The authors acknowledge the partial financial support provided by the Department of Electrical and Electronic Engineering of Eastern University, Dhaka, Bangladesh.

#### REFERENCES

- [1] M. Smit et al., "An introduction to InP-based generic integration technology," *Semiconductor Science and Technology*, vol. 29, no. 8, Mar. 2014, Art. no. 083001, <https://doi.org/10.1088/0268-1242/29/8/083001>.
- [2] F. L. Wang, X. W. Ma, Y. Z. Huang, Y.-D. Yang, J. Y. Han, and J. L. Xiao, "Relative intensity noise in high-speed hybrid square-rectangular lasers," *Photonics Research*, vol. 6, no. 3, pp. 193–197, Mar. 2018, <https://doi.org/10.1364/PRJ.6.000193>.
- [3] M. T. Chughtai, "A Realization of Stabilizing the Output Light Power from a Laser Diode: A Practical Approach," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7370–7374, Aug. 2021, <https://doi.org/10.48084/etasr.4276>.
- [4] B. Liu, Y. Jiang, and H. Ji, "Sensing by Dynamics of Lasers with External Optical Feedback: A Review," *Photonics*, vol. 9, no. 7, Jul. 2022, Art. no. 450, <https://doi.org/10.3390/photonics9070450>.
- [5] M. Seifouri and A. Faraji, "Simulation of the Optimized Structure of a Laterally Coupled Distributed Feedback (LC-DFB) Semiconductor Laser Above Threshold," *Engineering, Technology & Applied Science Research*, vol. 3, no. 5, pp. 522–525, Oct. 2013, <https://doi.org/10.48084/etasr.376>.
- [6] M. A. Humayun, M. A. Rashid, F. Malek, and Y. Namihiro, "Minimization of the performance degradation of laser parameters with operating time of quantum dot laser using indium nitride," *Physics Essays*, vol. 28, no. 1, pp. 134–137, Mar. 2015, <https://doi.org/10.4006/0836-1398-28.1.134>.
- [7] M. Seifouri, F. Shahshahani, and A. Faraji, "Numerical Analysis of SHB Effects in Laterally-Coupled Distributed Feedback (LC-DFB) Lasers," *Engineering, Technology & Applied Science Research*, vol. 2, no. 5, pp. 273–277, Oct. 2012, <https://doi.org/10.48084/etasr.196>.
- [8] F. P. García de Arquer, D. V. Talapin, V. I. Klimov, Y. Arakawa, M. Bayer, and E. H. Sargent, "Semiconductor quantum dots: Technological

- progress and future challenges," *Science*, vol. 373, no. 6555, Aug. 2021, Art. no. eaaz8541, <https://doi.org/10.1126/science.aaz8541>.
- [9] R. Carron, "Dilute-Nitride Low-Dimensional Nanostructures Formed on Non-Planar Substrates," Ph.D. dissertation, EPFL, Lausanne, Switzerland, 2013.
- [10] R. Pawlus, L. L. Columbo, P. Bardella, S. Breuer, and M. Gioannini, "Intensity noise behavior of an InAs/InGaAs quantum dot laser emitting on ground states and excited states," *Optics Letters*, vol. 43, no. 4, pp. 867–870, Feb. 2018, <https://doi.org/10.1364/OL.43.000867>.
- [11] S. Franchi, G. Trevisi, L. Seravalli, and P. Frigeri, "Quantum dot nanostructures and molecular beam epitaxy," *Progress in Crystal Growth and Characterization of Materials*, vol. 47, no. 2, pp. 166–195, Jan. 2003, <https://doi.org/10.1016/j.pcrysgrow.2005.01.002>.
- [12] M. A. Humayun, M. A. Rashid, F. A. Malek, and A. N. Hussain, "Effect of lattice constant on band-gap energy and optimization and stabilization of high-temperature InGa1-xN quantum-dot lasers," *Journal of Russian Laser Research*, vol. 33, no. 4, pp. 387–394, Jul. 2012, <https://doi.org/10.1007/s10946-012-9294-7>.
- [13] D. O'Brien, S. P. Hegarty, G. Huyet, and A. V. Uskov, "Sensitivity of quantum-dot semiconductor lasers to optical feedback," *Optics Letters*, vol. 29, no. 10, pp. 1072–1074, May 2004, <https://doi.org/10.1364/OL.29.001072>.
- [14] S. Kasap, *Principles of Electronic Materials and Devices*, 3rd ed. New York, NY, USA: McGraw-Hill, Inc., 2005.
- [15] T. Ito *et al.*, "Femtosecond pulse generation beyond photon lifetime limit in gain-switched semiconductor lasers," *Communications Physics*, vol. 1, no. 1, pp. 1–8, Aug. 2018, <https://doi.org/10.1038/s42005-018-0045-0>.
- [16] M. Gioannini and I. Montrosset, "Numerical Analysis of the Frequency Chirp in Quantum-Dot Semiconductor Lasers," *IEEE Journal of Quantum Electronics*, vol. 43, no. 10, pp. 941–949, Jul. 2007, <https://doi.org/10.1109/JQE.2007.904306>.
- [17] J. M. Senior and M. Y. Jamro, *Optical Fiber Communications: Principles and Practice*. Harlow, UK: Pearson Education, 2009.
- [18] M. M. Hossain, M. A.-A. Humayun, M. T. Hasan, A. G. Bhuiyan, A. Hashimoto, and A. Yamamoto, "Proposal of High Performance 1.55 $\mu$ m Quantum Dot Heterostructure Laser Using InN," *IEICE Transactions on Electronics*, vol. E95.C, no. 2, pp. 255–261, 2012, <https://doi.org/10.1587/transele.E95.C.255>.
- [19] M. A. Humayun, M. A. Rashid, F. A. Malek, A. Hussain, and I. Daut, "Design of Quantum Dot Based LASER with Ultra-Low Threshold Current Density," *Applied Mechanics and Materials*, vol. 229–231, pp. 1639–1642, 2012, <https://doi.org/10.4028/www.scientific.net/AMM.229-231.1639>.
- [20] M.-K. Lee, C.-H. Chu, Y.-H. Wang, and S. M. Sze, "1.55- $\mu$ m and infrared-band photoresponsivity of a Schottky barrier porous silicon photodetector," *Optics Letters*, vol. 26, no. 3, pp. 160–162, Feb. 2001, <https://doi.org/10.1364/OL.26.000160>.
- [21] M. A. Humayun, M. N. Hasan, M. A. Rashid, A. Kuwana, and H. Kobayashi, "Effect of optical fiber core diameter on Brillouin scattering loss," *Semiconductor Physics, Quantum Electronics and Optoelectronics*, vol. 24, no. 4, pp. 450–456, Nov. 2021, <https://doi.org/10.15407/spqeo24.04.450>.

## AUTHOR PROFILES



**Md. Abdullah Al Humayun** was born in 1987 and received his Ph.D. in Electrical Systems Engineering from Malaysia Perlis University in 2015. His research work focused on semiconductor device design. He is currently working as an Associate Professor in the Department of Electrical & Electronic Engineering at the Eastern University, Dhaka, Bangladesh. Previously, he was a postdoctoral fellow at the International Islamic University of Malaysia and worked in different private universities in Bangladesh. He has authored over 30 publications. The area of his scientific interests includes physics and technology of semiconductor materials, hetero and hybrid structures, and optoelectronic devices (lasers, LEDs, solar cells, photodetectors, etc.), and their applications in optical fiber communication.



**Mohd Abdur Rashid** (member, IEEE) received his Ph.D. degree in Electrical and Information Engineering from the University of the Ryukyus, Japan. He is a Professor in the Department of Electrical and Electronic Engineering (EEE) at NSTU. He is currently working as a visiting scholar at Gunma University, Japan. He has also work experience in Malaysia, Japan, and Canada in academic fields. He has authored more than 96 technical articles in journals and conferences. His research interests include multidisciplinary fields, including mathematical modeling, electronic devices, and biomedical engineering.



**Anna Kuwana** (associate member, IEEE) received her B.S. and M.S. degrees in information science and her Ph.D. from Ochanomizu University, in 2006, 2007, and 2011, respectively. She worked at Ochanomizu University as a member of the technical staff. She is currently an Assistant Professor in the Division of Electronics and Informatics at Gunma University. Her research interests include computational fluid dynamics and signal analysis.



**Haruo Kobayashi** (senior member, IEEE) received his B.S. and M.S. degrees in information physics from the University of Tokyo in 1980 and 1982, respectively, his M.S. in electrical engineering from UCLA in 1989, and his Ph.D. in electrical engineering from Waseda University in 1995. He joined Yokogawa Electric Corporation, Tokyo, Japan, in 1982, and was involved in research and development related to measuring instruments. In 1997, he joined Gunma University, where he is currently a Professor in the Division of Electronics and Informatics. His research interests include mixed-signal integrated circuit design and testing, and signal processing algorithms.

# Improved Torque Ripple of Switched Reluctance Motors using Sliding Mode Control for Electric Vehicles

**Vo Quang Vinh**

Faculty of Control and Automation, Electric Power University, Vietnam  
vinhvq@epu.edu.vn

**Vo Thanh Ha**

Faculty of Electrical and Electronics Engineering, University of Transport and Communications, Vietnam  
vothanhha.ktd@utc.edu.vn  
(corresponding author)

*Received: 13 December 2022 | Revised: 27 December 2022 and 5 January 2023 | Accepted: 7 January 2023*

## ABSTRACT

This study describes the direct torque control-DTC approach, based on the Sliding Mode Control (SMC) technology with chattering reduction, for reducing the torque ripple of the Switched Reluctance Motor (SRM). The SRM torque control loop has been given the SMC treatment to account for the low-frequency fluctuations in the torque output. To maintain a consistent motor speed, the sliding mode controller modifies the value of the reference current. The findings demonstrate that the constant sliding mode controller is superior to PI controllers at lowering the motor's torque ripple, compensating for its nonlinear torque characteristics, and rendering the drive insensitive to parameter changes. MATLAB/SIMULINK simulation has been used to show how well this SMC performs. The performance of the proposed SMC method has been demonstrated by simulation in MATLAB/SIMULINK with a three-phase 8/6 pole, and a 2kW SRM.

**Keywords**-switched reluctance motor; Sliding Mode Control (SMC); PI; EV; torque ripple

## I. INTRODUCTION

Switched Reluctance Motors (SRMs) have become a viable electrification option for the military, civic, agricultural, mining, and transportation industries, due to their low prices, high efficiency, and ability to run in hostile settings [1-3]. However, the fact that SRMs are vulnerable to torque ripple, which makes noise and shortens the motor's life, has been a problem [4-6]. The primary explanation is that SRMs align with magnetic fields using low-reluctance materials like iron and steel. The stator of the SRM features phased windings, while the rotor has low and high reluctance areas. The magnetic flux and materials' resistance creates a force that pushes the rotor pole into alignment with the closest stator pole when power is applied to the stator windings. The SRM control algorithms sequentially turn the stator windings on and off to change the magnetic field, which forces the rotor to revolve. The problem is that because of structural deformations and magnetic torque harmonics that cause the stator and rotor to interact, this process also generates vibration and noise. As a result, noise is produced, and the motor's lifetime is shortened by the interactions and oscillations that alter the torque or torque ripple [7-13]. This SRM torque ripple limits the

possibility of their direct-drive applications in the industry. Significantly, the fastness and accuracy of the torque response the engine needs to mobilize during driving are affected considerably. Therefore, structure and control strategy optimization consist the primary aim of research on torque ripple reduction. In terms of structural optimization, the torque ripple was reduced by optimizing the pole shape [14-16] and stepper air gap [17] using the Finite Element Modeling (FEM) in conjunction with optimization techniques. The improvement in control is often more flexible than the optimization in construction since SRMs have a straightforward structure. Numerous studies use sophisticated control techniques to reduce torque ripple [18], such as optimal harmonic current injection and current waveform tuning. In [19, 20], a PID fuzzy logic controller with the integration of the speed error, the sum of the speed error, and the derivative of the speed error are presented to decrease the torque ripple of SRM with two inputs.

A comparison of SMC and fuzzy-neural network approaches was reported in [21]. SMC is comparable to other popular nonlinear control techniques, including adaptive input-output feedback linearization [17] and adaptive back stepping

techniques [22]. Multiple industrial control systems have been used to study the development of SMC approaches [23-26]. Minimization of torque ripple in switched reluctance motor based on Model Predictive Control (MPC) and Torque Sharing Function (TSF) was conducted in [27]. A novel continuous terminal SMC technique has been devised to ensure that the system states reach the sliding surface in a finite amount of time. The suggested controller certifies robustness and continuity, making the control algorithm better suited for mechanical servo systems.

In this study, the speed loop is controlled by an SMC and a PI controller to enhance the performance of the DTC system by adequately determining the reference torque, further achieving accurate control, minimizing torque ripple, and improving the dynamic performance of the SRM.

## II. MODELING OF AN SRM MOTOR

The nonlinear, switching currents are used in the SRM to drive the phases. By correctly modeling the nonlinear flux-current-angle ( $\lambda$ - $i$ - $\theta$ ) characteristics of the machine, the magnetic nonlinearities of an SRM may be taken into consideration [25]. The three - phase 8/6 SRM is shown in Figure 1.

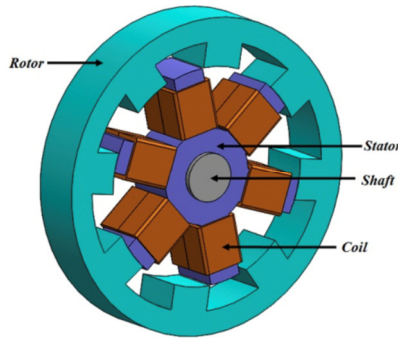


Fig. 1. The three - phase 8/6 SRM.

$$\frac{di}{dt} = \left( u - Ri - \frac{d\lambda(\theta, i)}{d\theta} \frac{d\theta}{dt} \right) \left( \frac{d\lambda(\theta, i)}{di} \right)^{-1} \quad (1)$$

The nonlinear torque dynamics are described by:

$$\frac{dT_e}{dt} = m + nu \quad (2)$$

where:

$$n = \left( \frac{dT_e}{di} \right) \left( Ri - \frac{d\lambda(\theta, i)}{d\theta} \frac{d\theta}{dt} \right) \left( \frac{d\lambda(\theta, i)}{di} \right)^{-1} + \left( \frac{dT_e}{d\theta} \frac{d\theta}{dt} \right) \quad (3)$$

$$m = \left( \frac{dT_e}{di} \right) \left( \frac{d\lambda(\theta, i)}{dt} \right)^{-1} \quad (4)$$

where  $T_e$  is the torque SRM motor,  $\lambda(\theta, i)$  is the nonlinear function of phase flux linkage,  $\theta$  is the rotor position, and  $i$  is the stator current.

The rate of energy change with rotor position may be used to express the SRM torque.

$$T_e(\theta, i) = \frac{1}{2} i^2 \frac{dL}{d\theta} \quad (5)$$

where  $L$  is the inductance at a position,  $\theta_r$  is the rotor position,  $\theta_1, \theta_2, \theta_3, \theta_4$  are the rotor positions at a phase (A, B, C, and D phases),  $L_{min}$  and  $L_{max}$  represent the minimum/maximum phase inductance with none aligned / aligned position.

$$L(\theta) = \begin{cases} L_{min} & \text{if } \theta_r < \theta_1 \\ L_{min} + \frac{dL}{d\theta}(\theta_r - \theta_1) & \text{if } \theta_1 < \theta_r < \theta_2 \\ L_{max} + \frac{dL}{d\theta}(\theta_r - \theta_1) & \text{if } \theta_2 < \theta_r < \theta_3 \end{cases} \quad (6)$$

## III. SLIDING MODE CONTROLLER DESIGN FOR SPEED CONTROL

The technique used in SMC is intrinsically resistant to changes in parameters, nonlinear models, outside disturbances, and uncertainty. Therefore, it is employed when the robustness requirement is crucial for in-vehicle applications, and there are significant uncertainties [22]. The following procedures are part of the torque controller design by the sliding mode controller.

- Step 1: Select the sliding surface function:

$$S = e_{Te}(t) + P_t \int_0^t e_{Te}(t) dt = 0 \quad (7)$$

where  $e_{Te}(t) = T_e^* - T_e$  is the error torque and  $P_t$  is a positive gain coefficient.

- Step 2: Taking the derivative of (7) and substituting from (2). The result of the sliding surface is expressed as:

$$\frac{dS}{dt} = -(m + nu) + P_t e_{Te} \quad (8)$$

- Step 3: The Lyapunov technique is often utilized to determine the conditions on the control rule that will cause the state to circle the equilibrium in SMC.

$$W = \frac{1}{2} S^T S \geq 0 \quad (9)$$

Then, the derivative of (8) and  $W$  is definitely negative without the sliding surface  $S=0$ . Therefore, the control law is:

$$u = n^{-1} \{ -m + P_t e_{Te} + \gamma \text{sgn}(S) \} \quad (10)$$

where  $\gamma$  is the positive gain coefficient and  $\text{sgn}(S)$  is the switch function. It is selected as:

$$\text{sgn}(S) = \begin{cases} 1, & \text{if } S > n \\ \frac{S}{n}, & \text{if } |S| \leq n \\ -1, & \text{if } S < -n \end{cases} \quad (11)$$

- Step 4: Based on (5) and (10), the control law is:

$$u = \frac{L_{min} + P_t \theta}{P_t} \left\{ -\frac{P_t i}{L_{min} + P_t \theta} \left( -Ri - P_t i \frac{d\theta}{dt} \right) + P_t e_{Te} + \gamma \text{sgn}(S) \right\} \quad (12)$$

Equation (12) has the  $\text{sgn}$  function which causes system chattering. Therefore, the main disadvantage in conventional SMC is chattering. Chattering adversely affects the performance of system significantly. The chattering in the sliding mode controller can be reduced by modifying the

control law in as  $u = k \text{sat}(\frac{s}{\theta})$  and constant factor  $\theta$  defines the thickness of the boundary layer around the switching surface. The  $\text{sat}(s/\theta)$  is the saturation function that is defined as:

$$\text{sat}\left(\frac{s}{\theta}\right) = \begin{cases} \frac{s}{\theta} & \text{if } \left|\frac{s}{\theta}\right| \leq 1 \\ \text{sgn}\left(\frac{s}{\theta}\right) & \text{if } \left|\frac{s}{\theta}\right| > 1 \end{cases} \quad (13)$$

#### IV. SIMULATION RESULTS

Simulations of the DTC system were conducted to check the proposed SMC controller's accuracy and efficacy. Figure 2 depicts the block diagrams of the SMC controller and the DTC system. MATLAB/Simulink was used to model the SRM drive's DTC system with a three-phase, 2-kW SRM. Simulations were run in the 2022 version of MATLAB/SIMULINK. The closed-loop SRM drive speed controller based on SMC in Figures 2, 3, and the SRM motor have the parameters shown in Table I.

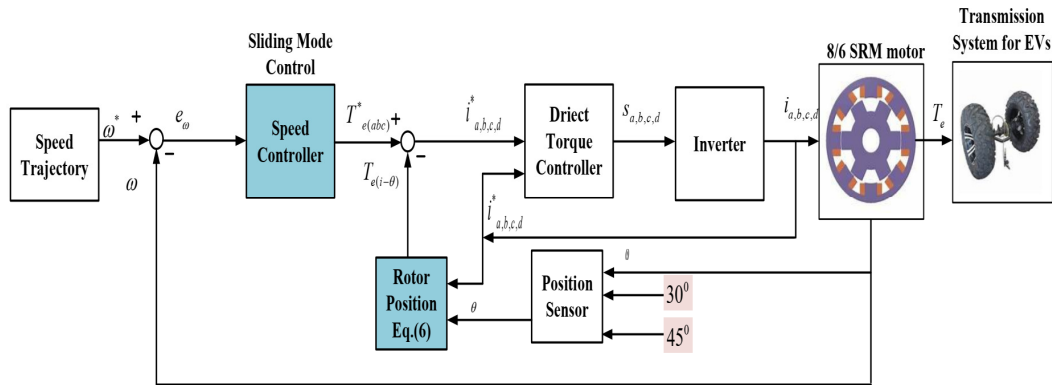


Fig. 2. Structure control of an improved torque ripple of an SRM for electric vehicles.

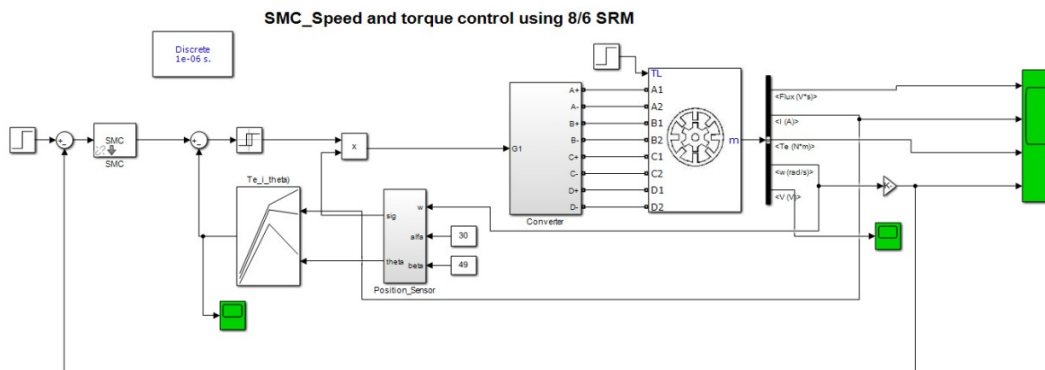


Fig. 3. Structure control reducing the torque ripple of the SRM motor in MATLAB/SIMULINK.

TABLE I. SRM MOTOR PARAMETERS

Parameters	Unit
Stator resistance ( $\Omega$ )	2
Inertia (kg.m.n)	0.082
Frication (N.m.s)	0.01
Inductance $L_{min}$	11
Inductance $L_{max}$	50
Stator pole number	8
Rotor pole number	6
Maximum current (A)	120
Maximum flux linkage (V.s)	0.486

The electromagnetic torque of the SRM with the classical PI controller and the proposed SMC for 1000 and 4000rpm are shown in Figure 4. The most significant advantage of the proposed SMC is its constant electromagnetic torque. Figure 4

shows the simulation results of the SRM at rotational speeds of 1000 and 4000rpm and a reference torque of 38Nm.

In addition, the SRMs have a significant ripple of electromagnetic torque. However, the average torque fluctuation is within 2.5%. Therefore, the speed ripple of the proposed SMC is smaller than that of a PI controller. Furthermore, the suggested SMC's electromagnetic torque is more uniform than the PI's. Therefore, reducing the electromagnetic torque ripple is the essential advantage of the proposed controller.

Figures 5 and 6 show the change of the flux linkage which should be made as smooth as possible when designing the torque, enhancing the movement following the ability of the SMC controller.



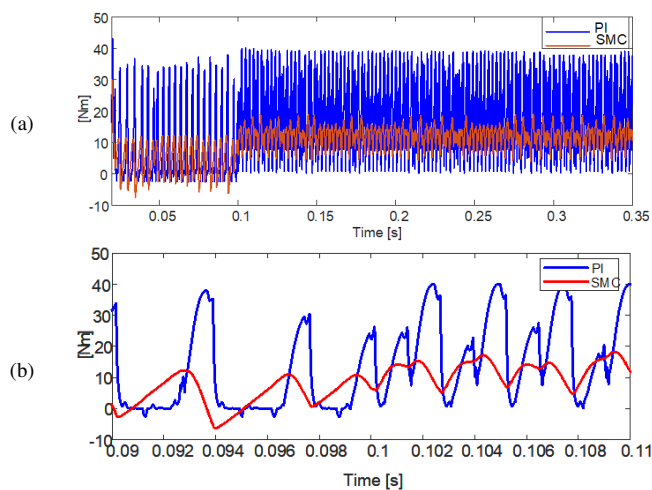


Fig. 4. Torque comparison, SMC and PI controllers. (a) 1000rpm, (b) 4000rpm.

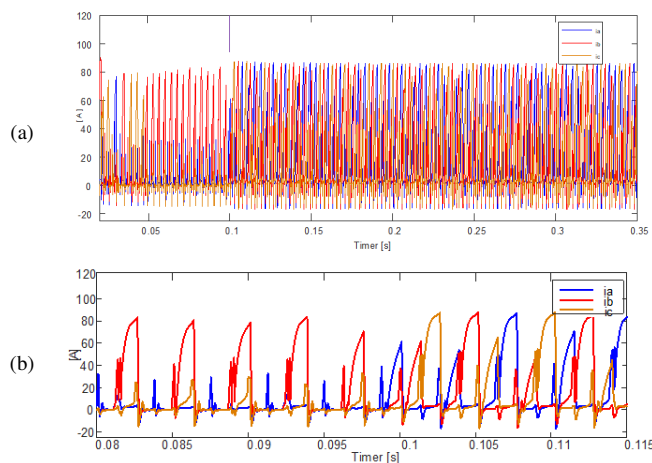


Fig. 5. The three-phase current response. (a) 1000rpm, (b) 4000rpm.

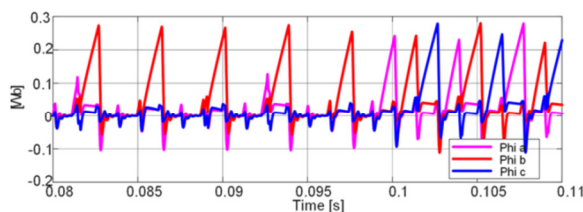


Fig. 6. The flux response.

Figure 7 shows that the actual speed response follows the set value. The short time is fast with  $t = 0.025s$ , and the static deviation is slight with  $\pm 1\%$ . Therefore, the proposed SMC method has the advantage of very small speed error of the SRM motor. Let's consider the case where the inductance parameter changes by 50%. The torque response is shown in Figure 8 and it does not change at 1000 and 4000rpm. Again, the short time is fast with  $t = 0.025s$ , and the static deviation is slight with  $\pm 1\%$ . In contrast, the increased torque response of the PI control method is 23%, and the setting time is slower than that of the SMC controller.

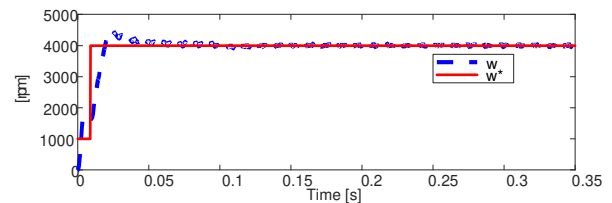


Fig. 7. Simulation results of the SRM control speed using SMC.

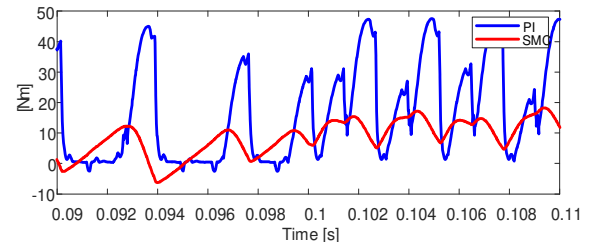


Fig. 8. Torque response when the stator resistance is increased by 50%.

## V. CONCLUSIONS

This paper proposes a torque ripple minimization of switched reluctance motors using direct torque control based on the sliding mode. The proposed torque control method combines DTC with the SMC controller. The stability of the proposed system is verified through the Lyapunov function. In addition, this control method has reduced the chattering of the torque response by using a saturation part in the control signal. As a result, the proposed control method performs better than the PI controller and it is robust to parameter changes.

To improve the torque ripple reduction in future research, a sliding controller combined with a fuzzy logic controller or connected with a neural network or a model predictive control, will be considered to optimize the duty ratio and to minimize the torque ripple. In addition, this control method should be tested experimentally in the future to prove its correctness.

## ACKNOWLEDGMENT

The authors would like to acknowledge the support from the University of Transport and Communications.

## REFERENCES

- [1] M. Takeno, A. Chiba, N. Hoshi, S. Ogasawara, M. Takemoto, and M. A. Rahman, "Test Results and Torque Improvement of the 50-kW Switched Reluctance Motor Designed for Hybrid Electric Vehicles," *IEEE Transactions on Industry Applications*, vol. 48, no. 4, pp. 1327–1334, Jul. 2012, <https://doi.org/10.1109/TIA.2012.2199952>.
- [2] X. Ling, J. Tao, B. Li, C. Qin, and C. Liu, "A Multi-Physics Modeling-Based Vibration Prediction Method for Switched Reluctance Motors," *Applied Sciences*, vol. 9, no. 21, Jan. 2019, Art. no. 4544, <https://doi.org/10.3390/app9214544>.
- [3] C. Gan, J. Wu, Q. Sun, W. Kong, H. Li, and Y. Hu, "A Review on Machine Topologies and Control Techniques for Low-Noise Switched Reluctance Motors in Electric Vehicle Applications," *IEEE Access*, vol. 6, pp. 31430–31443, 2018, <https://doi.org/10.1109/ACCESS.2018.2837111>.
- [4] E. Bostanci, M. Moallem, A. Parsapour, and B. Fahimi, "Opportunities and Challenges of Switched Reluctance Motor Drives for Electric Propulsion: A Comparative Study," *IEEE Transactions on Transportation Electrification*, vol. 3, no. 1, pp. 58–75, Mar. 2017, <https://doi.org/10.1109/TTE.2017.2649883>.

- [5] K. Lu, P. O. Rasmussen, S. J. Watkins, and F. Blaabjerg, "A New Low-Cost Hybrid Switched Reluctance Motor for Adjustable-Speed Pump Applications," *IEEE Transactions on Industry Applications*, vol. 47, no. 1, pp. 314–321, Jan. 2011, <https://doi.org/10.1109/TIA.2010.2091189>.
- [6] C. Mademlis and I. Kioskeridis, "Gain-Scheduling Regulator for High-Performance Position Control of Switched Reluctance Motor Drives," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 9, pp. 2922–2931, Sep. 2010, <https://doi.org/10.1109/TIE.2009.2038400>.
- [7] X. D. Xue, K. W. E. Cheng, and S. L. Ho, "Optimization and Evaluation of Torque-Sharing Functions for Torque Ripple Minimization in Switched Reluctance Motor Drives," *IEEE Transactions on Power Electronics*, vol. 24, no. 9, pp. 2076–2090, Sep. 2009, <https://doi.org/10.1109/TPEL.2009.2019581>.
- [8] D.-H. Lee, J. Liang, Z.-G. Lee, and J.-W. Ahn, "A Simple Nonlinear Logical Torque Sharing Function for Low-Torque Ripple SR Drive," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 8, pp. 3021–3028, Dec. 2009, <https://doi.org/10.1109/TIE.2009.2024661>.
- [9] X. D. Xue, K. W. E. Cheng, and N. C. Cheung, "Evaluation of torque sharing functions for torque ripple minimization of switched reluctance motor drives in electric vehicles," in *2008 Australasian Universities Power Engineering Conference*, Sydney, NSW, Australia, Sep. 2008.
- [10] V. D. Nhan, N. X. Bien, N. Q. Dich, and V. T. Ha, "An Experimental Study on the Control of Slotless Self-Bearing Motor Using Nonlinear Control," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8942–8948, Aug. 2022, <https://doi.org/10.48084/etasr.4997>.
- [11] A. Guediri and S. Touil, "Modeling and Comparison of Fuzzy-PI and Genetic Control Algorithms for Active and Reactive Power Flow between the Stator (DFIG) and the Grid," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8640–8645, Jun. 2022, <https://doi.org/10.48084/etasr.4905>.
- [12] N. Zerroug, K. Behih, Z. Bouchama, and K. Zehar, "Robust Adaptive Fuzzy Control of Nonlinear Systems," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8328–8334, Apr. 2022, <https://doi.org/10.48084/etasr.4781>.
- [13] W. Shang, S. Zhao, Y. Shen, and Z. Qi, "A Sliding Mode Flux-Linkage Controller With Integral Compensation for Switched Reluctance Motor," *IEEE Transactions on Magnetics*, vol. 45, no. 9, pp. 3322–3328, Sep. 2009, <https://doi.org/10.1109/TMAG.2009.2021264>.
- [14] T. Shi, L. Niu, and W. Li, "Torque-ripple minimization in switched reluctance motors using sliding mode variable structure control," in *Proceedings of the 29th Chinese Control Conference*, Beijing, China, Jul. 2010, pp. 332–337.
- [15] S. K. Sahoo, S. Dasgupta, S. K. Panda, and J.-X. Xu, "A Lyapunov Function-Based Robust Direct Torque Controller for a Switched Reluctance Motor Drive System," *IEEE Transactions on Power Electronics*, vol. 27, no. 2, pp. 555–564, Oct. 2012, <https://doi.org/10.1109/TPEL.2011.2132740>.
- [16] D.-H. Kim, H.-G. Jeong, and K.-B. Lee, "Torque ripple minimization of switched reluctance motors based on fuzzy logic and sliding mode control," in *2013 IEEE International Symposium on Industrial Electronics*, Feb. 2013, <https://doi.org/10.1109/ISIE.2013.6563640>.
- [17] N. Abut, B. Cakir, N. Inanc, A. B. Yildiz, and M. Z. Bilgin, "Switched reluctance motor drive by using fuzzy controller," in *Proceedings of the IECON'97 23rd International Conference on Industrial Electronics, Control, and Instrumentation (Cat. No.97CH36066)*, New Orleans, LA, USA, Aug. 1997, vol. 1, pp. 348–353, <https://doi.org/10.1109/IECON.1997.671075>.
- [18] W. Liu and S. Song, "Application of Fuzzy Control in Switched Reluctance Motor Speed Regulating System," in *2006 International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*, Aug. 2006, pp. 72–72, <https://doi.org/10.1109/CIMCA.2006.48>.
- [19] S. Paramasivam and R. Arumugam, "Hybrid fuzzy controller for speed control of switched reluctance motor drives," *Energy Conversion and Management*, vol. 46, no. 9, pp. 1365–1378, Jun. 2005, <https://doi.org/10.1016/j.enconman.2004.08.007>.
- [20] X. Zhang, Q. Yang, M. Ma, Z. Lin, and S. Yang, "A Switched Reluctance Motor Torque Ripple Reduction Strategy With Deadbeat Current Control and Active Thermal Management," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 317–327, Jan. 2020, <https://doi.org/10.1109/TVT.2019.2955218>.
- [21] M. A. J. Kondelaji and M. Mirsalim, "Segmented-Rotor Modular Switched Reluctance Motor With High Torque and Low Torque Ripple," *IEEE Transactions on Transportation Electrification*, vol. 6, no. 1, pp. 62–72, Mar. 2020, <https://doi.org/10.1109/TTE.2020.2969356>.
- [22] H. Hou, X. Yu, L. Xu, K. Rsetam, and Z. Cao, "Finite-Time Continuous Terminal Sliding Mode Control of Servo Motor Systems," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 7, pp. 5647–5656, Jul. 2020, <https://doi.org/10.1109/TIE.2019.2931517>.
- [23] H. Ríos, R. Falcón, O. A. González, and A. Dzul, "Continuous Sliding-Mode Control Strategies for Quadrotor Robust Tracking: Real-Time Application," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1264–1272, Oct. 2019, <https://doi.org/10.1109/TIE.2018.2831191>.
- [24] X. Zhang, L. Sun, K. Zhao, and L. Sun, "Nonlinear Speed Control for PMSM System Using Sliding-Mode Control and Disturbance Compensation Techniques," *IEEE Transactions on Power Electronics*, vol. 28, no. 3, pp. 1358–1365, Mar. 2013, <https://doi.org/10.1109/TPEL.2012.2206610>.
- [25] J. Castro, P. Andrada, and B. Blanqué, "Minimization of torque ripple in switched reluctance motor drives using an enhanced direct instantaneous torque control," in *2012 XXth International Conference on Electrical Machines*, Marseille, France, Sep. 2012, pp. 1021–1026, <https://doi.org/10.1109/ICEIMach.2012.6350001>.
- [26] N. Inanc and V. Ozbulur, "Torque ripple minimization of a switched reluctance motor by using continuous sliding mode control technique," *Electric Power Systems Research*, vol. 66, no. 3, pp. 241–251, Sep. 2003, [https://doi.org/10.1016/S0378-7796\(03\)00093-2](https://doi.org/10.1016/S0378-7796(03)00093-2).
- [27] P. Ren, J. Zhu, Z. Jing, Z. Guo, and A. Xu, "Minimization of torque ripple in switched reluctance motor based on MPC and TSF," *IEEE Transactions on Electrical and Electronic Engineering*, vol. 16, no. 11, pp. 1535–1543, 2021, <https://doi.org/10.1002/tee.23458>.

# The Influence of Slenderness in Steel and Composite Columns under Fire Conditions

**Farid Boursas**

Laboratory of Applied Civil Engineering (LGCA), Department of Civil Engineering, University of Tebessa, Algeria  
farid.boursas@univ-tebessa.dz  
(corresponding author)

**Rafik Boufarh**

Laboratory of Applied Civil Engineering (LGCA), Department of Civil Engineering, University of Tebessa, Algeria  
rafik.boufarh@univ-tebessa.dz

**Bilal Maghaghi**

Department of Civil Engineering, Mohamed Boudiaf University of M'sila, Algeria  
bilal.maghaghi@univ-msila.dz

*Received: 14 December 2022 | Revised: 2 January 2023 | Accepted: 5 January 2023*

## ABSTRACT

This study simulated two columns using the finite element software SAFIR. The first was a steel profile column and the second was a steel profile partially encased in a concrete column (SPPEC). Both columns were heated on four sides for one hour using the ISO834 standard fire curve. Two boundary conditions were considered for both columns, simply and doubly supported and loaded utilizing eccentric loading. A parametric analysis aimed to identify a thermal analysis of unprotected and protected steel columns in terms of temperature field distribution, and time-temperature curves of a few selected nodes were drawn from the SAFIR software. A thermomechanical study was also carried out using the results of the thermal analysis to determine the influence of slenderness on the resistance of the steel and the partially encased profile in concrete columns under fire solicitation. Finally, a comparison of the fire resistance of the two columns was made. The results proved that the slenderness negatively influences the fire resistance of the steel and composite columns and that the behavior of the composite columns is significantly better than the steel ones.

**Keywords-**steel profile column; SPPEC column; fire curve ISO834; slenderness; thermomechanical study

## I. INTRODUCTION

Nowadays, modeling of structures under fire conditions allows the evaluation of a structure's behavior and its resistance to fire without resorting to tests. Columns are structural elements that ensure structural stability. A fire can severely affect the structural performance of columns [1], and several theoretical approaches and tests have been proposed to examine the fire resistance of columns. A simplified method was proposed in [2] to examine the resistance of an element to fire conditions, specifying that the resistance to fire (uniform temperature) is ensured if the temperature of the element after a fire is lower than its critical temperature. A procedure to determine the critical temperature was proposed according to the EN 1993-1-2 standard [3] and presented according to different stresses on the element (element bent without risk of overturning, element compressed or bent with risk of

overturning, compressed element bent). This approach made it possible to check bare or protected elements, while if the element is protected, this procedure evaluated the thickness of the protective layer. In [4], the virtual work principle was used to establish a reinforced concrete beam element model under the influence of high temperature. The proposed model was evaluated under experimental tests on a frame. This configuration was chosen for parametric analysis, including the time factor. The study noted that the axial stress induces an axially compressive force which increases the deflection at the mid-span of the beam when the temperature increases, and the more the axial stress increases, the more the deflection increases. On the other hand, the deflection decreases if torsional rigidity increases. The experimental study presented in [5] investigated the influence of the circular shape on the behavior under fire conditions, describing the test procedure, observations, and the obtained fire resistance values.

Several theoretical methods have been developed and successfully applied for the rapid, safe, and efficient design of reinforced concrete columns under fire conditions. In [6], an experimental study was presented for the influence of spacing on the fire resistance of HSC circular columns under axial loading. The tested columns were modeled in the axis of symmetry with integrated reinforcements. The applied material model considered the influence of transient temperatures on the mechanical properties of concrete and steel, and a parametric study was carried out on the effect of cracking, the development of transient creep deformations, and plastic deformations for concrete. The results indicated that during the entire heating period, high thermal gradients generated tensile stresses in the section plane of the columns. The effects of axial and rotational stresses on the fire response of Concrete-Filled Tubular (CFT) columns were investigated in [7], using the fiber model to simulate the fire behavior of CFT columns in portal frames. A parametric study was carried out to analyze the influence of the retaining frame's adverse effects on the reaction to fire of CFT columns. The restraint thermal elongation induced restraint forces that negatively affected the column, however, a beneficial effect was produced by maintaining rotation which positively modified the boundary conditions of the heated column. In addition, another favorable effect came from the progressive redistribution of the internal forces as the heated column lost its mechanical capacity. In a second step, the current provisions given by Eurocode 4 Part 1.2 [8] were analyzed together with those given in the UK National Annex to Eurocode 4. The behavior of Concrete-Filled Steel Tubes (CFSTs) for slender columns subjected to eccentric loading under fire conditions was investigated in [9], simulating the destruction process of CFST slender columns under an ISO-834 standardized fire, and comparing the numerical to the experimental results. The comparison showed that the proposed model predicted the fire resistance of CFST slender columns subjected to a eccentric load. A parametric study was carried out to investigate the effect of load centering ratio, steel yield strength, concrete compressive strength, and CFST column reinforcement ratio under fire conditions.

The fire resistance of columns with double-filling concrete CFDST was studied in [10], using Finite Element Analysis (FEA) models and comparing the results with the experimental ones from the literature. Based on the developed numerical model, a study of the parameters affecting the resistance design of CFDST columns was carried out, using the experimental data to validate the accuracy of the proposed method and showing that it can predict the resistance of CFDST columns with reasonable accuracy. A three-dimensional Finite Element Model was proposed in [11] to predict the fire resistance of steel tubes filled with CFST cast-in-place concrete with preload, and was validated by experimental tests. The model was used to predict the fire resistance time of CFST columns with different slenderness, load, and preload ratios. The results showed that the preload of the steel tube had little influence on the fire resistance of short CFST columns, while the influence of the preload on the fire resistance can be significant on larger slenderness ratios and a further increase in slenderness ratio exceeding a certain range can reduce the effect of preload. In [12], a mathematical model using the fiber approach was

presented to calculate the resistance of eccentrically loaded Concrete-Filled thin Steel Tubes (CFST) under fire stress. The model used thermal simulation to determine the temperature distribution in the cross sections and nonlinear global buckling analysis. This model included the initial geometry imperfection, air space between the concrete and the steel tube, tensile strength of the concrete, deformations caused by preloads, and the behavior of materials as a function of temperature and it was examined against experimental tests and existing numerical results.

In [13], a study on the fire performance of square-section steel tubes filled with reinforced concrete columns (SRCFST) was presented, using ABAQUS to establish the numerical models of the members of the SRCFST to simulate fire behavior. The models were validated against the results of experimental tests in terms of temperature field, axial strain versus time curves, fire resistance, and failure modes. The fire performance of the SRCFST columns, including temperature distributions, axial strain-time ( $\Delta$ -t) curves, axial force distribution, and deformation, were investigated and compared with common CFST columns. The results showed that the fire behavior of SRCFST columns can be further improved with the integrated steel profile. Finally, the influences of various parameters on fire resistance, including fire load ratio, material strength, steel tube ratio, profile steel ratio, and slenderness ratio, were examined. In [14], eight full-scale Concrete-Encased Steel (CES) columns made of C120 concrete and an S500/S690 high-strength steel section were tested under concentrated load and heated under an ISO834 fire to failure, evaluating the fire resistance of each CES column. The axial displacement-time curve, the temperature-time curve, the angle of rotation of the end support, the post-fire conditions of the concrete surface, and the mode of rupture were examined. The experimental study showed that the addition of polypropylene fiber was effective in minimizing the explosive concrete spalling in high-strength CES columns to achieve a fire resistance time comparable to normal-strength CES columns. Furthermore, recommendations were provided for modifying the current methods for the fire-resistant design of high-strength concrete CES columns.

This study focused on the thermal and mechanical behavior of structural elements idealized by a steel column or partially encased with a concrete column (SPPEC) in the nonlinear domain. As composite columns differ from steel due to the combined action between concrete and steel that gives higher strength and more realistic behavior [15], the SPPEC column is challenging due to its sensitivity to parameters such as the cross-section size, material modeling, and boundary conditions [16]. To study the response of structures under the influence of fire, the SAFIR finite element program [17] developed at the University of Liège in Belgium solves the general Fourier heat transfer equation:

$$\left(k \frac{\partial^2 T}{\partial x^2} + K \frac{\partial^2 T}{\partial y^2}\right) + Q = \rho c \frac{\partial T}{\partial t} \quad (1)$$

Several methods can be used to transform the heat transfer equation into a form suitable for finite element analysis. The most common is the weighted residual method which requires that the weighted averages of the residuals is equal to zero:

$$\int_x w_i(x) e(x) dx = 0 \quad (2)$$

where  $w_i(x)$  is a set of weighting functions. To take convection into account, a modification must be made to the basic formulation already developed. Convection is quantitatively described by Newton's law of cooling:

$$q = h(T_s - T_\infty) \quad (3)$$

Applying it to this problem produces:

$$\begin{aligned} & \sum_{i=1}^n \left[ \int_A k \left( \frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} \right) dA \right] T_i \\ & + \sum_{i=1}^n \left[ \int_S h N_i N_j dS \right] T \\ & = \int_A Q N_j dA - \int_S q N_j dS + \int_S h T_\infty N_j dS \end{aligned} \quad (4)$$

as structural response analysis. The bar element is used to model the columns, and the stiffness matrix is:

$$K = K_s + K_u \quad (5)$$

Different deformations (mechanical and thermal) were considered in the numerical simulation. The total strain rate is the sum of the elastic strain rate, the plastic strain rate, the thermal strain rate, and the thermo-mechanical interaction rate:

$$\dot{\varepsilon} = \dot{\varepsilon}^\theta + \dot{\varepsilon}^P + \dot{\varepsilon}^\theta + \dot{\varepsilon}^{tm} \quad (6)$$

## II. THERMAL STUDY OF STEEL COLUMNS

For this study, an unprotected IPE120 steel column was solicited by eccentric loading of 5KN, according to different heights. Figure 1 shows the dimensions of the studied IPE120.

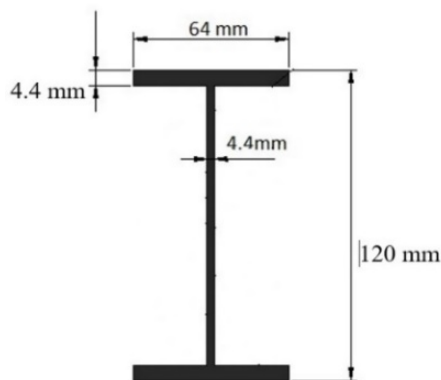


Fig. 1. Steel profile dimensions.

The discretization was based on a triangular meshing; which allowed to obtain 282 elements composed of 226 nodes for the steel section, as shown in Figure 2. Figure 3 shows the temperature evolution within the cross-section of the IPE120 exposed to an ISO-834 fire in the four accessible faces. This flow distribution shows a decreasing temperature towards the center of the web. For this purpose, three points were targeted:

- Node 1 on the solicited side of the upper flange
- Node 33 in terms of joint web-flange
- Node 109 in the center of the web

Figure 4 shows the representative curves of these points. The unprotected IPE 120 element reached a temperature of 60°C in 85s on the three nodes. Its temperature increased parabolically up to 1600s, and then rose linearly up to 940°C until 3600s. It can be noted that the three nodes take the same temperature assessment in the cross-section of the column solicited according to the four faces and that comes back to the finesse of the flange thickness and the chosen web.

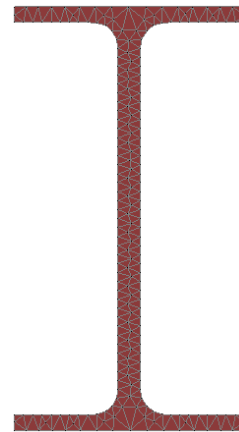


Fig. 2. Discretization of steel profile.

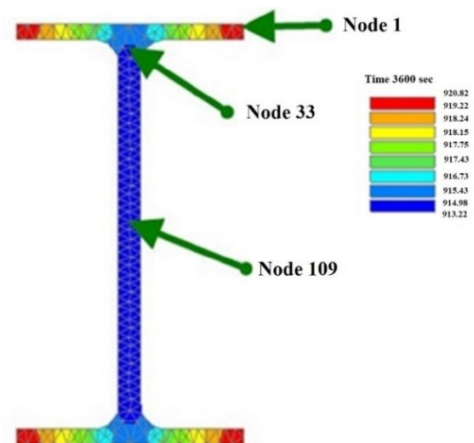


Fig. 3. Temperature distribution after 3600 sec of exposition.

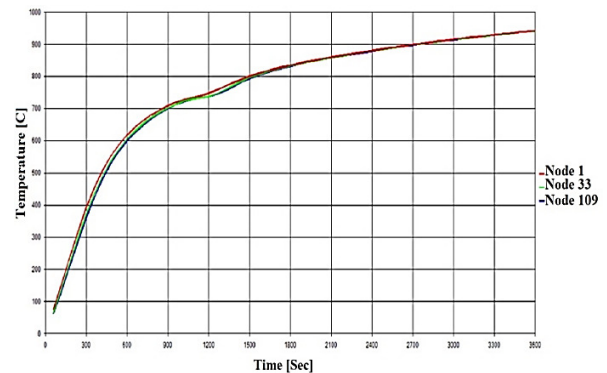


Fig. 4. Obtained temperatures vs time for nodes 1, 33, and 109.

### III. THERMAL STUDY OF COMPOSITE COLUMNS

A steel profile was partially encased with a concrete column (SPPEC) and solicited by eccentric loading of 45kN according to different heights, as shown in Figure 5.

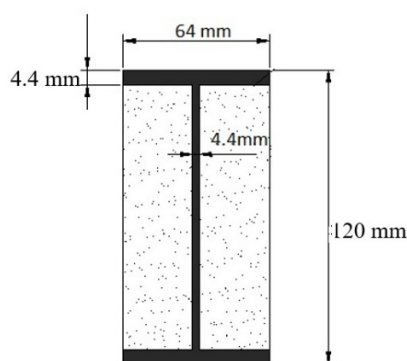


Fig. 5. SPPEC column dimensions.

The discretization was based on a triangular mesh, allowing to obtain the steel section partially coated with concrete, and consisted of 608 nodes or 1112 elements, as shown in Figure 6.

