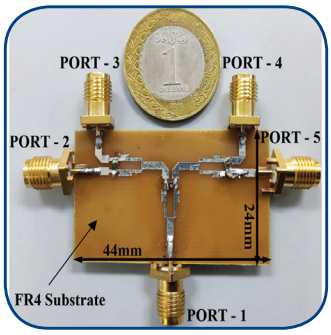
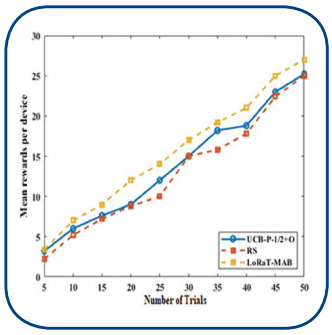
Volume 15, Number 9, 2024 ISSN 1847-6996

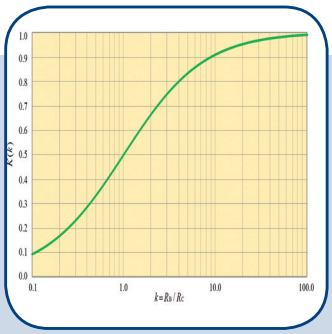
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY OSIJEK

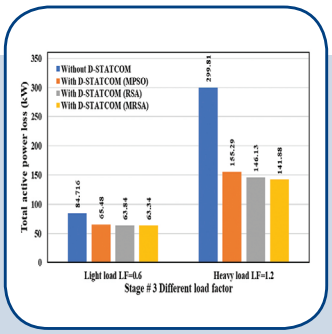
IJECES International Journal of Electrical and Computer Engineering Systems

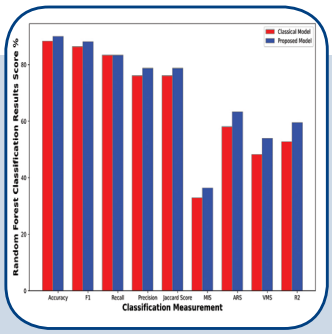
International Journal of Electrical and Computer Engineering Systems











ISSN: 1847-6996

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Volume 15, Number 9, 2024 | Pages 719 - 817

The International Journal of Electrical and Computer Engineering Systems is published with the financial support of the Ministry of Science and Education of the Republic of Croatia

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777, Croatian Postal Bank

Journal is referred in:

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003
Published: quarterly
Circulation: 300

IJECES online

https://ijeces.ferit.hr

TABLE OF CONTENTS

Energy Efficiency of the Passive Hybrid Battery Supercapacitor System: An Analytical Approach

Original Scientific Paper

Dalibor Buljić*
J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Kneza Trpimira 2B, HR-31000 Osijek, Croatia
dalibor.buljic@ferit.hr

Tomislav Barić
J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Kneza Trpimira 2B, HR-31000 Osijek, Croatia
tomislav.baric@ferit.hr

Hrvoje Glavaš
J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Kneza Trpimira 2B, HR-31000 Osijek, Croatia
hrvoje.glavas@ferit.hr

*Corresponding author

Abstract *– Under dynamic load, passive hybrid battery supercapacitor systems (HBSSs) have numerous advantages compared to stand-alone battery systems (SABS). The highlighted benefits include a lower voltage drop at the HBSS terminals, lower current and thermal stress on the battery, and higher energy efficiency. A novel set of analytical expressions is introduced for the analysis of energy efficiency of the passive HBSS in relation to the SABS. Using these expressions, energy efficiency of the passive HBSS was analyzed for different dynamic load parameters. Compared to existing approaches, the proposed approach that uses the introduced analytical expressions provides a better insight into the impact of battery and supercapacitor parameters on the energy efficiency of the passive HBSS in relation to the SABS. Additionally, the proposed approach is suitable for quick determination of HBSS behavior and energy efficiency under dynamic load. This contributes to the advancement of the sizing of HBSS components and optimization practices, ultimately leading to improved energy use. The analysis showed that in cases where the dynamic component of the load current is a very small part of the total load current, or slowly varying load currents, the passive HBSS in relation to the SABS has no advantages in terms of energy efficiency. The analysis showed that the HBSS is more suitable than the SABS in applications where the dynamic component of the load current is significant. The findings also indicate that there is an interval of the battery and supercapacitor parameters which is optimal for the sizing of the passive HBSS from a techno-economic aspect.*

Keywords: *analytical, battery, dynamic, energy efficiency, hybrid battery supercapacitor system, supercapacitor*

Received: June 4, 2024; Received in revised form: August 14, 2024; Accepted: August 19, 2024

1. INTRODUCTION

From the perspective of energy efficiency, energy sources such as photovoltaics (PV), batteries, wind turbines, diesel generators, supercapacitors, etc. have significantly different energy efficiency for dynamic and static loads. In general, system losses are higher with dynamic loads compared to static loads. Additionally, when supplying dynamic loads, energy sources experience increased current and thermal stresses and higher voltage drops. By hybridizing different energy sources, the desired characteristics of individual energy sources can be exploited in certain applications [1-3]. By applying optimization techniques, adequate power manage-

ment, the optimal configuration of such systems, dynamics and energy flows can be achieved [3, 4]. It is also possible to reduce current and thermal stresses of components within hybrid systems as well as voltage drops at the terminals of the hybrid system. One such hybrid system, the so-called passive hybrid battery and supercapacitor system, is described in detail in this paper.

As energy storage devices, batteries [5, 6] and supercapacitors (SCs) [7-12] are particularly suitable for static and dynamic loads, respectively. Despite the fact that supercapacitors are suitable for dynamic loads, they have low gravimetric density of stored energy compared to batteries (Fig. 1, a compilation from [13-15]). As energy storage devices, supercapacitors have approximately ten times higher gravimetric power density but ten times lower gravimetric energy density compared to batteries. By hybridizing these into a single energy storage system, the benefits of each component can be utilized. Such systems in which batteries and supercapacitors are hybridized are called battery-supercapacitor hybrid energy storage systems (HESS) (Fig. 2. A compilation from [16-22]). From a topological perspective, the simplest form of such an energy storage system is called a passive hybrid battery supercapacitor system (a passive HBSS), or a passive hybrid battery-supercapacitor energy storage systems (a passive HESS).

The basic difference between passive HBSS and all other HBSS topologies is that all other HBSS topologies use power electronics circuits to control energy flows. The use of power electronics circuits, such as bidirectional DC/DC converters, enables better adaptation and optimization of energy flows in variable operating conditions.

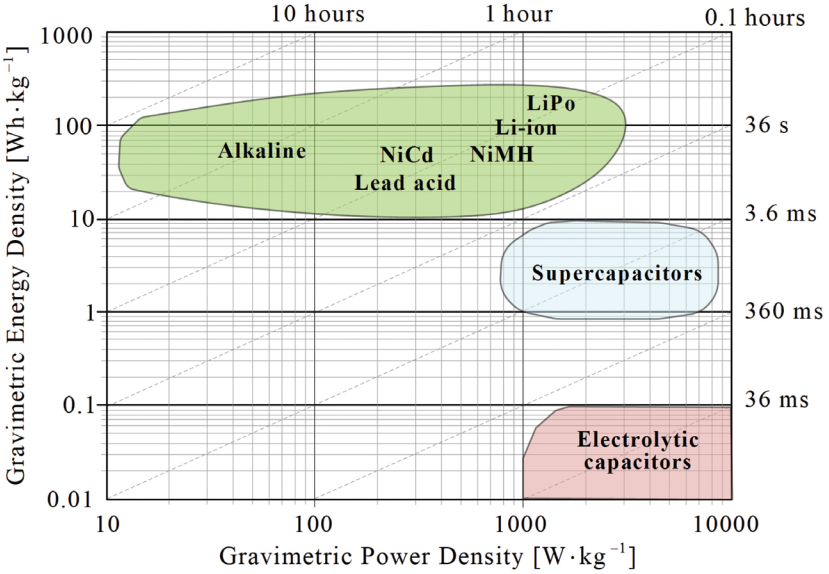


Fig. 1. Ragone plot with charge/discharge time for arbitrarily selected energy storage technologies

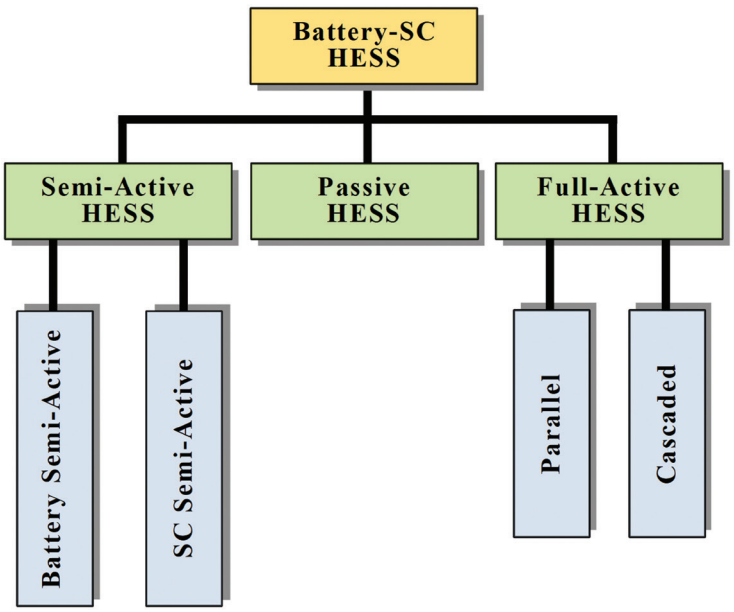


Fig. 2. Common classification of hybrid battery supercapacitor topologies

In a passive H BSS, the battery and the supercapacitor are connected in parallel, i.e. to common DC buses (Fig. 3. A compilation from [16-22]).

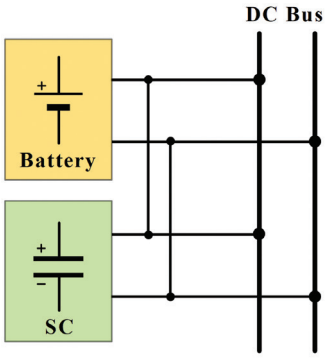


Fig. 3. Passive hybrid battery-supercapacitor topology

Compared to a stand-alone battery energy storage system, a hybrid battery supercapacitor system, or a "stand-alone battery system" (SABS), has several prominent advantages:

- Reduced battery current: In case of dynamic loads, the load current in the HBSS is distributed between HBSS components, i.e. between the battery and the supercapacitor. Therefore, the battery current inside the HBSS is reduced in comparison to the SABS [13, 23-28].
- Reduced battery temperature and thermal stress: Joule heating of the battery changes with the square of the current flowing through the battery. The released heat in the battery raises the temperature of the battery. Under dynamic load, during a significant increase in current, due to the fact that the released heat changes with the square of the current flowing through the battery, there is a sudden change in the temperature of the battery. Both a significant increase and a sudden change in battery temperature cause thermal stress on the battery. Thermal stress degrades battery performance and contributes to reduced battery lifespan. Given that the battery current inside the HBSS is reduced compared to the SABS, the temperature and thermal stress of the battery inside the HBSS are also reduced compared to the SABS [27-30].
- Increased energy efficiency and energy absorption: Regarding energy efficiency, the HBSS has higher energy efficiency under dynamic load compared to the SABS [13, 31, 32]. Furthermore, in relation to the stand-alone battery, the HBSS has an increased ability to absorb the energy delivered by the load during braking of the electric motor drive (recuperation).
- Less pronounced voltage sags: During sudden increases in load currents, the voltage at common DC buses, i.e. at the hybrid battery supercapacitor system terminals, experiences significantly less pronounced voltage sags compared to the SABS [13, 15, 28, 32].
- This paper focuses on energy efficiency, specifically analyzing the energy efficiency of the passive HBSS using analytical expressions. For this purpose, analytical expressions will be derived that enable analysis of the energy efficiency of the passive HBSS. First, the derivation of expressions for voltages and currents during dynamic loading of the passive HBSS will be presented. Using the obtained analytical expressions, the derivation of the expressions for losses in the HBSS will be presented for the scenario in which the load is dynamic. Then, the losses in the HBSS will be compared with the losses that would occur in a SABS under the same dynamic load. The energy efficiency of the passive HBSS was also analyzed for different parameters of the dynamic load. The contributions and findings of this paper can be summarized as follows:
- A novel set of analytical expressions is introduced for the analysis of the energy efficiency of the passive HBSS in relation to the SABS.
- Derived expressions are expressed through two additional parameters (K and k) that are related to the HBSS circuit parameters. This resulted in a form of equations that is suitable for rapid identification of the influence of HBSS parameters on current and voltage conditions and energy efficiency under dynamic load.
- By providing insights into the influence of HBSS parameters on current and voltage conditions and energy efficiency under dynamic load, this paper contributes to the advancement of sizing the HBSS components and optimization practices, ultimately leading to improved energy use.
- The proposed approach to analyzing the energy efficiency of the HBSS using the introduced analytical expressions resulted in the finding that there is an interval of the parameter k which is optimal for sizing the HBSS from a techno-economic aspect.

2. PASSIVE HBSS MODEL

In order for the derivation of analytical expressions to be possible, the HBSS model should be reduced to only necessary circuit elements. If the consideration is limited to very specific conditions under which the passive HBSS operates, then a straightforward model can effectively represent both the passive HBSS and the load. Such approach is often used in the performance analysis of hybrid battery supercapacitor systems [14, 22-24, 26, 27, 31-34]. We have already employed this approach in [28], and the same model will be utilized in this paper. For the purpose of presentation consistency, the main steps in modeling and derivation of analytical expressions are repeated from [28]. After repeating the steps from [28] in obtaining the analytical expressions for the HBSS currents, a derivation of analytical expressions is given for the energy efficiency specific to this paper.

Assuming that the analysis is carried out in a time interval during which the battery has not discharged significantly, the electromotive force of the battery (denoted by E), i.e. the open circuit voltage of the battery, can be considered constant.

Additionally, the battery equivalent series resistance (R_B) can be considered constant and equal for both battery charging and discharging processes. According to the above, the battery can be modeled as a real voltage source (Fig. 4 [13, 15, 22, 28]). Given that phenomena such as self-discharge and dielectric absorption are not expressed during the fast charge and discharge of a supercapacitor, the supercapacitor can be represented by a series connection of the SC capacitance (C) and the equivalent series resistance of the supercapacitor (R_C) (Fig. 4 [13, 15, 22, 28]).

According to the previously adopted models for battery and supercapacitor, the passive HBSS behaves as a first-order system during transients. Consequently, it has already been established that the transient will be described by a first-order differential equation. To further simplify the model of the entire system, it is essential to select an appropriate load model.

For this purpose, the dynamic load can be modeled with a current sink. The load current should have such a waveform that encompasses both steady and dynamic load time intervals. This can be achieved by representing the load current as a periodic pulse train that also contains a constant part (Fig. 5 [13, 15, 28]):

$$i_L = I_0 + \sum_{k=0}^{N} \left(I_p - I_0\right)\left(u(t-kT) - u(t-T_p-kT)\right), \quad (1)$$

where $u(t\text{-}kT)$ and $u(t\text{-}T_p\text{-}kT)$ are Heaviside unit step functions, I_0 is the amplitude of the constant part of the load, I_p is the amplitude of rectangular pulses, T_p is the duration of rectangular pulses, and T is the period of the waveform.

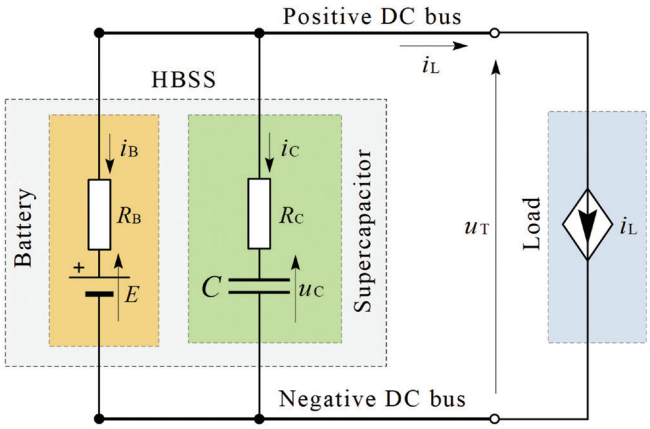


Fig. 4. Passive HBSS modeled as a first-order circuit

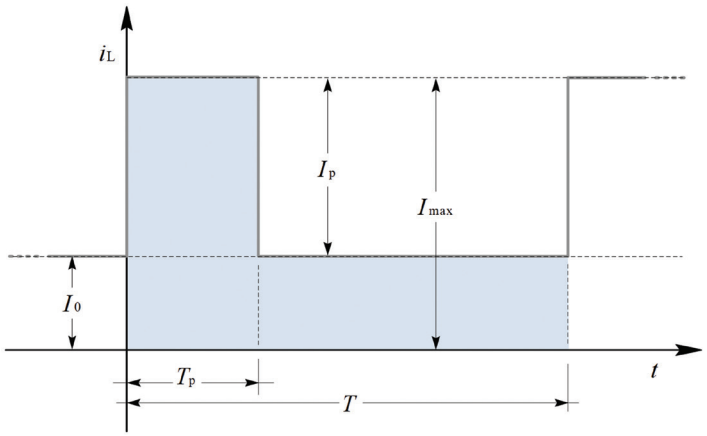


Fig. 5. Load current and relevant quantities

3. MATHEMATICAL DESCRIPTION OF TRANSIENTS

Based on the previously presented HBSS and the load model, the procedure used to obtain analytical expressions for battery and supercapacitor currents during transients will be described in this section.

According to Fig. 4, for each t there holds [28]:

$$i_L = -i_B - i_C, \quad (2)$$

$$E + i_B \cdot R_B - i_C \cdot R_C - u_C = 0, \quad (3)$$

$$i_C = C \cdot \frac{du_C}{dt}. \quad (4)$$

Combining the previous three expressions yields [28]:

$$(R_B + R_C)\cdot C\cdot\frac{du_C}{dt}+u_C = E - i_L\cdot R_B. \quad (5)$$

A first-order differential equation is obtained, which is usually written as follows [28]:

$$\tau\cdot\frac{du_C}{dt}+u_C = E - i_L\cdot R_B, \quad (6)$$

where the time constant is determined by the expression [28]:

$$\tau = (R_B + R_C)\cdot C. \quad (7)$$

Differential equation (6) holds for both SC transients (i.e. discharging and charging transients). The same time constant (7) applies to both transients. Depending on whether the SC transient is charging or discharging, the load current will vary in the differential equation (6).

For the SC discharge transient process ($0_+ \le t \le Tp$), the load current is

$$i_L = I_0 + I_p. \quad (8)$$

For the SC charging transient process ($Tp + \le t \le T$), the load current is

$$i_L = I_0. \quad (9)$$

To clearly distinguish between the voltages and currents in the HBSS associated with the charging and discharging processes of the supercapacitors, it is desirable to introduce the following notation. The voltages and currents related to the discharge process of the supercapacitor will be denoted with the index one, while those related to the charging process will be denoted with the index two.

According to the adopted notation, voltage u_C and currents i_L and i_B can be written as follows:

$$u_C = \begin{cases} u_{C1} \text{ for } 0_+ \le t \le T_{p-}, \\ u_{C2} \text{ for } T_{p+} \le t \le T. \end{cases} \quad (10)$$

$$i_L = \begin{cases} i_{L1} = I_0 + I_p \text{ for } 0_+ \le t \le T_{p-}, \\ i_{L2} = I_0 \quad \text{for } T_{p+} \le t \le T. \end{cases} \quad (11)$$

$$i_B = \begin{cases} i_{B1} \text{ for } 0_+ \le t \le T_{p-}, \\ i_{B2} \text{ for } T_{p+} \le t \le T. \end{cases} \quad (12)$$

3.1. SUPERCAPACITOR DISCHARGE PROCESS

The transient phenomenon of the supercapacitor discharge begins with the occurrence of a current pulse at $t=0_+$. For the SC discharge process ($0_+ \le t \le T_p$), the following holds:

$$\tau\cdot\frac{du_{C1}}{dt}+u_{C1} = E - (I_0 + I_p)\cdot R_B. \quad (13)$$

Applying the Laplace transform to Eq. (13) gives

$$\tau\cdot\left[s\cdot U_{C1}(s) - u_{C1}(0_-)\right] + U_{C1}(s) = \frac{E - (I_0 + I_p)\cdot R_B}{s}. \quad (14)$$

A minor rearrangement of the previous expression yields

$$U_{C1}(s) = \frac{u_C(0_-)}{s+\frac{1}{\tau}} + \frac{1}{T}\cdot\frac{1}{s}\cdot\frac{\left(E - (I_0 + I_p)\cdot R_B\right)}{s+\frac{1}{\tau}}. \quad (15)$$

According to the law of commutation [35], the voltage across the capacitor remains unchanged during commutation.

$$u_C 1(0_+) = u_{C1}(0_-) = E - i_L(0_-)\cdot R_B. \quad (16)$$

Since $i_L(0_-)=I_0$, it follows that

$$u_{C1}(0_-) = u_{C1}(0_+) = E - I_0\cdot R_B = U_0, \quad (17)$$

where U_0 indicates the steady-state voltage at the terminals of the HBSS. It is also convenient to define the theoretical voltage drop at the battery terminals due to the dynamic (peak) current I_p in the scenario when the SC is absent (i.e. as it would be in the SABS configuration):

$$\Delta U_p = I_p\cdot R_B. \quad (18)$$

Taking into account (17) and (18), expression (15) can be further simplified to the following form:

$$U_{C1}(s) = \frac{U_0}{s+\frac{1}{\tau}} + \frac{1}{\tau}\cdot\frac{1}{s}\cdot\frac{U_0 - \Delta U_p}{s+\frac{1}{\tau}}. \quad (19)$$

Applying the inverse Laplace transform to (19) gives

$$u_{C1} = U_0 - \Delta U_p\cdot\left(1 - e^{-\frac{t}{\tau}}\right). \quad (20)$$

Once the expression for the voltage u_{C1} is known, it is possible to determine the current i_{C1}. Substituting expression (20) into (4) gives

$$i_{C1} = C\cdot\frac{du_{C1}}{dt} = -\Delta U_p\cdot\frac{C}{\tau}e^{-\frac{t}{\tau}} = -\frac{\Delta U_p}{R_B + R_C}e^{-\frac{t}{\tau}}. \quad (21)$$

Taking into account (18) yields

$$i_{C1} = -I_p\frac{R_B}{R_B + R_C}e^{-\frac{t}{\tau}}. \quad (22)$$

To make further expressions concise, two auxiliary parameters, k and K, can be introduced [28]:

$$k = \frac{R_B}{R_C}, \quad (23)$$

$$K = K(k) = \frac{R_B}{R_B + R_C} = \frac{k}{1+k}. \quad (24)$$

Graphical representation of expression (24), i.e. K vs. k, is shown in Figure 6 [28].

Taking into account (24), expression (22) can be written as follows:

$$i_{C1} = -I_p\cdot K\cdot e^{-\frac{t}{\tau}}. \quad (25)$$

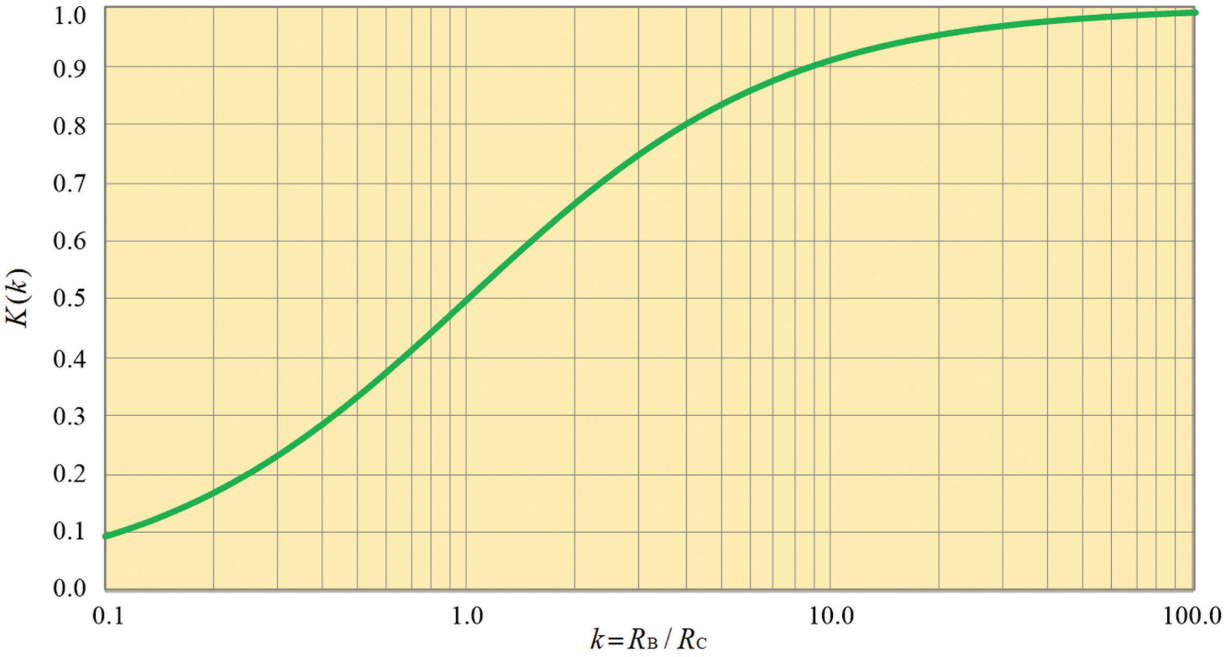


Fig. 6. Parameter K vs. parameter k

3.2. SUPERCAPACITOR CHARGING PROCESS

The transient phenomenon of supercapacitor charging begins when the current pulse ends at $t=Tp_+$. For the SC charging transient process ($T_p \le t < T$), there holds:

$$\tau \cdot \frac{\mathrm{d}u_{C2}}{\mathrm{d}t} + u_{C2} = E - I_0 \cdot R_B . \tag{26}$$

Taking into account that $E\text{-}I_0 \cdot R_B = U_0$, expression (26) turns into a more concise form:

$$\tau \cdot \frac{\mathrm{d}u_{C2}}{\mathrm{d}t} + u_{C2} = U_0 . \tag{27}$$

By applying the previously described procedure for solving the differential equation, the following is obtained:

$$u_{C2} = u_{C2}(T_{p-}) + \left(U_0 - u_{C2}(T_{p-})\right) \cdot \left(1 - \mathrm{e}^{-\frac{t-T_p}{\tau}}\right). \tag{28}$$

The initial condition $u_{C2}(T_{p-})$ is determined by applying the law of commutation. According to the law of commutation [35], the voltage across the capacitor remains unchanged during commutation:

$$u_{C2}(T_{p-}) = u_{C2}(T_{p+}) = u_{C1}(T_p) = U_0 - \Delta U_p \cdot \left(1 - \mathrm{e}^{-\frac{T_p}{\tau}}\right). \tag{29}$$

Once the expression for the voltage u_{C2} is known, it is possible to determine the current i_{C2}. Substituting expression (28) into (4) gives

$$i_{C2} = C \cdot \frac{\mathrm{d}u_{C2}}{\mathrm{d}t} = \left(U_0 - u_{C2}(T_p)\right) \cdot \frac{C}{\tau} \cdot \mathrm{e}^{-\frac{t-T_p}{\tau}} . \tag{30}$$

Substituting (7) into (30) gives

$$i_{C2} = \frac{U_0 - u_{C2}(T_p)}{R_B + R_C} \cdot \mathrm{e}^{-\frac{t-T_p}{\tau}} . \tag{31}$$

As was done for the current i_{C1}, it is beneficial to describe the current i_{C2} using the parameter K. From (24), it follows that

$$\frac{1}{R_B + R_C} = \frac{K}{R_B} . \tag{32}$$

Combining (31) and (32) gives

$$i_{C2} = \frac{U_0 - u_{C2}(T_p)}{R_B + R_C} \cdot \mathrm{e}^{-\frac{t-T_p}{T}} = \frac{K}{R_B} \cdot \left(U_0 - u_{C2}(T_p)\right) \cdot \mathrm{e}^{-\frac{t-T_p}{T}} . \tag{33}$$

Combining (29) and (33) gives

$$i_{C2} = \frac{K}{R_B} \cdot \Delta U_p \cdot \left(1 - \mathrm{e}^{-\frac{T_p}{\tau}}\right) \cdot \mathrm{e}^{-\frac{t-T_p}{T}} . \tag{34}$$

By respecting the definition of ΔU_p (expression (18)), expression (34) can be written as follows:

$$i_{C2} = I_p \cdot K \cdot \left(1 - \mathrm{e}^{-\frac{T_p}{T}}\right) \cdot \mathrm{e}^{-\frac{t-T_p}{T}} \tag{35}$$

3.3. BATTERY CURRENT IN THE HBSS

The battery current in the HBSS during both SC transients (discharging and charging) is determined from expression (2):

$$i_B = -i_L - i_C . \tag{36}$$

Taking into account the previously introduced notation, (36) and (11) yield the following:

$$i_B = \begin{cases} -(I_0 + I_p) - i_{C1} & \text{for } 0_+ \le t \le T_{p-}, \\ -I_0 - i_{C2} & \text{for } T_{p+} \le t \le T. \end{cases} \tag{37}$$

Inserting (25) and (35) into (37) gives

$$i_{B1} = -I_0 - I_p \cdot \left(1 - K \cdot \mathrm{e}^{-\frac{t}{\tau}}\right), \tag{38}$$

$$i_{B2} = -I_0 - I_p \cdot K \cdot \left(1 - e^{-\frac{T_p}{\tau}}\right) \cdot e^{-\frac{t-T_p}{\tau}}. \quad (39)$$

The expressions obtained for the supercapacitor and battery currents allow for an analysis of how the presence of the supercapacitor affects the redistribution of the current pulse between the supercapacitor and the battery in the passive HBSS. Specifically, this analysis focuses on the reduction of the battery's current stress in the passive HBSS. Battery current reduction and rectangular pulse current redistribution between the supercapacitor and the battery can be easily analyzed through the introduced parameter K [28]. For example, if the supercapacitor and the battery have equal equivalent series resistances, then k=1 follows according to (23). According to Fig. 6 for parameter k=1, the parameter K is equal to 0.5. According to expressions (25) and (38), at the moment a rectangular current pulse (t=0$_+$) appears, the pulse current I_p is evenly distributed between the supercapacitor and the battery. The supercapacitor absorbs half of the current pulse, while the battery absorbs the other half. Once all currents (i_{C1}, i_{C2}, i_{B1} and i_{B2}) are expressed through the parameter K, their waveform can be graphically represented qualitatively as a function of the parameter K (Fig. 7 [28]). The waveforms of battery and supercapacitor currents shown in Fig. 7 [28] are all consistent with the results given in the literature [14, 15, 26, 34, 36-38].

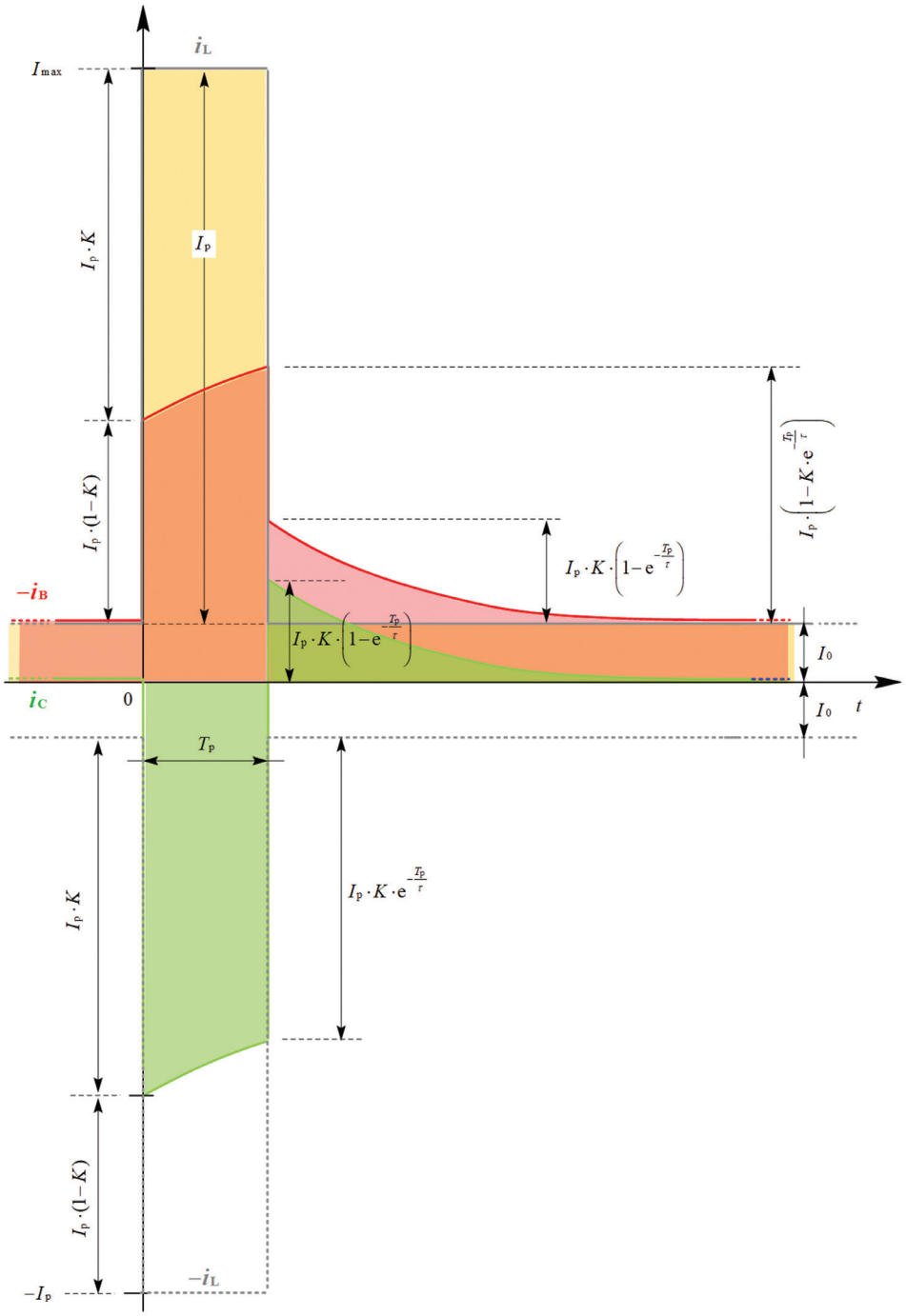


Fig. 7. Waveforms of the battery, load and SC current expressed through parameter K

If the parameter k is greater than one, then a larger portion of the rectangular pulse current is initially absorbed by the supercapacitor. More precisely, at $t=0_+$ the supercapacitor absorbs part of the pulse current, which is $Ip \cdot K$, and the battery absorbs the rest of the pulse current, which is $I_p \cdot (1\text{-}K)$ (see Fig. 7.). As long as the rectangular pulse current lasts, the current of the supercapacitor decreases, and the current of the batteries increases (see Fig. 7). As shown in Fig. 7, at the end of the rectangular pulse current, the direction of the supercapacitor current changes due to the supercapacitor charging process that occurs. The charging current of the supercapacitor is provided by the battery. Therefore, after the end of the current pulse, the battery is loaded with a current equal to the sum of the constant part of the load current I_0 and the supercapacitor charging current i_{C2}. The supercapacitor charging process takes approximately five time constants (Eq. 7). Therefore, when sizing the HBSS component, such HBSS parameters should be selected that the supercapacitor has enough time to charge before a new rectangular pulse current occurs.

4. ENERGY EFFICIENCY

To demonstrate that the HBSS is more energy efficient compared to the SABS under dynamic load, it is necessary to know the losses within both the HBSS and the SABS under the same operating conditions. The energy efficiency of the HBSS compared to a battery-only system (SABS) can be analyzed through a loss ratio:

$$\frac{E_{\mathrm{HBSS}}(0,T)}{E_{\mathrm{SABS}}(0,T)} = \frac{E_{\mathrm{HBSS}}(0,T_{\mathrm{p}}) + E_{\mathrm{HBSS}}(T_{\mathrm{p}},T)}{E_{\mathrm{SABS}}(0,T_{\mathrm{p}}) + E_{\mathrm{SABS}}(T_{\mathrm{p}},T)}, \tag{40}$$

where E_{HBSS} $(0,T)$ are losses in the *HBSS* during one period of the waveform T, i.e. total losses in the battery and the supercapacitor during one period, E_{SABS} $(0,T)$ are losses in the SABS during one period T.

For the sake of mathematical simplicity and to facilitate easier interpretation of the results, it is convenient to introduce the following substitutions:

$$\varepsilon = \frac{I_{\mathrm{p}}}{I_0}, \tag{41}$$

where I_p is the peak value of the rectangular pulse and I_0 is the constant part of the load current.

The ratio of the duration of the rectangular pulse to the time constant is given by

$$\alpha = \frac{T_{\mathrm{p}}}{\tau}. \tag{42}$$

where T_p is the pulse duration and τ is the time constant.

The ratio of the duration of the rectangular pulse to the period of the waveform, the so-called duty cycle of the pulsating (dynamic) part of the load current, is given by

$$\beta = \frac{T_{\mathrm{p}}}{T}, \tag{43}$$

where T_p is the pulse duration and T is the period of the waveform.

Using the previously introduced substitutions, the battery current and the SC current can be written as

$$i_{\mathrm{B1}} = -I_0 \cdot \left[1 + \varepsilon \cdot \left(1 - K \cdot \mathrm{e}^{-\frac{t}{\tau}}\right)\right], \tag{44}$$

$$i_{\mathrm{B2}} = I_0 \cdot \left[1 + \varepsilon \cdot K \cdot \left(1 - \mathrm{e}^{-\frac{T_{\mathrm{p}}}{\tau}}\right) \cdot \mathrm{e}^{-\frac{t-T_{\mathrm{p}}}{\tau}}\right], \tag{45}$$

$$i_{\mathrm{C1}} = -I_0 \cdot \varepsilon \cdot K \cdot \mathrm{e}^{-\frac{t}{\tau}}, \tag{46}$$

$$i_{\mathrm{C2}} = I_0 \cdot \varepsilon \cdot K \cdot \left(1 - \mathrm{e}^{-\frac{T_{\mathrm{p}}}{T}}\right) \cdot \mathrm{e}^{-\frac{t-T_{\mathrm{p}}}{T}}. \tag{47}$$

4.1. LOSSES IN A STAND-ALONE BATTERY SYSTEM

Losses due to the load current in a SABS over one period can be determined as follows:

$$E_{\mathrm{SABS}}(0,T) = R_{\mathrm{B}} \cdot \int_0^T i_{\mathrm{L}}^2 \mathrm{d}t = R_{\mathrm{B}} \cdot I_{\mathrm{L}}^2, \tag{48}$$

where I_L^2 is given by the expression:

$$I_{\mathrm{L}}^2 = \int_0^{T_{\mathrm{p}}} i_{\mathrm{L1}}^2 \cdot \mathrm{d}t + \int_{T_{\mathrm{p}}}^T i_{\mathrm{L2}}^2 \cdot \mathrm{d}t = I_{\mathrm{L1}}^2 + I_{\mathrm{L2}}^2. \tag{49}$$

Integrals I_{L1}^2 and I_{L2}^2 are given by the expressions:

$$I_{\mathrm{L1}}^2 = \int_0^{T_{\mathrm{p}}} (I_0 + I_{\mathrm{p}})^2 \cdot \mathrm{d}t = (I_0 + I_{\mathrm{p}})^2 \cdot \int_0^{T_{\mathrm{p}}} \mathrm{d}t, \tag{50}$$

$$I_{\mathrm{L2}}^2 = \int_{T_{\mathrm{p}}}^T I_0^2 \cdot \mathrm{d}t = I_0^2 \cdot \int_{T_{\mathrm{p}}}^T \mathrm{d}t. \tag{51}$$

Using the previously introduced substitutions, the previous integrals can be written as follows:

$$I_{\mathrm{L1}}^2 = (I_0 + I_{\mathrm{p}})^2 \cdot T_{\mathrm{p}} = I_0^2 \cdot (1+\varepsilon)^2 \cdot T_{\mathrm{p}}, \tag{52}$$

$$I_{\mathrm{L2}}^2 = I_0^2 \cdot (T - T_{\mathrm{p}}) = I_0^2 \cdot T_{\mathrm{p}} \cdot \left(\frac{T}{T_{\mathrm{p}}} - 1\right) = I_0^2 \cdot T_{\mathrm{p}} \cdot \left(\frac{1}{\beta} - 1\right). \tag{53}$$

4.2. LOSSES IN AN HBSS SYSTEM

Losses in the HBSS due to the load current consist of losses in the battery and the supercapacitor. The total energy loss in the HBSS from the moment $t=0$ to the moment T is given by

$$E_{\mathrm{HBSS}}(0,T) = E_{\mathrm{B}}(0,T) + E_{\mathrm{SC}}(0,T), \tag{54}$$

where E_B $(0,T)$ and E_{SC} $(0,T)$ denote the losses in the battery during one period and the losses in the supercapacitor during one period, respectively.

The losses in the battery during one period can be determined as

$$E_{\mathrm{B}}(0,T) = E_{\mathrm{B}}(0,T_{\mathrm{p}}) + E_{\mathrm{B}}(T_{\mathrm{p}},T), \tag{55}$$

where

$$E_{\mathrm{B}}(0,T_{\mathrm{p}})=R_{\mathrm{B}}\cdot\int_{0}^{T_{\mathrm{p}}}i_{\mathrm{B1}}^{2}\mathrm{d}t=R_{\mathrm{B}}\cdot I_{\mathrm{B1}}^{2}, \tag{57}$$

The losses in the supercapacitor during one period can be determined as

$$E_{\mathrm{SC}}(0,T)=E_{\mathrm{SC}}(0,T_{\mathrm{p}})+E_{\mathrm{SC}}(T_{\mathrm{p}},T), \tag{58}$$

where

$$E_{\mathrm{SC}}(0,T_{\mathrm{p}})=R_{\mathrm{C}}\cdot\int_{0}^{T_{\mathrm{p}}}i_{\mathrm{C1}}^{2}\mathrm{d}t=R_{\mathrm{C}}\cdot I_{\mathrm{C1}}^{2}, \tag{59}$$

$$E_{\mathrm{SC}}(T_{\mathrm{p}},T)=R_{\mathrm{C}}\cdot\int_{T_{\mathrm{p}}}^{T}i_{\mathrm{C2}}^{2}\mathrm{d}t=R_{\mathrm{C}}\cdot I_{\mathrm{C2}}^{2}, \tag{60}$$

where I_{B1}^2, I_{B2}^2, I_{C1}^2 and I_{C2}^2 denote the following integrals:

$$I_{\mathrm{B1}}^{2}=I_{0}^{2}\cdot T_{\mathrm{p}}\cdot\left[\begin{array}{l}(1+\varepsilon)^{2}-2\cdot\varepsilon\cdot(1+\varepsilon)\cdot K\cdot\left(\frac{1}{\alpha}\right)\cdot\left(1-\mathrm{e}^{-\alpha}\right)+\\+\varepsilon^{2}\cdot K^{2}\cdot\left(\frac{1}{2}\frac{1}{\alpha}\right)\cdot\left(1-\mathrm{e}^{-2\alpha}\right)\end{array}\right], \tag{61}$$

$$I_{\mathrm{B2}}^{2}=I_{0}^{2}\cdot T_{\mathrm{p}}\cdot\left[\begin{array}{l}\left(\frac{1}{\beta}-1\right)+2\varepsilon K\alpha\cdot\left(1-\mathrm{e}^{-\alpha}\right)\cdot\left(1-\mathrm{e}^{-\frac{1-\beta}{\beta}\alpha}\right)+\\+(\varepsilon\cdot K)^{2}\cdot\left(\frac{\alpha}{2}\right)\cdot\left(1-\mathrm{e}^{-\alpha}\right)^{2}\cdot\left(1-\mathrm{e}^{-2\frac{1-\beta}{\beta}\alpha}\right)\end{array}\right], \tag{62}$$

$$I_{\mathrm{C1}}^{2}=I_{0}^{2}\cdot T_{\mathrm{p}}\cdot(\varepsilon\cdot K)^{2}\cdot\left(\frac{1}{2}\frac{1}{\alpha}\right)\cdot\left(1-\mathrm{e}^{-2\alpha}\right), \tag{63}$$

$$I_{\mathrm{C2}}^{2}=I_{0}^{2}\cdot T_{\mathrm{p}}\cdot(\varepsilon\cdot K)^{2}\cdot\left(1-\mathrm{e}^{-\alpha}\right)^{2}\cdot\left(\frac{1}{2}\frac{1}{\alpha}\right)\cdot\left(1-\mathrm{e}^{-2\alpha\frac{1-\beta}{\beta}}\right). \tag{64}$$

Up to this point, all losses have been expressed by the corresponding integrals. By inserting the corresponding integrals into (40), the expression for the energy efficiency can be expressed in the form:

$$\frac{E_{\mathrm{HBSS}}(0,T)}{E_{\mathrm{SABS}}(0,T)}=\frac{R_{\mathrm{B}}\cdot I_{\mathrm{B1}}^{2}+R_{\mathrm{B}}\cdot I_{\mathrm{B2}}^{2}+R_{\mathrm{C}}\cdot I_{\mathrm{C1}}^{2}+R_{\mathrm{C}}\cdot I_{\mathrm{C2}}^{2}}{R_{\mathrm{B}}\cdot I_{\mathrm{L1}}^{2}+R_{\mathrm{B}}\cdot I_{\mathrm{L2}}^{2}}. \tag{65}$$

According to (23), i.e. $k=R_B/R_C$, expression (65) can be simplified and written as

$$\frac{E_{\mathrm{HBSS}}(0,T)}{E_{\mathrm{SABS}}(0,T)}=\frac{I_{\mathrm{B1}}^{2}+I_{\mathrm{B2}}^{2}+\frac{1}{k}\left(I_{\mathrm{C1}}^{2}+I_{\mathrm{C2}}^{2}\right)}{I_{\mathrm{L1}}^{2}+I_{\mathrm{L2}}^{2}}. \tag{66}$$

The previous expression allows for the determination of the energy efficiency of the HBSS in relation to the SABS. For practical application of this expression, it is necessary to adhere to the initial assumption under which it was derived. To ensure its applicability, it is necessary to determine the validity limits of the expression for energy efficiency.

4.3. VALIDITY LIMITS OF THE EXPRESSION FOR ENERGY EFFICIENCY

The initial assumption before deriving the expressions was that the duration of the rectangular current pulse T_p is relatively short compared to the period T. This means that after the rectangular current pulse ends, the supercapacitor manages to recharge during the time interval T-T_p.

According to the above, and given that the time required for the supercapacitor to recharge to 99% of the voltage on its terminals in a steady state is equivalent to five time constants, the following condition follows:

$$T-T_{\mathrm{p}}\geq 5\cdot\tau, \tag{67}$$

or expressed through the introduced substitutions:

$$\alpha\geq 5\cdot\frac{\beta}{1-\beta}. \tag{68}$$

The previous expression enables the calculation of the range of valid parameters for determining energy efficiency (Table 1).

Table 1. Valid parameter ranges

$\beta\leq 1/10$	0.01	0.02	0.03	0.04	0.05
$\alpha\geq 5(\beta/(1-\beta))$	0.051	0.102	0.155	0.208	0.263

Table 1. Valid parameter ranges - Continued

$\beta\leq 1/10$	0.06	0.07	0.08	0.09	0.1
$\alpha\geq 5(\beta/(1-\beta))$	0.319	0.376	0.435	0.495	0.556

By adhering to the limits specified in Table 1, within which the expression for determining the energy efficiency of the HBSS in relation to the SABS holds, a numerical calculation was carried out (Figs. 8 and 9).

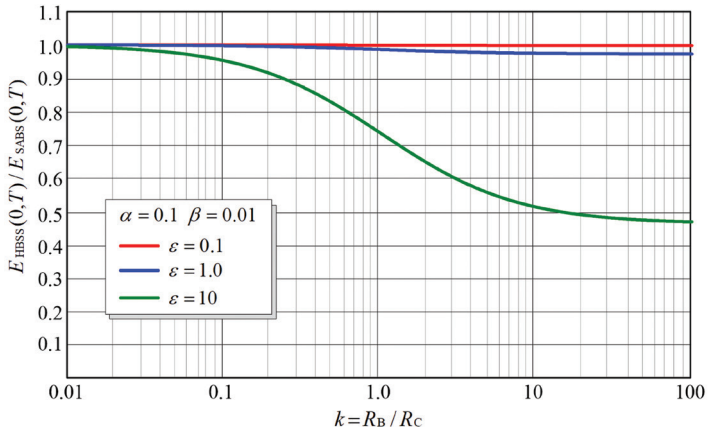


Fig. 8. Reduction of energy losses in the HBSS compared to a stand-alone battery system for α=0.1, β=0.01 and a variation of ε.

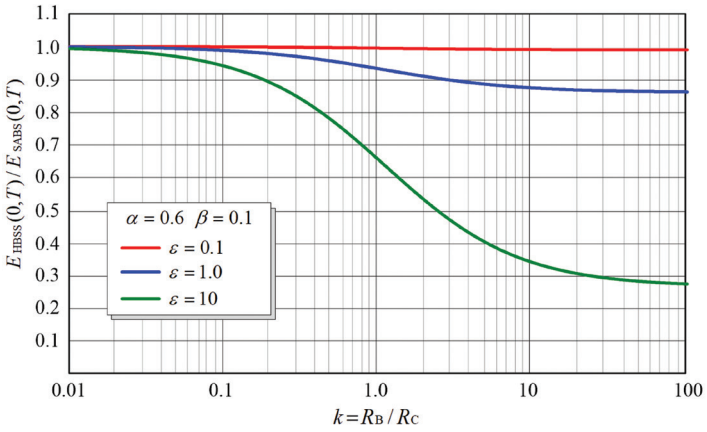


Fig. 9. Reduction of energy losses in the HBSS compared to a stand-alone battery system for α=0.6, β=0.1 and a variation of ε.

5. DISCUSION

In Section 3, analytical expressions are given that describe the battery and SC currents in the passive HBSS. The presented analytical expressions enable a physical explanation of the reasons why the HBSS is more energy efficient than the SABS under dynamic load. At the moment a rectangular current pulse appears, i.e. the dynamic component of the load current (see Fig. 7), the pulse current is distributed between the battery and the supercapacitor. Given that supercapacitors typically have a lower internal resistance than a battery (k>1), a larger part of the pulse current is absorbed by the SC (see Fig. 7). The current delivered by the SC to the load during the current pulse generates losses in the SC; however, since the internal resistance of the SC is lower than that of the battery, the losses in the HBSS are also reduced compared to the SABS.

For a more detailed analysis of the energy efficiency of the HBSS relative to the SABS, it is necessary to use the expressions provided in Section 4. This section includes analytical expressions that account for all losses in the HBSS. The expressions enable the numerical calculation of the energy efficiency of the HBSS relative to the SABS for various dynamic load parameters.

The numerical calculation of the energy efficiency of the HBSS relative to the SABS was performed over a range of parameter k from 0.01 to 100. Although the range of parameter k from 0.01 to 100 has no practical importance, as it would include extremes (a large battery and a small SC, a small battery and a large SC), the specified range was chosen for purely theoretical reasons. The selected range of parameter k enables recognition of the asymptotic values towards which the curves in figures converge.

The red curves in Figs. 8 and 9 refer to cases when there is a very small dynamic current component in the total load current, i.e. when $\varepsilon=I_p/I_0$=0.1. As expected, the ratio $E_{HBSS}(0,T)/E_{SABS}(0,T)$ is approximately one, i.e. the losses in the HBSS are not reduced compared to the SABS. The HBSS is as energy efficient as the SABS.

With an increase in the share of the dynamic current component (blue curves in Figs. 8 and 9), i.e. when $\varepsilon=I_p/I_0$=1.0, the advantages of the HBSS compared to the SABS become obvious, i.e. the losses in the HBSS compared to the SABS are smaller (a smaller $E_{HBSS}(0,T)/E_{SABS}(0,T)$ ratio). By comparing Figs. 8 and 9, it can be observed that the reduced losses in the HBSS compared to the SABS are not significant if the pulse duration is small compared to the period ($\beta=T_p/T$=0.01 in Fig. 8).

By increasing the duration of the pulse T_p compared to the period T (a greater $\beta=T_p/T$ ratio) and a simultaneous increase in the share of the dynamic current component in the load current (a greater $\varepsilon=I_p/I_0$ ratio), the advantages of the HBSS in terms of energy efficiency become obvious (the green curve in Fig. 9).

Furthermore, following the green curves in Fig. 8 and Fig. 9, asymptotic trends can be observed with the change of the parameter k. The physical interpretation is that there is a physical limit to how much the losses in the HBSS can be reduced, and consequently, how much the energy efficiency of the HBSS can be improved. For practical cases where the supercapacitor has a lower internal resistance than the battery (k>1), up to approximately $k\approx10$, the green curve shows significant changes with variations in the parameter k up to approximately $k\approx10$. Beyond this point, i.e. for k>10, the green curves exhibit pronounced asymptotic behavior, the so-called saturation. This indicates that there is an interval of the parameter k, i.e. $1\leq k\leq10$, which is optimal for the sizing of the HBSS from a techno-economic aspect. Reducing losses in the HBSS, or increasing the energy efficiency of the HBSS, by choosing a parameter k greater than 10 becomes uneconomical. This would lead to an HBSS design with a relatively large supercapacitor compared to the battery. In addition, the reduction of losses would be small compared to the reduction of losses with the range of k values between 1 and 10, i.e. for $1\leq k\leq10$.

The results obtained are consistent with those reported in the literature [13, 27]. However, the approach proposed in this paper offers a more straightforward inter-

pretation of the results and provides better insights into how HBSS parameters affect current and voltage conditions and energy efficiency under dynamic load. The conducted analysis points to the following conclusions.

If the dynamic component of the load current is a very small part of the total load current, or if load currents vary slowly, the passive HBSS does not offer any advantages over the SABS in terms of energy efficiency. The reduction in battery current within the HBSS compared to the SABS is negligible. Because of this, there are no techno-economic justifications for using the HBSS in these cases. The analysis also showed that the HBSS is more advantageous than the SABS in applications where the dynamic current component constitutes a significant portion in the total load current (a greater I_p/I_0 ratio). Additionally, from a techno-economic aspect, there is an interval of the parameter , i.e. ($1 \leq k \leq 10$), which is optimal for the sizing of the HBSS. Besides improving the energy efficiency compared to the SABS, the HBSS also offers other advantages compared to the SABS, including a lower voltage drop at the HBSS terminals, and reduced current and thermal stress on the battery inside the HBSS.

6. CONCLUSIONS

This paper presents an analytical approach to the analysis of energy efficiency of the passive HBSS in relation to the SABS. A novel set of analytical expressions is introduced for this analysis. Analytical expressions are derived using a first-order circuit model that represents both the HBSS and the dynamic load. These derived expressions are expressed through two additional auxiliary dimensionless parameters (K and k) that are related to the circuit parameters of the HBSS. The waveform of the load current is characterized by three dimensionless parameters (ε, α, and β), one of which (α) is also related to the HBSS parameters. Thanks to the above, general expressions for currents, voltages and energy efficiency have been achieved. Compared to existing approaches, the proposed approach offers better insight into the influence of battery and supercapacitor parameters on the energy efficiency of the passive HBSS in relation to the SABS. The proposed approach is also suitable for quick determination of HBSS behavior (voltage at the HBSS terminals, battery and SC currents) and energy efficiency under dynamic load. This enables the identification of HBSS parameters required for optimal behavior and optimal energy efficiency. In this way, it is possible to avoid successive simulations with varying model parameters to determine the behavior and energy efficiency of the HBSS under dynamic load. Instead, if necessary, more accurate results can be obtained through simulations based on a higher-order model.

The performed analysis illustrates the effectiveness of the proposed approach in determining the energy efficiency of the HBSS in relation to the SABS, as well as the rapid identification of the HBSS parameters required for desired behavior and energy efficiency. Regarding energy efficiency, the analysis reveals that there is an interval of the parameter k which is optimal for the sizing of the HBSS a techno-economic aspect. The advantages of the passive HBSS compared to the SABS are described in detail in the paper. The analysis shows that the passive HBSS has advantages over the SABS only under dynamic load. More specifically, the passive HBSS is more suitable than the SABS in applications where the dynamic current component constitutes a significant portion in the total load current (a greater I_p/I_0 ratio).

The presented set of analytical expressions can be further improved for greater accuracy by using a more complex circuit model to represent the HBSS. To pave the way for potential improvement, the derivation of this set of analytical expressions is provided in detail, accompanied by explanatory comments.

7. REFERENCES

[1] A. Djouahi, B. Negrou, Y. Touggui, M. M. Samy, "Optimal sizing and thermal control in a fuel cell hybrid electric vehicle via FC-HEV application", Journal of the Brazilian Society of Mechanical Sciences and Engineering, Vol. 45, No. 10, 2023.

[2] A. Djouahi, B. Negrou, B. Rouabah, A. Mahboub, M. M. Samy, "Optimal Sizing of Battery and Super-Capacitor Based on the MOPSO Technique via a New FC-HEV Application", Energies, Vol. 16, No. 9, 2023, p. 3902.

[3] A. F. Guven, A. Y. Abdelaziz, M. M. Samy, S. Barakat, "Optimizing energy Dynamics: A comprehensive analysis of hybrid energy storage systems integrating battery banks and supercapacitors", Energy Conversion and Management, Vol. 312, 2024, p. 118560.

[4] I. Jarraya, F. Abdelhedi, N. Rizoug, "An Innovative Power Management Strategy for Hybrid Battery–Supercapacitor Systems in Electric Vehicle", Mathematics, Vol. 12, No. 1, 2023, p. 50.

[5] D. Linden, T. Reddy, "Handbook of Batteries", 3rd Edition, McGraw Hill Professional, 2001.

[6] G. L. Plett, "Battery Management Systems, Volume II: Equivalent-Circuit Methods", 2nd Edition, Artech House Publishers, 2020.

[7] F. Beguin, E. Frackowiak, "Supercapacitors: Materials, Systems, and Applications", 1st Edition, Wiley-VCH, 2013.

[8] A. Yu, V. Chabot, J. Zhang, "Electrochemical Supercapacitors for Energy Storage and Delivery:

Fundamentals and Applications", 1st Edition, CRC Press, 2017.

[9] J. M. Miller, "Ultracapacitor Applications", The Institution of Engineering and Technology, 2011.

[10] B. E. Conway, "Electrochemical Supercapacitors: Scientific Fundamentals and Technological Applications", Springer Science & Business Media, 2013.

[11] R. P. Deshpande, "Ultracapacitors", McGraw- Hill Education, 2015.

[12] R. Kötz, M. Carlen, "Principles and applications of electrochemical capacitors", Electrochimica Acta, Vol. 45, No. 15-16, 2000, pp. 2483-2498.

[13] A. Kuperman, I. Aharon, "Battery-ultracapacitor hybrids for pulsed current loads: A review", Renewable & Sustainable Energy Reviews, Vol. 15, No. 2, Feb. 2011, pp. 981-992.

[14] T. Ma, H. Yang, L. Lu, "Development of hybrid battery-supercapacitor energy storage for remote area renewable energy systems", Applied Energy, Vol. 153, Sep. 2015, pp. 56-62.

[15] L. H. Seim, "Modeling, control and experimental testing of a supercapacitor/battery hybrid system: passive and semi-active topologies", http://brage.bibsys.no/xmlui/handle/ 11250/ 188829 (accessed: 2024)

[16] E. Wang, F. Yang, M. Ouyang, "A Hybrid Energy Storage System for a Coaxial Power-Split Hybrid Powertrain", InTech, 2017.

[17] C. Mi, M. A. Masrur, D. W. Gao, "Hybrid Electric Vehicles: Principles and Applications with Practical Perspectives", 1st Edition, Wiley, 2011.

[18] Z. Cabrane, M. Ouassaid, M. Maaroufi, "Analysis and evaluation of battery-supercapacitor hybrid energy storage system for photovoltaic installation", International Journal of Hydrogen Energy, Vol. 41, No. 45, 2016, pp. 20897-20907.

[19] Z. Dong et al. "A Survey of Battery-Supercapacitor Hybrid Energy Storage Systems: Concept, Topology, Control and Application", Symmetry, Vol. 14, No. 6, 2022, pp. 1-26.

[20] W. Jing, C. H. Lai, S. H. W. Wong, M. L. D. Wong, "Battery-supercapacitor hybrid energy storage system in standalone DC microgrids: a review", IET Renewable Power Generation, Vol. 11, No. 4, 2017, pp. 461-469.

[21] K. Jayasawal, A. K. Karna, K. B. Thapa, "Topologies for Interfacing Supercapacitor and Battery in Hybrid Electric Vehicle Applications: An Overview", Proceedings of the International Conference on Sustainable Energy and Future Electric Transportation, Hyderabad, India, 21-23 January 2021, pp. 1-6.

[22] W. Jing, C. H. Lai, W. S. H. Wong, M. L. D. Wong, "A comprehensive study of battery-supercapacitor hybrid energy storage system for standalone PV power system in rural electrification", Applied Energy, Vol. 224, 2018, pp. 340-356.

[23] V. Bolborici, F. P. Dawson, K. K. Lian, "Sizing considerations for ultracapacitors in hybrid energy storage systems", Proceedings of the IEEE Energy Conversion Congress and Exposition, Phoenix, AZ, USA, 17-22 September 2011, pp. 2900-2907.

[24] N. H. Liu, N. Z. Wang, N. J. Cheng, D. Maly, "Improvement on the Cold Cranking Capacity of Commercial Vehicle by Using Supercapacitor and Lead-Acid Battery Hybrid", IEEE Transactions on Vehicular Technology, Vol. 58, No. 3, 2009, pp. 1097-1105.

[25] I. Lahbib, A. Lahyani, A. Sari, P. Venet, "Performance analysis of a lead-acid battery/supercapacitors hybrid and a battery stand-alone under pulsed loads", Proceedings of the First International Conference on Green Energy ICGE 2014, Sfax, Tunisia, 25-27 March 2014, pp. 273-278.

[26] G. Gu, Y. Lao, Y. Ji, S. Yuan, H. Liu, P. Du, "Development of hybrid super-capacitor and lead-acid battery power storage systems", International Journal of Low Carbon Technologies, Vol. 18, 2023, pp. 159-166.

[27] R. A. Dougal, S. Liu, R. E. White, "Power and life extension of battery-ultracapacitor hybrids", IEEE Transactions on Components and Packaging Technologies, Vol. 25, No. 1, 2002, pp. 120-131.

[28] D. Buljić, T. Barić, H. Glavaš, "Analytical description of the dynamic behaviour of the passive battery supercapacitor hybrid system", Technical Gazette, Vol. 31, No. 4, 2024.

[29] C. T. Tshiani, P. Umenne, "The Impact of the Electric Double-Layer Capacitor (EDLC) in Reducing

Stress and Improving Battery Lifespan in a Hybrid Energy Storage System (HESS) System", Energies, Vol. 15, No. 22, 2022.

[30] D. Shin, M. Poncino, E. Macii, "Thermal management of batteries using a hybrid supercapacitor architecture", Proceedings of the Design, Automation & Test in Europe Conference & Exhibition, Dresden, Germany, 24-28 March 2014, pp. 1-6.

[31] S. Pay, Y. Baghzouz, "Effectiveness of battery-supercapacitor combination in electric vehicles", Proceedings of the IEEE Bologna Power Tech Conference Proceedings, Bologna, Italy, 23-26 June 2003, Vol. 3, p. 6.

[32] M. T. Penella, M. Gasulla, "Runtime Extension of Low-Power Wireless Sensor Nodes Using Hybrid-Storage Units", IEEE Transactions on Instrumentation and Measurement, Vol. 59, No. 4, 2010, pp. 857-865.

[33] E. Naderi, B. K. C, M. Ansari, A. Asrari, "Experimental Validation of a Hybrid Storage Framework to Cope With Fluctuating Power of Hybrid Renewable Energy-Based Systems", IEEE Transactions on Energy Conversion, Vol. 36, No. 3, 2021, pp. 1991-2001.

[34] A. W. Stienecker, T. Stuart, C. Ashtiani, "An ultracapacitor circuit for reducing sulfation in lead acid batteries for Mild Hybrid Electric Vehicles", Journal of Power Sources, Vol. 156, No. 2, 2006, pp. 755-762.

[35] A. L. Shenkman, "Transient Analysis of Electric Power Circuits Handbook", Springer, 2005.

[36] Y. Chuan, C. Mi, M. Zhang, "Comparative Study of a Passive Hybrid Energy Storage System Using Lithium Ion Battery and Ultracapacitor", World Electric Vehicle Journal, Vol. 5, No. 1, 2012, pp. 83-90.

[37] D. Cericola, R. Kötz, "Hybridization of rechargeable batteries and electrochemical capacitors: Principles and limits", Electrochimica Acta, Vol. 72, 2012, pp. 1-17.

[38] G. A. Turner, "US6836097B2 - Power supply for a pulsed load", Google Patents, https://patents.google.com/ patent/ US6836097B2 /en (accessed: 2024)

Miniaturized Ultra-Wide Band (UWB) Delta-Stub Based 4-Way Power Divider with Enhanced Isolation For Internet of Things (IoT) and Next-Generation Wireless Systems

Original Scientific Paper

Arshad Karimbu Vallappil*
Institute of Space Sciences (ICE, CSIC),
Campus UAB, Carrer de Can Magrans, s/n 08193, Cerdanyola del Vallès, Barcelona, Spain
kv.arshad@gmail.com, karimbu@ice.csic.es

Bilal A. Khawaja*
Faculty of Engineering, Department of Electrical Engineering,
Islamic University of Madinah, P. O. BOX 170 Madinah, 41411, Saudi Arabia
kjbmohammed@iu.edu.sa, bam.khawaja@gmail.com

Abdulmajeed M Alenezi
Faculty of Engineering, Department of Electrical Engineering,
Islamic University of Madinah, P. O. BOX 170 Madinah, 41411, Saudi Arabia
a.alenezi@iu.edu.sa

*Corresponding author

Abstract *– This paper proposes a novel wideband 1-to-4-way Wilkinson power divider (WPD) designed for operation in the ultra-wideband (UWB) frequency band. The proposed power divider operates in the frequency range of 3-8GHz. The design process started with conventional WPD design and evolved based on three stages of a 1-to-2 way power divider, incorporating 100Ω resistors and two delta-stubs for each stage. Delta-stubs are introduced in the WPD design for enhanced operational bandwidth and impedance-matching characteristics. The proposed power divider is designed using an FR4 substrate with a thickness of 0.76mm and ε_r of 4.3, respectively. The proposed design exhibits equal power split at all ports, good insertion-loss of approximately -6 ± 2dB, and return-loss of below -10 dB over the operating frequency band. Moreover, good impedance-matching, and isolation performance have been obtained over the desired frequency band. The proposed WPD was initially modeled mathematically and then designed and optimized using CST microwave studio (CST-MWS). The proposed design can be used in the next-generation UWB-Internet of Things (UWB-IoT) antenna array systems.*

Keywords: *UWB, Wilkinson Power Divider, Delta-Stub, Antenna Array Systems, Internet of Things (IoT)*

Received: April 16, 2024; Received in revised form: July 18, 2024; Accepted: August 9, 2024

1. INTRODUCTION

The last decade has seen a significant interest in ultra-wideband (UWB) technology [1-3] due to its enhanced bandwidth, radiation characteristics, and increased data-transmission features [2-4]. The Federal Communication Commission (FCC), USA, has designated a frequency band of 3.1-10.6 GHz (7.5 GHz bandwidth) for use in UWB applications [1-3]. Both wireless systems and sensor networks benefit greatly from UWB technology's enhanced data-transmission rates and low-energy consumption features [1-5]. Also, it is gaining acceptance in UWB-Internet of Things (UWB-IoT) applications for continuous data-transfer between low-cost sensor devices [6-9]. Due to its low-power emission characteristics, UWB technology is recognized as a viable solution for wireless body area networks (WBANs). This also renders the UWB technology suitable for wearable IoT devices and sensors [5-9].

For UWB systems, several authors have proposed a range of designs for UWB antennas and filters [1-11].

However, not a lot of UWB power dividers have been proposed in the past albeit the fact they are a key component in a variety of microwave systems. For example, power dividers are considered a critical component in the design of balanced mixers [12], phase shifters [13], and feeding networks for phased array antenna (PAA) systems [14-17].

The utilization of 3-port power dividers is particularly critical for the UWB-PAA systems [15-17] that make use of power splitters, for example, a corporate or parallel power splitting system. The corporate feeding network maintains equal path lengths between input and output ports while dividing power between n-output ports with a specific distribution. In situations where 3-port power dividers are frequently utilized, they can be implemented with an n-way power dividing network. The flexible configuration of the two-way power divider enables the employment of numerous stepped sections to produce power divisions with wideband operation capacity [12]. Although utilizing high-isolation power dividers minimizes the reliance on load matching, the operational bandwidth of the power divider network is principally constrained due to the matching required for the radiating elements [12].

The Wilkinson power divider (WPD) [12, 18], also known as the Wilkinson splitter, is a 3-port power divider network that was first introduced in 1960 by Ernest Wilkinson [18]. It has been extensively used in various microwave applications, including wireless communication, PAA, and radar systems. A WPD is a passive splitter that takes one input signal and divides it into n output signals with equal phases and amplitudes. Theoretically, WPD exhibits complete isolation among the output ports at the desired frequency. It is based on the principle of impedance transformation, which is achieved by using quarter-wave transmission-lines (TLs).

The WPD is a type of hybrid splitter that is fundamentally a narrowband device. Many different researchers [13, 14, 19-24] have implemented numerous techniques to enhance its bandwidth of operation. Some of these bandwidth enhancement techniques involve the addition of segmented structures [19], tapered lines [13], and stubs to each section [14, 20, 21]. Other researchers have proposed the use of defective ground structures of different shapes for this purpose [22, 23].

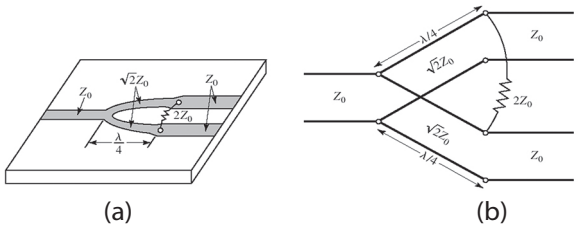
In [14], the authors present an innovative method for broadening the bandwidth of a power divider utilized in six-port interferometers by incorporating radial stubs within a conventional Wilkinson design. Their work across the 3.1-10.6 GHz UWB frequency band reveals a significant enhancement, showcasing a simulated bandwidth expansion of 37.5%. Yu et al. [19] introduce a novel broadband WPD topology utilizing a segmented structure, employing TL segments with shunt capacitors and series resistor-capacitor networks. The design reported in [20] presents a modified approach to the traditional two-section WPD by incorporating open stubs on each branch, demonstrating improved performance within the UWB frequency band. This design not only outperforms the traditional three-section power divider but also offers a 10% reduction in size and eliminates the need for one resistor, showcasing promising simulation and measurement results. Similarly, Dardeer et al. [21] introduced a cost-effective and compact microstrip-based WPD within the UWB range. The design uses two open-stubs to enhance the bandwidth. Peng et al. [24] demonstrated a UWB non-coplanar power divider design by adding two slots, a tapered and a fan-shaped slot, to enhance the power divider's overall performance.

In this paper, the authors present a study on the design, fabrication, and characterization of a compact and miniaturized UWB 4-Way WPD based on delta-stub with enhanced isolation for UWB-IoT and next-generation wireless system applications. The agreement between the simulated and measured results suggests that the proposed UWB WPD performs well and can be used in UWB-IoT and next-generation wireless systems.

The following sections are organized as follows: Section 2 briefly describes the conventional configuration of the WPD. Section 3 discusses the step-by-step design process of the proposed WPD geometry and its dimensions. Section 4 summarizes the WPD fabrication, characterization, and comparison of simulated and measured results in terms of parameters such as return-loss, insertion-loss, and isolation-loss. Additionally, it provides a comparative study between the design presented in this paper and the existing designs proposed by other researchers. Finally, Section 5 draws conclusions.

2. WPD CONVENTIONAL DESIGN

The WPD functions as a 3-port system, which is typically lossless in nature. When the output ports of a WPD are appropriately matched, the network becomes lossless with only the dissipation of the reflected power. The input power can be divided into two or maybe more in-phase signals of equal amplitude. A quarter-wave ($\lambda/4$) transformer with a characteristic impedance of $\sqrt{(2)}Z_0$ can be used to realize a 2-way WPD. Fig. 1(a-b), depicts the fundamental two-way design of the WPD. [12-13].



(a) (b)

Fig. 1: (a) An equal-split WPD showing the line-impedances and $\lambda/4$ transformers (b) WPD equivalent transmission-line circuit [12]

The initial dimensions of the WPD can be calculated using the well-known transmission-line model [12], as discussed in Eq. (1)-(6), respectively.

The effective dielectric constant (ε_{reff}) calculation in Eq. (1) leads to accurate modeling of the microstrip-line, and it depends upon the substrate thickness (h), relative permittivity (ε_r), and the conductor width (W).

$$\varepsilon_{reff} = \frac{(\varepsilon_r + 1)}{2} + \frac{(\varepsilon_r - 1)}{2}\left(1 + 12 * \frac{h}{W}\right)^{-\frac{1}{2}} \quad (1)$$

Eq. (2) is used to calculate the microstrip-line wavelength (λ in mm), which is correlated with the phase velocity. Where, f is the desired operating frequency. The center frequency is selected as 5.5GHz. It is important to note that Eq. (2) leads to the calculation of the guided wavelength factor, as it is used to determine the size of components such as stubs and the WPD itself at the intended frequency.

$$\lambda = \frac{300}{f\sqrt{\varepsilon_{reff}}} \quad (2)$$

The characteristic impedance (Z_o=50Ω) of a microstrip-line is computed using the dimensions from Eq. (3)-(4). The characteristic impedance is a crucial parameter for ensuring impedance matching. The width-to-height ratio (W/h) can be determined by using Eq. (5)-(6), respectively.

$$A = \frac{Z_o}{60}\left(\frac{\varepsilon_r + 1}{2}\right)^{1/2} + \frac{\varepsilon_r - 1}{\varepsilon_r + 1}\left(0.23 + \frac{0.11}{\varepsilon_r}\right) \quad (3)$$

$$B = \frac{60\pi^2}{Z_o\sqrt{\varepsilon_r}} \quad (4)$$

For $A > 1.52$:

$$W/h = \frac{8\,exp(A)}{exp(2A) - 2} \quad (5)$$

For $A < 1.52$:

$$\frac{W}{h} = \frac{2}{\pi}\{B - 1 - ln(2B - 1)\} + \frac{\varepsilon_r - 1}{2\varepsilon_r}\left[ln(B - 1) + 0.39 - \frac{0.61}{\varepsilon_r}\right] \quad (6)$$

It is important to note that the parameters in Eq. (3)-(4) are intermediate variables (A and B) that depend on Z_o. The initially achieved dimensions from the transmission-line model Eq. (1)-(6) are now used to simulate and optimize the WPD design using the CST microwave studio.

3. WPD PROPOSED DESIGN

Fig. 2(a-b) depicts the proposed design for the 1-to-4-way WPD based on delta-stub. The equivalent transmission-line circuit of the proposed WPD is shown in Fig. 2(c). Two delta-stubs, measuring 2.38 x 0.74 mm (length x width), are added on both sides of the $\lambda/4$ transformers to improve impedance matching and widen the bandwidth characteristics of the WPD. The labeling of WPD is shown in Fig. 2(a), and the optimum dimension of WPD is shown in Fig. 2(b). For ease of understanding, in the rest of the paper, Port-1 will be referred to as the input-port, whereas Ports 2-5 will be considered output ports, respectively. The impedance of the input and output ports is approximately 50Ω. The characteristic impedance of the first TLs is $Z_1 = \sqrt{(2)}Z_0 = 70.7\Omega$.

In the proposed design, at each output port, the power is 1/4th the power of the input-port. So, the transmission loss should be -6dB at each of the output ports. The proposed power divider is designed using an FR4-substrate with a thickness of 0.76mm and ε_r of 4.3, respectively. The FR4 substrate is used in this study because it is a readily available and cost-effective material for the proposed power divider design. Moreover, FR4 adheres to cost limitations in practical applications, particularly in IoT, and facilitates manufacturing and scalability [25]. The FR4 substrate size of 44 x 24 mm is taken in this study. The 50Ω TL width is taken as 1.5mm, and the length ($\lambda_g/4$) and width of 70.7Ω TL are taken as 7.17mm and 0.65 mm, respectively. The current design uses the delta-stub for better impedance matching and wide bandwidth characteristics. Table 1 lists the modified WPD's optimized dimensions.

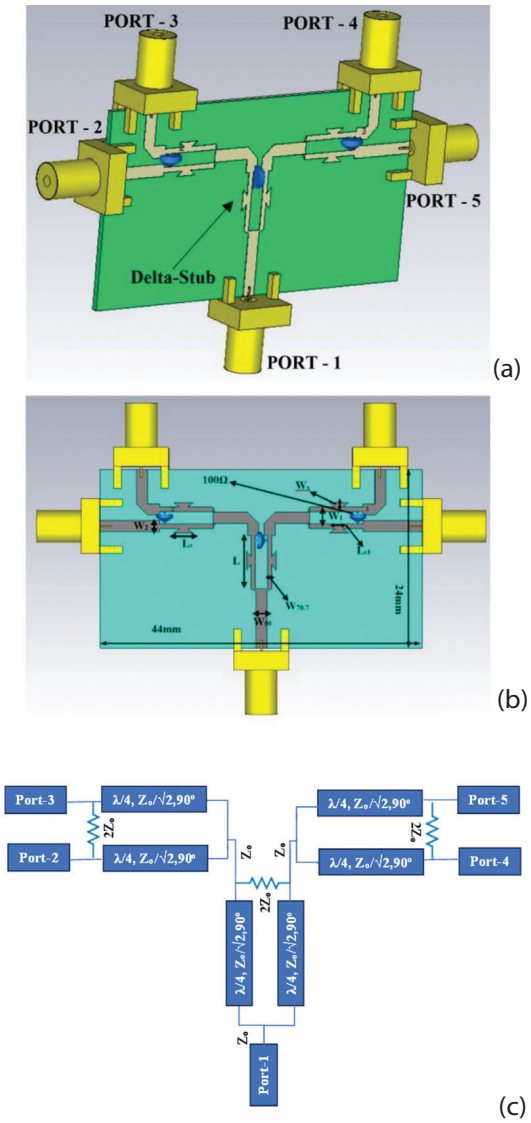


Fig. 2. CST simulation layout of the 1-to-4 way modified WPD (a) Port labeling (b) Optimized dimensions (c) Equivalent transmission-line circuit

Table 1. Parameters of modified 1-to-4 way WPD

W_{50} (mm)	$W_{70.7}$ (mm)	W_1 (mm)	W_2 (mm)
1.5	0.65	1.6	0.75
L (mm)	L_s (mm)	L_{s1} (mm)	W_s (mm)
7.17	2.62	1.1	0.6

4. WPD RESULT AND RELATED DISCUSSION

This section summarizes and validates the results of the proposed and designed WPD working in the UWB frequency range of 3-8GHz. The fabricated prototype of 1-to-4 way modified WPD is shown in Fig. 3. It was fabricated on FR4 substrate ($\varepsilon_r = 4.3$, $h = 0.76$ mm, and loss-tangent ($tan\ \delta$) of 0.019). The overall dimension of the fabricated prototype was 44 mm × 24 mm (1056 mm^2), as shown in Fig. 3, which demonstrates an 84% size reduction compared to the conventional TL-based WPD with the same frequency band [26]. The prototype of the proposed WPD was measured using Keysight Technologies FieldFox N9916B vector network analyzer (VNA), and the results are discussed in the following sections.

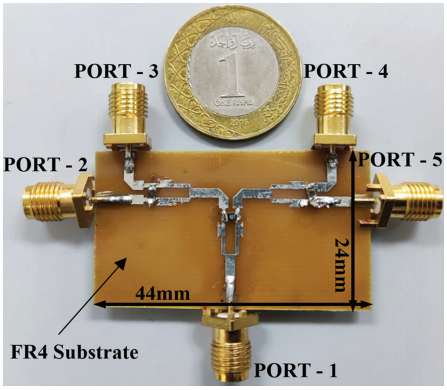
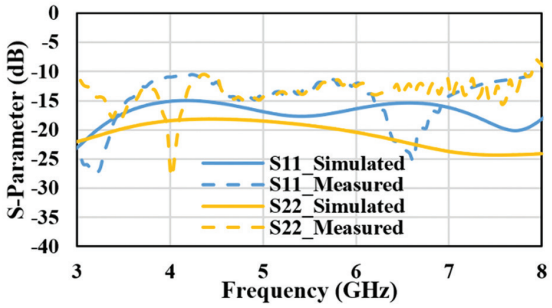


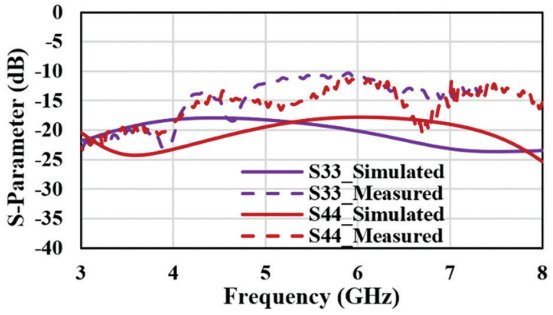
Fig. 3. Fabricated prototype of Proposed WPD

4.1. RETURN-LOSS

Fig. 4 shows the simulated and measured return-loss (S-parameter) results for all the ports of the designed WPD. The Port-1 of the WPD was connected to the Port-1 of the VNA using a high-frequency RF cable and connector to measure the return-loss of WPD. The remaining ports of WPD were terminated using 50 Ω terminators. Fig. 4 shows the simulated and measured return-loss performance across the 3-8GHz frequency band with a magnitude better than -10dB.



(a)

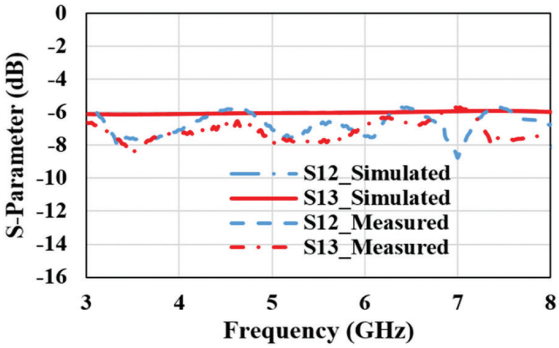


(b)

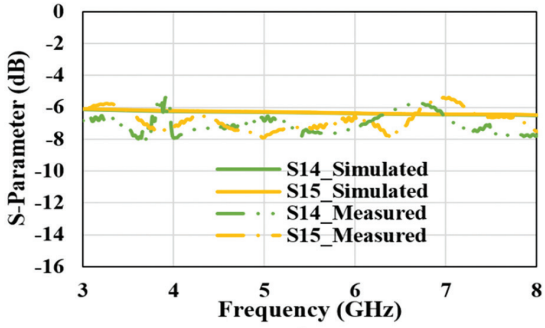
Fig. 4. Simulated and measured return-loss (a) S_{11}, S_{22} (b) S_{33}, S_{44}

4.2. INSERTION-LOSS

For the measurement of WPD insertion-loss, Port-1 of VNA was connected to the WPD Port-1, and Port-2 was connected to the WPD Port-2; meanwhile, WPD Port-3, Port-4, and Port-5 were terminated with 50Ω terminators. The simulated and measured insertion-loss results for the designed WPD for Port-2 to Port-5 with respect to the input Port-1 are shown in Fig. 5 (a-b).



(a)



(b)

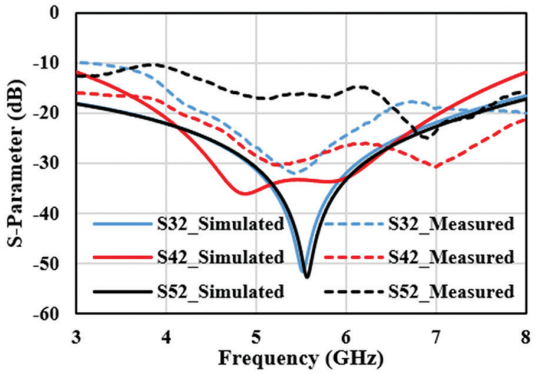
Fig. 5. Simulated and measured insertion-loss (a) S_{12}, S_{13} (b) S_{14}, S_{15}

The simulated and measured insertion-loss shows the value of -6 ± 0.2 dB and -6 ± 2 dB across the UWB frequency range of 3-8G Hz. Fig. 5 also shows the S_{12} results, which highly resemble S_{13}, S_{14}, and S_{15} results, respectively. This means that the input power (from WPD Port-1) has been split equally to all output ports (Port-2 to Port-5), and it performs well with minimal insertion-loss over the specified frequency range.

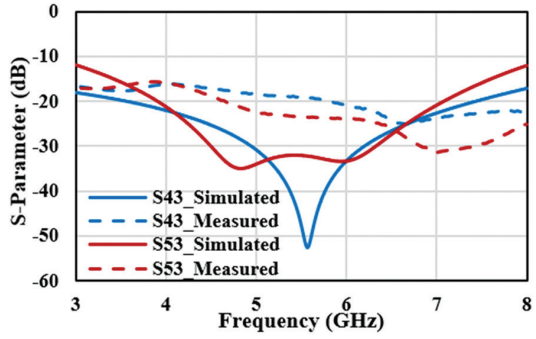
4.3. ISOLATION-LOSS

Fig. 6(a) shows the simulated and measured isolation-loss results of the designed WPD for Port-3 to Port-5 with respect to the input Port-2. As shown in Fig. 6 (a), the input power was distributed equally to all the power divider's output ports. The isolation-loss is below -10 dB, and it performs well across the required UWB frequency range of 3-8GHz.

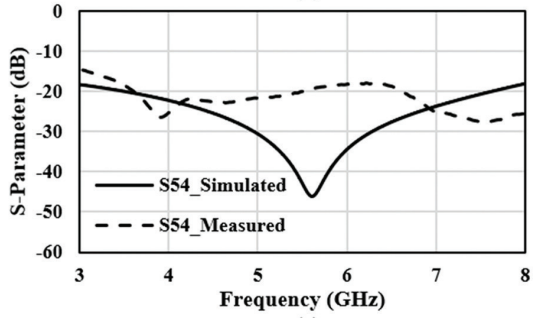
Similarly, the simulated and measured isolation-loss results for the designed WPD for Port-4 to Port-3 and Port-5 to Port-3 are shown in Fig. 6(b). The results show excellent isolation between the ports over the frequency range of 3-8GHz. Fig. 6(c) shows the simulated and measured isolation-loss results for the designed WPD for port-5 to port-4 with an isolation-loss below -10dB across the frequency range of 3-8GHz.



(a)



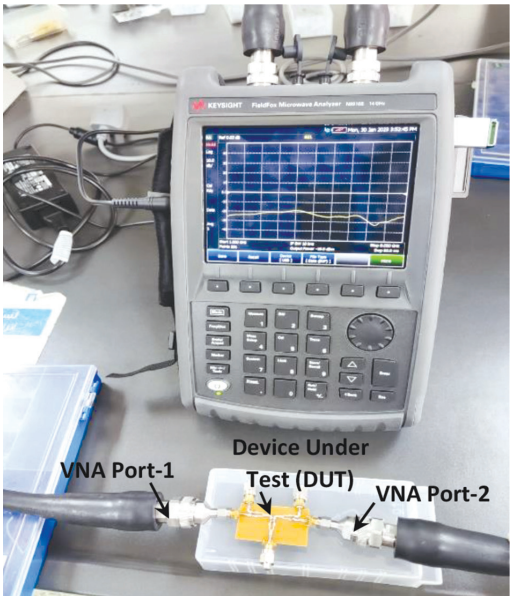
(b)



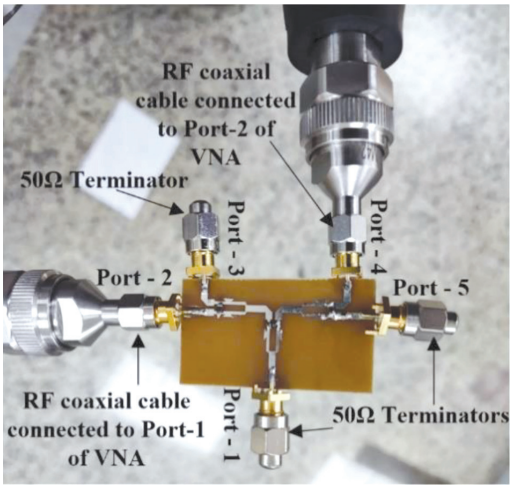
(c)

Fig. 6. Simulated and measured Isolation-loss (a) S_{32}, S_{42}, and S_{52}, (b) S_{43}, S_{53} (c) S_{54}

Fig. 7(a-b) shows the overall measurement setup of the WPD. Fig. 7(a) shows that the Port-2 and Port-5 of the device-under-test (DUT) are connected to VNA Port-1 and Port-2, respectively. For WPD isolation-loss measurement, VNA Port-1 was connected to the WPD Port-2, and VNA Port-2 was connected to the WPD Port-4; meanwhile, WPD Port-1, Port-3, and Port-5 were terminated using 50Ω terminators.



(a)



(b)

Fig. 7. Shows the WPD measurements setup using Keysight Technologies FieldFox N9916B Two-Ports vector network analyzer (VNA) and 50Ω Terminators (a) S_{25} measurement setup showing the DUT connected to Keysight VNA (b) Zoom-in on the S24 measurement setup

The results presented in Fig. 4 (a-b) and Fig. 5 (a-b) show a slight deviation between the simulated and measured responses of the WPD. The reason for this deviation is the lossy nature of the FR4 substrate, which has the loss-tangent ($tan\,\delta$) in the range of 0.019 - 0.025. Due to the lossy nature of the FR4 substrate, it is very difficult to get the measurement responses exactly the same as the simulated ones. However, it is important

to note that the shape of the simulated and measured *S*-parameter results is almost similar. This discrepancy can be improved by using high-end Roger's substrates. Another reason for this shift can be due to the amount of solder on the feed-line of the WPD. This problem can be solved in the future by employing precision soldering techniques.

4.4. COMPARATIVE STUDY

The performance of the proposed and designed WPD presented in this paper is compared with previous design works [27-37] and is summarized in Table 2. The designed WPD has an overall area of 44 mm × 24 mm = 1056 mm². In contrast, the previous design work [27] occupies an area of 86 mm × 42 mm = 3612 mm². This comparison shows that the proposed power divider occupies only 29% of the area compared to the design in [27]. This significant size reduction is achieved without any degradation in performance. The data in Table 2 suggest that the proposed design has excellent performance in terms of size reduction, bandwidth, return-loss, insertion-loss, and isolation-loss compared to previous designs. Although the article [28] shows better size reduction, its bandwidth is less compared to the proposed design.