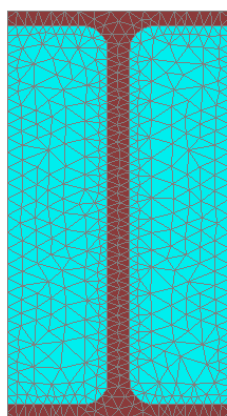


Fig. 6. Discretization of SPPEC column.

Similarly, as with the IPE120 column, the studied column was exposed to fire on four sides. The heat flow shows the formation of a symmetrical evolution compared to IPE120, and the temperatures varied between 932.7°C and 853.40°C. For this purpose, three points were targeted and their representative curves are shown in Figures 6 and 7. These chosen points were:

- Node 241 on the solicited face of the lower flange
- Node 33 on the exposed face of concrete
- Node 450 in the center of the cross-section

As shown in Figure 8, the temperature increased parabolically from 90s, where node 33 reached a temperature of 70°C, node 241 reached 40°C, and node 450 reached 20°C. After 3600s, the center of the section reached a temperature of 853.40°C, the side of the lower flange reached a temperature of

922.79°C, and the third node of the solicited face of concrete reached a temperature of 932.70°C.

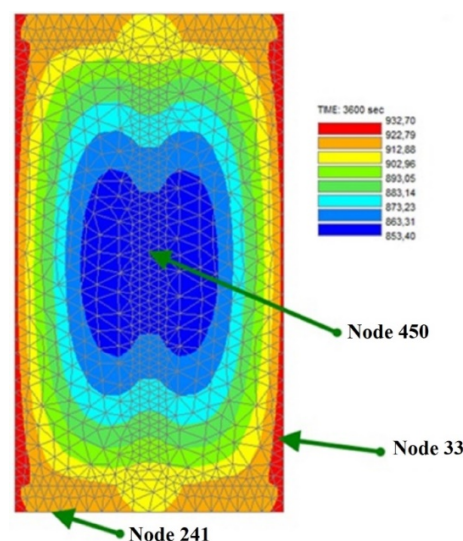


Fig. 7. Temperature distribution after 3600s of exposition.

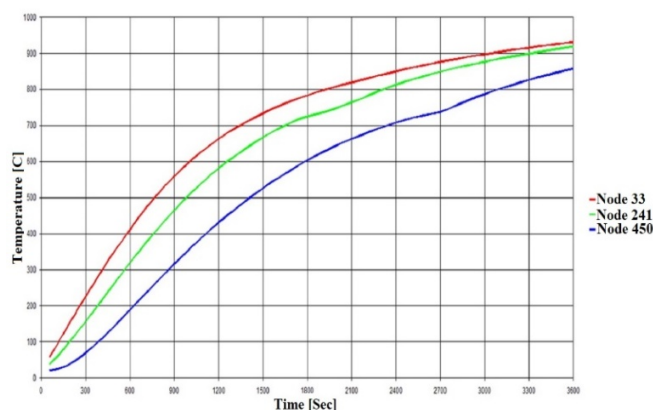


Fig. 8. Obtained temperatures vs time for nodes 33, 241, and 450.

### IV. MECHANICAL ANALYSIS

The mechanical behavior of a steel column and another composite bi-articulated with varied height (5m to 9m) was examined. The steel column was subjected to an eccentric load of 5kN, and the steel profile partially encased with the concrete column (SPPEC) was subjected to an eccentric load of 45kN. With:

$$F = \frac{\text{load applied}}{\text{limit load at room temperature}} = 50\%$$

The studied columns were discretized in 21 nodes and 10 elements, as shown in Figure 9. The slenderness parameter was evaluated to analyze its influence on the fire behavior of steel and composite columns.



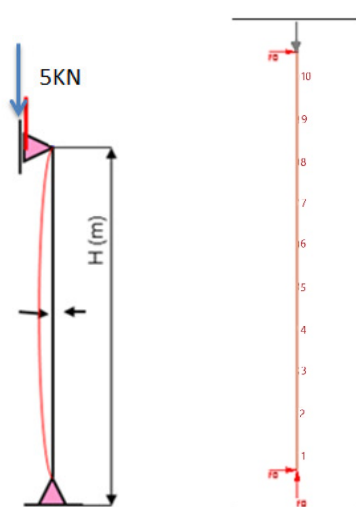


Fig. 9. Considered columns scheme.

TABLE I. FIRE RESISTANCE OF THE COLUMNS AS FUNCTION OF HEIGHT

Section type Steel profile Mixed	Fire resistance (Ref/s)				
	H=5m	H=6m	H=7m	H=8m	H=9m
Steel column	187	100	30.5	16.25	13.75
Composite column	2621	1989	1739	1580	1446

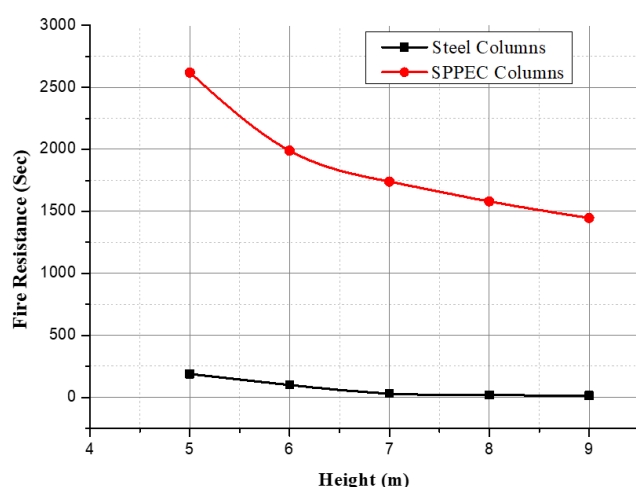


Fig. 10. Influence of columns' height for IPE120 protected in concrete and unprotected on fire resistance.

Figure 10 shows the fire resistance of IPE120 in both cases as a function of the height of the columns. Increasing the height of the columns hurts their resistance to fire. When the height of a column is increased, the fire resistance decreases for the two studied sections. Similarly, the temperature variation for the IPE120 profile has a low variation compared to the composite section. The influence of the protection of the steel profiles is very important in the mechanical behavior in terms of fire resistance.

## V. CONCLUSION

This study performed a numerical analysis of a steel profile column and a steel profile partially encased with concrete column (SPPEC) under fire conditions, using the SAFIR software based on the finite element method. The results showed the impact of the slenderness parameter on the fire resistance of this kind of column. From the results of this study, the following conclusions may be drawn:

- The numerical investigation indicated that the variation of temperature in the steel profile exposed to fire on the four sides is very minimal, and it undergoes a degradation of its mechanical characteristics abruptly and can reach failure.
- The protection of the steel profile with concrete in SPPEC columns creates a significant temperature difference within the cross-section after exposure to fire and causes the failure of the element to not be sudden.
- The increase in the height of the steel profile column significantly affects its fire resistance.
- The protection of steel profile columns with concrete has a positive and considerable influence on the fire resistance of IPE120 partially coated with concrete, and its resistance is 2621s.
- The increase in the height of the two types of columns harms their fire resistance. Increasing the height of the column results in a decrease in the fire resistance in the two section columns.

## ACKNOWLEDGMENT

This research was supported by Algeria's Directorate-General for Scientific Research and Technological Development (DGRSDT). The authors express special thanks to the University of Echahid Cheikh Larbi Tebessi and the Laboratory of Applied Civil Engineering (LGCA). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors.

## REFERENCES

- [1] P. C. Nguyen, D. D. Pham, T. T. Tran, and T. Nghia-Nguyen, "Modified Numerical Modeling of Axially Loaded Concrete-Filled Steel Circular-Tube Columns," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7094–7099, Jun. 2021, <https://doi.org/10.48084/etasr.4157>.
- [2] C. Renaud, "Calcul simplifié de la température critique selon la norme EN 1993-1-2," *Construction métallique*, no. 3, pp. 89–118, 2012.
- [3] "Eurocode 3: Design of steel structures - Part 1-2: General rules - structural fire design," European Committee for Standardization, Brussels, Belgium, European Standard EN 1993-1-2, 2005.
- [4] B. Wu and J. Z. Lu, "A numerical study of the behaviour of restrained RC beams at elevated temperatures," *Fire Safety Journal*, vol. 44, no. 4, pp. 522–531, May 2009, <https://doi.org/10.1016/j.firesaf.2008.10.006>.
- [5] J. M. Franssen and J. C. Dotreppe, "Fire Tests and Calculation Methods for Circular Concrete Columns," *Fire Technology*, vol. 39, no. 1, pp. 89–97, Jan. 2003, <https://doi.org/10.1023/A:1021783311892>.
- [6] S. Seręga, "Effect of transverse reinforcement spacing on fire resistance of high strength concrete columns," *Fire Safety Journal*, vol. 71, pp. 150–161, Jan. 2015, <https://doi.org/10.1016/j.firesaf.2014.11.017>.
- [7] C. Ibañez, M. L. Romero, and A. Hospitaler, "Effects of axial and rotational restraints on concrete-filled tubular columns under fire,"

- Journal of Constructional Steel Research*, vol. 125, pp. 114–127, Oct. 2016, <https://doi.org/10.1016/j.jcsr.2016.06.010>.
- [8] "Eurocode 4: Design of composite steel and concrete structures – Part 1-2: General rules - Structural fire design," European Committee for Standardization, Bruxelles, Belgium, European Standard EN 1994-1-2, 2005.
- [9] Y. Yao, H. Li, H. Guo, and K. Tan, "Fire resistance of eccentrically loaded slender concrete-filled steel tubular columns," *Thin-Walled Structures*, vol. 106, pp. 102–112, Sep. 2016, <https://doi.org/10.1016/j.tws.2016.04.025>.
- [10] Y. Yao, H. Li, and K. Tan, "Theoretical and numerical analysis to concrete filled double skin steel tubular columns under fire conditions," *Thin-Walled Structures*, vol. 98, pp. 547–557, Jan. 2016, <https://doi.org/10.1016/j.tws.2015.10.024>.
- [11] M. Yu, T. Wang, W. Huang, and J. Ye, "Fire resistance of concrete-filled steel tube columns with preload. Part II: Numerical and analytical investigation," *Composite Structures*, vol. 223, Sep. 2019, Art. no. 110995, <https://doi.org/10.1016/j.compstruct.2019.110995>.
- [12] G. M. Kamil, Q. Q. Liang, and M. N. S. Hadi, "Fire-Resistance of Eccentrically Loaded Rectangular Concrete-Filled Steel Tubular Slender Columns Incorporating Interaction of Local and Global Buckling," *International Journal of Structural Stability and Dynamics*, vol. 19, no. 08, Aug. 2019, Art. no. 1950085, <https://doi.org/10.1142/S0219455419500858>.
- [13] W.-J. Mao, W.-D. Wang, and W. Xian, "Numerical analysis on fire performance of steel-reinforced concrete-filled steel tubular columns with square cross-section," *Structures*, vol. 28, pp. 1–16, Dec. 2020, <https://doi.org/10.1016/j.istruc.2020.08.043>.
- [14] S. Li, J. Y. R. Liew, M.-X. Xiong, and B.-L. Lai, "Experimental investigation on fire resistance of high-strength concrete encased steel composite columns," *Fire Safety Journal*, vol. 121, May 2021, Art. no. 103273, <https://doi.org/10.1016/j.firesaf.2020.103273>.
- [15] A. N. Hassooni and S. R. A. Zaidee, "Behavior and Strength of Composite Columns under the Impact of Uniaxial Compression Loading," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8843–8849, Aug. 2022, <https://doi.org/10.48084/etasr.4753>.
- [16] M. Baghdadi, M. S. Dimia, and D. Baghdadi, "A Parametric Study of Fire-Damaged Reinforced Concrete Columns under Lateral Loads," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9113–9119, Oct. 2022, <https://doi.org/10.48084/etasr.5172>.
- [17] J. M. Franssen, "User's manual for SAFIR 2007a - A computer program for analysis of structures subjected to fire," University of Liege, Liege, Belgium, User Manual, Feb. 2008.

# On the Efficiency of the Advanced TWA Approach to the 60-GHz Microstrip Antenna Analysis for 5G Wireless Communication Systems

**Mohamed Ayari**

Faculty of Computing and Information Technology, Northern Border University, Saudi Arabia | Syscom Laboratory, National Engineering School of Tunis, University of Tunis El-Manar, Tunisia  
mohamed.ayari@nbu.edu.sa  
(corresponding author)

Received: 21 December 2022 | Revised: 5 January 2023 | Accepted: 7 January 2023

## ABSTRACT

The bandwidth demand in mobile wireless applications has grown at an astounding rate due to the fast evolution of technology, spurring the antenna design fields. The approaches associated with realizing antenna structure at mm-Wave frequency bands for future 5G cellular devices present advantages and disadvantages. This paper exhibits a fast and original approach based on the transverse wave formulation called-up Advanced Transverse Wave Approach (ATWA). A compact 5G patch antenna is designed, measured, simulated, and analyzed in the context of unlicensed mm-wave ISM-band applications. Compared to recently published works, the obtained result analysis proves the efficiency of the proposed method in terms of calculation accuracy, computational efficiency, and peak memory usage and the overall good performance of the proposed antenna.

*Keywords-5G antennas; advanced transverse wave approach; calculation accuracy, mm-wave analysis*

## I. INTRODUCTION

Nowadays, much research is directed at the fifth-generation mobile communications system (5G). Soon the traffic volume will be 1000 times over than the last decade with low cost and minimum power consumption, data rate communication will exceed 10Gbps making it 100 times faster than 4G, simultaneous connection of a large number of terminals will be available to support the future spread of M2M communications and internet of things (IoT), all thanks to the 5G technology. Undoubtedly, there are many challenges in this context, among them is to design, treat, and simulate mm-wave antennas used in 5G wireless in an effective way. Numerous recent scientific works have concentrated on the design and mm-wave investigation of 5G antennas with diverse size, from compact to large, either 2D or 3D [1-3]. Besides, 5G antenna shapes (i.e. circular, elliptical, rectangular, hexagonal, cross, etc.) [4-7] are considered as a significant parameter in the analysis process to achieve or not good quality in terms of performance and cost-effectiveness. In wireless communication systems, small patch antennas play an important role in the gain and bandwidth improvements due to features such as multiband properties, low profile topology and cost, good gain, and easy fabrication [8-10].

Several publications have been reported in this context. For instance, a compact size dual band patch antenna functioning at 10.15GHz and 28GHz for 5G mobile communications has been designed and studied in [11]. In [12], the authors presented a new capacitive coupled patch antenna array operating at the frequency band of 24-28GHz providing 360° coverage for 5G based smartphone services. In [13], a patch slotted antenna using FR4 substrate with the size of 18×18×1.6mm<sup>3</sup> providing at 10.15GHz a gain of 4.46dB and one of 6.6dB at 28GHz was reported in [14] using a small antenna with coplanar feeding for a Rogers RT5880 substrate and 5×5×0.254mm<sup>3</sup> size. Such antennas present low radiation efficiency, low gain, and some intricacy in the fabrication process. The mm-wave phased array 5G antenna which is running in the frequency range of 25 to 40GHz has been designed and analyzed in the context of Multiple-Input Multiple-Output (MIMO) applications.

The increased demand of frequency allocation coerced the Federal Communication Committee (FCC) to boost more than 18GHz of spectrum surrounding the mm-wave frequencies [15, 43] in order to enhance the development and scientific research in 5G technology. However, the 57-71GHz interval which is the unlicensed frequency band according to FCC covers an important set of applications, such as WiGig [16], a new technology derived from Wi-Fi but supporting faster wireless transfer speed than Wi-Fi. So, many recent publications have

been reported in the context of analysis and design of 5G antennas used for WiGig applications. In [17], a stacked circular patch antenna with monopole type pattern was developed for 60GHz WPAN application. A 60GHz microstrip patch antenna was investigated in [18] based on different dielectric materials to prove their impact on antenna operation. In [19], the authors proved the efficiency of a 60GHz virtual loop antenna built on low resistivity silicon substrate. These antennas have been fabricated and designed in different shapes, such as rectangular, circular, helical [20], double F-slot [21], double U-slot [22], T-slot [23], E- and H-slots [24]. Moreover, the feasibility of 5G connectivity beyond 70GHz was demonstrated using an experimental prototype in [25]. The investigation and analysis of mm-wave 5G antenna based on numerical methods is an unprecedented and great challenge in the full-wave electromagnetic (EM) simulation sphere.

Numerical EM methods belong to the core of the mm-wave analysis of 5G planar structures such as the Method of Moments (MOM) [26], or its alternative Fast Multipole Method (FMM) [27], and the full spectrum method named Partial Element Equivalent Circuit (PEEC)[28] for solving integral equations of EM waves. In the same vein, Finite Element Method (FEM)[29], Finite Difference Time Domain (FDTD) method [30], and Finite Integration Technique (FIT) [31] belong to differential equation solvers. These numerical methods necessitate a significant amount in memory and processor power as well as high accuracy in features and design for full-wave and mm-wave analysis of planar structures employed in wireless systems [32].

In the area of wireless applications, the Advanced Transverse Wave Approach (ATWA) developed by our research team has significant benefits over previous numerical EM approaches in terms of speed, compactness, and memory consumption. To this purpose, we adapt and refine our ATWA approach to work with mm-wave 5G antenna simulations, and we tested its efficiency and stability in this setting. In line with the above mentioned works, an appropriate 5G antenna was considered and designed to validate our approach in the context of mm-wave 5G applications and offer the possibility to investigate smart and powerful antennas for the next generation of innovative wireless technology. In the following, we present and develop our numerical EM method –advanced transverse wave approach– in the context of mm-wave applications.

## II. ATWA: THEORETICAL BACKGROUND

For most practical interest problems pertaining to EM radiation and scattering, an analytic solution of integral equations formulating the problem cannot be found. Therefore, researchers tend to utilize computational techniques to obtain a solution. We briefly present in this section the mathematical background and the theoretical foundation of the numerical EM method ATWA as well as its main characteristics and strong points.

The two-dimensional transverse wave approach – a fast numerical method based on an iterative process– has several potentialities that set it apart from other computational EM approaches. There are no matrix inversion calculations in the TWA process, no constraints are required on the component

forms, no numerical instabilities frequently arising from large matrices can be found, and the convergence is guaranteed independently of the interfaces of studied planar structures. In the following, we present a small theoretical background of our EM-approach 2D-ATWA for mmWave analysis of planar structures.

The general integral relation linking the electric field to current density and considering the boundary and edge conditions on sub-domains characterizing the air-dielectric interface  $\Omega$  can be written as follows:

$$E(u, v) = \int_{\Omega} G(u, v, u', v') J(u', v') du' dv' \quad (1)$$

where  $G(u, v, u', v')$  denotes the dyadic Green's function [33].

Referring to the wave concept already investigated and developed in [36], the general equation linking the transverse electric field ET to the transverse magnetic field HT can be expressed as:

$$W_{i,\tau} = (4Z_{0i})^{-\frac{1}{2}} (E_{T,i} + (-1)^{\tau} Z_{0i} J_{T,i}) \quad (2)$$

where  $\tau$  is a Boolean parameter referring to the wave nature (i.e.  $\tau=0$  for incident wave ( $W=A$ ) and  $\tau=1$  for reflected wave ( $W=B$ )), the superscript  $T$  refers to the tangential components, and  $Z_{0i}$  stands for the wave impedance of the homogeneous isotropic region  $i \in \{1,2\}$  given by:

$$Z_{0i} = \eta_0 \times (\mu_{r_i} / \epsilon_{r_i})^{1/2} \quad (3)$$

where  $\eta_0$  is the intrinsic impedance of free space defined as:

$$\eta_0 = 120\pi \text{ or } 377\Omega \quad (4)$$

where  $\epsilon_{r_i}, \mu_{r_i}$  denote the relative permittivity (capacitance) and relative permeability (inductivity) of the medium  $i$  and the vector  $J_T$  is defined from the tangential magnetic field  $H_T$  phased by  $\pi/2$  from  $v$  which denotes the outgoing normal vector oriented towards region  $i$ . It can be written as:

$$J_T = H_T \times v \quad (5)$$

Overall, both incident ( $A$ ) and reflected ( $B$ ) waves can be associated through the following relation:

$$\begin{cases} A = \hat{F} B & \text{In modal domain} \\ B = \hat{S} A + B^{(0)} & \text{In spatial domain} \end{cases} \quad (6)$$

where  $B$  represents the general excitation wave on the source.

Moreover, shifting from spatial to modal spaces is guaranteed by the transitional spectral space. In fact, at the discontinuity interface, the two-dimensional Fast Fourier Transform (2D-FFT) or the two-dimensional Non-Uniform Fast Fourier Transform (2D-NUFFT) [37-38] for uniform and non-uniform distribution, respectively, are intercepted in the transition between the spatial and spectral domains in order to accelerate the iterative process. The change from spectral to modal spaces can be ensured by the rotation of angle  $\theta$ , given by:

$$\theta = \arctan(K_{mn}^x / K_{mn}^y) \quad (7)$$

Undoubtedly, the passing from modal to spatial domains is achieved by the inverse 2D-FFT or 2D-NUFFT as a first step and the rotation of angle  $-\theta$  as the next step. The round trip from these spaces is repeated until the system's convergence. The 2D-TWA can be accelerated by applying the anisotropic mesh technique [39] which calls the 2D-NUFFT to manage precisely the non-uniform discretization. Memory and computational complexity for both the Forward and Backward of the ATWA process is guaranteed at  $O(N_T \log N_T)$  time, where  $N_T$  represents the meshing density applied to the structure to be analyzed, whereas the overall computational complexity of iterative solvers is  $O(N^2)$ , the same with MOM- and FEM-based solvers [40]. Table I recapitulates the iterative process of the ATWA approach, highlighting at each step the computational effort and the total time efficiency.

TABLE I. ITERATIVE PROCESS AND COMPUTATIONAL EFFORT OF ATWA

Step	Description	Complexity
Init	Computation of spectral-modal transition, reflection, and diffraction coefficients. Input complex vectors of bilateral excitation source (incident and reflected waves at $t=0$ ) $B_0 \neq 0$ and $A_0 = 0$ .	$O(N_T)$
1	Application of 2D-FFT/2D-NUFFT at iteration $k$ .	$O(N_T \log N_T)$
2	Application of modal transform at iteration $k$ .	$O(N_T)$
3	Application of reflection operator at iteration $k+1$ .	$O(N_T)$
4	Application of the inverse modal transform at iteration $k+1$ .	$O(N_T)$
5	Application of the inverse of 2D-FFT/2D-NUFFT at iteration $k+1$ .	$O(N_T \log N_T)$
6	Application of diffraction operator at iteration $k+1$ .	$O(N_T)$
7	Compute current density, electric field, and input impedance $Z_{in}$ at iteration $k+1$ .	$O(N_T)$
8	Test the convergence of $Z_{in}$ at iteration $k+1$ . If $Z_{in}$ converges, then keep all EM quantities Else goto Step 1.	$O(N_{Sce})$ ( $N_{Sce} \ll N_T$ : total number of pixels on the sub-domain source)

### III. ANTENNA DESIGN FOR 5G

The new 5G demand can open prospects for miscellaneous new applications and the design process can be evolved taking into consideration these requirements. Against this background, a 5G small patch antenna fed by a microstrip line has been designed in the context of wireless applications (Figure 1). Such an antenna can enhance the radiation directivity and be used to fix and relay signals around obstacles in the context of ISM-band applications. The proposed planar structure is composed by an  $l_p \times w_p$  radiating patch printed upon an  $l_s \times p_s \times h_s$  substrate with Rogers 5880. RT/duroid 5880 laminate is considered as a judicious choice and a perfect substrate, since it is isotropic and characterized by the lowest electrical loss for reinforced polytetrafluoroethylene (PTFE) [34], uniform electrical properties over frequency, low moisture absorption, and excellent chemical resistance. This substrate is more convenient to mm-wave applications, military radar systems, commercial airline broadband, and missile guidance systems.

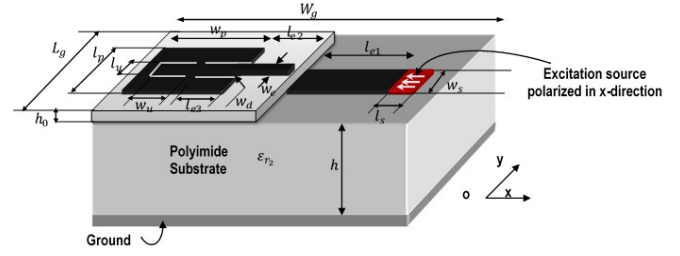


Fig. 1. The proposed 5G small patch antenna fed by microstrip line.

TABLE II. GEOMETRIC SIMULATION PARAMETERS OF THE PROPOSED 5G-ANTENNA

Meshing resolution	512x512
Box dimensions (mm)	$L_g = 8.192$
	$W_g = 8.192$
Patch	Length $l_p = 2.912 = 182$ pixels
	Width $w_p = 3.2 = 200$ pixels
Microstrip line (mm)	$l_{e1} = 2.56 = 160$ pixels
	$l_{e2} = 0.96 = 60$ pixels
	$l_{e3} = 1.12 = 70$ pixels
	$w_e = 0.736 = 46$ pixels
U-slot dimensions (mm)	$l_u = 1.12 = 70$ pixels
Width $w_d$ (mm)	$w_u = 1.024 = 64$ pixels
Source dimensions (mm)	$w_d = 0.192 = 12$ pixels
	$l_s = 0.736 = 46$ pixels
Form factor of source	$w_s = 1.792 = 112$ pixels
	$l_s/w_s = 0.41$
Substrate	Thickness $\delta = 130\mu\text{m}$
	Type: Rogers RT/duroid
	Height $h = 0.5\text{mm}$

In order to enhance the capacitance inhibiting the probe inductance which causes an expansion of the impedance bandwidth, the U-slot has been loaded on the patch. The U-slot technique has been applied in many recent 5G structures [15, 35] showing adeptness in wideband notion. The design parameters are detailed in Table II.

### IV. SIMULATION RESULTS AND DISCUSSION

This section provides a concise and precise description of the experimental results, their interpretation, and the conclusions that can be drawn.

In order to validate our numerical method in the EM simulation space of 5G antennas, we investigate the proposed planar structure shown in Figure 1 by analyzing the S11 parameter and comparing it with the results obtained in the literature. The simulation investigation was carried out by our EM simulator based on the ATWA method and developed in C++ environment. The modeling simulation constraints are given in Table III.

The advanced TWA approach gives certain significant potentialities to the iterative scheme. It detaches the continuities relations defining the boundary conditions from the integral one, taking into consideration the half-spaces describing the different metallic interfaces identifying the proposed structure. One of these potentialities are the above mentioned Heaviside functions associated to the 5G-antenna (i.e. sub-domains: metallic  $H_{\Omega_{SUB}} = Me$ , dielectric  $H_{\Omega_{SUB}} =$

$Di$ , source  $H_{\Omega_{SUB}} = Sce$ , and surface impedance  $H_{\Omega_{SUB}} = I_{surf}$ ). The reflection coefficient or insertion parameter S11 is considered as the standard parameter for antennas. It quantifies the power reflected from the antenna and identifies the different resonance frequencies. The simulation result as depicted in Figure 2 shows the evolution of the S11 of our proposed antenna.

TABLE III. MODELING SIMULATION PARAMETERS OF THE PROPOSED 5G-ANTENNA

Modeling parameters	Description
Nature of box	Periodic walls
Type of polarization	Bilateral in x-direction
Number of iterations	$N_{iter} = 500$
Value of surface impedance	$Z_s = 0$ (surface impedance sub-domain is equivalent to metallic sub-domain)
Permittivity of regions	$\epsilon_{r1} = 1$ $\epsilon_{r2} = 2.2$
Waveband	$F_{min} = 50\text{GHz}$ $F_{max} = 70\text{GHz}$ $Step\_F_{rq} = 0.5\text{GHz}$
Height $h_0$	$h_0 = 0$ (monolayer)

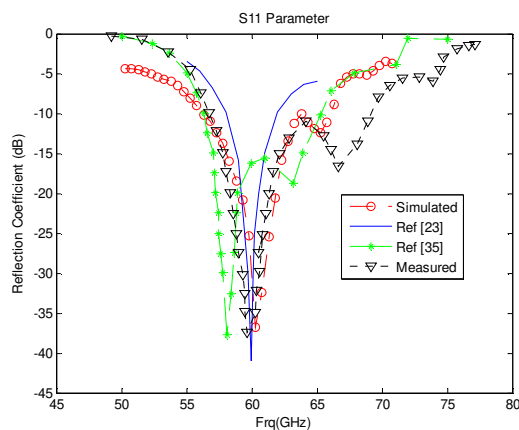


Fig. 2. Evolution of S11 coefficient.

Two bands have been detected at [57.62, 63.32] GHz and [64.29, 66.05] GHz, where the resonance frequencies are respectively depicted at 60.47GHz and 65.17GHz. The return losses are correspondingly -37.83dB and -12.44dB. The Fractional Bandwidths (FBWs) [36] are consequently 9.43% for the first and 2.7% for the second band of the proposed 5G-antenna. The Voltage Standing Wave Ratio (VSWR) is deduced at 1.0260dB for the 60GHz band which is considered as a good impedance for 5G wireless systems. The presence of double resonances is due to the U-slot loaded on the patch antenna that it is warranted by simulation result based on 2D-ATWA. Table III illustrates the comparison between the proposed 5G antenna and other known antennas.

It is noticed that the three selected reference 60GHz antennas are simulated using the Finite-Element Method (FEM)-based Frequency Domain Solver (FDS) of CST Microwave Technology [41] and the FEM based high-frequency structure simulator (HFSS) of ANSYS [24, 42].

Besides, different simulation results of the proposed antenna based on our ATWA method are successfully presented such as the evolution of the impedance as a function of frequency (Figure 3), making evident the stability of the system around the resonance frequency, the 3D-radiation pattern (Figure 4), the azimuth pattern (Figure 5), and the elevation pattern (Figure 6) at 60GHz proving the high and accurate radiation directivity and gain of the proposed 5G antenna.

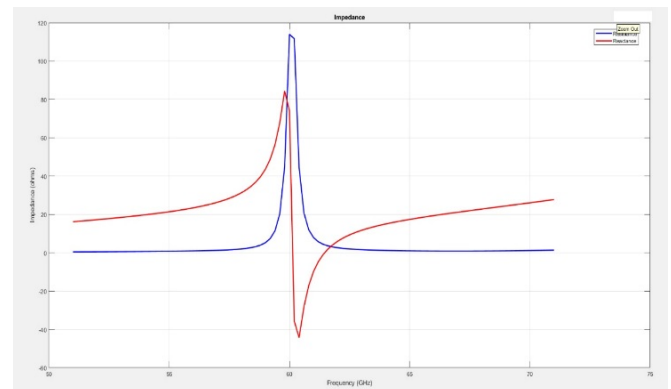


Fig. 3. Evolution of the impedance as a function of frequency.

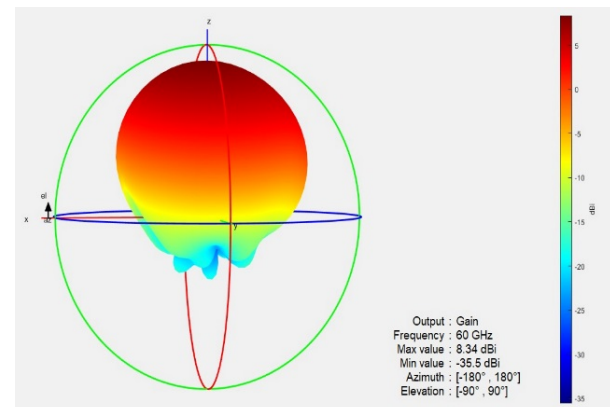


Fig. 4. Radiation pattern at 60GHz.

Overall, the obtained simulation results are in good agreement with the results of [24, 41, 42], with some preferences of our system in terms of size, multi-band notion, and impedance bandwidth. The computational statistics in terms of memory and CPU time for these simulations, using an Intel® Core™ i7-10875H Processor with 8 cores based on the Comet Lake architecture, are given in Table IV. The peak memory usage and total calculation time are significantly reduced in comparison with the other FEM-based solvers with the same computing resources. This proves the feasibility, plausibility, and EM-validity of our approach in 5G antenna simulations along with its accuracy and numerical stability in the compilation process.

In addition, considering the above mentioned features as the efficiency of the AMT technique presented by our approach (ATWA), a good workspace can be opened to investigate miniaturized and sensible structures used for new 5G wireless technologies.



TABLE IV. COMPARISON BETWEEN THE PROPOSED 5G-ANTENNA AND REFERENCE ANTENNAS

Parameter	Antenna			
	Proposed	[24]	[41]	[42]
Substrate	RT/duroid 5880	RT/duroid 5880	RT/duroid 5880	RT/duroid 5880
Shape-slot	U-slot	E and H slots	U-slot	U-slot
Size (mm <sup>2</sup> )	8.192×8.192	8×8	7.5×9.5	7.5×9.5
Simulator	Own simulator	HFSS	FDS	HFSS
Numerical EM method	ATWA	FEM	FEM	FEM
Return loss (dB)	-37.83	-40.99	-37.5	-29.36
Single/multiband	Double band	Single band	Double band	Single band
Impedance bandwidth (dB)	$S_{11} \leq -10$	$S_{11} \leq -10$	$S_{11} \leq -15$	$S_{11} \leq -15$
VSWR at 60GHz band (dB)	1.0260	1.0186	1.0270	1.038
Peak memory usage/GB	6.96	9.53	11.41	10.31

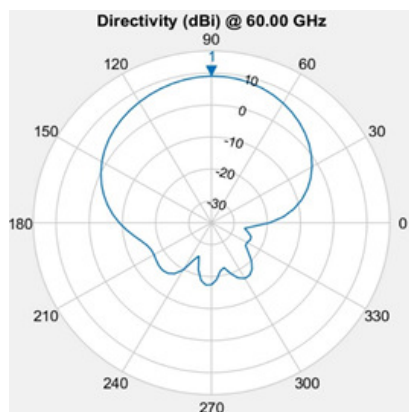


Fig. 5. Azimuth pattern at 60GHz.

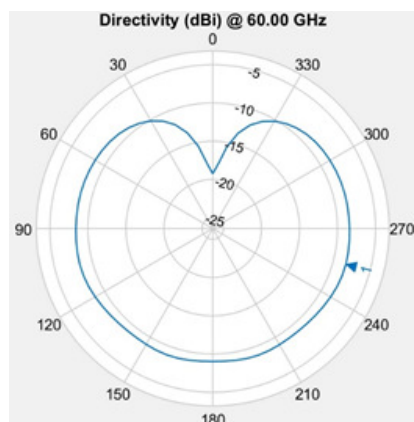


Fig. 6. Elevation pattern at 60GHz.

## V. CONCLUSIONS

In this paper, a brief theoretical background and the mathematical foundations of our two-dimensional Advanced Transverse Wave Approach (2D-ATWA) has been presented and developed for mm-wave applications. A 5G small patch antenna with U-slot fed by microstrip line was presented and effectively designed taking into account the presence of the U-shaped slot etched in the radiating patch and the isotropic substrate RT/duroid 5880. The obtained simulations were validated, evaluated, and compared to various 60GHz reference antennas already simulated and investigated using the Finite Element Method-based Frequency Domain Solver and ANSYS HFSS. The results prove the EM-stability and efficiency of our

approach in the context of mm-wave applications. Many antenna parameters such as shape, size, embedded shape-slots and slits, substrate material, dimensions, etc. have great impact on the operation of an antenna, such as increasing/decreasing bandwidth and gain or generating single or multiband frequencies. The marriage of numerical EM approaches with the game theory which is considered as a new trend in electromagnetics research can provide efficient solutions to stabilize and select good parameters for rigorous EM-investigation, mainly in the context of mm-wave applications.

## REFERENCES

- [1] W. A. Awan, S. I. Naqvi, A. H. Naqvi, S. M. Abbas, A. Zaidi, and N. Hussain, "Design and Characterization of Wideband Printed Antenna Based on DGS for 28 GHz 5G Applications," *Journal of Electromagnetic Engineering and Science*, vol. 21, no. 3, pp. 177–183, Jul. 2021, <https://doi.org/10.26866/jees.2021.3.r.24>.
- [2] D. T. T. My, H. N. B. Phuong, T. T. Huong, and B. T. M. Tu, "Design of a Four-Element Array Antenna for 5G Cellular Wireless Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6259–6263, Oct. 2020, <https://doi.org/10.48084/etasr.3771>.
- [3] H. Alsaif, "Compact Hexagonal Monopole Antenna for Lower 5G Bands," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4200–4202, Jun. 2019, <https://doi.org/10.48084/etasr.2714>.
- [4] A. Jamil, K. Rafique, M. D. Idrees, A. S. Rajput, A. Abdullah, and A. S. Saand, "Design and Geometric Transformations of Koch Curve Monopole Antennas," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8452–8457, Apr. 2022, <https://doi.org/10.48084/etasr.4767>.
- [5] G.-H. Jeon, P. A. Dzagbletey, and J.-Y. Chung, "A Cross-Joint Vivaldi Antenna Pair for Dual-Pol and Broadband Testing Capabilities," *Journal of Electromagnetic Engineering and Science*, vol. 21, no. 3, pp. 201–209, Jul. 2021, <https://doi.org/10.26866/jees.2021.3.r.27>.
- [6] D. V. Nandalal, A. Pavithra, S. Pavithra, and M. Kalaiselvi, "Performance Measure of Ultra wide Band Antenna for Hexagonal and Rectangular Shape for Wearable Application," *Asian Journal of Applied Science and Technology*, vol. 1, no. 3, pp. 80–84, 2017.
- [7] M. J. Hakeem and M. M. Nahas, "Improving the Performance of a Microstrip Antenna by Adding a Slot into Different Patch Designs," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7469–7476, Aug. 2021, <https://doi.org/10.48084/etasr.4280>.
- [8] M. Nahas, "A Super High Gain L-Slotted Microstrip Patch Antenna For 5G Mobile Systems Operating at 26 and 28 GHz," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8053–8057, Feb. 2022, <https://doi.org/10.48084/etasr.4657>.
- [9] S. Nelaturi and N. V. S. N. Sarma, "Compact Wideband Microstrip Patch Antenna based on High Impedance Surface," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3149–3152, Aug. 2018, <https://doi.org/10.48084/etasr.1971>.
- [10] D. Imran *et al.*, "Millimeter wave microstrip patch antenna for 5G mobile communication," in *International Conference on Engineering*

- and Emerging Technologies, Lahore, Pakistan, Feb. 2018, pp. 1–6, <https://doi.org/10.1109/ICEET1.2018.8338623>.
- [11] Y. Jandi, F. Gharnati, and A. Oulad Said, "Design of a compact dual bands patch antenna for 5G applications," in *International Conference on Wireless Technologies, Embedded and Intelligent Systems*, Fez, Morocco, Apr. 2017, pp. 1–4, <https://doi.org/10.1109/WITS.2017.7934628>.
  - [12] M. Stanley, Y. Huang, H. Wang, H. Zhou, A. Alieldin, and S. Joseph, "A Capacitive Coupled Patch Antenna Array With High Gain and Wide Coverage for 5G Smartphone Applications," *IEEE Access*, vol. 6, pp. 41942–41954, 2018, <https://doi.org/10.1109/ACCESS.2018.2860795>.
  - [13] N. Kalwit, P. Pawar, and P. Moghe, "Microstrip Patch Antenna for Future 5G Applications," in *Mobile Radio Communications and 5G Networks*, N. Marriwala, C. C. Tripathi, D. Kumar, and S. Jain, Eds. Singapore, Asia: Springer, 2021, pp. 469–474.
  - [14] O. M. Haraz, M. M. M. Ali, S. Alshebeili, and A.-R. Sebak, "Design of a 28/38 GHz dual-band printed slot antenna for the future 5G mobile communication Networks," in *International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, Vancouver, BC, Canada, Jul. 2015, pp. 1532–1533, <https://doi.org/10.1109/APS.2015.7305155>.
  - [15] A. Morgado, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, "A survey of 5G technologies: regulatory, standardization and industrial perspectives," *Digital Communications and Networks*, vol. 4, no. 2, pp. 87–97, Apr. 2018, <https://doi.org/10.1016/j.dcan.2017.09.010>.
  - [16] M. S. Ibrahim, "Low-Cost, Circularly Polarized, and Wideband U-Slot Microstrip Patch Antenna with Parasitic Elements for WiGig and WPAN Applications," in *13th European Conference on Antennas and Propagation*, Krakow, Poland, Mar. 2019, pp. 1–4.
  - [17] W. Byun, B.-S. Kim, K.-S. Kim, M.-S. Kang, and M.-S. Song, "Stacked circular patch antenna with monopole type pattern for 60GHz WPAN application," in *IEEE Antennas and Propagation Society International Symposium*, San Diego, CA, USA, Jul. 2008, pp. 1–4, <https://doi.org/10.1109/APS.2008.4620014>.
  - [18] M. Kamran Saleem, M. Xie, M. A. S. Alkanhal, and M. Saadi, "Effect of dielectric materials on integrated lens antenna for millimeter wave applications," *Microwave and Optical Technology Letters*, vol. 61, no. 4, pp. 1079–1083, 2019, <https://doi.org/10.1002/mop.31676>.
  - [19] D. Bereka, A. Gopinath, and B. Sainati, "Design of a 60GHz high efficiency virtual loop antenna on low resistivity silicon substrate," in *IEEE International Symposium on Antennas and Propagation*, Spokane, WA, USA, Jul. 2011, pp. 891–894, <https://doi.org/10.1109/APS.2011.5996419>.
  - [20] C. Liu, Y.-X. Guo, X. Bao, and S.-Q. Xiao, "60-GHz LTCC Integrated Circularly Polarized Helical Antenna Array," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 3, pp. 1329–1335, Mar. 2012, <https://doi.org/10.1109/TAP.2011.2180351>.
  - [21] F. Ahmad and B. Tlili, "Design and analysis of millimeter wave double F slot patch antenna for future 5G wireless communication," in *International Conference on Electrical and Computing Technologies and Applications*, Ras Al Khaimah, United Arab Emirates, Nov. 2017, pp. 1–4, <https://doi.org/10.1109/ICECTA.2017.8252049>.
  - [22] T. H. Jang, H. Y. Kim, D. M. Kang, S. H. Kim, and C. S. Park, "60 GHz Low-Profile, Wideband Dual-Polarized U-Slot Coupled Patch Antenna With High Isolation," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 7, pp. 4453–4462, Jul. 2019, <https://doi.org/10.1109/TAP.2019.2911623>.
  - [23] R. K. Goyal and K. K. Sharma, "T-slotted microstrip patch antenna for 5G Wi-Fi network," in *International Conference on Advances in Computing, Communications and Informatics*, Jaipur, India, Sep. 2016, pp. 2684–2687, <https://doi.org/10.1109/ICACCI.2016.7732465>.
  - [24] J. Saini and S. K. Agarwal, "Design a single band microstrip patch antenna at 60 GHz millimeter wave for 5G application," in *International Conference on Computer, Communications and Electronics (Comptelix)*, Jaipur, India, Jul. 2017, pp. 227–230, <https://doi.org/10.1109/COMPTELIX.2017.8003969>.
  - [25] G. Hattab, E. Visotsky, M. Cudak, and A. Ghosh, "Toward the Coexistence of 5G MmWave Networks with Incumbent Systems beyond 70 GHz," *IEEE Wireless Communications*, vol. 25, no. 4, pp. 18–24, Aug. 2018, <https://doi.org/10.1109/MWC.2018.1700436>.
  - [26] M. Ayari, Y. E. Touati, and S. Altowaijri, "Method of Moments versus Advanced Transverse Wave Approach for EM Validation of Complex Microwave and RF Applications," *Journal of Electromagnetic Engineering and Science*, vol. 20, no. 1, pp. 31–38, Jan. 2020, <https://doi.org/10.26866/jees.2020.20.1.31>.
  - [27] B. Wu and X.-Q. Sheng, "Fast Characteristic Mode Analysis for Half-space Platform Antennas with Multilevel Fast Multipole Algorithm," in *International Applied Computational Electromagnetics Society Symposium - China*, Beijing, China, Jul. 2018, pp. 1–2, <https://doi.org/10.23919/ACCESS.2018.8669181>.
  - [28] C.-C. Chou and T.-L. Wu, "Direct Simulation of the Full-Wave Partial Element Equivalent Circuit Using Standard SPICE [Application Notes]," *IEEE Microwave Magazine*, vol. 20, no. 6, pp. 22–34, Jun. 2019, <https://doi.org/10.1109/MMM.2019.2904376>.
  - [29] H. Santos, P. Pinho, and H. Salgado, "Patch Antenna-in-Package for 5G Communications with Dual Polarization and High Isolation," *Electronics*, vol. 9, no. 8, Aug. 2020, Art. no. 1223, <https://doi.org/10.3390/electronics9081223>.
  - [30] K. R. Mahmoud and A. M. Montaser, "Design of Compact mm-wave Tunable Filtenna Using Capacitor Loaded Trapezoid Slots in Ground Plane for 5G Router Applications," *IEEE Access*, vol. 8, pp. 27715–27723, 2020, <https://doi.org/10.1109/ACCESS.2020.291606>.
  - [31] A. Rahimian, S. F. Jilani, Q. H. Abbasi, A. Alomainy, and Y. Alfidhl, "A Millimetre-Wave Two-Dimensional 64-Element Array for Large-Scale 5G Antenna Subsystems," in *13th European Conference on Antennas and Propagation*, Krakow, Poland, Mar. 2019, pp. 1–2.
  - [32] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Kürner, "The Design and Applications of High-Performance Ray-Tracing Simulation Platform for 5G and Beyond Wireless Communications: A Tutorial," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 10–27, 2019, <https://doi.org/10.1109/COMST.2018.2865724>.
  - [33] A. Eroglu and J. K. Lee, "Dyadic Green's functions for an electrically gyrotropic medium," *Progress in Electromagnetics Research*, vol. 58, pp. 223–241, 2006, <https://doi.org/10.2528/PIER05070203>.
  - [34] T. Itoh, "Spectral Domain Imittance Approach for Dispersion Characteristics of Generalized Printed Transmission Lines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 28, no. 7, pp. 733–736, Apr. 1980, <https://doi.org/10.1109/TMTT.1980.1130158>.
  - [35] M. Titaouine, A. G. Neto, H. Baudrand, and F. Djahli, "WCIP method applied to active frequency selective surfaces," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 6, no. 1, pp. 1–16, Aug. 2007.
  - [36] H. Baudrand, M. K. Azizi, and M. Titaouine, "General Principles of the Wave Concept Iterative Process," in *The Wave Concept in Electromagnetism and Circuits: Theory and Applications*, New York, NY, USA: Wiley, 2016, pp. 1–42.
  - [37] M. Ayari, T. Aguil, H. Temimi, and H. Baudrand, "More efficiency of Transverse Wave Approach (TWA) by applying Anisotropic Mesh Technique (AMT) for full-wave analysis of microwave planar structures," *Progress In Electromagnetics Research B*, vol. 14, pp. 383–405, Jan. 2009.
  - [38] M. Ayari, T. Aguil, and H. Baudrand, "New Version of Twa Using Two-Dimensional Non-Uniform Fast Fourier Mode Transform (2D-Nuffmt) for Full-Wave Investigation of Microwave Integrated Circuits," *Progress In Electromagnetics Research B*, vol. 15, pp. 375–400, 2009, <https://doi.org/10.2528/PIERB09052301>.
  - [39] M. Ayari, T. Aguil, H. Temimi, and H. Baudrand, "On the Use of Anisotropic Mesh Technique (AMT) in Transverse Wave Approach (TWA) for the EM Analysis of Microwave Circuits in Wireless Applications," in *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing*, Brisbane, QLD, Australia, Jul. 2009, pp. 246–251, <https://doi.org/10.1109/UIC-ATC.2009.8>.
  - [40] S. Zuo, D. Garcia Donoro, Y. Zhang, Y. Bai, and X. Zhao, "Simulation of Challenging Electromagnetic Problems Using a Massively Parallel Finite Element Method Solver," *IEEE Access*, vol. 7, pp. 20346–20362, 2019, <https://doi.org/10.1109/ACCESS.2019.2896615>.

- [41] H. Liu, Y. He, and H. Wong, "Printed U-slot patch antenna for 60GHz applications," in *International Workshop on Electromagnetics, Applications and Student Innovation Competition*, Kowloon, Hong Kong, Aug. 2013, pp. 153–155, <https://doi.org/10.1109/iWEM.2013.6888796>.
- [42] G. Kumar and C. Kumar, "Design and analysis of u-slot microstrip patch antenna for mobile communication at 60 ghz," *Acta Technica Corviniensis - Bulletin of Engineering*, vol. 13, no. 1, pp. 35–38, 2020.
- [43] S. Ahmad *et al.*, "Design of a Tri-Band Wearable Antenna for Millimeter-Wave 5G Applications," *Sensors*, vol. 22, no. 20, Jan. 2022, Art. no. 8012, <https://doi.org/10.3390/s22208012>.

#### AUTHOR PROFILE

**Mohamed Ayari** –received the Dipl.-Ing., M.S., and Ph.D. degrees in Telecommunications in 2003, 2004, and 2009, respectively, from the National Engineering School of Tunis (ENIT), Tunisia, in collaboration with the National Polytechnic Institute of Toulouse, France, and Virginia Tech, USA. He has been teaching at several universities since 2003. He has been a permanent research member at the 6'COM laboratory at ENIT from 2003. In 2005, he joined RCEM-Inc. in Toulouse, France. Since 2010, he has been a tenure-track Associate Professor at the National Engineering School of Carthage (ENICAR), Carthage University, Tunisia. In 2015, he joined the IT Department of the Faculty of Computing and Information Technology, Northern Border University (NBU), Kingdom of Saudi Arabia, as an Assistant Professor. His current research interests are electromagnetic (EM) fields, numerical EM methods, and computer-aided design of microwave circuits and antennas. His research interests also include information security, image processing and wireless applications.

# Processor-in-the-Loop Validation of an Observer Current-based Dead-Beat Control for a Single-Phase UPS Inverter

**Mohammed Saoudi**

Laboratory of Electrical Engineering (LGE), University of M'sila, Algeria  
mohammed.saoudi@univ-msila.dz  
(corresponding author)

**Hani Benguesmia**

Laboratory of Electrical Engineering (LGE), University of M'sila, Algeria | Department of Electrical Engineering, Faculty of Technology, University of M'sila, Algeria  
hani.benguesmia@univ-msila.dz

**Aissa Chouder**

Laboratory of Electrical Engineering (LGE), University of M'sila, Algeria | Department of Electrical Engineering, Faculty of Technology, University of M'sila, Algeria  
aissa.chouder@univ-msila.dz

Received: 25 November 2022 | Revised: 29 December 2022 | Accepted: 3 January 2023

## ABSTRACT

This paper presents a dead-beat control algorithm for Uninterruptible Power Supply (UPS) applications of single-phase inverters. The proposed control method requires the measurement of capacitor current and output voltage in order to keep the output voltage sinusoidal ensuring high dynamic performance even under load changes. The dead-beat controller optimizes the behavior of the system by eliminating the error between the output and the reference voltage without increasing the number of current sensors, which are costly, and eliminates load voltage distortions and restores the system state in the event of external shutdown-loop road interference. In this paper, we propose a capacitor current estimation based on the Luenberger observer. Processor-In-the-Loop (PIL) is a test method that allows us to create and evaluate controllers by running built-in C code on the DSP scheduled for the controller during simulated PSIM power phase control. It can be seen that the simulation results match the PIL test results, which proves the validity of the proposed controller.

**Keywords-**dead-beat control; single-phase inverter; PWM; Luenberger observer; PIL; DSP; CCS

## I. INTRODUCTION

Using the output LC filter to control the inverters is important in applications such as distributed generation, renewable energy-based island applications, and Uninterruptible Power Supplies (UPSs) [1-3], where high quality voltage is required. The main control objective of UPS is to adjust the output voltage in the presence of parameter uncertainty and disturbance, and its characteristics depend on the quality of the output voltage. Disturbances from various loads, uncertainty due to inductive loads, non-linear loads, and model mismatches can lead to poor output voltage tracking and high Total Harmonic Distortion (THD). The effects of noise and uncertainty must be minimized to achieve excellent voltage performance. Any UPS system has two operating modes: backup mode and bypass mode. Ideally, a UPS should be able

to deliver a regulated sinusoidal output voltage with low THD during the two modes and includes a dead-beat controller [4-10]. In [6, 11], the dead-beat controller is used with the disturbance observer for proper response and robust controlled performance. In [12-14], the PWM technique using a modified dead-beat controller is explained clearly, but the PWM pattern in the presence of load uncertainty is not detailed. In [15], the dead-beat control for power converters using fractional-order time delay compensation is defined as an effective and powerful delay compensation approach that has been applied in various control systems. In [16-21], single-phase inverter's deadbeat-based Proportional-Integral (PI) controller was proposed and comparisons with conventional control techniques were made. However, the implementation of the deadbeat controller is very complicated, and due to the

difficulty of its design, it is not suited for low cost implementations. However, the technique uses relatively higher energy losses.

The main purpose of control is to obtain a nearly sinusoidal output voltage across the output capacitor of the LC filter. Dead-beat control is the most attractive control method in discrete time systems because it can reduce the error between the reference and measured values as long as the number of samples goes to infinity and the error goes to zero, giving the inverter a fast dynamic response. The inverter's output voltage is achieved with very low THD. Dead-beat control is sensitive to variations of the filter parameters, even when the load fluctuates. The proposed dead-beat control system can work stably when the load variation is within the allowable range. In order to decrease the number of sensors, an observer to estimate the capacitor current is proposed. However, conventional dead-beat controllers suffer from two delays, the first is inherent to the dead-beat control algorithm and causes a steady-state error [12], and the second arises when the PSIM simulates the algorithm after CCS builds it and installs it on a DSP-based platform. It has an impact on the system's stability and causes ripples and phase shifts in the output current. The dead-beat response has some advantages:

- Zero steady-state error.
- Shortest rise time.
- Minimum settling period.
- An over/undershoot of less than 2%.
- Very high control signal output [16].

## II. SYSTEM DESCRIPTION

The proposed single-phase inverter is shown in Figure 1 and consists of an H-bridge inverter connected to the load through LC filtering. Insulated Gate Bipolar Transistors (IGBTs) are chosen as the switching elements to perform energy transfer and power conversion. LC filters are used to filter higher harmonics before the load. The capacitor voltage and current are generated by the Luenberger observer. This system performs better with imbalanced loads and is more adaptable when it comes to working conditions.

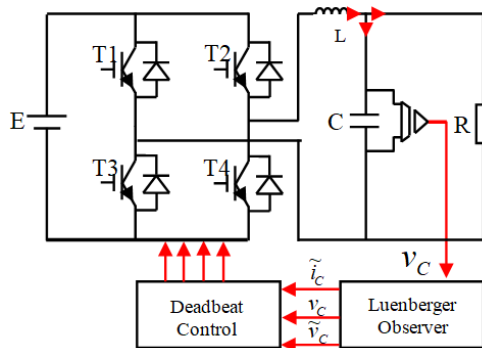


Fig. 1. Basic scheme of the proposed system.

## III. THE CONSIDERED SYSTEM

The single-phase inverter consists of two arms in a half-bridge. If the power switches S1 and S'1 are supposed to be perfect (and are negligible), the system (source-inverter-filter-charge) is then represented by the second-order linear model of Figure 2.

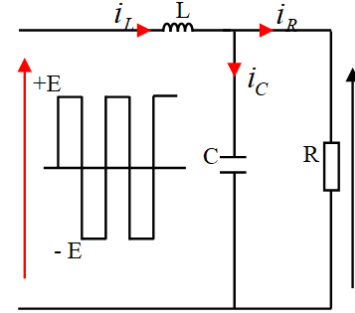


Fig. 2. Basic scheme of the proposed system.

This system's representation in state space is given by:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (1)$$

where  $x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$ .  $x(t)$  is the state vector, such that:

$$\begin{cases} x_1(t) = v_c(t) \\ x_2(t) = \dot{v}_c(t) \end{cases} \quad (2)$$

According to the second-order linear model of the half-bridge inverter, applying the laws of Kirchhoff and calculating the matrices A and B, we get:

$$v_{int}(t) = v_L(t) + v_C(t) \quad (3)$$

$$v_{int}(t) = LC \ddot{v}_C(t) + \frac{L}{R} \dot{v}_C(t) + v_C(t) \quad (4)$$

By (4) we have:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = \ddot{v}_C(t) = \frac{1}{LC} v_{int}(t) + \frac{1}{RC} x_2(t) - \frac{1}{LC} x_1(t) \end{cases} \quad (5)$$

So:

$$A = \begin{pmatrix} 0 & 1 \\ \frac{1}{LC} & \frac{1}{RC} \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 \\ \frac{1}{LC} \end{pmatrix}.$$

Figure 3 shows the waveforms of  $V_{inv}(t)$  during a sampling period  $T_s$ , where  $V_{inv}$  is the output voltage of the inverter and has two values,  $+E$  and  $-E$ .

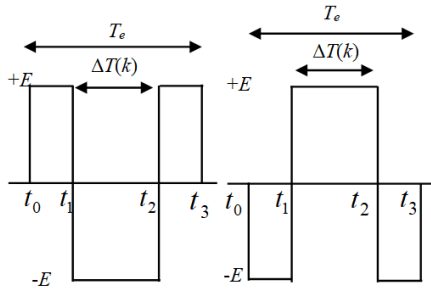


Fig. 3. Two-level PWM pattern.

The discrete-time system equation of the two-level switching patterns is [14]:

$$x(t) = \exp[A(t-t_0)]x(t_0) + \int \exp[A(t-\tau)]B v_{\text{int}}(\tau) d\tau \quad (6)$$

Thus, the discrete state equation from  $t_0$  to  $t_1$  and  $t_2$  to  $t_3$  is:

$$X[(k+1)T_e] = \exp[AT_e]X(k) + \exp\left[A\frac{T_e}{2}\right]BE\Delta T(k), E\Delta T(k) = v_{\text{int}}(\tau) \quad (7)$$

We will find from (8) the value of the duration  $\Delta T$  that we must apply at all times to generate the control signals of the switches.  $\Delta T(k)$  equals to the pulse-width in the  $k^{\text{th}}$  sampling interval. Assuming  $T \ll 2\pi\sqrt{LC}$ , the exponential quantities in (7) are approximated using the power series expansion:

$$A_k = \exp[AT] = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \quad (8)$$

$$A_k = \begin{bmatrix} 1 - \frac{T^2}{2LC} & T - \frac{T^2}{2RC} \\ -\frac{1}{LC} + \frac{T^2}{2RLC^2} & 1 - \frac{T}{RC} + \left(\frac{1}{R^2C^2} + \frac{1}{LC}\right)\frac{T^2}{2} \end{bmatrix}$$

In order to improve this problem, a predictive dead-beat control method is proposed, in which the control input at the  $k^{\text{th}}$  sampling time is predicted by using the system values at the  $(k-1)^{\text{th}}$  sampling time. PWM modulation signal, i.e. the reference output voltage of the inverter, calculated by the reference voltage  $V_{\text{ref}}$ , the actual capacitor current  $i_C$ , and the output load voltage of the inverter  $V_{\text{ch}}$  at the  $(k-1)^{\text{th}}$  sampling time calculated with the reference voltage can be described as:

$$X(k+1) = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} X(k) + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} E\Delta T(k) \quad (9)$$

The first line of (9) gives us the recurrent expression of the sampled output voltage:

$$\Delta T = \frac{1}{g_1 E} (V_{\text{ref}}(k+1) - \phi_{11}x_1 - \phi_{12}x_2) \quad (10)$$

#### IV. STATE OBSERVER

In the control system, observers can replace sensors. We can't estimate the state variables for which the observer is utilized. For the proposed control, capacitor current measurement is necessary and the principle of the observer is combining the feedback signal measured with the known components of the control system. The plant's behavior may be predicted with more precision than when the feedback signal is used alone. The observer augments the sensor output and provides a feedback signal to the control laws, as shown in Figure 4.

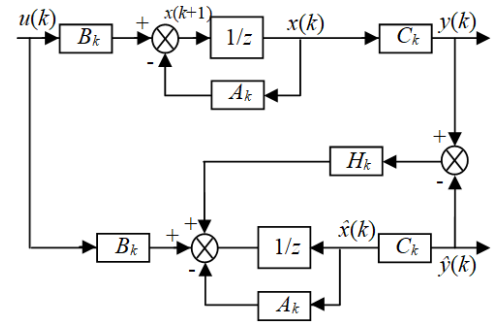


Fig. 4. Diagram of an observer.

##### A. Designing the State Observer

To introduce an observer to the system presented by (9), the system must be observable. For a system to be observable, it is sufficient that the size of the rows of the observability matrix  $O$  is equal to the number of state variables  $n$ . Hence the observability matrix is:

$$O = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \phi_{11} & \phi_{12} \end{pmatrix} \quad (11)$$

$\text{Rang}(O) = 2 = n$ , verified condition.

##### B. The Observer Model

The state representation for a linear discrete system is:

$$\text{SYS} = \begin{cases} x(k+1) = A_k x(k) + B_k u(k) \\ y(k) = C_k x(k) \end{cases} \quad (12)$$

where  $A_k = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$ ,  $B_k = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}$ ,  $C_k = \begin{pmatrix} 1 & 0 \end{pmatrix}$ .

The state vector of the system  $x(k)$ , is reconstructed using an observer whose dynamic equation, based on Luenberger's method, is expressed as:

$$\begin{aligned} \hat{x}(k+1) &= A_k \hat{x}(k) + B_k u(k) + H_k (y(k) - \hat{y}(k)) \\ \hat{y}(k) &= C_k \hat{x}(k) \end{aligned} \quad (13)$$

where  $u(k)$  is defined as input,  $\hat{y}(k)$  is the estimated output,  $H_k$  is the observer's gain, and  $\hat{x}(k)$  is the estimated state vector.



The error  $e(k) = x(k) - \hat{x}(k)$  must converge to zero for all values of  $k$  for the observer to be stable. When the observer error satisfies the equation  $e(k+1) = (A_k - H_k \cdot C_k) e(k)$ , then  $(A_k - H_k \cdot C_k)$  has its eigenvalue  $\lambda_i$  inside the unit circle. The determination of this matrix means that  $|\lambda_{i=1,...,n}| < 1$ .

The choice of the eigenvalues must be such that the module is lower than 1 in order to ensure stability, the real part is closer to  $z = 0$  to ensure a faster real part, and the imaginary part is faster if the angle to the real axis increases, but it should not be very close to the unit circle because otherwise the pole will be more resonant. Therefore  $\lambda_1$  and  $\lambda_2$  are chosen as:

$$\begin{cases} \lambda_1 = 0.1 + j0.1 \\ \lambda_2 = 0.1 - j0.1 \end{cases} \quad (14)$$

By calculating the determinant of the matrix  $(zI - A_k + H_k C_k)$  and by identification on the polynomial  $(z - \lambda_1)(z - \lambda_2)$  we can find  $h_1$  and  $h_2$ :

$$\det(zI - A_k + H_k C_k) = (z - \lambda_1)(z - \lambda_2) \quad (15)$$

The components of  $H_k$  are given by:

$$\begin{cases} h_1 = \phi_{11} + \phi_{22} - \lambda_1 - \lambda_2 \\ h_2 = \frac{1}{\phi_{12}} [\lambda_1 \lambda_2 + \phi_{22} (\phi_{22} - \lambda_1 - \lambda_2)] + \phi_{21} \end{cases} \quad (16)$$

## V. SIMULATION STUDY

Simulations and tests can be used to determine the performance of the above suggested control technique using the DSP F28335 to implant the control algorithm in the digital circuit and IGBT switching devices for the single phase inverter. The inverter's switching frequency is set to 10kHz, and the algorithm's sample period is set to 0.0001s. The inverter's design parameters are given in Table I.

TABLE I. PARAMETERS OF THE INVERTER

Parameter	Value
Rated output voltage $V_c$	220V
DC link voltage $E$	400V
Filter inductance $L$	2mH
Resistance $R$	20Ω
Filter capacitor $C$	20μF
Sampling frequency $F_s$	10kHz

The system's simulation results are presented with and without the presence of the observer, in order to show the role played by the latter in replenishing the  $V_c$  voltage as well as to ensure a good release of the disturbance when connecting charges occur.

## VI. PIL IMPLEMENTATION

The DSP is connected to the computer with a USB cable, which is the hardware configuration required for PIL simulation as shown in Figure 9.

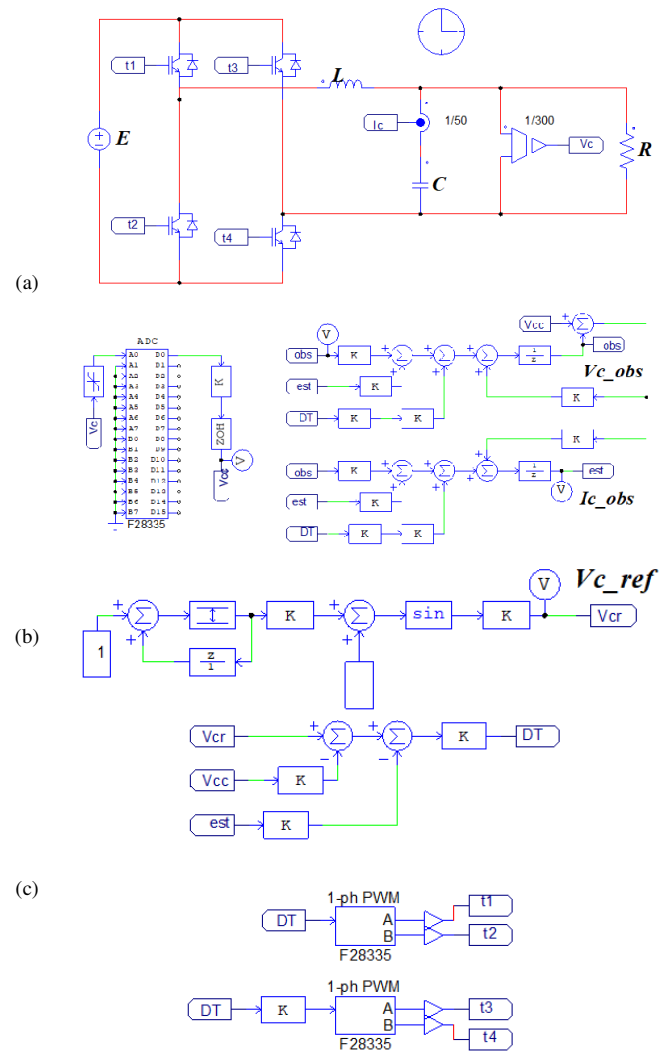


Fig. 5. Simulation of a deadbeat control with a PSIM observer. (a) Lyanberger observatory, (b) reference tension, (c) dead-beat control.

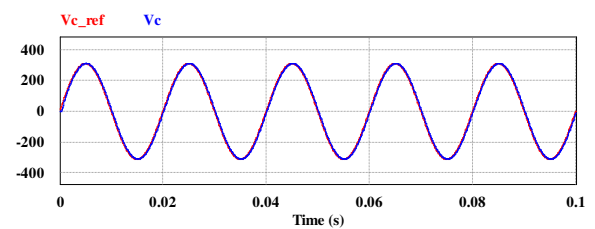


Fig. 6. Simulation result of the system without an observer.

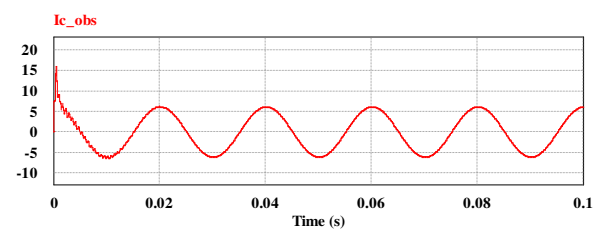


Fig. 7. Current observer of the capacitor.

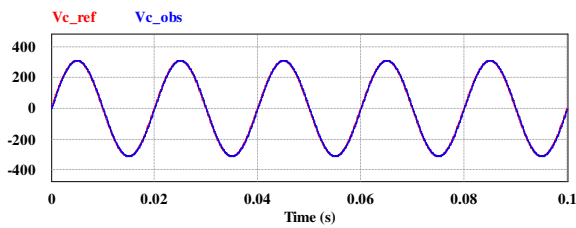


Fig. 8. Tension estimation with the reference tension.

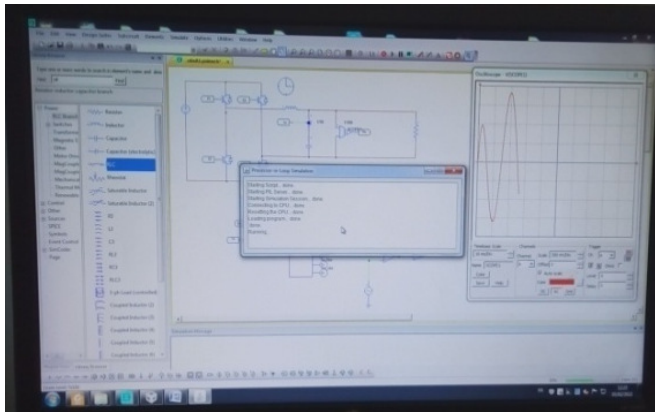


Fig. 9. PIL of dead-beat control whit observer.

Figure 10 shows a simulated block diagram with the PIL structure developed by PSIM in order to achieve safe and rapid prototyping. The control blocks (dead-beat) have been deleted and the creator code CCS has been added instead. This block uses the UART communication protocol to establish communication between the PSIM software and the FA28253. The inverter output voltage  $V_C$  values, which are input from the contact block, are supplied to the microcontroller at each step of the PSIM. The microcontroller gets the switching data ( $t_1$ ,  $t_2$ ,  $t_3$ , and  $t_4$ ) of the inverter control, which is the output of the contact block. The cooperative effort between FA28253 and PSIM is built using this structure.

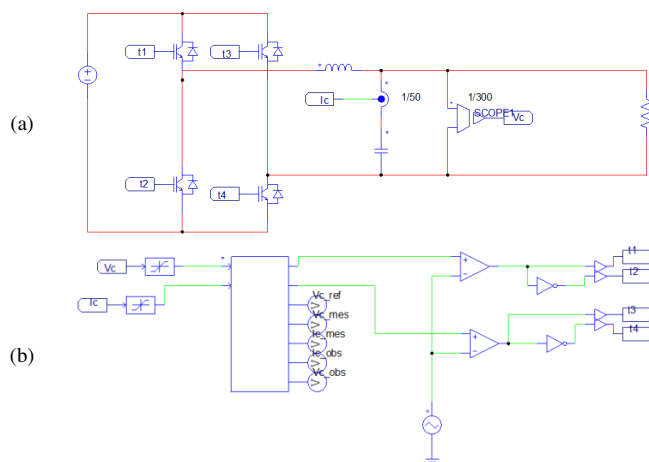
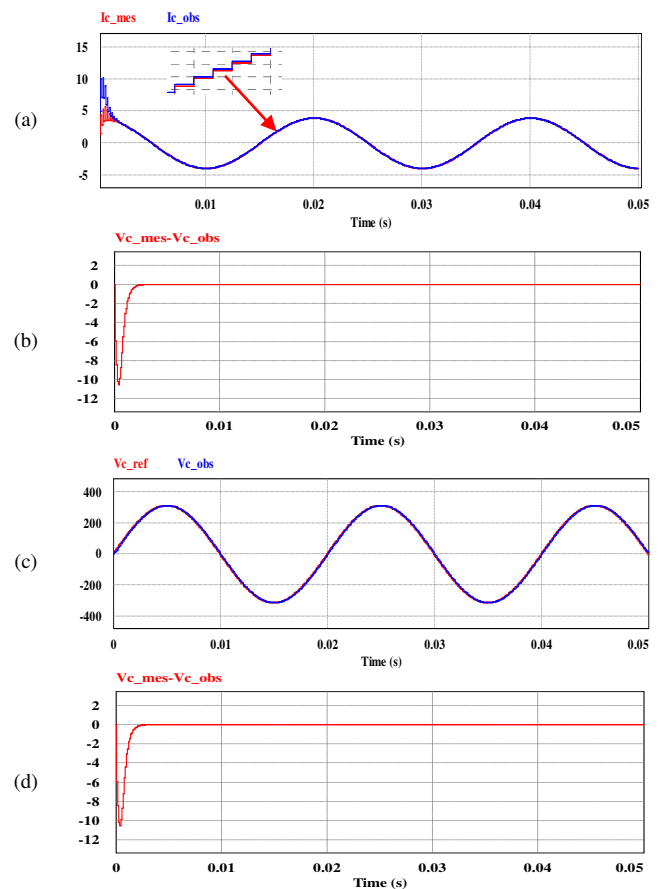


Fig. 10. PIL simulation of deadbeat control whit observer.

#### A. Linear Load with $R=20\Omega$

Figure 11 shows the simulation result of the deadbeat control with linear load equal to  $20\Omega$ . A current sensor and the observer's current are used to show the effect of load variation, nonlinear load, and purely inductive load on the controller. We start with preliminary tests to show the robustness of the linear overload command. Purely sinusoidal current and voltage are applied, the current being phase shifted  $90^\circ$  ahead. The error of the current and voltage between the measurement system and the estimation is close to zero, the feedback fault of the measured voltage is controlled by the observer and is injected in the current observer.

Fig. 11. PIL simulation of dead-beat control with observer: (a) current  $i_c$  (measured and observed), (b) error current, (c) voltage  $V_C$  (reference and observed), (d) voltage error.

#### B. Linear Load $R=20\Omega$ with Increasing Load Up to 100%

The PIL simulation of dead-beat control with increasing load up to 100% is presented in Figure 12. According to the obtained results, Figure 12 shows the PIL results for output voltage and current under a resistive load, at  $t = 0.045s$ . When the load is increased to 100%, we observe a voltage loss, which the control quickly recovers. In addition, the current is controlled relatively well, so it can be concluded that the controller is good.

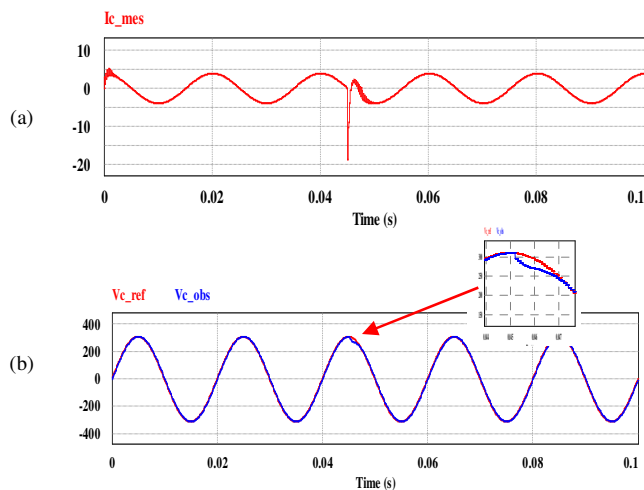


Fig. 12. PIL simulation of dead-beat control with increasing load up to 100%, equal to  $R=40\Omega$ : (a) current  $i_C$  (measured and observed), (b) voltage  $V_C$  (reference and observed).

### C. Nonlinear Load Diode Bridge with $R=20\Omega$ and $C=30\mu F$ in Parallel

The proposed controller was then tested using a distorting load (a diode bridge with a capacitive filter of  $30\mu F$  and a resistive load of  $20\Omega$ ). The output voltage remains sinusoidal, but the capacitor current waveform deviates from the ideal sinusoidal waveform, as seen in Figure 13.

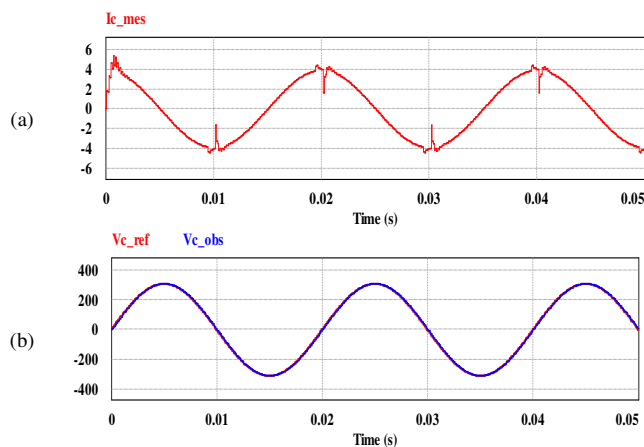


Fig. 13. PIL simulation of dead-beat control with nonlinear load diode bridge with  $R=20\Omega$  and  $C=30\mu F$  in parallel: (a) current  $i_C$  (measured and observed), (b) voltage  $V_C$  (reference and observed).

According to the obtained results, the sudden connection of a nonlinear load has caused a distortion in the output current. However, the proposed controller has eliminated this distortion in the voltage.

### D. Purely Inductive Load $L=0.10H$

Figure 14 presents the PIL simulation of dead-beat control with purely inductive load of  $L=0.10H$ . Figure 14 shows the PIL results for the output voltage and current for a purely inductive load.

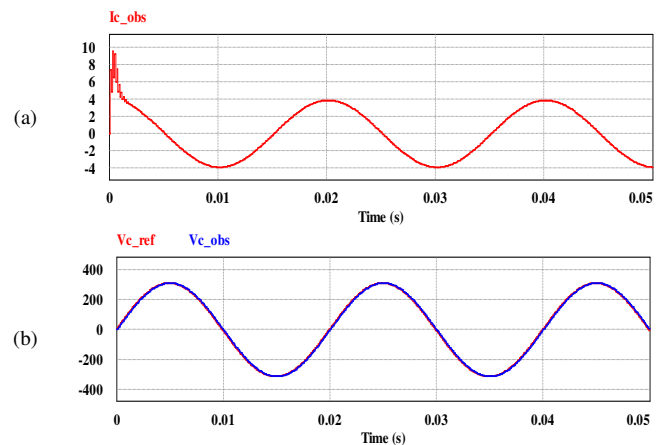


Fig. 14. PIL simulation of dead-beat control with purely inductive load  $L=0.10H$ : (a) current  $i_C$  (measured and observed), (b) voltage  $V_C$  (reference and observed).

We observe a voltage loss, which the control quickly recovers, but the capacitor current waveform deviates from the ideal sinusoidal waveform.

## VII. CONCLUSION

In this work, a dead-beat control scheme for a single-phase inverter with an observer is described. The feasibility of the proposed controller with an observer has been demonstrated by simulations and was implemented in a processor in the PIL loop. The results showed that the proposed scheme achieves good voltage regulation with linear and nonlinear loads. The proposed controller does not have any parameters that can be changed, it needs a model of the system to calculate the controlled variables, allowing fast dynamic response of voltage control. It has been proved that the use of an observer allows a better estimation of the unknown capacitor current.

Dead-beat control serves a different approach for the control of power converters, considering the discrete nature of the converters and the microprocessors used for control. Moreover, the high calculation power of the current existing DSPs makes this method very attractive for controlling power converters.

Since the dead-beat controlled system's efficiency is successfully proved, the next steps are to study the dead-beat control of a grid-connected inverter, the islanded operation of an inverter, and the application of the proposed control to other inverter structures.

## REFERENCES

- [1] J. M. Carrasco *et al.*, "Power-Electronic Systems for the Grid Integration of Renewable Energy Sources: A Survey," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 4, pp. 1002–1016, Jun. 2006, <https://doi.org/10.1109/TIE.2006.878356>.
- [2] F. Blaabjerg, R. Teodorescu, M. Liserre, and A. V. Timbus, "Overview of Control and Grid Synchronization for Distributed Power Generation Systems," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 5, pp. 1398–1409, Jul. 2006, <https://doi.org/10.1109/TIE.2006.881997>.
- [3] P. Cortes, G. Ortiz, J. I. Yuz, J. Rodriguez, S. Vazquez, and L. G. Franquelo, "Model Predictive Control of an Inverter With Output SLC Filter for UPS Applications," *IEEE Transactions on Industrial*

- Electronics*, vol. 56, no. 6, pp. 1875–1883, Jun. 2009, <https://doi.org/10.1109/TIE.2009.2015750>.
- [4] Y. Danayiyen, K. Lee, M. Choi, and Y. I. Lee, "Model Predictive Control of Uninterruptible Power Supply with Robust Disturbance Observer," *Energies*, vol. 12, no. 15, Jan. 2019, Art. no. 2871, <https://doi.org/10.3390/en12152871>.
- [5] D. A. Tuan, P. Vu, and N. V. Lien, "Design and Control of a Three-Phase T-Type Inverter using Reverse-Blocking IGBTs," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6614–6619, Feb. 2021, <https://doi.org/10.48084/etasr.3954>.
- [6] P. Mattavelli, "An improved deadbeat control for UPS using disturbance observers," *IEEE Transactions on Industrial Electronics*, vol. 52, no. 1, pp. 206–212, Oct. 2005, <https://doi.org/10.1109/TIE.2004.837912>.
- [7] M. Wang, F. Z. Li, L. P. Huang, and M. Sakane, "A Robust Dead-Beat Control Method for Ups Inverters," *Advanced Technology of Electrical Engineering and Energy*, vol. 26, no. 4, 2007.
- [8] G. Kecun and D. Yuxing, "DSP Control Method of Single-phase Inverters for UPS Applications," in *2007 Chinese Control Conference*, Zhangjiajie, China, Jul. 2007, pp. 670–672, <https://doi.org/10.1109/CHICC.2006.4346965>.
- [9] H.-T. Moon, H.-S. Kim, and M.-J. Youn, "A discrete-time predictive current control for PMSM," *IEEE Transactions on Power Electronics*, vol. 18, no. 1, pp. 464–472, Jan. 2003, <https://doi.org/10.1109/TPEL.2002.807131>.
- [10] L. Zhang *et al.*, "Improvement of Deadbeat Control for PV Converter," in *2017 2nd International Conference on Electrical, Control and Automation Engineering (ECAE 2017)*, Dec. 2017, pp. 202–205, <https://doi.org/10.2991/ecae-17.2018.42>.
- [11] J.-F. Stumper, V. Hagenmeyer, S. Kuehl, and R. Kennel, "Deadbeat Control for Electrical Drives: A Robust and Performant Design Based on Differential Flatness," *IEEE Transactions on Power Electronics*, vol. 30, no. 8, pp. 4585–4596, Dec. 2015, <https://doi.org/10.1109/TPEL.2014.2359971>.
- [12] A. Benyoucef, K. Kara, A. Chouder, and S. Silvestre, "Prediction-based Deadbeat Control for Grid-connected Inverter with L-filter and LCL-filter," *Electric Power Components and Systems*, vol. 42, no. 12, pp. 1266–1277, Sep. 2014, <https://doi.org/10.1080/15325008.2014.927031>.
- [13] I. Alhamrouni, N. Zainuddin, M. Salem, N. H. A. Rahman, and L. Awalin, "Design of single phase inverter for photovoltaic application controlled with sinusoidal pulse width modulation," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 2, pp. 620–630, Aug. 2019, <https://doi.org/10.11591/ijeecs.v15.i2.pp620-630>.
- [14] Z. Boumous and S. Boumous, "New Approach in the Fault Tolerant Control of Three-Phase Inverter Fed Induction Motor," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6504–6509, Dec. 2020, <https://doi.org/10.48084/etasr.3898>.
- [15] Z. Wang, K. Zhou, S. Li, and Y. Yang, "Fractional-order time delay compensation in deadbeat control for power converters," in *2018 IEEE International Power Electronics and Application Conference and Exposition (PEAC)*, Shenzhen, China, Aug. 2018, <https://doi.org/10.1109/PEAC.2018.8590593>.
- [16] T. L. Tiang and D. Ishak, "Modeling and simulation of deadbeat-based PI controller in a single-phase H-bridge inverter for stand-alone applications," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 22, no. 1, pp. 43–56, Jan. 2014, <https://doi.org/10.3906/elk-1206-45>.
- [17] S. Paul, A. Halder, and A. K. Nath, "Deadbeat Control of Linear and Nonlinear System using Signal Correction Technique," *MAYFEB Journal of Electrical and Computer Engineering*, vol. 2, pp. 1–23, Feb. 2017.
- [18] S. Bayhan, H. Komurcugil, and I. S. Bayram, "Deadbeat Control of a Three-Phase T-type Inverter with Output LC Filter for UPS Applications," in *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)*, Kyoto, Japan, Jun. 2021, <https://doi.org/10.1109/ISIE45552.2021.9576216>.
- [19] W. Yao, J. Cui, and W. Yao, "Single-Phase Inverter Deadbeat Control with One-Carrier-Period Lag," *Electronics*, vol. 9, no. 1, Jan. 2020, Art. no. 154, <https://doi.org/10.3390/electronics9010154>.
- [20] W. Guo, S. Duan, K. Xuejuan, Y. Kang, and J. Chen, "A modified deadbeat control for single-phase voltage-source PWM inverters based on asymmetric regular sample," in *2001 IEEE 32nd Annual Power Electronics Specialists Conference (IEEE Cat. No.01CH37230)*, Vancouver, BC, Canada, Jun. 2001, vol. 2, pp. 962–967, <https://doi.org/10.1109/PESC.2001.954244>.
- [21] A. S. Haider, A. Naz, F. Akhter, W. Azhar, and R. Mehmood, "A Novel Efficient Sine Wave Inverter with Custom Programmed PWM and Intelligent Control," *Engineering, Technology & Applied Science Research*, vol. 6, no. 2, pp. 956–963, Apr. 2016, <https://doi.org/10.48084/etasr.648>.

# Behavior of a Circular Footing resting on Sand Reinforced with Geogrid and Grid Anchors

**Rima Helis**

LGC-ROI Civil Engineering Laboratory, Department of Civil Engineering, Faculty of Technology,  
University of Batna 2, Algeria  
helis.r@univ-setif.dz  
(corresponding author)

**Tarek Mansouri**

LGC-ROI Civil Engineering Laboratory, Department of Civil Engineering, Faculty of Technology,  
University of Batna 2, Algeria  
t.mansouri@univ-batna2.dz

**Khelifa Abbeche**

LGC-ROI Civil Engineering Laboratory, Department of Civil Engineering, Faculty of Technology,  
University of Batna 2, Algeria  
k.abbeche@univ-batna2.dz

*Received: 23 December 2022 | Revised: 3 January 2023 | Accepted: 7 January 2023*

## ABSTRACT

This study used finite element analysis to investigate the influence of using two reinforcing systems, the geogrid and the grid anchor, on the bearing capacity of a circular footing resting on sand. The parameters studied were the effect of the number of reinforcement layers (N), the depth ratio of the topmost layer of reinforcement ( $u/d$ ), the vertical spacing ratio between consecutive layers ( $h/d$ ), and the effect of reinforcement length (L). The results showed that the reinforcement layout had a very significant effect on the behavior of the reinforced sand foundation. The maximum bearing capacity for single-layer inclusion was obtained when reinforcement was placed at a depth of  $u/d=0.42$ . Bearing capacity was also found to improve when increasing the number of reinforcement layers from 1 to 3. Additionally, the analysis showed that the sand reinforced by grid anchors performed better than that reinforced by geogrid. Finally, an improvement in load capacity was obtained by increasing the length of the inclusions, and the optimal length of the reinforcements was determined at 5d for both inclusions.

**Keywords-***geogrid; grid anchor; finite element analysis; circular footing; sand*

## I. INTRODUCTION

Geosynthetic reinforcing techniques is used to reinforce shallow foundations, improve bearing capacity, and reduce soil settlement below the foundation. Several studies investigated the bearing capacity of geosynthetic reinforced foundation soils using experimental, analytical, and numerical methods. One of the first experimental studies to analyze the bearing capacity of reinforced soils with metal strips was presented in [1-2]. Since then, many studies investigated the improvement of the load-bearing capacity of shallow foundations supported by sand and reinforced with different materials, such as metal strips and metal bars [3-5], rope fibers [6], geotextile [7], and geocells [8-10]. In addition, considerable studies were conducted to evaluate the bearing capacity of the reinforced soil by geogrid [11-18]. These studies confirmed the beneficial effect of reinforcement on improving the bearing capacity and reducing

the settlement of footing. More recently, the use of geogrid in geotechnical engineering applications was considerably increased due to advantages such as cost reduction, simplicity, and ease of construction [19-20]. Laboratory scale model tests on a circular embedded footing supported on geogrid-reinforced sand beds were presented in [21], reporting an increase in ultimate bearing capacity with the embedding depth ratio of the foundation. In [22], the behavior of circular footing on sand was studied, showing that bearing capacity increased when the number of reinforcement layers increased if the reinforcements were placed within a range of effective depths. This study also showed that increasing the stiffness of the reinforcement did not always have a better effect on bearing capacity. A numerical study was conducted in [23] using finite element analysis to investigate the behavior of circular footing resting over reinforced sand, showing that the depth of the top layer plays an important role in the behavior of the reinforced

soil, and reporting that the optimum depth of the top layer was 0.19 times the diameter of the footing. In [24], a small-scale laboratory model test was carried out on two closely spaced interfering footings resting on the surface of unreinforced and geogrid-reinforced sand bed, finding that the optimal depth of the geogrid layer for both interfering and isolated footings was one-third of the footing width. Therefore, footing interference had negligible or no effect on the optimum depth of the reinforcement layer for a single-layered reinforced sand bed. In [25], resting on a semi-infinite layer of reinforced sand with geotextiles was used to experimentally study the behavior of circular footings. Furthermore, analytical and numerical analyses were carried out to predict load-settlement behavior and compare them with the experimental observations. In [26], model plate load tests were conducted on various types of sand beds reinforced with geogrid, showing that substantial improvement in the load-settlement behavior can be obtained by increasing the number of geogrid layers (N) and decreasing the spacing between them. It was also shown that the load improvement ratio for the reinforced coarse sand was higher than that of the reinforced fine and medium sand. In [27], the behavior of geosynthetic-reinforced sandy soil foundations was investigated using laboratory model tests, showing that the settlement can be reduced by 20% at all footing pressure levels with two or more layers of geogrid. In [28], the upper bound theorem of limit analysis was used in conjunction with finite elements and linear optimization to determine the bearing capacity of a circular foundation embedded with horizontal layers of circular geogrid sheets. The optimal diameter and the critical positions of the reinforcement layers were established to achieve maximum bearing capacity, and a marked improvement in the bearing capacity was evident in the case of using two layers of reinforcement rather than a single. A laboratory model test of a surface strip footing on reinforced sand beds was presented in [29] to investigate the effects of reinforcement length with various types and numbers of reinforcements. An experimental study was conducted in [30] to assess the influence of the geogrid extension and embedment depth below strip footing rested on fine loose sand. In [31], a regression model was developed to determine the bearing capacity of a circular foundation supported on sand reinforced with geogrid. The results showed that the parameters studied had a significant influence on the performance of the footing in terms of bearing capacity. In [32-33], the bearing capacity of a strip footing subjected to inclined load and resting over a geogrid-reinforced sand bed was studied experimentally and numerically, showing that the footing performance could be substantially improved by including layers of geogrid, leading to an economic design of the footing. The effects of load eccentricity and inclination on the ultimate bearing capacity of shallow rectangular foundations placed over geogrid sand were studied in [34], finding that multiple geogrid-reinforced layers increased the ultimate bearing capacity by 75%. In [35], a strip foundation in weak soil was replaced with a granular trench and reinforced with geogrid, showing that the bearing capacity of a strip foundation could be significantly improved by replacing sand with granular materials up to 3 times. In addition, it was shown that placing the geogrid in the trench indicated a rise in the bearing capacity ratio. A new generation of reinforcement named grid-anchor was introduced in [36-37],

showing its effect on the increase of bearing capacity of the foundation.

This study aims to evaluate the performance of using ordinary geogrid and grid anchor reinforcement in increasing the bearing capacity and reducing the settlement. To achieve this objective, a numerical model was determined using the Plaxis finite element software to investigate the bearing capacity of a circular footing resting over reinforced sand. Different parameters that affect the behavior of the reinforcement sand layer are discussed.

## II. FINITE ELEMENT MODELING

The Plaxis software was utilized to perform a numerical finite element analysis by simulating a circular footing resting on sand reinforced by two reinforcement systems, GeoGrid (GG) and Grid Anchors (GA). Also, an axisymmetric analysis was performed. For all models, the boundary conditions in displacements were similar, such that the bottom boundary was assumed to be fixed and the vertical boundaries were constrained in motion in the horizontal direction. However, sand's behavior was supposed to be elastic and perfectly plastic, the Mohr-Coulomb rupture criterion was used, and the nonassociated flow rule was considered. A rigid circular footing with a 12cm diameter was simulated by applying a uniform downward displacement on the surface of the sandy soil. Table I shows the properties of the sand adopted in the model. Fifteen triangular plane strain elements were selected to model the soil, while the GG reinforcement was simulated with 5 node elastic elements. The GA was modeled using the fixed-end anchor option. Table II shows the physical and mechanical properties of GG and GA. The mesh refinement was adopted in the vicinity of the loading area around the foundation and GG layers to improve the accuracy of the numerical results.

TABLE I. SOIL PARAMETERS

Physical and Mechanical Property	Value
Maximum unit weight (kN/m <sup>3</sup> )	16.4
Minimum unit weight (kN/m <sup>3</sup> )	14.4
Maximum void ratio	0.890
Minimum void ratio	0.658
Specific gravity	2.72
Coefficient of uniformity	2.36
Coefficient of curvature	1.01
Classification	SP
Cohesion (kN/m <sup>2</sup> )	0
Internal friction angle	39°

TABLE II. PHYSICAL AND MECHANICAL PROPERTIES OF GEOGRID AND ANCHORS

Description	Geogrid CE 131
Polymer	High-density polyethylene
Form	Sheet
Color	Black
Mesh aperture size	27×27mm
Mesh thickness	5.2mm
Structural weight (+5%)	660g/m <sup>2</sup>
Elastic normal stiffness of geogrid	28.0KN/m
EA axial stiffness of anchors	0.18KN
Length of anchors (mm)	50mm



Figure 1 shows the prototype soil model with two systems of reinforcement, finite element mesh, and boundary conditions.

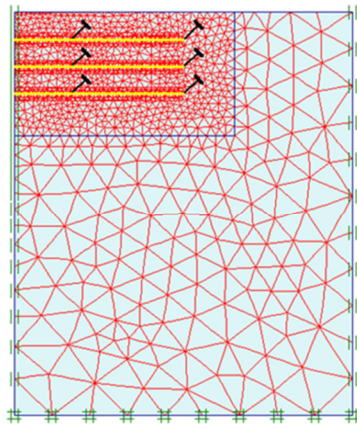


Fig. 1. The numerical model.

### III. RESULTS AND DISCUSSION

Numerical tests were carried out to study the effects of inclusion reinforcement elements on a circular footing, constructed on unreinforced and multi-layered reinforced sand beds, and investigate the improvement of bearing capacity. A non-dimensional factor called the Bearing Capacity Ratio (BCR) was considered, defined as the ratio of the reinforced soil bearing capacity to the unreinforced soil:

$$BCR = \frac{q_R}{q_U} \quad (1)$$

where  $q_R$  and  $q_U$  are the bearing capacity values for reinforced and unreinforced soil foundations, respectively.

#### A. Effect of Reinforcement's Top Spacing

A numerical study was carried out to investigate the effect of the depth of the first reinforcing layer from the footing on the bearing capacity for different depth ratio values ( $u/d$ ) with a single reinforcement layer in each reinforcing system, GG and GA. Figure 2 shows the variation of the BCR of the soil versus the different reinforcement depth ratios  $u/d$ . In the case of GG, as the depth ratio  $u/d$  increases from 0.2 to 0.42, the BCR also increases. However, between 0.42 and 0.8, a clear reduction in the BCR was found for both GG and GA.

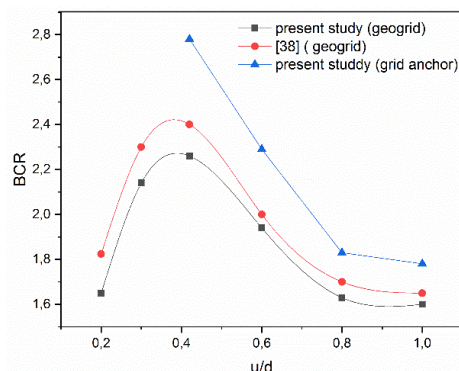


Fig. 2. Variation of BCR with depth ratio in single-layer reinforced sand.

Similar results were found in [31], where there is no increase in soil carrying capacity that exceeded  $u/d > 0.75$ . Hence, the optimal value of the depth ratio was obtained when the reinforcement was placed at  $u/d$  equal to 0.42 in both systems. Therefore, it can be concluded that the results obtained for GG reinforcement are in good agreement with [38]. Figure 2 also shows that the effect of the presence of GA reinforcement on the bearing capacity of the circular footing on sand becomes important compared to those obtained by GG. In addition, a considerable improvement of about 52% was observed for the anchorage of the grids.

#### B. Effect of Vertical Spacing of Reinforcement Layers

This study aims to investigate the effect of the spacing between the reinforcing elements on the performance of reinforced sand under the circular footing. The GG and GA layers were tested with a top layer spacing at  $0.42d$  and varied vertical spacing between the layers. Figure 3 shows the variation of the BCR with the vertical spacing ratio ( $h/B$ ). The results showed that for GG reinforcement, the BCR increased to a maximum value at  $h/d = 0.3d$ , but the GA had a critical value at  $u = 0.42d$ . Then, a remarkable decrease was observed for both reinforcements until  $0.6d$ . Beyond this value, BCR seems to stabilize, showing that adding inclusions is insignificant in this region. The trend of the curves is similar to that of [38]. Furthermore, in [28-31] it was shown that the increase in BCR was obtained when the vertical spacing between the reinforcement layers was between 0.25 and  $0.40d$ , which justifies the present case study. Therefore, the variation in amplitudes and the modest divergence can be attributed to the adapted reinforcement pattern.

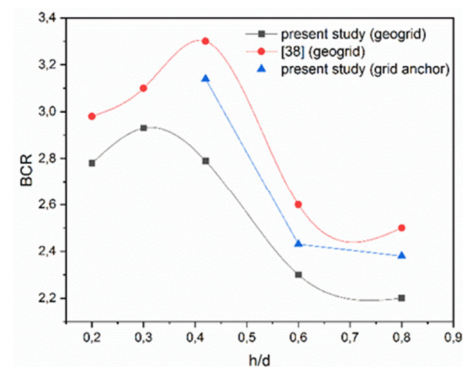


Fig. 3. Variation of BCR with the vertical spacing ratio of reinforced sand.

#### C. Effect of the Number of Reinforcing Layers

A series of numerical tests were conducted to study the influence of the variation of the number of reinforcement elements ( $N$ ) on the behavior of a circular footing on reinforced sand. The depth of the first layer ( $u$ ) was taken equal to  $0.42d$ , while the vertical distance between the reinforcement layers ( $h$ ) was equal to  $0.3d$  for the GG and  $0.42d$  for the GA. Figure 4 shows the variation of the BCR as a function of the number of reinforcement elements  $N$ . It can be observed that the increase in the BCR results from a considerable increase in the reinforcement elements up to an optimum value  $N=3$ , and a slight increase is observed over that. This confirms the findings

of several studies [22-34, 37] that showed that increasing the number of reinforcement layers beyond a certain number would not increase the BCR.

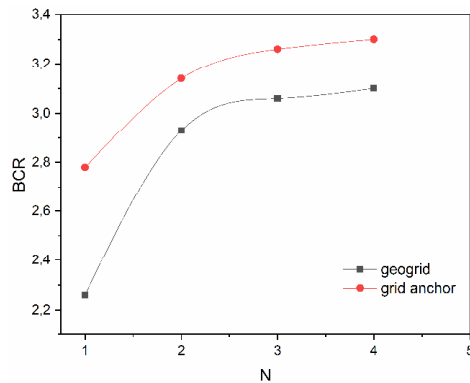


Fig. 4. Variation of BCR with the number of reinforcement layers.

#### D. Effect of the Length of Reinforcement

A circular foundation resting on sand was studied by keeping the number of geogrid layers  $N$  to 1 and depth of reinforcement at the optimal  $0.42d$ , while the length of the reinforcement layer ( $L$ ) varied between  $4d$ ,  $4.5d$ ,  $5d$ , and  $6d$  to investigate its effect on BCR. Figure 5 illustrates the variation of BCR with the different reinforcement length ratios ( $L/d$ ).