Table 2. Performance comparison of the proposed 1-to-4-way WPD with existing designs

Ref.	Bandwidth (GHz)	Return-loss (dB) (S_{11})	Insertion-loss (dB) (S_{21})	Isolation-loss (dB) (S_{32})	No. of Ports	Technique	Size (mm × mm)	Analysis
[27]	1-4	<-25	6 ± 1.5	<-28	4	Stub	1.43 λ_g × 0 .7 λ_g	High Isolation, Large Size
[28]	2.3-4.4	<-15	0.5	<-17	4	Microstrip-line Marchand balun	0.4 λ_g × 0.56 λ_g	High Isolation, Low Insertion-loss
[29]	1.08-1.93	<-15	– 6.8	<-13	4	Looped Coupled-Line Structures	0.32 λ_g × 0.32 λ_g	Narrow Bandwidth
[30]	0.54-1.08	<-15	0.3	<-15	4	Meandered Line	0.27 λ_g × 0.26 λ_g	Narrow Bandwidth, Better Return-loss
[31]	3.1-10	<-10	-6.5	<-13	4	Defected Ground Structure	1.97 λ_g × 2.5 λ_g	Wide Bandwidth, Large size
[32]	1.6-2.6	<-16	-7	<-20	4	Coshared Multi-Mode Resonators	0.63 λ_g × 0.33 λ_g	Narrow Bandwidth, Better Isolation
[33]	1.08-1.93	<-15	-6	<-13	4	Looped Coupled-Line Structures	0.5 λ_g × 0.5 λ_g	Narrow Bandwidth
[34]	0.83-1.39	<-10	-	<-12	4	Coupled-Lines	0.55 λ_g × 0.16 λ_g	Narrow Bandwidth, Compact Size
[35]	8.6-12.2	<-15	-7.8	-	4	Substrate Integrated waveguide method	4 λ_g × 4 λ_g	Wide Bandwidth, Large Size
[36]	7.92–9.53	<-28	0.37	<-20	4	Suspended Strip-line power divider	0.8 λ_g × 0.6 λ_g	Narrow Bandwidth, Low Insertion-loss, Better Return, and Isolation losses
[37]	17–19	<-12.5	2.25	<-20	4	Gysel Power Divider	3.29 λ_g × 0.8 λ_g	Large Size
[This Work]	3-8	<-10	6 ± 1	<-10	4	Delta-Stub	0.9 λ_g × 0.5 λ_g	Wide Bandwidth, Good Insertion and Return losses, Compact Size

The proposed WPD covers a moderate bandwidth of 3-8 GHz, making it suitable for UWB applications. In comparison, references [31] and [35] offer wide bandwidths but at the expense of significantly large sizes, while most other references cover narrow bandwidths, limiting their applicability in UWB systems. The return-loss achieved in this work is <-10 dB, which is acceptable for many applications. In contrast, references [30] and [36] demonstrate better return-loss performance (<-20 dB and <-28 dB, respectively), although they either have a much narrower bandwidth or a slightly larger size. The insertion-loss in the proposed design is approximately 6 ± 1 dB, comparable to most other references. Articles [28] and [36] have significantly lower insertion-loss but come with other trade-offs, such as isolation or application bandwidth.

In terms of isolation-loss, the proposed design demonstrates <-10 dB, which is decent but not as high as some other researchers [27], [32], and [36]. The designs presented in [27] has larger size, while [32] and [36] offer better isolation-loss but narrow bandwidth, making them more suitable for applications requiring high isolation.

Additionally, the proposed design presented in this work is relatively compact compared to most of the designs discussed in Table 2. A comparable compact design is presented in [34], but it does not offer the same bandwidth or performance metrics.

Articles [35-37] are based on substrate integrated, Gysel, and suspended-stripline power dividers. These articles present designs with lesser bandwidth and larger sizes compared to the proposed design. Furthermore, the proposed design's fabrication process is much easier compared to these works, as the latter uses multilayer techniques that increase fabrication cost and complexity. This highlights the advantage of the single-layer WPD method proposed in this paper.

Overall, the proposed work offers a balanced approach with moderate bandwidth, good return-loss, and insertion-loss, along with a compact size. The unique use of delta-stubs enhances operational bandwidth and impedance-matching characteristics, setting this design apart from other techniques. The specific focus on next-generation UWB-IoT antenna array systems highlights the practical applicability of this design in modern technologies.

5. CONCLUSION

In this study, we have designed, fabricated, and characterized a unique 1-to-4 way WPD employing the delta-stub technique for a wideband frequency range operation between 3-8 GHz. This study demonstrates that the proposed WPD outperforms previous designs by offering a smaller size, which leads to lower fabrication costs in practical applications. The proposed compact and wideband power divider has exhibited good performance with an overall size reduction of 71% as compared to previous work [27] by using stubs.

The compact yet wideband characteristics of the proposed WPD have yielded robust performance attributes, including equal power distribution, low insertion-loss, improved isolation, and adequate return losses across all ports. The unique use of delta-stubs enhances operational bandwidth and impedance-matching characteristics, setting this design apart from other techniques. These results indicate that the proposed WPD is an appropriate candidate for integration into UWB-IoT antenna array feeding networks and emerging next-generation wireless systems.

In the future, we plan to integrate the currently presented 1-to-4 way WPD with a similar metamaterial-based beamforming network and antennas [1, 3, 16] for real-world testing and implementation of such systems for 5G and next-generation wireless systems. This integration aims to explore further enhancements in beamforming capabilities and dynamic beam-shaping functionalities.

6. REFERENCES

[1] G. Adamiuk, T. Zwick, W. Wiesbeck, "UWB Antennas for Communication Systems", Proceedings of the IEEE, Vol. 100, No. 7, 2012, pp. 2308-2321.

[2] Z. G. Wang, R. You, M. Yang, J. Zhou, M. Wang, "Design of a Monopole Antenna for Wifi-UWB Based on Characteristic Mode Theory", Progress in Electromagnetics Research M, Vol. 125, 2024, pp. 107-116.

[3] A. K. Vallappil, B. A. Khawaja, M. K. A. Rahim, M. N. Iqbal, H. T. Chattha, M. F. B. Mohamad Ali, "A Compact Triple-Band UWB Inverted Triangular Antenna with Dual-Notch Band Characteristics Using SSRR Metamaterial Structure for Use in Next-Generation Wireless Systems", Fractal and Fractional, Vol. 6, No. 8, 2022, p. 422.

[4] K. Shafique, B. A. Khawaja, M. A. Tarar, B. M. Khan, M. Mustaqim, A. Raza, "A Wearable Ultra-Wideband Antenna for Wireless Body Area Networks", Microwave and Optical Technology Letters, Vol. 58, No. 7, 2016, pp. 1710-1715.

[5] J. R. Ojha, M. Peters, "Patch Antennas and Microstrip Lines", Microwave and Millimeter Wave Technologies Modern UWB Antennas and Equipment, InTechOpen, 2010, pp. 49-62.

[6] K. Shafique, B. A. Khawaja, M. D. Khurram, S. M. Sibtain, Y. Siddiqui, M. Mustaqim, H. T. Chattha, X. Yang, "Energy Harvesting Using a Low-Cost Rectenna for Internet of Things (IoT) Applications", IEEE Access, Vol. 6, 2018, pp. 30932-30941.

[7] O. Cetinkaya, O. B. Akan, "Electric-Field Energy Harvesting from Lighting Elements for Battery-Less Internet of Things", IEEE Access, Vol. 5, 2017, pp. 7423-7434.

[8] L. Davoli, L. Belli, A. Cilfone, G. Ferrari, "From Micro to Macro IoT: Challenges and Solutions in The Integration of IEEE 802.15. 4/802.11 And Sub-GHz Technologies", IEEE Internet of Things Journal, Vol. 5, No. 2, 2018, pp. 784-793.

[9] Z. Song et al. "Origami Metamaterials for Ultra-Wideband and Large-Depth Reflection Modulation", Nature Communications, Vol. 15, 2024, p. 3181.

[10] C.-H. Kang, S.-J. Wu, J.-H. Tarng, "A Novel Folded UWB Antenna for Wireless Body Area Network", IEEE Transactions on Antennas and Propagation, Vol. 60, No. 2, 2012, pp. 1139-1142.

[11] S. N. Mahmood, A. J. Ishak, A. Ismail, A. C. Soh, Z. Zakaria, S. Alani, "ON-OFF Body Ultra-Wideband (UWB) Antenna for Wireless Body Area Networks (WBAN): A Review", IEEE Access, Vol. 8, 2020, pp. 150844-150863.

[12] D. M. Pozar, "Microwave Engineering", 4th Edition, John Wiley and Sons, 2011.

[13] F. Razzaz, S. M. Saeed, M. A. S. Alkanhal, "Compact Ultra-Wideband Wilkinson Power Dividers Using Linearly Tapered Transmission Lines", Electronics, Vol. 11, No. 19, 2022, pp. 1-10.

[14] N. H. A. Rahim, M. F. A. H. Saari, S. Z. Ibrahim, M. S. Razalli, G. S. Tan, "Wideband Power Divider Using Radial Stub for Six-Port Interferometer", Proceedings of the IEEE Asia-Pacific Conference on Applied Electromagnetics, Langkawi, Malaysia 11-13 December 2016, pp. 127-131.

[15] A. K. Vallappil, B. A. Khawaja, M. K. A. Rahim, M. Uzair, M. Jamil, Q. Awais, "Minkowski-Sierpinski Fractal Structure-Inspired 2 × 2 Antenna Array for Use in Next-Generation Wireless Systems", Fractal and Fractional, Vol. 7, No. 2, 2023, pp. 1-18.

[16] A. K. Vallappil, M. K. A. Rahim, B. A. Khawaja, M. N. Iqbal, "A Miniaturized Metamaterial-Loaded Switched-Beam Antenna Array System with Enhanced Bandwidth for 5G Applications", IEEE Access, Vol. 12, 2024, pp. 6684-6697.

[17] B. Khawaja, "Eight-Port Tapered-Edged Antenna Array with Symmetrical Slots and Reduced Mutual-Coupling for Next-Generation Wireless and Internet of Things (IoT) Applications", International Journal of Electrical and Computer Engineering Systems, Vol. 14, No. 6, 2023, pp. 695-702.

[18] E. J. Wilkinson, "An N-way Power Divider", IEEE Transactions on Microwave Theory and Techniques, Vol. 8, No. 1, 1960, pp. 116-118.

[19] T. Yu, "A Broadband Wilkinson Power Divider Based on the Segmented Structure", IEEE Transactions on Microwave Theory and Techniques, Vol. 66, No. 4, 2018, pp. 1902-1911.

[20] X.-P. Ou, Q.-X. Chu, "A Modified Two-Section UWB Wilkinson Power Divider", Proceedings of the International Conference on Microwave and Millimeter Wave Technology, Nanjing, China, 21-24 April 2008, pp. 1258-1260.

[21] O. Dardeer, T. Abouelnaga, A. Mohra, H. Elhennawy, "Compact UWB Power Divider, Analysis and Design", Journal of Electromagnetic Analysis and Applications, Vol. 9, No. 2, 2017, pp. 9-21.

[22] S. H. Ramazannia, S. Chamaani, S. A. Mirtaheri, F. K. M. Yazdi, M. Yazdani, "A UWB 1 to 4 Wilkinson Power Divider with Triple High-Q Band-Notched Characteristic Using U-Shaped DGS", Proceedings of the IEEE International Conference on Ultra-Wideband, Syracuse, NY, USA, 17-20 September 2012, pp. 302-305.

[23] M. Salman, Y. Jang, J. Lim, D. Ahn, S.-M. Han, "Novel Wilkinson Power Divider with an Isolation Resistor on a Defected Ground Structure with Improved Isolation", Applied Sciences, Vol. 11, No. 9, 2021, pp. 1-21.

[24] H. Peng, Z. Yang, Y. Liu, T. Yang, K. Tan, "An Improved UWB Non-Coplanar Power Divider", Progress in Electromagnetics Research, Vol. 138, 2013, pp. 31-39.

[25] P. K. Malik, P. N. Shastry, "Internet of Things Enabled Antennas for Biomedical Devices and Systems: Impact, Challenges and Applications", Springer Tracts in Electrical and Electronics Engineering, Springer Singapore, 2023.

[26] F. A. Shaikh, S. Khan, A. Z. Alam, M. H. Habaebi, O. O. Khalifa, T. A. Khan, "Design And Analysis Of 1-To-4 Wilkinson Power Divider for Antenna Array Feeding Network", Proceedings of the IEEE International Conference on Innovative Research and Development, Bangkok, Thailand, 11-12 May 2018, pp. 1-4.

[27] H. B. Tila et al. "Novel low cost 1-to-4 Microstrip Wilkinson Power Divider", Proceedings of the IEEE MTT-S International Wireless Symposium, Shanghai, China, 14-16 March 2016, pp. 1-4.

[28] K. Song, J. Yao, C. Zhong, S. R. Patience, "Miniaturised Wideband Four-Way Out-Of-Phase Power Divider Based on Marchand Balun", IET Microwaves, Antennas & Propagation, Vol. 13, No. 15, 2019, pp. 2682-2686.

[29] H. Zhu, A. M. Abbosh, L. Guo, "Wideband Four-Way Filtering Power Divider with Sharp Selectivity and Wide Stopband Using Looped Coupled-Line Structures", IEEE Microwave and Wireless Components Letters, Vol. 26, No. 6, 2016, pp. 413-415.

[30] Ö. Kasar, "Design and Implementation of Compact Four Way Wilkinson Power Divider for UHF Applications", Sigma Journal of Engineering and Natural Sciences, Vol. 38, No. 4, 2020, pp. 2193-2203

[31] M. Squartecchia, B. Cimoli, V. Midili, T. K. Johansen, V. Zhurbenko, "Design of a Planar Ultra-Wideband Four-Way Power Divider/Combiner Using Defected Ground Structures", Proceedings of the 47th European Microwave Conference, Nuremberg, Germany, 10-12 October 2017, pp. 9-12.

[32] G. Zhang, Z. Qian, J. Yang, J.-S. Hong, "Wideband Four-Way Filtering Power Divider with Sharp Selectivity and High Isolation Using Coshared Multi-Mode Resonators", IEEE Microwave and Wireless Components Letters, Vol. 29, No. 10, 2019, pp. 641-644.

[33] H. Zhu, A. M. Abbosh, L. Guo, "Wideband Four-Way Filtering Power Divider with Sharp Selectivity and Wide Stopband Using Looped Coupled-Line Structures", IEEE Microwave and Wireless Components Letters, Vol. 26, No. 6, 2016, pp. 413-415.

[34] K. Song, Y. Mo, Y. Fan, "Wideband Four-Way Filtering-Response Power Divider with Improved Output Isolation Based on Coupled Lines", IEEE Microwave and Wireless Components Letters, Vol. 24, No. 10, 2014, pp. 674-676.

[35] X. Zou, C.-M. Tong, D.-W. Yu, "Y-Junction Power Divider Based on Substrate Integrated Waveguide", Electronics Letters, Vol. 47, No. 25, 2011, pp. 1375-1376.

[36] S. Guo, K. Song, Y. Fan, "Compact Four-Way Suspended-Stripline Power Divider with Low Loss and High Isolation", International Journal of Microwave and Wireless Technologies, Vol. 12, No. 8, 2020, pp. 749-753.

[37] H. Chen, W. Che, X. Wang, W. Feng, "Size-Reduced Planar and Nonplanar SIW Gysel Power Divider Based on Low Temperature Co-Fired Ceramic Technology", IEEE Microwave and Wireless Components Letters, Vol. 27, No. 12, 2017, pp. 1065-1067.

Enhancement of Active Distribution Network Performance with Multiple Distributed Generators and DSTATCOMs using Reptile Search Algorithm

Original Scientific Paper

Ahmed Y. Hatata*
Shaqra University,
Department of Electrical Engineering, College of Engineering
Al-Dawadmi, Riyadh 11911, Saudi Arabia
ahmed_hatata@su.edu.sa

Rabab R. Eiada
Higher Institute of Engineering and Technology,
Communication and Electronics Engineering Department
New Damietta, Egypt
rabab.reda@ndeti.edu.eg

Thabet M. Alzahrani
Shaqra University,
Department of Electrical Engineering, College of Engineering
Al-Dawadmi, Riyadh 11911, Saudi Arabia
talzahrani@su.edu.sa

*Corresponding author

Abstract *– The integration of multiple DGs can introduce power quality issues and instability in the (DN) due to their intermittent and fluctuating nature. Distributed Static Compensators (DSTATCOMs) are devices that effectively manage and regulate both active and reactive powers, thereby maintaining the desired levels of reactive power in the DNs. This paper investigates the problem of determining the optimal number and placement of multiple DSTATCOMs within active DNs with multiple Distributed Generators (DGs) connected to them. To achieve the optimal sizing and allocation of multiple DSTATCOMs, a novel heuristic method called the reptile search algorithm (RSA) is introduced in this study. A combination of the RSA method and the loss sensitivity factor (LSF) are utilized to identify the optimal number, sizes, and locations of DSTATCOMs to enhance the performance of active DNs with different types of DGs and loads. The desired improvements include mitigating voltage deviation at nodes, minimizing system power losses, and alleviating overloading in feeders. The effectiveness of the algorithm is evaluated using an IEEE-33 bus DN, which is modified to incorporate different DGs and load types. To assess the efficiency of the proposed method, a modified particle swarm optimization (MPSO) algorithm is used for comparison. The results demonstrate that the RSA approach presented in this paper is robust in obtaining optimal solutions, offers fast and easy implementation, can be applied to large DNs, and outperforms the MPSO algorithm.*

Keywords: *DSTATCOM allocation and sizing, active distribution networks, RSA, MPSO, Active power losses*

Received: February 7, 2024; Received in revised form: August 13, 2024; Accepted: August 26, 2024

1. INTRODUCTION

Distribution networks (DNs) are traditionally configured to facilitate the one-way flow of energy from the transmission/generation systems to the end consumers. However, in active DNs, the energy can also be transmitted in the reverse direction from the consumers to the grids. This bidirectional power flow, attributed to Distributed Generators (DGs), introduces challenges related to the power quality of the DNs. The impact on power losses, whether decrease or increase, is predominantly contingent on factors such as network structure, configuration, DG type, capacity, and location. Furthermore, these dynamics influence the overall voltage profile of the system [1].

DSTATCOM, a component of distributed FACTS, serves as a shunt-connected apparatus for reactive power compensation, distinguishing itself among various compensation devices. Recently, DSTATCOM has gained prominence due to its operational flexibility and control capabilities within power systems. It effectively facilitates reactive power flow control, reduction of power losses, and regulation of voltage and current. Unlike series or shunt capacitors, DSTATCOM units operate without encountering operational challenges such as harmonics. DSTATCOM units offer several benefits, such as minimal harmonic distortion, low cost, high regulatory capability, and low power losses [2].

Recently, there has been significant scientific attention given to determining the optimal placement and capacity of renewable energy source (RES) based DGs to overcome the problems that were mentioned previously. The strategic placement of multiple DGs to simultaneously reduce the total active power losses (TACPL) and generation costs for both conventional generators and DGs was employed by [3]. The particle swarm optimization (PSO) method was implemented to attain the best locations and sizes of the RESs-based DGs (Solar, fuel cell, wind units) in the DNs. The main objective was to decrease the cost of DG units, the TACPLs, and Total Harmonic Distortion (THD) [4]. The ant lion optimizer was utilized to obtain the best sizes and settings of DGs in DNs while considering multiple objective functions such as reducing the cost of energy, TACPLs, and voltage drop and improving reliability [5]. A hybrid method based on Grey Wolf Optimizer method was introduced in [6] to place the REEs in the DNs. Another study utilized a hybrid optimization method based on combining the grasshopper optimization and cuckoo search algorithms to determine the best locations and capacities of DGs [7]. The optimal network reconfiguration and DG allocation were used to reduce TACPLs and enhance the stability of voltage using the adaptive modified whale optimization algorithm [8]. The quasi-oppositional grey wolf optimization technique with the Pareto concept were employed to optimally placement of DGs in the radial DNs to reduce the TACPLs and voltage drop [9]. The transient search optimizer (TSO) method was implemented to find the optimal DGs allocation to reduce the power losses and voltage deviation (VD) and enhance the stability of voltage [10]. Another study used the Whale Optimization technique (WOT) to obtain efficient DGs ratings and locations [11]. An analytical method to attain the best locations and rating of RESs based DGs in the DNs for minimizing the power flow limits, DG capacities, bus voltage limits and power loss using coot bird optimization method (CBOM) [12]. Although previous research laid the foundation for determining the best locations for DGs in the electrical DN, it was insufficient to solve the problems that arise from the presence of these DGs within the electrical DN.

Numerous research endeavors have leveraged DSTATCOM in distribution networks, reaping various advantages. Nevertheless, the size and placement of the DSTATCOM units are crucial factors that must be carefully considered to maximize the benefits from their installation. The optimal determination of distribution system tie switches, coupled with the optimal locations and sizes of DSTATCOM units, has been a key focus to achieve favorable operational conditions. To address the intricacies of this combinatorial nonlinear optimization problem, the Differential Evolution Algorithm (DEA) was employed [13] [14]. Additionally, the bio-inspired Cuckoo search method was used to obtain the best sizes of the DSTATCOM units in the radial DNs, while the loss sensitivity factor (LSF) was utilized to find the optimal locations [15]. Furthermore, the Ant-Lion optimization (ALO) method was applied in [16] to appropriately allocate the DSTATCOM. The objective was to increase the cost-benefit resulting from reducing the expense of purchasing power from the electrical power system, and the cost of DSTATCOM. The location strategy of DSTATCOM units within the radial DNs was accomplished using the black widow optimization technique [17]. The main objective was to reduce the TACPLs while considering different technical and economic factors such as annual cost savings and the voltage stability index (VSI). Furthermore, the modified capuchin search algorithm was introduced to address multiple issues related to optimal DSTATCOM allocations [18]. A multi-objective approach was employed to compensate the reactive power in radial DNs [19]. It focused on finding the optimal simultaneous allocation of DSTATCOMs and static capacitors by employing fuzzy decision making.

Other researchers have concurrently utilized DGs along with DSTATCOMs in the DNs to achieve both power loss minimization and power quality enhancement. In [20], an optimal placement strategy for multiple DG and DSTATCOM units in DN resulted in reduced line losses and improved power quality, as indicated by THD. Considering load models, ref. [21] utilized Genetic Algorithm (GA) to improve the total VD and minimize the TACPLs by integrating DGs with DSTATCOM units. The allocations of both DSTATCOM units and DGs were executed by implementing the LSF [22]. The bacterial foraging optimizer technique was introduced to ascertain the optimal size of DGs and DSTATCOM, considering different models of loads to minimize the VD, reduce TACPLs, improve VSI, and enhance the security of the DN [23]. Meanwhile, a novel Lightning Search Algorithm was applied to simultaneous allocate the DGs and DSTATCOM units in radial DNs to minimize TACPLs, Total VD, and maximize the value of VSI [24]. The whale optimization algorithm (WOA) was used to simultaneous allocate DGs and DSTATCOMs in the radial DNs [25]. The main goal of the algorithm was to reduce both the TACPLs of the system and the operating costs associated with DGs and DSTATCOMs. A planning strategy was introduced to optimally reconfigure the feeders, allocate and sizing the DSTATCOM in unbalance radial DNs [26]. The seagull optimization algorithm (SOA) was applied to address this mixed-integer nonlinear planning problem to achieve minimum TAC-

PLs. A hybrid technique based on Sine Cosine Algorithm (SCA) and Moth Flame Optimization (MFO) method was proposed for performing the exploitation and exploration phases in finding the best locations of the capacitors and DGs in DNs [27]. The TACPLs were applied as an objective function to integrate DGs and capacitors in DNs. The DSTATCOM and the PV DGs were integrated into the 33-bus and 69-bus DNs [28]. The modified homonuclear molecules optimizer technique is used to attain the optimal positions of them while minimizing the cost of devices' integrations, total voltage deviations (TVD), and TACPLs. The results demonstrated that the algorithm could reduce the TACPLs by 94.27% for IEEE 33-bus and 97.87% for IEEE 69-bus systems. The honey badger method was utilized to attain the optimal location and size of various types of DGs to reduce the TACPLs of the DNs [29]. The power LSF was used to order the buses for optimally installing the SCs and DGs in the radial DNs to accelerate conversion of the honey badger algorithm. Also, two types of DGs were applied in the IEEE 69-bus standard radial DN. The Slime Mould algorithm was used to locate the PVDGs and the distribution static Var compensators (DSVCs) to satisfy economic and technical objectives [30]. The selected objectives consisted of minimizing the TACPL, TVD, Total Reactive Power Loss (TRPL), Operating times of the overcurrent relays, and investment costs of both PVDGs and DSVCs. Previous studies did not consider the optimal number of DSTATCOMs, and the issue of relieving distribution systems by increasing DG penetration levels through optimal DSTATCOM allocation was not explored. Additionally, similar research did not investigate the impact of different load types and various DG types on DSTATCOM optimal allocation.

The proposed algorithm in this paper tackles the combinatorial non-linear optimization problem of identifying the optimal number, location, and size for multiple DSTATCOM units in the DNs. This innovation is advantageous for distribution system operators seeking to enhance operational performance. The algorithm considers the presence of already allocated DGs and has the capability to remove DG penetration level restrictions. Furthermore, this work delves into the behavior of various load and DG types. The main contributions of this paper can be listed as:

- Propose a new methodological approach for achieving the best number, location, and sizing of DSTATCOMs in DNs.
- Determine the optimum location and sizing by the proposed Reptile Search Algorithm (RSA) and MPSO.
- Examine the impacts of various load types, DG types, and load factors on the optimal allocation and sizing of DSTATCOMs.
- The results of the RSA are compared with Modified PSO algorithms.

The structure of this paper is outlined as: The comprehensive review of modeling the DSTATCOM as a highly effective remedy in DNs is provided in section 2. Illustrating problem formulation and introducing the LSF as an indicator of potential bus locations for DSTATCOM placement are presented in section 3. Moreover, the proposed Reptile Search Algorithm (RSA) and MPSO method are introduced. The IEEE-33 bus radial DN is identified in section 4 as a test distribution system. Section 5 encompasses numerical analysis and simulation studies conducted using the proposed RSA and MPSO algorithms. Finally, the paper concludes with closing remarks in section 6.

2. DSTATCOM MODELING

The DSTATCOM functions as a variable current source (VCS) and is connected as a shunt element. It is able to absorb or inject reactive and active currents. To enhance the dynamic rating of the capacitive range, the DSTATCOM can be connected in parallel with a fixed capacitor. By incorporating energy storage apparatus on the DC side via DC/DC converter, the active power is temporarily interchanged with the grid, especially during events such as large voltage sags or momentary interruptions. The DSTATCOM can be composed of voltage source converter and energy storage, enabling injection of both reactive and active power. However, injecting active power for an extended duration is constrained by voltage regulation and the energy storage system's capacity limit. In the second model, the DSTATCOM is composed of a voltage source converter and small capacitor. In this configuration, only reactive power can be exchanged between the DSTATCOM and the AC system [31].

In an electric DN comprising two nodes labeled *j* and *k*, a DSTATCOM unit is linked to bus k to regulate voltage. The installation of the DSTATCOM unit resulted in an increase in the voltage at node k from V_k to V_k^n. The amount of reactive power injected from the DSTATCOM denoted as QDSTATCOM, can be given by:

$$Q_{DSTATCOM} = Img\ (V_k^n . I_{DSTATCOM}{}^*) \quad (1)$$

$$V_k^n = V_k^n \angle\, \alpha_{new} \quad (2)$$

$$I_{DSTATCOM} = |I_{DSTATCOM}| \angle \left(\alpha_n + \frac{\pi}{2} \right) \quad (3)$$

where, $I_{DSTATCOM}$ represents the injected current from the DSTATCOM and α_{new} represents the voltage angle at node k after connecting DSTATCOM device.

3. DSTATCOM OPTIMAL SIZING AND ALLOCATION

The main objectives of the optimal locations and sizes of DSTATCOM units in the DN with DGs are to minimize the TACPLs and enhance the TVD in the system. These objectives are subject to constraints that ensure the system's feeder currents and voltage profile remain within specified bounds. The identifying of candidate buses for DSTATCOM units' allocation is accomplished by applying the LSF method, implemented through a MATLAB code. Subsequently, the optimal number, location, and size for multiple DSTATCOM units are obtained by ap-

plying the proposed RSA method, which is also implemented through a MATLAB code specifically designed for this purpose. The optimal solution is defined as the best among the candidate buses, where the TACPL is minimized, and the TVD is enhanced. The proposed algorithm can be divided into the following key stages.

- Stage 1, Using the LSF method for defined the DSTATCOM candidate buses.
- Stage 2, Using RSA and MPSO as optimization methods for calculating the best number of DSTATCOM units and their sizes.
- Stage 3, Comparing the values of the TACPLs and TVD of various combinations for defined the optimal DSTATCOM number.

3.1. DSTATCOM CANDIDATE BUSES BY USING LSF

The LSF method is employed to obtain the optimal locations of DSTATCOMs in the DN regarding power loss minimization. LSF has the ability to forecast the bus where the installation of DSTATCOM will result in the most significant reduction in power losses. Consequently, these critical locations serve as potential bus candidates for the allocation of DSTATCOM. To illustrate the LSF concept, let's examine a two-bus distribution system with a given impedance $R+jX$ connected between the two buses and a load of P_e+jQ_e as shown in Fig. 1. The TACPLs through the line could be expressed as given in Eq. 4 [31].

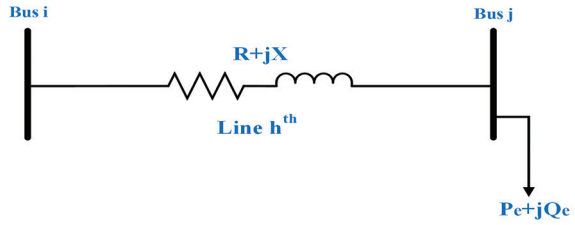


Fig. 1. Simple two-bus distribution system

$$P_{LL}(h) = \frac{[P_e^2(j) + Q_e^2(j)] * R(h)}{[V(j)]^2} \tag{4}$$

where, $P_{LL}(h)$ is the real power loss in distribution line hth, $V(j)$ is the value of voltage at bus jth, and $R(h)$ is the distribution line hth resistance. So, the LSF is calculated by:

$$LSF = \frac{\partial P_{LL}(j)}{\partial Q_e(j)} = \frac{[2 * Q_e(j)] * R(h)}{[V(j)]^2} \tag{5}$$

The LSF is computed using the Backward-Forward Distribution Load Flow technique [31]. Their values are arranged in descending sequence, and the respective buses indices are documented in the position vector, $Pp(i)$. Furthermore, the arranged components of $(\partial P_{LL}(j)/\partial Q_e(j))$ in $P_p(i)$ dictate the sequence of candidates for the optimal allocation of the DSTATCOMs. The normalized voltage magnitudes at the buses corresponding to $P_p(i)$ are computed by:

$$Norm\,V(i) = \frac{|V(i)|}{0.95} \tag{6}$$

where the division of $|V(i)|$ by 0.95 refers to the voltage acceptable tolerance of 5%. The $Norm(i)$ function assesses the requirement for reactive power compensation at the arranged buses. Therefore, for DSTATCOM placement, buses with $Norm\ V(i)$ greater than 1.01 are considered as DSTATCOM candidates. The process of identifying DSTATCOM candidates is outlined as follows:

Step 1: perform load flow analysis on the DN in the base case.

Step 2: LSF is calculated for each feeder in the DN.

Step 3: the values of the LSF are organized in descending sequence, and the associated buses are documented in the vector $Pp(i)$.

Step 4: for the corresponding feeders, the values of the normalized buses voltage are calculated.

Step 5: according to the $Norm(i)$, the reactive power compensation is established by selecting buses with a value less than 1.01 as candidates for installing DSTATCOM units. Conversely, buses with a Norm(i) exceeding 1.01, indicating that their voltage profile is within acceptable limits, do not require reactive power compensation.

3.2. DSTATCOM OPTIMAL SIZES DETERMINATION

In this paper, part of the proposed method involves obtaining the optimal sizes of DSTATCOMs. Mathematical formulation is employed to determine the sizing of DSTATCOM, posing it as an optimization problem. The objective function is optimized while adhering to specific equality and inequality constraints as illustrated below.

3.2.1. Objective function

The objective function (OF) is nonlinear, aiming to minimize the TACPs and enhance the TVD in the DN. It can be expressed as follows:

$$TACPL = \sum_{h=1}^{N_f} P_{LL}(h) \tag{7}$$

$$TVD = \sum_{j=1}^{n} (V_{ref} - V_j)^2 \tag{8}$$

$$OF = \min(\omega_1 TACPL + \omega_2 TVD) \tag{9}$$

where, N_f and n are the total number of feeders and nodes, respectively. V_{ref} and V_j are the reference and node voltage magnitude. ω_1 and ω_2 are the weighted factors where the summation these factors should be equal one. In this paper, many combinations of the weighted factors are evaluated by applying them to the IEEE 33-bus test system to reach the best one. Finally, the best one is used to obtain the OF. For this analysis, the power losses have the higher weight (0.63)

since it is important for distribution system. The total voltage deviations receive a weight of 0.37.

3.2.2. Constraints

The objective function is bound by multiple operational constraints, encompassing voltage limitations, feeder current restrictions, constraints on reactive power compensation, power balance constraints, and limits on both active and DSTATCOM sizes. These constraints can be articulated by:

1. - Feeder current limits

The primary feeders within the DNs are capable of providing a maximum magnitude of current, defined by:

$$\left|I_{fi}\right| \leq I_{fi}^{max}\ ;\ \ i = 1,2, \dots \dots, N_f \tag{10}$$

where, I_{fi}^{max} represents the maximum current in feeder ith and I_{fi} represent the flow current in feeder i^{th}.

2. - Bus voltage constraints

$$V_{j\,min} \leq V_j \leq V_{j\,max} \quad , j = 1,2, \dots \dots, n \tag{11}$$

where, V_{jmax} and V_{jmin} represent the upper and lower voltage limits at busbar j^{th}, respectively.

3. - Generator reactive power limits

$$Q_G \leq Q_G^{max} \tag{12}$$

where, Q_G and Q_G^{max} are the current value and maximum permissible value of generator reactive power.

4. - DSTATCOM size

$$Q_{DSTATCOM}^{min} \leq Q_{DSTATCOM} \leq Q_{DSTATCOM}^{max} \tag{13}$$

where, $Q_{DSTATCOM}^{max}$ and $Q_{DSTATCOM}^{min}$ represent the upper and lower limits of injected reactive power from DSTATCOM.

5. - Slack bus active power limit

$$P_S \leq P_S^{max} \tag{14}$$

where, P_S and P_S^{max} are the current and maximum permissible active power generated at the slack bus, respectively.

6. - Power balance constraints

Maintaining a balance among power generation, power demand, and power loss, is of utmost importance in DNs to uphold system stability and reliability.

$$P_{LOSS} = \sum_{j=1}^{n} P_{Gj} - \sum_{j=1}^{n} P_{Dj} \tag{15}$$

$$Q_{LOSS} = \sum_{j=1}^{n} Q_{Gj} - \sum_{j=1}^{n} Q_{Dj} \tag{16}$$

The succeeding section outlines the optimization methods that are employed to determine the optimal size of DSTATCOMs. The main problem is to obtain the best locations and sizes of DSTATCOM units. The subsequent section outlines the principles and implementation strategies of the proposed MPSO and RSA.

3.2.3. The proposed optimization algorithms

1. - Modified particle swarm optimization Algorithm (MPSO)

The PSO operates on the premise that every particle in the population represents a potential solution to the problem. It relies on the concept of parallel search within a group of particles. These particles, working collectively, converge towards the optimal value by leveraging their current velocity and positions [32]. The fundamental rules for updating the position and velocity are outlined as follows [32-33]:

$$V_i^{t+1} = wV_i^t + A_1R_1 \times (P_i^t - X_i^t) + A_2rR_2 \times (G^t - X_i^t) \tag{17}$$

$$X_i^{t+1} = X_i^t + V_i^{t+1} \tag{18}$$

where V_i^t and V_i^{t+1} are the velocity of particle ith at iteration t and t+1, respectively. P_i^t and P_i^{t+1} represent the position of particle ith at iteration t and t+1, respectively. $R1$ and $R2$ are random numbers between 0, and 1, and w is the weighed inertia coefficient. $A1$ and $A2$ are accelerating factors.

When the PSO method is dealing with a large search space, its convergence speed tends to slow down, resulting in suboptimal solutions. Additionally, when dealing with large and complex datasets, the algorithm may produce unsatisfactory results in terms of accuracy. In order to overcome these constraints and improve both convergence speed and accuracy of the PSO, several variants of the PSO have been proposed [32]. This paper utilizes various variants of the PSO algorithm, PSO with time-varying acceleration coefficient (PSO-TVAC) and damped inertial weight (PSO-DIW) to conduct a comparative analysis. TVAC and DIW values in the PSO algorithm are updated by applying Eqs. (19)-(22) [34].

$$\alpha_1 = (\alpha_{1f} - \alpha_{1i})\frac{iter}{Max_{iter}} + \alpha_{1i} \tag{19}$$

$$\alpha_2 = (\alpha_{2f} - \alpha_{2i})\frac{iter}{Max_{iter}} + \alpha_{2i} \tag{20}$$

$$w = (w_{max} - w_{min})\frac{t_m - t}{t_m} + w_{min} \tag{21}$$

$$w = w \times w_{damp} \tag{22}$$

where α_{1f} and α_{1i} represent the final and initial values, respectively, of the first acceleration coefficients. α_{2f} and α_{2i} represent the final and initial values, respectively, of the second acceleration coefficients. w_{min}, w_{max} and w_d represent the lower, upper, and damped values, respectively, of the weighed inertia.

In this research, Eqs. 19-21 are employed to dynamically adjust the acceleration and weighted coefficients in the PSO-TVAC variant. However, Eq. 22 is specifically used to adjust the coefficients in the PSO-DIW variant. The flowchart in Fig. 2 illustrates the different PSO variants, including the traditional PSO algorithm, employed to optimize sizes of the DSTATCOM units.

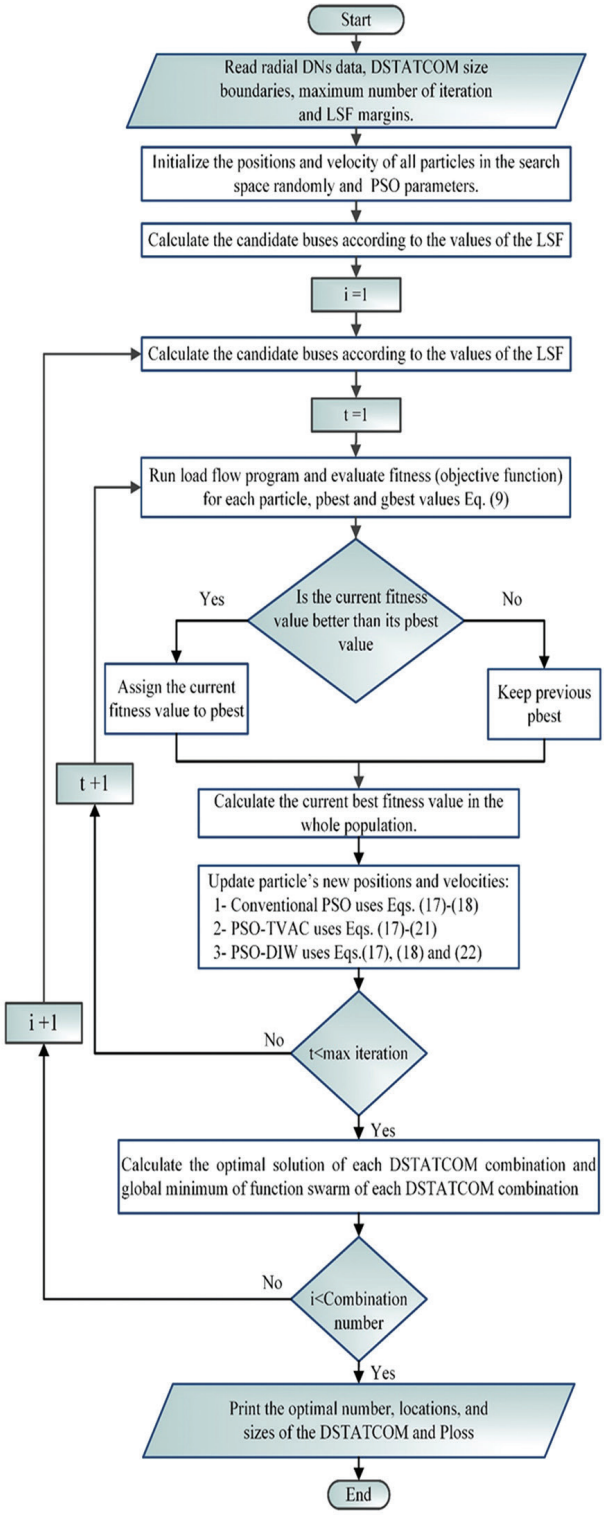


Fig. 2. Proposed MPSO algorithm flowchart for DSTATCOM optimal sizing

2- Reptile search algorithm (RSA)

The RSA is a metaheuristic method that derives its inspiration from the natural behaviors of crocodiles. It incorporates elements such as the hunting behavior, enveloping mechanism, and social dynamics observed in crocodiles. By employing a swarm-based approach, the RSA algorithm leverages these principles to effectively guide its search process [35]. The RSA initialization formula is employed to generate the initial solution in a random manner, ensuring it falls within the solution domain. The solution domain encompasses the entire range of potential solutions for the given optimization problem [35].

$$L_{ij} = L_v + r_1.(U_v - L_v) \tag{23}$$

where L_{ij} is the value of the jth dimension of the ith crocodile. L_v and U_v minimum and maximum boundary values of the search domain, respectively. r_1 is random number in range 0 and 1, and n is the size of populations. The mathematical modeling of the RSA is expressed in the following subsections.

- Enveloping Mechanism

The RSA incorporates the concept of enveloping to explore the search space in a manner that prioritizes promising regions while avoiding less favorable ones. This approach enables the algorithm to effectively navigate complex optimization landscapes. In the global search phase, crocodiles engage in extensive and wide-ranging walks, which can be quantified through iterations. The mathematical models capturing this mechanism are described in references [35, 36].

$$\boldsymbol{L_{ij}(t+1)} = \begin{cases} \boldsymbol{L^t_{j,best}} \cdot \left(-\boldsymbol{\vartheta^t_{ij}} \times \boldsymbol{\beta} \times \left(\dfrac{\boldsymbol{L^t_{j,best}} - \boldsymbol{L^t_{ij}}}{\boldsymbol{L^t_{j,best}} + \boldsymbol{\varepsilon}}\right) \times \boldsymbol{r_1}\right) & \boldsymbol{t \le \dfrac{T_m}{4}} \\ \boldsymbol{L^t_{j,best}} \times \boldsymbol{L^t_{j,rand}} \times \boldsymbol{2r_2}\left(\dfrac{\boldsymbol{T_m - 1}}{\boldsymbol{T_m}}\right) \times \boldsymbol{r_1} & \boldsymbol{\dfrac{T_m}{4} \le t \le \dfrac{2T_m}{4}} \end{cases} \tag{24}$$

where t and T_{max} are the current and maximum number of iterations, respectively. $\vartheta_{ij}^{\ t}$ denotes the hunting operator value for ith solution at jth location during the tth iteration. $L^t_{j,best}$ represents the optimum solution at jth location during the tth iteration. The hunting operator is determined by the following expression[44]:

$$\vartheta^t_{ij} = L^t_{j,best} \times \left(\sigma + \frac{L^t_{ij} - Avr\left(L^t_{ij}\right)}{L^t_{j,best} \times (U_v - L_v) + \varepsilon}\right) \tag{25}$$

where r_2 represents a random number ranging from -1 to 1, σ is a constant to controls the accuracy of the exploration, ε is a minimum number that guarantees the denominator does not reach zero, and Avr refers to the average value.

- Hunting Mechanism

The hunting mechanism within the RSA is analogous to the enveloping mechanism and comprises two distinct phases: hunting coordination and cooperation. These phases are designed to explore the search domain and facilitate the identification of the best possible solution while capturing the prey.

The specific definition of these phases is based on the number of iterations. Hunting coordination is applied for iterations in range $0.5\ T_{max} < t \le 0.75\ T_{max}$, whereas hunting cooperation is employed in range $0.75\ T_{max} < t \le T_{max}$.

The hunting mechanism is mathematically represented by the following equations as described in references [33, 34].

$$L_{ij}(t+1) = \begin{cases} L^t_{j,best} \times P_{ij} \times r_1 & \frac{T_m}{2} < t \le \frac{3T_m}{4} \\ L^t_{j,best} - \vartheta^t_{ij} \times \varepsilon - \left(\frac{L^t_{j,best} - L^t_{ij}}{L^t_{j,best} + \varepsilon}\right) \times r_1 & \frac{3T_m}{4} < t \le T_m \end{cases} \quad (26)$$

- Modified Reptile Search Algorithm (MRSA)

The RSA algorithm, despite its effectiveness, is subject to certain limitations, including the trapping of local minima, high computational complexity, and slow convergence speed. To overcome these challenges, several modifications have been proposed for the original RSA. One particular adjustment involves incorporating a sine operator into the high walking phase of the previous RSA algorithm. This modification draws inspiration from the SCA [37]. By introducing the sine operator, the algorithm becomes capable of avoiding local minimum trapping and enhancing global exploration. The sin operator is inserted into Eq. 24 and is modified as follows.

$$L_{ij}(t+1) = \begin{cases} L^t_{j,best} + \left(r_2 \times \sin(r) \times \left|r_3 \times L^t_{j,best} - L^t_{ij}\right|\right) & t \le \frac{T_m}{3} \\ L^t_{j,best} \times L^t_{j,rand} \times 2r_2 \left(\frac{T_m - 1}{T_m}\right) \times r_1 & \frac{T_m}{4} \le t \le \frac{2T_m}{4} \end{cases} \quad (27)$$

In the given context, the variables r_2 and r_3 represent randomly selected numbers within the range of [0, 1]. The adoption of the chaotic inverse learning strategy by all individuals results in increased computing-related to costs and hampers algorithm convergence. To tackle this problem, the paper introduces the linear decreasing population strategy. As the iteration continues, there is a gradual decrease in the number of individuals utilizing the chaotic backward learning strategy. The precise mathematical formula for implementing this strategy is described by:

$$P = r_1 \times \left(\frac{(P_{min} - P_{max}) \times t}{T_m} + P_{max}\right) \quad (28)$$

where P_{max} and P_{min} represent the maximum and minimum number of populations, respectively. P represents the number of chaotic backward learning strategy populations.