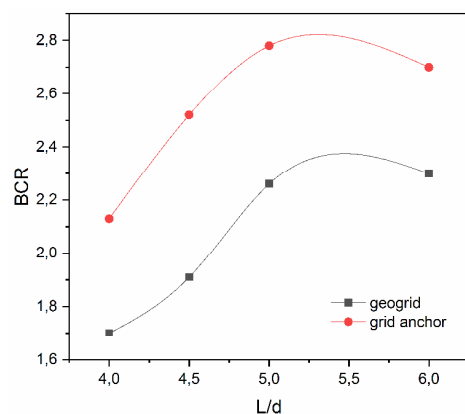


Fig. 5. Variation of BCR with the reinforcement layer's length.

As can be observed, BCR increases linearly with reinforcement length up to  $L/d=5$ , while reinforcement length beyond this value is ineffective on BCR for both reinforcement types. Therefore, the optimal length of reinforcement is obtained at 5 times the length of the footing, as in [39].

#### IV. CONCLUSION

This study caused finite element analysis to assess the behavior of a circular footing constructed on unreinforced and reinforced sand soil, drawing the following conclusions:

- An increase in bearing capacity was obtained when the depth of the first reinforcing layer to the footing diameter

was equal to  $0.42d$ . This was considered an optimal depth of the top reinforcement layer from the bottom of the footing.

- A visible reduction was noted in the bearing capacity beyond  $0.42d$  for both types of reinforcement.
- The effect of reinforcement cannot be seen when the installation depth is deeper than a certain depth ( $u/d \geq 0.80$ ).
- There is an optimal value for the vertical spacing of the reinforcement layer where the BCR was the highest. This optimal value was found to  $0.3d$  for Geogrid and  $0.42d$  for Grid Anchors.
- The bearing capacity of reinforced soil increases with increasing the number of layers. In this study, the optimal number of layers obtained was 3 in both reinforcement types.
- The analysis clearly showed that using the Grid Anchor system reinforcement of circular footing on a sand bed causes a significant increase in the bearing capacity in comparison with the ordinary Geogrid.
- An improvement in the load capacity was obtained by increasing the length of the inclusions. The optimal length of the reinforcements was determined at  $5d$  for both inclusions.

#### REFERENCES

- [1] J. Binquet and K. L. Lee, "Bearing Capacity Tests on Reinforced Earth Slabs," *Journal of the Geotechnical Engineering Division*, vol. 101, no. 12, pp. 1241–1255, Dec. 1975, <https://doi.org/10.1061/AJGEB6.0000219>.
- [2] J. Binquet and K. L. Lee, "Bearing Capacity Analysis of Reinforced Earth Slabs," *Journal of the Geotechnical Engineering Division*, vol. 101, no. 12, pp. 1257–1276, Dec. 1975, <https://doi.org/10.1061/AJGEB6.0000220>.
- [3] R. J. Frigaszy and E. Lawton, "Bearing Capacity of Reinforced Sand Subgrades," *Journal of Geotechnical Engineering*, vol. 110, no. 10, pp. 1500–1507, Oct. 1984, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1984\)110:10\(1500\)](https://doi.org/10.1061/(ASCE)0733-9410(1984)110:10(1500)).
- [4] C. C. Huang and F. Tatsuoka, "Prediction of bearing capacity in level sandy ground reinforced with strip reinforcement," presented at the International geotechnical symposium on theory and practice of earth reinforcement, Fukuoka, Japan, 1988, pp. 191–196.
- [5] C. C. Huang and F. Tatsuoka, "Bearing capacity of reinforced horizontal sandy ground," *Geotextiles and Geomembranes*, vol. 9, no. 1, pp. 51–82, Jan. 1990, [https://doi.org/10.1016/0266-1144\(90\)90005-W](https://doi.org/10.1016/0266-1144(90)90005-W).
- [6] J. O. Akinmusuru and J. A. Akinbolade, "Stability of Loaded Footings on Reinforced Soil," *Journal of the Geotechnical Engineering Division*, vol. 107, no. 6, pp. 819–827, Jun. 1981, <https://doi.org/10.1061/AJGEB6.0001153>.
- [7] V. A. Guido, G. L. Biesiadecki, and M. J. Sullivan, "Bearing capacity of a geotextile-reinforced foundation," presented at the International conference on soil mechanics and foundation engineering 11, San Francisco, CA, USA, 1985, pp. 1777–1780.
- [8] S. K. Dash, S. Sireesh, and T. G. Sitharam, "Model studies on circular footing supported on geocell reinforced sand underlain by soft clay," *Geotextiles and Geomembranes*, vol. 21, no. 4, pp. 197–219, Aug. 2003, [https://doi.org/10.1016/S0266-1144\(03\)00017-7](https://doi.org/10.1016/S0266-1144(03)00017-7).
- [9] A. Hegde and T. G. Sitharam, "3-Dimensional numerical modelling of geocell reinforced sand beds," *Geotextiles and Geomembranes*, vol. 43, no. 2, pp. 171–181, Apr. 2015, <https://doi.org/10.1016/j.geotextmem.2014.11.009>.

- [10] S. N. Moghaddas Tafreshi, T. Shaghagh, Gh. Tavakoli Mehrjardi, A. R. Dawson, and M. Ghadrán, "A simplified method for predicting the settlement of circular footings on multi-layered geocell-reinforced non-cohesive soils," *Geotextiles and Geomembranes*, vol. 43, no. 4, pp. 332–344, Aug. 2015, <https://doi.org/10.1016/j.geotexmem.2015.04.006>.
- [11] V. A. Guido, D. K. Chang, and M. A. Sweeney, "Comparison of geogrid and geotextile reinforced earth slabs," *Canadian Geotechnical Journal*, vol. 23, no. 4, pp. 435–440, Nov. 1986, <https://doi.org/10.1139/t86-073>.
- [12] C. R. Patra, B. M. Das, and C. Atalar, "Bearing capacity of embedded strip foundation on geogrid-reinforced sand," *Geotextiles and Geomembranes*, vol. 23, no. 5, pp. 454–462, Oct. 2005, <https://doi.org/10.1016/j.geotexmem.2005.02.001>.
- [13] D. Loukidis and R. Salgado, "Bearing capacity of strip and circular footings in sand using finite elements," *Computers and Geotechnics*, vol. 36, no. 5, pp. 871–879, Jun. 2009, <https://doi.org/10.1016/j.compgeo.2009.01.012>.
- [14] A. A. Lavasan and M. Ghazavi, "Behavior of closely spaced square and circular footings on reinforced sand," *Soils and Foundations*, vol. 52, no. 1, pp. 160–167, Feb. 2012, <https://doi.org/10.1016/j.sandf.2012.01.006>.
- [15] A. Abdi, K. Abbeche, D. Athmania, and M. Bouassida, "Effective Width Rule in the Analysis of Footing on Reinforced Sand Slope," *Studia Geotechnica et Mechanica*, vol. 41, no. 1, pp. 42–55, Apr. 2019, <https://doi.org/10.2478/sgem-2019-0005>.
- [16] S. Bildik and M. Laman, "Effect of geogrid reinforcement on soil - structure – pipe interaction in terms of bearing capacity, settlement and stress distribution," *Geotextiles and Geomembranes*, vol. 48, no. 6, pp. 844–853, Dec. 2020, <https://doi.org/10.1016/j.geotexmem.2020.07.004>.
- [17] J.-Q. Wang, L. L. Zhang, Y. Tang, and S.-B. Huang, "Influence of reinforcement-arrangements on dynamic response of geogrid-reinforced foundation under repeated loading," *Construction and Building Materials*, vol. 274, Mar. 2021, Art. no. 122093, <https://doi.org/10.1016/j.conbuildmat.2020.122093>.
- [18] B. Mazouz, T. Mansouri, M. Baazouzi, and K. Abbeche, "Assessing the Effect of Underground Void on Strip Footing Sitting on a Reinforced Sand Slope with Numerical Modeling," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9005–9011, Aug. 2022, <https://doi.org/10.48084/etasr.5131>.
- [19] A. Lazizi, H. Trouzine, A. Asroun, and F. Belabdelouhab, "Numerical Simulation of Tire Reinforced Sand behind Retaining Wall Under Earthquake Excitation," *Engineering, Technology & Applied Science Research*, vol. 4, no. 2, pp. 605–611, Apr. 2014, <https://doi.org/10.48084/etasr.427>.
- [20] M. Touahmia, "Performance of Geosynthetic-Reinforced Soils Under Static and Cyclic Loading," *Engineering, Technology & Applied Science Research*, vol. 7, no. 2, pp. 1523–1527, Apr. 2017, <https://doi.org/10.48084/etasr.1035>.
- [21] T. G. Sitharam and S. Sireesh, "Model studies of embedded circular footing on geogrid-reinforced sand beds," *Proceedings of the Institution of Civil Engineers - Ground Improvement*, vol. 8, no. 2, pp. 69–75, Jan. 2004, <https://doi.org/10.1680/grim.2004.8.2.69>.
- [22] J. H. Boushehrian and N. Hataf, "Experimental and numerical investigation of the bearing capacity of model circular and ring footings on reinforced sand," *Geotextiles and Geomembranes*, vol. 21, no. 4, pp. 241–256, Aug. 2003, [https://doi.org/10.1016/S0266-1144\(03\)00029-3](https://doi.org/10.1016/S0266-1144(03)00029-3).
- [23] A. F. Zidan, "Numerical Study of Behavior of Circular Footing on Geogrid-Reinforced Sand Under Static and Dynamic Loading," *Geotechnical and Geological Engineering*, vol. 30, no. 2, pp. 499–510, Apr. 2012, <https://doi.org/10.1007/s10706-011-9483-0>.
- [24] S. Saha Roy and K. Deb, "Effect of aspect ratio of footing on behavior of two closely-spaced footings on geogrid-reinforced sand," *Geotextiles and Geomembranes*, vol. 48, no. 4, pp. 443–453, Aug. 2020, <https://doi.org/10.1016/j.geotexmem.2020.02.003>.
- [25] P. K. Basudhar, S. Saha, and K. Deb, "Circular footings resting on geotextile-reinforced sand bed," *Geotextiles and Geomembranes*, vol. 25, no. 6, pp. 377–384, Dec. 2007, <https://doi.org/10.1016/j.geotexmem.2006.09.003>.
- [26] B. R. Phanikumar, R. Prasad, and A. Singh, "Compressive load response of geogrid-reinforced fine, medium and coarse sands," *Geotextiles and Geomembranes*, vol. 27, no. 3, pp. 183–186, Jun. 2009, <https://doi.org/10.1016/j.geotexmem.2008.11.003>.
- [27] M. Abu-Farsakh, Q. Chen, and R. Sharma, "An experimental evaluation of the behavior of footings on geosynthetic-reinforced sand," *Soils and Foundations*, vol. 53, no. 2, pp. 335–348, Apr. 2013, <https://doi.org/10.1016/j.sandf.2013.01.001>.
- [28] M. Chakraborty and J. Kumar, "Bearing capacity of circular foundations reinforced with geogrid sheets," *Soils and Foundations*, vol. 54, no. 4, pp. 820–832, Aug. 2014, <https://doi.org/10.1016/j.sandf.2014.06.013>.
- [29] E. Cicek, E. Guler, and T. Yetimoglu, "Effect of reinforcement length for different geosynthetic reinforcements on strip footing on sand soil," *Soils and Foundations*, vol. 55, no. 4, pp. 661–677, Aug. 2015, <https://doi.org/10.1016/j.sandf.2015.06.001>.
- [30] S. Abu El-Soud and A. M. Belal, "Bearing capacity of rigid shallow footing on geogrid-reinforced fine sand—experimental modeling," *Arabian Journal of Geosciences*, vol. 11, no. 11, May 2018, Art. no. 247, <https://doi.org/10.1007/s12517-018-3597-0>.
- [31] D. Useche-Infante, G. Aiassa Martinez, P. Arrúa, and M. Eberhardt, "Experimental study of behaviour of circular footing on geogrid-reinforced sand," *Geomechanics and Geoengineering*, vol. 17, no. 1, pp. 45–63, Jan. 2022, <https://doi.org/10.1080/17486025.2019.1683621>.
- [32] R. Sahu, C. R. Patra, B. M. Das, and N. Sivakugan, "Bearing capacity of shallow strip foundation on geogrid-reinforced sand subjected to inclined load," *International Journal of Geotechnical Engineering*, vol. 10, no. 2, pp. 183–189, Mar. 2016, <https://doi.org/10.1080/19386362.2015.1105622>.
- [33] R. Sahu, C. R. Patra, N. Sivakugan, and B. M. Das, "Behavior of Inclined Loaded Strip Footings Resting on Geogrid-Reinforced Sand," *Geotechnical and Geological Engineering*, vol. 38, no. 5, pp. 5245–5256, Oct. 2020, <https://doi.org/10.1007/s10706-020-01360-z>.
- [34] S. Gupta and A. Mital, "A comparative study of bearing capacity of shallow footing under different loading conditions," *Geomechanics and Geoengineering*, vol. 17, no. 4, pp. 1338–1349, Jul. 2022, <https://doi.org/10.1080/17486025.2021.1940310>.
- [35] M. M. Hajitaheriha, D. Akbarimehr, A. Hasani Motlagh, and H. Damerchilou, "Bearing capacity improvement of shallow foundations using a trench filled with granular materials and reinforced with geogrids," *Arabian Journal of Geosciences*, vol. 14, no. 15, 1431, Jul. 2021, Art. no. 1431, <https://doi.org/10.1007/s12517-021-07679-y>.
- [36] A. H. Boushehrian, N. Hataf, and A. Ghahramani, "Numerical Study of Cyclic Behavior of Shallow Foundations on Sand Reinforced with Geogrid and Grid-Anchor," *International Journal of Civil and Environmental Engineering*, vol. 3, no. 10, pp. 390–393, Oct. 2009.
- [37] S. Alamshahi and N. Hataf, "Bearing capacity of strip footings on sand slopes reinforced with geogrid and grid-anchor," *Geotextiles and Geomembranes*, vol. 27, no. 3, pp. 217–226, Jun. 2009, <https://doi.org/10.1016/j.geotexmem.2008.11.011>.
- [38] E. Badakhshan and A. Noorzad, "Effect of footing shape and load eccentricity on behavior of geosynthetic reinforced sand bed," *Geotextiles and Geomembranes*, vol. 45, no. 2, pp. 58–67, Apr. 2017, <https://doi.org/10.1016/j.geotexmem.2016.11.007>.
- [39] K. M. Lee, V. R. Manjunath, and D. M. Dewaikar, "Numerical and model studies of strip footing supported by a reinforced granular fill - soft soil system," *Canadian Geotechnical Journal*, vol. 36, no. 5, pp. 793–806, Nov. 1999, <https://doi.org/10.1139/t99-053>.

# A Fault Diagnosis Technique for Wind Turbine Gearbox: An Approach using Optimized BLSTM Neural Network with Undercomplete Autoencoder

**M. Sreenatha**

JSS Academy of Technical Education, India  
sreenath.jssate@gmail.com  
(corresponding author)

**P. B. Mallikarjuna**

JSS Academy of Technical Education, India  
pbmalli2020@gmail.com

Received: 21 December 2022 | Revised: 3 January 2023 and 7 January 2023 | Accepted: 9 January 2023

## ABSTRACT

The gearbox is one of the critical components of a wind turbine. Proactive maintenance of wind turbine gearboxes is crucial to decrease maintenance and operational costs and the long downtime of the complete system. As the gearbox is a significant part of the wind turbine, a fault in the gearbox leads to the breakdown of the wind turbine system. Hence, it is important to study and analyze the faults in wind turbine gearbox systems. In this article, a neural network-based model, a Bidirectional Long Short-Term Memory (BLSTM) fused with an autoencoder is intended to categorize the condition of the gearbox into good or bad (broken tooth) condition. Feature learning and reduction are achieved extensively through the autoencoder. This improves the performance of the BLSTM model regarding time complexity and classification accuracy. This model has been applied with time series vibration data of the gearbox in a wind turbine system. The suggested model's performance is analyzed using an openly available wind turbine gearbox vibration dataset. The result showed that BLSTM accuracy with an under-complete autoencoder is highly robust and appropriate for the health monitoring of wind turbine gearbox systems using time series data. Also, in order to illustrate the advantage of the projected model for fault analysis and diagnosis in wind turbine gearbox, the throughput or time complexity of training and testing of the split dataset is compared with the conventional BLSTM model.

**Keywords-**autoencoder; bidirectional long short-term memory; fault detection; vibration data; wind turbine gearbox

## I. INTRODUCTION

Wind energy is very popular, since it is form of clean energy. For the effective production of wind energy, the wind turbines must be maintained with less downtime. Gearbox failure directly impacts the reliability of the gearbox in the wind turbine. The operation and the cost towards the maintenance of wind turbines installed at remote locations is around 15–35% of the installation cost. Around 80% of this is spent on unplanned maintenance issues resulting from defects in the wind turbine's various components [1]. Wind turbine fault analysis and diagnosis are usually performed manually from individuals with a high level of technical expertise. This method is inefficient and incapable of meeting the needs of wind farm operations and maintenance. It also leads to production loss due to the prolonged unavailability of the

energy production system and requires a large number of fault diagnosis analysts. The increase in demand for wind energy needs reliable wind farms. To reach the demands, the design of low-cost advanced proactive intelligent fault detection systems is crucial for better performance. The gearbox is one of the most vital and frequently worn parts of the mechanical rotatory system of a wind turbine. Wind turbine gearbox failure diagnosis is critical in mechanical rotating systems, and unpredicted breakdown of this component results in prolonged system downtime. Wind turbine energy production unit maintenance and operating may determine whether the system is cost-efficient or not. Therefore, an intelligent and proactive fault detection system, which reduces the downtime of the wind turbine energy production system, is essential, since it reduces the number of skilled professionals required for maintenance. An expert system enables timely intervention and

early warnings and hence reduces production loss due to prolonged unavailability. As a result, a smart system for condition monitoring of wind turbines improves the reliability by reducing downtime significantly. In the current work, the proposed hybrid Bidirectional Long Short-Term Memory (BLSTM) model with an autoencoder achieves an accuracy of 98.68% in wind turbine gearbox healthiness classification and improves the performance by 71.73% in testing and 65.61% in training when compared with the original BLSTM model. In the next section, various efforts to diagnose and analyze wind turbine gearboxes are discussed.

## II. RELATED WORK

A few efforts have been made to use traditional machine learning algorithms to classify the healthiness of the gearbox. The time-domain sequence Approximate Entropy (ApEn) adaptive strategy, a Wavelet Packet Transform (WPT) filter, and a Cross-validated Particle Swarm Optimized (CPSO) kernel extreme learning machine were used to develop gearbox fault analysis models in [1]. The Discrete Wavelet Transformation (DWT) was used to compute wavelet coefficients for vibration signals obtained from wind turbines. Wavelet coefficients are applied to Phase Space Reconstruction (PSR) and singular value decomposition to extract fault attributes in [2]. The Supervisory Control and Data Acquisition (SCADA) process delivers the most frequently used input data for wind turbine anomaly detection [4]. By reducing the feature dimension of the time-series data, Dynamic Principal Component Analysis (DPCA) was used to identify faults in the gearbox of wind turbines in [5]. To extract gearbox vibration features from oscillated vibration signals of gearbox fault diagnosis, a method combining the Empirical Mode Decomposition (EMD) and Time Synchronous Averaging (TSA) was used in [6]. To detect anomalies in the wind turbine gearbox, Twin Support Vector Machine (TWSVM) and an adaptive threshold were used in [7]. To extract features from three-axial vibration data for fault diagnosis of a wind turbine gearbox, a Deep Enhanced Fusion Network (DEFN) was used in [8]. The deep joint variational autoencoder method was used in conjunction with wind farm supervisory control and data acquisition to diagnose faults in the wind turbine gearbox in [9]. By decomposing vibration signals with a wavelet packet, a fast deep graph convolutional network model was used to analyze the wind turbine gearbox in [10]. The fused vibration signals were classified with a multiclass SVM model in [11]. Electrical signals from generator terminals were used to find faults in the gearbox of the wind turbine in [12]. Using fault features of convolution channels and frequency bands of wavelet coefficients, the residual network can be used to identify a fault in the gearbox of a wind turbine [13]. The methods of convolutional neural networks and isolation forests were applied to classify the health of the gearbox of a wind turbine in [14]. The neighborhood component analysis technique for best feature collection was used to evaluate the healthiness of wind turbine gearboxes in [15]. The remaining useful life of a wind turbine gearbox and its failure can be predicted using machine learning methods such as artificial neural networks, SVM, and logistic regression [16]. CNN is used for feature extraction and representation, and LSTM is used to learn the latent time series relationship between

features in various periods of time [17]. An optimized LSTM neural network with cosine loss was used to analyze wind turbine gearbox faults in [17]. The Cos-LSTM networks were analyzed using the wavelet energy entropy and energy sequence features of the vibration signals in [18]. The traditional LSTM model was improved using multiple swarm intelligence models for classifying failures in the wind turbine gearbox using vibration signal data acquired from the faulty gearbox in [19]. The Simulated Annealing (SA) algorithm was used to optimize the vibration of the powertrain system in [21]. To analyze the functional schemes of the selected gears, the method for trying to generate a mechanism of kinematics equations for signal dependency graphs was used in [22]. Industrial bearing, fault detection, and isolation using time frequency domain has been applied and compared with the theoretical results in [23]. The following are the main contributions of the current article:

- Design of a gearbox diagnostic model using an undercomplete autoencoder and the BLSTM deep learning model.
- Analysis of vibration data collected through sensors.
- Comparison of the proposed model's training and testing times to those of the conventional BLSTM model.

## III. PROPOSED MODEL

The projected model was created on undercomplete autoencoder along with a BLSTM-based hybrid model to diagnose and classify wind turbine gearbox health conditions. The undercomplete autoencoder and BLSTM architectures are described below.

### A. Undercomplete Autoencoder

Autoencoders are unsupervised learning methods used for representation learning. The neural network architecture denotes autoencoders to impose a bottleneck in the neural network, resulting in a compressed feature representation of the original input. Autoencoders contain four layers, namely input layer, hidden layer, bottleneck layer, and output layer. The objective of autoencoders is to minimize the number of nodes located in the hidden layer in order to reduce the information flow through the neural network. An autoencoder model discovers the most important characteristics of the input data. The compressed and essential features are extracted at the bottleneck layer in order to recreate the original data with minimal loss. The number of neurons in the hidden layer is lesser than the number of neurons in the input layer. The bottleneck layer contains fewer nodes than the hidden layer. The reduced features are extracted from the bottleneck layer.

### B. Bidirectional LSTM

The bidirectional LSTM recurrent neural network is made up of LSTM cells, which are memory blocks with a hidden unit. Such states have the responsibility to transfer immediately preceding time step data in the network from the input state to the next cell. These cells are made up of input, forget, and output gates. The forget gate forgets irrelevant information, the input gate updates new information, and the output gate passes the updated information to the next cell.



### C. Design

Figure 1 depicts the construction of an autoencoder with a BLSTM-based model. To design the proposed model for time series gearbox vibration data analysis, input samples of 3600 (1600×2) with a sample size of 500 time steps were fed to the undercomplete autoencoder. The output of the encoder tapped at the bottleneck has reduced features and contains 350 time steps in each sample. The compressed data are loaded into the BLSTM model, which is trained for 500 epochs. To achieve better results, the model employs the rmsprop optimizer and the sigmoid perceptron at the output layer. The metrics are recorded and discussed below.

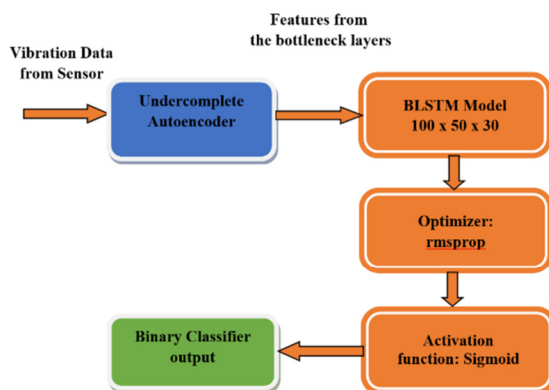


Fig. 1. Proposed autoencoder and BLSTM-based model.

### IV. EXPERIMENTATION AND RESULTS

The developed autoencoder and BLSTM-based model were tested on experimental data collected in a publicly available wind turbine gearbox vibration dataset [20]. Spectra Quest's Gearbox Fault Diagnostics Simulator was used to generate vibration data for two conditions of the gearbox, one for good health and the other for a broken tooth, and both were subjected to bearing capacity ranging from 0% to 90%. Four sensor nodes were positioned in 4 different locations on the body of the gearbox. The dataset contains 10 samples with different loads of wind turbine gearbox vibration data in good and bad conditions. The dataset was created with a 30Hz frequency for

a total of 6.9s. Each sample was collected at a rate of 12,800 time steps/s. A total of 69s (6.9×10) for data generation for a broken tooth or bad condition and data generation for a healthy condition was considered. Table I shows the raw recorded data.

TABLE I. WIND TURBINE GEARBOX VIBRATION DATASET

Wind turbine gearbox condition	Number of samples	Time steps per second	Time steps per sample
Good	10	12800	88320
Bad (broken tooth)	10		

Each sample is made up of 88320 time steps. Because the time series sequence is too long, the data have been divided into subsamples of 500 time steps. As a result, 176 samples from each sample of length 88320 time steps were derived by considering 500 time steps for each individual sample. As a result, we acquired 1760 (176×10) good-condition gearbox samples and 1760 (176×10) bad-condition gearbox samples. The derived dataset contains 3520 samples, as shown in Table II. The purpose of choosing 500 as the subsample's time step is to facilitate experiments, however the size of the sample set can vary. We chose 3200 sample data at random from a total of 3520 for experiments.

TABLE II. GENERATED WIND TURBINE GEARBOX VIBRATION DATASET

Wind turbine gearbox condition	Number of samples	Total number of samples
Good	1760	3520
Bad	1760	

To collect vibration data, 4 sensors were placed on the body of the gearbox in 4 different directions. Table III displays the Accuracy, Precision, Recall, F1-score, Training Time (time taken to train a total 80% of the 3200 (i.e. 2560) training samples, and Testing or Execution Time (time taken to test a total of 640 testing samples (20% of 3200)). It is tested with data frequency (dataset sample size in time steps) equal to 500 time steps and a trained dataset with 500 epochs. The results were captured on a machine with the following architecture: GPU: NVIDIA-SMI 460.32.03, CUDA Version:11.2, Tensor Core GPU: A100-SXM4-40GB with a runtime memory of 89.6GB.

TABLE III. PERFORMANCE OF THE BLSTM MODEL

Sensor	Precision	Recall	F1-Score	Training Time (s)	Testing Time (s)	Classification Accuracy
S-1	0.9875	0.9783	0.9828	8697.74	9.76	98.28%
S-2	0.9838	0.8892	0.9341	8121.10	8.17	93.28%
S-3	0.6451	0.8695	0.8480	8722.22	13.30	78.12%
S-4	0.9062	0.6904	0.7837	8200.48	19.03	75.00%

TABLE IV. PERFORMANCE OF THE UNDERCOMPLETE AUTOENCODER WITH THE BLSTM MODEL TESTED WITH VIBRATION DATA FROM SENSOR-1

No. of time steps tapped at the bottleneck	Precision	Recall	F1-Score	Training Time (s)	Testing Time (s)	Classification Accuracy
50	0.9617	0.9123	0.9364	1175.35	2.16	93.59%
100	0.9593	0.8924	0.9246	1334.07	2.68	92.19%
150	0.9693	0.9080	0.9376	1776.15	2.70	93.44%
200	0.9660	0.9260	0.9456	1997.61	2.61	94.38%
250	0.9670	0.9376	0.9571	2188.04	2.31	95.12%
300	0.9793	0.9776	0.9421	2792.95	2.46	96.53%
350	0.9895	0.9783	0.9857	2991.74	2.76	98.68%



Since the Classification Accuracy using Sensor-1 and Sensor-2 vibration data is optimum, the significance of Sensor-3 and Sensor-4 vibration data analysis has less importance. Hence, the diagnosis of the gearbox fault can be achieved with an exemption of Sensor-3 and Sensor-4 vibration data as these two sensors yield less Accuracy. Hence, the proposed model was tested only on Sensor-1 vibration data.

Table IV shows different test cases with varied numbers of features extracted at the bottleneck layer of the autoencoder with data frequency equal to 500 (number of time steps in each sample), tested with Sensor-1 vibration data. The results clearly show that at the bottleneck output of 350 time steps outperforms the model with respect to Accuracy and Training and Testing Times.

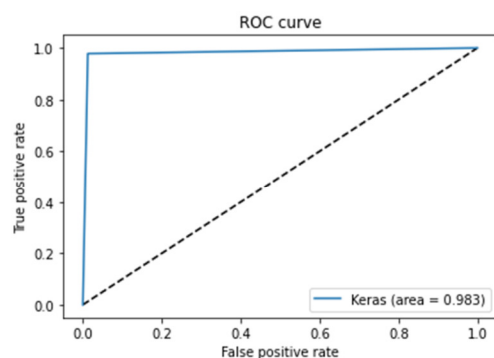


Fig. 2. ROC curve for vibration data from Sensor-1 with 80-20% split ratio BLSTM model.

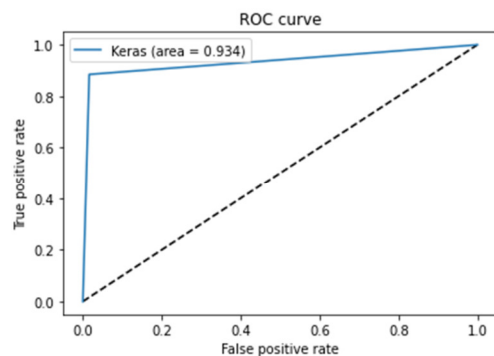


Fig. 3. ROC curve for vibration data from Sensor-2 with 80-20% split ratio BLSTM model.

The proposed method outperforms the BLTSM algorithm model in Classification Accuracy, Training Time, and Testing Time. Based on the experimental observations; the results indicate that the gearbox fault vibration data features can be learned to improve the generalization and the model's accuracy by fusing the undercomplete autoencoder and BLSTM models. However, the result comparison in Tables V-VI shows that the proposed model achieves 98.68% Accuracy with reduced features (350 time steps) in the samples, and the performance is increased by 71.73% in testing time and 65.61% in training time when compared to the BLSTM model. As a result, the fused model of autoencoder with BLSTM outperforms the conventional BLSTM.

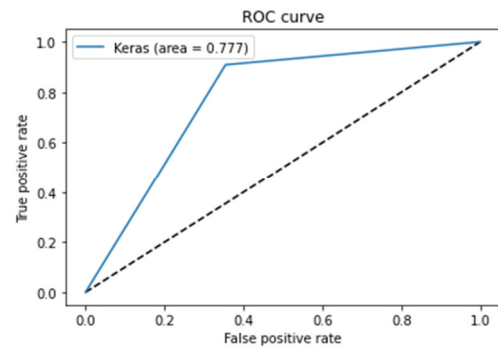


Fig. 4. ROC curve for vibration data from Sensor-3 with 80-20% split ratio BLSTM model.

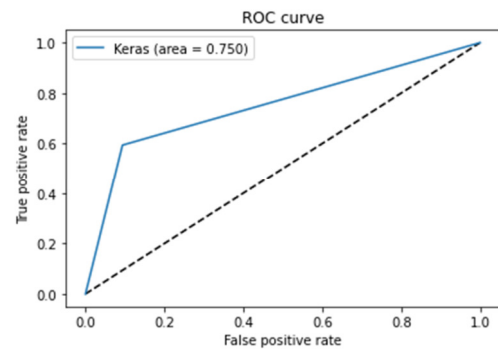


Fig. 5. ROC curve for vibration data from Sensor-4 with 80-20% split ratio BLSTM model.

TABLE V. PERFORMANCE COMPARISON OF THE PROPOSED MODEL WITH BLSTM

Sensor	Conventional BLSTM model with 500 time steps per sample		
	Training Time (s)	Testing Time (s)	Classification Accuracy
S-1	8697.74	9.76	98.28%
	Undercomplete autoencoder with BLSTM model with 350 time steps per sample		
	Training Time (s)	Testing Time (s)	Classification Accuracy
S-1	2991.74	2.76	98.68%

TABLE VI. GEARBOX HEALTH DIAGNOSIS METHOD COMPARISON

Ref.	Method	Accuracy
[2]	WPT-PSO-KELM	94.17%
[18]	Cos-LSTM	98.55% with 550 samples
[19]	LSTM with firefly, cuckoo, PSO and ACO and relu activation function	87.5%
[5]	Support vector machine model used to detect and isolate gear faults. It performs better than the Dynamic Principle Component Analysis (DPCA) using MLP	91.24%
[24]	A technique for feature extraction based on CNN and LSTM for categorization was used to estimate gearbox faults using a better prediction method based on the early fusion of multisource sensing data	97.9%
Proposed	Undercomplete autoencoder with BLSTM	98.68%

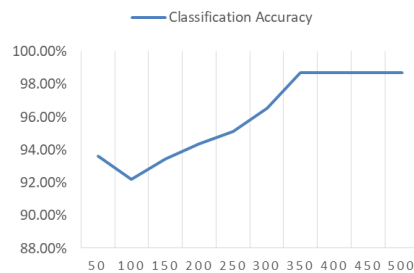


Fig. 6. Performance of the undercomplete autoencoder with BLSTM.

## V. CONCLUSION

This article provides the analysis of wind turbine gearbox vibration data for fault classification. The proposed hybrid model, combines the bidirectional LSTM neural network algorithm with the undercomplete autoencoder and is used for wind turbine gearbox fault detection and diagnosis. The experimental results demonstrate that the suggested model gives an Accuracy of 98.68% in wind turbine gearbox fault classification and improves the performance by 71.73% in the dataset testing samples and 65.61% in the dataset training samples, when compared to the traditional bidirectional LSTM model. With respect to Testing and Training Time, as well as Classification Accuracy, the proposed model outperforms the known models that was compared with.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Yogesh Pandya, Symbiosis University of Applied Sciences, Indore for his help and for providing the dataset for public access.

## REFERENCES

- [1] M. Ruiz *et al.*, "Wind turbine fault detection and classification by means of image texture analysis," *Mechanical Systems and Signal Processing*, vol. 107, pp. 149–167, Jul. 2018, <https://doi.org/10.1016/j.ymssp.2017.12.035>.
- [2] W. Hu, H. Chang, and X. Gu, "A novel fault diagnosis technique for wind turbine gearbox," *Applied Soft Computing*, vol. 82, Sep. 2019, Art. no. 105556, <https://doi.org/10.1016/j.asoc.2019.105556>.
- [3] Y. Pang, L. Jia, X. Zhang, Z. Liu, and D. Li, "Design and implementation of automatic fault diagnosis system for wind turbine," *Computers & Electrical Engineering*, vol. 87, Oct. 2020, Art. no. 106754, <https://doi.org/10.1016/j.compeleceng.2020.106754>.
- [4] G. de N. P. Leite, A. M. Araújo, and P. A. C. Rosas, "Prognostic techniques applied to maintenance of wind turbines: a concise and specific review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1917–1925, Jan. 2018, <https://doi.org/10.1016/j.rser.2017.06.002>.
- [5] M. Kordestani, M. Rezamand, M. Orchard, R. Cariveau, D. S. K. Ting, and M. Saif, "Planetary Gear Faults Detection in Wind Turbine Gearbox Based on a Ten Years Historical Data From Three Wind Farms," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 10318–10323, Jan. 2020, <https://doi.org/10.1016/j.ifacol.2020.12.2767>.
- [6] S. Mukherjee, V. Kumar, S. Sarangi, and T. K. Bera, "An Improved Signal Pre-Processing Method for Gearbox Fault Features Extraction," *Procedia Computer Science*, vol. 167, pp. 1604–1614, Jan. 2020, <https://doi.org/10.1016/j.procs.2020.03.371>.
- [7] H. S. Dhiman, D. Deb, S. M. Muyeen, and I. Kamwa, "Wind Turbine Gearbox Anomaly Detection Based on Adaptive Threshold and Twin Support Vector Machines," *IEEE Transactions on Energy Conversion*, vol. 36, no. 4, pp. 3462–3469, Sep. 2021, <https://doi.org/10.1109/TEC.2021.3075897>.
- [8] Z. Pu, C. Li, S. Zhang, and Y. Bai, "Fault Diagnosis for Wind Turbine Gearboxes by Using Deep Enhanced Fusion Network," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021, <https://doi.org/10.1109/TIM.2020.3024048>.
- [9] L. Yang and Z. Zhang, "Wind Turbine Gearbox Failure Detection Based on SCADA Data: A Deep Learning-Based Approach," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021, <https://doi.org/10.1109/TIM.2020.3045800>.
- [10] X. Yu, B. Tang, and K. Zhang, "Fault Diagnosis of Wind Turbine Gearbox Using a Novel Method of Fast Deep Graph Convolutional Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021, <https://doi.org/10.1109/TIM.2020.3048799>.
- [11] Y. Peng, W. Qiao, F. Cheng, and L. Qu, "Wind Turbine Drivetrain Gearbox Fault Diagnosis Using Information Fusion on Vibration and Current Signals," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021, <https://doi.org/10.1109/TIM.2021.3083891>.
- [12] D. Lu, W. Qiao, and X. Gong, "Current-Based Gear Fault Detection for Wind Turbine Gearboxes," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1453–1462, Jul. 2017, <https://doi.org/10.1109/TSTE.2017.2690835>.
- [13] K. Zhang, B. Tang, L. Deng, and X. Liu, "A hybrid attention improved ResNet based fault diagnosis method of wind turbines gearbox," *Measurement*, vol. 179, Jul. 2021, Art. no. 109491, <https://doi.org/10.1016/j.measurement.2021.109491>.
- [14] A. Meyer, "Vibration Fault Diagnosis in Wind Turbines Based on Automated Feature Learning," *Energies*, vol. 15, no. 4, Jan. 2022, Art. no. 1514, <https://doi.org/10.3390/en15041514>.
- [15] H. S. Dhiman, D. Deb, J. Carroll, V. Muresan, and M.-L. Unguresan, "Wind Turbine Gearbox Condition Monitoring Based on Class of Support Vector Regression Models and Residual Analysis," *Sensors*, vol. 20, no. 23, Jan. 2020, Art. no. 6742, <https://doi.org/10.3390/s20236742>.
- [16] J. Carroll, S. Koukoura, A. McDonald, A. Charalambous, S. Weiss, and S. McArthur, "Wind turbine gearbox failure and remaining useful life prediction using machine learning techniques," *Wind Energy*, vol. 22, no. 3, pp. 360–375, 2019, <https://doi.org/10.1002/we.2290>.
- [17] Y. Xie, J. Zhao, B. Qiang, L. Mi, C. Tang, and L. Li, "Attention Mechanism-Based CNN-LSTM Model for Wind Turbine Fault Prediction Using SSN Ontology Annotation," *Wireless Communications and Mobile Computing*, vol. 2021, Mar. 2021, Art. no. e6627588, <https://doi.org/10.1155/2021/6627588>.
- [18] A. Yin, Y. Yan, Z. Zhang, C. Li, and R.-V. Sánchez, "Fault Diagnosis of Wind Turbine Gearbox Based on the Optimized LSTM Neural Network with Cosine Loss," *Sensors*, vol. 20, no. 8, Jan. 2020, Art. no. 2339, <https://doi.org/10.3390/s20082339>.
- [19] G. Durbhaka, B. Selvaraj, M. Mittal, T. Saba, A. Rehman, and L. Goyal, "Swarm-LSTM: Condition Monitoring of Gearbox Fault Diagnosis Based on Hybrid LSTM Deep Neural Network Optimized by Swarm Intelligence Algorithms," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 2041–2059, 2020, <https://doi.org/10.32604/cmc.2020.013131>.
- [20] "Gear Box Fault Diagnosis Data Set - dataset by gearbox," *data.world*, <https://data.world/gearbox/gear-box-fault-diagnosis-data-set>.
- [21] M. O. Genc and N. Kaya, "Vibration Damping Optimization using Simulated Annealing Algorithm for Vehicle Powertrain System," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5164–5167, Feb. 2020, <https://doi.org/10.48084/etasr.3242>.
- [22] A. Deptula, "Application of the Dependency Graph Method in the Analysis of Automatic Transmission Gearboxes," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 7033–7040, Apr. 2021, <https://doi.org/10.48084/etasr.4098>.
- [23] Y. Bella, A. Oulmane, and M. Mostefai, "Industrial Bearing Fault Detection Using Time-Frequency Analysis," *Engineering, Technology & Applied Science Research*, vol. 8, no. 4, pp. 3294–3299, Aug. 2018, <https://doi.org/10.48084/etasr.2135>.
- [24] H. Habbouche, T. Benkedjouh, Y. Amirat, and M. Benbouzid, "Gearbox Failure Diagnosis Using a Multisensor Data-Fusion Machine-Learning-Based Approach," *Entropy*, vol. 23, no. 6, Jun. 2021, Art. no. 697, <https://doi.org/10.3390/e23060697>.

# Best Fit versus Default Distribution and the Impact on the Reliability over the Design Lifetime of Hydraulic Structures

**Farida Krimil**

Civil and Hydraulic Engineering Department, Sciences and Technology Faculty, University of Mohamed Khider Biskra, Algeria  
f.krimil@univ-batna2.dz  
(corresponding author)

**Nora Bouchahm**

Technical and Scientific Center on Arid Regions, CRSTRA, Algeria  
bouchahm.nora@crstra.dz

**Fatima Zohra Tebbi**

Natural Hazards and Territory Planning Laboratory (LRNAT), Earth Sciences and Universe Institute, University of Mustapha Benboulaïd Batna 2, Algeria  
f.tebbi@univ-batna2.dz

Received: 16 December 2022 | Revised: 29 December 2022 and 6 January 2023 | Accepted: 7 January 2023

## ABSTRACT

In the present study, Flood Frequency Analysis (FFA) is performed on the daily inflows of a reservoir dam taken as a case study. The Peaks-Over-Threshold (POT) approach was adopted. A comparison between the default generalized Pareto distribution and the best distribution fitted to the data has been carried out. After the risk analysis, the reliability of the structure decreases to 25.60% for the chosen threshold values if the best distribution is adopted instead of the default fit.

**Keywords**-flood frequency analysis; peaks-over-threshold; generalized Pareto distribution; best fit; risk analysis

## 1. INTRODUCTION

Flood Frequency Analysis (FFA) [1, 2] plays a major role in the design of hydraulic structures because it affects both safety and cost of the structure. The principal objective of FFA is to construct a relationship between the flood magnitude and the return period by the estimation of the probability of exceedance [3]. Hydrologists generally apply two types of approaches to perform FFA, Annual Maximum (AM) and Peaks-Over-Threshold (POT), also called partial duration series approach. The AM series approach is the most used method in FFA due to its simplicity in the sampling process [4, 5]. It uses only one maximum discharge value from each year. However, defining the samples in such a way eliminates a large portion of the data and results in loss of useful information, e.g. the second highest flow data in a year (which could be higher than many data points in the AM series) is not selected [4, 5]. The POT approach consists in retaining all peak values that are above a certain level usually called the threshold. The POT approach is more useful to the analysis of extreme values. The

main advantage of the POT approach is that it allows the selection of an enriched series of events to be considered as floods [7] and controls the number of flood occurrences to be used in the analysis, unlike the AM approach which includes only one event per year. Authors in [7] prepared a guide for the use of the POT approach. Authors in [6] indicate that the POT method can provide adequate and comparable estimates of N-year discharges for more stations with short temporal coverage. Authors in [8] reviewed recent advances in the POT approach from a statistical perspective. Authors in [4] reported that the POT method has better results than the AM method after comparing the results of both methods for data from the Litija 1 gauging station on the Sava River in Slovenia. Authors in [5] reviewed and summarized the current status of the POT model and identified the difficulties in applying it in FFA. Ensuring the independence of the data series and choosing an appropriate threshold value are two of the main difficulties associated with the POT approach [4]. The threshold should be high enough to maintain the assumption of flood independence, but should not be so high as to increase the variance by reducing the number

of events needed for flood analysis [9]. Therefore, the selection of an appropriate threshold is the subject of several studies and many methods, either graphical or analytical, have been reported. Graphical methods are widely used [10]. There are several probability distributions for modeling flood extreme events, the selection of the best fit distribution and the associated parameter estimation is an important step in FFA [11]. Generalized Pareto Distribution (GPD) is broadly used to model extreme floods over a threshold and the distribution of POT can be approximated by the GPD [12]. The annual number of events above the threshold can be approximated by either Poisson or Negative binomial distributions [13, 14].

In the current study, the FFA method, applied to the daily inflows of the Koudiat Medouar reservoir and the POT approach are adopted. Several additional distributions are tested to find the best one that fits the data. A risk analysis is then conducted.

## II. STUDY AREA

Located 7km north-east of Timgad and 35km from Batna, the Koudiat Medouar dam was built in 1994 on the Oued Rebôa. It has a height of 48m. The reservoir is a part of the great strategic hydraulic complex Beni Haroun Transfer, with a capacity of 70 million m<sup>3</sup>. It supplies drinking and irrigation water to several towns. Figure 1 shows the Koudiat Medouar reservoir location.

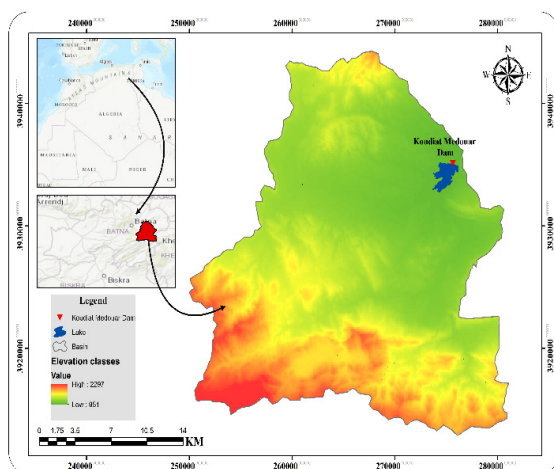


Fig. 1. Location of the Koudiat Medouar reservoir.

## III. METHODOLOGY

FFA is a statistical method that consists of studying past events and characteristics of a given process (hydrological or other) to define the probability of future occurrence by a probability distribution using the AM or POT series. The POT approach consists in the constitution of a series of data by extracting the values superior to a well determined threshold. The selection of an appropriate threshold value is an important phase, Mean Residual Life Plot (MRLP) and the stability graph of the parameters are recommended in [15]. The threshold value is chosen upon a certain stability of the scale and shape graphs. Threshold is taken at the tail of the MRLP when the function of the mean value begins to be almost linear. The POT

flood series are adjusted by the distribution of GPD [16]. The adjustment of excesses above a threshold according to the GPD model hypothesis must be ensured by two discrete distributions, Poisson or negative binomial. The main step in the application of the POT method is the choice of the threshold. Indeed, selecting a very high threshold has the effect of decreasing the size of the series and thus increasing the sampling uncertainties. The Number of Maximums per Year (NMpY  $\approx 1$ ) subsequently gives a very short AM series [17]. On the other hand, a threshold that is too low (NMpY  $\geq 6$ ) gives series artificially enriched with information but one will be confronted on the problem of obtaining dependent series. Therefore, a trade-off must be found between acquiring information and obtaining independent series (Figure 2). However, there is no universal criterion for identifying such a threshold value [18]. In this paper, we focus on the following criteria for threshold selection:

- MRLPs.
- The stability of scale and modified shape graphs.
- Fixing the NMpY.

Once the preselection has been performed according to the above criteria, the basic hypothesis tests of FFA, i.e. homogeneity, independence, and stationarity, must be verified. To verify these hypotheses, three non-parametric statistical tests are used: [19] to verify the independence, [20, 21] to verify whether the data come from the same distribution or not, and [22, 23] to verify the stationarity of the data.

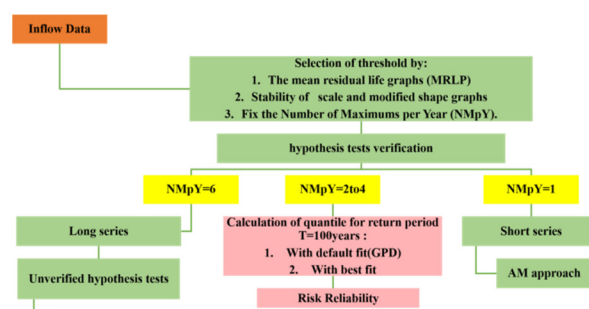


Fig. 2. Methodology chart.

### A. Model Fitting

Several types of distribution are used to estimate flow extremes. The reasons for the operational use of a particular distribution type in many countries are often subjective or historical [24]. Table I illustrates the probability distributions mostly used to fit inflows. The choice of the appropriate estimator is one of the most important issues in FFA [25, 28]. The most commonly used are: the Method Of Moments (MOM), the Method of L-moments (ML), and the Maximum Likelihood Estimation (MLE). The ML is more used in statistical hydrology for estimating various hydro-meteorological variables [29, 30]. The objective of the FFA is to find reliable estimates of quantiles that will help us in the design of structures. After the selection of the thresholds, and defining the best fit, the quantiles are evaluated for a return period of 100 years.

TABLE I. USUAL PROBABILITY DISTRIBUTIONS FUNCTIONS

Distribution	PDF
Normal	$f(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}$
Log normal	$f(x; x_0, \mu, \sigma) = \frac{1}{(x - x_0) \sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left[ \ln \left( \frac{x - x_0}{\sigma} \right) - \mu \right]^2}$
Gumbel	$f(x) = \frac{\exp \left[ -e^{-\frac{x-x_0}{\alpha}} \right]}{\alpha} \exp \left( -\frac{x-x_0}{\alpha} \right)$
Pearson 3	$f(x; \lambda, \beta, \gamma) = \frac{1}{\lambda \Gamma(\beta)} \left( \frac{x-\gamma}{\lambda} \right)^{\beta-1} e^{-\left( \frac{x-\gamma}{\lambda} \right)}$
Generalized extreme value	$f(x; \alpha, \kappa, \mu) = \frac{1}{\alpha} \left[ 1 - \frac{\kappa}{\alpha} (x - \mu) \right]^{\frac{1}{\kappa}-1} \exp \left\{ - \left[ 1 - \frac{\kappa}{\alpha} (x - \mu) \right]^{\frac{1}{\kappa}} \right\}$
Generalized Pareto distribution	$f(x; \lambda, \beta, \gamma) = \begin{cases} \frac{1}{\lambda} \left[ 1 + \gamma \left( \frac{x-\beta}{\lambda} \right) \right]^{-1-\frac{1}{\gamma}}, & \gamma \neq 0 \\ \frac{1}{\lambda} \exp \left[ - \left( \frac{x-\beta}{\lambda} \right) \right], & \gamma = 0 \end{cases}$
Kappa	$f(x; \alpha, \kappa, \mu, h) = \alpha^{-1} \left[ 1 - \frac{\kappa(x-\mu)}{\alpha} \right]^{\frac{1}{\kappa}-1} [F(x)]^{1-h}, \text{ where}$ $F(x; \alpha, \kappa, \mu, h) = \left\{ 1 - h \left[ 1 - \frac{\kappa(x-\mu)}{\alpha} \right]^{\frac{1}{\kappa}} \right\}^{\frac{1}{h}}$

$\alpha$ : scale parameter,  $\mu$ : location parameter,  $\kappa$ : shape parameter,  $\gamma$ : second scale parameter,  $h$ : second shape parameter,  $\Gamma(\cdot)$ : gamma function,  $F(\cdot)$ : cumulative distribution function

### B. Risk Analysis

Water-control design requires the consideration of risk [31]. Risk and reliability analysis have great importance when it comes to natural phenomena such as floods. Risk analysis of future events requires a probabilistic approach [32]. This natural hydrological risk can be calculated using the following equations:

$$p(X \geq x_T) \text{ at least once in } N \text{ years} = 1 - \left(1 - \frac{1}{T}\right)^N \quad (1)$$

$$\bar{R} = 1 - [1 - p(X \geq x_T)]^N \quad (2)$$

where:

$$p(X \geq x_T) = \frac{1}{T} \quad (3)$$

$$\bar{R} = 1 - \left[ 1 - \frac{1}{T} \right]^N \quad (4)$$

where  $N$  is the expected life of the structure (in our case  $N=50$  years),  $T$  is the return period (usually the life expectancy of a hydraulic structure is 100 years), and  $\bar{R}$  is the probability that an event  $X \geq x_T$  will occur at least once in  $N$  years. The reliability  $Re$  is defined by:

$$Re = 100 - \bar{R} \quad (5)$$

## IV. RESULTS AND DISCUSSION

### A. Data Description

The used database in this study was collected from the National Agency for Dams and Transfer (ANBT). Daily inflows of Koudiat Medouar reservoir spanning the period from 2004 to 2019 were calculated according to the following water balance equation:

$$\text{Inflow} = (V_{ini} - V_{fin}) + (DEV + DWS + DRV + DSV + DFV + DLV) \quad (6)$$

where Inflow is the daily inflow,  $V_{ini}$  and  $V_{fin}$  represent the initial and the volume in the reservoir,  $DEV$  is the daily evaporated volume,  $DWS$  the daily volume allocated to water supply,  $DRV$  the daily volume allocated to irrigation,  $DSV$  the daily spilled volume,  $DFV$  the daily flashed volume, and  $DLV$  is the daily leakage volume.

Daily inflows are not normally distributed. The box plot in Figure 3 shows that two values  $117.608 \text{ m}^3/\text{s}$  and  $110.915 \text{ m}^3/\text{s}$  appear as outliers, but in fact they are real occurred events. Figure 4 shows the interannual variations of daily inflows at the Koudiat Medouar reservoir. Even if the stability of scale and shape parameters are confirmed (Figure 5), it is important to verify the hypothesis tests (homogeneity, independence, and stationarity). Between the thresholds of  $5 \text{ m}^3/\text{s}$  and  $11 \text{ m}^3/\text{s}$ , even with that of  $14 \text{ m}^3/\text{s}$ , the tests were not verified.



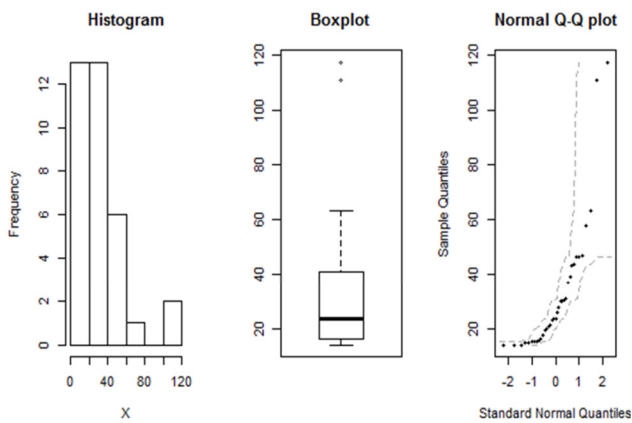


Fig. 3. Daily inflows description of Koudiat Medouar reservoir.

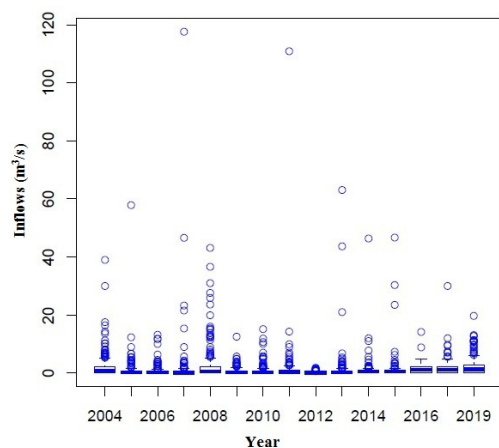


Fig. 4. Time series of the daily inflows at Koudiat Medouar reservoir (2004-2019).

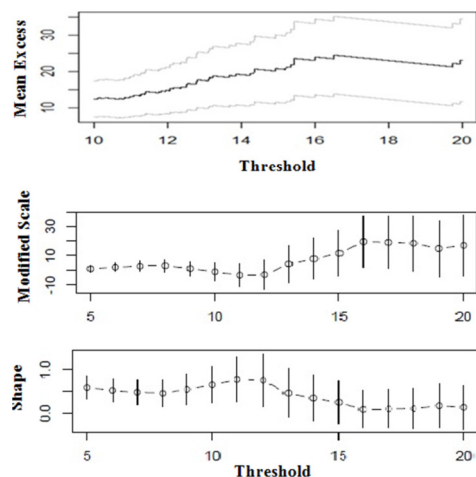


Fig. 5. Mean residual life plot for the daily inflows of the Koudiat Medouar reservoir.

This led us to keep only the values of 12, 13, and 15 m<sup>3</sup>/s respectively (Table II). Using the R code RStudio [33] and for the previously chosen thresholds, we tested whether the different distributions fit. The results show that GPD

distribution is not the best fit for 12 m<sup>3</sup>/s and 13 m<sup>3</sup>/s threshold (Figure 6). The threshold of 15 m<sup>3</sup>/s just ensures an exceedance number of 2 per year, for this purpose, this threshold is presented as a limit of choice beyond which we find the statistical situation of the AM. For thresholds below 12 m<sup>3</sup>/s, the opposite situation arises: these are samples artificially enriched in size with a greater chance of losing the satisfaction of the hypothesis tests, in particular the homogeneity and the randomness required for any reliable frequency analysis without statistical violation (Table III). The frequency distribution results are shown in Figure 6.

TABLE II. STATISTICAL PARAMETERS CORRESPONDING TO POT TIME SERIES

Statistic	$u=12\text{m}^3/\text{s}$	$u=13\text{m}^3/\text{s}$	$u=15\text{m}^3/\text{s}$
Sample size	50	40	31
Min	12.08	13.12	15.22
1 <sup>st</sup> quantile	13.34	15.31	19.85
Median	17.06	22.41	27.67
Mean	26.81	30.40	35.20
3 <sup>rd</sup> quantile	30.27	37.26	43.42
Max	117.60	117.60	117.60
Wald-Wolfowitz	p-value	0.086	0.337
Wilcoxon-Mann-Whitney	p-value	0.851	0.814
Mann-Kendal	p-value	0.536	0.954

TABLE III. ESTIMATED QUANTILES FOR RETURN PERIOD T=100 YEARS

Threshold	Type of fit	Distribution	Quantiles (m <sup>3</sup> /s)	RMSE (m <sup>3</sup> /s)
$u=12\text{m}^3/\text{s}$	Best fit	Pearson 3	159.285	0.024
	Default fit	GPD	211.924	0.049
$u=13\text{m}^3/\text{s}$	Best fit	Kappa	169.318	0.026
	Default fit	GPD	190.416	0.037
$u=15\text{m}^3/\text{s}$	Best fit = default fit	GPD	174.899	0.028

The frequency distribution results are shown in Figure 6. The plots indicate that the Pearson 3 (RMSE=0.024) and Kappa (RMSE=0.026) distributions are the best fit for  $u=12$  and 13 m<sup>3</sup>/s, respectively. Using (4) and (5), the risk reliability is calculated for a return period of 100 years and is presented for best fit and GPD distribution results. The quantile corresponding to 100 years return period for Pearson 3 distribution is equivalent to a 48 years return period of the GPD for the 12 m<sup>3</sup>/s threshold. Respectively, quantile corresponding to 100 years return period for Kappa distribution is equivalent to a return period of 70 years of the GPD fit for the 13 m<sup>3</sup>/s threshold (Table IV).

By selecting the best fit, the reliability of the structure decreases by 25.60% and 11.80% for thresholds of 12 m<sup>3</sup>/s and 13 m<sup>3</sup>/s, respectively, compared to the default fit. Quantile Pearson 3= 159.285 m<sup>3</sup>/s for  $T=100$  years, corresponding to  $T=48$  years quantile in GPD (reliability=-25.60%) and quantile Kappa= 169.318 m<sup>3</sup>/s for  $T=100$  years in GPD corresponding to  $T=70$  years (reliability=-11.80%). We note that when the reliability decreases, the risk increases, so we have under dimensioned hydraulic structures.



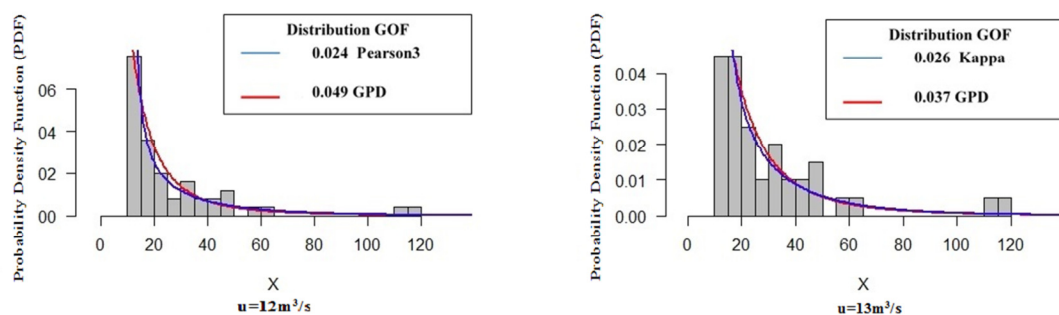


Fig. 6. Probability density function of best and default fit for the selected thresholds.

TABLE IV. RISK RELIABILITY

			Pearson 3	GPD
$u$ (m <sup>3</sup> /s)	12	Return period ( $T$ )	100	48
		Risk (for $N=50$ years)	39.50%	65.10%
		Reliability	60.50%	34.90%
	13		Kappa	GPD
		Return period ( $T$ )	100	70
		Risk (for $N=50$ years)	39.50%	51.30%
	Reliability	60.50%	48.70%	

## V. CONCLUSION

This article presents high flood analysis in relation to the data from daily inflows of the Koudiat Medouar reservoir for the period from 2004 to 2019. FFA was carried out with the POT series. One of the main disadvantages concerning the POT approach is the selection of the threshold value. In FFA, choosing a very high threshold Number of Maximums per Year ( $NMPY \approx 1$ ), has the effect of decreasing the size of the series and therefore increasing the sampling uncertainty which subsequently gives a very short series. On the other hand, a too small threshold gives series artificially enriched information ( $NMPY \geq 6$ ). For this reason, we have retained only the series of 2 to 4 max per year. After the verification of the basic hypothesis tests, the thresholds retained were 12, 13, and 15 m<sup>3</sup>/s. In this study, the GPD distribution is not systematically the best fit of the daily inflows. The study showed that the Pearson 3 and Kappa distributions are better suited than GPD for thresholds equal to 12 and 13 m<sup>3</sup>/s, respectively, for the estimation of flood discharge.

The carried out risk analysis shows that the choice of GPD distribution decreases the reliability of the structure compared to the case of choosing the best fit distribution. The results presented in the current paper are helpful to the planning and optimization of the dimensions of hydraulic structures.

## REFERENCES

- [1] N. Harkat, S. Chaouche, and M. Bencherif, "Flood Hazard Spatialization Applied to The City of Batna: A Methodological Approach," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5748–5758, Jun. 2020, <https://doi.org/10.48084/etasr.3429>.
- [2] K. Loumi and A. Redjem, "Integration of GIS and Hierarchical Multi-Criteria Analysis for Mapping Flood Vulnerability: The Case Study of M'sila, Algeria," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7381–7385, Aug. 2021, <https://doi.org/10.48084/etasr.4266>.
- [3] S. Mkhandi, A. Opere, and P. Willems, "Comparison between annual maximum and peaks over threshold models for flood frequency prediction," in *International Conference of UNESCO Flanders FIT FRIEND/Nile Project - "Towards a better cooperation"*, Sharm El-Sheikh, Egypt, Nov. 2005.
- [4] N. Bezak, M. Brilly, and M. Sraj, "Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis," *Hydrological Sciences Journal*, vol. 59, no. 5, pp. 959–977, May 2014, <https://doi.org/10.1080/02626667.2013.831174>.
- [5] X. Pan, A. Rahman, K. Haddad, and T. B. M. J. Ouarda, "Peaks-over-threshold model in flood frequency analysis: a scoping review," *Stochastic Environmental Research and Risk Assessment*, vol. 36, no. 9, pp. 2419–2435, Sep. 2022, <https://doi.org/10.1007/s00477-022-02174-6>.
- [6] V. Bacova-Mitkova and M. Onderka, "Analysis of extreme hydrological Events on THE danube using the Peak Over Threshold method," *Journal of Hydrology and Hydromechanics*, vol. 58, no. 2, pp. 88–101, Jun. 2010, <https://doi.org/10.2478/v10098-010-0009-x>.
- [7] M. Lang, T. B. M. J. Ouarda, and B. Bobee, "Towards operational guidelines for over-threshold modeling," *Journal of Hydrology*, vol. 225, no. 3, pp. 103–117, Dec. 1999, [https://doi.org/10.1016/S0022-1694\(99\)00167-5](https://doi.org/10.1016/S0022-1694(99)00167-5).
- [8] C. Scarrott and A. MacDonald, "A Review of Extreme Value Threshold Estimation and Uncertainty Quantification," *REVSTAT-Statistical Journal*, vol. 10, no. 1, pp. 33–60, Apr. 2012, <https://doi.org/10.57805/revstat.v10i1.110>.
- [9] A. Gharib, E. G. R. Davies, G. G. Goss, and M. Faramarzi, "Assessment of the Combined Effects of Threshold Selection and Parameter Estimation of Generalized Pareto Distribution with Applications to Flood Frequency Analysis," *Water*, vol. 9, no. 9, Sep. 2017, Art. no. 692, <https://doi.org/10.3390/w9090692>.
- [10] B. Bader, J. Yan, and X. Zhang, "Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate," *The Annals of Applied Statistics*, vol. 12, no. 1, pp. 310–329, Mar. 2018, <https://doi.org/10.1214/17-AOAS1092>.
- [11] 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>.
- [12] J. Pickands, "Statistical Inference Using Extreme Order Statistics," *The Annals of Statistics*, vol. 3, no. 1, pp. 119–131, 1975.
- [13] C. Cunnane, "A note on the Poisson assumption in partial duration series models," *Water Resources Research*, vol. 15, no. 2, pp. 489–494, 1979, <https://doi.org/10.1029/WR015i002p00489>.
- [14] B. Onoz and M. Bayazit, "Effect of the occurrence process of the peaks over threshold on the flood estimates," *Journal of Hydrology*, vol. 244, no. 1, pp. 86–96, Apr. 2001, [https://doi.org/10.1016/S0022-1694\(01\)00330-4](https://doi.org/10.1016/S0022-1694(01)00330-4).
- [15] S. Coles, *An introduction to statistical modeling of extreme values*. New York, NY, USA: Springer, 2001.

- [16] S. D. Grimshaw, "Computing Maximum Likelihood Estimates for the Generalized Pareto Distribution," *Technometrics*, vol. 35, no. 2, pp. 185–191, May 1993, <https://doi.org/10.1080/00401706.1993.10485040>.
- [17] P. Dubreuil, *Initiation a l'analyse hydrologique: dix exercices suivis des corriges*. Paris, France: Masson & Cie, 1974.
- [18] S. Lachance-Cloutier, "Series de durees partielles: application en contexte non homogene et non stationnaire," M.S. thesis, Universite du Quebec, Quebec, Canada, 2011.
- [19] A. Wald and J. Wolfowitz, "An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 378–388, 1943.
- [20] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, Dec. 1945, <https://doi.org/10.2307/3001968>.
- [21] H. B. Mann and D. R. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947.
- [22] H. B. Mann, "Nonparametric Tests Against Trend," *Econometrica*, vol. 13, no. 3, pp. 245–259, 1945, <https://doi.org/10.2307/1907187>.
- [23] M. G. Kendall, *Rank Correlation Methods*. Oxford, UK: Oxford University Press, 1975.
- [24] C. Cunneane, *Statistical distributions for flood frequency analysis*. Geneva, Switzerland: World Meteorological Organization, 1989.
- [25] R. W. Vogel and D. E. McMartin, "Probability Plot Goodness-of-Fit and Skewness Estimation Procedures for the Pearson Type 3 Distribution," *Water Resources Research*, vol. 27, no. 12, pp. 3149–3158, 1991, <https://doi.org/10.1029/91WR02116>.
- [26] B. Merz and A. H. Thielen, "Flood risk curves and uncertainty bounds," *Natural Hazards*, vol. 51, no. 3, pp. 437–458, Dec. 2009, <https://doi.org/10.1007/s11069-009-9452-6>.
- [27] F. Laio, G. Di Baldassarre, and A. Montanari, "Model selection techniques for the frequency analysis of hydrological extremes," *Water Resources Research*, vol. 45, 2009, Art. no. 07416, <https://doi.org/10.1029/2007WR006666>.
- [28] K. Haddad and A. Rahman, "Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia," *Stochastic Environmental Research and Risk Assessment*, vol. 25, no. 3, pp. 415–428, Mar. 2011, <https://doi.org/10.1007/s00477-010-0412-1>.
- [29] T. S. Gubareva and B. I. Gartsman, "Estimating distribution parameters of extreme hydrometeorological characteristics by L-moments method," *Water Resources*, vol. 37, no. 4, pp. 437–445, Jul. 2010, <https://doi.org/10.1134/S0097807810040020>.
- [30] J. R. M. Hosking, "L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 52, no. 1, pp. 105–124, 1990, <https://doi.org/10.1111/j.2517-6161.1990.tb01775.x>.
- [31] V. T. Chow, *Applied Hydrology*. New York, NY, USA: McGraw-Hill, 2010.
- [32] B. Bobee and Ashkar, *Gamma Family and Derived Distributions Applied in Hydrology*. Littleton, CO, USA: Water Resources Pubns, 1991.
- [33] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.

# Evaluation of Confining Pressure Models for Circular Concrete-Filled Steel Tube Short Columns under Concentric Loading

**Hoang An Le**

Ho Chi Minh City University of Transport, Vietnam

hoangan.le@ut.edu.vn

(corresponding author)

Received: 5 December 2022 | Revised: 2 January 2023 | Accepted: 5 January 2023

## ABSTRACT

It is well known that the confinement effect provided by the steel tube significantly increases the strength and ductility of circular Concrete Filled Steel Tube (CFST) columns under concentric loading. The lateral pressure is an important factor to calculate the strength enhancement of CFST columns. To reliably predict the ultimate strength of circular CFST columns, many models have been developed for predicting the lateral pressure due to the confinement effect. This paper aims to evaluate some of these confining pressure models. The values of the compressive strength of confined concrete and lateral confining pressure of circular CFST short columns were calculated using these existing models and were compared with those obtained from previous test results. In addition, a comparison between the ultimate loads predicted by these models and by Eurocode 4 (EC4) was carried out. Based on the comparison results, some suitable models for circular CFST short columns with the use of normal-strength, high-strength, and ultra-high-strength concrete were indicated.

**Keywords**-concrete filled steel tube columns; confinement effect; strength enhancement; steel tube; confining pressure

## I. INTRODUCTION

The composite action between the steel tubes and the concrete core leads to the strength enhancement of Concrete Filled Steel Tube (CFST) columns [1, 19]. The steel tube provides confining pressure to the concrete core and shares the axial load, which puts the concrete core under a triaxial state of stress while the steel tube is stressed biaxially [1-3]. In addition, the steel tube is more stiffened by the concrete core. The inward buckling of the steel tube is prevented by the concrete core, thus leading to an increase in the stability as well as the strength of the column [4]. Circular CFST columns exhibit better gain of load capacity due to the effectiveness of the confinement effect as compared with square or rectangular CFST columns [1]. Therefore, the circular CFST columns are recommended to be used in practice to obtain better strength and ductility [19-21]. To calculate the ultimate strength of circular CFST columns, the lateral pressure induced by the steel tube on the concrete core should be accurately estimated. Because the measurement of lateral pressure in the real test is very complicated, many researchers have proposed models to predict the lateral pressure. Authors in [3] tried to find the stress-strain relationship and the ultimate strength of concrete inside CFST columns with considering the Poisson's ratios. The lateral confining pressure of the steel tube on the concrete was calculated at the point that the steel is in the plastic range and the concrete core is completely crushed. Authors in [15]

proposed a confining pressure model for concrete in CFST columns by taking into account the effect of geometry and material properties on the strength enhancement and post-peak behavior. Authors in [3] modified the models in [4, 5] for CFST columns. Authors in [6] developed a uniaxial stress-strain relationship for concrete confined by various shaped steel tubes. In [1], the efficiency of the steel tube in confining the concrete core was examined by investigating the effects of concrete strength and steel tube thickness. Authors in [7] also developed a confining pressure model for concrete in circular, square, and octagonal CFST columns by employing the nonlinear finite element method in ABAQUS. Authors in [8] suggested a model to calculate the ultimate strength of the concrete core in CFST columns depending on the hoop stress of steel tube at yield condition. Authors in [9] proposed models for predict the ultimate load of circular CFST under extreme loading conditions with the use a new empirical equation for estimate the hoop stress of the steel tube. Authors in [10] proposed an accurate model for confined concrete and a new design formula for determining the ultimate axial loads of CFST columns using Normal Strength Concrete (NSC) and High Strength Concrete (HSC). Authors in [11] presented a full elasto-plastic model and a simplified model for CFST stub columns with concrete strength ranging from 30 to 120MPa and diameter-to-wall thickness ( $D/t$ ) greater than 20. All collected models have been mainly developed for NSC and

HSC filled in steel tubes. Nevertheless, the application of these models for Ultra-High Strength Concrete (UHSC) is still questionable. Therefore, there is a need to investigate the suitability of these models for circular CFST short columns with various concrete strengths, i.e. NSC, HSC, and UHSC.

Based on the aforementioned issues, this paper aims to evaluate the existing models [1, 2, 6, 8-11] through the comparison between the predictions and the test results. The ultimate loads and lateral confining pressures of circular CFST columns under the concentric compression on the entire section from some previous experimental tests were calculated using the equations of the existing models. Furthermore, the

prediction of ultimate loads from Eurocode 4 (EC4) was also compared with the predictions obtained from the existing models. The findings of this paper indicate suitable models for the estimation of the ultimate load and confining pressure of circular CFST short columns.

## II. CONFINING PRESSURE MODELS AND TEST DATABASE

The considered confining pressure models for circular CFST short columns [1, 2, 6, 8-11] were collected and are shown in Table I. These models were used for the prediction of confining pressure and ultimate load of the previous tests.

TABLE I. EXPRESSIONS OF CONFINING PRESSURE MODELS FOR CIRCULAR CFST SHORT COLUMNS

Model	Expression	Explanation
[6]	$v'_c = \frac{0.881}{10^6} \cdot \left(\frac{D}{t}\right)^3 - \frac{2.58}{10^4} \cdot \left(\frac{D}{t}\right)^2 + \frac{1.953}{10^2} \cdot \left(\frac{D}{t}\right) + 0.4011$ $v_c = 0.2312 + 0.3528 \cdot v'_c - 0.1524 \cdot \left(\frac{f'_c}{f_y}\right) + 4.843 \cdot v'_c \cdot \left(\frac{f'_c}{f_y}\right) - 9.169 \cdot \left(\frac{f'_c}{f_y}\right)^2; \beta = v_c - v_s;$ $f_{rp} = \beta \cdot \frac{2t}{D-2t} \cdot f_y; f_{cc} = f_c + 4 \cdot f_{rp}; N_u = A_c \cdot f_{cc} + A_a \cdot f_y$	$v'_c$ : Empirical factor. $v_s$ : Poisson ratio of a steel tube, taken equal to 0.5. $f_{rp}$ : Lateral pressure at the peak load. $f_{cc}$ : Confined compressive strength of concrete. $N_u$ : Axial capacity of CFST column.
[9]	$\sigma_h = f_y \cdot \exp \left[ \ln \left( \frac{D}{t} \right) + \ln \left( \frac{f_y}{f_c} \right) - 11 \right] \cdot f_y; f_{rp} = \frac{2t}{D-2t} \sigma_h; f'_{cc} = f'_c + k_1 \cdot f_{rp}$ $f_y = 0.5 \cdot \left( \sigma_h - \sqrt{4 \cdot f_y^2 - 3 \cdot \sigma_h^2} \right); N_u = A_c \cdot f_{cc} + A_a \cdot f_y$	$\sigma_h$ : Hoop stress of the steel. $f_y$ : Yield strength of steel.
[1]	$v_a = 0.3, v_c = 0.2, \varepsilon_v = 0.002; \varepsilon_{ahr} = \frac{\varepsilon_v \cdot (v_a - v_c)}{1 + \frac{2t}{D-2t} \cdot \frac{E_a}{E_c}}$ $\varepsilon_{ah} = -v_a \cdot \varepsilon_v + \varepsilon_{ahr}; \sigma_h = \frac{E_a}{1 - v_a^2} \cdot (\varepsilon_{ah} + v_a \cdot \varepsilon_{al});$ $\sigma_{al} = \frac{E_a}{1 - v_a^2} \cdot (\varepsilon_v + v_a \cdot \varepsilon_{ah}); \sigma_{lat} = \sigma_{ah} \cdot \frac{2t}{D-2t}; k = 1.25 \cdot \left( 1 + 0.062 \cdot \frac{\sigma_{lat}}{f_{ct}} \right) \cdot f_c^{-0.21} \cdot f_c$ $f_{cc} = f_c \cdot \left( \frac{\sigma_{lat}}{f_{ct}} + 1 \right)^k; N_u = A_c \cdot f_{cc} + A_a \cdot \sigma_{al}$	$v_a, v_c, \varepsilon_v$ : Initial considered values. $\varepsilon_{ahr}$ : Restrained steel strain. $\varepsilon_{ah}$ : Final lateral strain of steel. $\sigma_{al}$ : Steel's longitudinal stress. $\sigma_{lat}$ : Compressive confining pressure. $k$ : Reflects the effectiveness of confinement. $f_{ct}$ : Tensile strength of concrete.
[11]	$f_{cc} = f_c + 3.4 \cdot \sigma_{rc}; \sigma_{rc} = \frac{\theta(k-1)}{\sqrt{9+3(k-1)^2}}; F_{Fa} = f_c \cdot A_c \left( 1 + \frac{k\theta}{2} \right)$	$\sigma_{rc}$ : Confining pressure around the concrete. $\theta$ : confinement index, $\frac{f_s A_s}{f_c A_c}$ $k=3 \div 4.3$
[8]	$\sigma_{ccB} = \gamma_U \cdot f'_c + 4.1 \cdot \sigma_r; \gamma_U = 1.67 D_c^{-0.112}; \sigma_r = \frac{-2t}{D-2t} \sigma_{s\theta}$ $\sigma_{s\theta} = \alpha_u \cdot \sigma_{sy} \text{ with } \alpha_u = -0.19$ $\sigma_{sz} = \beta_{uc} \cdot \sigma_{sy} \text{ with } \beta_{uc} = 0.89$ $N_u = (A_c \cdot \sigma_{ccB} + A_s \cdot \sigma_{sz})$	$\sigma_{ccB}$ : Strength of confined concrete. $\gamma_U$ : Strength reduction factor for concrete. $\sigma_r$ : Lateral pressure. $\sigma_{s\theta}$ : Hoop stress of steel tube in yield condition. $\sigma_{sy}$ : Tensile yield stress of steel tube. $\sigma_{sz}$ : Axial yield stress of steel tube.
[10]	$f'_{cc} = \gamma_c \cdot f'_c + k_1 \cdot f_{rp}$ $f_{rp} = \begin{cases} 0.7(v_e - v_s) \frac{2t}{D-2t} f_{sy} & \text{for } \frac{D}{t} \leq 47 \\ \left( 0.006241 - 0.0000357 \frac{D}{t} \right) f_{sy} & \text{for } 47 \leq \frac{D}{t} \leq 150 \end{cases}$ $v_e = 0.2312 + 0.3528 \cdot v'_e - 0.1524 \cdot \frac{f'_c}{f_{sy}} + 4.843 \cdot v'_e \cdot \frac{f'_c}{f_{sy}} - 9.169 \cdot \left( \frac{f'_c}{f_{sy}} \right)^2$ $v'_e = \frac{0.881}{10^6} \cdot \left(\frac{D}{t}\right)^3 - \frac{2.58}{10^4} \cdot \left(\frac{D}{t}\right)^2 + \frac{1.953}{10^2} \cdot \left(\frac{D}{t}\right) + 0.4011$ $\gamma_c = 1.85 D_c^{-0.135} \text{ for } (0.85 \leq \gamma_c \leq 1.0); \gamma_s = 1.458 \left( \frac{D}{t} \right)^{-0.1} \text{ for } (0.9 \leq \gamma_s \leq 1.1)$ $N_u = (\gamma_c f'_c + 4.1 \cdot f_{rp}) A_c + \gamma_s f_{sy} A_s$	$f'_{cc}$ : Strength of confined concrete. $f_{rp}$ : Lateral pressure. $\gamma_c$ : Strength reduction factor for concrete. $\gamma_s$ : Strength factor for steel tube. $f_{sy}$ : Tensile yield strength of steel. $v'_e$ : Empirical factor. $v_e$ : Poisson ratio of a steel tube filled with concrete
[3]	$p_0 = -\frac{\alpha}{2} \frac{2\mu' + 1}{[3(\mu'^2 + \mu' + 1)]^{\frac{1}{2}}} f_y; \mu' = -\frac{1}{2} - \frac{1}{2(\xi + 1)}; \xi = \alpha \frac{f_y}{f_c}$ $\alpha = \frac{A_s}{A_c}; N_u = N_c + N_s; N_s = \frac{\mu' + 2}{[3(\mu'^2 + \mu' + 1)]^{\frac{1}{2}}} f_y A_s$ $N_c = (f_c + 4p_0) A_c$	$p_0$ : Lateral confining pressure of the steel tube on concrete. $N_c$ : Sectional strength of concrete core. $N_s$ : Sectional strength of steel tube.

A total of 178 circular CFST columns were chosen for the calculation. These columns were tested under concentric compression on the entire section. Dimensions and material properties of the selected columns are summarized in Tables II and III. The main dimensions of the columns are described by the outer diameter (D), the thickness of steel tube (t), and the length of the columns (L). The material properties include the

unconfined concrete strength ( $f_c$ ) and the yield strength of the steel tube ( $f_y$ ). The selected columns were divided into 2 groups: (1) group 1 includes CFST short columns with NSC ( $f_c \leq 60$ MPa), (2) group 2 includes CFST short columns with HSC ( $60$ MPa  $< f_c \leq 120$ MPa) and UHSC ( $f_c > 120$ MPa). The test database covers a wide range of concrete strengths.

TABLE II. TEST DATABASE OF CIRCULAR CFST COLUMNS USING NSC (GROUP 1)

Tested by	D (mm)	t (mm)	L (mm)	f <sub>c</sub> (MPa)	f <sub>y</sub> (MPa)
[13], 18 specimens	76.4-152.6	1.68-4.09	152.3-304.9	20.9-40.9	363.3-451.6
[14], 26 specimens	101.0-318.5	3.03-10.37	305-955	23.2-52.2	371-452
[15], 42 specimens	76.5-300	1.5-4.5	276-1000	20.54-45.77	232.3-433.2

TABLE III. TEST DATABASE OF CIRCULAR CFST COLUMNS USING HSC AND UHSC (GROUP 2)

Tested by	D (mm)	t (mm)	L (mm)	f <sub>c</sub> (MPa)	f <sub>y</sub> (MPa)
[15], 18 specimens	108-133	1.0-7.0	378-465	106-116	232-429
[16], 26 specimens	60-250	1.87-2.0	180-750	85.2-90.0	282-404
[17], 21 specimens	101.6-139.8	2.37-3.0	304.8-419.4	62.4-135.6	341-462.6
[18], 27 specimens	149-165	1.0-6.0	500	87.1-91.8	338-438

The models proposed in [1, 2, 6, 8-11] were used to predict the axial capacity and lateral confining pressure of circular CFST columns. However, the model of [6] is only applicable for  $f_c/f_y$  ranging from 0.04 to 0.2. The columns with HSC and UHSC have  $f_c/f_y > 0.2$ , thus the model of [6] was not further considered.

For the assessment of the accuracy of the prediction, two statistical indicators were determined, the Mean Square Error (MSE) and the Average Absolute Error (AAE):

$$MSE = \frac{\sum_{i=1}^n \left( \frac{pre_i - exp_i}{exp_i} \right)^2}{N} \quad (1)$$

$$AAE = \frac{\sum_{i=1}^n \left| \frac{pre_i - exp_i}{exp_i} \right|}{N} \quad (2)$$

where  $pre_i$  represents values of ultimate load from model predictions,  $exp_i$  represents the values of ultimate load from the experimental results, and  $N$  is the total number of test data.

The confinement ratio  $\emptyset$  is defined by:

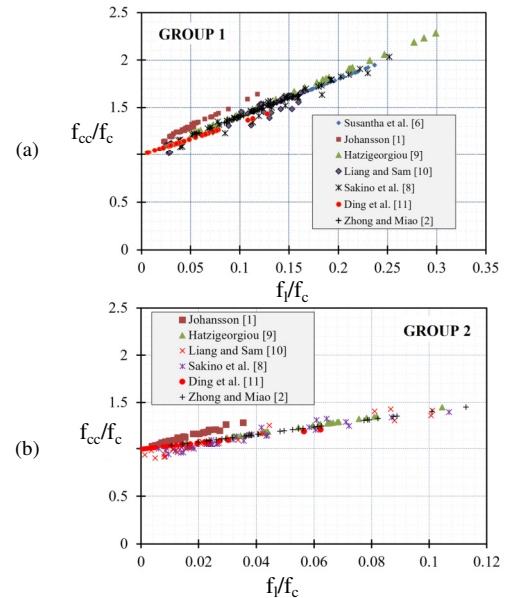
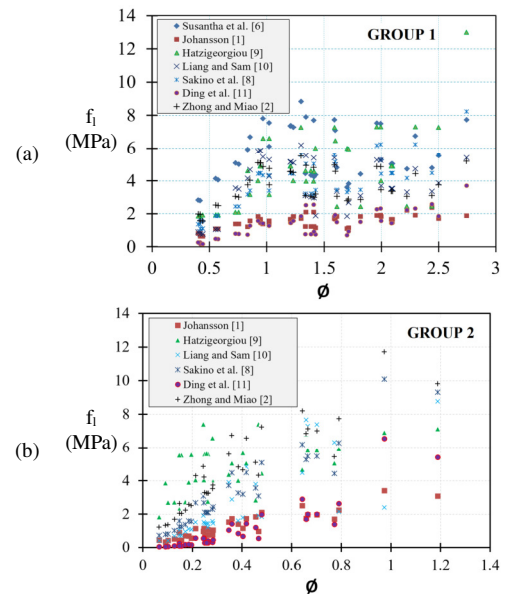
$$\emptyset = \frac{A_s f_y}{A_c f_c} \quad (3)$$

where  $A_s$  is the area of steel section, while  $A_c$  is the area of concrete section.

### III. RESULTS AND DISCUSSION

The relation between the ratios  $f_{cc}/f_c$  and  $f_l/f_c$ , where  $f_l$  is the lateral confining pressure and  $f_{cc}$  is the confined concrete strength, is shown in Figure 1. It can be seen that all the models describe a quite similar trend. This relation can be represented by a linear equation. As the ratio  $f_l/f_c$  increases, the ratio  $f_{cc}/f_c$  also increases considerably in the case of group 1 while the increment of  $f_{cc}/f_c$  is slighter in the case of group 2, due to the higher confinement when employing NSC for CFST columns. It could be observed that in group 1, the models of [6, 8, 9] exhibit the same predicting value of  $f_{cc}/f_c$  at a given value of  $f_l/f_c$ , whereas the scatter plots of the other models show a different tendency. Furthermore, in group 2, the models suggested by [2, 8-11] are rather close to each other, while the

model of [1] gives an unlikely trend. Figure 2 illustrates the predicted value of confining pressure  $f_l$  versus the confinement ratio  $\emptyset$ .

Fig. 1. Comparison of models in prediction of  $f_{cc}/f_c$  and  $f_l/f_c$ .Fig. 2. Comparison of models in prediction of confining pressure  $f_l$ .

Overall, the graph indicates the variation of confining pressure  $f_l$  corresponding to each model. It can be seen that when the confinement ratio increases, the confinement pressure shows an upward trend. The models of [1, 11] exhibited a very low value of confining pressure in comparison with the other models, while the models of [6, 9] presented the highest values. It is evident from the data that the confining pressure predicted from [10] was very close to [2] in the case of group 1. Moreover, the model of [2] gives higher values of  $f_l$  in the case

of group 2. It is apparent from the scatter plot that the confinement ratio  $\emptyset$  has no correlation with the confining pressure. The above results can be explained by the fact that except for the model of [10], all models consider high confining pressure of steel tube in circular CFST stub columns. However, the confining pressure of steel tube decreases with increasing concrete strength.

Figure 3 shows the comparison of ultimate load between model predictions ( $N_{pre}$ ) and EC4 [12] predictions ( $N_{EC4}$ ). In order to better reflect the deviations of prediction as compared to experimental results and EC4 [12], the -10% and +10% error bounds depicted in the Figure are presented in the following sub-sections. It has been demonstrated in group 1 that the predictions from [8-10] are quite close to those from EC4 [12] (around 10%), while the models from [2, 11] give less accurate values (about 20% higher and 20% lower on average, respectively). At the same time, the predictions by [1, 6] have varying values. Interestingly, it could be observed in group 2 that almost all models except the one of [11] have very close values (within 10%) to the ones of [12]. The agreement of EC4 and the predicted models is generally good. It should be noted that EC4 is applicable with concrete strength up to 70MPa. Therefore, the prediction by EC4 is less accurate for the circular CFST columns having concrete strength higher than 70MPa.

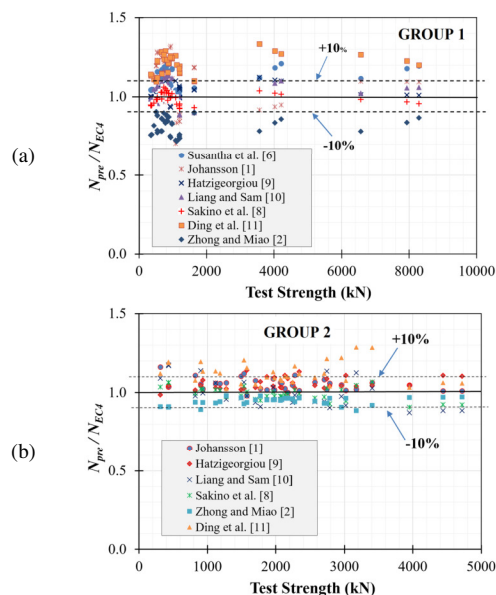


Fig. 3. Comparison of ultimate loads between the model predictions and EC4 [12].

The comparison between the ultimate strength from the real tests ( $N_{test}$ ) and the calculated strength from the models ( $N_{pre}$ ) is illustrated in Figure 4. Regarding group 1, the predictions of [6, 9, 10] give a quite good correlation with the test data (around 10%), whereas the models from [2, 11] exhibit large differences between the experimental and the predicted values. In addition, all models except the ones from [6, 11] generally tend to underestimate the strength of CFST columns when the test strength increases. Figure 4 clearly indicates that almost all

models show a large variability in individual prediction. However, the models from [8, 10] give a better precision than the other models in estimating the ultimate strength. On the other hand, the models from [1, 9, 11] overestimate these values. The models from [8, 10] use the strength reduction factor for concrete depending on the outer diameter of the concrete core. Therefore, these models give ultimate load prediction smaller than the other models.

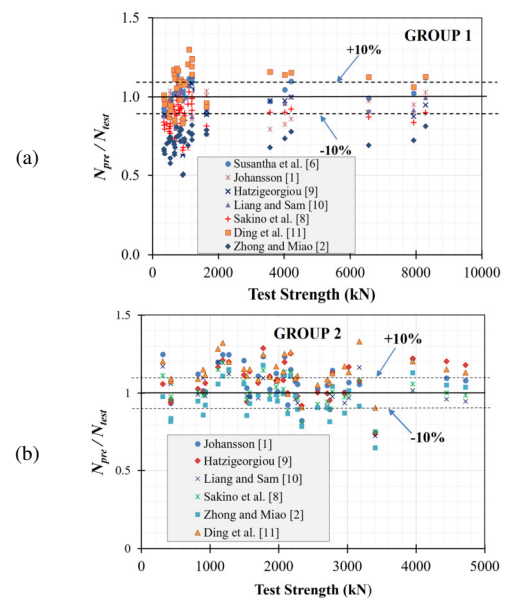


Fig. 4. Comparison of ultimate loads between the model predictions and the test results.

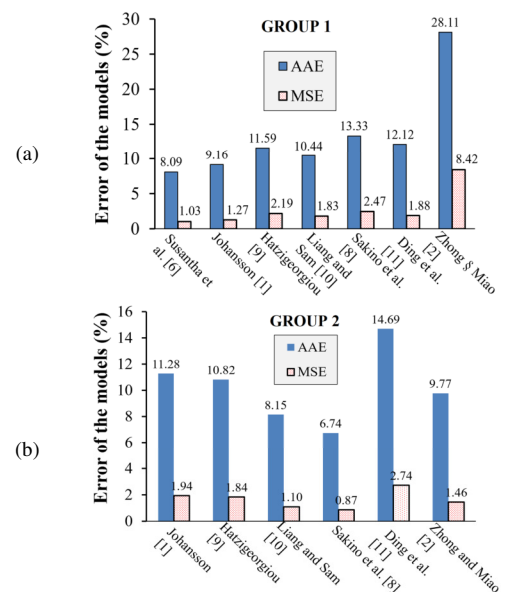


Fig. 5. Comparison of statistical indicators AAE and MSE.

Based on the two statistical indicators, AAE and MSE, shown in Figure 5 for the CFST stub columns employing NSC, the models from [1, 6] provide the best prediction of ultimate



strength because they have the smallest error (just under 9%), followed by the model from [10] with error around 10%. The models proposed in [8, 9, 11] show a good prediction, with error greater than 10%. Conversely, the error of the model from [2] is larger, thus this model should not be used for estimating the ultimate strength of NSC filled in steel tube stub columns.

Regarding group 2, Figure 5 shows that the models from [8, 10] give the smallest error. The errors in the prediction of ultimate strength in CFST columns using HSC and UHSC are much larger than those in CFST columns using NSC. This may partly be caused by the less effective confinement of CFST columns using higher strength concrete core, while almost all existing models consider the passive confinement in calculating the confined strength of concrete core.

#### IV. CONCLUSION

In this paper, models for predicting the lateral confining pressure of circular CFST short columns were considered and used for the calculation of the lateral stress and ultimate load of the tested columns. The main conclusions derived from the findings of this paper are:

- Almost all models give more accurate predictions for circular CFST short columns using NSC than HSC and UHSC.
- The prediction is less accurate with increasing concrete strength.
- There are significant differences among the models in lateral stress prediction.
- The models suggested in [8, 10] are suitable for predicting the ultimate strength of CFST columns with HSC and UHSC.
- The models of [1, 6] give accurate prediction of the ultimate strength for CFST columns using NSC.
- The ultimate load prediction obtained by the model of [11] is close to the ultimate load predicted by EC4 [12].

#### REFERENCES

- [1] M. Johansson, "The efficiency of passive confinement in CFT columns," *Steel and Composite Structures*, vol. 2, no. 5, pp. 379–396, Oct. 2002, <https://doi.org/10.12989/scs.2002.2.5.379>.
- [2] Z. Shan-Tong and M. Ruo-Yu, "Stress-Strain Relationship and Strength of Concrete Filled Tubes," presented at the Composite Construction in Steel and Concrete, Henniker, NH, USA, 1988, pp. 773–785.
- [3] M. D. O'Shea and R. Q. Bridge, "Design of Circular Thin-Walled Concrete Filled Steel Tubes," *Journal of Structural Engineering*, vol. 126, no. 11, pp. 1295–1303, Nov. 2000, [https://doi.org/10.1061/\(ASCE\)0733-9445\(2000\)126:11\(1295\)](https://doi.org/10.1061/(ASCE)0733-9445(2000)126:11(1295)).
- [4] J. B. Mander, M. J. N. Priestley, and R. Park, "Theoretical Stress-Strain Model for Confined Concrete," *Journal of Structural Engineering*, vol. 114, no. 8, pp. 1804–1826, Aug. 1988, [https://doi.org/10.1061/\(ASCE\)0733-9445\(1988\)114:8\(1804\)](https://doi.org/10.1061/(ASCE)0733-9445(1988)114:8(1804)).
- [5] M. M. Attard and S. Setunge, "Stress-Strain Relationship of Confined and Unconfined Concrete," *Materials Journal*, vol. 93, no. 5, pp. 432–442, Sep. 1996, <https://doi.org/10.14359/9847>.
- [6] K. A. S. Susantha, H. Ge, and T. Usami, "Uniaxial stress-strain relationship of concrete confined by various shaped steel tubes," *Engineering Structures*, vol. 23, no. 10, pp. 1331–1347, Oct. 2001, [https://doi.org/10.1016/S0141-0296\(01\)00020-7](https://doi.org/10.1016/S0141-0296(01)00020-7).
- [7] H.-T. Hu, C.-S. Huang, M.-H. Wu, and Y.-M. Wu, "Nonlinear Analysis of Axially Loaded Concrete-Filled Tube Columns with Confinement Effect," *Journal of Structural Engineering*, vol. 129, no. 10, pp. 1322–1329, Oct. 2003, [https://doi.org/10.1061/\(ASCE\)0733-9445\(2003\)129:10\(1322\)](https://doi.org/10.1061/(ASCE)0733-9445(2003)129:10(1322)).
- [8] H.-T. Hu, C.-S. Huang, M.-H. Wu, and Y.-M. Wu, "Nonlinear Analysis of Axially Loaded Concrete-Filled Tube Columns with Confinement Effect," *Journal of Structural Engineering*, vol. 129, no. 10, pp. 1322–1329, Oct. 2003, [https://doi.org/10.1061/\(ASCE\)0733-9445\(2003\)129:10\(1322\)](https://doi.org/10.1061/(ASCE)0733-9445(2003)129:10(1322)).
- [9] G. D. Hatzigeorgiou, "Numerical model for the behavior and capacity of circular CFT columns, Part I: Theory," *Engineering Structures*, vol. 30, no. 6, pp. 1573–1578, Jun. 2008, <https://doi.org/10.1016/j.engstruct.2007.11.001>.
- [10] Q. Q. Liang and S. Fragomeni, "Nonlinear analysis of circular concrete-filled steel tubular short columns under axial loading," *Journal of Constructional Steel Research*, vol. 65, no. 12, pp. 2186–2196, Dec. 2009, <https://doi.org/10.1016/j.jcsr.2009.06.015>.
- [11] F. Ding, Z. Yu, Y. Bai, and Y. Gong, "Elasto-plastic analysis of circular concrete-filled steel tube stub columns," *Journal of Constructional Steel Research*, vol. 67, no. 10, pp. 1567–1577, Oct. 2011, <https://doi.org/10.1016/j.jcsr.2011.04.001>.
- [12] *Eurocode 4, Design of Composite Steel and Concrete Structures, Part 1.1, General Rules and Rules for Buildings*. Brussels, Belgium: CEN, 2004.
- [13] N. J. Gardner and E. R. Jacobson, "Structural Behavior of Concrete Filled Steel Tubes," *Journal Proceedings*, vol. 64, no. 7, pp. 404–413, Jul. 1967, <https://doi.org/10.14359/7575>.
- [14] T. Yamamoto, J. Kawaguchi, and S. Morino, "Experimental Study of Scale Effects on the Compressive Behavior of Short Concrete-Filled Steel Tube Columns," in *Composite Construction in Steel and Concrete IV Conference 2000*, Apr. 2012, pp. 879–890, [https://doi.org/10.1061/40616\(281\)76](https://doi.org/10.1061/40616(281)76).
- [15] J. Tang, S. Hino, I. Kuroda, and T. Ohta, "Modeling of Stress - Strain Relationships for Steel and Concrete in Concrete Filled Circular Steel Tubular Columns," *Steel Construction Engineering*, vol. 3, no. 11, pp. 35–46, 1996, [https://doi.org/10.11273/jssc1994.3.11\\_35](https://doi.org/10.11273/jssc1994.3.11_35).
- [16] L.-H. Han, W. Li, and R. Bjorhovde, "Developments and advanced applications of concrete-filled steel tubular (CFST) structures: Members," *Journal of Constructional Steel Research*, vol. 100, pp. 211–228, Sep. 2014, <https://doi.org/10.1016/j.jcsr.2014.04.016>.
- [17] S. Motoo and M. Kazuaki, "Experimental Study on Ultimate Behavior of Ultra-High-Strength Concrete-Filled Steel Pipe Shear-Bending Columns," *Proceedings of the Department of Construction of the Architectural Institute of Japan*, vol. 64, no. 520, pp. 133–140, 1999, [https://doi.org/10.3130/aijs.64.133\\_2](https://doi.org/10.3130/aijs.64.133_2).
- [18] Q. Yu, Z. Tao, and Y.-X. Wu, "Experimental behaviour of high performance concrete-filled steel tubular columns," *Thin-Walled Structures*, vol. 46, no. 4, pp. 362–370, Apr. 2008, <https://doi.org/10.1016/j.tws.2007.10.001>.
- [19] P. C. Nguyen, D. D. Pham, T. T. Tran, and T. Nghia-Nguyen, "Modified Numerical Modeling of Axially Loaded Concrete-Filled Steel Circular-Tube Columns," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7094–7099, Jun. 2021, <https://doi.org/10.48084/etasr.4157>.
- [20] A. N. Hassooni and S. R. A. Zaidee, "Behavior and Strength of Composite Columns under the Impact of Uniaxial Compression Loading," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 8843–8849, Aug. 2022, <https://doi.org/10.48084/etasr.4753>.
- [21] Z. H. Abdulghafoor and H. A. Al-Baghaddi, "Static and Dynamic Behavior of Circularized Reinforced Concrete Columns Strengthened with Hybrid CFRP," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9336–9341, Oct. 2022, <https://doi.org/10.48084/etasr.5162>.