In this research the RSA technique is implemented, aiming to calculate the optimal size of the DSTATCOM. A detailed depiction of the proposed RSA can be observed in Fig. 3, showcasing the flowchart of the MRSA.

4. DSTATCOM OPTIMAL NUMBER DETERMINATION

To identify the optimal number of DSTATCOMs, the solution with the minimum objective function is sought. Consequently, all conceivable combinations of DSTATCOM units at candidate buses are compared based on their respective objective function values. The algorithm proposed in this study calculates the minimum TACPLs and enhance the TVD in the DN and determines the optimal sizes of DSTATCOMs at each potential location. These values are saved for comparison, making it easier to identify the combination that results in the lowest power loss. This combination represents the optimal number, locations, and sizes of DSTATCOMs.

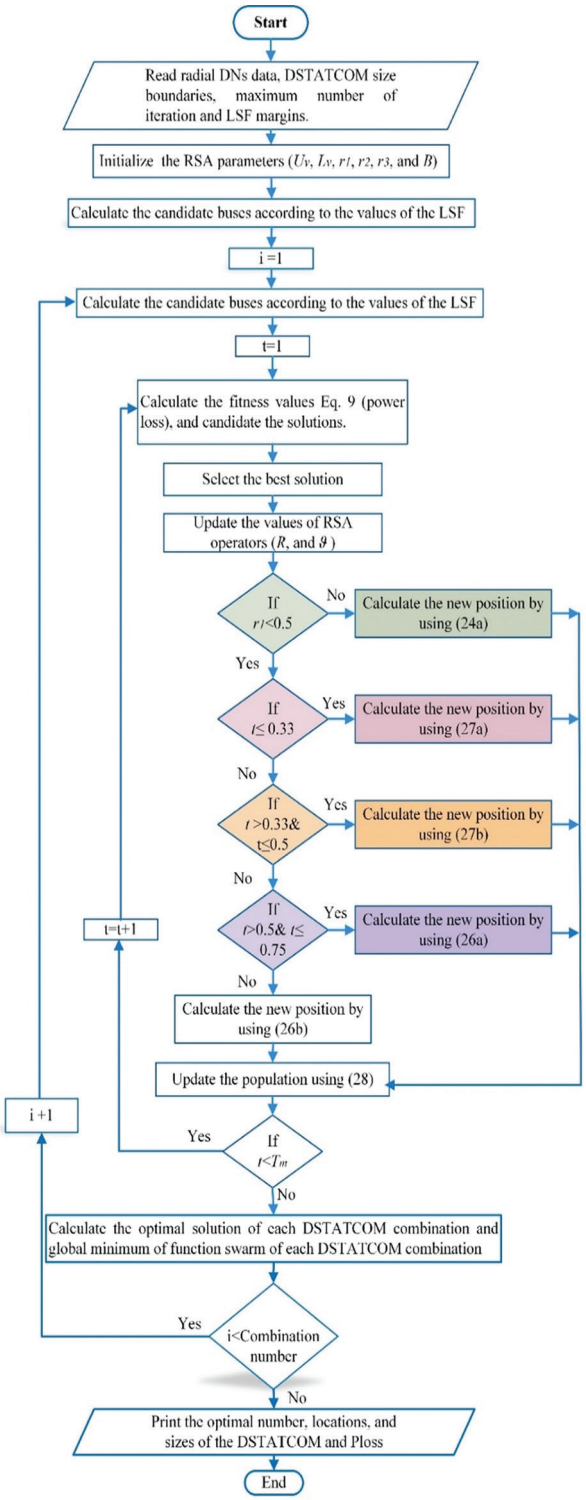


Fig. 3. Flowchart of the MRSA method

5. TEST SYSTEM: IEEE-33 BUS SYSTEM.

It comprises 33 buses and 32 feeders, with base voltage value of 12.66 kV and 100 MVA for apparent power. The total power consumed by the system is 2300 kVAR

for reactive power and 3715 kW for real power. In the base scenario, the TACPLs are 210.9 kW and 135.03 kVAR for reactive power. Bus 1 serves as the power feeder linked to the transmission network, while the other buses can potentially accommodate DSTATCOM [21]. The maximum permissible value of Q_G^{max} and P_S^{max} are 2.5 MVAR and 4 MW, rspectively. The single-line diagram of the system is depicted in Fig. 4. The line and bus data in IEEE-33 bus distribution network in light loading are shown in Table A1 in the appendix.

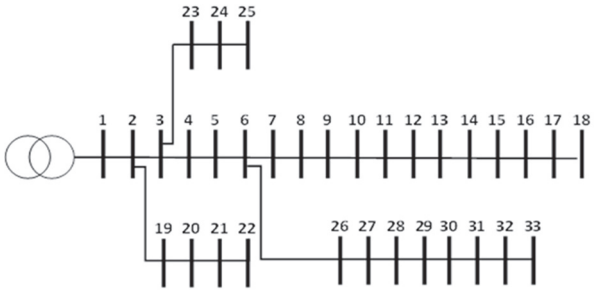


Fig. 4. Single line diagram of IEEE-33 bus radial DN

The permissible range for DSTATCOM sizes is set between 0.1 MVAR and 10 MVAR. The maximum allowable voltage deviation at the buses is limited to 10%. Branches 1 to 5 are designed to carry a maximum current of 400 A, while branches 6 to 7 and branches 25 to 27 have a constraint of 300 A. The current carrying capacity for all other branches, including tie lines, is set at 200 A.

6. SIMULATION RESULTS AND DISCUSSION

To perform load flow analysis and compute power loss for the optimal placement of multiple DSTATCOM units, a custom-coded program is introduced utilizing the MATLAB software. Both LSF and the proposed optimization algorithms (MPSO and RSA) are implemented to an IEEE-33 bus DN. This showcases the efficacy of the proposed algorithm in improving system performance. The simulation results prioritize TVD enhancement and TACPLs minimization while adhering to specified constraints. The influence of various DG models and various load types on the optimal placement of DSTATCOM units is also examined.

A population size of 50 particles is chosen for the proposed MPSO and RSA algorithms, and the maximum number of iteration criterion is defined as 250 iterations. The parameters of the MPSO and the MRSA algorithms are illustrated in Table 1.

Table 1. MPSO and MRSA algorithms parameters

Parameter	Value
MPSO	
Max. No. of iterations	250
Population size	50
Inertia weight (w)	0.7
Cognitive parameter ($c1$)	1.5
Social parameter ($c2$)	1.5
RSA	
Max. No. of iterations	250
Population size	50
Dimensions	6
α	0.1
β	0.05

Assuming an extension to the IEEE 33-bus radial DN with a new community installed at buses 24, 25, and 30, the system's topology is taken into account. The objective is to investigate the influence of three factors: varying load factors (LFs), load types, and DG types on the optimal placement of multiple DSTATCOM units in the DN. To streamline the exploration of various combinations of these factors and reduce the number of scenarios, the simulations are organized into three stages as outlined below:

Stage 1: Simulating various load types and selecting the worst case.

Stage 2: Simulating various DG types and selecting the worst case.

Stage 3: Simulating various LF.

Upon completion of each stage, the least favorable type is selected, and the following factors are systematically varied. This approach aims to facilitate a comprehensive assessment of the effectiveness of DSTATCOM units in addressing various operational scenarios within active distribution systems.

6.1. STAGE #1 LOAD TYPE IMPACT ON OPTIMAL LOCATION AND SIZING OF DSTATCOM.

During the first stage, a modification in the load type for the load center of modified configuration at buses 24, 25, and 30 is performed in the IEEE 33-bus radial DN, while the other buses are constant power load type. Additionally, the characteristics of DSTATCOM units, encompassing the optimal number, locations, and sizes, are compared across these load types. Furthermore, the study explored the influence of load types on voltage profiles and power losses. In this stage, the intention is to keep all other variables constant, enabling a focused examination of the impact of load types. Accordingly, the LF is set to unity, and no DG is integrated into the test radial DN. Four distinct cases are chosen, each representing one of the four load types: residential load (RL), commercial load (CL), industrial load (IL), and constant power load (CPL). These four cases can be illustrated in the following subsections.

6.1.1. Case# 1: CPL type (unity load factor (LF))

Before installing the DSTATCOM units, the TACPLs is 210.9 kW, and the bus with the minimum voltage value is 0.903 p.u at bus 18 while no overloaded feeders. In the LSF computation, the values for test system branches are determined, leading to the identification

of potential candidate buses. The top five candidate buses are ranked as follows: 6, 28, 8, 29, and 30.

The main objective is to improve the performance of the IEEE 33-bus radial DN. After implementing the proposed MPSO algorithm, the TACPL is decreased from 210.9 kW to 143.27 kW, resulting in a significant reduction of 32.1%. The best number of DSTATCOM units is determined to be four, allocated at buses 6, 8, 29, and 30, with sizes of 0.53743, 0.44425, 0.37655, and 0.43043 MVAR, respectively. Furthermore, the MRSA algorithm has the ability to decrease the TACPL to 132.27 kW leading to a substantial decrease of 37.31% in case of installing four DSTATCOM units at the same buses.

The comprehensive voltage profile, both before and after DSTATCOM allocation, is detailed in Fig. 5. Consequently, the system's voltage profile is improved. The MPSO algorithm can improve the minimum voltage to 0.925 p.u, while the MRSA algorithm produced a slightly higher values of 0.929 p.u. These values satisfy the criteria for acceptable voltage limits. The proposed MPSO and MRSA algorithms exhibit effectiveness in balancing feeder loadings, which is evident in the reduction of the TACPL in the test system.

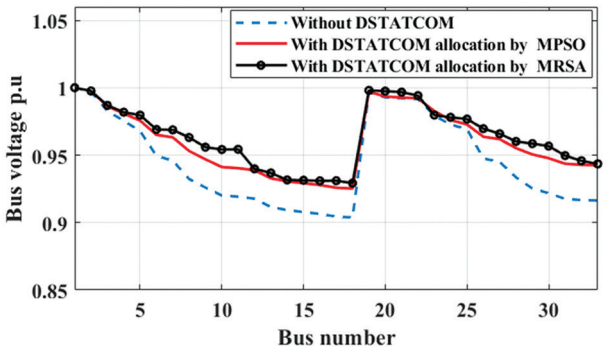


Fig. 5. Bus voltage for DSTATCOM installation in case#1 of CPL type

6.1.2. Case# 2: Industrial load type

The radial DN of IEEE 33-bus configuration is equipped with an IL type, which is installed to the new load center of modified configuration. This load type is connected to buses 24, 25, and 30, and there is no DG installed to the test system. The TACPLs in the system before allocating DSTATCOM units is 188.5 kW. The bus with minimum voltage value is 0.906 p.u at bus 18 while no overloaded feeders.

In the LSF computation, the values for test system branches are determined, leading to the identification of potential candidate buses. The candidate buses are ranked as follows: 6, 8, 28, 29 and 9. After applying the proposed optimization algorithms (MPSO and MRSA). It is observed that the TAPLs were decreased to 142.38 kW, reflecting a 24.47% reduction in the TACPL using the proposed MPSO algorithm. The optimal number of DSTATCOM units is determined to be four, strategically placed at buses 6, 9, 28, and 29, with sizes of 0.39862, 0.32428, 0.34033, and 0.3864 MVAR, respectively. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 134.54 kW leading to a substantial decrease of 28.31% in case of installing four DSTATCOM units at the same buses.

As a result of the allocation of DSTATCOM units, the system voltage profile has improved. The bus with the lowest voltage level is now bus 18, which has a voltage of 0.924 p.u in case of using the proposed MPSO algorithm while the proposed MRSA algorithm can produce a slightly higher value of 0.932 p.u . The complete voltage profile, both before and after DSTATCOM allocation, is provided in Fig. 6.

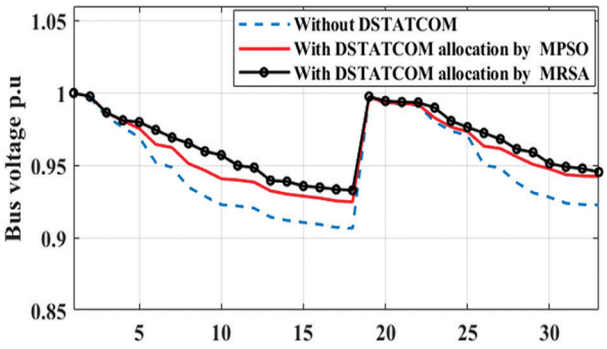


Fig. 6. Bus voltage for DSTATCOM installation in case#2 of IL type

6.1.3. Case# 3: Residential load type

In this scenario, the IEEE 33-bus radial DN incorporates a RL type, which is installed to buses 24, 25, and 30 and there is no DG installed to the test system. The proposed optimization algorithms (MPSO and MRSA) for optimal DSTATCOM allocation are employed to identify the worst load type. Prior to the allocation of DSTATCOM units, the TACPL in the test system is recorded at 191.87 kW. The bus with the minimum voltage value is 0.906 p.u at bus 18 while no overloaded feeders.

After implementing the proposed MPSO algorithm for DSTATCOM allocation, it is observed that the TACPLs decreased to 141.43 kW, resulting in a significant reduction of 26.29% in the TACPL. The optimal number of DSTATCOM units is attained to be four, strategically placed at buses 6, 9, 28, and 29. The respective sizes of these units are 0.44275, 0.43531, 0.4203, and 0.2623 MVAR. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 139.34 kW leading to a substantial decrease of 27.38% in case of installing four DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the minimum voltage level is bus 18, with 0.927 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.937 p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation, can be illustrated in Fig.7.

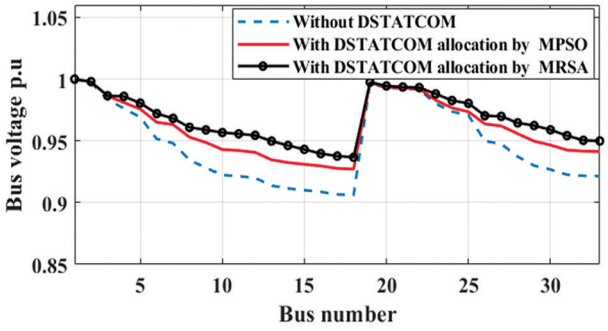


Fig. 7. Bus voltage for DSTATCOM installation in case#3 of RL type

6.1.4. Case# 4: system with commercial load

In this scenario, the IEEE 33-bus radial DN incorporates a CL type, which is installed to buses 24, 25, and 30 and there is no DG installed to the test system. Prior to DSTATCOM allocation, the power loss was 192.34 kW. The bus with the minimum voltage value is 0.906 p.u at bus 18 while no overloaded feeders.

Following implementing the proposed MPSO algorithm for optimal DSTATCOM allocation, the five candidate buses are sequenced as 6, 28, 8, 29, and 30. The TACPLs were decreased to 137.94 kW, representing a 28.28% reduction in the TACPL. The optimal configuration includes five DSTATCOM units placed at buses 6, 8, 28, 29, and 30, with sizes of 0.28306, 0.39505, 0.244, 0.4114, and 0.4254 MVAR, respectively. Furthermore, the proposed RSA algorithm has the ability to decrease the TACPL to 132.46 kW leading to a substantial decrease of 31.13% in case of installing five DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the lowest voltage level is bus 18, with 0.926 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.935 p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation, can be illustrated in Fig.8.

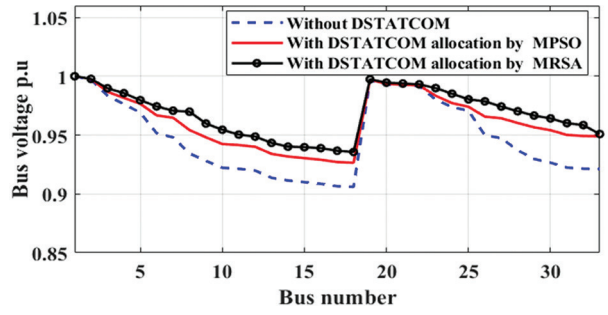


Fig. 8. Bus voltage for DSTATCOM installation in case#4 of CL type

The proposed optimization algorithms (MPSO and MRSA) demonstrate their effectiveness in balancing feeder loadings, leading to a decrease in the TACPL in the system. Additionally, Fig. 9 illustrates the values of the TACPL in the aforementioned cases.

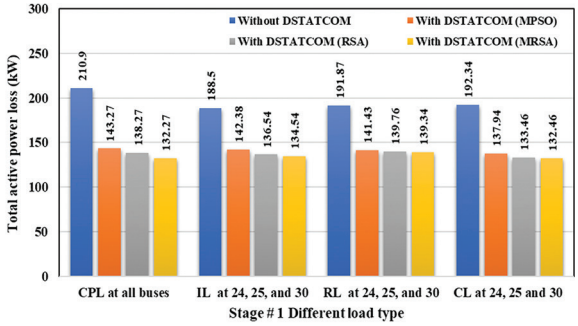


Fig. 9. TACPLs comparison before& after DSTATCOM with different load types

6.2. STAGE TWO: DG TYPE IMPACT ON OPTIMAL LOCATION AND SIZING OF DSTATCOMS

In the previous stage, the CPL type emerged as the worst load type. Consequently, this load type remains constant, while the second variable, DG type, is subjected to variation.

In the subsequent stage, changes are made to the DG type connected to the modified configuration at buses 24, 25, and 30 in the IEEE 33-bus radial DN. All other buses maintained a CPL type at ULF. Furthermore, the analysis includes a comparison of different aspects of DSTATCOM units, such as the optimal sizes, locations, and number with respect to these types of loads. The examination also covers the evaluation of the impact of DG types on TACPLs and voltage profiles. Three specific cases, representing DG types 1, 3, and 4, have been chosen. These cases are further elaborated in the following cases.

6.2.1. Case# 1: DG type 1 and CPL

In this scenario, the IEEE 33-bus radial DN features a CPL across all buses and DG type 1, which exclusively supplies only active power (at p.f=1). The DGs with capacities of 0.25 MW, 1 MW, and 0.5 MW, are installed to the modified configuration at buses 24, 25, and 30, respectively.

The objective is to improve the performance of the IEEE 33-bus radial DN by incorporating DSTATCOM, facilitating a comparative analysis of the three previously mentioned DG types. Consequently, the proposed algorithms (MPSO and MRSA) for optimal DSTATCOM allocation and sizing are tested using the least favorable DG type.

Before installing the DSTATCOM, the TACPL for the IEEE 33-bus radial DN is 135 kW. The bus with the minimum voltage value is 0.917 p.u at bus 18 while no overloaded feeders. The level of DG penetration in this particular scenario is at 45.45%. The goal is to improve system performance using DSTATCOM, considering five candidate buses in the order of 28, 8, 29, 30, and 9.

After allocating and sizing the DSTATCOM using the proposed MPSO algorithm, the TACPL is reduced to 73.08 kW, marking a 45.87% reduction. The optimal configuration involves installing five DSTATCOM units at buses 8, 9, 28, 29, and 30, with respective sizes of 0.33135, 0.45581, 0.4342, 0.3075, and 0.2708 MVAR. Consequently, the DG penetration level increases to 46.17% with the integration of DSTATCOM. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 72.56 kW leading to a substantial decrease of 46.25% in case of installing five DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the lowest voltage level is bus 18, with 0.945 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.955 p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation are illustrated in Fig.10.

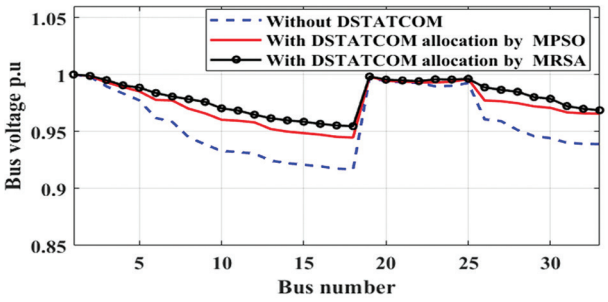


Fig. 10. Bus voltage for DSTATCOM installation in case#1 of DG type 1

6.2.2. Case# 2: DG type 3 and CPL

In this scenario, DG units with capacities of 0.25 MVA, 1 MVA, and 0.5 MVA are installed at buses 24, 25, and 30, respectively. These DG units supply both reactive and active power at p.f equal 0.8, as determined through preliminary sensitivity analysis. The initial TACPL before installing the DSTATCOM units is 110.6 kW, with bus 18 exhibiting the lowest level voltage at 0.919 p.u while no overloaded feeders. The selected level of the DG penetration is 36.55%. Five candidate buses ordered as 8, 28, 29, 9, and 13 undergo a

By applying the proposed MPSO algorithm, the TACPL is reduced to 78.87 kW, marking a 28.69% decrease. The optimal configuration involves four DSTATCOM units installed at buses 8, 13, 28, and 29, with respective sizes of 0.28646, 0.4872, 0.2556, and 0.4863 MVAR. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 78.61 kW leading to a substantial decrease of 28.93% in case of installing four DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the lowest voltage level is bus 18, with 0.949 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.957p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation are illustrated in Fig.11.

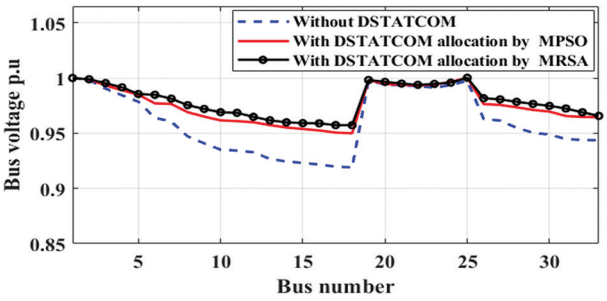


Fig.11. Bus voltage for DSTATCOM installation in case#2 of DG type 3

6.2.3. Case# 3: DG type 4 and CPL

The DG units are installed at buses 24, 25, and 30, supplying active power and consuming reactive power at p.f equal 0.8. Their capacities are 0.25 MVA, 1 MVA, and 0.5 MVA, respectively. The initial TACPL, prior to installing the DSTATCOM units, is 206.4 kW, with bus 18 having the lowest level voltage of 0.909 p.u., while no overloaded feeders. The current level of DG penetration in this scenario is 35.71%. The five candidate buses are ranked in the following order: 6, 28, 29, 8, and 30. By applying the proposed MPSO algorithm, the TACPL is reduced to 98.91 kW, marking a 52.08% decrease. The optimal DSTATCOM configuration involves installing five units at buses 6, 8, 28, 29, and 30, with sizes of 0.38721, 0.51646, 0.33417, 0.48635, and 0.52263 MVAR, respectively. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 97.85 kW leading to a substantial decrease of 52.59% in case of installing five DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the lowest voltage level is bus 18, with 0.937 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.943 p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation are illustrated in Fig.12.

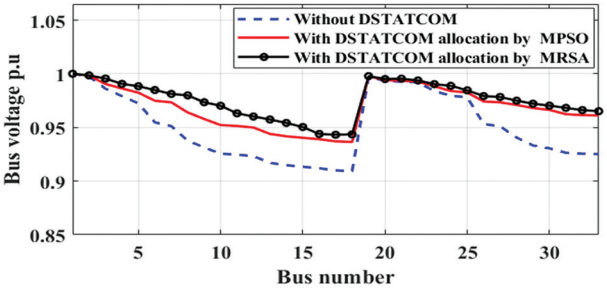


Fig. 12. Bus voltage for DSTATCOM installation in case#3 of DG type 4

The proposed optimization algorithms (MPSO and MRSA) demonstrate their effectiveness in balancing feeder loadings, leading to a decrease in the TACPL in the system for different DG types. Additionally, Figure 14 illustrates the changes in TACPL in the aforementioned cases.

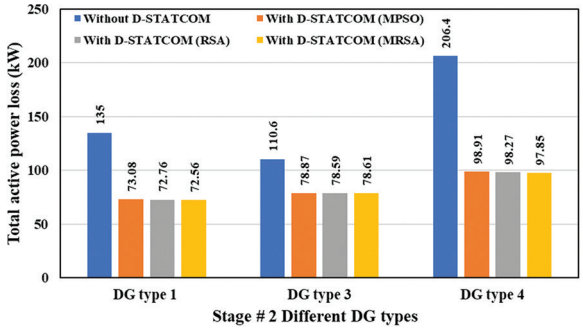


Fig. 13. TACPLs comparison before & after DSTATCOM with different DG types

6.3. STAGE THREE: LOAD FACTOR IMPACT ON OPTIMAL LOCATION AND SIZING OF DSTATCOMS

The most challenging DG and load types identified in the previous stages are the DG type 4 and CPL type, respectively. Consequently, these DG and load types will be kept constant in this stage, while the final variable, the LF, will be systematically altered. This variation is essential to assess the proposed algorithm's efficacy in optimizing the allocation and sizing the multiple DSTATCOM units under different LFs. Additionally, the study will investigate the influence of load factors on TACPL and voltage profiles. This stage serves as a crucial verification of the algorithm's effectiveness in this study, as it is anticipated that a heavily loaded system may negatively impact overall system performance. Two load levels LFs (0.6 and 1.2), are chosen to represent different LFs at distinct times. The outcomes of these two cases are discussed in the following sections.

6.3.1. Case# 1: DG type 4 and CPL at lightly loaded.

In this particular scenario, DG units with capacities of 0.25 MVA, 1 MVA, and 0.5 MVA are installed at buses 24, 25, and 30, respectively. These DG units supply active power while consuming reactive power at p.f equal 0.8. The system is operating with a light load, characterized by a LF of 0.6. Prior to installing the DSTATCOM, the TACPL value is 84.716 kW. Among the buses, bus 18 has the lowest voltage level at 0.95 p.u., while no overloaded feeders. The current level of DG penetration in this scenario is 60.51%. The five candidate buses are ranked in the following order: 13, 10, 14, 12, and 17.

By applying the proposed MPSO algorithm, the TACPL is reduced to 65.48 kW, reflecting a significant 22.71% reduction. The optimal configuration involves installing two DSTATCOM units at buses 10 and 14, with sizes of 0.57558 and 0.13375 MVAR, respectively. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 63.34 kW leading to a substantial decrease of 25.23% in case of installing two DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the lowest voltage level is bus 18, with 0.97 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.98 p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation are illustrated in Fig.14.

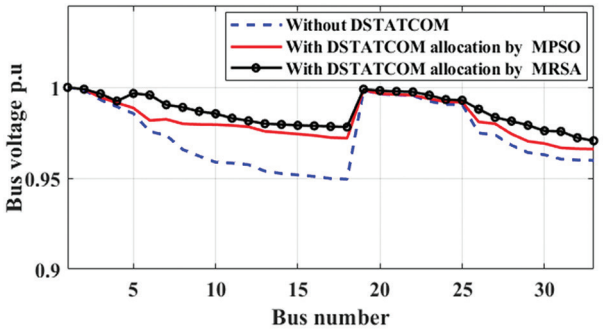


Fig. 14. Bus voltage for DSTATCOM installation in case#1 of LF=0.6

6.3.2. Case# 2: DG type 4 and CPL at heavily loaded

The DG units remain consistent with the previous case, but the system is under heavy load with LF of 1.2. The initial TACPL before installing the DSTATCOMs units is 299.81 kW, and bus 18 registers the minimum level voltage at 0.888 p.u, while no overloaded feeders. The current level of DG penetration in this scenario is 29.43%, The five candidate buses are ranked in the following order: 6, 28, 29, 8, and 30, chosen to enhance system performance.

By applying the proposed MPSO algorithm, the TACPL is reduced to 155.29 kW, reflecting a significant 48.2% reduction. The optimal configuration involves installing five DSTATCOM units at buses 6, 28, 29, 8, and 30, with respective sizes of 0.36844, 0.51691, 0.3574, 0.49, and 0.505 MVAR. Furthermore, the proposed MRSA algorithm has the ability to decrease the TACPL to 141.88kW leading to a substantial decrease of 52.67% in case of installing five DSTATCOM units at the same buses.

The system voltage profile is enhanced as a consequence of the DSTATCOM allocation. The bus with the lowest voltage level is bus 18, with 0.915 p.u. voltage magnitude using the proposed MPSO. While the proposed MRSA algorithm can achieve the minimum bus voltage level of 0.926 p.u. A comprehensive voltage profile, both before and after the DSTATCOM allocation are illustrated in Fig.15.

Figure 16 visually depicts the variations in TACPLs across the discussed cases.

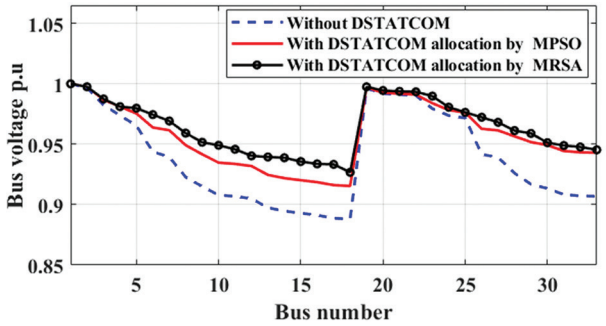


Fig. 15. Bus voltage for DSTATCOM installation in case#1 of LF=1.2

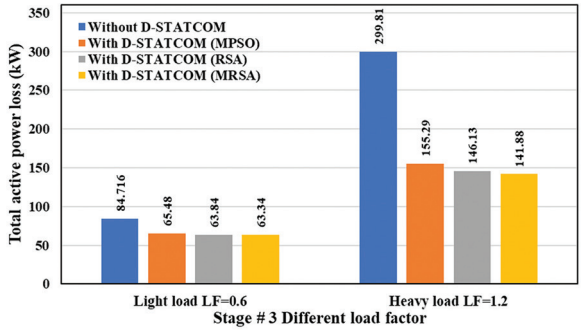


Fig.16. TACPLs comparison before & after DSTATCOM at different LF

7. CONCLUSION

This paper presents the optimal allocation of DSTATCOM units in IEEE-33 bus radial DN with DG, utilizing the LSF and a proposed MPSO and the MRSA methods. The primary objective is to enhance the distribution system's performance by enhancing the voltage profile, reducing TACPL, and maximizing DG penetration to meet consumer load requirements. The simulations were conducted using MATLAB, where the DSTATCOM units model underwent testing and verification. A comparative analysis was performed by assessing the TACPL and voltage profile of the tested radial DN with and without the DSTATCOM.

The optimal determination of the number, size, and location of the multiple DSTATCOM units holds significant importance in the DNs. The study concludes that, among various load types and DG types, CPL type and DG type 4 exhibit unfavorable characteristics in terms of power loss and voltage profile, especially in heavily loaded systems as illustrated below:

- The MRSA obtained the optimum location and size of the DSTATCOM in the DN to minimuze the TACPLs for the different load types. It reduced the TACPLs by 37.31%, 28.31%, 27.38%, and 27.38% for load types CPL, IL, RL, and CL, respectively. While the minimum bus voltage level was improved to 0.929 p.u., 0.932 p.u., 0.937 p.u., and 0.935 p.u. for load types CPL, IL, RL, and CL, respectively.
- For different types of DGs, the TACPLs were reduced by 46.25%, 28.93%, and 52.59% for DG type 1, type 3, and type 4, respectively. While the minimum bus voltage level was improved to 0.955 p.u., 0.957 p.u., and 0.943 p.u. for for DG type 1, type 3, and type 4, respectively.

These findings highlight the effectiveness of the proposed algorithm in addressing challenges posed by the worst load type (CPL) and DG type 4. Furthermore, the algorithm proves its utility in handling different combinations of DG types connected to DNs under varying LFs. The results emphasize the practical benefits of the proposed algorithm for DN operators, offering an efficient approach to allocate DSTATCOM units and optimize system parameters across various operating scenarios in distribution networks. Additionally, the results demonstrated that the MRSA approach was applicable to large DNs, offered fast and easy implementation, was robust in obtaining optimal results, and outperforms the MPSO algorithm.

ACKNOWLEDGMENT

The authors extend their appreciation to the deanship of scientific research at Shaqra university for funding this research work through the project number (SU-ANN- 2023028).

8. REFERENCES

[1] A. Y. Hatata, Eman. O. Hasan, M. A. Alghassab, B. E. Sedhom, "Centralized Control Method for Voltage Coordination Challenges With OLTC and D-STATCOM in Smart Distribution Networks Based IoT Communication Protocol", IEEE Access, Vol. 11, 2023, pp. 11903-11922.

[2] T. Yuvaraj, K. Ravi, K. R. Devabalaji, "DSTATCOM allocation in distribution networks considering load variations using bat algorithm", Ain Shams Engineering Journal, Vol. 8, No. 3, 2017, pp. 391-403.

[3] A. J. Gil Mena, J. A. Martín García, "An efficient approach for the siting and sizing problem of distributed generation", International Journal of Electrical Power & Energy Systems, Vol. 69, 2015, pp. 167-172.

[4] H. H. Fard, A. Jalilian, "Optimal sizing and location of renewable energy-based DG units in distribution systems considering load growth", International Journal of Electrical Power & Energy Systems, Vol. 101, 2018, pp. 356-370.

[5] M. J. Hadidian-Moghaddam, S. Arabi-Nowdeh, M. Bigdeli, D. Azizian, "A multi-objective optimal sizing and siting of distributed generation using ant lion optimization technique", Ain Shams Engineering Journal, Vol. 9, No. 4, 2018, pp. 2101-2109.

[6] S. A. Nowdeh et al. "Fuzzy multi-objective placement of renewable energy sources in distribution system with objective of loss reduction and reliability im-

provement using a novel hybrid method", Applied Soft Computing, Vol. 77, 2019, pp. 761-779.

[7] M. C. V. Suresh, J. B. Edward, "A hybrid algorithm based optimal placement of DG units for loss reduction in the distribution system", Applied Soft Computing, Vol. 91, 2020, p. 106191.

[8] A. Uniyal, S. Sarangi, "Optimal network reconfiguration and DG allocation using adaptive modified whale optimization algorithm considering probabilistic load flow", Electric Power Systems Research, Vol. 192, 2021, p. 106909.

[9] S. Kumar, K. K. Mandal, N. Chakraborty, "Optimal placement of different types of DG units considering various load models using novel multiobjective quasi-oppositional grey wolf optimizer", Soft Computing, Vol. 25, No. 6, 2021, pp. 4845-4864.

[10] J. S. Bhadoriya, A. R. Gupta, "A novel transient search optimization for optimal allocation of multiple distributed generator in the radial electrical distribution network", International Journal of Emerging Electric Power Systems, Vol. 23, No. 1, 2022, pp. 23-45.

[11] S. Mehroliya, A. Arya, A. Verma, S. Tomar, "Optimized Placement of Distributed Generator in Radial Distribution System Using Whale Optimization Technique", SN Computer Science, Vol. 4, No. 5, 2023, p. 626.

[12] G. Memarzadeh, M. Arabzadeh, F. Keynia, "A new optimal allocation of DGs in distribution networks by using coot bird optimization method", Energy Informatics, Vol. 6, No. 1, 2023, p. 30.

[13] J. Sanam, S. Ganguly, A. K. Panda, C. Hemanth, "Optimization of Energy Loss Cost of Distribution Networks with the Optimal Placement and Sizing of DSTATCOM Using Differential Evolution Algorithm", Arabian Journal for Science and Engineering, Vol. 42, No. 7, 2017, pp. 2851-2865.

[14] J. Sanam, "Optimization of planning cost of radial distribution networks at different loads with the optimal placement of distribution STATCOM using differential evolution algorithm", Soft Computing, Vol. 24, No. 17, 2020, pp. 13269-13284.

[15] D. K. Rukmani et al. "A New Approach to Optimal Location and Sizing of DSTATCOM in Radial Distribution Networks Using Bio-Inspired Cuckoo Search Algorithm", Energies, Vol. 13, No. 18, 2020, p. 4615.

[16] S. A. Salimon, Q. O. Lawal, O. W. Adebiyi, M. O. Okelola, "Cost-Benefit of Optimal Allocation of DSTATCOM in Distribution Networks Using Ant-Lion Optimization Algorithm", Periodica Polytechnica Electrical Engineering and Computer Science, Vol. 66, No. 4, 2022, pp. 350-360.

[17] S. A. Salimon, I. G. Adebayo, G. A. Adepoju, O. B. Adewuyi, "Optimal Allocation of Distribution Static Synchronous Compensators in Distribution Networks Considering Various Load Models Using the Black Widow Optimization Algorithm", Sustainability, Vol. 15, No. 21, 2023, p. 15623.

[18] M. A. Tolba, E. H. Houssein, M. H. Ali, F. A. Hashim, "A new robust modified capuchin search algorithm for the optimum amalgamation of DSTATCOM in power distribution networks", Neural Computing and Applications, Vol. 36, 2024, pp. 843-881.

[19] A. Noori, Y. Zhang, N. Nouri, M. Hajivand, "Hybrid Allocation of Capacitor and Distributed Static Compensator in Radial Distribution Networks Using Multi-Objective Improved Golden Ratio Optimization Based on Fuzzy Decision Making", IEEE Access, Vol. 8, 2020, pp. 162180-162195.

[20] A. R. Gupta, "Effect of optimal allocation of multiple DG and D-STATCOM in radial distribution system for minimizing losses and THD", Proceedings of the 7th International Symposium on Embedded Computing and System Design, Durgapur, India, 18-20 December 2017, pp. 1-5.

[21] B. Singh, M. K. Yadav, "GA for enhancement of system performance by DG incorporated with D-STATCOM in distribution power networks", Journal of Electrical Systems and Information Technology, Vol. 5, No. 3, 2018, pp. 388-426.

[22] F. Iqbal, M. T. Khan, A. S. Siddiqui, "Optimal placement of DG and DSTATCOM for loss reduction and voltage profile improvement", Alexandria Engineering Journal, Vol. 57, No. 2, 2018, pp. 755-765.

[23] K. R. Devabalaji, K. Ravi, "Optimal size and siting of multiple DG and DSTATCOM in radial distribution system using Bacterial Foraging Optimization Algorithm", Ain Shams Engineering Journal, Vol. 7, No. 3, 2016, pp. 959-971.

[24] Y. Thangaraj, R. Kuppan, "Multi-objective simultaneous placement of DG and DSTATCOM using novel lightning search algorithm", Journal of Applied Research and Technology, Vol. 15, No. 5, 2017, pp. 477-491.

[25] T. Yuvaraj, K. R. Devabalaji, S. B. Thanikanti, "Simultaneous Allocation of DG and DSTATCOM Using Whale

Optimization Algorithm", Iranian Journal of Science and Technology, Transactions of Electrical Engineering, Vol. 44, No. 2, 2020, pp. 879-896.

[26] P. Samal, D. Panigrahy, "Simultaneous feeder reconfiguration, DSTATCOM allocation, and sizing using seagull optimization algorithm in unbalanced radial distribution systems", Soft Computing, Vol. 28, 2024, pp. 6403-6421.

[27] H. Abdel-Mawgoud, S. Kamel, A. A. A. El-Ela, F. Jurado, "Optimal Allocation of DG and Capacitor in Distribution Networks Using a Novel Hybrid MFO-SCA Method", Electric Power Components and Systems, Vol. 49, No. 3, 2021, pp. 259-275.

[28] N. Belbachir, S. Kamel, F. A. Hashim, J. Yu, H. Zeinoddini-Meymand, S. F. Sabbeh, "Optimizing the hybrid PVDG and DSTATCOM integration in electrical distribution systems based on a modified homonuclear molecules optimization algorithm", IET Renewable Power Generation, Vol. 17, No. 12, 2023, pp. 3075-3096.

[29] M. A. Elseify, S. Kamel, H. Abdel-Mawgoud, E. E. Elattar, "A Novel Approach Based on Honey Badger Algorithm for Optimal Allocation of Multiple DG and Capacitor in Radial Distribution Networks Considering Power Loss Sensitivity", Mathematics, Vol. 10, No. 12, 2022, p. 2081.

[30] M. Zellagui, N. Belbachir, R. A. El-Sehiemy, C. Z. El-Bayeh, "Multi-Objective Optimal Allocation of Hybrid Photovoltaic Distributed Generators and Distribution Static Var Compensators in Radial Distribution Systems Using Various Optimization Algorithms", Journal of Electrical Systems, Vol. 18, No. 1, 2022, pp. 1-22.

[31] M. Frahat, A. Y. Hatata, M. M. Saadawi, S. S. Kaddah, "Grasshopper Optimization-based Optimal Sizing of DG/DSTATCOM in Distribution Networks", Mansoura Engineering Journal, Vol. 47, No. 2, 2022, pp. 6-16.

[32] S. Rana, S. Jasola, R. Kumar, "A review on particle swarm optimization algorithms and their applications to data clustering", Artificial Intelligence Review, Vol. 35, No. 3, 2011, pp. 211-222.

[33] D. Iqbal, T. Ahmad, I. Pervez, I. H. Malick, A. Sarwar, M. Tariq, "Performance of PSO Based Variants in Tracking Optimal Power in a Solar PV based Generation System under Partial Shading Condition", Smart Science, Vol. 8, No. 1, 2020, pp. 1-13.

[34] M. Talaat, B. E. Sedhom, A. Y. Hatata, "A new approach for integrating wave energy to the grid by an efficient control system for maximum power based on different optimization techniques", International Journal of Electrical Power & Energy Systems, Vol. 128, 2021, p. 106800.

[35] L. Abualigah, M. A. Elaziz, P. Sumari, Z. W. Geem, A. H. Gandomi, "Reptile Search Algorithm (RSA): A nature-inspired meta-heuristic optimizer", Expert Systems with Applications, Vol. 191, 2022, p. 116158.

[36] M. K. Khan, M. H. Zafar, S. Rashid, M. Mansoor, S. K. R. Moosavi, F. Sanfilippo, "Improved Reptile Search Optimization Algorithm: Application on Regression and Classification Problems", Applied Sciences, Vol. 13, No. 2, 2023, p. 945.

[37] Q. Yuan, Y. Zhang, X. Dai, S. Zhang, "A Modified Reptile Search Algorithm for Numerical Optimization Problems", Computational Intelligence and Neuroscience, Vol. 2022, 2022, pp. 1-20.

APPENDIX

Table A1 Line and bus data in IEEE-33 bus distribution network

Bus		$R(\Omega)$	$X(\Omega)$	$P_{Recieve}$ (kW)	$Q_{Recieve}$ (kVAR)
Send	Receive				
1	2	0.0922	0.0470	100	60
2	3	0.4930	0.2511	90	40
3	4	0.3660	0.1864	120	80
4	5	0.3811	0.1941	60	30
5	6	0.8190	0.7070	60	20
6	7	0.1872	0.6188	200	100
7	8	0.7114	0.2351	200	100
8	9	1.0300	0.7400	60	20
9	10	1.0440	0.7400	60	20
10	11	0.1966	0.0650	45	30
11	12	0.3744	0.1238	60	35
12	13	1.4680	1.1550	60	35
13	14	0.5416	0.7129	120	80
14	15	0.5910	0.5260	60	10
15	16	0.7463	0.5450	60	20
16	17	1.2890	1.7210	60	20
17	18	0.7320	0.5740	90	40
2	19	0.1640	0.1565	90	40
19	20	1.5042	1.3554	90	40
20	21	0.4095	0.4784	90	40
21	22	0.7089	0.9373	90	40
3	23	0.4512	0.3083	90	50
23	24	0.8980	0.7091	420	200
24	25	0.8960	0.7011	420	200
6	26	0.2030	0.1034	60	25
26	27	0.2842	0.1447	60	25
27	28	1.0590	0.9337	60	20
28	29	0.8042	0.7006	120	70
29	30	0.5075	0.2585	200	600
30	31	0.9744	0.9630	150	70
31	32	0.3105	0.3619	210	100
32	33	0.3410	0.5302	60	40

Performance of TVWS-based LoRa Transmissions using Multi-Armed Bandit

Original Scientific Paper

Anjali R. Askhedkar
Department of Electrical and Electronics Engineering
Dr. Vishwanath Karad MIT World Peace University, Pune 411038, India
anjali.askhedkar@mitwpu.edu.in

Bharat S. Chaudhari*
Department of Electrical and Electronics Engineering
Dr. Vishwanath Karad MIT World Peace University, Pune 411038, India
bsc@ieee.org

Rashid A. Saeed
Department of Computer Engineering
College of Computers and Information Technology,
Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
abdulhaleem@tu.edu.sa

*Corresponding author

Hesham Alhumyani
Department of Computer Engineering
College of Computers and Information Technology,
Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
h.alhumyani@tu.edu.sa

Abdullah Alenizi
Department of Information Technology
College of Computer and Information Sciences,
Majmaah University, Al-Majmaah 11952, Saudi Arabia
aalenizi@mu.edu.sa

Abstract *– Low Power Wide Area Networks (LPWANs) support long-range communication that suits them for varied IoT applications such as asset tracking, environmental monitoring, agricultural management, work site monitoring, fleet management, and smart cities. Even with the large number of connected devices, LPWANs are more power efficient than traditional IoT networks. Long Range (LoRa) is a noteworthy LPWAN technology that employs the ISM band, which experiences congestion due to the growing IoT and smart network applications. LoRa networks can utilize the available TV white spaces to overcome the congestion problem. Performance enhancement of the LoRa network in terms of energy efficiency is a significant concern. This paper proposes, for the first time, Multi-Armed Bandit (MAB) to improve the performance of the LoRa network using TVWS. We have developed a novel algorithm, LoRaT-MAB, for TVWS-based LoRa, and results show that the mean rewards increase by about 12.5% over conventional LoRa and the energy consumption for TVWS-based LoRa using LoRaT-MAB decreases by about 11.7% over conventional LoRa. The findings are encouraging and provide a basis for further research on the TVWS-based LoRa and other IoT applications.*

Keywords: *Internet of Things, LPWANs, LoRa, TV White Spaces, Multi-Armed Bandit, UCB*

Received: May 13, 2024; Received in revised form: July 25, 2024; Accepted: July 27, 2024

1. INTRODUCTION

The Internet is still expanding due to advancements in wireless and mobile technology. The rapid and easy adoption of the Internet of Things (IoT), thanks to the emerging new technologies, has changed the way we live and work. Low Power Wide Area Networks (LPWAN) are a leading network paradigm for IoT as they support range connectivity for tiny edge nodes with lower power and cost than conventional wireless networks. LPWANs provide scalability and can accommodate the varied IoT requirements, enabling applications including smart metering, smart homes, and smart cities, tracking and monitoring applications such as wildlife, agriculture, industrial assets, infrastructure, and so on. Long Range (LoRa), Weightless-N, Weightless-P, SigFox, SNOW, IQRF, DASH7, RPMA, NB-IoT, and 5G are some of the LPWAN technologies. Semtech's LoRa LPWAN technology enables low power, low throughput, and long-range communication. It employs spread spectrum modulation at the physical layer, which improves link budget and interference resistance. It also takes advantage of Forward Error Correction (FEC). It functions in the 433, 868, or 915 MHz Industrial, Scientific, and

Medical (ISM) band. It employs Chirp Spread Spectrum (CSS) modulation with quasi-orthogonal spreading factors and supports up to 50 kbps data rate [1].