# Towards an IOT Approach for Smart Waste Management based on Context Ontology: A Case Study

**Nadia Aloui**

Department of Software Engineering, College of Computer Sciences and Engineering, University of Jeddah, Saudi Arabia | Sfax University, Tunisia  
naaloui@uj.edu.sa

**Wafa Almukadi**

Department of Software Engineering, College of Computer Sciences and Engineering, University of Jeddah, Saudi Arabia  
wsalmukadi@uj.edu.sa

**Aymen Belghith**

College of Computing and Informatics, Saudi Electronic University, Saudi Arabia  
a.belghith@seu.edu.sa  
(corresponding author)

*Received: 28 November 2022 | Revised: 10 December 2022 | Accepted: 12 December 2022*

## ABSTRACT

Nowadays, waste management faces the challenge of providing effective and efficient solutions for waste collection, disposal, and recycling while respecting health and environmental standards. This challenge also includes the lack of understanding of the diverse factors that influence the various stages of waste management, inefficient route planning, insufficient resources, etc. Faster collection, management, and processing of waste are possible with smart containers and IoT technologies allowing waste real-time data provision. Thus, this research proposes a waste management system based on generic and comprehensive generic context ontology and smart containers. The context ontology is conceived to solve the limits and the insufficiencies of waste management by covering all the waste facets for all the stakeholders, optimizing, analyzing, and reusing the waste data and conditions. For given smart container and waste management context, we need to have a global view of the relevant contextual data according to a unified model such as the waste environment data, the waste activity context, the waste computing context, the user context, the collaboration context, etc. One significant advantage of our system is that it provides a unified model for waste management contextual data that could be reused for other waste management systems covering all the properties of this domain. The proposed solution implements an intelligent and adaptive IoT system for waste management according to different waste contexts, waste objectives, and waste activities. The proposed system has been successfully tested under different scenarios in Jeddah City Municipality.

**Keywords-**smart waste management; waste context ontology; IoT; smart container

## I. INTRODUCTION

Waste management, from generation to disposal, is one of the biggest challenges that municipal organizations face today. Due to the growth in waste, containers positioned around cities in open spaces are overflowing, putting the citizens in an unhygienic environment. The implementation of smart waste management systems remains in an embryonic stage, something that is expected to change with urban development, similarly to other engineered systems [1]. To overcome this, a smart waste management system is proposed, based on smart

containers that enable context-specific waste management via a general exhaustive context ontology. Real-time monitoring through smart devices or containers is furthering the digital era of waste management by enhancing transparency, dependability, agility, security, resilience, connectivity, and sustainability of waste chains. The proposed system, based on smart containers and general context ontology, is leading the way in developing smart context-based waste management systems, increasing the quality of life today and assuring a greener world for future generations. The adaptation of a system context can take many aspects, such as behavior,

content, or presentation adaptation. In this approach, focus is given on the adaptation of a smart waste management system to different waste contexts of different stakeholders. It is possible to manage the waste process situations in different contexts, i.e. the waste management domain, waste objectives, and waste activities, such as waste generation, collection, transformation, segregation, cleaning, etc. The proposed solution implements an intelligent and adaptive system for waste management according to the context and the different objectives according to various stakeholders. The main contributions of this paper are:

- A unique way to combine two technologies, namely IoT and ontological engineering, ensuring an optimal and general based context in the waste management field.
- An architectural development process of the smart garbage box and the process of complete waste management in addition to the objectives, activities, reuse, and intelligent learning of the waste management system, as well as a smart way to monitor waste in real time.

## II. LITERATURE REVIEW

Waste management stakeholders looking for a real-time monitoring, smart, and adaptive solution, will eventually have to use smart containers that are permanently equipped with IoT devices. In this section, a brief overview of smart containers and smart devices is presented.

### A. The Impact of Smart Containers and IOT on Waste Management

It is worth noting that technological IoT advancements [2] are a key part of the waste problem that we face in industrial processes. Fortunately, technological advancements also provide innovative solutions that can help us meet such challenges when they emerge. IoT is completely changing the game regarding industrial waste management best practices. Firstly, by using IoT services and solutions, we can significantly improve the efficacy of waste collection and recycling. We can automate, improve, and interpret the entire waste management process in new ways while reducing cost. Smart containers save time and effort [3], as operators know the number of containers that must be collected per day and the number of trucks that go to collect containers, in addition to saving time in classifying waste by type, making it easier for workers to send them to the competent authorities for recycling. The main benefit of IoT is the ability to collect huge

amounts of data and update them in real time utilizing cost-effective sensors. These inexpensive sensors, recording and monitoring equipment, send all their data to the cloud for storage and review. Then, AI models can analyze the data and draw actionable conclusions. In a series of case studies, we'll look at how companies in the food, agriculture, transportation, and energy sectors have implemented smart waste management.

### B. Related Work

The literature includes a variety of waste management-related research initiatives. Through a design that uses solar energy to feed the system and sensors to keep track of how much waste has been gathered inside the enclosure, authors in [3] introduce the idea of intelligent disposal. If necessary, the container can compress the waste, reducing its volume up to 10 times, even before it is collected. The waste monitoring solution SENSONEO [4] combines smart sensors, a smart waste management system, and a citizen app. Smart sensors use ultrasound technology to measure the filling levels in boxes and containers several times a day and send data to the waste management system. It divides the container into several sections. Authors in [5] offer a novel method of implementing an integrated sensing system, which automates the solid waste management process, using cloud technology and mobile app-based monitoring. The suggested smart trash can is built on ultrasonic sensors and a variety of gas sensors. Authors in [6] suggest an IoT-based smart waste management system that aids surveying the estimation of garbage in trash cans and then transmits data through the Internet to a server. A specific suggestion focusing on the intelligent container was originally made in [7]. In order to prevent trash disposal outside the container, the authors suggest a method in which monitoring takes place both inside and outside the compartment with the help of infrared sensors. The system in [8] includes a method for monitoring dumpsters around-the-clock. A smart and well-organized mechanism is in place here for selective clearing. The level of waste in the dumpster is determined using the ultrasonic sensor.

### C. Comparison of Smart Containers/Waste Management Systems

To make the system innovative and optimized, a comparison research based on a variety of criteria has been conducted (Table I). The primary factors considered in this study are the detection of container fullness, and the used technologies.

TABLE I. COMPARISON BETWEEN THE PROPOSED SYSTEM AND OTHER, KNOWN, SYSTEMS

Criteria	Reference						
	[3]	[4]	[5]	[6]	[7]	[8]	Proposed
Intelligent route planning	√	√	x	x	x	√	√
Container fullness	√	√	√	√	√	√	√
Used technologies	Solar energy and sensors	IoT	Cloud technology and mobile app	IoT	Infrared sensors	IoT and cloud	IoT and ontological engineering
Includes smart bin	√	√	x	√	√	√	√
Recycling and disposal	x	x	x	x	√	√	√
Full waste management process	x	√	x	x	x	√	√
Context awareness	x	x	x	x	√	x	√
Context model	x	x	x	x	x	x	√

There are many waste classification techniques currently available, but many of them require human intervention. If a fully automatic system is deployed in a company, it will be a win-win situation for government, companies, and industry. The majority of the research discussed here focuses on the collection system when discussing waste management and doesn't give interest to the whole process of waste management. Previous studies did not take into account the waste management context, objectives, and activities according to the whole waste management process and all the waste management factors. They are not based on context models specific and exhaustive for the waste management field. It is assumed that the best waste management approach has to be based on context and reuse. Thus, the services offered for waste management will be more adapted, optimized and context aware. The proposed system in this research takes into account the contributions in the literature and covers their limits.

#### D. Context and Context-Awareness of Ubiquitous and Smart Systems

Context is all that surrounds and gives a meaning to something [9, 10]. A definition of context cannot be given without taking into account the element concerned by the context. This definition shows that the context must be external to the element related to this context. In the field of ubiquitous computing (computing pervasive), context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves [11-14]. There are mainly two types of context [16]. The first one is active context use (automatically adapting the application's behavior in accordance with the state of the context) and the second one is passive context use (visualizing the state of the context for the user or saving this state for its future use).

In the literature, many definitions of context-awareness have been proposed. Authors in [16] consider context-awareness as the set of context-aware applications. Another definition given in [10] specifies that a system is considered as context-aware, if it uses the context to give pertinent information or a service to a user, knowing that the relevance depends on the user's activity. Furthermore, context-awareness emerged in the fields of mobile and pervasive computing as a technique to design applications with a conscience of the environment, to ensure high level of autonomy and flexibility. The context-awareness or the conscience of the context is known under other synonyms like adaptive or reactivate [17]. In order to model the smart container context, we use the concept of ontology, which provides an explicit specification of a conceptualization and an important step for reuse [18-20]. Moreover, "an ontology is expressed in language (representation) based on a theory (semantics) that guarantees the properties of the ontology in terms of consensus, consistency, reuse, and sharing." [21]. The proposed generic context ontology will be able to guarantee unified, coherent, shared, and reused smart containers management in different contexts.

### III. PROPOSED GENERIC CONTEXT ONTOLOGY FOR SMART WASTE MANAGEMENT

There are several methods for developing ontologies (top-down, bottom-up, and combined). To design the context ontology for waste management, we adopt the iterative method recommended in [22].

#### A. Field and Scope of the Ontology

We proceed to the ontology development process by defining its domain and scope. Indeed, ontology is the structured study of a specific field [23], which is accomplished by answering the following questions: What domain will the ontology cover? What are the objectives of ontology creation? Who will utilize the ontology? Our ontology field is the waste management context-based IoT of all actors. Our context ontology therefore includes concepts related to the waste management activity through smart containers. This ontology is designed to integrate both the active and passive context (smart container context). Moreover, our ontology is not intended to be used by a human, but by the different components of context management to provide a better IoT waste management in the city in a given context.

#### B. Generic Context Ontology

A current representation of the context object is available for waste management via smart containers according to the generic context ontology. By using this context model, the system can become aware of the current situation and adapt its services for various stakeholders in response to environmental changes. To give the waste management system exact values for these attributes, this generic ontology identifies the characteristics of the current context. These attributes are used to specify the smart container's contextual aspect in an active waste management situation and to adapt the content to that context and the current waste management activity and objective. We offer a general context model based on the different state-of-the-art models, and its attributes are organized into six facets:

##### 1) The Actor Context

All participants in a waste management environment are referred to as actors. We define the identifier and the actor as the two elements that constitute the actor context aspect. The first element uniquely identifies the actor's smart containers for waste management environments. The second element is utilized to link the waste management purpose and action to the actor who used, will use, or reuse the smart container in another waste management context.

##### 2) The Waste Activity Context

Waste management domain, waste management activity, waste activity objective, generation, collection, transformation, segregation, cleaning, and the collaborative waste activities are the elements the waste activity context.

##### 3) The Waste Environment Context

The environment context of a smart container waste management situation is described by a variety of factors that affect waste, including time, place, temperature, humidity,

weather conditions, level, type, status, priority, classification and duration of waste, etc.

#### 4) The Waste Computing Context

This aspect of context describes the equipment and tools the actor used to carry out his waste management task, including hardware, waste management platform, organizer, and operating system.

#### 5) The Waste Reuse Context

Two elements constitute the reuse context: the reuse identifier, which determines how much waste may be reused in each smart container, and the reuse percentage, which increases the value and effectiveness of waste management.

#### 6) The Context of Waste Management Objectives

The waste management objectives facet can be specified through the waste management platform or environment using the context-specific waste management objective aspect. Knowledge of waste management objective, objective type, and waste management evaluation are the parts of the process.

### IV. PROPOSED APPROACH FOR SMART WASTE MANAGEMENT BASED ON SMART CONTAINERS

Traditional containers that have been fitted with electronics to allow them to detect, interact, and communicate are known as smart containers. The additional electronics allow for tracking and monitoring of a container as well as the circumstances surrounding the management and transportation of the container's waste. The smart container can be made to exchange physical information close to real-time, including location, door opening and closing events, shocks and vibrations, temperature, humidity, and any other pertinent physical factors (Figure 1).

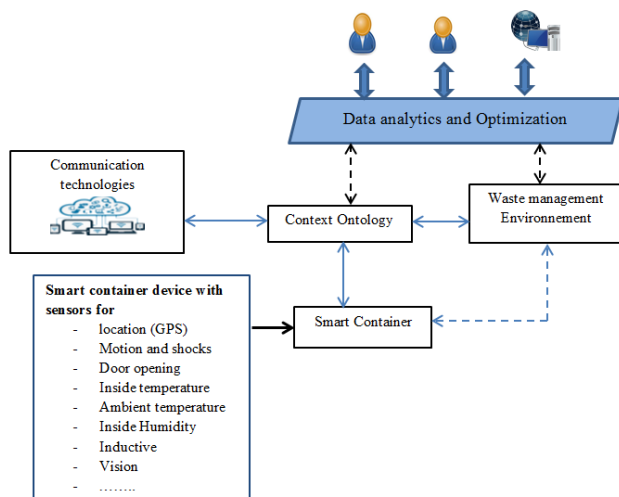


Fig. 1. The proposed approach for smart waste management.

In order to achieve shared visibility amongst diverse stakeholders during the whole smart container journey, the smart container serves as the common data source. In order to clarify the roles and responsibilities of each actor in the waste management chain and to improve collaboration and

coordination among stakeholders some essential tasks must be accomplished: (i) shared visibility, (ii) assessing the activities involved for the purpose and determining the type, nature, and estimated volumes of waste to be generated, (iii) identifying any potential environmental impact from the generation of waste at the site, and (iv) recommending appropriate waste handling and disposal measures/routings in accordance with the current legislative and administrative requirements. It is suggested that smart containers could be integrated with other technological advancements to benefit the waste management, trade, and transportation communities even more.

### V. A CASE STUDY OF THE PROPOSED SMART WASTE MANAGEMENT SYSTEM

#### A. Study Area

The study case chosen for developing this waste management system based on smart containers and context model is the Municipality of Jeddah. For the development of the smart waste management system, the work was divided into the following stages.

#### B. Functional and Non-functional Requirements

The functional requirements outline the tasks that the suggested solution must complete in order to function and the non-functional requirements are those aspects of the system that are visible as it is being used. The required specifications are:

- The administrator can:
  1. Edit the existing containers in the system
  2. Establish the waste collection plan according to the waste management activities.
  3. Select the best route to collect containers and send information to the collector.
  4. Track smart containers and trucks locations.
  5. View the full container status and context features affecting the waste management activities.
  6. View most frequently utilized portion in each container, the full section, the most used section, the total number of containers that have been collected and the section percentage.
- The collector can:
  1. Collect containers that change color on the map.
  2. View the location of smart containers.
  3. Collect containers according to smart containers priority and waste management activities and objectives
- The Jeddah municipality can:
  1. Create waste collection report in a daily and or monthly basis according to different waste activity objectives such as waste generation, collection, transformation, segregation, etc.

2. Generate peak day and time of waste collection and prevent waste risks.

### C. Hardware Components

This section represents the hardware used throughout the implementation of the smart container for waste management. We used NEO-6M GPS to detect the location, an inductive sensor to detect metal, an ultrasonic sensor to detect the level of trash, an ESP8266 NodeMCU to connect with the firebase, an Arduino Uno to do programming over the other products, etc.

### D. Prototype

Figure 2 depicts the view of the finalized hardware prototype of the smart container of the waste management system.

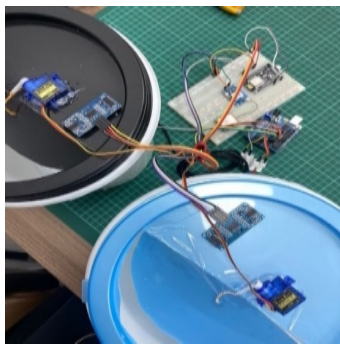


Fig. 2. Smart container prototype.

### E. Interface

The smart waste management system based on smart containers has been implemented by the students of the College of Computer Science and Engineering of Jeddah University. Screenshots of the User Interface are exhibited below. Figure 3, the system displays several optimal paths and the best path is indicated as the "best choice", and the administrator will determine the path that suits him.

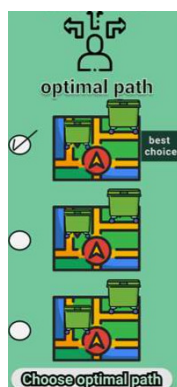


Fig. 3. Choose optimal path page.

As can be seen in Figure 4, the waste management system users have access to improved visibility identifying the waste inbound flow to optimize waste management and react to unexpected risks. In Figure 5, the administrator must be able to

see if the waste in the section has the same type for example. A warning sign will appear if there are different types of waste in a section. The percentage and quantity of each type also appear. The smart sensors and IOT technologies utilized in this waste management system that gathers and tracks real-time data allow the administrator to, dispose of, reduce, reuse, and prevent waste collecting peaks (Figure 6).

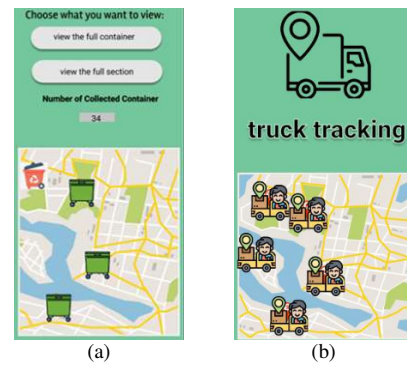


Fig. 4. (a) Container data in real time, (b) truck tracking.

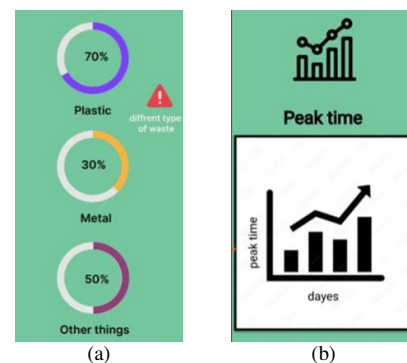


Fig. 5. Contained types of waste, (b) peak time and waste risk.



Fig. 6. Waste management report and statistics.

## VI. CONCLUSION AND FUTURE WORK

The context ontology is opening up a variety of data sources that could give the waste management system a context-aware picture. Any documentation of the contextual information contained in a container is helpful to the coordination and predictability assuring quality and optimizing



waste management. It is possible to predict when a container will be ready for the next operation and effectively manage risk in accordance with the waste management process by simply being aware of the status of a specific container, its contents, all relevant contextual information, and the waste management objective at a given time. As a perspective of this work, we suggest the use of deep learning methods as an intelligent way to classify waste through image classification. Machine learning algorithms can be very beneficial in the waste management data analytics in order to take the best decisions and optimize smart waste management.

#### ACKNOWLEDGMENT

The authors would like to thank Aizah Alsomai, Njoud Barayan, Raghad Alghamdi, Shroug Alzhrani for their contribution in implementing the case study application of smart waste management during their senior project.

#### REFERENCES

- [1] D. M. S. H. Dissanayaka and S. Vasanthapriyan, "Smart Solid Waste Management — a Scientific Literature Review," *The Journal of Solid Waste Technology and Management*, vol. 46, no. 4, pp. 606–619, Nov. 2020, <https://doi.org/10.5276/JSWTM/2020.606>.
- [2] B. E. Sabir, M. Youssfi, O. Bouattane, and H. Allali, "Towards a New Model to Secure IoT-based Smart Home Mobile Agents using Blockchain Technology," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5441–5447, Apr. 2020, <https://doi.org/10.48084/etasr.3394>.
- [3] I. Sosunova and J. Porras, "IoT-Enabled Smart Waste Management Systems for Smart Cities: A Systematic Review," *IEEE Access*, vol. 10, pp. 73326–73363, 2022, <https://doi.org/10.1109/ACCESS.2022.3188308>.
- [4] *Waste Monitoring*. SENSONEO, 2019.
- [5] D. Misra, G. Das, T. Chakraborty, and D. Das, "An IoT-based waste management system monitored by cloud," *Journal of Material Cycles and Waste Management*, vol. 20, no. 3, pp. 1574–1582, Jul. 2018, <https://doi.org/10.1007/s10163-018-0720-y>.
- [6] C. S. Srikanth, T. B. Rayudu, J. Radhika, and R. Anitha, "Smart Waste Management using Internet-of-Things (IoT)," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 9, pp. 2518–2522, Jul. 2019, <https://doi.org/10.35940/ijitee.G5334.078919>.
- [7] T. S. Vasagade, S. S. Tamboli, and A. D. Shinde, "Dynamic solid waste collection and management system based on sensors, elevator and GSM," in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, India, Mar. 2017, pp. 263–267, <https://doi.org/10.1109/ICICCT.2017.7975200>.
- [8] A. Doshi, B. Shah, and J. Kamdar, "BINTERNET: Smart Waste Management System," *International Journal for Research in Applied Science and Engineering Technology*, vol. 9, no. 6, pp. 703–707, Nov. 2021, <https://doi.org/10.22214/ijraset.2021.38882>.
- [9] "FOLDDOC - Computing Dictionary," *Free On-Line Dictionary of Computing*. <https://foldoc.org/>.
- [10] A. K. Dey, "Understanding and Using Context," *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, Feb. 2001, <https://doi.org/10.1007/s007790170019>.
- [11] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context Aware Computing for The Internet of Things: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 414–454, 2014, <https://doi.org/10.1109/SURV.2013.042313.00197>.
- [12] U. Alegre, J. C. Augusto, and T. Clark, "Engineering context-aware systems and applications: A survey," *Journal of Systems and Software*, vol. 117, pp. 55–83, Jul. 2016, <https://doi.org/10.1016/j.jss.2016.02.010>.
- [13] F. Azouaou and C. Desmoulin, "Taking Teaching Context into Account for Semantic Annotation Patterns," in *Innovative Approaches for Learning and Knowledge Sharing*, 2006, pp. 543–548, [https://doi.org/10.1007/11876663\\_50](https://doi.org/10.1007/11876663_50).
- [14] P. Rosenberger and D. Gerhard, "Context-awareness in industrial applications: definition, classification and use case," *Procedia CIRP*, vol. 72, pp. 1172–1177, Jan. 2018, <https://doi.org/10.1016/j.procir.2018.03.242>.
- [15] Song Yanwei, Zeng Guangzhou, and Pu Haitao, "Research on the Context Model of Intelligent Interaction System in the Internet of Things," *2011 IEEE International Symposium on IT in Medicine and Education*, pp. 379–382, Dec. 2011, <https://doi.org/10.1109/ITIME.2011.6132129>.
- [16] G. Chen and D. Kotz, "A Survey of Context-Aware Mobile Computing Research," Dartmouth College, Hanover, NH, USA, Computer Science Technical Report TR2000-381, 2000.
- [17] B. N. Schilit and M. M. Theimer, "Disseminating active map information to mobile hosts," *IEEE Network*, vol. 8, no. 5, pp. 22–32, Sep. 1994, <https://doi.org/10.1109/65.313011>.
- [18] T. Chaari and F. Laforest, "L'adaptation dans les systèmes d'information sensibles au contexte d'utilisation: approche et modèles," presented at the Conférence Génie Électrique et Informatique (GEI'2005), Sousse, Tunisia, Mar. 2005, pp. 56–61.
- [19] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?," *International Journal of Human-Computer Studies*, vol. 43, no. 5, pp. 907–928, Nov. 1995, <https://doi.org/10.1006/ijhc.1995.1081>.
- [20] S. R. Mashwani and S. Khuro, "The Design and Development of a Semantic File System Ontology," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2827–2833, Apr. 2018, <https://doi.org/10.48084/etasr.1898>.
- [21] C. Roche, "Terminologie et ontologie," *Langages*, no. 157, pp. 48–62, 2005.
- [22] N. F. Noy et al., "Protégé-2000: an open-source ontology-development and knowledge-acquisition environment," *AMIA Annual Symposium Proceedings*, 2003.
- [23] B. Heydari and M. Aajami, "Providing a New Model for Discovering Cloud Services Based on Ontology," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2268–2272, Dec. 2017, <https://doi.org/10.48084/etasr.1577>.

# A Dynamic Adaptive Bio-Inspired Multi-Agent System for Healthcare Task Deployment

**Hamza Reffad**

LRSD, Faculty of Technology, Technology Department, University Ferhat Abbas Sétif 1, Algeria  
hamza.reffad@univ-setif.dz (corresponding author)

**Adel Alti**

LRSD, Faculty of Sciences, Computer Science Department, University Ferhat Abbas Sétif 1, Algeria  
alti.adel@univ-setif.dz

**Ahmed Almuhrat**

Department of Management Information Systems, College of Business & Economics, Qassim University, Saudi Arabia  
a.almuhrat@qu.edu.sa

*Received: 16 December 2022 | Revised: 30 December 2022 | Accepted: 3 January 2023*

## ABSTRACT

The use of the Internet of Things (IoT) in healthcare is increasing significantly, bringing high-quality health services, but it still generates massive data with massive energy consumption. Due to the limited resources of fog servers and their impact on limiting the time needed for health data analysis tasks, the need to handle this problem in a fast way has become a necessity. To address this issue, many optimization and IoT-based approaches have been proposed. In this paper, a dynamic and adaptive healthcare service deployment controller using hybrid bio-inspired multi-agents is proposed. This method offers optimal energy costs and maintains the highest possible performance for fog cloud computing. At first, IGWO (Improved Grey Wolf Optimization) is used to initialize the deployment process using the nearest available fog servers. Then, an efficient energy-saving task deployment was achieved through Particle Swarm Optimization (PSO) to reduce energy consumption, increase rewards across multiple fog servers, and improve task deployment. Finally, to ensure continuous control of underloaded and overloaded servers, the neighborhood multi-agent coordination model is developed to manage healthcare services between the fog servers. The developed approach is implemented in the iFogSim simulator and various evaluation metrics are used to evaluate the effectiveness of the suggested approach. The simulation outcome proved that the suggested technique provides has better performance than other existing approaches.

**Keywords-***IoT; multi-agent; energy consumption; PSO; grey wolf optimization; fog-cloud*

## I. INTRODUCTION

We live in the era of the digital healthcare revolution. Every day, healthcare organizations generate billions of gigabytes of data from various sources of information such as medical sensors (IoT), social networks, telemedicine, etc. For this reason, data production and energy consumption are growing dramatically [1]. For example, a smart hospital consists of different actors (i.e. physicians, patients, and nurses) with various smart devices (i.e. smartphones, tablets), exchanging different health-related data (temperature and humidity levels, user glucose, user movement state, etc.), analysis results, and technical reports throughout various health services. Due to the explosion of services and massive energy consumption, the need to handle a large number of healthcare services and minimize their energy consumption increases sharply. Nowadays, most developers focus on improving healthcare

task scheduling with optimum cost and reduced execution time. Interesting service configurations have been produced, satisfying the user requirements, but without considering the change of their needs and the incoming loads of applications on the limited resources of devices.

On the other hand, a great deal of data is collected through many different devices (wearables, medical/vital monitors, smartphones, etc.) and their heterogeneity is related to the type of data (text, video, image, sound) or the type of service such as healthcare services (glucose level, heart beating, etc.), spatiotemporal services (location, time, etc.), and social services (Facebook, Twitter, etc.). All these factors make service management more complex, and their placement is certainly very difficult [2]. Thus, many researchers work on service placement through a deployment component called service controller. They aim to deploy service on external data

centers, build deployment scripts, and dynamically deploy it to the desired data centers [3].

Cloud computing is an interesting solution to efficiently deploy large candidate services. It allows interconnecting and integrating IoT devices and fogs according to the current services and users' needs. Besides, it is essential to know that the deployment of service is not a standardized solution. It can vary according to specific needs [4]. For this reason, service controllers must use advanced and intelligent techniques to deploy large candidates of services. In recent years, several approaches for efficient service deployment have been proposed and are generally classified into 3 main categories [5-19]: (1) green IoT approaches [5-9], (2) optimization approaches [11-14], and (3) hybrid approaches combining mobile devices, fog, and cloud optimization techniques [15-19]. Authors in [8] presented a new model called Health-Fog based on deep learning for saving energy on the fog and the cloud while the data are processed and transmitted from the fog to the cloud. Authors in [15] introduced a microservice-based system for Industrial Internet of Things (IIoT) in distributed fog cloud environments. Authors in [19] introduced a new priority traffic-conscious task placement algorithm that aimed to reduce the energy use of incoming IoT traffic servers. These techniques do not react to frequent moving scenarios of the users and none of them could get an optimum consensus between energy saving, processing load, computational cost, and fast waiting time. The most striking case is the health mobile applications where the mobility causes them to lose valuable time in providing accurate diagnosis results to their patients. There is a high computing required to select the right server to deploy services dynamically.

In the current work, a novel bio-inspired adaptive and dynamic multi-agent system for providing optimal deployment of healthcare services and applications under context changes (e.g. user's location, workload, bandwidth, reduced total time of execution) is presented. It enables much less deployment cost of health services in an optimal and context-aware fog cloud design and ensures effective solutions to deploy complex services by subdividing them into smaller ones. The main contributions of the current paper are:

- Providing an optimal deployment map for incoming loads of healthcare applications using a hybrid bio-inspired system by combining the Improved Grey Wolf Optimization (IGWO) and Particle Swarm Optimization (PSO) in IoT fog-cloud systems.
- Developing an adaptive and dynamic multi-agent-based controller that offers redeployment map for the ongoing needs of various users to the corresponding neighboring fog servers while minimizing energy consumption and waiting time, maximizing the gain of service.
- Evaluation of the proposed approach on the real healthcare data set of Qassim hospital in terms of energy consumption and execution time using the iFogSim simulator.

## II. THE CASE STUDY: QASSIM SMART HOSPITAL

Qassim Hospital consists of many floors. Each floor consists of many patient rooms and a fog server. Each room is equipped with sensors (e.g. IP cameras). Any patient can be equipped with biosensors (glucose meter, heart tracking, smart bracelet, and GPS) that monitor medical data. The monitored information covers patient identifier, glucose level, temperature, patient location, etc. The monitored information is collected in a gateway. The aggregated context health data of the patient in a given room are transmitted to the fog server and the cloud controller for data analysis and storage. The local fog server triggers an immediate response to the device that gives an order to release some quantity of the drug in the body of a patient. The IoT healthcare application consists of connected IoT and various health tasks. Each task is defined by arrival time, priority, deadline, service time, and requested resources. Connected things are constantly consuming energy and sending huge data flows. This leads to low latency and consumer pressure on the amount of data flow on the fog and the Internet. Whether for patient healthcare, examinations or treatments, or hospitalization, an optimized task deployment approach of data collected by increasingly smart sensors can shed new light, improve flows, and reduce energy consumption. This case study raises several needs: (1) ensuring an optimal deployment map for incoming loads of healthcare applications and (2) determining their (re-) deployment into fog-cloud systems while reducing energy consumption and minimizing waiting time.

## III. THE PROPOSED APPROACH

To develop an efficient and effective task deployment approach that can provide minimum energy costs and maximum benefits for users (i.e. patients, medical personnel, and administrators), we propose a new adaptive and dynamic system based on bio-inspired and cooperative agents with capacities like mobility, scalability, and efficiency. Figure 1 shows an overview of the proposed approach. Our goal is to ensure the coordination between the different agents of the healthcare system to determine the nearest best fog server between all feasible servers and select the server with the lowest processing load, waiting tasks, and energy consumption.

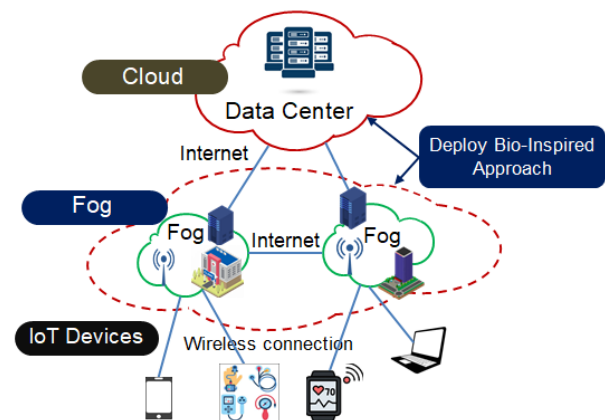


Fig. 1. An overview of the proposed approach.

The proposed approach consists of 3 static agents (Interact Agent, Deployer Agent, and Execution Agent) and two mobile agents (Monitor Agent and Bio-Research Agent) that interact and cooperate for realizing the deployment process. When a user sends a request, it consists of an application deployment to Interact Agent who plays the role of a mediator. Bio-Research Agent is a mobile bio-inspired agent that is able to migrate to fog servers to look for the requested deployment application. The Bio-Research Agent searches for local fog servers able to make partial or full deployment of the requested application. The Bio-Research Agent gives all the possible solutions and sends its result to the Deployer Agent. The Deployer Agent selects the best solution (a single fog server or a set of fog servers). Various politics based on weights (user's location, CPU load, energy-saving, high bandwidth) are used. Once the deployment map has been generated and the application has been deployed, the deployment results are sent back to the user. If the Bio-Research Agent does not find any solution on local fogs, it sends requests to the Deployer Agent to deploy tasks involved in the application on cloud servers. The Monitor Agent is responsible for handling the environment context during the execution of tasks such as changes of users' needs and incoming new urgent tasks. If it is about a single fog server, it is going to move and if it is about more than one fog servers, it clones itself, moves towards one fog server, and sends its clones to the other fog servers. The context execution parameters (CPU load, current bandwidth) will be transmitted to the Deployer Agent. The Deployer Agent selects the best fog server and generates the re-deployment map.

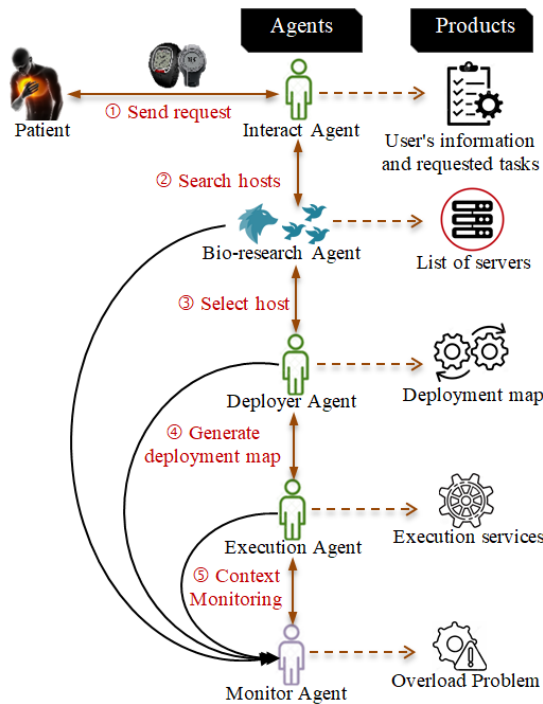


Fig. 2. The multi-agent functional model of AD-GWOPSO.

#### A. Problem Formulation

The problem consists of finding out an optimal deployment of IoT applications on IoT-Fog-Cloud infrastructure for

minimizing energy consumption while maximizing the benefits of users (i.e. response time, reliability, and security). We consider three hierarchical levels of fog infrastructure  $\mathcal{I}$  consisting of  $\mathcal{N}$  physical nodes. The IoT layer (level 0) regroups a set of  $\mathcal{T}$  smart devices, sensors, actuators, mobile devices, and edge devices. The fog computing layer (level 1) is a set of  $\mathcal{F}$  fog servers and gateways. The cloud computing layer (level 2) consists of a set of  $\mathcal{C}$  large-scale data centers.

The physical network is denoted as  $\mathcal{G}_f = (\mathcal{N}, \mathcal{E})$  which consists of a set of nodes (i.e. servers)  $\mathcal{N}$  and a set of edges  $\mathcal{E}$ . Each edge  $e_{i,j}$  represents a link between a server  $i$  and a server  $j$ . Each node  $s_i$  ( $1 \leq i \leq \mathcal{N}$ ) is localized by a position  $(x_{s_i}, y_{s_i})$ . A node  $s_i$  has a queue  $q_i$  defined by the occupation rate  $\rho_i$  and maximum size  $max\_q_i$ . Each node  $s_i$  has a current workload  $W_{s_i}^{CPU}$ .  $E_{s_i}^{CPUa}$  denotes the power consumed by the server  $s_i$  in the active state and  $E_{s_i}^{CPUidle}$  denotes the power consumed in the idle state.  $BW_{ij}$  represents the available bandwidth and  $TP_{ij}$  the data transferred (throughput) between servers  $i$  and  $j$ . The energy bandwidth consumed between servers  $i$  and  $j$  in their active state is  $E_{ij}^{BWa}$ , and the energy bandwidth consumed between  $i$  and  $j$  in the idle state is  $E_{ij}^{BWidle}$ . To solve the problem, we use a four step approach. At first, we sort all tasks by priority. A task is designed with a name and its priority value. Then, we initiate and sort all possible solutions (i.e. servers' nodes) by their weights based on the user's location and resource availability. Thirdly, we launch a bio-inspired search algorithm that gives all possible server chains providing the deployment. The best solution is based on the weight (CPU charge, energy-saving, low bandwidth, and security) of each deployment. Finally, we ensure an automated task transfer between neighboring nodes for incoming loads of applications. The problem is formalized as follows:

- The set  $U = \{u_1, u_2, \dots, u_M\}$  of  $M$  users. Each user  $U_i$  requests the deployment of the application  $a_i$  and broadcasts his current location is  $(x_{u_i}, y_{u_i})$ .
- Each task  $t_i$  ( $1 \leq j \leq M$ ) can be deployed to a device  $S_k$  ( $j = 1, 2, \dots, N$ ) by the projection  $\mathcal{P} : T \times U \rightarrow S$ . This projection is called the deployment map.
- $TT_{s_i}^{t_j}$  is the time taken by a server  $s_i$  to perform a task  $t_j$ .

Equation (1) calculates the Euclidian distance separating the person's location from  $m$  servers of deployment map  $\mathcal{P}$ :

$$Dist(\mathcal{P}) = \sum_{i=1}^m \sqrt{(x_u - x_{s_i})^2 + (y_u - y_{s_i})^2} \quad (1)$$

where  $x_u, y_u$  are the current location coordinates of the user  $u$  and  $x_{s_i}, y_{s_i}$  are the coordinates of the fog server  $s_i$ . Fogs with the shortest distance closer to a person's current device tend to be selected as best fogs by minimizing  $Dist(\mathcal{P})$ . The best fog has the smallest distance from the person. Equation (2) calculates the total consumed computing energy  $E_{CPU}$  of the deployment map:

$$E_{CPU}(\mathcal{P}) = \sum_{s_i \in \mathcal{X}} W_{s_i}^{CPU} \times E_{s_i}^{CPU_a} + (1 - W_{s_i}^{CPU}) \times E_{s_i}^{CPU_{idle}} \quad (2)$$

By reducing  $E_{CPU}$ , lower computing energy fogs tend to be chosen as the best deployment maps. Equation (3) calculates the total consumed network energy to transfer  $r$  tasks between fogs of deployment map  $\mathcal{P}$ :

$$E_{BW}(\mathcal{P}) = \sum_{(i,j) \in \mathcal{E}} TP_{ij} \times E_{ij}^{BW_a} + (BW_{ij} - TP_{ij}) \times E_{ij}^{BW_{idle}} \quad (3)$$

Equation (4) calculates the total network energy consumption, the total CPU power consumption applied to the fog servers for the task requests on deployment map:

$$f_{energy} = E_{CPU}(\mathcal{P}) + E_{BW}(\mathcal{P}) \quad (4)$$

Equation (5) is used to minimize the waiting time of the performed task.

$$f_{wt} = T_{wait} = \sum_{i=1}^m (\rho_i) \quad (5)$$

where  $\rho_i$  is the occupancy rate of fog  $i$ . Equation (6) is used to maximize the reward time of the servers:

$$f_{gain} = G_{rewards} = \sum_{i=1}^m rewards \times e^{-\gamma \times TT_{s_i}^t} \quad (6)$$

where  $\gamma$  is the discount rate, ranging between 0 and 1.

Our goal is to establish an objective function ( $f$ ) such that: (1) the total energy is minimized, (2) the global occupation rate of a person's tasks is minimized, and (3) the global gain of the fog server is maximized. From such a perspective, we define the objective function as follows:

$$f = \frac{w_{gain} \times f_{gain}}{w_{energy} \times f_{energy} + w_{time} \times f_{wt}} \quad (7)$$

where  $w_{energy}$ ,  $w_{time}$ , and  $w_{gain}$  are the weights associated with the objectives energy, time, and gain, respectively. All are normalized in  $[0, 1]$ . The higher the weight value, the more important this criterion is. The objective function is customizable according to the different customer needs.

### B. The Dynamic Adaptive Service Placement Algorithm

- **Solutions coding:** Each solution is represented by an array of  $n$  fog servers.

- **Generate the initial population based on IGWO:** GWO [20] is the most well-known optimization approach, and is used to generate the initial population (i.e. list of fog servers). The wolf agent moves towards the nearest server and provides information on server CPU load, available storage, and bandwidth communication. The best list of servers is assigned a high number of particles according to its objective functions whereas a good server has a higher fitness value than a bad one. We build the vector  $W = [W_1, \dots, W_n]$ , where  $W_i = \frac{U_i^r}{\sum_{i=1}^n U_i^r}$  is used to allocate

particles for each server and  $U_i^r$  is the utility of the given fog server  $i$  based on  $r$  attributes that describe the fog. The number of particles for all services is the total number of particles multiplied by their normalized weight to better valorize the suitable server and the fitness of each particle is

evaluated according to (7) of their related energy, time, and gain functions.

- **Fitness evaluation:** From the initial particles' position that defines the service placement map, a particle changes iteratively from one map to another with respect to velocity. The velocity is defined between two maps: the local best map ( $P_{Best}$ ) and the global best map ( $G_{Best}$ ). Fog servers with the best score (lowest energy consumption, lowest occupation rate, and highest rewards by server) are assigned a high-priority value.
- **Optimal deployment map generation.** The system generates an optimal deployment map to those selected fog servers.
- **Copy deployment map and service execution.** The generated deployment map in the previous phase will be copied to fog servers and deploy services.

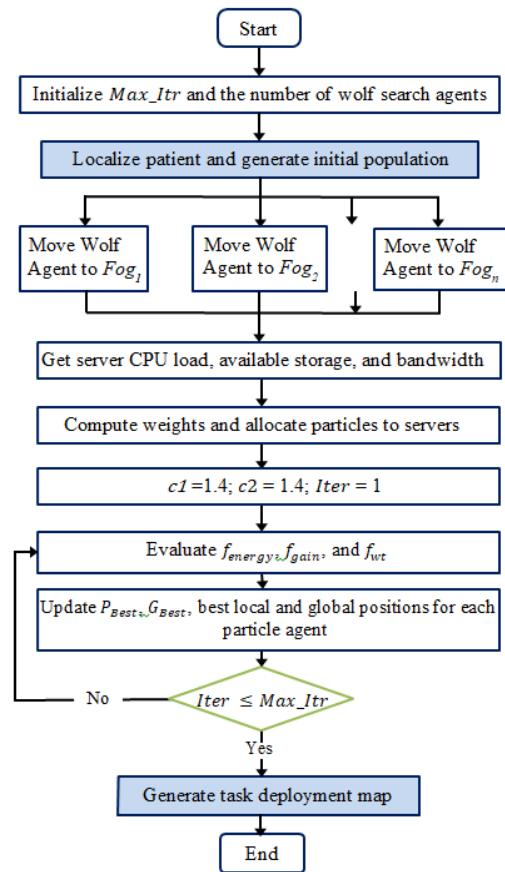


Fig. 3. Multi-agent functional model of the dynamic adaptive IGWO-PSO.

When unpredicted events occur, such as patient mobility or low latency on the fog server, the service controller will be switched to the adaptive mode, and the neighboring multi-agent technique is employed to tackle such events. Each service controller agent  $ag_{S_i}$  has a set of neighboring agents  $ag_{S_i^j}$ . The agent  $ag_{S_i}$  of server  $S_i$  selects the best neighbors



depending on occupancy rate, processing load, and transmission consumption energy. It checks the occupation rate  $\rho$  from a queue of the fog server  $S_i$ , and if it is greater than a threshold  $T_h$  then selects the tasks that are going to be transferred. The selection of a task is based on the type of task and the requested resources. A deployment request is created and sent to every neighbor. Everyone can accept or reject the request according to its occupation rate and processing load. The list of neighbors will be sorted by their queue size, available computing resources, and bandwidth. The controller transfers the requested tasks to the top ranked neighbors.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

The developed approach was implemented in Eclipse using iFogSim [21]. The number of rejected/waited tasks is used to assess the efficiency of the proposed approach. Various evaluation metrics were considered, including the number of rejected tasks, the number of waiting tasks, and the average makespan. The number of fog servers is 10 and the size of the waiting queue is 10. We simulated and compared different priority strategies: (1) Shortest Arrival Time (P-SAT) which selects the task with high priority and the shortest arrival time and (2) Least Requested Resources (P-LRR) that selects the task with high priority and the fewest requested resources.

##### A. Evaluation of Average Makespan

Figure 4 shows the average makespan time (ms) of the existing approaches (P-SPT, P-LRR), and the proposed approach with varying number of tasks. We can see that the proposed approach demonstrated optimum makespan compared to P-SPT and P-LPR due to the optimal deployment time using IGWO-PSO with a good energy saving while exploiting performance metrics.

##### B. Evaluating the Number of Rejected and Waited Tasks

Figure 5 shows the performance of the proposed approach with varying number of tasks from 87 to 2697. The number of rejected and waited tasks is computed for different numbers of tasks and different algorithms (P-SAT, P-LRR and proposed approach). The proposed approach obtained the optimum number of rejected tasks and the optimum number of waiting tasks for all the number of task values, due to the pre-selection

of the best servers using IGWO and applying an adaptive task transfer to neighboring fog servers with a good server utility while exploiting performance metrics.

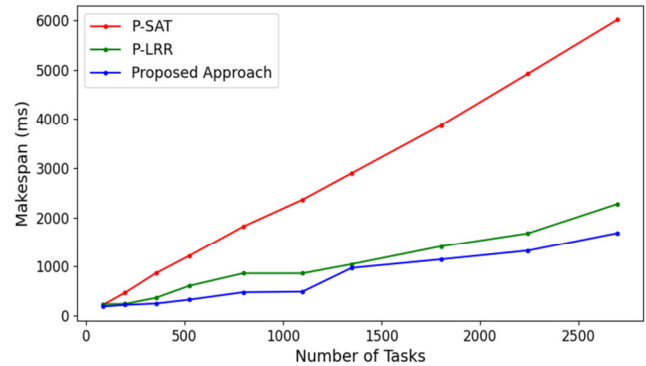


Fig. 4. The average makespan of the proposed and existing approaches.

##### C. Evaluation of Energy Consumption

To evaluate the energy consumption of our approach, we measure the average bandwidth/processing energy consumption of the proposed approach in cloud + fog and fog only. The proposed approach in fog environments is better in terms of bandwidth and processing energy consumption compared to the same approach in fog + cloud environments (Figure 6).

##### D. Discussion

In this paper, a dynamic and fast-to-deploy DS-IGWOPSO is proposed, one that ensures dynamic task deployment and applies a transfer to neighboring fog servers. Compared to other popular deployment algorithms (P-SPT and P-LRR), DS-IGWOPSO demanded less bandwidth and processing energy to deploy tasks. It also reduces the makespan average on varying number of tasks. To have IGWOPSO with its best performance, the neighbor selection should be based on the optimization of occupation rate and waiting time. However, it still depends on the number of requests to ensure optimal computational time.

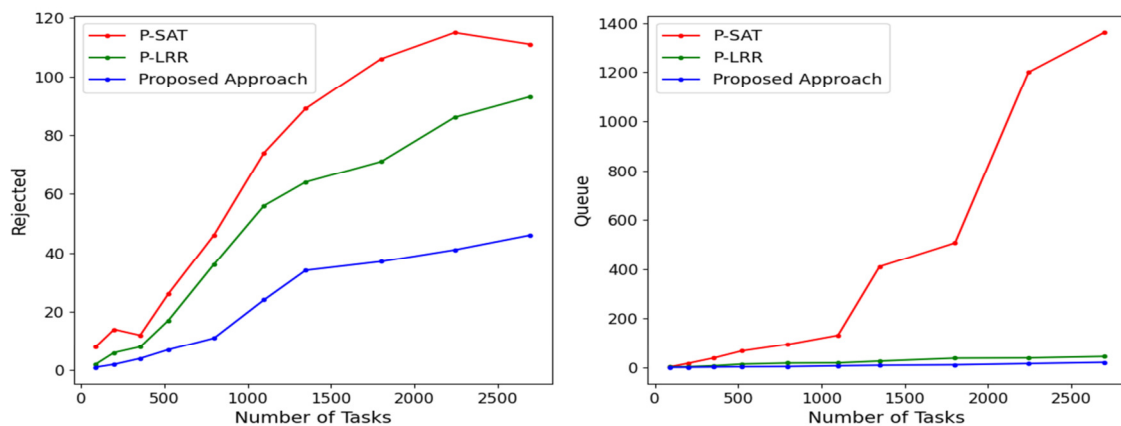


Fig. 5. Performance comparison.



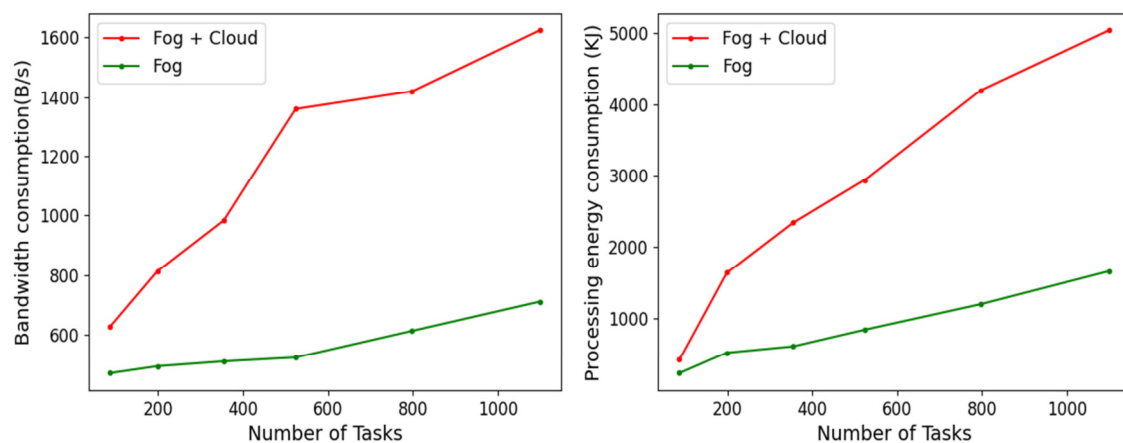


Fig. 6. Bandwidth and processing energy consumption of the proposed approach in fog and fog+cloud.

## V. CONCLUSION

An adaptive and dynamic bio-inspired multi-agent approach for deploying healthcare tasks in IoT-fog-cloud environments was developed and presented in this paper. Hybrid particle swarm and grey wolf optimizers were utilized for the determination of a set of optimal hosts when deploying health services to optimize deployment cost and improve response time at run-time. The approach is based on neighboring agents for ensuring the continued functioning of health services by using both service priority and high servers' resource of availability. The approach was evaluated by deploying different types of tasks. The evaluation of the obtained results was achieved through well-known metrics such as waiting tasks and average makespan time. The results show that the proposed approach is economic in terms of energy consumption. The obtained results are very encouraging and confirm the efficiency of the approach.

## REFERENCES

- [1] S. Kallam, R. Patan, T. V. Ramana, and A. H. Gandomi, "Linear Weighted Regression and Energy-Aware Greedy Scheduling for Heterogeneous Big Data," *Electronics*, vol. 10, no. 5, Jan. 2021, Art. no. 554, <https://doi.org/10.3390/electronics10050554>.
- [2] H. H. A. Valera, M. Dalmau, P. Roose, J. Larracoechea, and C. Herzog, "DRACeo: A smart simulator to deploy energy saving methods in microservices based networks," in *2020 IEEE 29th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, Bayonne, France, Sep. 2020, pp. 94–99, <https://doi.org/10.1109/WETICE49692.2020.00026>.
- [3] A. Adel, S. Laborie, and P. Roose, "Towards a Context-Aware Service and Quality Multimedia Adaptation for Healthcare Applications," in *International Conference on Digital Information Processing, E-Business and Cloud Computing*, 2013.
- [4] S. Saxena and D. Saxena, "Green Cloud Computing Architecture with Efficient Resource Allocation System," *International Journal of Trend in Research and Development*, vol. 3, no. 6, pp. 248–251, 2016.
- [5] H. H. A. Valera, P. Roose, M. Dalmau, C. Herzog, and K. Respicio, "KaliGreen: A distributed Scheduler for Energy Saving," *Procedia Computer Science*, vol. 141, pp. 223–230, Jan. 2018, <https://doi.org/10.1016/j.procs.2018.10.172>.
- [6] A. S. H. Abdul-Qawy, N. M. S. Almurisi, and S. Tadisetty, "Classification of Energy Saving Techniques for IoT-based Heterogeneous Wireless Nodes," *Procedia Computer Science*, vol. 171, pp. 2590–2599, Jan. 2020, <https://doi.org/10.1016/j.procs.2020.04.281>.
- [7] S. Tuli *et al.*, "HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments," *Future Generation Computer Systems*, vol. 104, pp. 187–200, Mar. 2020, <https://doi.org/10.1016/j.future.2019.10.043>.
- [8] K. Haseeb, N. Islam, Y. Javed, and U. Tariq, "A Lightweight Secure and Energy-Efficient Fog-Based Routing Protocol for Constraint Sensors Network," *Energies*, vol. 14, no. 1, Jan. 2021, Art. no. 89, <https://doi.org/10.3390/en14010089>.
- [9] A. A. Brincat, F. Pacifici, S. Martinaglia, and F. Mazzola, "The Internet of Things for Intelligent Transportation Systems in Real Smart Cities Scenarios," in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, Limerick, Ireland, Apr. 2019, pp. 128–132, <https://doi.org/10.1109/WF-IoT.2019.8767247>.
- [10] M. N. Hasan, R. N. Toma, A.-A. Nahid, M. M. M. Islam, and J.-M. Kim, "Electricity Theft Detection in Smart Grid Systems: A CNN-LSTM Based Approach," *Energies*, vol. 12, no. 17, Jan. 2019, Art. no. 3310, <https://doi.org/10.3390/en12173310>.
- [11] N. Sasikaladevi and L. Arockiam, "Genetic Approach for Service Selection problem in Composite Web Service," *International Journal of Computer Applications*, vol. 44, no. 4, pp. 22–29, Apr. 2012, <https://doi.org/10.5120/6252-8396>.
- [12] W. Song, W. Ma, and Y. Qiao, "Particle swarm optimization algorithm with environmental factors for clustering analysis," *Soft Computing*, vol. 21, no. 2, pp. 283–293, Jan. 2017, <https://doi.org/10.1007/s00500-014-1458-7>.
- [13] F. Choukairy, "Optimization of energy consumption in a Cloud environment," Ph.D. dissertation, Laval University, Québec, QC, Canada, 2018.
- [14] F. H. Khoso, A. Lakhani, A. A. Arain, M. A. Soomro, S. Z. Nizamani, and K. Kanwar, "A Microservice-Based System for Industrial Internet of Things in Fog-Cloud Assisted Network," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 7029–7032, Apr. 2021, <https://doi.org/10.48084/etasr.4077>.
- [15] S. Omer, S. Azizi, M. Shojafar, and R. Tafazolli, "A priority, power and traffic-aware virtual machine placement of IoT applications in cloud data centers," *Journal of Systems Architecture*, vol. 115, May 2021, Art. no. 101996, <https://doi.org/10.1016/j.sysarc.2021.101996>.
- [16] D. Mills, S. Sivarajah, T. L. Scholten, and R. Duncan, "Application-Motivated, Holistic Benchmarking of a Full Quantum Computing Stack," *Quantum*, vol. 5, Mar. 2021, Art. no. 415, <https://doi.org/10.22331/q-2021-03-22-415>.
- [17] S. F. Issawi, A. A. Halees, and M. Radi, "An Efficient Adaptive Load Balancing Algorithm for Cloud Computing Under Bursty Workloads," *Engineering, Technology & Applied Science Research*, vol. 5, no. 3, pp. 795–800, Jun. 2015, <https://doi.org/10.48084/etasr.554>.
- [18] M. E. Hassan and A. Yousif, "Cloud Job Scheduling with Ions Motion Optimization Algorithm," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 7029–7032, Apr. 2021, <https://doi.org/10.48084/etasr.4077>.

- Research*, vol. 10, no. 2, pp. 5459–5465, Apr. 2020, <https://doi.org/10.48084/etasr.3408>.
- [19] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, Mar. 2014, <https://doi.org/10.1016/j.advengsoft.2013.12.007>.
- [20] H. Gupta, A. Vahid Dastjerdi, S. K. Ghosh, and R. Buyya, "iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, Edge and Fog computing environments," *Software: Practice and Experience*, vol. 47, no. 9, pp. 1275–1296, 2017, <https://doi.org/10.1002/spe.2509>.

# The Impact of Operational Scenarios and Concrete Aging Factor on the Freeboard Height of an Irrigation Canal

**Masoud Kazem**

Faculty of Civil Engineering, Iran University of Science and Technology, Iran  
masoud\_kazem@cmps2.iust.ac.ir

**Mohammad Nazari-Sharabian**

Department of Mathematics, Engineering, and Computer Science, West Virginia State University, USA  
m.nazari@wvstateu.edu  
(corresponding author)

**Hossein Afzalimehr**

Faculty of Civil Engineering, Iran University of Science and Technology, Iran  
hafzali@iust.ac.ir

**Nader Darban**

Faculty of Civil Engineering, Iran University of Science and Technology, Iran  
n\_darban@cmps2.iust.ac.ir

**Moses Karakouzian**

Department of Civil and Environmental Engineering and Construction, University of Nevada, USA  
mkar@unlv.nevada.edu

*Received: 21 December 2022 | Revised: 8 January 2023 | Accepted: 11 January 2023*

## ABSTRACT

The prediction of operational freeboard in irrigation canals is a complicated issue, particularly when the cumulative effects of time-dependent factors, such as maloperation and concrete aging, are considered. While most classic approaches consider a fixed freeboard due to uniform flow as a fundamental assumption. This study proposes a flowchart considering the effects of nonuniform flow to determine the adequacy of the freeboard of irrigation canals under different conditions, including time-dependent and operational scenarios. The results of this method indicated that the freeboard values obtained by classical methods may not be sufficiently reliable in providing the appropriate level of performance in the operating conditions of agricultural canals. Regarding the case study of this paper, an irrigation canal in Kurdistan-Iran, the results showed that the formation of the M1 profile is the most critical scenario and the initial freeboard must be extended by 20cm at a distance of about 2.3Km at the end of the canal towards upstream.

*Keywords-freeboard; concrete aging; irrigation canal; operational scenarios*

## I. INTRODUCTION

Despite the prevalent traditional methods that are widely used in the initial design phase of irrigation canals, water delivery performance and economically optimal design of canal sections are still hot topics in current engineering works [1-4]. The geometry of the canal section plays a key role in the total cost of an irrigation system. In [5], an optimal parabolic section

with a freeboard was proposed to provide optimal flow capacity. However, mathematically designed sections are generally difficult to construct. In addition to the need for specific machinery, such parabolic sections may not perform well in geotechnical stability analysis without a thick concrete lining, especially in large canals. In [6], a power-law section was proposed that had similar disadvantages to the parabolic section. Therefore, traditional trapezoidal sections with

concrete lining are still the most common sections in many irrigation systems. A trapezoidal section is an optimal choice in the design phase, not only compared to curved sections, but to other simple geometries. In [7], triangular, rectangular, and trapezoidal sections were compared, concluding that the optimal trapezoidal section has the lowest seepage loss and cross-sectional area among the three optimal sections. Consequently, practical approaches to refine the design procedure of trapezoidal sections can help engineers to consider various affecting factors which are usually ignored in the initial design phases.

In general, canal lining and banks are extended above a canal's normal water surface as a safety measure to protect the conveyance system from overtopping. Freeboard in a canal provides a water surface higher than normal, which may be caused by sedimentation in the canal, temporary maloperation of the canal system, excess flows caused by storm runoff entering the canal through drain inlets, additional water depth resulting from a higher friction coefficient than used in the design, and waves produced by wind or surges which accompany sudden changes in flow. In the initial design phase, the freeboard height is generally recommended to be 1/6 of the total depth in uniform flow conditions [8]. Figure 1 illustrates the freeboard of an irrigation canal. However, as a prevalent issue, overestimating or underestimating the freeboard can increase construction and operational costs [9, 10]. An overestimated freeboard can also increase the water loss in an irrigation canal, particularly in the compacted earth lining [11, 12]. On the other hand, underestimation of the freeboard can lead to difficulties in the operation or even to catastrophes in some cases [13-15]. The efficiency of the freeboard critically depends on the precise estimation of canal roughness. The effect of concrete aging on Manning's roughness coefficient ( $n$ ) has always been a controversial issue. An uncertain assumption for this factor not only causes problems such as overflow, but also can disturb the controllability of an automated irrigation canal [16]. Although engineering codes recommend a standard range of 0.012-0.017 for Manning's  $n$ , there is evidence that the value could be increased to 0.023 in aged concrete-lined canals [17]. Increasing roughness can occur at a higher rate in smaller canals compared to large ones.

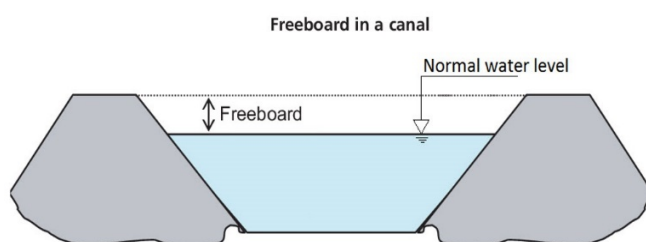


Fig. 1. Freeboard in a typical irrigation canal.

To increase the efficiency of water transport in irrigation canals, engineers offer slight modifications in the design phase that can influence construction prices and long-term maintenance costs. There is not a certain rule to guide water authorities in selecting the best option, and the selection procedure is performed case by case. On the other hand, it must

be well thought-out that in many water transport projects, such as the studied case, the initial investment and construction costs are covered by governmental budgets while the operational costs rely on the stakeholders' payments, and most of them use leased facilities and cannot handle high costs of major rehabilitation for many years. This view is also amplified when considering that this project aims to improve the economy of a relatively poor region on the aging issue and the effectiveness of the freeboard depends on the appropriate prediction of operational scenarios. Most operational canals are designed based on a normal flow regime and predefined demand scenarios. Therefore, they would transfer the design flow plus a fixed freeboard to cover up fluctuations and waves. However, in constant geometric and roughness conditions, any change in demand nodes would disturb the water surface profile. Therefore, this factor plays an important role in the fundamental hydraulic analysis that determines the crest level of emergency spillways or the installation level of automated intake gates and is rarely considered in determining the general freeboard. The operational flow regime variation is more critical in long canals with multi-demand points, because of the effects of back-water and slight steepness in long distances in addition to the multi-stakeholders policies. In [18], it was shown that the conventional freeboard design did not suit a long-distance water transfer project [18]. There are also examples of aged canals, such as the Chashma Right Bank Canal in Pakistan, where the freeboard, adequate for the normal operational flow regime, could be over-stressed under specific flow conditions [19-22]. Therefore, unforeseen operational conditions must be considered for each irrigation canal. As there is no common rule for generally addressing this issue, it should be investigated case by case.

The current study investigated the following questions:

- To what extent a constant freeboard height calculated using classic approaches is reliable under different operational and time-dependent scenarios?
- How can some modifications help engineers to evaluate the freeboard height by considering operational and time-dependent scenarios?

## II. CASE STUDY

The Garan water transport canal is part of an extensive irrigation project in the Kurdistan region, west Iran, next to the international borderline with Iraq. Having a length of 17Km, the Garan agro-industrial water transport system consists of a concrete-lined trapezoidal canal that conveys approximately 44 MCM of water annually. According to the design data, the maximum flow from the dam intake is  $7\text{m}^3/\text{s}$ . However, after supplying a pumping station, the flow decreases to  $5\text{m}^3/\text{s}$  in the MPC2 canal, as shown in Figure 2, before arriving at a reverse-siphon pipeline. The main subject of this case study was the evaluation of the adequacy of the freeboard, as there are extreme operational scenarios in emergency conditions. Figure 3 shows the studied part of the canal and the control sill at the endpoint, which warrants a minimum water level for industrial demands upstream during winter when the irrigation demand is zero. The MPC2 canal was designed later and despite the supposition for the Garan canal, where Manning's roughness is

0.015, its roughness was assumed to be 0.016 in the design phase. During the construction phase, a design review was considered to evaluate the effects of different roughness coefficients assigned to these two canals. An optional retarder sill is planned at the endpoint of the Garan canal to maintain a minimum water level during winter to supply the constant industrial demand when there is no irrigation demand and the MPC2 canal must be shut down. This obstruction will create an M1 G.V.F flow profile towards the upstream that will raise the water level. However, there is a question to what extent it would affect the upstream. In addition, particular operational scenarios were considered to evaluate the capacity of the canal with different roughness values by considering the concrete aging effects. Based on a flowchart developed, shown in Figure 4, the surface profile of the water was determined using the G.V.F equations. This profile was then compared to the canal geometry, and the freeboard was increased accordingly in critical regions.



Fig. 2. Aerial view of the Garan irrigation canal and the location of the Garan dam and the MPC2 canal.

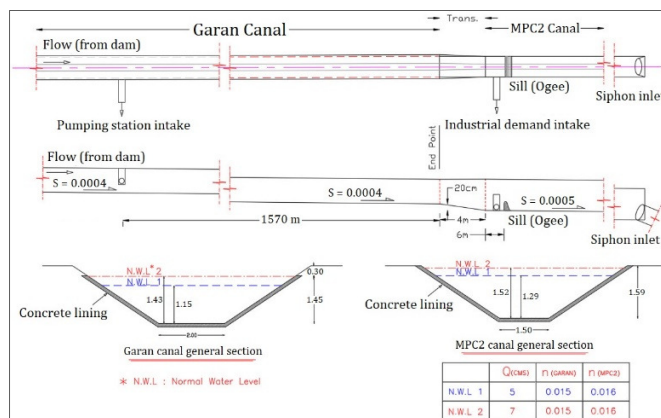


Fig. 3. The profile and cross-sections of the Garan and MPC2 irrigation canals in the attentive region.

### III. METHODOLOGY AND CALCULATIONS

This study considered the change in roughness and the shutdown/running of the pumping station as the main factors to derive the critical scenarios shown in Table I. The G.V.F equation was used to develop the water surface profile:

$$\frac{dy}{dx} = \frac{S_0 - S_f}{(1 - Fr^2)} \quad (1)$$

where  $S_0$  is the bottom slope which is positive in the downward direction,  $S_f$  is the friction slope which is positive in the downward direction,  $y$  is the water depth measured from the canal bed to the water surface,  $x$  is the longitudinal distance measured along the canal from the endpoint toward upstream, and  $Fr$  is the Froude number. The friction slope was approximated using Manning's equation:

$$S_f = \frac{n^2 V^2}{\phi R^3} \quad (2)$$

where  $S_f$  is the friction slope which is positive in the downward direction,  $n$  is Manning's roughness coefficient,  $V$  is the cross-sectional mean velocity,  $\phi$  is a constant equal to 1.49 for English units and 1.00 for SI units, and  $R$  is the hydraulic radius.

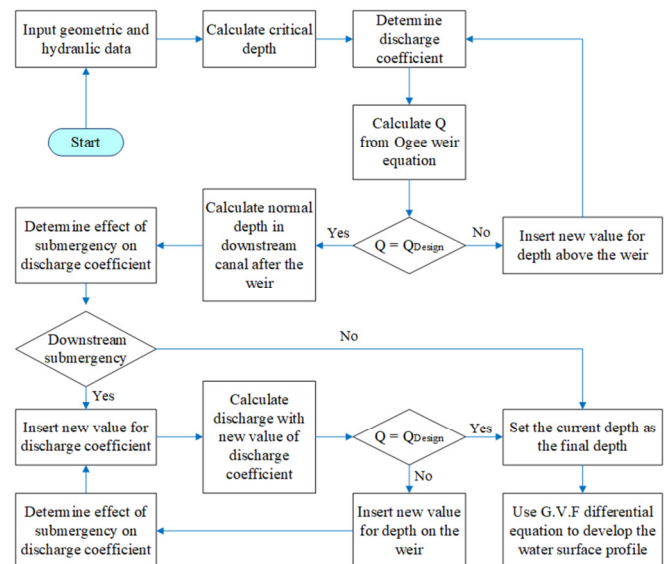


Fig. 4. Determination of the Ogee discharge coefficient and the G.V.F boundary condition flowchart.

TABLE I. OPERATION AND AGING SCENARIOS

Scenario No.	Canal roughness		Pumping station condition
	Garan	MPC2	
1	0.011	0.011	Running
2	0.014	0.014	Running
3	0.015	0.016	Running
4	0.017	0.017	Running
5	0.011	0.011	Shut Down
6	0.014	0.014	Shut Down
7	0.015	0.016	Shut Down
8	0.017	0.017	Shut Down

A variety of mathematical techniques can be applied in a traditional method to generate the results [23]. In this case, due to the effect of the Ogee weir, the boundary condition of the G.V.F equation had to be determined carefully. The effect of downstream conditions on weir function was determined by the Ogee spillway discharge coefficient diagram recommended in [8]. Consequently, a flowchart was developed to perform the



calculations considering the effects of different scenarios, as shown in Figure 4. The final goal was to evaluate the performance of the canal and its freeboard upstream of the Ogee sill. Table II presents the results and Figure 5 illustrates the water surface profiles generated by G.V.F for each scenario. According to the diagram, for the worst-case scenario (no. 8), a freeboard extension must be installed throughout the entire length of the canal. For scenarios 6 and 7, the water surface level never exceeds the berm level upstream of the pumping station. However, the desirable freeboard was not achievable for the entire length of the canal. For scenario 5, a minimum of 20cm freeboard was attainable if the berm level increased for 2,300m of the canal length from the sill, upward to the pumping station. In the other group of scenarios, where the pumping station was assumed to be operational, the current berm level was sufficient. However, for scenario 4, the aging effect would cause the maximum freeboard not to exceed 17cm. As an outcome of this calculation, the water surface profile converged to the normal water level in a shorter distance when the roughness increased. However, the final water level was higher.

TABLE II. RESULTS OF FREEBOARD EXTENSION

Scenario No.	Q (CMS)	Garan roughness	MPC2 roughness	Freeboard extension span (m)
1	5	0.011	0.011	120
2	5	0.014	0.014	180
3	5	0.015	0.016	300
4	5	0.017	0.017	960
5	7	0.011	0.011	2,300
6	7	0.014	0.014	Up to the pumping station
7	7	0.015	0.016	Up to the pumping station
8	7	0.017	0.017	The entire canal

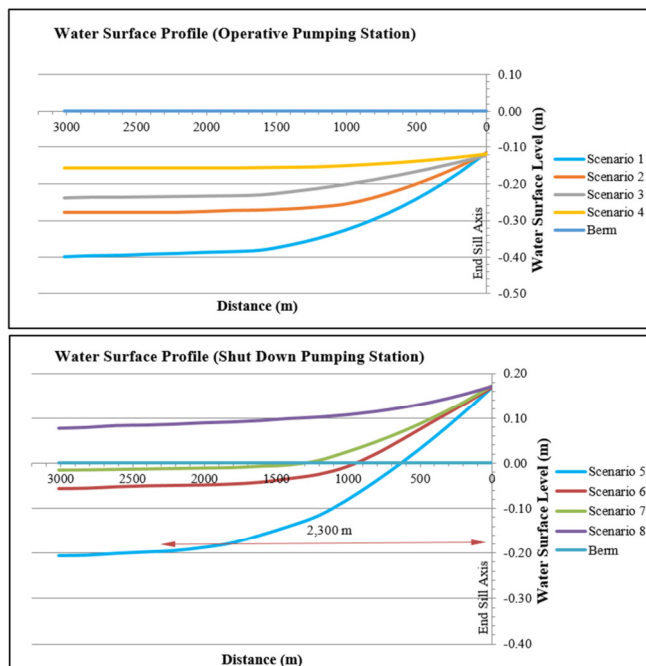


Fig. 5. G.V.F water surface profile for the Garan canal. Note scenario 5 where a freeboard extension is needed for a length of 2300m.

#### IV. DISCUSSION

Many challenging factors influence the hydraulic geometry of canals, including roughness, discharge coefficients, and the application of G.V.F. Although these factors have been used in many studies, the results show that there is no unique and theoretical method to solve the challenges of canal design. This calls for more empirical data to apply to specific sites. Any change in the estimation of the roughness coefficient influences the energy slope estimation, and any change in the discharge coefficient affects the velocity estimation. Consequently, the changes in these coefficients affect the water surface determination when applying (2). The effect of climate change (including changes in rainfall and evaporation), drought, and economic considerations force the designers to work on practical methods for canal restoration and modify the inputs in the G.V.F equation. One suitable and easy approach to canal design is to modify inputs including freeboard and hydraulic coefficients based on specific regional considerations. Figure 4 presents a practical approach to obtaining suitable results in irrigation canals.

#### V. CONCLUSIONS

Relying on the classic approaches to determine the freeboard height can lead to technical difficulties in multiple irrigation canals, particularly when there is a large gap between the demands during various seasons. In addition to the classic approaches followed in the initial design phase, operational scenarios must be considered through more sophisticated approaches. This study proposed a flowchart to determine the adequacy of the freeboard of a particular irrigation canal in different scenarios. The freeboard of the Garan canal was designed using the conventional approach that considers normal water level as the basic water surface profile. However, the results showed that there is a need for freeboard extension up to 20cm for a distance of about 2.3 kilometers from the next-added sill towards the upstream in a moderate scenario. The pessimistic scenarios indicated that a freeboard extension must be installed for the entire length of the canal. This output emphasizes the fact that the initial assumption about the freeboard and the conventional design approach must be revised. The sill and aging effects could influence the freeboard, which could be sufficient for the normal range of variation, but could be over-stressed under specific operation sets. The results of this study indicate that engineers should use the freeboard values obtained by traditional methods with more caution. In this regard, the development and use of methods such as the proposed algorithm can help engineers make the necessary revisions.

#### REFERENCES

- [1] A. R. Vatankhah, "Normal depth and wetted perimeter in general power-law channels," *Flow Measurement and Instrumentation*, vol. 64, pp. 234–241, Dec. 2018, <https://doi.org/10.1016/j.flowmeasinst.2018.11.003>.
- [2] M. M. Wilsnack, S. Yue, and M. A. Ansar, "A Canal Capacity Evaluation Program for the South Florida Water Management District," in *World Environmental and Water Resources Congress 2019*, Pittsburgh, PA, USA, May 2019, pp. 200–207, <https://doi.org/10.1061/9780784482353.019>.
- [3] M. T. Shamaa and H. A. Abdel-Gawad, "Minimum Cost Design of Irrigation Canal Sections.," *MEJ. Mansoura Engineering Journal*, vol.



- 29, no. 1, pp. 134–155, Dec. 2020, <https://doi.org/10.21608/bfemu.2020.132664>.
- [4] Y. Terefe and P. Singh, "East-bank canal water delivery performance evaluation: case study of Finchaa Sugar Estate, Ethiopia," *ISH Journal of Hydraulic Engineering*, vol. 28, no. 28:sup1, pp. 518–526, Nov. 2022, <https://doi.org/10.1080/09715010.2019.1708817>.
- [5] B. R. Chahar, N. Ahmed, and R. Godara, "Optimal Parabolic Section with Freeboard," *Journal of Indian Water Works Association*, vol. 39, no. 1, pp. 43–48, Jan. 2007.
- [6] A. S. Hussein, "Simplified Design of Hydraulically Efficient Power-Law Channels with Freeboard," *Journal of Irrigation and Drainage Engineering*, vol. 134, no. 3, pp. 380–386, Jun. 2008, [https://doi.org/10.1061/\(ASCE\)0733-9437\(2008\)134:3\(380\)](https://doi.org/10.1061/(ASCE)0733-9437(2008)134:3(380)).
- [7] P. K. Swamee, G. C. Mishra, and B. R. Chahar, "Design of Minimum Seepage Loss Canal Sections," *Journal of Irrigation and Drainage Engineering*, vol. 126, no. 1, pp. 28–32, Jan. 2000, [https://doi.org/10.1061/\(ASCE\)0733-9437\(2000\)126:1\(28\)](https://doi.org/10.1061/(ASCE)0733-9437(2000)126:1(28)).
- [8] W. Duncan, C. Huntley, J. Hokenstrom, A. Cudworth, and T. McDaniel, "Design of small dams (third edition). A water resources technical publication. Final report," Bureau of Reclamation, Denver, CO (United States), Engineering and Research Center, PB-95-176368/XAB, Dec. 1987. [Online]. Available: <https://www.osti.gov/biblio/29108>.
- [9] H. A. El-Ghandour, E. Elbeltagi, and M. E. Gabr, "Design of irrigation canals with minimum overall cost using particle swarm optimization – case study: El-Sheikh Gaber canal, north Sinai Peninsula, Egypt," *Journal of Hydroinformatics*, vol. 22, no. 5, pp. 1258–1269, Jun. 2020, <https://doi.org/10.2166/hydro.2020.199>.
- [10] S. Pourbakhshian and M. Pouraminian, "Analytical Models for Optimal Design of a Trapezoidal Composite Channel Cross-Section," *Civil and Environmental Engineering Reports*, vol. 31, no. 1, 2021, <https://doi.org/10.2478/ceer-2021-0009>.
- [11] D. A. El-Molla and M. A. El-Molla, "Reducing the conveyance losses in trapezoidal canals using compacted earth lining," *Ain Shams Engineering Journal*, vol. 12, no. 3, pp. 2453–2463, Sep. 2021, <https://doi.org/10.1016/j.asej.2021.01.018>.
- [12] P. Varjavand, S. Absalan, N. Salamati, A. Azizi, M. Goosheh, and I. Lakzadeh, "Field Investigation of Operational Management Effect on Water Losses and Sedimentation in Irrigation Channels," *Water and Soil Science (Agricultural Science)*, vol. 30, no. 2, pp. 75–89, Jan. 2020.
- [13] A. s. Shakir and N. Maqbool, "Remodelling of the Upper Chenab Canal: A Case Study from Pakistan," *Irrigation and Drainage*, vol. 60, no. 3, pp. 285–295, 2011, <https://doi.org/10.1002/ird.579>.
- [14] M. J. Reddy and S. Adarsh, "Overtopping Probability Constrained Optimal Design of Composite Channels Using Swarm Intelligence Technique," *Journal of Irrigation and Drainage Engineering*, vol. 136, no. 8, pp. 532–542, Aug. 2010, [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000217](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000217).
- [15] R. K. Bhattacharjya and M. Satish, "Flooding Probability-Based Optimal Design of Trapezoidal Open Channel Using Freeboard as a Design Variable," *Journal of Irrigation and Drainage Engineering*, vol. 134, no. 3, pp. 405–408, Jun. 2008, [https://doi.org/10.1061/\(ASCE\)0733-9437\(2008\)134:3\(405\)](https://doi.org/10.1061/(ASCE)0733-9437(2008)134:3(405)).
- [16] D. Lozano, D. Dorchie, G. Belaud, X. Litrico, and L. Mateos, "Simulation study on the influence of roughness on the downstream automatic control of an irrigation canal," *Journal of Irrigation and Drainage Engineering*, vol. 138, no. 4, pp. 285–293, 2012.
- [17] E. Akkuzu, H. B. Unal, B. S. Karatas, M. Avci, and S. Asik, "Evaluation of Irrigation Canal Maintenance according to Roughness and Active Canal Capacity Values," *Journal of Irrigation and Drainage Engineering*, vol. 134, no. 1, pp. 60–66, Feb. 2008, [https://doi.org/10.1061/\(ASCE\)0733-9437\(2008\)134:1\(60\)](https://doi.org/10.1061/(ASCE)0733-9437(2008)134:1(60)).
- [18] Z. Y. Wu, W. G. Xiao, and L. M. Han, "Research on Canal Freeboard Design Criteria," *South-to-North Water Transfers and Water Science & Technology*, no. 6, pp. 359–361, 2009.
- [19] Z. Habib, S. K. Shah, M. K. Ullah, A. Vabre, D. A. Mobin, and A. Sophyani, "Hydraulic simulations to evaluate and predict design and operation of the Chasma Right Bank Canal," International Irrigation Management Institute, Mexico city, Mexico, 1999.
- [20] A. Liaghat, A. Adib, and H. R. Gafouri, "Evaluating the Effects of Dam Construction on the Morphological Changes of Downstream Meandering Rivers (Case Study: Karkheh River)," *Engineering, Technology & Applied Science Research*, vol. 7, no. 2, pp. 1515–1522, Apr. 2017, <https://doi.org/10.48084/etasr.969>.
- [21] A. Bashir, R. Chaudhry, A. L. Qureshi, U. Memon, and N. Bheel, "Understanding the Seepage Behavior of Nai Gaj Dam through Numerical Analysis," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8085–8089, Feb. 2022, <https://doi.org/10.48084/etasr.4560>.
- [22] A. A. Mahessar *et al.*, "Sediment Transport Dynamics in the Upper Nara Canal Off-taking from Sukkur Barrage of Indus River," *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6563–6569, Dec. 2020, <https://doi.org/10.48084/etasr.3924>.
- [23] C. D. Jan, "Basic Equations for the Gradually-Varied Flow," in *Gradually-varied Flow Profiles in Open Channels: Analytical Solutions by Using Gaussian Hypergeometric Function*, C. D. Jan, Ed. Berlin, Heidelberg: Springer, 2014, pp. 1–20.

# Groundwater Quality Assessment and Health Risks from Fluoride in Jamui, Bihar

**Krishna Neeti**

Department of Civil Engineering, National Institute of Technology Patna, India  
krishnan.phd19.ce@nitp.ac.in

**Reena Singh**

Department of Civil Engineering, National Institute of Technology Patna, India  
reena@nitp.ac.in  
(corresponding author)

Received: 17 December 2022 | Revised: 9 January 2023 | Accepted: 11 January 2023

## ABSTRACT

This study aimed to determine the fluoride concentration in drinking water and assess its health risks by analyzing 12 physicochemical parameters, including fluoride, pH, EC, TDS, chloride, carbonate and bicarbonate (alkalinity), sulfate, nitrate, calcium, and magnesium hardness. Correlation analysis, WQI, and HRA were used to determine whether groundwater in the study area was suitable for drinking. Correlation analysis showed that fluoride was negatively correlated with EC (-0.649),  $\text{CO}_3^{2-}$  (-0.855) and positively correlated with  $\text{Mg}^{2+}$  (+0.559). All water samples exceeded the permissible fluoride limits according to BIS (IS 10500:2012). The WQI for all water samples was more than 100, indicating that the water was not suitable for drinking. Health risk assessment was also performed to determine the risks of non-carcinogenic diseases. The Hazard Index (HI) was determined as greater than 1. The HI ranged from 1.275 to 3.346 for adult men, 1.431 to 3.954 for adult women, and 1.986 to 5.4864 for children. Fluoride concentrations in drinking water pose a greater health risk to children than to adults. The fluoride level in drinking water is an essential parameter that must be monitored as a preventive measure against dental and skeletal fluorosis.

**Keywords-groundwater; fluoride; correlation analysis; health risks**

## I. INTRODUCTION

A healthy lifestyle requires clean and safe water, as poor water quality causes health problems and social difficulties. In certain parts of the world, access to clean drinking water is difficult due to economic, environmental, and budgetary constraints [1, 2]. Many countries prioritize providing safe drinking water, but undeveloped and many developing countries cannot meet regulatory drinking water requirements [3-5]. Clean water services are not even accessible to 783 million people, according to the World Health Organization (WHO). More than half of humanity is expected to suffer from a drinking water crisis by 2025 [6]. Drinkable freshwater consumed is more often derived from groundwater than surface water due to the higher microbiological activity of the latter [7]. However, several geological processes increase groundwater contamination [8]. It is common for rural water sources to be contaminated by illegally discharged industrial, agricultural, and urban effluents. These contaminated water bodies harm humans and aquatic life [9, 10], and several efforts are being made to improve the quality of drinking water. The burden of waterborne diseases from consuming contaminated

water is measurable and considerable [11], and it negatively affects society on an economic level [12].

Many studies have shown that fluoride, nitrate, and arsenic ions adversely affect living organisms [13]. The presence of excessive levels of fluoride in drinking water (>1mg/L) has raised concerns about public health and has been associated with several harmful health effects [4, 14-18]. When people drink excessive amounts of fluoride in water, they are more likely to have fluorosis, which can range from dental (1.5-4.0mg/L) to severe (>10mg/L), depending on the amount of  $\text{F}^-$  consumed [5-6, 18-19]. Some areas have significant levels of  $\text{F}^-$  in their groundwater due to natural sources present in those areas, making drinking water enriched with geogenic  $\text{F}^-$  ions [20]. In sedimentary and igneous rocks, fluoride exists naturally in the minerals sellaite ( $\text{MgF}_2$ ), cryolite ( $\text{Na}_2\text{AlF}_6$ ), fluorite ( $\text{CaF}_2$ ), and fluorapatite ( $\text{Ca}_5(\text{PO}_4)_3\text{F}$ ) [21]. Few countries have established allowed limits for fluoride in drinking water in cooperation with the WHO to reduce overexposure to  $\text{F}^-$ . Fluorosis poses a significant risk to approximately 200 million people worldwide [22].

## II. STUDY AREA

The Bihar, Jamui district, located at latitude 24°55'N and longitude 86°13'E, was selected as the study area, while most of the district has hilly topography. Table I shows the geographical sampling locations.

TABLE I. LOCATIONS OF GEOGRAPHICAL SAMPLING

Sampling locations	Latitude	Longitude
Harna (S1)	86.41	24.7593
Bishanpur (S2)	86.316	24.984
Nargango (S3)	86.3544	24.7454
Jokatia (S4)	86.3544	24.7454
Lalmatia (S5)	86.3	24.9829
Chandrasekhar Nagar (S6)	86.3553	24.8916
Primary School Chooljeevan Tola (S7)	86.38	24.8
Bhadwaria (S8)	86.4624	24.7635
Prakash Nagar (S9)	86.388	24.8141
Panch Pahari (S10)	86.3362	24.7469
Nabinagar (S11)	86.1572	25.0292
Malaypur (S12)	86.2625	24.9717
Majhwe (S13)		

## III. MATERIALS AND METHODS

### A. Sample Collection

The samples were collected in May-June 2022. A total of 13 groundwater samples were collected from 13 different sampling locations, as shown in Table I, according to the APHA (2005) guidelines. The water samples were collected after pumping for 10 to 15 minutes to eliminate stagnant water. Each water sample was collected in a 1L pre-cleaned plastic bottle and stored at 40°C in a dark box. The store box was carried to the laboratory for further analysis.

### B. Water Analysis

The APHA (2005) protocol was followed for all procedures. A portable water analyzer-371 (Systronics India Ltd.) was used to evaluate pH, TDS, and EC at the sampling locations. The argentometric method was used to determine chloride levels. The fluoride concentration was measured using SPADNS. The concentrations of nitrate and sulfate were determined using ultraviolet (UV) spectrophotometry. Calcium, magnesium, carbonate, and bicarbonate (alkalinity) concentrations were measured using titration.

### C. Water Quality Index (WQI) Calculation

The Water Quality Index (WQI) was established to evaluate groundwater suitability for drinking [32]. Table II shows the WQI ratings.

TABLE II. WATER QUALITY INDEX RATINGS

Water Quality Index	Water Quality Status
0-25	Excellent
26-50	Good
51-75	Poor
76-100	Very Poor
>100	Unsuitable for consumption

### D. Calculating Health Risk Assessment (HRA)

The United States Environmental Protection Agency (US EPA) developed an approach to evaluate the risk to human

health, which has since been extensively used [23-25]. The HRA is an essential reference for conserving and managing a groundwater environment [25-27]. Water and air are the primary sources of health risks for humans. This study focused on drinking water consumption as a primary route, using F<sup>-</sup> to compute HRA and noncarcinogenic risks by:

$$Intake_{oral} = \frac{C_w \times IR \times EF \times ED}{BW \times AT} \quad (1)$$

where  $Intake_{oral}$  (mg/kg/day) is the average daily exposure dosage by groundwater consumption,  $C_w$  is the amount of a specific contaminant (fluoride) present in groundwater (mg/L),  $IR$  is the intake rate for groundwater ( $IR=1.5L/day$  for adult men/women and  $0.7L/day$  for children),  $EF$  is the exposure frequency (365days/year),  $ED$  is the exposure duration ( $ED=30$  years for adults and 12 years for children) obtained from EPA,  $BW$  is the average body weight (55kg and 65kg for women and men, respectively, and 18.5kg for children) [29], and  $AT$  is the average exposure time (4380 days for children and 10950 days for adults). The Hazard Quotient (HQ) was applied to evaluate the fluoride risk as:

$$HQ_{oral} = \frac{Intake_{oral}}{RfD_{oral}} \quad (2)$$

where  $HQ_{oral}$  and  $RfD_{oral}$  are the hazard quotients and reference doses for noncarcinogenic pollutants through drinking water intake, and the oral reference doses of F<sup>-</sup> were obtained from the EPA as 0.04mg/kg/day. An  $HQ_{oral}$  greater than 1 is considered a noncarcinogenic risk, while  $HQ_{oral}$  equal to 1 is considered safe.

## IV. RESULTS AND DISCUSSION

### A. Physicochemical Analysis for Drinking Water

Table III shows the statistical analysis of the physicochemical parameters and a comparison with BIS (IS 1050:2012).

TABLE III. PHYSICOCHEMICAL PARAMETERS ANALYSED STATISTICALLY

Variable	Unit	Minimum	Maximum	Mean	StdDev	BIS acceptable
F	mg/L	2.100	5.800	3.483	1.266	1
pH	-	7.590	8.570	8.167	0.290	6.5-8.5
EC	μS/cm	200.000	1850.000	723.077	460.351	-
TDS	mg/L	125.000	920.000	601.538	279.385	500
Ca <sup>2+</sup>	mg/L	0.180	60.000	20.960	20.853	75
Mg <sup>2+</sup>	mg/L	3.510	47.500	17.368	15.236	30
Cl <sup>-</sup>	mg/L	7.560	350.800	107.152	82.175	250
HCO <sub>3</sub> <sup>-</sup>	mg/L	115.340	460.600	316.306	103.397	-
CO <sub>3</sub> <sup>2-</sup>	mg/L	1.500	8.000	6.310	1.788	-
SO <sub>4</sub> <sup>2-</sup>	mg/L	5.550	75.500	34.981	25.336	200
NO <sub>3</sub> <sup>-</sup>	mg/L	7.550	36.850	16.390	7.949	45

### B. Analysis of Water Quality Index (WQI)

Figure 2 shows the WQI of the sampling locations. As can be seen, the water quality rating of all sampling locations was greater than 100, showing that the drinking water at the selected locations was not suitable for consumption. The fluoride was mainly responsible for the higher value of WQI.

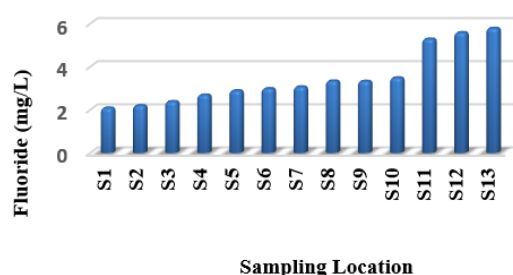


Fig. 1. Fluoride concentrations in the study area.

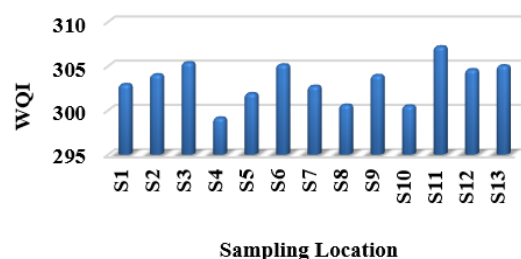


Fig. 2. WQI of the sampling locations

### C. Correlation Analysis

A correlation analysis measures and establishes the relationship between two factors [30]. A value less than 0.5 indicates a low correlation, 0.5 indicates a fair correlation, and greater than 0.5 indicates a very high correlation. Pearson's correlation coefficient reveals the essential connection between the unique variables obtained in the non-parametric form [31]. An inducement of zero demonstrates no relationship between the two variables, while a correlation coefficient of +1 or -1 indicates a strong correlation. Figure 3 shows the association between fluoride and the other water quality measures of the current study. Correlation analysis showed that fluoride was negatively and strongly correlated with  $\text{Ca}^{2+}$  (-0.649) and  $\text{CO}_3^{2-}$  (-0.855) and positively with  $\text{Mg}^{2+}$  (0.559).

Variables	pH	EC	TDS	$\text{Ca}^{2+}$	$\text{Mg}^{2+}$	Cl	$\text{HCO}_3^-$	$\text{CO}_3^{2-}$	$\text{SO}_4^{2-}$	$\text{NO}_3^-$	$\text{F}^-$
pH	1										
EC	-0.145	1									
TDS	-0.015	-0.005	1								
$\text{Ca}^{2+}$	-0.493	0.221	-0.360	1							
$\text{Mg}^{2+}$	0.030	0.209	0.269	-0.419	1						
Cl	-0.121	0.573	-0.025	-0.110	0.278	1					
$\text{HCO}_3^-$	0.289	0.063	-0.004	-0.357	0.259	0.081	1				
$\text{CO}_3^{2-}$	-0.220	-0.180	-0.070	0.465	-0.177	-0.287	0.196	1			
$\text{SO}_4^{2-}$	0.117	-0.025	-0.066	0.024	0.413	0.247	0.362	0.384	1		
$\text{NO}_3^-$	0.270	-0.115	-0.474	-0.175	-0.181	0.276	0.227	0.314	0.132	1	
$\text{F}^-$	0.201	0.225	0.308	-0.649	0.559	0.470	-0.015	-0.855	-0.060	-0.245	1

Fig. 3. Fluoride correlation with the physicochemical parameters.

### D. Analysis of HRA

The population studied was classified into two groups:

- Children, aged between 2-16 years. Data from 50 children were collected.
- Adults, male or female, aged over 16 years. Data from 40 males and 25 females were collected.

Fluoride concentrations in drinking water pose a greater health risk to children than adults. The Hazard Index (HI) ranged from 1.275 to 3.346 for adult men, 1.431-3.954 for adult women, and 1.986-5.4864 for children. Table IV shows the noncarcinogenic risk for the sample in the study region.

### Hazard Index (HI)

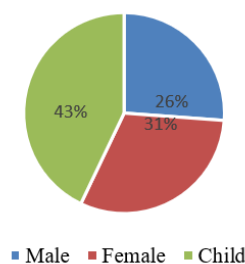


Fig. 4. Hazard index of males, females, and children.

TABLE IV. NONCARCINOGENIC RISK FOR MALES, FEMALES, AND CHILDREN IN THE STUDY REGION

$C_w$ Fluoride	$\text{Intake}_{\text{oral}}$ (Male)	$\text{Intake}_{\text{oral}}$ (Female)	$\text{Intake}_{\text{oral}}$ (Child)	$HQ_{\text{oral}}$ (Male)	$HQ_{\text{oral}}$ (Female)	$HQ_{\text{oral}}$ (Child)
2.1	0.121153846	0.057273	0.079459	3.028846154	1.431818182	1.986486486
2.21	0.051	0.060273	0.083622	1.275	1.506818182	2.090540541
2.4	0.055384615	0.065455	0.090811	1.384615385	1.636363636	2.27027027
2.7	0.062307692	0.073636	0.102162	1.557692308	1.840909091	2.554054054
2.9	0.066923077	0.079091	0.10973	1.673076923	1.977272727	2.743243243
3	0.069230769	0.081818	0.113514	1.730769231	2.045454545	2.837837838
3.08	0.071076923	0.084	0.116541	1.776923077	2.1	2.913513514
3.35	0.077307692	0.091364	0.126757	1.932692308	2.284090909	3.168918919
3.34	0.077076923	0.091091	0.126378	1.926923077	2.277272727	3.159459459
3.5	0.080769231	0.095455	0.132432	2.019230769	2.386363636	3.310810811
5.3	0.122307692	0.144545	0.200541	3.057692308	3.613636364	5.013513514
5.6	0.129230769	0.152727	0.211892	3.230769231	3.818181818	5.297297297
5.8	0.133846154	0.158182	0.219459	<b>3.346153846</b>	<b>3.954545455</b>	<b>5.486486486</b>

### V. CONCLUSION

This study investigated the fluoride contamination of groundwater samples in the Jamui district of Bihar, India, and examined the relationship between fluoride contamination and

other water quality measures. The observed physicochemical data were compared with the Bureau of Indian Standards (BIS) limits, and it was found that fluoride exceeded the allowed limits. The fluoride was the leading cause of contamination in the selected area, responsible for a higher value of WQI. That

indicates that drinking water at the selected locations is not suitable for consumption. Therefore, appropriate filtration units should be installed in hand pumps to remove the excess fluoride. Correlation analysis showed that fluoride is strongly negatively correlated with  $\text{Ca}^{2+}$  (-0.649) and  $\text{CO}_3^{2-}$  (-0.855) and positively with  $\text{Mg}^{2+}$  (0.559). Regression analysis showed that fluoride has a decreasing trend with calcium and carbonate and an increasing trend with magnesium. Drinking water containing fluoride causes greater health risks in children than in adults, according to published risk assessments for human health. The range of possible values for the Hazard Index (HI) for adult men ranged from 1.275 to 3.346, for adult females from 1.431 to 3.954, and for children from 1.986 to 5.4864. In the study area, children are more susceptible to health risks than adults due to the highly polluted drinking water with high fluoride levels.