With the expansion of wireless applications and due to the restricted availability, fixed resource allocation techniques are unable to meet the growing demand of frequency spectrum from the expansion. With cognitive radio, underutilized TV bands and TV White Spaces (TVWS) may be made available for these kinds of applications. Without interfering with the primary licensed users' ability to communicate, cognitive radio makes it easier for secondary, unlicensed users to take advantage of the unused licensed bands. The White Space Devices (WSD) identify and use the unoccupied spectrum using spectrum sensing. These days, a lot of applications using 2.4 GHz ISM protocols, such as Wi-Fi, Bluetooth, and others, are prevalent, particularly in indoor and urban environments. Herein, the data rate and Quality of Service (quality of service) are affected by interference and coexistence. Using the 5 GHz range for indoor communication raises the issue of fading brought on by obstructions. For such scenarios, TVWS can be used instead. Compared to the ISM band, the VHF and UHF bands show superior signal propagation and greater obstacle penetration.

Various TVWS applications have been researched in the last few years. Regulations governing the maximum duty cycle in the unlicensed ISM bands have a significant im-pact on network capacity. Performance is affected when node density is high. According to studies in [2], using TVWS for LPWAN may lessen coexistence and interference problems. One potential solution for LPWANs operating in the unlicensed yet restricted spectrum is to employ non-ISM spectrum, including whitespaces [3]. Using whitespaces for LPWAN can significantly minimize ISM band conflict, even if it could require improved time synchronization, listen-before-talk functionality, and channel information transmission. Since available spectrum is limited and unlicensed spectrum is susceptible to interference, TVWS offers an excellent alternative to LPWANs. Studies have shown that current LPWAN technologies face challenges such as coexistence, coverage, lack of spectrum, mobility, scalability, and security. As LPWAN drives the visibly growing IoT domain, addressing these issues is imperative. The unlicensed ISM band is commonly used in most current LPWAN solutions. Access to this frequency range is not controlled at the global level except by duty cycle guidelines. An available TVWS can be used to deal with interference, coexistence, and scalability issues in LPWAN networks.

A LoRa-based LPWAN is constructed using stars topology and consists of several nodes that use the CSS modulation technique and the LoRaWAN MAC layer protocol to communicate with a gateway. The gateways send packets received from end devices to the network server [4]. Transmission channels, spreading factors, transmit power, channel bandwidth, and transmission rate are the essential parameters that can be customized with LoRa modulation. The network's overall performance, coverage, capacity, time-on-air, transmission energy, and range are all impacted by the selection of these parameters [5]. One of the six spreading factors and an available subchannel are used by the end devices to communicate with the gateway. When multiple devices use the same channel and spreading factor at the same time, a collision could happen. The likelihood of a collision increases as the number of end devices in the network increases, which leads to a decline in network performance. In this situation, choosing the best parameters to reduce interference and increase energy efficiency may be done using machine learning techniques, which will ultimately improve network performance [6]. Adaptive transmission and efficient resource use are the two strategies for improving low power IoT energy efficiency [7].

In LoRa networks, resource distribution and parameter selection can be done via centralized or distributed methods. Devices have two options: either they allow the network to control the transmission power and data rate, or they take control themselves. The network server manages the end node's transmission parameters. By adjusting the data rate, it lowers a node's transmit power. In this scenario, the network needs to know the node's transmitted power for roughly the last twenty transmissions. It then adjusts the data rate to estimate the transmit power for the upcoming transmission and sends it to the node. The node then modifies its parameters based on the data it receives from the server. This approach's drawback is that it can only be used in stable radio frequency scenarios in which the end nodes remain stationary [4]. In practice, the end node can be mobile, and hence, for the low-complex network with uniformly distributed nodes, considering a single frequency channel and uniform transmit power, the best parameter selection is still challenging.

Additionally, the adaptive data rate approach has certain drawbacks where it as-signs SF to a node based on the uplink signal-to-noise ratio (lower SF for the nodes close to the gateway and higher SF for nodes away from the gateway). ADR may assign the same SF to all nodes that are closer together, which could result in collisions from using one SF more and not using the other SFs [8]. ADR also tends to use less energy but has significant packet losses [5]. In such situations, distributed learning algorithms could be employed so that the edge nodes can select the best parameters for enhancing the performance.

The aim is to optimize energy efficiency and reduce interference at the edge nodes of LoRa-LPWANs that use the TVWS band for transmissions. These two LoRa network performance metrics are affected by the selection of the spreading factor. The network performance is also influenced by several other factors, including channel frequency, bandwidth, coding rate, and transmission power. Multi-Armed Bandit (MAB) is reinforce-

ment learning algorithms that conform to such a structure. A new strategy employing the MAB algorithms is proposed to achieve energy efficiency in this work.

The significant contributions of this work are as given. First is the use of licensed TVWS for LoRa transmissions in contrast to the typical use of the unlicensed spectrum to avoid possible congestion and improve LoRa network performance. Second is the use of MAB algorithm such as DUCB for TVWS-based LoRa is experimented first time in this work, as per our literature study. The third contribution is that we have also developed a novel algorithm, LoRaT-PLM, based on MAB, for the use of TVWS in LoRa, which demonstrates improved performance and enhanced energy efficiency.

The structure of the paper is as follows. Section 2 discusses the relevant works. A brief description of LoRa technology and TVWS is given in Section 3. In Section 4, a novel policy for TVWS-based LoRa using MAB algorithms is proposed. In Section 5, simulations conducted and results obtained are discussed, and Section 6 presents the conclusions.

2. RELATED WORK

Different approaches for the selection of transmission parameters that boost energy efficiency and enhance performance for IoT and LoRa-based LPWANs are being studied and investigated. There are roughly 6720 possible configurations for a LoRa device based on the different transmit power levels, coding rates (CR), spreading factors (SFs), and bandwidths that can be used. As a result, choosing the optimal course of action to maximize network performance is extremely difficult. The technique developed by [9] examines the link and effectively decides a suitable transmission parameter value. The method performs channel estimation based on the data extraction rate and modifies the spreading factor to adapt to the changing channel. In dense networks, experiments show that the suggested scheme outperforms other spreading factor provisioning strategies in terms of capacity and reliability [10]. Utilizing the K-means clustering algorithm for LoRa SF al-location offers added flexibility, enhancing coverage likelihood and enabling uniform resource distribution [11].

Allocating resources at the end node through decentralized learning is an intriguing strategy [12]. To improve energy efficiency and reliability, the end device can choose various parameters for each packet transmission, including sub-channel, spreading factor, transmission power, and others. This method focuses on applying MAB algorithms. To lessen collisions with other nodes, the first application of learning algorithms on LoRa network devices is suggested. The MAB-based upper confidence bound (UCB) algorithm is used for channel selection in LoRa, and the experimental results show that it is possible to double the device's battery life with less memory and processing requirement and achieve better outcomes as compared to random selection. These algorithms are lightweight and can be used to avoid interference coming from other gateways. MAB-based GNU radio implementation also illustrates how such approaches help improve network connectivity [13]. It suggests that both the UCB1 as well as TS are effective and attain convergence quickly in stationary environments; UCB1 learns more quickly than TS, while TS provides slightly superior average performance. If the end nodes in a network are based on learning algorithms, it is possible to accommodate more nodes. Recent works also analyze TS and UCB1 in conjunction with a time and frequency slotted ALOHA, validating an increase in packet delivery ratio even in non-stationary scenarios [14].

The EXP3 algorithm takes into account inter-spreading factor collision, and adversarial MABs are used in the design of a simulator for allocating the resources in LoRa-based LPWANs [15] and improving the overall performance. The EXP3 algorithm's lengthy convergence time is one of its limitations. Compared to EXP3, the improved version, the EXP3.S algorithm, requires less convergence time and is computationally efficient. It performs well for the non-uniform distribution of devices, but the convergence rate might become worse as the number of parameters increases [16].

Reinforcement learning-based resource management techniques that take into account the channel and energy correlation are also developed that exhibit improved energy efficiency [17]. As dense LoRa network deployments experience more packet collisions, a deep reinforcement learning-based transmission parameter assignment algorithm for LoRaWAN is proposed that demonstrates an enhanced packet delivery ratio [18]. A multi-agent cooperative Q-learning approach for resource allocation in LoRa networks demonstrates an improved packet delivery ratio and reduces energy consumption in a deep reinforcement learning-based PHY layer transmission parameter assignment algorithm for LoRaWAN [19]. A multi-agent Q-learning algorithm for dynamic allocation uplink power and SF in LoRa is designed to decrease the power requirement and improve reliability giving an advantage for signal-to-interference noise ratio (SINR) and data rate [20].

Several studies have suggested employing stochastic and adversarial-based distributed learning like updated UCB (UUCB) and its variations, along with updated EXP3 (UEXP3), to fine-tune the communication parameters of devices according to the surrounding conditions. The simulations yield encouraging results for enhancing low-power IoT networks' dependability and energy efficiency [6]. In recent times, scholars have also investigated the UCB for channel selection and various retransmission strategies based on UCB. The technique is equally efficient and raises the transmission rate in dense networks [21].

LoRa specifies the centralized adaptive data rate (ADR) algorithm. The studies show that various MAB

algorithms perform better in terms of energy consumption and packet loss than the conventional ADR algorithm. From a cognitive radio (CR) perspective, MAB learning algorithms are also being studied for spectrum sensing and MAC perspectives. To improve detection efficiency in varying scenarios, a discounted UCB algorithm is proposed for cooperative spectrum sensing [22]. In terms of energy transmission, the results demonstrate higher throughput when compared to the current ADR method [23]. DUCB policy for frequency band selection in a non-stationary CR is also studied, and according to the application requirements, discount functions and exploration bonuses are taken into account; as a result, the policy offers reduced regret [24]. In one of our previous works, new discount functions and exploration bonuses for DUCB were developed to meet LoRa requirements. In comparison to other existing algorithms, the developed algorithm exhibits superior performance and lower complexity [25].

MAB has been used in literature for various wireless network applications for dynamic spectrum access, and modified algorithms have also been proposed. In [26], the authors investigate a dynamic spectrum access problem as a budget-constrained MAB. A modified UCB-MAB algorithm is proposed for dynamic spectrum access and transmission power selection for data rate maximization, resulting in improved performance. Two Thompson sampling-based methods that detect the channel variations and adjust the channel access policy for dynamic spectrum access are suggested [27]. The methods proposed do not consider any information exchange between the end nodes but display a better success rate. A deep learning-based approach for CR results in improved channel access success probability and reduced interference probability [28]. Implementation of UCB-based Reinforcement Learning (RL) algorithm for opportunistic spectrum access on real radio environment using USRP N210 platforms is demonstrated [29]. The UCB algorithm favors the best solution and converges faster, validating the use of RL for dynamic spectrum access. A new approach to a non-stationary MAB problem that uses the predictive abilities of a Large Language Model (LLM) to guide the decision-making process is introduced [30]. Conventional bandit strategies such as epsilon greedy and UCB struggle in case of dynamic variations. An LLM-informed policy that provides guidance dynamically exhibits improved performance. Wireless networks are emerging as self-evolving networks where the use of Generative AI (GenAI) can be beneficial. LLMs, a subfield of GenAI promise to facilitate autonomous wireless networks. A large model trained over various network data can be adapted to accomplish tasks, thus leading to what can be termed artificial general intelligence-enabled wireless networks. The fast growth of LLM offers vast opportunities for network optimization and management in future networks [31]. A TVWS database with a prediction feature that is suitable for different TV frequencies is suggested. It forecasts TVWS availability using RL depending on the time, day, location, and device [32]. Studies show that there are several tools employing geo-location spectrum databases to estimate and guide the TVWS availability to promote efficient radio frequency utilization and dynamic spectrum access [33].

In summary, there are databases available that provide information about the TVWS at a particular location and time. MAB algorithms are shown to perform better for parameter selection in LoRa networks. TVWS database can be exploited in an LLM-like manner along with MAB for parameter selection with additional channels, giving the advantage of faster learning and enhanced success rate. We used a combination of an informed strategy along with the developed UCB-P-1/2+O MAB algorithm [25] for parameter selection in TVWS-based LoRa.

3. LORA, TV WHITE SPACES AND MAB

3.1. LORA TECHNOLOGY

LoRa employs the low-power CSS modulation technique [9] and LoRaWAN medium access control (MAC) [4]. LoRa can operate over different frequency ranges [34]. Although it typically utilizes unlicensed ISM bands like 433 MHz, 868 MHz, and 915 MHz, it operates in licensed bands as well [35]. The packets transmitted by an end node can be received by several gateways in the neighborhood, as shown in Fig. 1.

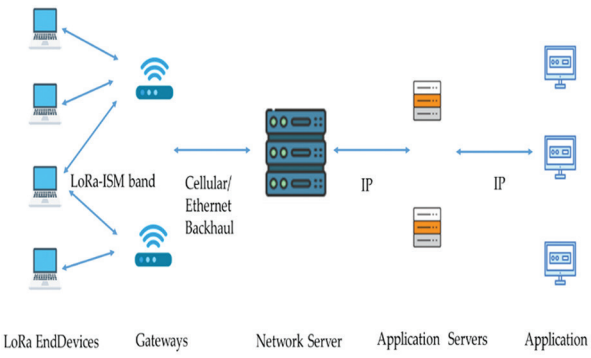


Fig. 1. LoRa Network Architecture

Every gateway uses a backhaul (either satellite, cellular, Ethernet, or Wi-Fi) to forward the packets to the network server for handling sophisticated and intricate tasks like adaptive data rates, sending acknowledgments via the appropriate gateway, and eliminating unnecessary packets. The LoRa network's asynchronous nodes only communicate utilizing pure ALOHA. The number of channels, data rate, and frequency of transmission by the nodes all affect capacity. Variations in the data rate and orthogonal signals are caused by various spreading factors. As a result, the gateway can receive several packets with different data rates on the same channel concurrently [4, 36]. There are also other options being investigated to improve coverage with minimal interference [37].

Different parameters, including SF, CR, transmit power, frequency, and channel bandwidth, can be set for any LoRa device [10]. According to the rules, transmission power can be adjusted roughly in steps of 1 dB, ranging from 2 dBm to 17 dBm. A higher spreading factor corresponds to increased SNR, sensitivity, range, and packet airtime. Spreading factor is defined as the ratio of symbol rate to chip rate. Any value between 7 and 12 can be chosen for SF [36]. A bandwidth of 125 kHz, 250 kHz, or 500 kHz is typically used by LoRa networks. A larger bandwidth increases data rate but at the expense of decreased sensitivity. Forward error correction (FEC) is used by the LoRa modem, and its coding rate can be adjusted to 4/5, 4/6, 4/7, or 4/8. Although it lengthens time on air, a higher coding rate improves error prevention. For transmission, a LoRa packet consumes an average amount of energy, as given by

$$E_{avg} = P_t T_{pkt} N_p \quad (1)$$

where P_t is transmit power, T_{pkt} is transmission time, and N_p is per packet transmissions required for successful transmission. T_{pkt} depends on SF [36]; thus, selecting optimal SF can facilitate optimal energy consumption. If the SINR at the desired LoRa node is higher than the receiver sensitivity for a given SF, a signal is detected at the gateway. The signal power must be high, and the interference power must be low for a high SINR value. It is implied by the above equations that when SF rises, sensitivity improves, and a lower SINR is needed. Time on Air (ToA) or packet time, average energy, and throughput are likewise low for lower SF. The likelihood of a successful transmission decreases as the size of the network grows because there are more devices with the same SF. Achieving energy efficiency and interference avoidance are always trade-offs.

3.2. TV WHITE SPACES

TVWS includes free bands at different times and some TV broadcast frequency bands that are unutilized as a result of TV services being digitalized. TVWS's lower frequency ranges (50–698 MHz) are better at passing through obstructions and are, therefore, less susceptible to fading as well as multipath, allowing for indoor and other applications [35]. Because TVWS offers much bandwidth, it can also support several high-bit-rate applications. Sufficient TVWS may be generally accessible, in contrast to ISM bands, particularly in rural areas, because of the small number of broadcast stations [38, 39]. TVWS presents a promising option for numerous critical indoor and outdoor wireless applications due to its superior indoor penetration, higher spectral efficiency, and good propagation characteristics. Applications requiring a broad transmission range could use TVWS. The Indian government has granted experimental licenses in the 470–590 MHz band, unlocking the possibility of using the TVWS spectrum. Numerous TVWS applications for home networks, smart metering, WLAN, and rural wireless broadband access are demonstrated by literature reviews. For LPWAN, several hardware and software platforms operate over different frequency ranges [35]. Presently, LoRa transceivers for 137 MHz to 1020 MHz frequency bands (*Semtech SX126** and *SX127** series) are available [36]. We have investigated the use of TVWS frequencies, primarily the licensed bands for LoRa, using MAB for optimal parameter selection.

3.3. MULTI-ARMED BANDIT ALGORITHM

Multi-armed bandit is a reinforcement learning structure where an agent must choose arms or actions to maximize its cumulative reward. The end device must choose SF or a strategy $s(t) = \{SFs\}$ from the provided set of SFs. The devices are unaware of their location or the state of the channel. Therefore, any SF that is a part of the set, $s \in S$ may be chosen by the device. Each end device selects a strategy $s(t)$ at each packet arrival time t based on a specific distribution over S, yielding a reward of $rs(t) \in \{0, 1\}$. The transmission may be successful or unsuccessful after the device transmits a packet after choosing a specific value for SF. The LoRa gateway notifies the device of its successful packet reception by sending an acknowledgment. The selection of SF that leads to a successful transmission and receipt of acknowledgment can be modeled as the reward, while SF value can be modeled as the arm or action. It is apparent that the end device receives a reward of 1 if it receives an acknowledgment; otherwise, the reward is 0. The end device chooses an optimal value of SF based only on locally available information, i.e., the received acknowledgment, and experiences the fewest collisions. Since the end nodes are dynamic, it is possible to model the SF selection problem as a non-stationary MAB problem.

Discounted UCB for LoRa:

Stochastic MAB algorithms such as TS and UCB are applicable for stationary distribution scenarios, whereas the advanced DUCB algorithm is suitable for a non-stationary problem. By using an appropriate discount factor, the UCB algorithm can be modified to suit a non-stationary problem. This is the idea behind the Discounted UCB algorithm. The discount factor gives more weightage to the most recent plays and averages past rewards in the DUCB policy. This approach fits the time-varying wireless environment. Therefore, the DUCB policy can also be optimized by modifying the discount factor and exploration bonus to adapt to the varying and complex LoRa network environment. The DUCB algorithm core index $U_k(t)$ is given as

$$U_k(t) = X_k(t) + B_k(t) \quad (2)$$

$X_k(t)$ is the discounted average for exploitation, $B_k(t)$ is the exploration bonus [24]. If the discount function is a power function that is defined as $f(x)=\gamma^x$, then the term $X_k(t)$ can be written as

$$X_k(t) = \frac{\sum_{s=1}^{t} \gamma^{t-s} X_t^k 1_{i_s=i}}{\sum_{s=1}^{t} \gamma^{t-s} 1_{i_s=i}}. \quad (3)$$

Here, $X_k(t)$ gives the average reward of action k at time step t, s is the sample, B is the upper bound, γ^{t-s} denotes

the discount function, 1_{i_s} is the indicator function with value 1 if true and 0 if false. The $B_k(t)$ can be written as

$$\boldsymbol{B_k(t) = 2B\sqrt{\frac{\xi\,(\log\sum_{i=1}^{k} Ni(t))}{N_k(t)}}} \quad (4)$$

where N is the maximum number of trials, i is the index of actions, X_k is the average reward for action k, N_k is number of times action k is chosen, ξ is the bias parameter. $N_k(t)$ is given as

$$\boldsymbol{N_k(t) = \sum_{s=1}^{t} \gamma^{t-s} 1_{i_s=i}} \quad (5)$$

Multi-armed bandit algorithms such as DUCB can be modified to suit the LoRa networks.

4. MAB ALGORITHM FOR TVWS-BASED LORA

The use of MAB algorithms such as modified Discounted Upper Confidence Bound for LoRa using TVWS bands is discussed below subsections.

4.1. UCB-P-1/2+O ALGORITHM FOR LORA

In our previous work, an exhaustive study related to DUCB for LoRa is carried out and a modified DUCB policy for LoRa, UCB-P-1/2+O is proposed [26]. The core index of this policy is as given.

$$\boldsymbol{U_k(t) = X_k(t) + 0.5\sqrt{\frac{(X_k(t)-X_k(t)^2)}{N_k(t)}}} \quad (6)$$

where $X_k(t)$ is the discounted average with the discount function as $[(N\text{-}x)/N]^{1/2}$. Fig. 2 illustrates the flowchart of the developed UCB-P-1/2+O policy, which is utilized by an intelligent node to choose the SF.

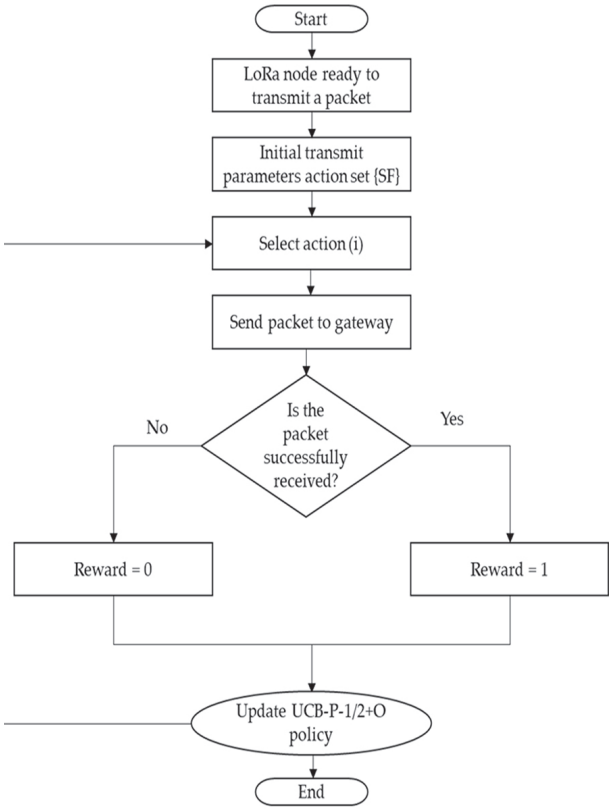


Fig. 2. Flowchart of UCB-P-1/2+O Policy

The node first chooses any SF value from $\{SF\}$ to transmit a packet. The reward is equal to one if a packet is successfully received at the gateway, which sends an acknowledgment. If the packet is not received successfully, the reward is equal to zero and there is no acknowledgment. The UCB-P-1/2+O policy is therefore updated, and the subsequent selection is carried out as per the revised policy.

4.2. LORA OVER TVWS – PATH LOSS MODEL (LORAT-PLM)

The received signal strength varies as a result of hills, trees, buildings, and other similar objects standing between the transmitter and the receiver; this phenomenon is known as shadowing. This effect is seen in wireless networks and TV broadcasts, and it must be taken into account when determining the appropriate transmission power by adding a margin. A hybrid path loss model for LoRa transmissions over TVWS frequencies is developed after a thorough literature review and analysis of the results of various existing path loss models for TV and LoRa transmissions [39]. Based on the Okumura Hata model, the proposed LoRa over TVWS – Path Loss Model (LoRaT-PLM) takes shadow fading into account as given. For urban settings

$$L_u = 69.5 + 26.16\,log_{10} f - 13.82\,log_{10} h_B - C_H + [44.9 - 6.55\,log_{10} h_B]\,log_{10}\,d + C_s \quad (7)$$

where L_u is the path loss in urban regions (dB), f is the transmission frequency (MHz), h_B is the base station antenna height (m), h_M is the mobile station antenna height (m), C_H is a factor for antenna height correction, d is the distance between the transmitter and receiver in kilometers. For rural settings

$$L_0 = L_u - 4.78(log_{10} f)^2 + 18.33(log_{10} f) - 40.94 + Cs \quad (8)$$

where L_o is the path loss in open regions (dB), L_u is the average path loss from a small city form of the model (dB), and f is the transmission frequency in MHz. In (7) and (8), shadow fading is taken into account by the newly designed C_s, a correction factor with a log-normal distribution. When there is an obstruction in the line of sight, when one turns a corner, passes behind a large building, or enters a building, shadow fading can be experienced. The C_s for urban settings would be higher than those for rural ones because these cases are more frequent in urban settings than in rural ones.

4.3. LORAT- MAB ALGORITHM

We developed a modified MAB algorithm, UCB P-1/2+O algorithm, for parameter selection in a LoRa network [24]. The developed algorithm is analyzed for SF selection and gives better performance in terms of mean rewards and execution time. In terms of energy cost, the algorithm demonstrates enhanced energy efficiency as compared to other algorithms. This algorithm is designed and the analysis is carried out in the

context of LoRa operating in the 867 MHz frequency band in India. The use of 470 - 590 MHz TVWS bands, specifically in the Indian context for LoRa, is further investigated by using the developed LoRaT-PLM model [39]. It is demonstrated that TVWS-based LoRa performs better for path loss, energy consumption, and uplink delivery rate. Both the proposed approaches (MAB and TVWS) utilize different techniques for energy efficiency, be it parameter selection or the use of TVWS. The rationale is to combine the two approaches to get the advantages of both. It has been shown previously that the UCB P-½+O algorithm for LoRa at ISM bands performs better than conventional LoRa. Performance analysis of UCB P-½+O algorithm for LoRa at 470 MHz (TVWS) is carried out and compared with the conventional LoRa at ISM band for mean rewards and energy consumption. It is observed that the UCB P-½+O algorithm for LoRa at 470 MHz (TVWS) gives better rewards and the energy consumption is less as compared to the conventional LoRa at the ISM band. This corroborates the use of the UCB P-½+O algorithm for LoRa at TVWS frequencies.

LoRaT-MAB algorithm uses the core index of the UCB-P-½ +O algorithm. It works on the strategy of explore and exploit. The additional information on TVWS availability can be obtained from an authorized database. This information is used by the LoRaT-MAB algorithm to exploit the available channels, thus reducing the exploration requirements. The developed UCB-P-½+O algorithm, as given in (6), has the core index, which is derived from the standard DUCB policy as given in (2). It consists of the addition of two terms: $X_k(t)$, which decides the exploitation of the action depending on the discounted averages, and $Bk(t)$, which decides the exploration of actions done by the policy. Based on this, a novel algorithm is proposed that exploits the database-assisted information effectively to its benefit. This presents an important contribution as a new method to assist the decision-making process of complex stochastic MAB problems. This type of decision-making can adapt to the changing rewards and their distribution patterns and perform better in a non-stationary scenario. This approach can ensure better selection by the algorithm, which will finally lead to improved performance [31]. The strategy to select the transmit parameters can be defined as explore or exploit according to the MAB concept as Strategy $S = \{explore, exploit\}$. The strategy to explore or exploit as decided by the UCB-P-½+O algorithm is S_{actual}. The strategy to explore or exploit according to the information from the TVWS database is Sinformed. The decision D to explore or exploit is made by the LoRaT-MAB algorithm, depending on the strategies S_{actual} and $S_{informed}$, given as

$Decision\ D(S_{actual}, S_{informed})=$
$\{'exploit'$, if both S_{actual} and $S_{informed}$ are $'exploit'\}$
$Or\ \{'explore'$, if S_{actual} &/ or $S_{informed}$ are $'explore'\}$

The flowchart of the developed LoRaT-MAB algorithm is illustrated in Fig. 3.

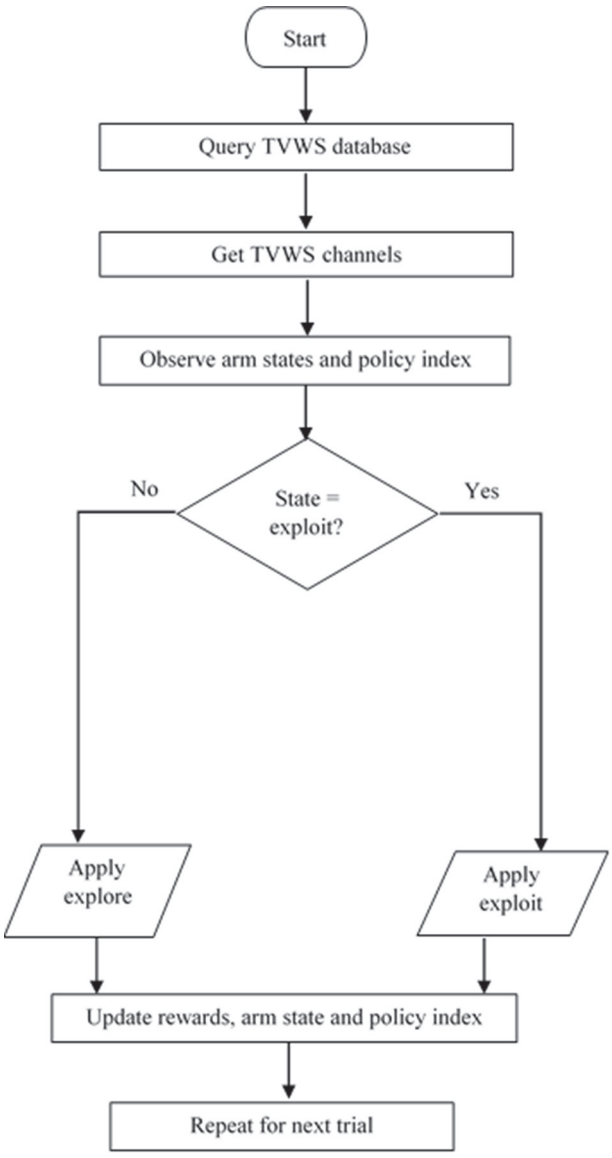


Fig. 3. Flowchart of LoRaT-MAB Algorithm

The important steps of LoRaT-MAB algorithm are as given. At first, a query to the TVWS database is made. The database provides information about TVWS availability. The MAB policy state and the arm states are observed, and then it is verified whether the channel selected by the UCB-P-½+O MAB algorithm is free or busy and whether it should be used at that time instant or not. Depending on the data from the TVWS database, the channel is exploited if free; if the state is exploiting. Otherwise, a new channel is explored. The process is repeated for the subsequent trial.

5. RESULTS AND DISCUSSION

This section presents the performance analysis of the proposed LoRaT-MAB algorithm for LoRa at 470 MHz (TVWS) and a comparison with the UCB-P-½+O algorithm and conventional LoRa using the Random Selection (RS) method. Table 1 lists the various simulation parameters and their settings. All the methods compared consider the selection of SF as the primary

LoRa transmit parameter for selection. Uplink delivery rate (UDR) is the percentage of packets correctly received at the gateway. The simulations also consider the designed LoRaT-PLM path loss model and it shows an improved uplink delivery rate [37, 39]. Results show that LoRaT-MAB gives better rewards, better success rate, and lower energy consumption.

Table 1. Simulation Parameters for TVWS-based LoRa

Parameters	Values
End Devices	5
Area Radius	1 km
Bandwidth	125 kHz
Preamble Symbols	8
Packet Length	11
Header Disabled (H)	1
Data Rate Optimization Disabled (D)	0
Coding Rate	4/5

To investigate the performance of TVWS-based LoRa using the LoRaT-MAB algorithm, LoRa transmissions are simulated using 470 MHz transmission frequency. The results are compared with the performance of TVWS-based LoRa using 470 MHz frequency and UCB-P-½+O algorithm, as well as conventional LoRa using ISM band transmission frequency. Fig. 4 shows the mean rewards per device in a multiple intelligent node scenario as a function of a number of trials.

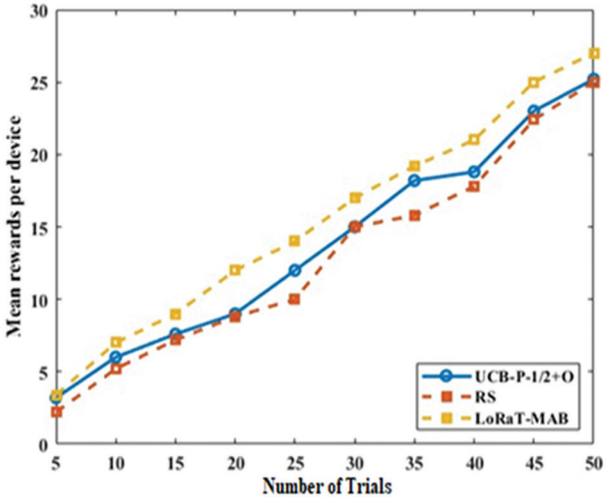


Fig. 4. Mean Rewards vs. Number of Trials for LoRa with LoRaT-MAB and UCB-P-½+O, and Conventional LoRa

The results in Fig. 4 show that TVWS-based LoRa using the LoRaT-MAB algorithm gives better mean rewards per device than TVWS-based LoRa using the UCB-P-½+O algorithm and the conventional LoRa using the ISM band. It is also observed that as the number of trials increases, the mean rewards also increase, as expected. For example, the mean rewards for TVWS-based LoRa using LoRaT-MAB increase by about 12.5% over conventional LoRa and 8.3% over TVWS-based LoRa using UCB-P-½+O for 50 trials. It is also seen that the proposed LoRaT-MAB algorithm gives consistent rewards and outperforms the other two methods. Fig. 5 depicts the changes in average energy consumption for TVWS-based LoRa using the LoRaT-MAB algorithm. It is seen that when the LoRaT-MAB algorithm is applied for TVWS LoRa, it results in lesser energy consumption than TVWS LoRa employing the UCB-P-½+O algorithm and also the conventional LoRa using the ISM band. For example, the energy consumption for TVWS-based LoRa using LoRaT-MAB decreases by about 11.7% over conventional LoRa and 5.8% over TVWS-based LoRa using UCB-P-½+O for 50 trials. It is also seen that the proposed LoRaT-MAB algorithm consistently demonstrates less energy consumption and outperforms the other two methods.

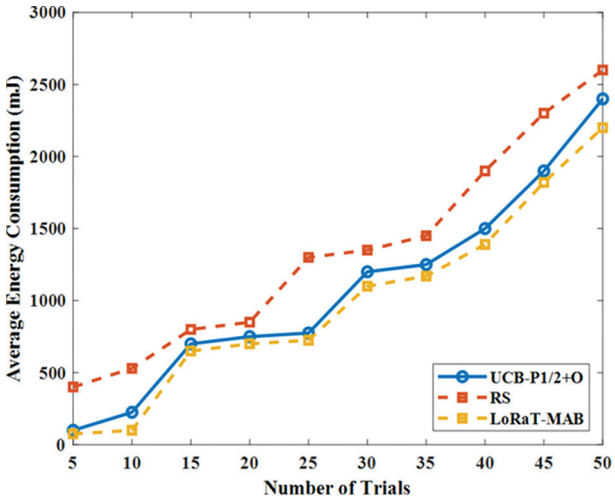


Fig. 5. Average Energy Consumption for LoRa with LoRaT-MAB and UCB-P-½+O, and Conventional LoRa

The consistent and enhanced performance of the proposed LoRaT-MAB algorithm can be attributed to the fact that the algorithm benefits both the MAB and the TVWS database. The database query reduces unnecessary explorations and exploitations. Thus, this additional input from the TVWS database ensures that the decision to exploit or explore is more tolerant and robust to the non-stationary wireless channel environment.

The analysis discussed in the previous sections is carried out for conventional LoRa using ISM band and TVWS-based LoRa using UCB-P-½+O and LoRaT-PLM algorithms separately. Instead of being executed separately, these methodologies can be combined. This would enable the LoRa end device to operate in the ISM band and TVWS band. Initially, the LoRa device would work in the ISM band using the UCB-P-½+O algorithm. If the ISM band is occupied and the selected transmit parameters yield lesser rewards than the required threshold, the transmissions would shift to the TVWS frequencies according to the LoRaT-PLM algorithm. In case the TVWS band operations also yield lesser rewards than the required threshold, the system will reset and repeat the process. The entire procedure is repeated for further trials. A combination of the

methodologies, such as making use of ISM and TVWS band channels and MAB algorithms, may increase the probability of successful transmissions, thus improving energy efficiency and network performance.

6. CONCLUSIONS

To improve the energy efficiency in LoRa along with the network performance, the UCB-P-½+O algorithm has been developed and investigated for ISM band LoRa. It is employed on TVWS-based LoRa transmissions and exhibits improved performance compared to other studied methods. TVWS is beneficial in terms of interference avoidance, reduced path loss, and reduced energy consumption for LoRa networks. These approaches are combined, and a modified MAB algorithm is developed, the LoRaT-MAB algorithm. Simulation results validate the enhanced performance of the LoRaT-MAB algorithm for TVWS-based LoRa transmissions. The methods are compared for the selection of SF as the transmit parameter and can be easily extended for multiple parameter selection. LoRaT-MAB also takes into account the channel frequency selection and displays better performance in terms of rewards obtained and energy consumption. Further, work can be carried out for the selection of multiple parameters simultaneously to increase the network performance. The findings serve as a foundation for future study of MAB-based algorithms, TVWS-based LoRa, and the use of such techniques for other machine-to-machine and 6G applications.

7. ACKNOWLEDGMENT

The authors extend their appreciation to Taif University, Saudi Arabia, for supporting this work through project number (TU-DSPP-2024-132) and Dr. Vishwanath Karad MIT World Peace University, Pune, India for their support and encouragement.

8. FUNDING

The research was funded by Taif University, Taif, Saudi Arabia (TU-DSPP-2024-132).

9. REFERENCES:

[1] J. Finnegan, "A Comparative Survey of LPWA Networking", Zhejiang University International Doctoral Students Conference, China, 2018.

[2] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, T. Watteyne, "Understanding the Limits of LoRaWAN", IEEE Communications Magazine, Vol. 55, No. 9, 2017, pp. 34-40.

[3] A. Dongare et al. "OpenChirp: A Low-Power Wide-Area Networking Architecture", Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops, Kona, HI, USA, 13-17 March 2017, pp. 569-574.

[4] LoRaWAN for Developers, https://lora-alliance.org/resource_hub (accessed: 2024)

[5] R. Kerkouche, R. Alami, R. Féraud, N. Varsier, P. Maillé, "Node-Based Optimization of LoRa Transmissions with Multi-Armed Bandit Algorithms", Proceedings of the 25th International Conference on Telecommunications, Saint-Malo, France, 26-28 June 2018, pp. 521-526.

[6] A. Azari, C. Cavdar, "Self-organized low-power IoT networks: A distributed learning approach", Proceedings of the IEEE Global Communications Conference, Abu Dhabi, United Arab Emirates, 9-13 December 2018, pp. 1-7.

[7] S. N. Ghorpade, M. Zennaro, B. S. Chaudhari, "IoT-based hybrid optimized fuzzy threshold ELM model for localization of elderly persons", Expert Systems with Applications, Vol. 184, 2021, p. 115500.

[8] M. N. Ochoa, A. Guizar, M. Maman, A. Duda, "Toward a Self-Deployment of LoRa Networks: Link and Topology Adaptation", Proceedings of the International Conference on Wireless and Mobile Computing, Networking and Communications, Barcelona, Spain, 21-23 October 2019, pp. 1-7.

[9] M. Bor, U. Roedig, "LoRa Transmission Parameter Selection", Proceedings of the 13th International Conference on Distributed Computing in Sensor Systems, Ottawa, ON, Canada, 5-7 June 2017, pp. 27-34.

[10] Q. Zhou, J. Xing, L. Hou, R. Xu, K. Zheng, "A Novel Rate and Channel Control Scheme Based on Data Extraction Rate for LoRa Networks", Proceedings of the IEEE Wireless Communications and Networking Conference, Marrakesh, Morocco, 15-18 April 2019, pp. 1-6.

[11] M. A. Ullah, J. Iqbal, A. Hoeller, R. D. Souza, H. Alves, "K-Means Spreading Factor Allocation for Large-Scale LoRa Networks", Sensors, Vol. 19, p. 4723.

[12] C. Moy, "IoTligent: First World-Wide Implementation of Decentralized Spectrum Learning for IoT Wireless Networks", Proceedings of the URSI Asia-Pacific Radio Science Conference, New Delhi, India, 9-15 March 2019, pp. 1-4.

[13] L. Besson, R. Bonnefoi, C. Moy, "GNU Radio Imple-

mentation of MALIN: Multi-Armed bandits learning for Internet-of-Things Networks", Proceedings of the IEEE Wireless Communications and Networking Conference, Marrakesh, Morocco, 15-18 April 2019.

[14] R. Bonnefoi, L. Besson, C. Moy, E. Kaufmann, J. Palicot, "Multi-Armed Bandit Learning in IoT Networks: Learning Helps Even in Non-stationary Settings", Cognitive Radio Oriented Wireless Networks, Springer International Publishing, Vol. 228, 2018, pp. 173-185.

[15] D. T. Ta, K. Khawam, S. Lahoud, C. Adjih, S. Martin, "LoRa-MAB: A Flexible Simulator for Decentralized Learning Resource Allocation in IoT Networks," Proceedings of the 12th IFIP Wireless and Mobile Networking Conference, Paris, France, 11-13 September 2019, pp. 55-62.

[16] S. Gupta, B. Chaudhari, B. Chakrabarty, "Vulnerable network analysis using war driving and security intelligence," Proceedings of the International Conference on Inventive Computation Technologies, Coimbatore, India, 26-27 August 2016, pp. 1-5.

[17] R. Hamdi, E. Baccour, A. Erbad, M. Qaraqe, M. Hamdi, "LoRa-RL: Deep Reinforcement Learning for Resource Management in Hybrid Energy LoRa Wireless Networks", IEEE Internet Things Journal, Vol. 9, No. 9, 2022, pp. 6458-6476.

[18] I. Ilahi, M. Usama, M. O. Farooq, M. U. Janjua, J. Qadir, "LoRaDRL: Deep Reinforcement Learning Based Adaptive PHY Layer Transmission Parameters Selection for LoRaWAN", Proceedings of the IEEE 45th Conference on Local Computer Networks, Sydney, NSW, Australia, 16-19 November 2020, pp. 457-460.

[19] A. Scarvaglieri, S. Palazzo, F. Busacca, "A lightweight, fully-distributed AI framework for energy-efficient resource allocation in LoRa networks", Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing, Taormina (Messina), Italy, December 2023, pp. 1-6.

[20] Y. Yu, L. Mroueh, S. Li, M. Terre, "Multi-Agent Q-Learning Algorithm for Dynamic Power and Rate Allocation in LoRa Networks", Proceedings of the IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, London, United Kingdom, 31 August - 3 September 2020, pp. 1-5.

[21] R. Bonnefoi, L. Besson, J. Manco-Vasquez, C. Moy, "Upper-Confidence Bound for Channel Selection in LPWA Networks with Retransmissions", Proceedings of the IEEE Wireless Communications and Networking Conference Workshop, Marrakech, Morocco, 15-18 April 2019, pp. 1-7.

[22] W. Ning, X. Huang, K. Yang, F. Wu, S. Leng, "Reinforcement learning enabled cooperative spectrum sensing in cognitive radio networks", Journal of Communications and Networks, Vol. 22, No. 1, 2020, pp. 12-22.

[23] G. Park, W. Lee, I. Joe, "Network resource optimization with reinforcement learning for low power wide area networks", EURASIP Journal on Wireless Communications and Networking, Vol. 2020, No. 1, 2020, p. 176-184.

[24] Y. Chen, S. Su, J. Wei, "A Policy for Optimizing Sub-Band Selection Sequences in Wideband Spectrum Sensing", Sensors, Vol. 19, No. 19, 2019, p. 4090.

[25] A. R. Askhedkar, B. S. Chaudhari, "Multi-Armed Bandit Algorithm Policy for LoRa Network Performance Enhancement", Journal of Sensor and Actuator Networks, Vol. 12, No. 3, 2023, p. 38.

[26] A. Amrallah, E. M. Mohamed, G. K. Tran, K. Sakaguchi, "Enhanced Dynamic Spectrum Access in UAV Wireless Networks for Post-Disaster Area Surveillance System: A Multi-Player Multi-Armed Bandit Approach", Sensors, Vol. 21, No. 23, 2021, p. 7855.

[27] S. Ye, T. Wang, S. Wang, "Thompson Sampling-Based Dynamic Spectrum Access in Non-Stationary Environments", IEEE Transactions on Cognitive Communications and Networking, Vol. 9, No. 3, 2023, pp. 593-603.

[28] X. Wang, Y. Teraki, M. Umehira, H. Zhou, Y. Ji, "A Usage Aware Dynamic Spectrum Access Scheme for Interweave Cognitive Radio Network by Exploiting Deep Reinforcement Learning", Sensors, Vol. 22, No. 18, 2022, p. 6949.

[29] C. Moy, A. Nafkha, M. Naoues, "Reinforcement learning demonstrator for opportunistic spectrum access on real radio signals", Proceedings of the IEEE International Symposium on Dynamic

Spectrum Access Networks, Stockholm, Sweden, 29 September - 2 October 2015, pp. 283-284.

[30] J. De Curtò, I. De Zarzà, G. Roig, J. C. Cano, P. Manzoni, C. T. Calafate, "LLM-Informed Multi-Armed Bandit Strategies for Non-Stationary Environments", Electronics, Vol. 12, No. 13, 2023, p. 2814.

[31] A. Rao, B. Chaudhari, "Development of LoRaWAN based Traffic Clearance System for Emergency Vehicles," Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), Palladam, India, 7-9 October 2020, pp. 217-221.

[32] A. Pakzad, R. M. Manuel, L. Materum, "TVWS Geolocation Database for Secondary-User TVWS Devices for Spectrum Forecasting", EPSTEM, Vol. 21, 2022, pp. 188-195.

[33] A. Lysko, L. Mfupe, "Television Whitespace enabling rural and utility connectivity with CSIR geolocation spectrum database technology", Presented in CSIR Conference, November 2020.

[34] LoRa SX1276 Datasheet, https://www.semtech.com/products/wireless-rf/lora-connect/sx1276 (accessed: 2024)

On-Tree Mango Fruit Count Using Live Video-Split Image Dataset to Predict Better Yield at Pre-Harvesting Stage

Original Scientific Paper

Devender Nayak Nenavath
Vellore Institute of Technology, Vellore
School of Computer Science and Engineering-SCOPE
Vellore, Tamilnadu, India-632014
devendernayak.n@vit.ac.in

Boominathan Perumal*
Vellore Institute of Technology, Vellore
School of Computer Science and Engineering-SCOPE
Vellore, Tamilnadu, India-632014
boominathan.p@vit.ac.in

*Corresponding author

Abstract *– This study introduces a method for fruit counting in agricultural settings using video capture and the YOLOv7 object detection model. By splitting captured videos into frames and strategically selecting representative frames, the approach aims to accurately estimate fruit counts while minimizing the risk of double counting. YOLOv7, known for its efficiency and accuracy in object detection, is employed to analyze selected frames and detect fruits on trees. Demonstrated the method's effectiveness through its ability to provide farmers with precise yield estimations, optimize resource management, and facilitate early detection of orchard issues such as pest infestations or nutrient deficiencies. This technological integration reduces labor costs and supports sustainable agricultural practices by improving productivity and decision-making capabilities. The scalability of the approach makes it suitable for diverse orchard sizes and types, offering a promising tool for enhancing agricultural efficiency and profitability. The researcher compared YOLOv5n, YOLOv5s, YOLOv7, and YOLOv7-tiny with eight-sided imaging techniques around the tree. The experimental results of YOLOv7 with the eight-sided technique performed best and achieved a count accuracy of 97.7% on a single tree in just 17.112 ms of average inference time. On multiple trees, it is 95.48% in just 17 ms of average inference time, with the help of an eight-sided method on tree images.*

Keywords: *Computer Vision, Deep Learning, Image Processing, Agricultural Technology, Horticulture Fruit*

Received: May 10, 2024; Received in revised form: July 18, 2024; Accepted: August 16, 2024

1. INTRODUCTION

The famous Persian poet "Amir Khusrau named the mango Naghza Tarin Mewa Hindustan," which means the fairest fruit of Hindustan (India). Mango is a member of the Anacardiaceae family and includes several other species, notably cashews, sumac, and pistachio, which are traditionally grown in different climates [1]. Mango fruit is a seasonally available fruit, particularly in summer in India. Mango is rich in polyphenols, predominantly gallic acid, and has antioxidant, anticancer, and anti-inflammatory activities that improve chronic inflammation. Polyphenols and mango fiber may serve as prebiotics to increase probiotic bacteria in the intestines. Anti-inflammation prevents other symptoms, such as colon cancer, chronic intestinal diseases, and leaky intestines, and improves intestinal health [2]. Mango peel and bagasse are rich sources of dietary fiber, which is beneficial for cardiovascular diseases, type 2 diabetes, metabolic syndrome, and cancer. Mango seeds are rich in vegetable oils, proteins, and antioxidants with antibiotic potential [3].

Our focus is only on mangoes because of their shape, structure, and color. Compared to various fruits, man-

goes are different in shape, with more than 1000 varieties, a few varieties shown in (Fig 1). Although we have a model with a specific dataset of mango types for detection and counting, the same model cannot work for another type of mango. To make this work possible for all varieties, different varieties of images are required to create a dataset for training, testing, and validation to make an algorithm with reasonable accuracy, whereas, in other fruit scenarios such as apple, lemon, and sapodilla, this is not necessary for training, testing, and validation for all varieties.

Horticulture is a subdivision of agriculture that involves plants, flowers, turf, fruits, and nuts. Different storage techniques, transportation facilities, and marketing strategies are available for pulses and cereals, but these facilities are for something other than fruits. Different storage mechanisms are available, such as cold storage and traditional storage mechanisms constructed with the help of wheat straw, paddy straw, grass, bamboo, wood, bricks, mud, and cow dung for grains. However, cold storage can change the fruit's taste and natural fragrance within a few days [4].

Fig. 1. Various mango varieties. Each mango is different in shape, structure, color, and size

Detection is a technique related to computer vision, which locates and identifies objects within a given image or video. Various algorithms detect an object, such as R-CNN, Fast R-CNN, Faster R-CNN, and Histogram of Oriented Gradients (HOG), which was used in improved YOLOv2 to detect immature mango fruit [5], R-FCN, SSD, SPP-net, and YOLO. Counting harvestable and non-harvestable fruits, if the fruit count is well known before plucking from the tree or at maturity of the fruit, a farmer can predict their outcome based on the fruit count and can participate in the online sale or sell the fruits at a confident price based on the number of fruits. Using this procedure, the farmer can obtain the expected yield. The central research concept of counting tree fruit will help farmers obtain better yields [6].

Counting is a technique that will give the number of detected objects based on the detection technique. Counting the number of mangoes on trees and branches is the best method for obtaining production data. At the beginning of the harvest cycle, performing the count when measuring all fruits of the productive cycle on the tree will result in the number of fruits per tree [6]. The farmer will come to know the expected outcomes from an orchard. Only a few counting methods are available to count fruits; the primary and traditional method is manual counting, where humans count the fruit by their eye vision on the tree; it is a very high-cost and time-consuming technique [7]. Bounding box counting is the following counting method used in image processing. Where the total fruit count is the observed bounding box count. [8]. The third method uses Vertical and Horizontal line-based counting, which is performed based on a line. If an object passes through the line, counting starts and produces the total passed count [9]. The fourth method of counting uses SORT techniques. SORT is helpful in deep learning concepts, where the fruit is counted based on the SORT technique [10]. Another counting technique is ROS core counting, in which the fruits are counted based on a Robot Operating System [11]. Another counting method is the region-of-interest (ROI) and unique object identification (ID) methods [12].

The contribution of this procedure involves developing an innovative and efficient method to count fruits on a tree using video capture and object detection with YOLOv7. We strategically selected representative frames from the different segments by capturing a 360-degree video around the tree and splitting it into individual frames to avoid double counting. This approach ensures comprehensive coverage of the tree while minimizing redundancy. Applying the YOLOv7 model to these frames enabled accurate and realtime fruit detection. Our method optimizes computational resources and enhances the accuracy of fruit counting, thereby providing a practical solution for agricultural and horticultural applications. This contribution is significant for improving the monitoring and estimation of fruit production and supporting better yield management and resource allocation in the field.

2. RELATED WORK

Counting the fruits and flowers was performed manually, which is very expensive; to overcome this expensive problem, they proposed a simulated deep convolutional neural network for yield estimation. In this study [7], they created a 26400 image dataset in which they used 24000 for training and the remaining 2400 for testing, and the error decreased by using the Adam optimizer. They used a modified version of the Inception-ResNet architecture to capture features at multiple scales. Finally, this network tested on authentic images. It achieved 91% accuracy. The advantage of this work is that it can be applied to other fruits because the dataset preparation involves filling the entire blank image with green and brown colored circles and simulating the background.

Ref. [9] proposed a lightweight YOLOV5-CS (Citrus Sort) object detection model with 3000 original images used to detect and count citrus fruits in the natural environment. First, to improve generalization image rotation, a convolutional layer with a block next to the

backbone and the subsequent detection layer was embedded for accuracy improvement. Both loss function full Intersection over Union and Cosine annealing applied for improved training. The developed model moves to implement an edge artificial intelligence system. For the counting scene segmentation method with the virtual region and the formed embedded system, mAP@.5 is 98.23%, and recall is 97.66% with a frame rate of 28 FPS.

In [13], the multi-scale multilayer perceptrons (MLP) and CNN were used to overcome the previous fruit segmentation performance of a benchmarked MLP network for fruit detection and counting in orchard image data. They incorporated metadata in these architectures to explicitly capture the relationships between meta-parameters and object classes. Watershed Segmentation and Circular Hough Transform algorithms were used to post-process pixel-wise image segmentation and achieved a computing and detection F1-score of 0.858. This model has the advantage of detecting partially circular regions, thereby enabling the merging of disjointed fruit regions into a single detection. In addition, it is not possible to visualize all fruits in the image data owing to occlusions and clustering.

To estimate the yield of citrus fruits under natural lighting conditions [14], a computer vision algorithm using a hybrid watershed transform was proposed to detect and count citrus trees and performed the image on 84 images from 21 trees. These images were noisy because they included some other tree parts. Therefore, some input images were subtracted from the background and resized to 1824:1028 to improve data processing speed. They converted these from RGB to HSV and evaluated the marker-controlled watershed and distance transform algorithms for automated watershed segmentation, obtaining an R^2 of 93%.

In [15], night-acquired images of 1515 trees across five orchards for single-stage architectures such as YOLOv3, YOLOv2, YOLOv2-tiny, SSD, and two-stage architectures such as Faster R-CNN with VGG, Faster R-CNN with ZF were train with an original resolution of 512 × 512 for a total of 11 models. Compared to a previous poor study on fruit, leaf color, shape, and texture, a hybrid model named MangoYOLO was developed. The MangoYOLO of 33 layers model was constructed based on the better features of YOLOv2-tiny's fewer layers and higher speed as advantages, as well as the multiple detection layers and high-speed features of YOLOv3. MangoYOLO achieved 0.97 of the F1-score for fruit detection in an image. They have proposed a new MangoYOLO model based on YOLOv3 and YOLOv2-tiny features and compared all models to obtain better results with the new MangoYOLO-512-pt model.

In [16] an R-CNN model performed training on 1160 unmanned aerial vehicles (UAV)- based data images of two years captured in different directions and at different distances from the ground level to detect and count the number of apple fruits on individual trees. The proposed model's results compare with the agro technician in situ apple counts; the acquired R-square value was 0.86, with a Mean Absolute Error of 10.35 and a Root Mean Square Error of 13.56. In the top-view images, the number of total images acquired R-square value was 0.80, with MAE: 128.56 and RMSE = 130.56. According to [16], using a colab is the main advantage.

In [17] a study based on a single-shot multi-box detector with MobileNet and a faster R-CNN with Inception V2 architectures for detection. Training and testing were performed on three different fruits, avocado, lemon, and apple, with two architectures, under different field conditions. For video-based fruit counting multi-object tracking with the Gaussian estimation algorithm, Faster R-CNN with Inception V2 achieved 93% of the result and 90% using SSD with MobileNet. A disadvantage of this study is that the results could be more conclusive for other fruits.

In recent years, deep learning has been widely applied in agricultural fields [18]. Using YOLO model techniques, detection was applied under various imaging and illumination situations to estimate the load of orange fruit in an orchard. They used 1115 trees for examination, conducted in three steps: creating an orange-tree dataset under different illumination conditions, evaluating the selected model on 100 sample trees, and finally extracting the yield based on detecting and counting the oranges of every image taken. Using this method, they observed some two-sided differences for thin canopy and four-sided differences for dense canopy imaging. With the help of the YOLOv4 model, the precision was 91.23%, the recall was 92.8%, F1-score was 92%, and mAP was 90.8%.

A novel methodology was developed for apple fruit detection and counting using deep learning with apple fruit trunk tracking. They [19] constructed their dataset using images and videos and divided the image data 800 into 80% and 20% ratios for training and testing. In early studies, these algorithms mismatched or lost their targets because of the large number of similar fruits. However, in this study of apple fruit trunk, which is usually more significant than the fruit in appearance, and YOLOv4-tiny with the channel spatial reliability-discriminative correlation filter (CSR-DCF) algorithm. The developed method was tested using the ID-switched number of fruits, MIDE, and RMSE to assess the performance of matching fruit in a video frame and observed an mAP of 99.35% for fruit trunk detection, 91.49% counting accuracy, and R-square of 0.9875. The advantage of the proposed method is that it provides the possibility of realtime yield estimation of the orchard using a CPU at 2–5 fps. They found some drawbacks with this procedure; this study considered only a single-sided row of the tree at the time of counting; if the practitioner uses both sides of the tree, it will provide a double count. Therefore, they suggested that some investigation is required to perform counting of both sides of the fruit.

For early crop load estimation of the apple fruit canopy, the [20] YOLOv4-based model used on 480 raw apple tree images split into three growing stages: early, mid, and harvest. Their previous research modified the YOLOv4 network architecture and fine-tuned it to adapt it for apple fruit detection. In this study, they designed the CA-YOLOv4 model with three significant improvements, which specifically addressed some major challenges: small fruit size, dense canopy conditions, and severe canopy occlusion. The first improvement is the convolutional block attention module (CBAM) mechanism, which learns to improve the detection accuracy based on target features and surpasses nontarget features. In the second stage, we added an adaptive layer and a large-scale map d together. The regression box loss function was optimized, and the last phase included the densely connected network structure. The results showed that CA-YOLOv4 had a lower final loss value, more excellent recall, f1-score, and precision. CA-YOLOv4 performed better than the Faster R-CNN and SSD. Finally, this proposed CA-YOLOv4 study performed a superior detector for fruit counting and can perform near realtime with an average detection time of 0.1 s per image with the described hardware.

Fruit counting is essential in orchard management and plantation science. Ref. [21] early studies showed a need for robust and accurate fruit-counting methods in complex orchards, such as covering, shadows, clustering images, and complete fruit counting on whole trees. This study proposed and validated a panoramic method based on deep learning object detection for complete yield estimation for holy fruit. This method used a holly fruit dataset of 640×640 samples divided into 75% and 25% ratios for training and testing purposes. To form a complete panoramic unfolding map of the fruit tree surface, the images surrounding the fruit trees were captured using a UAV, and SIFT-based image matching was performed. Tested the accuracy and effectiveness of this method at different scales and scenarios and observed that high-quality built panoramic images for an accurate fruit count. The statistical rate between the detected and actual number is more than 96% when the ring shot parameter of the holly tree is less than or equal to 1.2 m; when the shot ring parameter is less than or equal to 1.6 m, then the statistical rate is 95%. The detection rate between the detected and captured numbers in the panorama image is over 99% when R≤1.2 m and over 97% when R≤2.0 m. Even though the model has a high detection rate, the current confidence threshold is still missing. These missed fruits are difficult to identify because of incomplete fruit contours, fake pixel values, insufficient pixels, and mutual interference between highly similar targets.

An automatic apple counting system for modern orchards was developed by [22], where they acquired ten sets of original videos and 1600 images with 720×1280 pixels of two consequent harvesting seasons. These 1600 images were divided randomly into 1280 images and 320 images for training and testing, in which they were labeled manually to the fruits and trunks. In the third year, labele 93050 samples with two classes: fruit and trunk. They then performed a regular detection-matched fruit counting system (NDMFCS) test on ten sets of original videos. In NDMFCS, based on YOLOv4-tiny performed object detection, abnormal fruit detection was abatement based on a threshold, and fruit counting was performed based on trunk tracking and identity document (ID) assignment. Finally, the results indicated that the average fruit detection precision was improved from 89.1% to 93.3% based on ten sets of original videos. Implementing CPU at 3-5 FPS is the advantage of this model. However, it was developed based on Intel RealSense D435 camera-based videos, which are challenging to use widely because they require computing equipment.

Ref. [23] picking the litchi fruit failed but was located successfully due to random obstruction in early studies. In this study, with the help of 1000 training sets and 100 test sets with the YOLOv8-seg model on litchi and its branches classified images with binocular vision technology picked points, they proposed a picking point framework for the robot system. This procedure achieved 88.1% precision for segmenting litchi fruit branches with an 88% picking-point success rate, and the overall success rate was 81.3%, with an average error of 2.8511 mm. The advantage of this study is that it is quick and accurate in identifying target points, which is a realtime operation.

In this [24] article, they investigated fruit detection methods, including traditional and deep learning methods. However, they focused on deep learning and optimization strategies for fruit detection in two ways: optimization strategies for fruit detection on pre-image sampling and optimization strategies after image collection to overcome the unstructured background challenge in the orchard field environment. They studied complex background factors and adverse effects, lighting conditions, occluded fruits, fruits with different degrees of maturity, and complex backgrounds in outdoor orchard environments and suggested future work.

We created a dataset of 1021 pictures as every 20 frames of video data from the UAV obtained 304 images. The remaining 717 images were captured using a mobile phone, scaled to 640 × 640, and labeled using the LabelImg tool, divided into a ratio of 80:10:10. Existing works fail to balance speed and accuracy and perform well when features are distinct and occlusion minimized. Therefore, they [25] introduced a novel lightweight network architecture based on the YOLOv5 foundation. First, they pruned YOLOv5 using filter pruning and then introduced an adaptive BN layer to identify the best-pruned subnet based on the score. Finally, an ECA module is appended to the optimal network to form and fine-tune a new one. They observed that the proposed YOLOv5_E has 24.2% parameters as 26.2% of YOLOv5 size; it runs at 178 FPS, with only a 0.9% loss in accuracy. This model pruning is advantageous for efficiency, incurring a subtle decrease in accuracy.

Ref. [26] the citrus detection and dynamic counting method was proposed based on the lightweight target detection network YOLOv7-tiny, Kalman filter tracking, and the Hungarian algorithm. This work uses the YOLOv7-tiny algorithm for predictive tracking of discovered fruits utilizing the Kalman filter to recognize citrus fruit in a video. Added the Euclidean distance, overlap matching to the Hungarian technique, and a two-stage life filter. Finally, drawing line counting was proposed. The average detection accuracy of YOLOv7-tiny was 97.23%, dynamic detection was 95.12%, multi-target tracking accuracy was 67.14%, and the improved dynamic counting algorithm was 67.14%.

This study [27] introduces a novel blueberry ripeness and count detection methodology that integrates an attention mechanism with a bidirectional feature pyramid network (BiFPN) within the YOLOv5 framework. In 192 images of the 2515 blueberries, 1612 immature and 903 mature blueberries were divided 192 images into 80% for training, 10% for testing, and 10% for validation in this study. Their proposed YOLOv5-CA model achieved an mAP at an IoU threshold of 0.5, recall of 88.2%, precision of 88.8%, and culminating mAP of 91.1%. As the model was YOLOv5-SE+BiFPN, the mAP was 90.5%, the recall was 88.5%, and the precision was 88.4%.

Previous studies treated the detection of clusters and berries as separate tasks to count the grape, owing to the clustered nature of the grape. In this study, they [28] proposed a probability map-based grape detection and counting framework, where first detects two intermediate maps through a neural network and uses three stages to finish the three grape detection and counting subtasks; for this study, they used the WGISD dataset; Chengdu dataset, and BpGC dataset three different types of datasets. They used the WGISD dataset and combined 100 more datasets from the Chengdu dataset and tested the proposed framework, achieving a localization performance of AP of 0.851, counting performance of MAE of 1.845, RMSE of 2.142 for grape clusters, 23.414 for MAE and 31.391 for RMSE of counting performance for grape berries, and MRD of 0.142, 1-FVU 0.865 of counting performance for berries per grape cluster. However, this study also has certain limitations. First, our method detects only visible grape clusters and grape berries, while occluded, invisible grape berries still need to be discussed. Second, our research only focuses on grape detection and counting, upstream tasks in digital viticulture. However, we do not apply the grape detection results to downstream tasks such as predicting grape picking points and actual grape productions, which are more relevant to practical production activities.

Developing an efficient control method for each basic module and constructing its internal conditions is vital to transitioning a harvesting robot from a functional prototype to a practical machine. Therefore, this study [29] tackles efficient locomotion, picking, and seamless integration. They built a system and proposed a set of algorithms for locomotion-destination estimation, realtime self-positioning, and dynamic harvesting. They have established a solid coordination mechanism for continuous locomotion and picking behavior. As a result, the success rate of positioning was 95.8%; at 17 destinations of dragon fruit in an orchard, the robot carried out 24 positioning operations and obtained 14 successful movements, where the time consumption was 7.71 s. whereas the fig orchard had a 76.9% picking success rate, where at 11 destinations, seven successful movements, one collided with the branch, and three lost visual tracking. Each method offers distinct advantages such as improved accuracy, adaptability to varying conditions, and enhanced picking efficiency to operate a robot autonomously and continuously. This method is limited because the module works better in daytime conditions than in night vision.

3. MATERIALS AND METHODS

Counting fruit on trees where it is hidden under a leaf or branch and occluded with another fruit avoids double counting [30]. To solve this research problem, we used multiclass classification with YOLOv5 and YOLOv7 versions and compared each. We will obtain better detection and counting accuracy with a clear solution using a better YOLO version.

3.1. ACQUISITION OF DATA

The images in this study are of Vikarabad District, Telangana State, India. The variety of mangoes is 'Banganapalle,' also known as 'Benishan' [31]. The video captured the mango fruit's maturity stage between 06:00 AM and 08:00 AM, which is a perfect time for capturing images under natural lighting conditions. They used an iQOO Z3 5G mobile phone to capture the video, with a duration of 45 s to 75 s around the tree in a clockwise direction, which covered 360 °of the tree. The data storage was from 70MB to 160 MB with 1080 pixels ×1920 pixels of resolution in a portrait capturing way and saved as mp4 video. The approximate speed of the captured video was 30 fps. The video was acquired under natural daylight, while the outdoor environment was warm. (Fig 2) shows a few images of the data frames. Split the mp4 video into frames, and a video of 152 MB of storage data splits into 2228 image frames of on-tree mango fruit at approximately 30 fps. After splitting the 152 MB video into 2228 frames, the size of these 2228 frames was 2.03 GB of storage.

3.2. ANNOTATION

To create the images as a dataset, we must annotate every image into a .jpg file and a .txt file. Among these 360 image frames, 288 were for training, and 72 were for testing and validation purposes, with an 80:20 percent ratio. An open-source tool, LabelImg [32], was used to annotate the images in this research. It is straightforward to use and has better options in labeling formats.

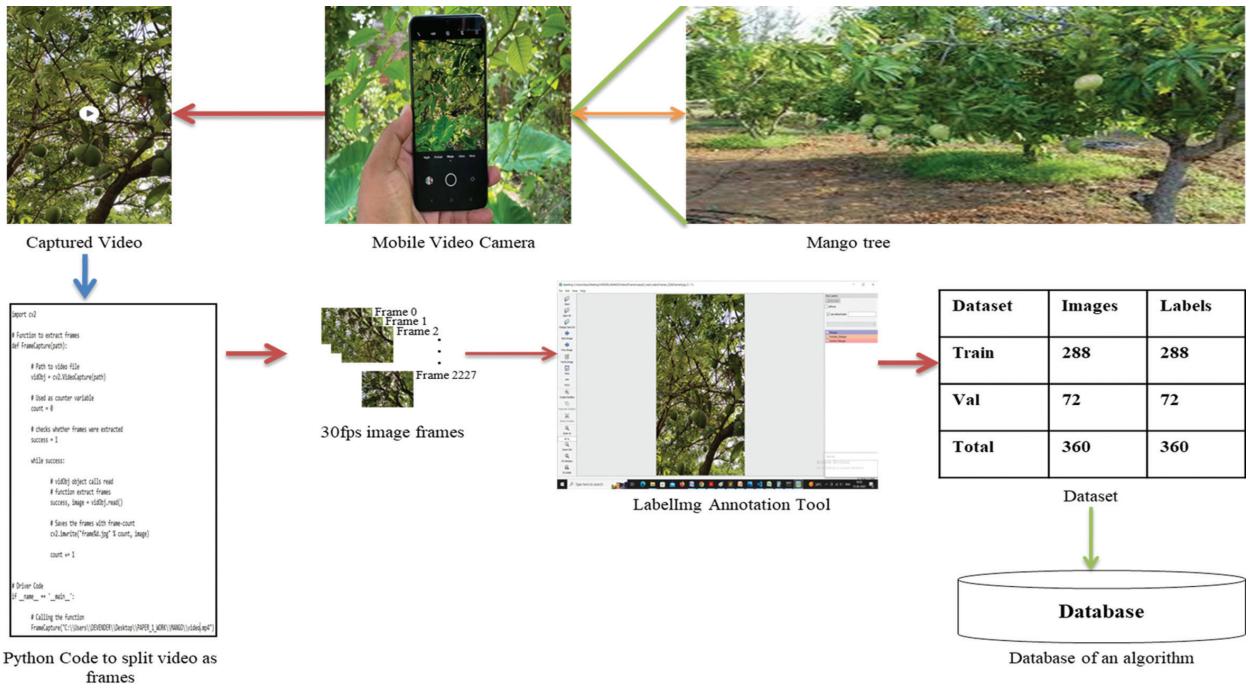


Fig. 2. Create a dataset and store it in YOLO format data to make predictions

3.2.1. LabelImg: LabelImg is written in Python with Qt as its graphical interface and is a graphical image annotation tool. Tzutalin created the popular image annotation tool LabelImg with some contributors, and now it is a developed tool. It is a part of the Label Studio Community [32]. Annotations of image data saved in XML files in the PASCAL VOC format (the tool LabelImg-1.8.6, released on October 10, 2021) also support the YOLO and CreateML formats. The working procedure in Fig. 2 shows that the image will be selected and asked for its format to store for future research. After proper format allocation, the user must draw a label for the object to annotate. Therefore, we used YOLO format with the text file content as "class; x-center; y-center; width; height." This text file will saved in the system, including images, and the user will use it to train the data using pictures.

3.3. METHODS

The flow diagram contained the input, algorithm, database, detection, counting, and output steps. The first step in Fig. 3 is the input part, which obtains the image as input for the algorithm, starts the process on the given input image, and works for detection with the help of a trained dataset from the database created by the researcher. After training the dataset with the help of the YOLOv7 algorithm, mango fruits were detected on the input image and counted using the DeepSORT-count algorithm.

Step 1: Input image.

Step 2: Use the given image for testing and validation based on the dataset already generated by the user and stored in the database.

Step 3: Check whether the given input image class is trained perfectly.

Step 4: Use the trained dataset and the input image with YOLOv5 and YOLOv7 to detect and count the classes available in the input image.

Step 5: If there are more than two classes, the output is given in a multiclass classification of the detected object and count.

Step 6: Output is the Number of Classes with the number of fruits, as shown in (Fig. 4) in a single frame.

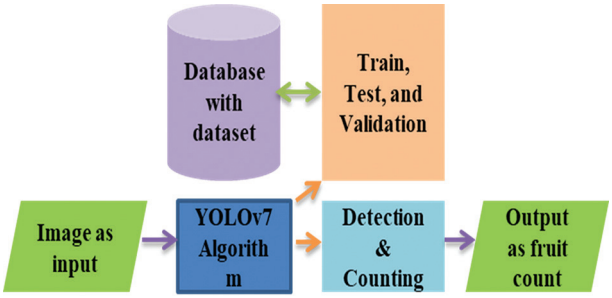


Fig. 3. Counting fruits using YOLOv7: an image as input, and the number of fruits is output

3.3.1. Experiment Platform: Windows 10 Pro 64-Bit, Core 19-9900KF CPU @ 3.60GHz, 32 GB RAM, Dedicated 8 GB Memory, NVIDIA GeForce RTX 2080 SUPER GPU. The model framework was PyTorch with CUDA 11.6, cudnn 11.6, and Python 3.9.0. The original YOLOv7 model used YOLOv7.pt and YOLOv7-tiny.pt for pretraining and retraining based on the pretraining results. The parameters for training were as follows: image input size 640×640, epoch 500, initial learning rate 0.01, and batch size 8.

Fig. 4. Hidden-Mango=20; Mango=4; and Corner-Mango=4; Total=28 Mangoes on tree.

3.3.2. Evaluation and performance of the model: Several evaluation indicators are available to help ascertain and validate the model's functionality, including the confusion matrix, precision, recall, mAP, tradeoff, F1-score, mean, median, mode, variance, standard deviation, and root mean square error. Accuracy is applicable only for binary classes, not for multiclass classification, and the confusion matrix will take care of multiclass classification.

3.3.3. Confusion Matrix: A confusion matrix defines a classification algorithm's performance. It visualizes and summarizes its performance. *TP*: True Positive, *FP*: False Positive, *FN*: False Negative, *TN*: True Negative.

Accuracy: Accuracy is the ratio of the exact classified on-tree mango samples to the number of actual on-tree mango samples in the image for binary class classification.

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

Precision is the ratio of correctly predicted positive on-tree mango classes for all items to be positive.

$$\text{Pr}\ ecision\ = \frac{(TP\ )}{(TP\ +\ FP\ )} \quad (2)$$

*F*1-*Score*: The *F*1-*Score* is a performance metric considering precision and recall values. It will calculated using the two metrics' harmonic mean.

$$F1 - Score = \frac{(2 * \text{Pr}ecision * \text{Re}call)}{(\text{Pr}ecision + \text{Re}call)} \quad (3)$$

3.3.4. SORT—Simple Online and Realtime Tracking: SORT stands for simple online and realtime tracking, an approach for tracking multiple objects with the help of any deep learning algorithm. To obtain the count of the on-tree fruit, we used YOLOv5 and YOLOv7 with SORT. SORT can track an object for extended periods to determine its occlusions [10].

3.3.5. YOLOv5n Algorithm for on-tree fruit counting: After a few days of YOLOv4, YOLOv5, a PyTorch-based approach, was released on May 27, 2020. In YOLOv5, some sub-variants based on 640 image size YOLOv5n (Nano), YOLOv5s (small), YOLOv5m (medium), YOLOv5l (large), and YOLOv5x (extra-large); based on an image size of 1280 are YOLOv5n6, YOLOv5s6, YOLOv5m6, YOLOv5l6, and YOLOv5×6 [33]. All YOLOv5 versions of the first two sub-variant models, called YOLOv5n and YOLOv5s, are used in this study. They then worked on both models and found that YOLOv5n works better for their self-prepared dataset with better accuracy and inference time than YOLOv5s.

Divide the YOLOv5n model into four regions: the input, backbone, neck, and head regions. Here, in the input region, the model takes an image of 1×3×640×640 and calculates the best-fit anchor box value according to the custom dataset. The convolutional layers and spatial pyramid pooling fast [34] were the backbone of this model. The combination of Feature Pyramid Networks (FPN) [35] and Path Aggregation Network (PAN) network layers [36] acts as the neck region in this model. The three detection heads with 1×3×80×80×8, 1×3×40×40×8, and 1×3×20×20×8 scale integration will give the predicted bounding box information to the final output [37].

3.3.6. YOLOv7 Algorithm for counting: YOLOv7 works excellently as an object detector with a high speed from 5 to 160FPS and has the highest accuracy of 56.8% AP using the MS-COCO dataset with 30FPS or higher on a GPU machine. YOLOv7 is a very balanced object detector compared with all known object detectors in speed and accuracy. YOLOv7 architecture [38], the pipeline has three significant parts: backbone,

encoder, and decoder. Again, the architecture of YOLOv7 consists of three parts: the input, backbone, and head. The 640×640 image was used in the input part, whereas, in backbone feature extraction, the image at the head strengthened the feature extraction network and made it ready for prediction.

3.4. COUNTING MANGO ON THE TREE

Fig. 5 shows a combination of several procedures: the input, output, annotation, database, and processing components, respectively.

Input: Eight image frames of each tree are considered input images; these images are 1080 pixels wide by 1920 pixels high and between 900 KB and 1100 KB in size.

Annotation: Based on chapter (3.2), the researcher labeled fruit annotations.

Database: Based on the above annotation procedure, the created dataset was stored in this database for future reference and use during training, testing, and validation. If the dataset images are not maintained and the labels should be the same, then algorithm failure is possible.

- **Processing:** The backbone and neck network-related procedures will processed as a processing model, which is the immediate step to the input part.
- **Output:** The output part is immediately adjacent to the neck region of the YOLO version. In these models, the three head-bounding boxes provide output predictions.

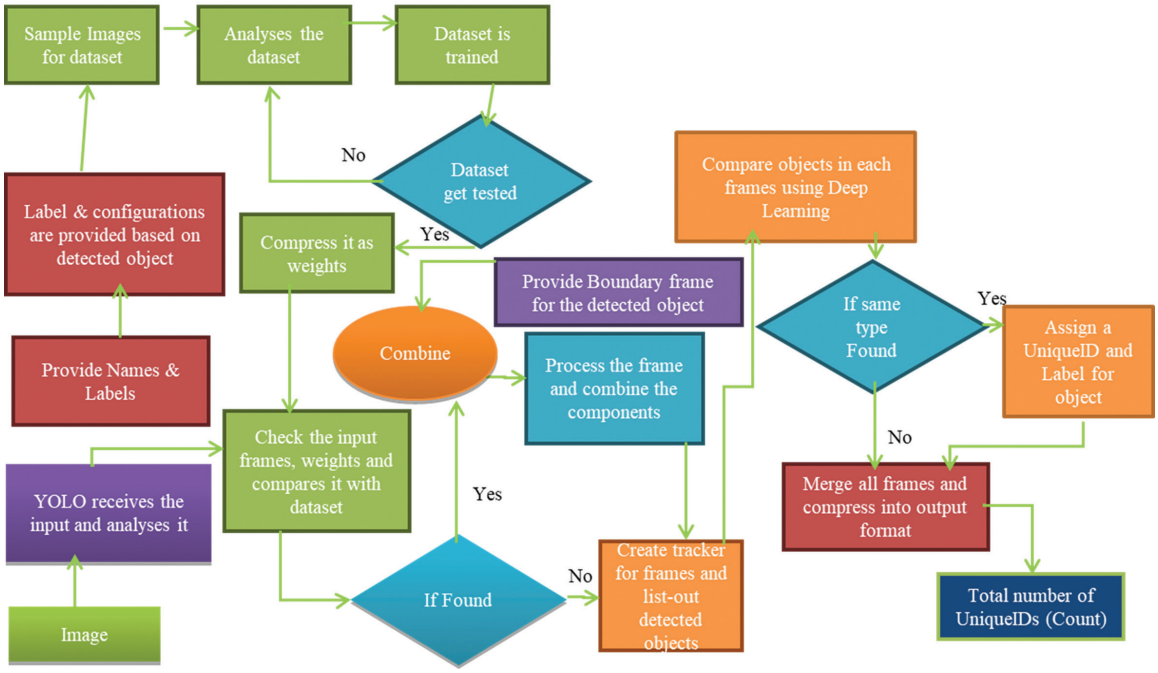


Fig. 5. Fruit counting with YOLOv7 algorithm using an image as input and number of fruits as output based on stored database data

4. RESULTS

4.1. ONE TREE FRUIT COUNT IN ORCHARD

One tree was selected, and 130 fruits were counted manually on May 30, 2022. Based on (Chapter 3.3.1), we used multiclass classification, a 4-side image model on YOLOv5n, YOLOv5s, YOLOv7, and YOLOv7-tiny and worked; the results were stored and performed the same with 8-side model got the results (Eq. (5)) and finally evaluated the data for a more suitable model. Among the four YOLO sub-versions, YOLOv7 is very close to the reality of 97.7% accuracy and achieves a lower average inference of just 17.112 ms of time. Counting the fruit using the 4-side image of the tree procedure is too far from the actual result, but the 8-side image of the tree is too close to the reality of the tree fruit count. Consequently, employing the 8-side model is preferable to the 4-side model when counting the fruit on the tree. Selected One hundred seventy-nine images for training from 224 images; the remaining 20% were for testing and validating the results in Table 1.

Table 1. On-tree fruit counting with YOLOv5 and YOLOv7 models of a single tree.

Models	Manual Count	4-Side Count	Inference (s)	8-Side Count	Inference (s)
YOLOv5n	130	66	0.0384	135	0.03927
YOLOv5s	130	71	0.07155	145	0.07292
YOLOv7	130	60	0.01575	133	0.01711
YOLOv7-tiny	130	71	0.01075	141	0.01200

$$4 - Side\_Count = \sum_{i=1}^{4} (F_{TM} - F_{CM}) \quad (4)$$

$$8 - Side\_Count = \sum_{i=1}^{8} (F_{TM} - F_{CM}) \quad (5)$$

F_{TM} is the frame's total mango count, and F_{CM} is the corner mango of the frame.

4.2. MULTIPLE TREE FRUITS COUNT IN AN ORCHARD

4.2.1. Manual Counting of On-Tree: Some trees in an orchard were counted based on our eight-sided model. Then, the number of fruits available on each tree and the resulting count can be validated. They picked a few fruits off the tree, and some were in different structures, so there were ups and downs in the graph in the manual count flow, as shown in (Fig. 6).

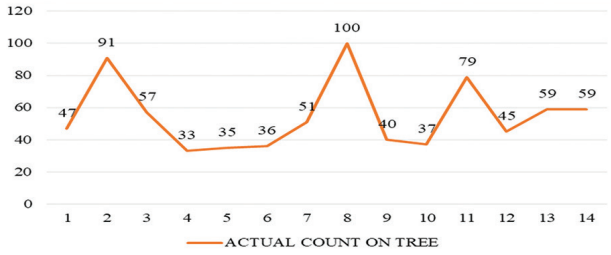


Fig. 6. On trees manually, there is a fruit count of fourteen trees

4.2.2. Avoiding Double Count using the Eight-Sided Model: If images are taken precisely on the eight sides of the tree, then the procedure below will work. Because eight exact images of the tree based on eight directions, such as east, south, west, north, south-east, south-west, north-east, and north-west, will give only eight frames. At the corner, mangoes will counted twice; to avoid this problem, we used an eight-sided model (Eq. (5)). However, in this research, we split a video into frames to obtain eight-sided images; thus, the total frames were divided by eight, and then eight were obtained. Then, the frame in the east is the first frame, and the last in the northeast frame is the eighth frame used as the eight-direction images.

- YOLOv7: Considered 14 trees of an orchard and performed the task using the YOLOv7 model, and observed as in Fig. 7, which has a significant impact on the count of the tree using the YOLO model, where the predicted count is very close to the actual count of fruits on the tree with a good accuracy of 95.48% with 17 ms of inference time.
- YOLOv7-tiny: YOLOv7-tiny also used the same dataset and performed the same task as in the YOLOv7 model, as shown in Fig. 7, where YOLOv7-tiny also performed the same as YOLOv7, with some differences. The detection and counting accuracy was only 94.1%, with an average inference time of 16 ms.
- YOLOv5n: YOLOv5n also used the same dataset and performed the same task as in the YOLOv7 model, as shown in Fig. 7, where YOLOv5n also performed well, with an average inference time of 94.1% with 104.5 milliseconds of average inference time.
- YOLOv5s: With the help of the same dataset, performed the same task as in the YOLOv7 model and observed as in Fig. 7, where YOLOv5s also performed well, with an average inference time of 97.2% with 85.69 milliseconds of average inference time.

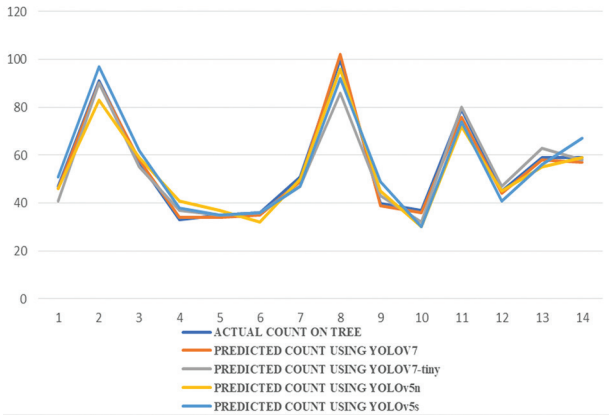


Fig. 7. Comparison of four models and validation of fourteen trees with actual count using eight-sided models

5. DISCUSSION

The researchers used four models and compared them in different ways. With the help of 360 image frames and their annotated labels, only the model trained at different epochs with batch sizes of 8 and 16. Then, 288 image frames were used only for training purposes, and the remaining for testing and validating the YOLOv7 and YOLOv7-tiny models. Trained these models, compared them at 100 and 500 epochs, and observed that YOLOv7 is the best model with good accuracy and a meager inference time for tree fruit count prediction.

Three procedures determine the best time to capture an image under natural lighting conditions. In these three procedures, we used 96 images of datasets; however, based on their lighting conditions, the storage size differed automatically, as shown in Table 2.

Table 2. Lighting conditions validation accuracy in a day to capture images

Time	Dataset Size	Storage size	Detection Accuracy
Morning	96 Images	363MB	84%
Day	96 Images	375MB	42%
Evening	96 Images	359MB	57%

We considered 06:00 AM to 09:00 AM as the morning time, 10:00 AM to 03:00 PM as the daytime, and 04:00 PM to 06:00 PM as the evening time. We captured images, prepared a new dataset for each time of the day, and trained and evaluated them.

Natural lighting was impossible at night, except during the full moon day. It is necessary to use artificial light for capture, which is compatible with the device's camera.

Consider three different time-captured images and the number of fruits in these images counted manually. Images were compared with the best training result of each timing separately, and then a table was created for these results. We then compared the images and the timing of the picture capture, as shown in Table 3, and observed tremendous results with the help of YOLOv7. Based on the results of Tables 2 and 3, we propose that morning image capturing is the best time to capture images of trees in an orchard.

Table. 3. Comparison of best fruit image capturing time of on-tree

Time	Morning	Day	Evening	Actual Count
Morning Images	21	9	11	25
Day Images	8	6	5	14
Evening Images	16	8	11	19

6. CONCLUSION

Implementing video capture and YOLOv7 for fruit counting offers significant advancements in agricultural technologies. Employing a strategic frame selection process ensures accurate and efficient fruit counts while minimizing redundancy and double counting. This innovation enhances yield estimation and optimizes resource allocation and early issue detection, reducing labor costs and increasing overall productivity. Its scalability and adaptability make it suitable for various orchards and commercial agricultural operations. Ultimately, this approach empowers farmers with precise, realtime data, enabling informed decision-making and contributing to improved profitability and sustainable farming practices. Owing to the rapid explosion of data in agriculture and horticulture sciences, a new trending computer science area, deep learning technology, has become a hot research focus for a new era in artificial intelligence. To determine the actual count of on-tree mango fruits, the researcher performed experiments with four algorithms, YOLOv5n, YOLOv5s, YOLOv7, and YOLOv7-tiny, using an eight-sided imaging technique around the tree, which showed that YOLOv7 performed the best about accuracy and inference. In this study, a deep learning framework was compared and applied to a computer vision algorithm for fruit detection and counting of trees using videos and images. It also presented the most suitable time to capture the images for better detection in the morning, daytime, and evening. It proposed that capturing morning-time images under natural conditions is the best time for on-tree fruits. In this on-the-spot situation, the video split image-converted dataset model helps count the object using the most miniature image training with better accuracy. Even though this method counts ideally, it also has some limitations; the video should be captured only in the forward direction, not in slow motion, and the reverse direction while capturing the video.

7. REFERENCES

[1] P. Wang, Y. Luo, J. Huang, S. Gao, G. Zhu, Z. Dang, J. Gai, M. Yang, M. Zhu, H. Zhang, X. Ye, "The genome evolution and domestication of tropical fruit mango", Genome Biology, Vol. 21, 2020, pp. 1-17.

[2] H. Kim, M. J. Castellon-Chicas, S. Arbizu, S. T. Talcott, N. L. Drury, S. Smith, S. U. Mertens-Talcott, "Mango (Mangifera indica L.) polyphenols: Anti-inflammatory intestinal microbial health benefits, and associated mechanisms of actions", Molecules, Vol. 26, No. 9, 2021, p. 2732.

[3] A. Wall-Medrano, F. J. Olivas-Aguirre, J. F. Ayala-Zavala, J. A. Domínguez-Avila, G. A. Gonzalez-Aguilar, L. A. Herrera-Cazares, M. Gaytan-Martinez, "Health benefits of mango by-products", Food wastes and by-products: Nutraceutical and Health Potential, 2020, pp. 159-191.

[4] A. J. Mobolade, N. Bunindro, D. Sahoo, Y. Rajashekar, "Traditional methods of food grains preservation and storage in Nigeria and India", Annals of Agricultural Sciences, Vol. 64, No. 2, 2019, pp. 196-205.

[5] Y. Xue, N. Huang, S. Tu, L. Mao, A. Yang, X. Zhu, X. Yang, P. Chen, "Immature mango detection based on improved YOLOv2", Transactions of the Chinese Society of Agricultural Engineering, Vol.34, No. 7, 2018, pp. 173-179.

[6] P. J. Ramos, F. A. Prieto, E. C. Montoya, C. E. Oliveros, "Automatic fruit count on coffee branches using computer vision", Computers and Electronics in Agriculture, Vol. 137, 2017, pp. 9-22.

[7] M. Rahnemoonfar, C. Sheppard, "Deep count: fruit counting based on deep simulated learning", Sensors, Vol. 17, No. 4, 2017, p. 905.

[8] S. Kamkar, R. Safabakhsh, "Vehicle detection, counting and classification in various conditions", IET Intelligent Transport Systems, Vol. 10, No. 6, 2016, pp. 406-413.

[9] S. Lyu, R. Li, Y. Zhao, Z. Li, R. Fan, S. Liu, "Green citrus detection and counting in orchards based on YOLOv5-CS and AI edge system", Sensors, Vol. 22, No. 2, 2022, p. 576.

[10] N. Wojke, A. Bewley, D. Paulus, "Simple online and realtime tracking with a deep association metric," Proceedings of the IEEE International Conference

on Image Processing, Beijing, China, 17-20 September 2017, pp. 3645-3649.

[11] Y. Qiao, Y. Hu, Z. Zheng, H. Yang, K. Zhang, J. Hou, J. Guo, "A counting method of red jujube based on improved YOLOv5s", Agriculture, Vol. 12, No. 12, 2022, p. 2071.