### REFERENCES

- [1] S. Ahmad, R. Singh, T. Arfin, and K. Neeti, "Fluoride contamination, consequences and removal techniques in water: a review," *Environmental Science: Advances*, vol. 1, no. 5, pp. 620–661, 2022, <https://doi.org/10.1039/D1VA00039J>.
- [2] T. Arfin, P. Ranjan, S. Bansod, R. Singh, S. Ahmad, and K. Neeti, "Organic Electrodes: An Introduction," in *Organic Electrodes: Fundamental to Advanced Emerging Applications*, R. K. Gupta, Ed. Cham: Springer International Publishing, 2022, pp. 1–26.
- [3] T. Poonia, N. Singh, and M. C. Garg, "Contamination of Arsenic, Chromium and Fluoride in the Indian groundwater: a review, meta-analysis and cancer risk assessment," *International Journal of Environmental Science and Technology*, vol. 18, no. 9, pp. 2891–2902, Sep. 2021, <https://doi.org/10.1007/s13762-020-03043-x>.
- [4] T. Onipe, J. N. Edokpayi, and J. O. Odiyo, "A review on the potential sources and health implications of fluoride in groundwater of Sub-Saharan Africa," *Journal of Environmental Science and Health, Part A*, vol. 55, no. 9, pp. 1078–1093, Jul. 2020, <https://doi.org/10.1080/10934529.2020.1770516>.
- [5] S. Ali *et al.*, "Concentration of fluoride in groundwater of India: A systematic review, meta-analysis and risk assessment," *Groundwater for Sustainable Development*, vol. 9, Oct. 2019, Art. no. 100224, <https://doi.org/10.1016/j.gsd.2019.100224>.
- [6] C. F. Z. Lacson, M. C. Lu, and Y. H. Huang, "Fluoride-containing water: A global perspective and a pursuit to sustainable water defluoridation management -An overview," *Journal of Cleaner Production*, vol. 280, Jan. 2021, Art. no. 124236, <https://doi.org/10.1016/j.jclepro.2020.124236>.
- [7] K. Neeti, R. Singh, and S. Ahmad, "The role of green nanomaterials as effective adsorbents and applications in wastewater treatment," *Materials Today: Proceedings*, Nov. 2022, <https://doi.org/10.1016/j.matpr.2022.11.300>.
- [8] A. N. Laghari, Z. A. Siyal, D. K. Bangwar, M. A. Soomro, G. D. Walasai, and F. A. Shaikh, "Groundwater Quality Analysis for Human Consumption: A Case Study of Sukkur City, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2616–2620, Feb. 2018, <https://doi.org/10.48084/etasr.1768>.
- [9] X. Zhang, H. Gao, T. Yang, H. Wu, Y. Wang, and X. Wan, "Al<sup>3+</sup>-promoted fluoride accumulation in tea plants (*Camellia sinensis*) was inhibited by an anion channel inhibitor DIDS," *Journal of the Science of Food and Agriculture*, vol. 96, no. 12, pp. 4224–4230, 2016, <https://doi.org/10.1002/jsfa.7626>.
- [10] K. Percy, J. Elphick, and C. Burnett-Seidel, "Toxicity of fluoride to aquatic species and evaluation of toxicity modifying factors," *Environmental Toxicology and Chemistry*, vol. 34, no. 7, pp. 1642–1648, 2015, <https://doi.org/10.1002/etc.2963>.
- [11] R. Singh and N. S. Maurya, "Microbiological drinking water quality and its relation to human health - A case of Patna municipal area," *Indian Journal of Environmental Protection*, vol. 37, pp. 41–47, Jan. 2017.
- [12] *International standards for drinking-water*, 3rd ed. Geneva, Switzerland: World Health Organization, 1971.
- [13] N. Kumar, A. A. Mahessar, S. A. Memon, K. Ansari, and A. L. Qureshi, "Impact Assessment of Groundwater Quality using WQI and Geospatial tools: A Case Study of Islamabad, Tharparkar, Pakistan," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5288–5294, Feb. 2020, <https://doi.org/10.48084/etasr.3289>.
- [14] K. L. Vandana, B. Srishti Raj, and R. Desai, "Dental Fluorosis and Periodontium: an Original Research Report of In Vitro and In Vivo Institutional Studies," *Biological Trace Element Research*, vol. 199, no. 10, pp. 3579–3592, Oct. 2021, <https://doi.org/10.1007/s12011-020-02494-0>.
- [15] S. J. Wimalawansa, "Does fluoride cause the mysterious chronic kidney disease of multifactorial origin?," *Environmental Geochemistry and Health*, vol. 42, no. 9, pp. 3035–3057, Sep. 2020, <https://doi.org/10.1007/s10653-019-00503-3>.
- [16] M. Skórka-Majewicz *et al.*, "Effect of fluoride on endocrine tissues and their secretory functions -- review," *Chemosphere*, vol. 260, Dec. 2020, Art. no. 127565, <https://doi.org/10.1016/j.chemosphere.2020.127565>.
- [17] N. R. Johnston and S. A. Strobel, "Principles of fluoride toxicity and the cellular response: a review," *Archives of Toxicology*, vol. 94, no. 4, pp. 1051–1069, Apr. 2020, <https://doi.org/10.1007/s00204-020-02687-5>.
- [18] A. Raja and T. Gopikrishnan, "Drought Analysis Using the Standardized Precipitation Evapotranspiration Index (SPEI) at Different Time Scales in an Arid Region," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9034–9037, Aug. 2022, <https://doi.org/10.48084/etasr.5141>.
- [19] P. K. Mohapatra, R. Vijay, P. R. Pujari, S. K. Sundaray, and B. P. Mohanty, "Determination of processes affecting groundwater quality in the coastal aquifer beneath Puri city, India: a multivariate statistical approach," *Water Science and Technology*, vol. 64, no. 4, pp. 809–817, Aug. 2011, <https://doi.org/10.2166/wst.2011.605>.
- [20] A. Chowdhury, M. K. Adak, A. Mukherjee, P. Dhak, J. Khatun, and D. Dhak, "A critical review on geochemical and geological aspects of fluoride belts, fluorosis and natural materials and other sources for alternatives to fluoride exposure," *Journal of Hydrology*, vol. 574, pp. 333–359, Jul. 2019, <https://doi.org/10.1016/j.jhydrol.2019.04.033>.
- [21] S. K. Jha, V. K. Mishra, D. K. Sharma, and T. Damodaran, "Fluoride in the Environment and Its Metabolism in Humans," in *Reviews of Environmental Contamination and Toxicology Volume 211*, D. M. Whitacre, Ed. New York, NY: Springer, 2011, pp. 121–142.
- [22] H. Kabir, A. K. Gupta, and S. Tripathy, "Fluoride and human health: Systematic appraisal of sources, exposures, metabolism, and toxicity," *Critical Reviews in Environmental Science and Technology*, vol. 50, no. 11, pp. 1116–1193, Jun. 2020, <https://doi.org/10.1080/10643389.2019.1647028>.
- [23] P. Li, H. Qian, J. Wu, J. Chen, Y. Zhang, and H. Zhang, "Occurrence and hydrogeochemistry of fluoride in alluvial aquifer of Weihe River, China," *Environmental Earth Sciences*, vol. 71, no. 7, pp. 3133–3145, Apr. 2014, <https://doi.org/10.1007/s12665-013-2691-6>.
- [24] A. Narsimha and S. Rajitha, "Spatial distribution and seasonal variation in fluoride enrichment in groundwater and its associated human health risk assessment in Telangana State, South India," *Human and Ecological Risk Assessment: An International Journal*, vol. 24, no. 8, pp. 2119–2132, Nov. 2018, <https://doi.org/10.1080/10807039.2018.1438176>.
- [25] P. Li, X. Li, X. Meng, M. Li, and Y. Zhang, "Appraising Groundwater Quality and Health Risks from Contamination in a Semiarid Region of Northwest China," *Exposure and Health*, vol. 8, no. 3, pp. 361–379, Sep. 2016, <https://doi.org/10.1007/s12403-016-0205-y>.
- [26] P. Li, H. Qian, and J. Wu, "Conjunctive use of groundwater and surface water to reduce soil salinization in the Yinchuan Plain, North-West China," *International Journal of Water Resources Development*, vol. 34, no. 3, pp. 337–353, May 2018, <https://doi.org/10.1080/07900627.2018.1443059>.
- [27] J. Wu and Z. Sun, "Evaluation of Shallow Groundwater Contamination and Associated Human Health Risk in an Alluvial Plain Impacted by Agricultural and Industrial Activities, Mid-west China," *Exposure and Health*, vol. 8, no. 3, pp. 311–329, Sep. 2016, <https://doi.org/10.1007/s12403-015-0170-x>.

- 
- [28] B. Means, "Risk-assessment guidance for Superfund. Volume 1. Human Health Evaluation Manual. Part A. Interim report (Final)," Environmental Protection Agency, Washington, DC (USA). Office of Solid Waste and Emergency Response, PB-90-155581/XAB; EPA-540/1-89/002, Dec. 1989. [Online]. Available: <https://www.osti.gov/biblio/7037757>.
- [29] Nutrient Requirements and Recommended Dietary Allowances for Indians: A Report of the Expert Group of the Indian Council of Medical Research. New Delhi, India: Indian Council of Medical Research, 2010.
- [30] N. Kazakis, C. Mattas, A. Pavlou, O. Patrikaki, and K. Voudouris, "Multivariate statistical analysis for the assessment of groundwater quality under different hydrogeological regimes," *Environmental Earth Sciences*, vol. 76, no. 9, May 2017, <https://doi.org/10.1007/s12665-017-6665-y>, Art. no. 349.
- [31] M. Vasanthavigar, K. Srinivasamoorthy, and M. V. Prasanna, "Identification of groundwater contamination zones and its sources by using multivariate statistical approach in Thirumanimuthar sub-basin, Tamil Nadu, India," *Environmental Earth Sciences*, vol. 68, no. 6, pp. 1783–1795, Mar. 2013, <https://doi.org/10.1007/s12665-012-1868-8>.
- [32] R. M. Brown, N. I. McClelland, R. A. Deininger, and M. F. O'Connor, "A Water Quality Index — Crashing the Psychological Barrier," in *Indicators of Environmental Quality*, Boston, MA, USA, 1972, pp. 173–182, [https://doi.org/10.1007/978-1-4684-2856-8\\_15](https://doi.org/10.1007/978-1-4684-2856-8_15).



# Slope Reliability Analysis in Locating and Observing the Direction of Failure Propagation

**Houcine Djefal**

Tissemsilt University, Algeria | Ecole Nationale Polytechnique, Algeria  
houcine.djefal@yahoo.com  
(corresponding author)

**Smain Belkacemi**

Ecole Nationale Polytechnique, Algeria  
smain.belkacemi@g.enp.edu.dz

Received: 22 November 2022 | Revised: 5 December 2022 | Accepted: 12 December 2022

## ABSTRACT

Slope stability analysis is traditionally carried out with the ultimate equilibrium methods. In these approaches, the global assessment of slope stability is obtained through the evaluation of the factor of safety. The strength parameters that allow the evaluation of resistant forces along the hypothetical slip surface are considered unique and constant. The shape of the stress-strain curve of the soil may substantially affect the stability of the soil mass over, particularly in the case of the potential of a progressive failure. As a part of this work, the influence of uncertainties on the peak and post-peak strength parameters and the impact of a progressive failure on the failure probability of a hypothetical slip surface are discussed. The conventional method of Morgenstern-Price of slope stability analysis consisted of locating the critical slip surface by evaluating the minimum factor of safety. The probability of failure was estimated using three different techniques, i.e. FORM, FOSM, and Monte Carlo simulation. This study improved the assessment of the impact of progressive failure on the risk of slope failure. The reliability analysis of slices allowed for locating the failure area and observing the direction of failure propagation.

*Keywords-limiting equilibrium; slope stability; slip surface; progressive failure; uncertainties; reliability index*

## I. INTRODUCTION

Slope failure has become one of the most important problems in geotechnical engineering. The limit equilibrium method is commonly used in slope stability analysis due to its simplicity and high efficiency. The assumptions and simplifications of conventional limit equilibrium slope stability methods may not be sufficient to represent the behavior of complex slope problems. Since this method is a statically indeterminate problem, assumptions on the inter-slice shear forces are employed to render the problem statically determinate. The classical limit equilibrium method only considers the ultimate limit state of the slope and provides no information on the development of progressive failure [1]. Some researchers [1-3] extended the limit equilibrium method to analyze the stability of strain-softening slopes, and assumed that the soil strength decreases directly to the residual value from the peak value. With the recent advancements in computational approaches, it is possible to model slope stability problems more realistically by adopting numerical simulation methods [4-6].

Due to depositional and post-depositional processes, soil characteristics such as density, mineral composition, moisture

content, stress history, and shear strength change geographically. In addition, there are other factors, including the insufficient number of samples and measurement errors, which make more difficult to determine the soil properties precisely. In practice, measurements are taken only at selected locations, and thus, the soil properties are known only at these locations and can therefore be considered as random quantities. Then, utilizing their correlation structures, soil characteristics may be efficiently described in terms of their spatial variations [7, 8]. True security is unknown because the simplifications of these methods take into account neither the variability of soil strength nor the role of the deformation behavior of the mass of slip soil. The progressive failure mechanism is also dominant for the slope stability, because the interfaces between components in slopes exhibit strain-softening behavior [7]. The shape of the stress-strain curve of the soil may substantially affect the stability of the soil mass over, particularly in the case of a potential progressive failure. The many sources of uncertainty that emerge in these types of problems are systematically evaluated using a reliability approach to a slope stability analysis that takes into account the spatially correlated soil variables. Measurement errors and bias that often occur during soil investigations are now incorporated into the

probabilistic model of soil properties in addition to the spatial variability and the effects of spatial averaging. Probabilistic methods are increasingly recognized as a useful tool to explicitly consider the effect of uncertainties on slope stability assessment. Many methods regarding slope reliability analysis have been developed [8-13]. Such approaches allow the engineers to better understand and quantify the risks.

In order to analyze slope stability, the current research suggests integrating several mechanical methodologies with reliability analysis methods. This paper examines the issue of slope stability in a probabilistic framework and aims to quantify the influence of uncertainties of strength parameters, and the incidence of progressive failure on stability. The analysis employed the Morgenstern-Price slope stability deterministic approach. Probabilistic methods FORM (First Order Reliability Method), FOSM (First Order Second Moment), and the Monte Carlo Simulation (MCS) technique were used to perform slope reliability analysis at different uncertainty levels of the basic parameters.

## II. RELIABILITY INDEX AND PROBABILITY OF FAILURE

Soils are heterogeneous materials, created by complex geological processes. Their properties change from one point to another. Due to the uncertainty inherent spatial variability, and the limited information available, soil properties can be considered as random variables. The probability theory provides a mathematical framework for quantifying uncertainty. The probabilistic approach is based on a deterministic model, in which the various uncertain parameters are modeled as random variables  $X = [X_1, X_2, X_3, \dots, X_n]$ . Considering the simple case of a resistance  $R$  subjected to an  $S$ , consisting of independent random variables or dependent and distributed according to given laws of probability, the failure probability free operation, reliability, is equal to the probability of occurrence of the event  $R \geq S$ .  $R$  and  $S$  are characterized by their first moments, i.e.  $\mu_R$  and  $\mu_S$ ,  $\sigma_R$  and  $\sigma_S$ , respectively [14].

A widely accepted and simple definition of the reliability index is the ratio of the expected safety margin to the standard deviation of the performance function. Several numerical methods have been developed for estimating the statistical moments of a performance function dependent on multiple random variables. For practical purposes, the two important statistical moments to be estimated are the expected value and the variance (square of standard deviation) [14-21]. The fundamental problem of reliability theory lies on the calculation of the integral multiple of probability to estimate the probability of failure  $P_f$ :

$$P_f = P[g(X) \leq 0] = \int_{g(X) \leq 0} f(x) dx \quad (1)$$

where  $g(X)$  is the performance function associated with an operating rule, the relation  $g(X) = 0$  constitutes the state limit.

Various solution methods have been proposed to estimate  $P_f$  and the Reliability Index. Among the most widely used methods are FOSM, FORM, and MCS [14, 15].

## III. SLOPE STABILITY ANALYSIS

Once the geometry and conditions of the basement of a slope were determined, slope stability can be assessed. The main objectives of slope stability analysis include the assessment of rupture risk, through the calculation of the overall safety factor for a slope and the locations along the potential slip surface areas with high potential rupture [16]. There are several methods of slope stability analysis, based on the calculation of the limit equilibrium. Most of these methods use a technique called slices. In these methods, the safety factor is calculated using one or more equations of static equilibrium applied to the soil mass. The limit equilibrium methods need first to define the surface for which the safety factor will be evaluated. The safety coefficient calculations are repeated for a sufficient number of slip surfaces arbitrarily selected to locate the surface having the minimum safety factor. The slip surface shape depends on the geometry, stratigraphy of the problem, physical characteristics, and capabilities of the analytical method used. Forms of the slip surface may be circular or non-circular. The methods of static equilibrium decompose the soil mass above the sliding surface, in a finite number of slices. The analysis problem of slope stability is statically indeterminate. The various limit equilibrium methods use different assumptions or constraints to remove the indeterminacy.

## IV. PROGRESSIVE FAILURE

Since the sixties, several authors have observed that the failure of slopes is a consequence of the development of large shear stresses well before the start of the failure, eventually leading to what is called the progressive failure [16-18]. More recently, many authors confirm that the rupture of the slopes is often a progressive nature [22]. The analysis of slope stability is often associated with the stress-strain curve of the soil for very large deformation which mobilizes a residual strength. If the material, along with a sliding surface, is best represented by a stress-strain curve having a peak, it will be impossible to engage the same shear strength on the entire surface. After the mobilization of the shear strength available to the peak, the mobilized shear strength decreases and tends towards the peak with shear strength increasing ground deformation. For simplicity, we assume that the strength reduction, the peak to residual, is instantaneous (Figure 1).

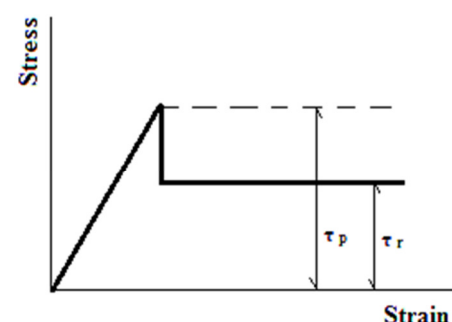


Fig. 1. Stress-strain curve used in the model.

Assessing the reliability of a slope in a progressive failure is a difficult task. When the shear strength mobilized locally

corresponds to the maximum shear strength of the soil, there is a failure due to local stress concentrations likely to spread. The differential shear force is systematically transferred to the neighboring elements.

## V. PROGRESSIVE FAILURE PROBABILITY

The slope stability problem can be considered as a system with multiple slip surfaces. The failure probability of the critical slip surface is considered as an estimate of the failure probability of the system. This approach assumes that the failure probabilities along the different slide surfaces are strongly correlated [23, 24]. The common approach in the probabilistic analysis is to locate the critical surface and then calculate the failure probability corresponding to this surface. However, the surface having a minimum safety factor may not be the surface of maximum failure probability [25]. The critical surface probability is associated with the highest failure probability or the lowest reliability [24, 26-29]. A model based on the bivariate failure probability of successive slices was presented in [17, 18]. This model has been developed to include the gradual overall probability by considering the slope of the spread out along the entire sliding surface. Fracture propagation is only a consequence of the interdependence of different age breakdowns of a sliding surface. The bivariate calculated probabilities relate to progressive successive events. This probability depends on the order in which the various tranches occur in sequence during the rupture propagation.

### A. The Proposed Model

The proposed model treats the probabilistic progressive failure of a slope based on conditional probabilistic calculations of failure. The formulation of the statistical parameters of the safety factor, and in particular the correlation coefficient between the performance functions  $G_i$  and  $G_{i-1}$  of slices  $i$  and  $i-1$ , is laborious. Extending the problem to 4 strength parameters ( $C_p, \phi_p$ ), ( $C_r, \phi_r$ ), representing respectively the peak and the residual (post-peak) state, and their spatial variability is particularly cumbersome. Obviously, it is desirable to have a simpler model, in which complex calculations involving such correlations and joint probabilities of events are not required. The performance function  $G_i$  of a slice  $i$  is defined as the safety margin calculated assuming that the shear strength has reached the maximum peak:

$$G_i = (R_p)_i - S_i \quad (2)$$

The performance function  $G_j^*$  of progressive failure is defined as the conditional safety margin related to failure propagation until the  $j$ th slice of the slip surface:

$$G_j^* = (R_p)_j + (R_r)_{1,j-1} - S_{1,j} \quad (3)$$

where  $(R_p)_j$  is the shear strength of the slice  $j$  in peak,  $(R_r)_{1,j-1}$  is the sum of the residual shear resistance of units 1 to  $(j-1)$ ,  $S_j$  is the shear stress of slice  $j$ ,  $S_{1,j}$  is the sum of shear stresses along the slices 1 to  $j$ .

### B. Progressive Failure Probability

Let us assume  $F_j$  as the conditional event for failure, the mobilization of peak strength, developed at the slice  $j$ , given that the  $(j-1)$  previous installments mobilize the peak strength

after the residual strength. The conditional event  $F_j$  is expressed by  $F_j = (G_j^* < 0)$ .  $E_j$  is the event of progressive failure of slices 1 to  $j$  in the order listed and about 1 to  $j$  represents the order of reaching the breaking of each slice. The probability that the failure develops gradually from slice 1 to slice  $j$  is a conditional probability. If we consider the situation in which the failure is propagated to the  $j$ th slice, this occurs only if the slices 1 to  $(j-1)$  are beyond the post-peak, this particular situation allows to write:

$$P[F_{j-1} / F_j] = 1 \quad (4)$$

knowing that:

$$P[F_{j-1} / F_j] = \frac{P[F_j F_{j-1}]}{P[F_j]} \quad (5)$$

$$P[F_j / F_{j-1}] = \frac{P[F_j F_{j-1}]}{P[F_{j-1}]} \quad (6)$$

Equations (4)-(6) are used to acquire (7):

$$P[F_j / F_{j-1}] = \frac{P[F_j]}{P[F_{j-1}]} \quad (7)$$

Hence:

$$P[E_j] = \frac{P[F_j]}{P[F_{j-1}]} P[E_{j-1}] \quad (8)$$

Failure propagation, based on conditional events, is a process in which a number of slices are already at the state of failure, having a failure probability that should be considered equal to unity. The likelihood of progressive failure of the system composed of  $n$  slices knowing that the first  $j$  slices are already beyond the state of failure is:

$$P[E_{(j+1)n}] = P[E_n | E_1 = E_2 = \dots = E_j = 1] \quad (9)$$

Hence:

$$P[E_{(j+1)n}] = P[E_n / E_{n-1}] P[E_{n-1} / E_{n-2}] P[E_{n-2} / E_{n-3}] \dots P[E_{j+1} / E_j] \quad (10)$$

Calculations are performed in 3 distinct stages. In the first step, successive calculations of the probability or reliability index of the performance function  $G_i$  of each slice are made. After evaluating the failure probability of each slice, the direction of the rupture propagation is identified. A new index  $j$  is used to order the slices in descending order of failure probability of the slice. In the second stage, the calculation of conditional probabilities  $P[F_j / F_{j-1}]$  and the probabilities  $P[E_j]$  is performed. The final step in the simulation of the probability of progressive failure of the system composed of  $n$  slices, considering that for an increasing proportion, defined by a number of slices, the previously located surface potential shift, occurs beyond the peak strength and thereby engages only a residual strength.

## VI. ILLUSTRATIVE EXAMPLE

The evaluation of the reliability index of slopes, taking into account the progressive failure, is made in two stages. In the first step, the limit equilibrium method of Morgenstern-Price is used to locate the critical slip surface. The geometrical parameters are assumed deterministic and their average values are shown in Table I. In the second step, a probabilistic analysis of the critical slip surfaces is carried out by means of physical models of the objective function for each of the developed considered deterministic methods. The analysis is made by taking into account the variation of the strength parameters and includes the impact of a progressive failure through an evolution of strength parameters of their initial values ( $C_p, \phi_p$ ) at their peak to residual values ( $C_r, \phi_r$ ). The probabilistic methods FORM, FOSM, and MCS were used. The parameters considered as random variables in the reliability analysis of slopes are shown in Table I. The change in the reliability index or the failure probability as a function of the variation coefficient of random variables is obtained while keeping the average random variables constant, and setting the coefficients of change in the other random variables at an arbitrary fixed value of 30%, and by changing the value of the variation coefficient of the variable into account. Random variable distributions are considered either normal or lognormal. Strength parameters to peak and post-peak are either independent or negatively correlated with a correlation coefficient  $\rho = -0.5$ .

TABLE I. PARAMETERS AND MEAN VALUES CONSIDERED IN THE SLOPE STABILITY ANALYSIS

Parameter	Symbol	Variable	Mean value
Cohesion term in peak	$C_p$	Random	20kN/m <sup>2</sup>
Friction angle in peak	$\phi_p$	Random	25°
Cohesion term in post-peak	$C_r$	Random	10kN/m <sup>2</sup>
Friction angle in post-peak	$\phi_r$	Random	10°
Weight	$\gamma$	Deterministic	18kN/m <sup>3</sup>
Slope angle	$\beta$	Deterministic	26.56°
Slope height	$H$	Deterministic	20m

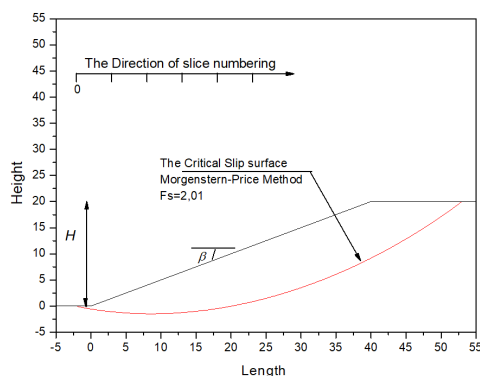


Fig. 2. The critical slip surface Morgenstern-Price method.

### A. Deterministic Analysis of Slope Stability

The results of the deterministic study of the slope stability obtained by Morgenstern-Price method are represented below. The equilibrium of forces and moments is examined for each slice and the whole system, by dividing the soil mass into 250

slices. The critical slip surface is shown in Figure 2. The results of the deterministic analysis of slope stability have yielded a value of safety factor  $F_s = 2.01$ .

### B. Probabilistic Analysis of Slope Stability

The results of the reliability analysis of each slice are shown in Figures 3-8. These figures show, for each value of the variation coefficient of strength parameters, variations in the reliability index of each slice of the system depending on the physical number of slices and the number of slices ordered in ascending order of reliability index. The main results are:

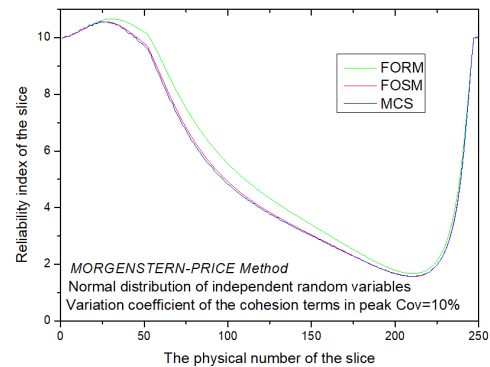


Fig. 3. Variation of the reliability index of the slice as a function of the physical number of the slice (normal distribution of independent random variables). Variation coefficient of the cohesion terms in peak  $Cov = 10\%$ .

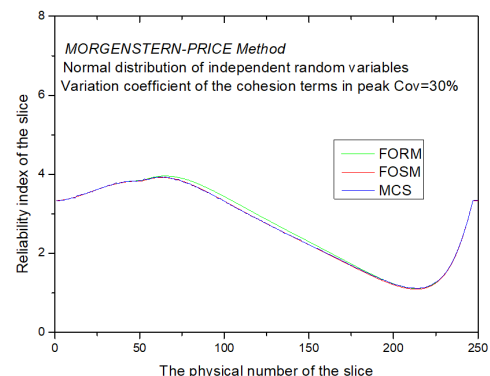


Fig. 4. Variation of the reliability index of the slice as a function of the physical number of the slice (normal distribution of independent random variables). Variation coefficient of the cohesion terms in peak  $Cov = 30\%$ .

- The change in the shape of the reliability index as a function of the physical number of slices can locate the first slice passed in the failure state. It corresponds to the slice of the least reliability index.
- The change of the reliability index takes place simultaneously in both directions, relative to the first phase triggering failure.
- A strong decrease in the reliability index of slices is observed for high variation coefficients of strength parameters in post-peak. This decrease reflects the high sensitivity of the reliability index of a slice to the strength parameters in post-peak relatively to peak strength parameters.

- Taking into account the correlation between the strength parameters results in an increase in the reliability index of slices. This increase is more sensitive to low variation coefficients.
- For all cases considered, the FORM, FOSM and MCS methods give very similar results, regardless of the distribution type of random variables, normal or lognormal, and regardless of the correlation matrix considered.

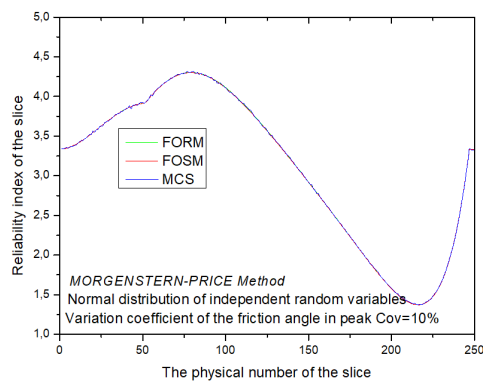


Fig. 5. Variation of the reliability index of the slice as a function of the physical number of the slice (normal distribution of independent random variables). Variation coefficient of the cohesion terms in peak  $Cov = 10\%$ .

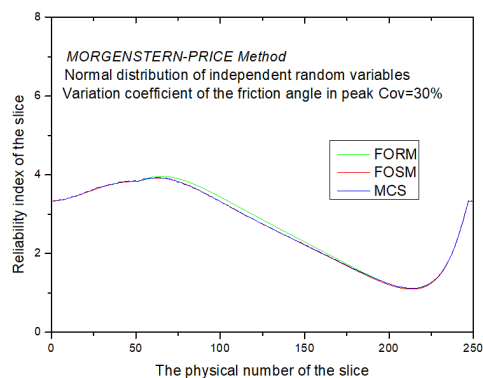


Fig. 6. Variation of the reliability index of the slice as a function of the physical number of the slice (normal distribution of independent random variables). Variation coefficient of the cohesion terms in peak  $Cov = 30\%$ .

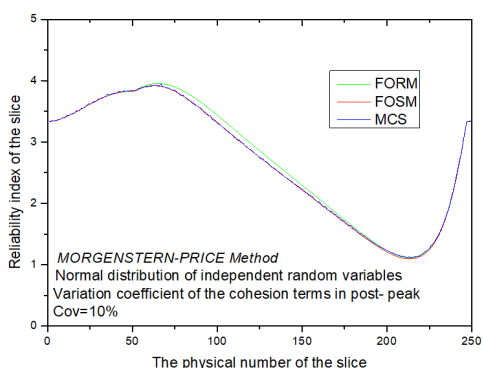


Fig. 7. Variation of the reliability index of the slice as a function of the physical number of the slice (normal distribution of independent random variables). Variation coefficient of the cohesion terms in post-peak  $Cov = 10\%$ .

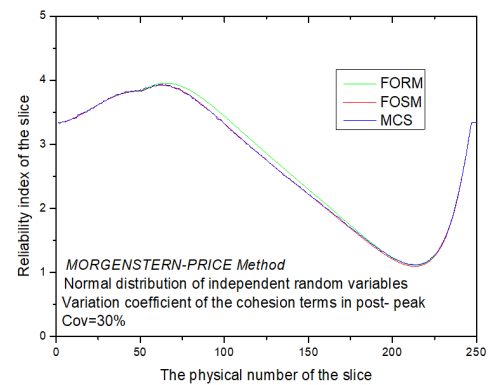


Fig. 8. Variation of the reliability index of the slice as a function of the physical number of the slice (normal distribution of independent random variables). Variation coefficient of the cohesion terms in post-peak  $Cov = 30\%$ .

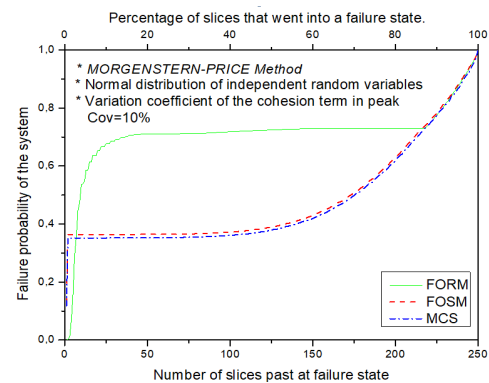


Fig. 9. Variation of the progressive failure probability of the system as a function of the number of slices past failure state (normal distribution of independent random variables). Variation coefficient of the cohesion term in peak  $Cov = 10\%$ .

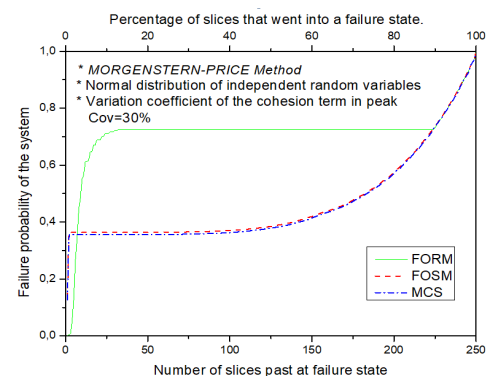


Fig. 10. Variation of the progressive failure probability of the system as a function of the number of slices past failure state (normal distribution of independent random variables). Variation coefficient of the cohesion term in peak  $Cov = 30\%$ .

The probabilistic analysis results of the progressive failure of the system are shown in Figures 9-14. These figures show, for each value of the variation coefficient of strength parameters peak or post-peak changes in the failure probability of the system based on the number of slices past at failure state. The main derived results are:

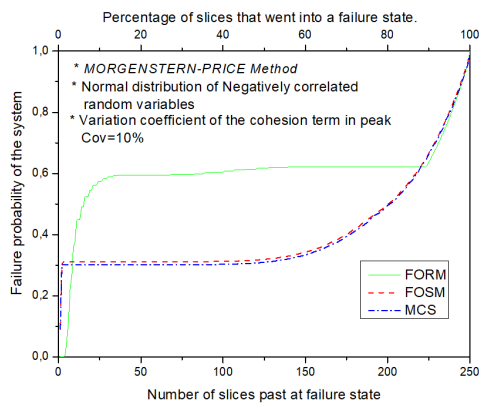


Fig. 11. Variation of the progressive failure probability of the system as a function of the number of slices past failure state (normal distribution of independent random variables). Variation coefficient of the cohesion term in peak  $Cov = 10\%$ .

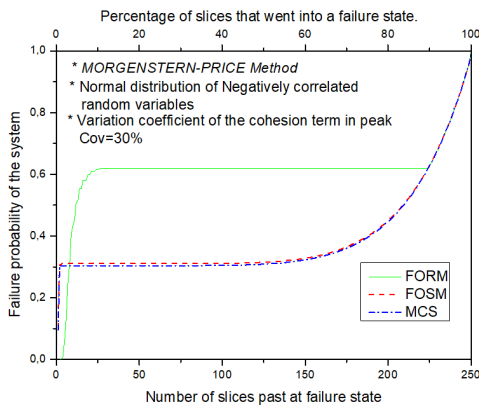


Fig. 12. Variation of the progressive failure probability of the system as a function of the number of slices past failure state (normal distribution of independent random variables). Variation coefficient of the cohesion term in peak  $Cov = 30\%$ .

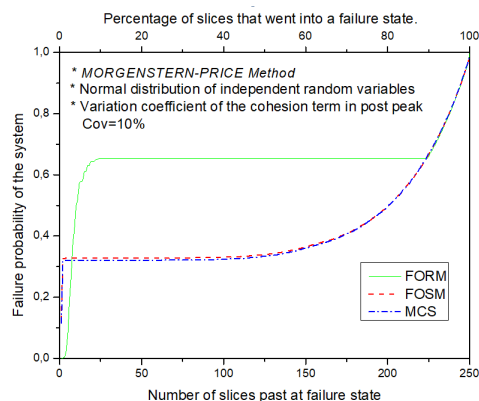


Fig. 13. Variation of the progressive failure probability of the system as a function of the number of slices past failure state (normal distribution of independent random variables). Variation coefficient of the cohesion term in post-peak  $Cov = 10\%$ .

- For all the considered cases, FOSM and MCS give very similar results, regardless of the type of distribution of the random variables.

- The failure probabilities of the system obtained by the reliability methods FOSM and MCS differ from those obtained by FORM, and they become similar when over 90% of the slices of the system are under the condition of failure.
- The inclusion of a negative correlation between the strength parameters in probabilistic analysis, results in a decrease in the failure probability of the system.

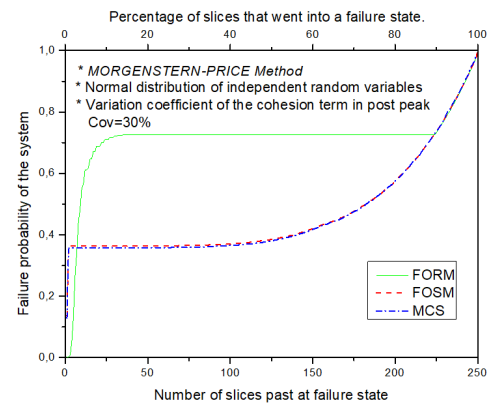


Fig. 14. Variation of the progressive failure probability of the system as a function of the number of slices past failure state (normal distribution of independent random variables). Variation coefficient of the cohesion term in post-peak  $Cov = 30\%$ .

The results of the system reliability analysis are shown in Figures 15-20. These figures show, for each strength parameter of the limit state function, changes in the reliability index of the system depending on variation coefficient. The main derived results are:

- The reliability index is strongly dependent on the strength parameters in post-peak relative to the peak strength parameters, regardless of the type of distribution of the random variables and regardless of the considered correlation matrix.

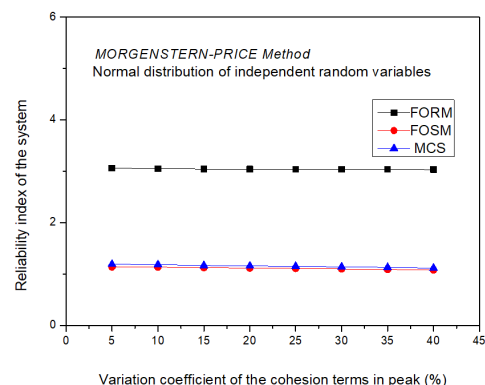


Fig. 15. Variation of the reliability index of the system as a function of the variation coefficient of the cohesion terms in peak. Independent random variables.



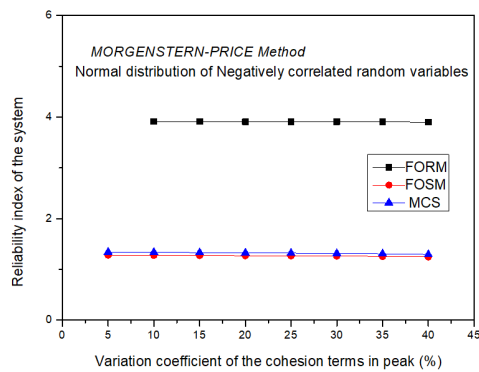


Fig. 16. Variation of the reliability index of the system as a function of the variation coefficient of the cohesion terms in peak. Negatively correlated random variables.

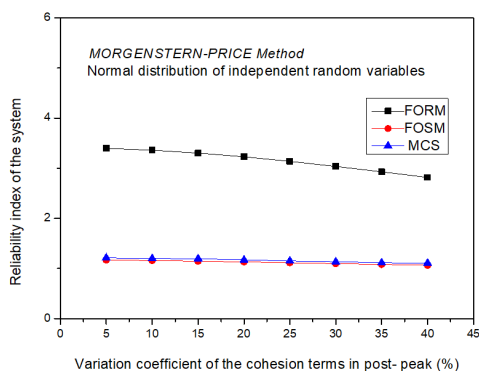


Fig. 17. Variation of the reliability index of the system as a function of the variation coefficient of the cohesion terms in post-peak. Independent random variables.

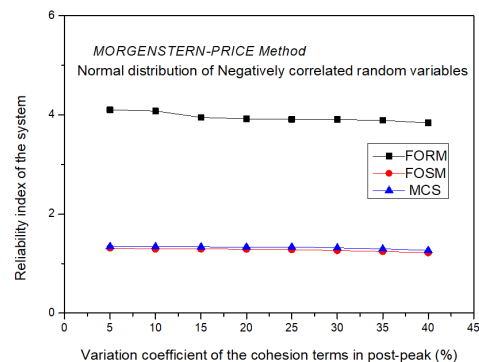


Fig. 18. Variation of the reliability index of the system as a function of the variation coefficient of the cohesion terms in post-peak. Negatively correlated random variables.

- In post peak, the cohesion term has a strong influence on the reliability index of the system with respect to the friction angle.
- The distribution of strength parameters has a strong influence on the reliability index.
- For all the considered cases, FOSM and MCS give very similar results, regardless of the distribution of the random variables.

- The reliability index of the system obtained by FOSM and MCS differs from that obtained by FORM.

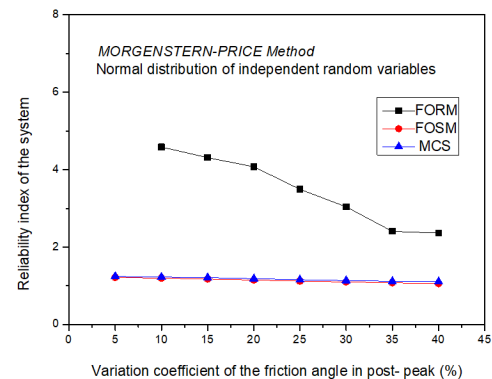


Fig. 19. Variation of the reliability index of the system as a function of the variation coefficient of the cohesion terms in post-peak. Independent random variables.

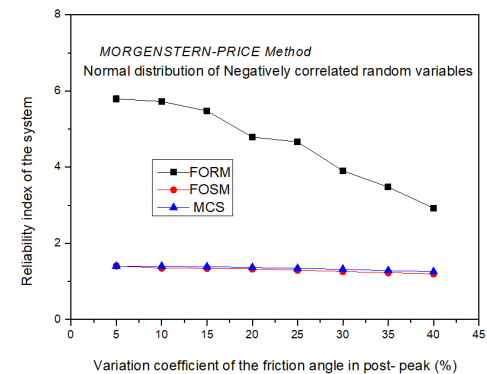


Fig. 20. Variation of the reliability index of the system as a function of the variation coefficient of the cohesion terms in post-peak. Negatively correlated random variables.

## CONCLUSION

The study of reliability analysis of slope stability has, in the case of a progressive failure, to assess the effects of the uncertainties of strength parameters and the reliability index. The main conclusions from this paper are:

- The traditional design of slopes that uses safety factors can be misleading. The slopes designed with large safety factors are not exempt from failure risk.
- Reliability analysis provides a logical framework to introduce uncertainties in the design process and improve the perception of risk in traditional design methods.
- The limit state function considered in the probabilistic analysis of slope stability depends more on post-peak strength parameters than in peak strength parameters.
- The integration of the residual strength parameters of ( $C_r$ ,  $\phi_r$ ) in the proposed model has better simulation results than the progressive failure process of slopes.

- The reliability analysis of slices allowed locating the area triggering a failure and observing the direction of failure propagation.
- The existence of a subsystem which limits the evaluation of the system behavior to that of the subsystem.
- The residual strength parameters of the soil are the most dominant variables and the order of influence of the other variables depends on the choice of distribution.
- The Monte Carlo Simulation method is simple and robust, but requires a large number of simulations.
- The difference between the results obtained by the reliability methods FORM, FOSM, and MCS decreases with increasing number of slices in failure.
- The probabilistic analysis is a tool for decision support under uncertainty. It helps the problem's structure and guides the engineer in his judgment.

## REFERENCES

- [1] Y. M. Cheng and C. K. Lau, *New Methods and Insight*. London, UK: CRC Press, 2008, <https://doi.org/10.4324/9780203927953>.
- [2] K. T. Law and P. Lumb, "A limit equilibrium analysis of progressive failure in the stability of slopes," *Canadian Geotechnical Journal*, vol. 15, no. 1, pp. 113–122, Feb. 1978, <https://doi.org/10.1139/t78-009>.
- [3] L. Hu, A. Takahashi, and K. Kasama, "Effect of spatial variability on stability and failure mechanisms of 3D slope using random limit equilibrium method," *Soils and Foundations*, vol. 62, no. 6, Dec. 2022, Art. no. 101225, <https://doi.org/10.1016/j.sandf.2022.101225>.
- [4] M. Havaej, D. Stead, E. Eberhardt, and B. R. Fisher, "Characterization of bi-planar and ploughing failure mechanisms in footwall slopes using numerical modelling," *Engineering Geology*, vol. 178, pp. 109–120, Aug. 2014, <https://doi.org/10.1016/j.enggeo.2014.06.003>.
- [5] G. Firpo, R. Salvini, M. Francioni, and P. G. Ranjith, "Use of Digital Terrestrial Photogrammetry in rocky slope stability analysis by Distinct Elements Numerical Methods," *International Journal of Rock Mechanics and Mining Sciences*, vol. 48, no. 7, pp. 1045–1054, Oct. 2011, <https://doi.org/10.1016/j.ijrmm.2011.07.007>.
- [6] D. A. Mangnejo, S. J. Oad, S. A. Kalhor, S. Ahmed, F. H. Laghari, and Z. A. Siyal, "Numerical Analysis of Soil Slope Stabilization by Soil Nailing Technique," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4469–4473, Aug. 2019, <https://doi.org/10.48084/etasr.2859>.
- [7] R. b. Gilbert and R. j. Byrne, "Strain-Softening Behavior of Waste Containment System Interfaces," *Geosynthetics International*, vol. 3, no. 2, pp. 181–203, Jan. 1996, <https://doi.org/10.1680/gein.3.0059>.
- [8] J. T. Christian, C. C. Ladd, and G. B. Baecher, "Reliability Applied to Slope Stability Analysis," *Journal of Geotechnical Engineering*, Vol. 120, pp. 2180–2207, 1994, <https://doi.org/10.1016/j.compgeo.2017.07.020>.
- [9] D. A. Crum *et al.*, "Search Algorithm for Minimum Reliability Index of Earth Slopes," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 127, no. 2, pp. 194–200, Feb. 2001, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2001\)127:2\(194.2\)](https://doi.org/10.1061/(ASCE)1090-0241(2001)127:2(194.2)).
- [10] D. V. Griffiths and G. A. Fenton, "Probabilistic Slope Stability Analysis by Finite Elements," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 130, no. 5, pp. 507–518, May 2004, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2004\)130:5\(507\)](https://doi.org/10.1061/(ASCE)1090-0241(2004)130:5(507)).
- [11] H. El-Ramly, N. R. Morgenstern, and D. M. Cruden, "Probabilistic slope stability analysis for practice," *Canadian Geotechnical Journal*, vol. 39, no. 3, pp. 665–683, Jun. 2002, <https://doi.org/10.1139/t02-034>.
- [12] K. Winkelmann, L. Zabuski, J. Przewłócki, and J. Górski, "Reliability-Based Stability Analysis of a Baltic Cliff by the Combined Response Surface Method," *Geotechnical and Geological Engineering*, vol. 38, no. 5, pp. 5549–5563, Oct. 2020, <https://doi.org/10.1007/s10706-020-01384-5>.
- [13] S.-H. Jiang, D.-Q. Li, Z.-J. Cao, C.-B. Zhou, and K.-K. Phoon, "Efficient System Reliability Analysis of Slope Stability in Spatially Variable Soils Using Monte Carlo Simulation," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 141, no. 2, Feb. 2015, Art. no. 04014096, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0001227](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001227).
- [14] G. Baecher and J. Christian, *Reliability and Statistics in Geotechnical Engineering*, 1st ed. Chichester, West Sussex, UK; Hoboken, NJ, USA: Wiley, 2003.
- [15] H. AM and N. Lind, "An Exact and Invariant First Order Reliability Format," *Journal of Engineering Mechanics*, vol. 100, Jan. 1974.
- [16] A. W. Skempton, "Long-Term Stability of Clay Slopes," *Géotechnique*, vol. 14, no. 2, pp. 77–102, Jun. 1964, <https://doi.org/10.1680/geot.1964.14.2.77>.
- [17] R. N. Chowdhury and D. W. Xu, "Reliability index for slope stability assessment—two methods compared," *Reliability Engineering & System Safety*, vol. 37, no. 2, pp. 99–108, Jan. 1992, [https://doi.org/10.1016/0951-8320\(92\)90002-3](https://doi.org/10.1016/0951-8320(92)90002-3).
- [18] R. N. Chowdhury, W. H. Tang, and I. Sidi, "Reliability model of progressive slope failure," *Géotechnique*, Vol. 37, pp. 467–481, 1987, <https://doi.org/10.1680/geot.1987.37.4.467>.
- [19] S. S. Kar and L. B. Roy, "Reliability Analysis of a Finite Slope Considering the Effects of Soil Uncertainty," *International Journal of Performability Engineering*, vol. 17, no. 5, May 2021, Art. no. 473, <https://doi.org/10.23940/IJPE.21.05.P7.473483>.
- [20] Z. Cao, Y. Wang, and D. Li, "Probabilistic Approaches for Geotechnical Site Characterization and Slope Stability Analysis," 1st ed. New York, NY, USA: Springer, 2016.
- [21] N. M. Okasha, "Reliability-Based Design Optimization of Trusses with Linked-Discrete Design Variables using the Improved Firefly Algorithm," *Engineering, Technology & Applied Science Research*, vol. 6, no. 2, pp. 964–971, Apr. 2016, <https://doi.org/10.48084/etasr.675>.
- [22] A.-R. Yarahmadi-Bafghi, "La méthode des groupes-clef probabiliste," Ph.D. dissertation, Institut National Polytechnique de Lorraine, Nancy, France, 2003.
- [23] G. M. Filz, J. J. B. Esterhuizen, and J. M. Duncan, "Progressive Failure of Lined Waste Impoundments," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 127, no. 10, pp. 841–848, Oct. 2001, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2001\)127:10\(841\)](https://doi.org/10.1061/(ASCE)1090-0241(2001)127:10(841)).
- [24] R. N. Chowdhury and D. W. Xu, "Geotechnical system reliability of slopes," *Reliability Engineering & System Safety*, vol. 47, no. 3, pp. 141–151, Jan. 1995, [https://doi.org/10.1016/0951-8320\(94\)00063-T](https://doi.org/10.1016/0951-8320(94)00063-T).
- [25] A. I. Husein Malkawi, W. F. Hassan, and F. A. Abdulla, "Uncertainty and reliability analysis applied to slope stability," *Structural Safety*, vol. 22, no. 2, pp. 161–187, Jun. 2000, [https://doi.org/10.1016/S0167-4730\(00\)00006-0](https://doi.org/10.1016/S0167-4730(00)00006-0).
- [26] R. Y. Liang, O. K. Nusier, and A. H. Malkawi, "A reliability based approach for evaluating the slope stability of embankment dams," *Engineering Geology*, vol. 54, no. 3, pp. 271–285, Oct. 1999, [https://doi.org/10.1016/S0013-7952\(99\)00017-4](https://doi.org/10.1016/S0013-7952(99)00017-4).
- [27] R. Rackwitz, "Reliability analysis—a review and some perspectives," *Structural Safety*, vol. 23, no. 4, pp. 365–395, Oct. 2001, [https://doi.org/10.1016/S0167-4730\(02\)00009-7](https://doi.org/10.1016/S0167-4730(02)00009-7).
- [28] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, 2nd ed. Hoboken, N.J, USA: Wiley-Interscience, 2007.
- [29] S. S. Kar and L. B. Roy, "Probabilistic Based Reliability Slope Stability Analysis Using FOSM, FORM, and MCS," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8236–8240, Apr. 2022, <https://doi.org/10.48084/etasr.4689>.