[12] A. I. B. Parico, T. Ahamed, "Real time pear fruit detection and counting using YOLOv4 models and deep SORT", Sensors, Vol. 21, No. 14, 2021, p. 4803.

[13] S. Bargoti, J. P. Underwood, "Image segmentation for fruit detection and yield estimation in apple orchards", Journal of Field Robotics, Vol. 34, No. 6, 2017, pp. 1039-1060.

[14] U. O. Dorj, M. Lee, S. S. Yun, "An yield estimation in citrus orchards via fruit detection and counting using image processing", Computers and Electronics in Agriculture, Vol. 140, 2017, pp. 103-112.

[15] A. Koirala, K. B. Walsh, Z. Wang, C. McCarthy, "Deep learning for realtime fruit detection and orchard fruit load estimation: Benchmarking of 'MangoYOLO", Precision Agriculture, Vol. 20, No. 6, 2019, pp. 1107-1135.

[16] O. E. Apolo-Apolo, M. Pérez-Ruiz, J. Martinez-Guanter, J. Valente, "A cloud-based environment for generating yield estimation maps from apple orchards using UAV imagery and a deep learning technique", Frontiers in Plant Science, Vol. 11, 2020, p. 1086.

[17] J. P. Vasconez, J. Delpiano, S. Vougioukas, F. A. Cheein, "Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation", Computers and Electronics in Agriculture, Vol. 173, 2020, p. 105348.

[18] H. Mirhaji, M. Soleymani, A. Asakereh, S. A. Mehdizadeh, "Fruit detection and load estimation of an orange orchard using the YOLO models through simple approaches in different imaging and illumination conditions", Computers and Electronics in Agriculture, Vol. 191, 2021, p. 106533.

19] F. Gao, W. Fang, X. Sun, Z. Wu, G. Zhao, G. Li, R. Li, L. Fu, Q. Zhang, "A novel apple fruit detection and counting methodology based on deep learning and trunk tracking in modern orchard", Computers and Electronics in Agriculture, Vol. 197, 2022, p. 107000.

[20] S. Lu, W. Chen, X. Zhang, M. Karkee, "Canopy-attention-YOLOv4-based immature/mature apple fruit detection on dense-foliage tree architectures for early crop load estimation", Computers and Electronics in Agriculture, Vol. 193, 2022, p. 106696.

[21] Y. Zhang, W. Zhang, J. Yu, L. He, J. Chen, Y. He, "Complete and accurate holly fruits counting using YOLOX object detection", Computers and Electronics in Agriculture, Vol. 198, 2022, p. 107062.

[22] Z. Wu, X. Sun, H. Jiang, W. Mao, R. Li, N. Andriyanov, L. Fu, "NDMFCS: An automatic fruit counting system in modern apple orchard using abatement of abnormal fruit detection", Computers and Electronics in Agriculture, Vol. 211, 2023, p. 108036.

[23] C. Wang, C. Li, Q. Han, F. Wu, X. Zou, "A performance analysis of a litchi picking robot system for actively removing obstructions, using an artificial intelligence algorithm", Agronomy, Vol. 13, No. 11, 2023, p. 2795.

[24] Y. Tang, J. Qiu, Y. Zhang, D. Wu, Y. Cao, K. Zhao, L. Zhu, "Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: A review", Precision Agriculture, Vol. 24, No. 4, 2023, pp. 1183-1219.

[25] G. Yu, R. Cai, Y. Luo, M. Hou, R. Deng, "A-pruning: a lightweight pineapple flower counting network based on filter pruning", Complex & Intelligent Systems, Vol. 12, No. 2, 2024, pp. 2047-2066.

[26] Y. Feng, W. Ma, Y. Tan, H. Yan, J. Qian, Z. Tian, A. Gao, "Approach of Dynamic Tracking and Counting for Obscured Citrus in Smart Orchard Based on Machine Vision", Applied Sciences, Vol. 14, No. 3, 2024, p. 1136.

[27] X. Zhai, Z. Zong, K. Xuan, R. Zhang, W. Shi, H. Liu, Z. Han, T. Luan, "Detection of maturity and counting of blueberry fruits based on attention mechanism and bi-directional feature pyramid network", Journal of Food Measurement and Characterization, 2024, pp. 1-16.

[28] C. Yang, T. Geng, J. Peng, Z. Song, "Probability map-based grape detection and counting", Computers and Electronics in Agriculture, Vol. 224, 2024, p. 109175.

[29] M. Chen, Z. Chen, L. Luo, Y. Tang, J. Cheng, H. Wei, J. Wang, "Dynamic visual servo control methods for continuous operation of a fruit harvesting robot working throughout an orchard", Computers and Electronics in Agriculture, Vol. 219, 2024, p. 108774.

[30] M. Stein, S. Bargoti, J. Underwood, "Image based mango fruit detection, localisation and yield estimation using multiple view geometry", Sensors, Vol. 16, No. 11, 2016, p. 1915.

[31] C. Zheng, P. Chen, J. Pang, X. Yang, C. Chen, S. Tu, Y. Xue, "A mango-picking vision algorithm on instance segmentation and key point detection from RGB images in an open orchard", Biosystems Engineering, Vol. 206, 2021, pp. 32-54.

[32] L. Tzutalin, "LabelImg", https://github. com/tzutalin/labelImg (accessed: 2020)

[33] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, J. Fang, K. Michael, D. Montes, J. Nadar, P. Skalski, Z. Wang, "ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference", Zenodo, 2022.

[34] D. Xu, H. Zhao, O.M. Lawal, X. Lu, R. Ren, S. Zhang, "An automatic jujube fruit detection and ripeness inspection method in the natural environment", Agronomy, Vol. 13, No. 2, 2023, p. 451.

[35] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017, pp. 936-944.

[36] W. Wang, E. Xie, X. Song, Y. Zang, W. Wang, T. Lu, G. Yu, C. Shen, "Efficient and Accurate Arbitrary-Shaped Text Detection With Pixel Aggregation Network", Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October - 2 November 2019, pp. 8439-8448.

[37] W. Yang, X. Ma, W. Hu, P. Tang, "Lightweight blueberry fruit recognition based on multi-scale and attention fusion NCBAM", Agronomy, Vol. 12, Vo. 10, 2022, p. 2354.

[38] C.-Y. Wang, A. Bochkovskiy, H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17-24 June 2023, pp. 7464-7475.

Minimum Skewness based Myocardial Infarction Detection Model using Classification Algorithms

Original Scientific Paper

Satish Kumar Tewalker*
Department of IT and Computer Application
Dr. C. V. Raman University
Kota Bilaspur,India
satishkumarphd88@gmail.com

Ragini Shukla
Department of IT and Computer Application
Dr. C. V. Raman University
Kota Bilaspur,India
raginishukla008@gmail.com

*Corresponding author

Abstract *– Myocardial infarction is one of the most dangerous public health issues in the world. The accurate prediction of myocardial infarction disease aids in disease diagnosis and biological analysis of the patient's health. The classification algorithms are one of the solutions that predict accurate diseases based on the symptoms (attributes) in patients' details. The ability to predict accurately reduces the risk of causality and decision-making time. This study proposed the Minimum Skewness-Based Myocardial Infarction Detection Model (MSMIDM) with the help of a statistical and feature selection-based approach. Minimum skewness is a feature selection statistical approach that selects essential attributes of a dataset. The MSMIDM provides accurate results with the highest accuracy among the six classification algorithms. The experimental analysis makes use of the most widely used Cleveland dataset for myocardial infarction detection. The experimental results are analyzed through a confusion metric, statistical, and partitional validation approach. The proposed model obtains an accuracy of 90%, 87.037%, 87.037%, 83.238%, 81.481%, and 85.556% with respect to Random Forest, K-Nearest Neighbor, Support Vector Machine, Naive Bayes, Decision Tree and Neural Network classification algorithms. According to the experimental analysis, this study suggests the MSMIDM-based Random Forest algorithm is excellent for myocardial infarction disease detection.*

Keywords: *Heart Diseases, Statistical Classification, Feature Selection, Skewness Classification, Myocardial Infarction Detection Model*

Received: April 20, 2024; Received in revised form: June 17, 2024; Accepted: July 2, 2024

1. INTRODUCTION

Heart disease is the most dangerous disease nowadays that impacts numerous individuals in their middle or late years and leads to deadly consequences for many causes. According to WHO figures, cardiac disorders account for 24% of all non-communicable disease fatalities in India [1]. Heart disorders account for one-third of all global fatalities [2] [3]. Around 17 million individuals die each year as a result of cardiovascular disease (CVD), with the ailment being especially prevalent in Asia [4]. Tobacco use, cholesterol, high blood pressure, family history, poor diet, alcohol consumption, eating habits, physical inactivity, diabetes and obesity are all known heart disease risk aspects [5].

Myocardial infarction (MI) is the most widespread heart disease and it is the most prevalent cause of death and morbidity, as well as a significant cost of care because it is a major cause of mortality and disability all around the world [6]. Myocardial Infarction generally recognized as a heart attack, which is one of the most dangerous and deadly cardiovascular diseases [7]. Existing research indicates a relationship between the formation of MI disease and a blockage in the

coronary artery. The first clinical accounts arose in the early twentieth century as clinical symptoms were connected to the coronary artery. Denmark et al. (2019) [6] depicted the five principal varieties of MI. The first is a coronary incident. The second cause of MI is caused by an oxygen supply. The third case is a fatal cardiac death. Percutaneous coronary intervention is the fourth cause of MI, and the last significant type is connected to coronary artery bypass grafting.

Manually determination of risk factors for myocardial infarction and other diseases is a challenging task [8]. Machine learning techniques help extract hidden risk factors for myocardial infarction heart diseases with the help of existing data and high accuracy. Machine learning is used to extract hidden information, facts, patterns, and knowledge related to diseases from a variety of disease-related data sources. Machine learning determines these sorts of patterns by employing intelligible functions and procedures. In general, machine learning algorithms are classified as either predictive or descriptive. Predictive learning employs inference to drive conclusions from data sources. Descriptive learning is the differentiation of broad properties in any data storage [9].

Classification is one of the most essential and effective approaches in predictive learning, which is known as supervised machine learning. Classification is essential for dealing with massive amounts of data, and it is used for data analysis to predict class labels in the form of categorical labels [10]. The classifier employed two steps for predicting the class label. The first step is learning, where a predicated class is produced from a dataset of known classes. The second step predicts the class labels based on the dataset. The second step is an important task that is used by the constructed classifiers [11]. There are several data mining classification techniques including Rule-Based Classifier, Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naive Bayes (NB) and Neural Network (NN).

Myocardial infarction is a leading cause of mortality and disability around the world. The prediction of myocardial infarction disease is beneficial for reducing the death ratio, and machine learning assists in identifying the voluntary history of many diseases, including myocardial infarction diseases of the heart. Researchers employed well-known biostatistical approaches to predict the relationship between MI and associated risk factors in patients. Biostatistical approaches are inappropriate for identifying correct data and risk indicators. There is a global need to use more appropriate machine learning algorithms with statistical techniques in the medical field for disease prediction. The objective of this research is to predict MI among chronic diseases through a classification approach to extract knowledge from data with the help of statistical techniques.

The rest of the paper is structured as follows: Section 2 reviews the various classification models for myocardial infarction detection. Section 3 details the objective function of classification and proposes the classification model through a feature selection statistical approach. Section 4 implements the proposed classification model and compares its performance with various classical classification models and algorithms. Section 5 concludes the study and describes further work.

2. RELATED WORK

The classification model is essential in predicting myocardial infarction heart disease using various statistical approaches, feature selection algorithms, optimization metrics, etc. This section examined the literature related to several investigations and classification models for myocardial infarction heart disease prediction.

Parthiban and Srivatsa [12] select the attributes with the help of correlation and predict myocardial infarctions through SVM and NB classification methods.

Sun et al. [13] constructed an automatic detection system for myocardial infarction in patients utilizing supervised machine learning techniques known as multiple instance learning (MIL). The experimental results of the proposed MIL algorithm are assessed using the classical MIL method with the help of the ECG and PTB diagnostic databases.

Safdar et al. [14] investigated seven classification and association rules learning algorithms to produce an accurate risk analysis prediction model in myocardial infarction. The experiment findings are compared utilizing validation models for accuracy, sensitivity, specificity and precision. The CHAID decision tree model outperforms other classification techniques by obtaining 93.4% precision.

Seenivasagam et al. [7] suggested a computer-aided detection method that assesses a supervised machine learning classifier to assess the risk level of myocardial infarction. The experimental investigation demonstrated that the PSONN classifier beat other algorithms regarding risk prediction, while the neural network classifier exceeded other algorithms regarding dataset training accuracy.

Otoom et al. [15] applied hold-out and cross-validation tests and used NB, SVM, and DT classification algorithms to predict myocardial infarction heart diseases. Vembandasamy et al. [16] employ the various computational parameters of the NB method to predict heart diseases.

Daraei et al. [17] developed the model for MI detection using the cost-sensitive J48 approach. It employed a hybrid feature selection approach for feature selection, whereas the meta cost classifier to prediction. The experimental research revealed that combining hybrid feature selection and metacost classifier enhances sensitivity and obtains a greater propensity for MI prediction compared to classical classification.

Verma et al. [18] use a hybrid method that includes

correlation-based feature selection, K-means clustering, particle swarm optimization, multinomial logistic regression, multi-layer perceptrons and fuzzy unordered rule induction algorithm to predict myocardial infarction. Chadha and Mayank [19] extract disease patterns from vital data and utilize NN, DT, and NB classification algorithms to extract disease patterns.

Dwivedi et al. [20] established a classification framework for detecting heart disease utilizing tenfold cross-validation and machine-learning algorithms. The experimental analysis shows that the logistic regression classification algorithm achieves excellent levels of classification accuracy (85%), sensitivity (89%) and specificity (81%) as compared to other classification algorithms.

Mokeddem et al. [21] developed a clinical decision support system (CDSS) to predict MI and provide diagnosis for Coronary Artery Diseases (CAD). The proposed diagnosis model used RF, C5.0 decision tree and fuzzy classification algorithms. The RF method is used to extract features based on rank, the C5.0 decision tree is used to generate crisp rules, and the fuzzy algorithm is used to forecast disease and diagnosis. The CDSS model has the capability to handle missing and noisy data. The experimental research revealed that the CDSS optimized processing time and provided a classification accuracy of up to 90.50%.

Ed-daoudy and Maalmi [22] uses the well-known SVM, DT, RF and Logistic Regression classification algorithms and implement inside the Apache Spark for big data processing. The proposed model improved accuracy and computing time while suggesting a logical resolution for detecting myocardial infarction disease. Anitha and Sridevi [23] used SVM, KNN, and NB classification algorithms to improve the accuracy of the heart disease detection framework. The experiment results indicate that NB is excellently suited for detecting myocardial infarction disease.

Christalin Latha et al. [1] enhance the accuracy of classification algorithms to detect heart disease using ensemble classifiers. It uses ensemble techniques such as stacking, bagging, majority voting, and boosting. The bagging method increases accuracy by 6.92%. The boosting method increases accuracy to 5.94. The majority voting method increases accuracy by 7.26% and the stacking method increases accuracy to 6.93%.

Garate-Escamila et al. [24] suggested a dimensionality reduction approach for detecting risk characteristics of heart disease. The suggested dimensionality reduction approach is the CHI-PCA algorithm. It employs PCA for feature extraction and CHI for feature selection. The experimental findings demonstrated that CHI-PCA with the random forest classification algorithm achieves the maximum accuracy in various data sets.

Shah et al. [25] assessed the effectiveness of several supervised machine learning methods to detect the probability of heart disease in patients, such as myocardial infarction. The experimental results demonstrated that the KNN classifier outperformed the NB, DT, and RF algorithms in accuracy.

Mandair et al. [26] developed a machine-learning model to predict MI utilizing random sampling and a deep neural network. The experimental study employed harmonized Electronic Health Record data and seven layers of one hundred neurons per layer. According to the experimental results, the suggested model has a specificity of 73.3%, a sensitivity of 82%, a recall of 82%, and an accuracy of 05%.

Ibrahim et al. [27] developed the active MI prediction framework using a synthetic minority oversampling technique (SMOTE), recurrent neural network (RNN), decision-tree, convolutional neural network (CNN) and XGBoost classification methods. The suggested model achieves 89.9%, 84.6%, and 97.5% accuracy.

Reddy et al. [28] developed a forecasting system for myocardial infarction utilizing hybrid machine learning methods. The suggested approach outperforms the KNN, DT, NN, NB, RF and SVM classifiers in terms of accuracy. The experiment findings reveal that the random forest method achieves a superior accuracy compared to other classical algorithms.

Patro et al. [29] designed a heart disease prediction framework utilizing risk attributes and Salp Swarm Optimized Neural Network (SSA-NN), NB, KNN, and Bayesian Optimized Support Vector Machine (BO-SVM) classification algorithms. The experimental results revealed that the proposed framework optimized the classification algorithm and provided an excellent and effective healthcare monitoring system.

Kondababu et al. [30] integrated the properties of RF and linear method and constructed a hybrid random forest-linear approach to predict myocardial infarction disease. Liu et al. [31] presented a risk prediction model for actively monitoring patients with myocardial infarction disease. Top-layer and bottom-layer algorithms are used in the suggested risk prediction model. The top layer contains the recursive feature elimination technique, while the bottom layer has the gradient boosting decision tree (GBDT), RF, SVM, and logistic regression algorithms.

Nagavelli et al. [32] improved the machine-learning model that detects heart disease. The improved model employs the Naïve Bayes classifier with weighted models, SVM classifier with XGBoost algorithm. The experimental investigation demonstrates that the single XGBoost algorithm obtains excellence in permanence, whereas the Naïve Bayes classifier obtains the lowest accuracy when compared to other algorithms.

Sadiyamole et al. [33] recommended combining the advantages of genetic algorithms with ensemble deep learning to predict heart disease. The genetic algorithm retrieves relevant attributes from the dataset, while the deep learning technique predicts the presence of heart disease. The proposed method achieves excellent accuracy as compared to the classical deep learning approach.

Ahmad et al. [34] used the jellyfish method to optimize the machine-learning model. The jellyfish method extracts the dataset's essential features and uses machine learning to predict heart illnesses. The experiment analysis demonstrates that the SVM classifier outperforms other machine learning techniques. The suggested model addresses model complexity and overfitting-related issues.

Bhatt et al. [35] employ a clustering and classification classifier to accurately predict cardiovascular diseases. The suggested model first employs K-mode clustering to classify the data before beginning the classification procedure. In classification, the data was divided into 80:20 ratios for building models, and a number of classifiers were utilized to predict heart disease. The experimental results demonstrate that K-mode clustering and multilayer perceptrons are more accurate.

Subramani et al. [36] implemented a cardiovascular disease detection model in a real-time environment using the Internet of Things. The proposed model first collects the data from IoT devices and then preprocesses the data according to the classification model. The classification model observed the data and automatically decided the training and testing classifiers. The experimental analysis shows that the deep learning model achieves excellent performance as compared to other classifiers.

Al Alshaikh et al. [37] improved the cardiovascular heart disease prediction model by combining different feature extraction-related algorithms and machine learning classifiers. The improved model initially employs feature-related algorithms. The proposed model's second stage includes classification tasks performed by a multilayer deep convolutional neural network. The multilayer deep convolutional neural network outperforms another classifier using the adaptive elephant herd optimization approach.

3. PROPOSED ALGORITHM

This section states the classification objective and describes the Minimum Skewness Myocardial Infarction Detection Model (MSMIDM) using a different classification algorithm.

3.1. CLASSIFICATION OBJECTIVE FUNCTION

The preprocessed MDS myocardial infarction diseases dataset consists of $X=\{x_1, x_2, x_3 \ldots x_N\}$ data points with $X_d=\{x_{1d}, x_{2d}, x_{3d} \ldots x_{Nd}\}$ attributes of myocardial infarction diseases. The dimensions' attributes defined the disease's behavior, symptoms, and nature of the sample data. The classification algorithms divide the MDS dataset into various classes based on the symptoms of diseases. The predicted class describes the severity of myocardial infarction diseases and suggests a diagnosis through a consultation with a physician. The goal of classification is to maximize predicted class accuracy while minimizing diagnosis and analysis time for myocardial infarction disease patients.

Suppose $C=\{c_1, c_2, c_3 \ldots c_k\}$ is the predicted class of the classification and $P=\{p_1, p_2, p_3 \ldots p_{k'}\}$ is the predefined class of the myocardial infarction diseases dataset. The K and K' represented the number of classes related to diseases. The number of classes of classification results and predefined class is equal and their intersection is always null. The $N_{i,j}$ is the number of correctly classified data that belongs to both predicted and predefined classes. The N_i is the number of data points of the predicted class of the classification. The N_j is the number of data points that belong to only a pre-defined class. Thus, the objective of the classification is to maximize the $N_{i,j}$ number of data points to obtain better accuracy. The formulation of the objective function is shown in Equation (1) to (3), respectively {Formatting Citation}.

$$Max = \sum_{i=1}^{k} max_{j=1}^{K'} N_{i,j} \quad (1)$$

$$s.t. \quad min\, N_j \; and \; N_i \quad (2)$$

$$N_j \cap N_i = \emptyset \quad (3)$$

3.2. CLASSIFICATION ALGORITHM

Machine learning algorithms utilize a diverse range of statistical, mathematical, probabilistic, and optimization-related methods to learn from previous experience and detect meaningful patterns, knowledge, and trends in large, unstructured, and complex disease datasets. This study uses Random Forest (RF), Support Vector Machine (SVM), Naive Bayes(NB), K-Nearest Neighbor (KNN), Decision Tree (DT), Neural Network (NN) machine learning algorithms for myocardial infarction disease detection [20, 25, 38-41].

3.3. PROPOSED MYOCARDIAL INFARCTION DETECTION MODEL

This section describes MSMIDM in three stages. The first phase of MSMIDM is to select the most significant adverse skewness attributes by utilizing the mean and standard deviation of the myocardial infarction disease datasets. The selected attributes avoid outlining and noise in the dataset during the classification. The second phase used a different classification algorithm to predict myocardial infarction disease with the help of selected negative skewness attributes.

The third phase selects the results with the highest accuracy across all classification algorithms. Thus, the first phase identifies the probable attributes of the dataset, the second phase assesses the accuracy of all selected classification algorithms, and the last phase generates correct results for myocardial infarction disease.

3.3.1 Skewness

Skewness is a third-moment statistical approach that analyzes the symmetry and accuracy of a dataset utilizing the normal distribution's probability distribution. The skewness of the data is classified as positive or negative depending on its skewed curve. Inside

a positive (maximum) skewness, the mean value is more significant than the mode value. Therefore, the data is skewed towards the right side of the curve. The mode value exceeds the mean within the negative (minimal) skewness.

Therefore, the data is skewed towards the left side of the curve. The normal distribution selects those data points that are shifted to higher side. If most of the data is moved inside a positive curve, it indicates that data is acceptable for classification. If most of the data is moved into a negative curve, it indicates that data is used for classification [42, 43]. Here, the most negative data are selected for classification that reason this study divided the data into 80:20 ratios. If we select the 80 % negative data during the classification, then the data is divided into 70:30 ratios. Thus, the data division ratio is dependent on the skewness value and it is validated by the K-fold cross-validation.

The Fisher-Pearson coefficient formulation of skewness was employed in this investigation. The equation (4) depicts the formulation of skewness.

$$S = \frac{\sum(x_i - \bar{x})^3/n}{\sigma^3} = \frac{3\,(\mu - m)}{\sigma} \quad (4)$$

Where x_i is the data point of the selected attributes of the dataset, $\bar{x}$ and μ is the mean of the selected attribute of the dataset, n is the total number of data points inside a selected attribute of the dataset, m is the median of the selected attribute of the dataset, σ_3 is the third moment of standard deviation for the selected attributes, σ is the standard deviation of a specific attribute of the dataset.

3.3.2. Algorithm Description

The Minimum Skewness-based Myocardial Infarction Detection Model (MSMIDM) is described in this subsection through skewness, classification algorithms, and accuracy measurement for myocardial infarction detection. The algorithm and flow chart of the proposed MSMIDM algorithm are outlined in Algorithm 1 and Fig. 1, which find excellent accuracy in the myocardial infarction disease prediction.

4. EXPERIMENTAL ANALYSIS

This section implements the proposed model for myocardial infarction heart disease prediction with the help of an experimental environment, myocardial infarction disease dataset, characteristics of the dataset, evaluation criteria and results. The experiment analysis of this study examined the accuracy of the proposed model and state-of-the-art models.

The implementation code of MSMIDM is written using Python 3.5.3, a programming language inside the Jupyter Notebook. The computational environment of the system is configured with a 320 GB hard disk, 4 GB main memory, Intel I3 processor, CPU M350@2.27 GHz, and Windows 7 operating system.

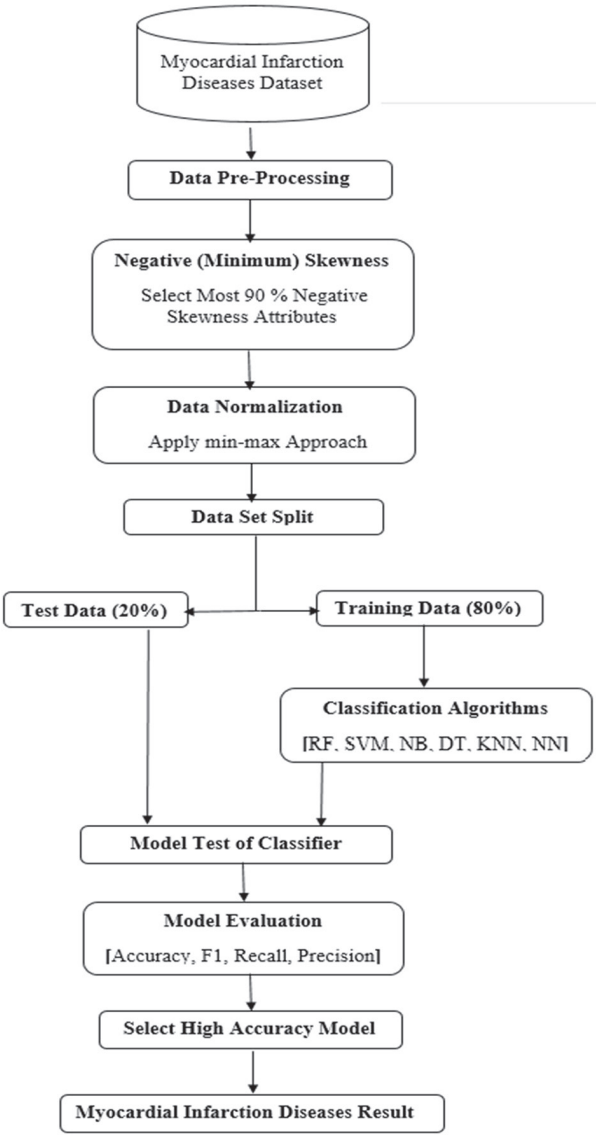


Figure 1. Flow Chart of the Proposed Minimum Skewness-based Myocardial Infarction Detection Model (MSMIDM)

4.1. EXPERIMENT DATASET AND ALGORITHM

This study utilizes the Cleveland Myocardial Infarction Diseases dataset from the UCI Machine Learning Repository. The used dataset consists of 300 samples with 14 dimensions, where 13 dimensions define predictors and 1 dimension describes the actual class. The details of this dataset are described in various studies, which is described in Table 1. The proposed Myocardial Infarction Detection Model (MSMIDM) was compared to various state-of-the-art models based on classification objectives. The description of the selected classification model is described in Table 1 references.

Algorithm 1: Minimum Skewness based Myocardial Infarction Detection Model (MSMIDM)

Input:

1. *MDS*={$x_1, x_2, \ldots x_n$} *MI Diseases Dataset with d dimensions*.

2. *CA = Set of Classification Algorithms* [*RF, SVM, NB, DT, KNN, NN*]

Output:

1. *Myocardial Infarction Diseases results basis of CA.*

Method:

1. *Perform data preprocessing tasks based on missing values, and so on.*
2. *Extract most of the 90 % negative skewness attributes for classification through Equation (1).*
3. *Perform the min-max normalization approach for data normalization.*
4. *Splitting the data into test and training data based on 80:20 ratios as per selection of negative skewness attributes.*
5. *Incorporate the CA classifier set into the training data.*
6. *Test the model based on the classifier results.*
7. *Evaluating the model based on classification validation approaches.*
8. *Select a highly accurate model for myocardial infarction disease analysis.*

Table 1. Description of Selected Algorithms for Accuracy Comparison

ID	Algorithms/Authors	References
A01	(Parthiban and Srivatsa, 2012)	[13]
A02	(Otoom et al., 2015)	[16]
A03	(Vembandasamy et al.,2015)	[17]
A04	(Verma et al.,2015)	[20]
A05	(Chadha and Mayank, 2016)	[21]
A06	(Seenivasagam and Chitra, 2016)	[7]
A07	(Dwivedi, 2018)	[23]
A08	(Ed-daoudy and Maalmi, 2019)	[25]
A09	(Anitha and Sridevi, 2019)	[26]
A10	Bagging (Latha and Jeeva, 2019)	[1]
A11	Boosting (Latha and Jeeva, 2019)	[1]
A12	(Shah et al., 2020)	[28]
A13	(Patro et al., 2021)	[33]
A14	(Reddya and G,2021)	[32]
A15	(Kondababu et al.,2021)	[34]

4.2. PERFORMANCE MEASURES

The performance metrics assess the perfection of algorithms and models. This study utilized accuracy, F1-measure, recall, precision, Jaccard score, Mutual Info Score (MIS), Adjusted Rand Score (ARS), V Measure Score (VMS), Mean Squared Error (MSE) and R2 Score (R2) performance metrics to evaluate the proposed model. The formulation of all employed measurements is described in respectively [20], [25], [38-41], [44], and [45].

4.3. RESULTS AND DISCUSSION

The performance of numerous classification methods is compared in Tables 2-4. Table 2 depicts the confusion matrix-based accuracy, F1 measure, recall and precision measurement of various classification algorithms employing the proposed model. The RF algorithm obtains the maximum accuracy of 90%, the F1 measure of 88.1%, the recall of 83.333%, and the precision of 78.72% in Table 2 compared to the state-of-the-art algorithms.

Table 3 compares the confusion matrix and statistically based ARS and VMS measurements for the proposed model. The RF algorithm achieves the highest ARS of 63.3% and VMS of 53.97% in Table 3 as compared to the state-of-the-art algorithm.

Table 2. Classification results of the MSMIDM-based Machine Learning(ML) approach using confusion matrix component

ML Algorithms	Accuracy (%)	F1 (%)	Recall (%)	Precision (%)
RF	90	88.101	83.333	78.727
SVM	87.037	84.444	79.167	73.077
NB	87.037	84.444	79.167	73.077
DT	83.238	79.475	79.167	65.977
KNN	81.481	77.273	70.833	62.963
NN	85.556	82.976	79.167	70.912

Table 3. Classification results of the MSMIDM-based Machine Learning(ML) approach using confusion matrix and statistical component

ML Algorithms	ARS (%)	VMS (%)
RF	63.305	53.971
SVM	54.027	44.973
NB	54.027	44.973
DT	39.538	31.437
KNN	38.524	31.241
NN	49.664	40.441

Table 4 compares the Jaccard score, MIS, MSE, and R2 classification results of the RF, SVM, NB, DT, KNN, and NN classification algorithms applying a partitional validation method to the proposed model. In this case, the RF method obtains the most robust Jaccard score of 78.73%, MIS of 36.36%, and R2 of 59.5% compared to the state-of-the-art approach. The SVM and NB achieve outstanding classification results within an MSE validation index. The MSE for RF results demonstrates that randomization influences classification outcomes over non-randomization-based classification techniques.

Tables 5-10 examine the comparative accuracy of the proposed model to that of state-of-the-art models. Table 5 reveals that the proposed MSMIDM-based RF algorithm delivers more robust accuracy results than the existing work on RF-based models. The MSMIDM-based RF classifier obtained 90% as compared to the state-of-the-art model. The MSMIDM improved accuracy by 3.16% when compared to the excellent A12 algorithm. The MSMIDM improves worst-case RF accuracy by 11.12% compared to the A11 model.

Table 4. Classification results of the MSMIDM-based Machine Learning(ML) approach based on partitional validation approach

ML Algorithms	Jaccard Score (%)	MIS (%)	MSE (%)	R2 (%)
RF	78.727	36.358	13.334	59.5
SVM	73.077	30.474	12.963	47.5
NB	73.077	30.474	12.963	47.5
DT	65.977	21.57	18.148	26.5
KNN	62.963	21.027	18.519	25
NN	70.912	27.523	14.444	41.5

Table 5. Accuracy assessment of the proposed MSMIDM and conventional models for RF algorithm

Algorithms	Accuracy (%)
A08	87.5
A10	80.53
A11	78.88
A12	86.84
A15	86.1
MSMIDM	**90**

Table 6 illustrates that the proposed MSMIDM-based SVM algorithm produces excellent classification accuracy results compared to other SVM-based models. The MSMIDM-based SVM classifier outperformed the most recently developed model by 87.037%. The MSMIDM improved the accuracy by 0.937% when compared to the superior A15 algorithm. The MSMIDM resolves the worst-case accuracy of SVM using the A14 model and improves the accuracy by 12.037% over the A14 model.

Table 6. Accuracy assessment of the proposed MSMIDM and conventional models for SVM algorithm

Algorithms	Accuracy (%)
A06	82.02
A02	84.50
A07	82
A08	85.82
A09	77.7
A13	80
A14	75
A15	86.1
MSMIDM	**87.037**

Table 7 shows that the proposed MSMIDM-based NB algorithm achieves reasonable accuracy when compared to earlier NB-based models. The MSMIDM-based NB classifier outperformed the state-of-the-art model by 87.037%. The MSMIDM improved the accuracy by 0.037% against the outstanding A13 algorithm. The MSMIDM improves the worst-case accuracy of NB by 13.037% as compared to the A01 model.

Table 8 shows that the proposed MSMIDM-based DT algorithm achieves superior classification results than former DT-based models in terms of accuracy validation. The MSMIDM-based DT classifier outperformed the state-of-the-art model by 83.238%, while MSMIDM improved accuracy by 0.438% over the effective A08 strategy. The MSMIDM resolves the worst-case accuracy of DT using the A11 model and improves accuracy by 7.338% with the A11 model.

Table 7. Accuracy assessment of the proposed MSMIDM and conventional models for NB algorithm

Algorithms	Accuracy (%)
A01	74
A02	84.50
A03	86.42
A05	85.86
A07	83
A09	86.6%
A10	84.16
A11	84.16
A13	86.7
A14	78
A15	75.8
MSMIDM	**87.037%**

Table 8. Accuracy assessment of the proposed MSMIDM and conventional models for DT algorithm

Algorithms	Accuracy (%)
A04	80.6
A07	77
A08	82.8
A10	79.87
A11	75.9
A12	80.263
MSMIDM	**83.238**

Table 9 shows that the proposed MSMIDM-based KNN algorithm has the highest accuracy outcomes when compared to other KNN-based models. The MSMIDM-based KNN classifier outperformed the state-of-the-art model by 81.481%. The MSMIDM boosted accuracy by 1.484% when compared to the most effective A07 and A13 algorithms. The MSMIDM resolves the worst-case accuracy of KNN using the A14 model and improves the accuracy by 11.481% over the A14 model.

Table 9. Accuracy assessment of the proposed MSMIDM and conventional models for KNN algorithm

Algorithms	Accuracy (%)
A07	80
A09	76.67
A13	80
A14	70
MSMIDM	**81.481**

Table 10 shows that the proposed MSMIDM-based NN algorithm produces better classification results than other NN-based models regarding accuracy validation. The MSMIDM-based NN classifier outperformed the most advanced model by 85.556%. The MSMIDM im-

proved accuracy by 1.553% when compared to the A07 algorithm. The MSMIDM improves the worst-case accuracy of NN by 3.016% when compared to the A11 model.

Table 10. Accuracy assessment of the proposed MSMIDM and conventional models for NN algorithm

Algorithms	Accuracy (%)
A07	84
A10	81.52
A11	79.54
A13	80
A14	82
MSMIDM	**85.556**

Figs. 2–7 depict a comparison of conventional and proposed model-based algorithms utilizing accuracy, F1-measure, recall, precision, Jaccard score, MIS, ARS, VMS, MSE, and R2 classification metrics.

Fig. 2 illustrates that the MSMIDM-based RF algorithm maximizes the confusion, statistical, and partitional-based validation metrics compared to the standard RF method. The MSMIDM-based RF classifier improves the 1.67 % accuracy, 1.69 % F1-measure, 2.64 % precision, 2.64% Jaccard score, 3.46% MIS, 5.24% ARS, 5.74% VMS, and 6.75% R2 validation metrics as compared to the classical RF classifier.

Fig. 3 shows that the proposed MSMIDM-based SVM method outperforms classical SVM regarding all considered validation approaches. The MSMIDM-based SVM enhanced the 5.56 % accuracy, 5.28 % F1-measure, 0.0 % recall, 7.56 % precision, 7.56 % Jaccard score, 9.55 % MIS, 15.52 % ARS, 14.52 % VMS, and 22.5 % R2 assessment as compared to the classical SVM classifier.

Fig. 4 depicts that the MSMIDM-based NB algorithm outperforms the classical NB approach based on classification results. The MSMIDM-based NB outperformed the standard NB classifier in terms of 7.41% accuracy, 8.89% F1-measure, 8.33% recall, 12.36% precision, 12.36% Jaccard score, 12.29% MIS, 20.12% ARS, 18.14% VMS, and 30.0% R2 authentication metric.

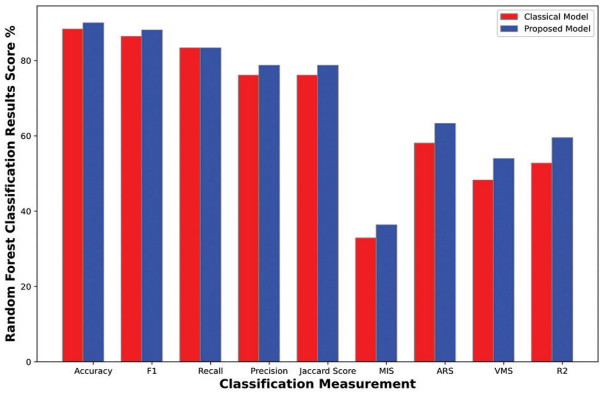


Fig. 2. Classification results assessment between classical-RF and proposed MSMIDM based RF algorithm

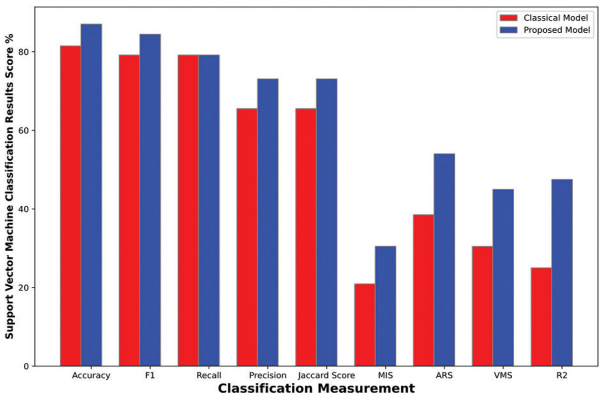


Fig. 3. Classification results assessment between classical-SVM and proposed MSMIDM based SVM algorithm

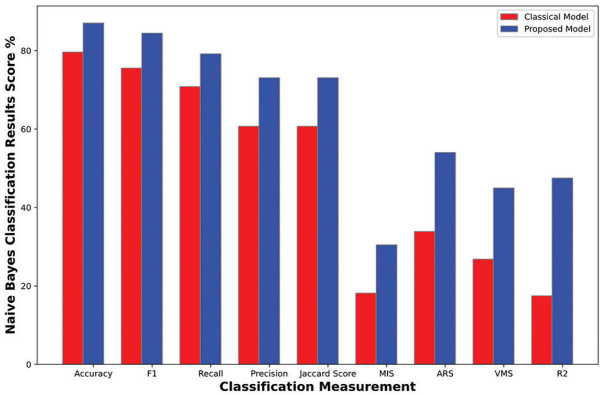


Fig. 4. Classification results assessment between classical-NB and proposed MSMIDM based NB algorithm

Fig. 5 shows that the proposed MSMIDM-based DT algorithm maximizes the confusion, statistical, and partitional-related classification results over the classical DT algorithm. The MSMIDM-based DT boosted the 2.87 % accuracy, 1.57 % F1-measure, 1.25 % recall, 2.16 % precision, 2.16 % Jaccard score, 2.16 % MIS, 3.8% ARS, 3.16 % VMS, and 6.0% R2 validation parameter as compared to the classical DT classifier.

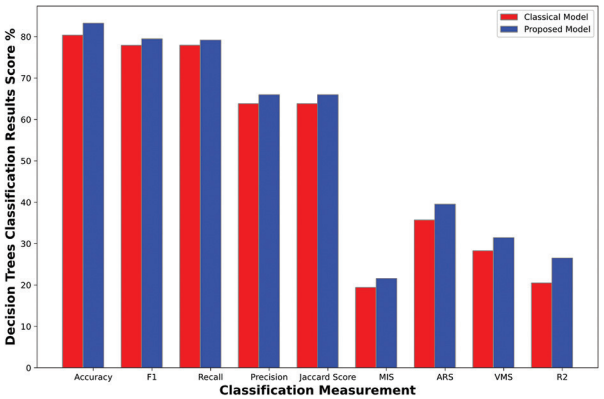


Fig. 5. Classification results assessment between classical-DT and proposed MSMIDM based DT algorithm

Fig. 6 illustrates that the MSMIDM-based KNN algorithm achieves excellent classification results compared to the classical KNN algorithm. The MSMIDM-based KNN enhanced 16.67% accuracy, 21.46% F1-measure, 65.83% recall, 24.25% precision, 24.25% Jaccard score, 17.16% MIS, 31.43% ARS, 25.45% VMS, and 20.5% R2 validation process in comparison to the classical KNN classifier.

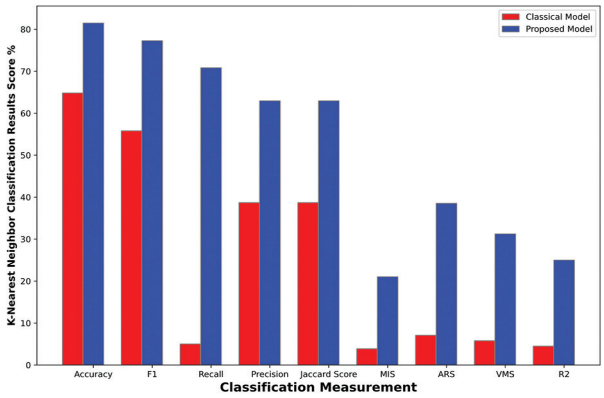


Fig. 6. Classification results assessment between classical-KNN and proposed MSMIDM based KNN algorithm

Fig. 7 demonstrates that the proposed MSMIDM-based NN algorithm obtains more robust classification results as compared to classical NN based on considered validation metrics. The MSMIDM-based NN outperformed the standard NN classifier in terms of 6.67% accuracy, 5.62% F1-measure, 0.0% recall, 7.64% precision, 7.64% Jaccard score, 6.9% MIS, 15.71% ARS, 9.58% VMS, and 27.0% R2 validation metrics.

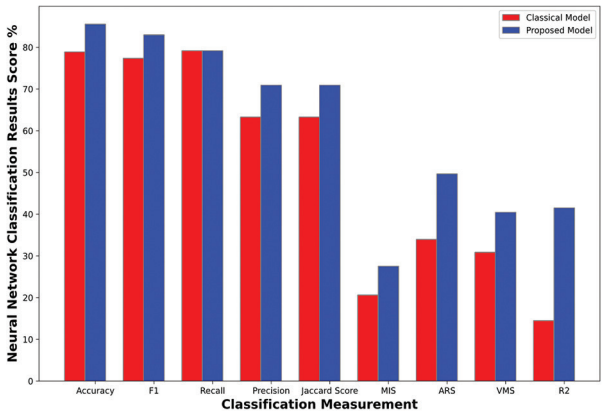


Fig. 7. Classification results assessment between classical-NN and proposed MSMIDM based NN algorithm

5. CONCLUSION

This study addresses and analyzes classification algorithms for identifying myocardial infarction heart disease. The detection of diseases is one of the tasks of the machine-learning algorithm. The classification approach extracts the hidden risk and disease-related information from the data set based on the attributes. The first section of this study describes the challenges and issues with the machine learning algorithm for detecting myocardial infarction. The second section examines various classification model literature and identifies statistical and feature selection-based approaches to improve the classification model accuracy. The selection of attributes plays a major role in disease detection. Therefore, this study uses the statistical and feature selection-based minimum skewness approach and proposes a Minimum Skewness-based Myocardial Infarction Detection Model (MSMIDM). The proposed algorithm selects the attributes based on the minimum skewness value and performs classification tasks for disease prediction through RF, NB, SVM, DT, KNN and NN classification algorithms. The experimental analysis indicates that the proposed MSMID model improves the prediction accuracy of classification algorithms. In comparison to the RF, SVM, NB, DT, KNN, and NN classification algorithms, the proposed MSMID model achieves 90%, 87.037%, 83.238%, 81.481%, and 85.556% accuracy. According to the classification results, the MSMIDM-based RF algorithm is excellent for detecting myocardial infarction and other heart diseases. Future research will use a statistical and feature selection-based approach to ensemble classification.

6. REFERENCES

[1] C. B. C. Latha, S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques", Informatics in Medicine Unlocked, Vol. 16, 2019.

[2] K. Uyar, A. İlhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks", Procedia Computer Science, Vol. 120, 2017, pp. 588–593.

[3] S. B. Patil, Y. S. Kumaraswamy, "Extraction of Significant Patterns from Heart Disease Warehouses for Heart Attack Prediction", International Journal of Computer Science and Network Security, Vol. 9, No. 2, 2009, p. 228.

[4] B. Fida, M. Nazir, N. Naveed, S. Akram, "Heart disease classification ensemble optimization using Genetic algorithm", Proceedings of the EEE 14th International Multitopic Conference, Karachi, Pakistan, 22-24 December 2011, pp. 19-24.

[5] A. Rairikar, V. Kulkarni, V. Sabale, H. Kale, A. Lamgunde, "Heart disease prediction using data mining techniques", Proceedings of the International Conference on Intelligent Computing and Control, Coimbatore, India, 23-24 June 2017, pp. 1-8.

[6] K. Thygesen et al. "Fourth Universal Definition of Myocardial Infarction (2018)", Circulation, Vol. 138, No. 20, 2018.

[7] V. Seenivasagam, R. Chitra, "Myocardial Infarction Detection Using Intelligent Algorithms", Neural Network World, Vol. 26, No. 1, 2016, pp. 91-110.

[8] A. Khazaee, "Heart Beat Classification Using Particle Swarm Optimization", International Journal of Intelligent Systems and Applications, Vol. 5, No. 6, 2013, pp. 25-33.

[9] D. K. Singh, V. Swaroop, "Data security and privacy in data mining: Research issues & preparation", International Journal of Computer Trends and Technology, Vol. 4, No. 2, 2013, pp. 194–200.

[10] S. M. Gorade, P. A. Deo, "A Study Some Data Mining Classification Techniques", International Journal of Modern Engineering Research, Vol. 4, No. 1, 2017, pp. 210-215.

[11] Y. Sun, J. Han, "Mining Heterogeneous Information Networks", Springer International Publishing, 2012.

[12] G. Parthiban, S. K. Srivatsa, "Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients", International Journal of Applied Information Systems, Vol. 3, No. 7, 2012, pp. 25-30.

[13] L. Sun, Y. Lu, K. Yang, S. Li, "ECG analysis using multiple instance learning for myocardial infarction detection", IEEE Transactions on Biomedical Engineering, Vol. 59, No. 12, 2012, pp. 3348-3356.

[14] R. Safdari, M. Ghazisaeedi, G. Arji, M. Gharooni, M. Soraki, M. Nasiri, "A model for predicting myocardial infarction using data mining techniques", Iranian Journal of Medical Informatics, Vol. 2, No. 4, 2013.

[15] A. F. Otoom, E. E. Abdallah, Y. Kilani, A. Kefaye, M. Ashour, "Effective diagnosis and monitoring of heart disease", International Journal of Software Engineering, Technology and Applications, Vol. 9, No. 1, 2015, pp. 143-156.

[16] K. Vembandasamy, R. Sasipriya, E. Deepa, "Heart Diseases Detection Using Naive Bayes Algorithm", International Journal of Innovative Science Engineering and Technology, Vol. 2, No. 9, 2015, pp. 441-444.

[17] A. Daraei, H. Hamidi, "An efficient predictive model for myocardial infarction using cost-sensitive J48 model", Iranian Journal of Public Health, Vol. 46, No. 5, 2017, pp. 682-692.

[18] L. Verma, S. Srivastava, P. C. Negi, "A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data", Journal of Medical Systems, Vol. 40, No. 7, 2016.

[19] R. Chadha, S. Mayank, "Prediction of heart disease using data mining techniques", CSI Transactions on ICT, Vol. 4, No. 2-4, 2016, pp. 193-198.

[20] A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease", Neural Computing and Applications, Vol. 29, No. 10, 2018, pp. 685-693.

[21] S. A. Mokeddem, "A fuzzy classification model for myocardial infarction risk assessment", Applied Intelligence, Vol. 48, No. 5, 2018, pp. 1233-1250.

[22] A. Ed-Daoudy, K. Maalmi, "Performance evaluation of machine learning based big data processing framework for prediction of heart disease", Proceedings of the International Conference on Intelligent Systems and Advanced Computing Sciences, Taza, Morocco, 26-27 December 2019, pp. 1-5.

[23] S. Anitha, N. Sridevi, "Heart Disease Prediction Using Data Mining Techniques Using Data Mining Techniques", Journal of Analysis and Computation, Vol. 13, No. 2, 2019, pp. 48-55.

[24] A. K. Gárate-Escamila, A. H. El Hassani, E. Andrès, "Classification models for heart disease prediction using feature selection and PCA", Informatics in Medicine Unlocked, Vol. 19, 2020.

[25] D. Shah, S. Patel, S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques", SN Computer Science, Vol. 1, No. 6, 2020, p. 345.

[26] D. Mandair, P. Tiwari, S. Simon, K. L. Colborn, M. A. Rosenberg, "Prediction of incident myocardial infarction using machine learning applied to harmonized electronic health record data", BMC Medical Informatics and Decision Making, Vol. 20, No. 1, 2020, pp. 1-10.

[27] L. Ibrahim, M. Mesinovic, K. W. Yang, M. A. Eid, "Explainable Prediction of Acute Myocardial Infarc-

tion using Machine Learning and Shapley Values", IEEE Access, Vol. 8, 2020, pp. 210410-210417.

[28] V. N. Reddy et al. "Myocardial Infarction Prediction Using Hybrid Machine Learning Techniques", Turkish Journal of Computer and Mathematics Education, Vol. 12, No. 3, 2021, pp. 4251-4260.

[29] S. P. Patro, G. S. Nayak, N. Padhy, "Heart disease prediction by using novel optimization algorithm: A supervised learning prospective", Informatics in Medicine Unlocked, Vol. 26, 2021.

[30] A. Kondababu, V. Siddhartha, B. B. Kumar, B. Penumutchi, "A comparative study on machine learning based heart disease prediction", Materials Today: Proceedings, 2021, pp. 1-5.

[31] R. Liu et al. "An artificial intelligence-based risk prediction model of myocardial infarction", BMC Bioinformatics, Vol. 23, No. 1, 2022, pp. 1-17.

[32] U. Nagavelli, D. Samanta, P. Chakraborty, "Machine Learning Technology-Based Heart Disease Detection Models", Journal of Healthcare Engineering, Vol. 2022, 2022.

[33] P. A. Sadiyamole, S. M. Priya, "A Robust Cardiovascular Disease Predictor Based on Genetic Feature Selection and Ensemble Learning Classification", International Journal of Electrical and Computer Engineering Systems, Vol. 14, No. 7, 2023, pp. 799-809.

[34] A. A. Ahmad, H. Polat, "Prediction of Heart Disease Based on Machine Learning Using Jellyfish Optimization Algorithm", Diagnostics, Vol. 13, No. 14, 2023.

[35] C. M. Bhatt, P. Patel, T. Ghetia, P. L. Mazzeo, "Effective Heart Disease Prediction Using Machine Learning Techniques", Algorithms, Vol. 16, No. 2, 2023, p. 88.

[36] S. Subramani et al. "Cardiovascular diseases prediction by machine learning incorporation with deep learning", Frontiers in Medicine, Vol. 10, No. April, 2023, pp. 1-9.

[37] H. A. Al-Alshaikh et al. "Comprehensive evaluation and performance analysis of machine learning in heart disease prediction", Scientific Reports, Vol. 14, No. 1, 2024, pp. 1-15.

[38] X. Zhao, J. Liang, C. Dang, "A stratified sampling based clustering algorithm for large-scale data", Knowledge-Based Systems, Vol. 163, 2019, pp. 416-428.

[39] S. I. Ayon, M. M. Islam, M. R. Hossain, "Coronary Artery Heart Disease Prediction: A Comparative Study of Computational Intelligence Techniques", IETE Journal of Research, Vol. 68, No. 4, 2022, pp. 2488-2507.

[40] S. Mohan, C. Thirumalai, G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access, Vol. 7, 2019, pp. 81542-81554.

[41] R. Katarya, S. K. Meena, "Machine Learning Techniques for Heart Disease Prediction: A Comparative Study and Analysis", Health and Technology, Vol. 11, No. 1, 2021, pp. 87-97.

[42] J. P. Verma, "Normal Distribution and Its Application", Statistics and Research Methods in Psychology with Excel, Springer, 2019, pp. 201-235.

[43] S. K. Yadav, S. Singh, R. Gupta, "Measures of Dispersion", Biomedical Statistics, Springer, 2019, pp. 59-70.

[44] K. K. Pandey, D. Shukla, "Min–max kurtosis mean distance based k-means initial centroid initialization method for big genomic data clustering", Evolutionary Intelligence, Vol. 16, 2022, pp. 1055-1076.

[45] K. K. Pandey, D. Shukla, "Maxmin distance sort heuristic-based initial centroid method of partitional clustering for big data mining", Pattern Analysis and Applications, Vol. 25, No. 1, 2022, pp. 139-156.

Designing a Prototype for Smartphone Footage Stabilization

Original Scientific Paper

Bharathi R.B.*
Department of Electrical & Electronics Engineering,
Manipal Institute of Technology, Manipal Academy of Higher Education
Manipal, Karnataka, India
bharathi.rb@manipal.edu

VijayaLaxmi
Department of Electrical & Electronics Engineering,
Manipal Institute of Technology, Manipal Academy of Higher Education
Manipal, Karnataka, India
vijaya.laxmi@manipal.edu

Shashank Bhat
Department of Data Science and Engineering
Manipal Institute of Technology, Manipal Academy of Higher Education
Manipal, Karnataka, India
shashank2.mitmpl2023@learner.manipal.edu

*Corresponding author

Abstract *– Hardware stabilizers are on high demand due to the increasing use of media capture devices such as smartphones and portable cameras in market. Hence an effort is made to develop a stabilizer prototype with Inertial measurement Unit (MPU 6050 sensor), arduino Uno microcontroller and servo motor. Using existing libraries of arduino, a code is developed to read the position of the smartphone by the sensor and is fed to arduino microcontroller. The microcontroller generates an appropriate PWM signal based on the output of sensor and sends it to two servo motors to nullify the effect of jitter produced by the movement of the smartphone along roll and pitch axes respectively. Stabilized footage is produced by smartphones compensating the movement along roll and pitch axes respectively. The cost of this developed prototype is very low compared to existing gimbal systems in the market.*

Keywords: *Hardware stabilizer, Inertial measurement Unit, servo motor*

Received: March 2, 2024; Received in revised form: May 2, 2024; Accepted: May 27, 2024

1. INTRODUCTION

Video stabilization is a method used to minimize blurring caused by camera movement during video recording [1]. Shaky videos have always been a problem for professional photographers. The rapid advancement of smartphone technology is closely linked to the development of photography and image processing. The camera is among the most used features. Poor quality results are produced by shocks and hand movements. A stabilizer for camera position is required to enhance the quality of camera images and video [14]. Using a 2- or 3-axis (roll, tilt, and yaw) gimbal system to separate the rotation of the UAV platform from the camera is the most popular method for video stabilization [2].

Full-frame films can be created using the technique suggested by Matsushita et. Al. [3], which works by locally matching the picture data of nearby frames to automatically fill in any missing image segments. Motion inpainting is suggested as a means of enforcing temporal and spatial consistency of the completion in both dynamic and static image sections.

Over the past 20 years, research on video stabilization (VS) has been ongoing. Guilluy et. Al. [4] focused particularly on the Video Stabilization Quality Assessment (VSQA) and introduced a new methodology that was influenced by the findings of research on Image Quality Assessment (IQA) in general. When image stitching techniques are directly applied to shake films, significant temporal and spatial distortions are

frequently present [5]. To address this issue, Guo et. Al. [5] provides a unified framework that combines video stabilization and stitching.

There have been three phases of video stabilization: digital, optical, and mechanical [6]. By using an invisible tripod built into the camera to correct the recorded video and produce stabilized videos, a device called a gyroscopic stabilizer is used in mechanical stabilization. The use of optical stabilization is used in both high-end SLR cameras like Nikon and Canon and smartphones like Samsung, Apple, and others. Optical stabilization counteracts image vibration and achieves image stabilization by moving the lens group on a plane perpendicular to the optical axis. Without the need for hardware, digital stabilization directly determines the motion vectors of the subsequent frames. Motion correction and filtering are used to isolate and eliminate the shaky component. Ultimately, the picture warp reconstructs frames.

To increase the estimation efficiency, the suggested approach [7] employs differential global motion estimation with Taylor series expansion. The inter-frame error between successive frames is defined by the Affine Motion Model. Through the analytical solution of the derivatives of the inter-frame error, motion vectors have been computed. Gaussian kernel filtering has been used to smooth out computed motion parameters following motion estimation. To eliminate the rotation effect from the smoothed transformation chain, inverse rotation smoothing has been used. As a result, the accumulation error has been decreased by [7] and the missing image area has been much reduced.

To address the issue of video stabilization, Liu et. Al. [8] suggests using an additional depth sensor, such as the Kinect camera. Even with its low resolution, noise, and incompleteness, the depth image helps with frame warping and camera motion estimation, making video stabilization a much better-posed problem.

A unique method for stabilizing videos is suggested by Yang et. Al. [9] and is based on the particle filtering framework. Yang et. Al. [9] added tracking of the projected affine model of the camera motions to the conventional application of particle filters in object tracking. Through simulated experiments, the algorithm's improved performance is shown by the author.

Image stabilization, whether software or hardware based, can significantly enhance images, particularly in hand-held photography, slower shutter speed shooting, and other unsteady conditions. Compared to SLR lenses smart phone cameras are far slower. They leave the shutter open for an extended period, which blurs any movement [10]. There are many techniques available in the present-day market that resolve this problem. They are Optical steady shot (OSS) by Sony, Vibration Reduction (VR) by Nikon and Image stabilizer (IS) by Canon [11]. Stabilization of footage is acquiring a greater significance day by day, and the form factor of the cameras that are used by the users is also rapidly decreasing [12].

Stabilizing footage either through specialized hardware or by software is now becoming an industrial norm for use either in personal vlogs or in commercial videography applications. With the increasing penetration of media capture devices such as mobile phones and portable cameras, the market for stabilizers in small appliances is increasing day by day [13]. Osmo mobile, manufactured by DJI is one of the solutions for the same, currently costing nearly 130 dollars. Thus, the primary goal of the research is to reduce the cost of the prototype.

Hence, a hardware prototype is designed that can be used for video stabilization which builds upon the existing technologies that are already present in the camera.The smartphone video stabilizer has Arduino microcontroller, gyro sensor MPU-6050 [14] and 2 servo motors for stabilization along two axes roll and pitch respectively. The change in angular velocity or displacement of the phone is detected by the MPU-6050 sensor and this data will be transmitted to the microcontroller. The servo motor then rotates the base on which the smartphone is placed by the angle detected by the gyro sensor, hence, compensating for the motion. Therefore, resulting in a stabilized footage. A servo motor is used to stabilize the smartphone for smoother footage. Fig. 1 summarizes the block diagram of the proposed solution.

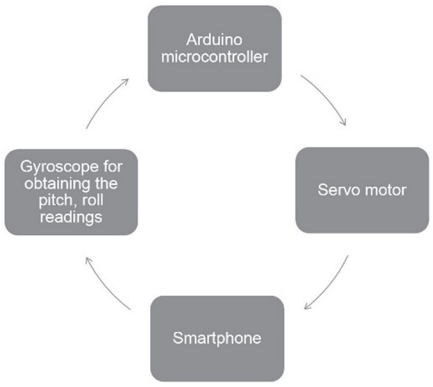


Fig. 1. Block Diagram of the proposed solution

2. METHODOLOGY

The algorithm of the working of the prototype is summarized below:

Step 1: Wake up the MPU 6050 and initiate communication via I2C protocol using address 0x68.

Step 2: Assign the servo motors to pins 9,10 (PWM pins) of the Arduino controller. Initialize the Pulse width values corresponding to 0 and 180 degrees as obtained from calibration to separate variables.

Step 3: Send a command to the Arduino to obtain the first values from the accelerometer and the gyroscope.

Step 4: Calculate pitch and roll and assign them as initial values to the complementary filter.

Step 5: Wait for an interrupt and hence obtain the raw values from the sensor again. Calculate values of roll and pitch.

Step 6: Update the complementary filter with the new values of roll and pitch.

Step 7: Use this filtered value and map it to obtain the value of pulse width based on the values obtained from the calibration code.

Step 8: Send the appropriate pulse width to the motors to compensate for the motion caused by the jitter.

A brief description of the components used in the development of the prototype is below:

2.1. MPU 6050 SENSOR

To stabilize the platform on which the smartphone was placed, an Inertial Measurement Unit (IMU) was required to identify the degree to which the setup deviated from its initial position [13]. The MPU-6050 is an advanced chip that has two on-board sensors, a gyroscope and an accelerometer, each with its advantages [14]. The individual components of the IMU (MPU 6050) are extremely susceptible to noise. The gyroscope component is accurate for short duration of time but tends to drift after time (because it is integrating), while the accelerometer is sensitive to forces and gives inaccurate reading for short duration of time. The gyroscope gives a high quality signal in a short term and the accelerometer is good for measuring changes in speed. The complementary filter is a powerful and simple tool to improve the shortcomings of MPU 6050 sensor. Combining the accelerometer and gyroscope together gives the basic complementary filter. A complementary filter combines a high-pass and a low-pass filter. In this case, a high-pass filter is used on the gyroscope readings, and a low-pass filter is applied over the accelerometer readings. A complementary filter eliminated the gyroscope drift and reduced the accelerometer sensitivity. Thus, the output becomes more robust against noise factors [15]. The coefficients of the high-pass and low-pass filters add up-to one, and hence filter is said to as complimentary. To appropriately calibrate the filter, the coefficients of filter are changed by trial-and-error method. The high-pass filter design with a coefficient of 0.98 and low-pass filter with a coefficient of 0.02 is found to be appropriate for the given application. Therefore, 98% of the values for each update come from the gyroscope, and 2% come from the accelerometer.

The MPU-6050 is tested at the initial stage using an external I2C library developed by Jeff Rowberg and has been interfaced by a self-authored code using a complementary filter. The library had in-built calls for obtaining values from the MPU-6050 and made use of the on-board DMP (data management platform) to do the same. The values from the accelerometer, gyroscope stored individually on the MPU 6050 on different registers. Arduino is open source and is used for a large variety of projects. The filtered values from sensor are read by Arduino using the I2C communication protocol. The I2C communication address used by the sensor is a 7-bit address 0x68. The sensor supports a communication speed of 400Khz. Every time the data is ready, the sensor makes the interrupt (INT) pin high. MPU-6050 sensors are placed in the unstabilized portion of the setup, and the degree of instability, as communicated by the sensor, is then sent to the Arduino controller after filtering to control the servo motor accordingly [14]. The Sultana et. Al. [16] uses the MPU 6050 with a complementary filter for the stabilization of the video. The performance of various wearable cameras in the surgical setting is carried out and has concluded that IMU based sensors are appropriate in the operating room to objectively quantify camera motion [17]. The results are validated by comparing the positional sensing of the device to a geared tripod head that allows for fine, measured manipulations of the sensor in three orthogonal axes.

2.2. SERVO MOTOR

The servo motor is connected to the output of the microcontroller and has metal gears with a torque of 15.5kg cm to 17 kg cm. It has a built-in gearbox, position feedback mechanism, and motor controller. The servo motor can be controlled to move to any position by using PWM using in-built resistive feedback. This motor has a three-wire interface, one for control and two for power supply. The weight of the motor is 77 grams and has an operating voltage of 4.8V-6V. The biggest advantage of using a servo motor is the presence of robust pre-existing library functions for controlling the servo motor [18]. The servo motor is also extremely lightweight and cost-effective.

For controlling the position of the servo motor, a PWM signal is to be sent to it. The width of this PWM signal determines the angle at which the shaft of the servo motor moves. Based on the manufacturer, the width of the PWM signal that is required to move the shaft from 0 to 180 degrees varies. To calibrate the servo motor to move exactly by 180 degrees, a pre-existing Arduino library for controlling servo motors was used to send a PWM signal of a width starting from an extremely low value of 100 microseconds. At 540 microseconds, it was noticed that the shaft of the motor started to move. This was marked as the pulse width corresponding to 0 degrees. A paper was placed below the servo motor and this position was marked. Then at 2140 microseconds pulse width, it was noticed that the shaft moved by exactly 180 degrees. A mapping function was then used to linearly map 540 to 2140 microseconds pulse width to -90 to +90 degrees, to match the output obtained from the MPU-6050. The motors

drew approximately 400mA of current and a peak current of 4A. The use of servo motors, on account of their easier PWM-based position control, originated directly from the Arduino website itself [19].

2.3. ARDUINO UNO BOARD

An Arduino Uno was used in the development of the prototype. The I2C communication protocol was used to communicate with the MPU-6050 sensor, while two PWM pins of the Arduino Uno microcontroller were used to communicate with two motors. The software implementation was performed in the Arduino IDE.

2.4. OVERALL CIRCUIT

Connection of MPU 6050 with arduino Uno:

- Vcc to +5V
- Gnd to Gnd of Arduino
- SDA to Analog pin A4 (I2C communication)
- SCL to Analog pin A5 (I2C communication)
- INT to Digital pin 2

Connections of Servo motors with Arduino Uno:

- Red wire to positive terminal of external supply
- Black wire to negative terminal of external supply
- Yellow wire of motor B to Digital pin 9 of Arduino (PWM)
- Yellow wire of motor A to Digital pin 10 of Arduino (PWM)
- GND of Arduino to negative terminal of external supply

The overall circuit connection is shown in Fig. 2.

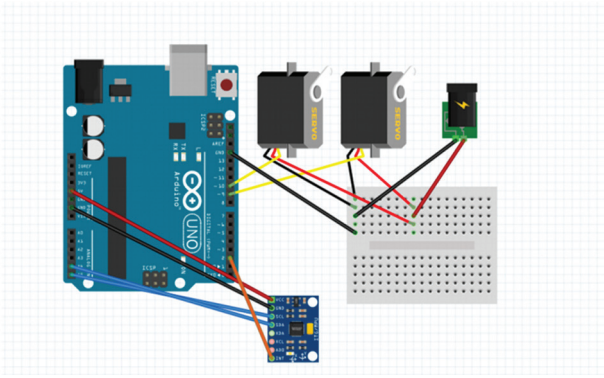


Fig. 2. Overall Circuit Diagram

The implementation of the chassis and the software is discussed below:

2.5. CHASSIS DESIGN

The servo motor (A), which is stabilizing along the X axis. The IMU is guaranteed to be on the unstabilized section of the chassis because the user is holding this motor. On one side of motor A, the servo motor (B) for Y-axis stabilization is attached, and on the other, the rotating shaft of this motor is connected to the platform holding the smartphone. A servo horn drilled into a wooden block has been used to mount Motor (A) on Motor (B), which is subsequently attached onto Motor (A) using epoxy resin. Additionally, Motor (B) is attached to a servo horn that is drilled into a wooden block that is positioned 90 degrees from the initial block. The revolving shaft of this motor is attached to the platform holding the smartphone. To provide smartphone stabilization, motors (A) and (B) work in tandem.

The chassis design of the prototype is indicated in Fig. 3.

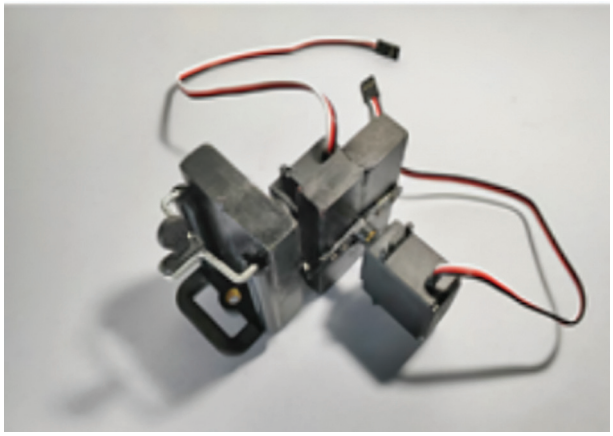


Fig. 3. Chassis design of the proposed solution

3. RESULTS AND DISCUSSION

The prototype designed will compensate for the motion along pitch and roll axes by employing the sensor along with two servo motors. The Gyroscope +accelerometer sensor is interfaced with the Arduino microcontroller. After calibrating the servo motor, they are placed on the chassis as described in Fig. 3. Upon running the code, the IMU sensor continuously monitors for any change in angle. If there is a change in the angular position of the motor, caused by a movement in the user's hand or by jitter, this change is filtered using the complementary filter sent to the Arduino microcontroller for better stabilization. If unfiltered readings are sent to the motor, the motor shall not be able to stabilize as efficiently as with the filtered readings. The change in angle is mapped to the corresponding value of pulse width by the Arduino microcontroller. This pulse width modulation signal is fed to the servo motor and the motor then moves to the corresponding change in angle. This process is carried on iteratively. Note that this is an open-loop control system.

A graph of unfiltered values of Roll & pitch vs filtered values of Roll & pitch are explained in plotted in Fig. 4 and Fig. 5 respectively. The illustrated values include the angle in pitch (Y-axis) and roll (X-axis), both of which are required to achieve stabilization. Fig. 4 and Fig. 5 denote the results in the self-authored code, with the X-axis illustrating time and the Y-axis illustrating angle in degrees along the pitch and roll direction respectively. The MPU-6050 sensor is moved randomly

in space, and the resulting values are illustrated on the graph. The blue curves denote the unfiltered values while the red curves denote the filtered values.

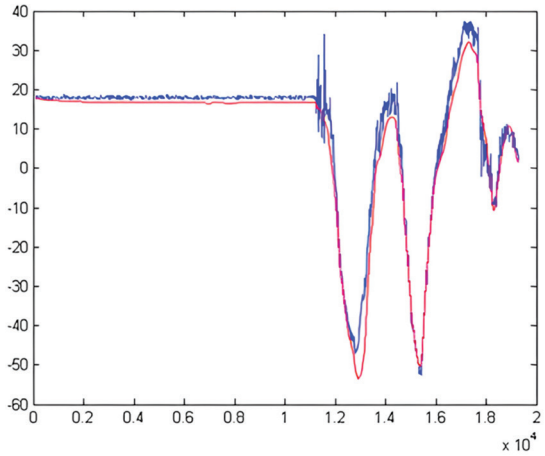


Fig. 4. Self-authored code with the obtained values for pitch (Y-axis)

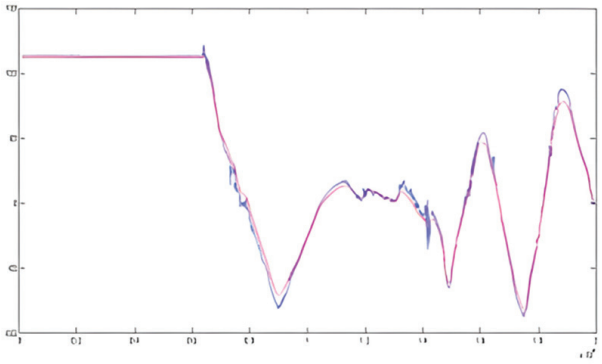


Fig. 5. Self-authored code with the obtained values for roll (X-axis)

To demonstrate the stabilization, the setup is moved first along the X axis (roll) and the resulting stabilization is shown in Fig 6. The motion of the setup is caused along Y-axis (pitch) is shown in Fig. 7. A snapshot from a video showing the stabilization along both axes is shown in Fig. 6 and Fig. 7 respectively. Hence, the movement of the person holding the chassis shall not matter, the platform housing the smartphone will remain stable.

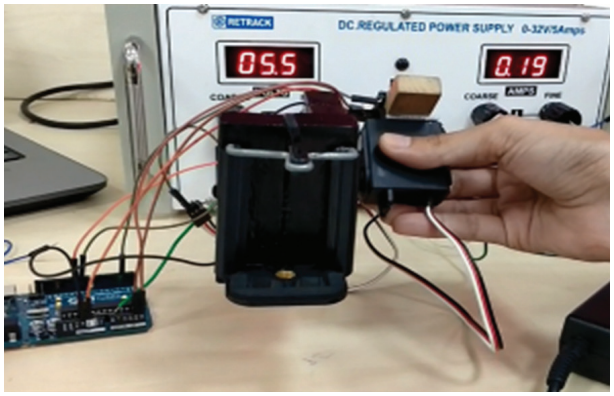


Fig. 6. Stabilization along X-axis (roll)

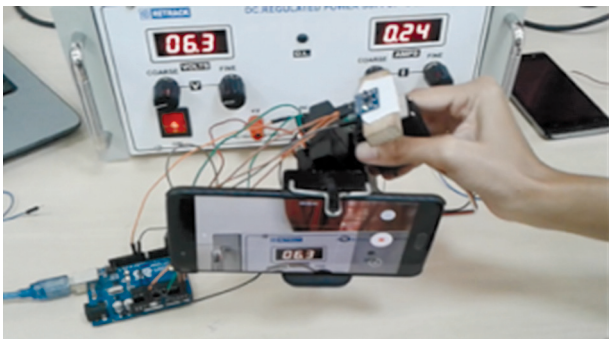


Fig. 7. Stabilization along Y-axis (pitch)

To illustrate the movement of the motor upon sending the required final angle through the Arduino code, an app developed by Google called 'Science Journal' was used as indicated in Fig. 8. This app used the inbuilt compass of a smartphone to graphically illustrate the current angle as read from the phone's compass. Further phone is attached to the top of the servo motor to obtain these readings. Thus, the command was given to the motor to move by different angles, and the readings obtained are shown in Fig. 8.

Fig. 8. Motor angle with respect to time, as obtained using the compass sensor of a smartphone, using the science journal app by Google

The angle detected by the compass sensor of a smartphone is around 16° Fig. 8 (part A) and the command was given to the motor to provide compensation is 15° as indicated in Fig. 8 (part B). Hence the error found to be around 1°.

Rafiq [14] and others develop a gimbal smartphone, which is made simple and less money consuming by utilizing microcontroller and MPU 6050 sensor. This MPU 6050 sensor is optimized to detect sway in axis X, Y, and Z or roll, pitch, and yaw. Gyroscope and accelerometer provide input to the microcontroller, which will process output on 3 servomotors that function to maintain the camera's position at a specified set point. The results show that MPU 6050 sensor can response angle reading error of 1.34° of roll, 0.25° of pitch, and 0.78° of yaw. Error in maximum servomotor movement is 1.5° [14]. The performance of the designed prototype cannot be compared with the reference [14] in terms of performance as the present design works on;y roll and pitch axes.

4. CONCLUSIONS

Smartphone are embedded with software stabilizer (like OSS and VR), but hardware stabilizers are more accurate specially for photographic and good quality video recording purposes. In this work,a prototype is developed to compensate the jitter in shaky videos along roll and pitch axes. The values from a gyroscope and an accelerometer of MPU-6050 are sent to complementary filter. The arduino receives the filtered values and this will generate a PWM signal. The width of this PWM signal determines the angle at which the shaft of the servo motor moves. Two PWM signals are used to control the position of motor along 2 axes. The prototype explains a new method with less cost to improve the quality of image from smartphone. The error will be introduced by the sensor and DC motor operation. The proposed solution can be scaled up to stabilize along the third axis yaw (Z axis). Stabilization along all three axes will ensure better and good quality footage. The total cost of the developed prototype including the component cost is around Rs. 3630.00.

5. REFERENCES:

[1] F. Vella, A. Castorina, M. Mancuso, G. Messina, "Robust digital image stabilization algorithm using block motion vectors", Proceedings of the International Conference on Consumer Electronics, Los Angeles, CA, USA, 18-20 June 2002, pp. 234-235.

[2] M. Odelga, N. Kochanek, H. H. Bülthoff, "Efficient real-time video stabilization for UAVs using only IMU data", Proceedings of the Workshop on Research, Education and Development of Unmanned Aerial Systems, Sweden, 3-5 October 2017, pp. 210-215.

[3] Y. Matsushita, E. Ofek, X. Tang, H.-Y. Shum, "Full-frame video stabilization", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20-25 June 2021, pp. 50-57.

[4] W. Guilluy, L. Oudre, A. Beghdadi, "Video stabilization: Overview, challenges and perspectives", Signal Processing: Image Communication, Vol. 90, 2021.

[5] H. Guo, S. Liu, T. He, S. Zhu, B. Zeng, M. Gabbouj, "Joint video stitching and stabilization from moving cameras", IEEE Transactions on Image Processing, Vol. 25, No. 11, 2016, pp. 5491-5503.

[6] Y. Wang, Q. Huang, C. Jiang, J. Liu, M. Shang, Z. Miao, "Video stabilization: A comprehensive survey", Neurocomputing, Vol. 516, 2023, pp. 205-230.

[7] H.-C. Chang, S.-H. Lai, K.-R. Lu, "A robust and efficient video stabilization algorithm", Proceedings of the IEEE International Conference on Multimedia and Expo, Taipei, 27-30 June 2004, pp. 29-32.

[8] S. Liu, Y. Wang, L. Yuan, J. Bu, P. Tan, J. Sun, "Video stabilization with a depth camera", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16-21 June 2012, pp. 89-95.

[9] J. Yang, D. Schonfeld, M. Mohamed, "Robust video stabilization based on particle filter tracking of projected camera motion", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, No. 7, 2009, pp. 945-954.

[10] D. Har, Y. Son, S. Lee, "SLR digital camera for forensic photography", Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications, Vol. 5301, 2004, pp. 276-284.

[11] R. Furchgott, "Camcorders allow a steady shot, even without a steady hand", The New York Times, 2000.

[12] I. Rodriguez-Padilla, B. Castelle, V. Marieu, D. Morichon, "A simple and efficient image stabilization method for coastal monitoring video systems", Remote Sensing, Vol. 12, No. 1, 2019.

[13] K. N. Truong, S. N. Patel, J. W. Summet, G. D. Abowd, "Preventing camera recording by designing a capture-resistant environment", Proceedings of the 7th International Conference UbiComp, Tokyo, Japan, 2005, pp. 73-86.

[14] A. A. Rafiq, W. N. Rohman, S. D. Riyanto, "Development of a simple and low-cost smartphone gimbal with MPU-6050 sensor", Journal of Robotics and Control, Vol. 1, No. 4, 2020, pp. 136-140.

[15] S. Ortega, F. Javier, "Microelectromechanical Systems Inertial Measurement Unit As An Attitude And Heading Reference System", Open Access Theses & Dissertations, 2017, https://scholarworks.utep.edu/open_etd/546 (accessed: 2024)

[16] J. M. Sultan, N. H. Zani, M. Azuani, S. Z. Ibrahim, A. Md Yusop, "Analysis of inertial measurement accuracy using complementary filter for MPU6050 sensor", Jurnal Kejuruteraan, Vol. 34, No. 5, 2022, pp. 959-964.

[17] T. J. Saun, T. P. Grantcharov, "Design and validation of an inertial measurement unit (IMU)-based sensor for capturing camera movement in the operating room", HardwareX, Vol. 9, 2021.

[18] Instructables, "Making a brushless DIY gimbal with Arduino", http://www.instructables.com/id/DIY-Brushless-Gimbal-with-Arduino/ (accessed: 2024)

[19] J. A. Langbridge, "Arduino sketches: tools and techniques for programming wizardry", John Wiley & Sons, 2015.

Efficacy of Continued Fraction Expansion technique in the approximation of fractional order systems

Original Scientific Paper

Nitisha Shrivastava*
Ajay Kumar Garg Engineering College,
Department of Electrical and Electronics Engineering, Ghaziabad, Uttar Pradesh, India, 201015
nitishashrivastav@gmail.com

Arjun Baliyan
Jamia Millia Islamia
Department of Electrical Engineering, New Delhi, India, 110025
arjunbaliyaneee@gmail.com

*Corresponding author

Abstract *– At a macroscopic level derivatives and integrals are the usual mathematical tools to model real time processes and to perform the basic control and signal processing actions. However, the analysis, design, synthesis and implementation of fractional order differentiator and integrator is a difficult task because of its irrational behaviour. Therefore, for mathematical evaluation of any fractional order system, conversion to its approximate integer order equivalent is essential. In this paper approximated integer order models of fractional differentiator and integrator are developed using the continued fraction expansion technique. A continued fraction is an expression obtained through an iterative process. For any iteration to terminate, a finite numerical value is assigned, which in this paper is equal to the number of frequency points within the desired frequency band. It includes both the lower and upper limit values. A set of coefficients are obtained by finding the gains of the fractional term at respective frequencies and thereby applying the recursive formula. The coefficients thus obtained are substituted in the expression of continued fraction which results in a polynomial function of finite order. The developed models can be directly applied for analysis and realization of fractional order systems. The models are developed for fractional terms 0.1 to 0.9 in steps of 0.1, and also for 0.25 and 0.75. A detailed discussion on the sensitivity analysis is presented, which includes the influence of variable parameters on the accuracy and length of the order. Simulations have been performed in MATLAB. A comparison with both, the ideal values and also with existing methods is performed and tabulated to validate the correctness of the developed models both in terms of accuracy and integer order of the model. It shows that the Matsuda method yield very good results both in terms of magnitude and phase. And, is most suitable for linear phase circuits. Also, the proposed models can be directly used for the realization of customized fractional order Proportional Integral (PI), Proportional-Derivative (PD) and PID controllers. To establish the correctness of CFE based technique for hardware realization, the integer order approximated model of one-tenth and seven-tenth differentiator is decomposed to obtain the circuit parameters resistor (R) and capacitor (C). Then its implementation in OrCAD Capture CIS is performed. It can be seen that the results of realization closely match the actual response.*

Keywords: *Matsuda method, Continued Fraction Expansion, Fractional order differentiator, Fractional order integrators, Frequency band*

Received: March 2, 2024; Received in revised form: May 2, 2024; Accepted: May 27, 2024

1. INTRODUCTION

In recent years, researches have been able to explore many potential applications of fractional calculus in science, engineering and business administration [1-7]. It has also been shown that modeling and controlling many financial, biological, chemical, physical, electrical and control phenomena is better done using fractional order calculus [8-10]. Some physical phenomena which show fractional behaviour are; spectral densities of music, viscoelasticity (modelling of cement, gels, polymers), cardiac rhythm, diffusion in plasmas, transport of substances by water in soil, muscle activities, flexible transmission lines, path planning and tracking

for a mobile robot, heat diffusion in the soil, relaxation behaviour of polarized impedances in dielectrics and interfaces and hydraulic actuator.

In order that richness of the dynamic features exhibited by a system or process be properly modified, we need a model of the system, tools for its analysis, ways to specify the required behaviour, methods to design the controller, and techniques to implement them. Since the usual tools to model dynamic systems at a macroscopic level are integrals and derivatives, the algorithms that implement the controllers are mainly composed of such tools. Also, the basic control actions are proportional, derivative and integral. It is quite natural to conclude that by introducing more specific control actions and mathematical substitutes of the form s^{α} and $1/s^{\alpha}$, $(0 < \alpha < 1)$ we could achieve more satisfactory compromises between positive and negative effects and combining the actions we could develop more powerful and flexible design methods to satisfy the controlled system specifications [11-15]. In systems theory the analysis of dynamical behaviour is often made by means of transfer functions. The physical systems which exhibit fractional order (f-o) dynamic behaviour are described by f-o transfer functions in the s-domain. The simple form of a f-o transfer function is $F(s)=s^{\pm\alpha}$ $(0 < \alpha < 1)$. The form s^{α} is continuous time (c-t) f-o differentiator and the form $s^{-\alpha}$ is c-t f-o integrator. To implement transfer functions of this form is not easy, due to its infinite dimensional nature. Therefore, for implementation, it is required to convert fractional functions into integer order functions using different approximation techniques [16-24]. These techniques which are available in literature are based on rational approximations in the frequency domain developed to approximate the arbitrary order with low level of error and wide bandwidth, such as Oustaloup's [16], refined Oustaloup's [17], Charef's [18], Carlson's [19, 20], Matsuda's [21], etc. Both Oustaloup and refined Oustaloup methods are based on pole and zero recursion and are useful where a frequency band of interest is set initially. Also the desired order of the developed model can be chosen a priori. Charef method is also based on interlacing of pole-zero technique, but in this technique the desired order of the developed model cannot be chosen a priori. And moreover, the order of the developed model depends on the fractional order α. The Carlson method is based on Newton iterative process. Similar to Charef method here also the order of the developed model depends on fractional order α and is not uniform for all values of α. In Matsuda method the approximation is obtained using Continued fraction expansion technique and the frequency band as well as the order of the approximated model is set initially. The approximation of fractional differentiator $s^{0.5}$ using different methods is summarized in a survey paper [25]. There are many more applications of fractional order calculus. For example, a synthesis methodology of fractional-order chaotic systems was discussed in [26]. The analysis and analog design of fractional-order charge/flux controlled memristor emulators of incremental/decremental type was described in [27]. In [10, 28, 29] the analog design of fractional-order proportional-integral-derivative controllers was reported. The fractional-order lead/lag compensator was investigated in [30, 31]. The design of double exponent fractional-order filters and power law filters were investigated in [32, 33], respectively, and so on. The dynamic analysis, and subsequently the hardware implementation and realization of such systems can be performed after representing them with finite number of poles and zeros. Usually, hardware implementation of finite order transfer functions is done easily using electronic components/devices. Therefore, in order to study the dynamical behaviour, and for hardware realization purposes, the integer order approximation of the f-o system is required.

Here, the authors have proposed rational approximations of fractional order operator s^{α} for all values of α in the range (-1 to 1) upto one decimal place and for $\alpha = \pm$¼ & ±¾. This forms one of the major contributions of the paper. While analysing fractional order systems, if the integer order approximation of the different fractional orders is readily available, then the fractional order transfer function is converted to its integer order equivalent by merely substituting the approximations of the fractional orders. The resulting transfer function then best approximates the original fractional order system. The Matsuda approximation technique has been utilized for this purpose. Though, this technique is well established, the effect of variation of the parameters used to develop the approximation has not been considered yet. So, another major contribution of the paper is to highlight the variation in frequency response of the approximated integer order models of fractional operator by varying its parameters. The findings are as follows: In Matsuda method there are two important parameters to be selected. One is the number of frequency points (n) and the other is the frequency range for which the developed model is to be used. The relation between differentiator and integrator is that they are inversely proportional to each other and this concept is true for most of the models developed using the different approximation techniques. But this is not applicable to all the models developed by Matsuda method.

The approximate fractional order derivatives obtained as ratio of polynomials in the Laplace domain can be implemented using analog and digital electronics. However, in both cases the exactness depends on the approximation to solve the fractional order function. The correctness of the circuit is therefore dependent on the numerical method and approximated Laplace model used for synthesis. The Field Programmable Analog Array (FPAA) and Field Programmable Gate Array (FPGA) can be used for fast verification and prototyping of fractional order dynamical systems. The circuit whose output displays the behaviour of an irrational function is known as fractional order element (FOE) or fractance device in literature [34-39]. The analog realization with different topologies viz., ladder network, nested ladders, first order RC filters, CMOS OTA based filters have been presented in [40, 41]. The authors in [42] present implementation of resistor less fractional order filters. A review of all the recent developments on the realization of frac-

tance devices can be found in [43]. A comparative study of discrete component realization is given in [44]. The FPGA realization of fractional order chaotic systems is explained in [45, 46]. The drawback in FPGA implementation is that the hardware resources are dependent on the length of the digital word that is used, and this can degrade the desired response due to the finite number of bits to perform computer arithmetic. Therefore implementation with analog electronics using FPAA is good alternative to achieve desired accuracy as presented in [47]. This paper demonstrates a generalized design procedure to develop a passive element R-C structure of fractional order differentiator. As an example, two fractional order differentiators of powers 0.7 (seven-tenth differentiator) and 0.1 (one-tenth differentiator) are simulated using the circuit simulator OrCAD Capture PSpice. A general formula for magnitude and frequency scaling is also presented. The magnitude and phase plots are discussed.

The section wise sequence is as follows: Introduction is covered in Section I. Matsuda method is explained in Section II. In this section integer order models of fractional order differentiator developed using the Continued Fraction Expansion (CFE) formula is listed. Performance and simulation results are discussed in Section III. Seven-tenth and one-tenth differentiator is realized using OrCAD Capture CIS and is presented in Section IV. The conclusion of the paper is in Section V.

2. METHOD

In this section the mathematical steps involved to develop the approximated integer order (AIO) models of fractional differentiators and integrators of order α,($\alpha \in$([±0.1, ±0.9] *in steps of* 0.1 *and* ±1/4, ±3/4)) based on CFE technique are discussed. This method is popularly known as Matsuda method among the fractional community.

The fractional order (f-o) operator is given as [22]

$$F(s) = s^{\pm\alpha}; (0 < \alpha < 1) \tag{1}$$

The f-o differentiator (FOD) is defined as [22]

$$G(s) = s^{\alpha}; (0 < \alpha < 1) \tag{2}$$

The f-o integrator (FOI) is defined as [22]

$$H(s) = s^{-\alpha}; (0 < \alpha < 1) \tag{3}$$

In Matsuda method, the approximated integer order transfer functions of the f-o system are obtained by the use of CFE [21]. For the fractional order α, there are two parameters to be set before applying the formula of continued fraction expansion as given in Eqn. (4). The two parameters are: number of frequency points (n) and the specific Frequency Band (FB) in which the approximation is to developed. ω_i's are the numerical values of the frequency points. A detailed explanation of the frequency band and frequency points is given in section 3.1.1. a_i's are the set of coefficients which are obtained using Eqn. (5). The f-o operator in (1) can be replaced by its equivalent integer order function given as [21, 22]

$$F'(s) = a_1 + \cfrac{s-\omega_1}{a_2 + \cfrac{s-\omega_2}{a_3 + \cfrac{s-\omega_3}{a_4 + \ldots.}}} \tag{4}$$

In Eqn. (4) a_i's are the set of coefficients at different frequencies ωi ($1 \le i \le n$), which are defined within a desired FB [21, 22].

$$a_i = \frac{\omega_i - \omega_{i-1}}{q_{i-1}(\omega_i) - q_{i-1}(\omega_{i-1})}; \text{ for } i = 2,3,\ldots,n \tag{5}$$

where $q_1(\omega_i) = |F(j\omega_i)|$ for $i = 1,2,3,\ldots..,n$ $|F(j\omega_i)|$ are the gains at respective frequencies.

The expression to find gain $|F(j\omega_i)|$ at a specific frequency ω_i is equal to $\pm(\alpha) \times 20\ log\ \omega_i$

Rest of the q's ($q_2, q_3 \ldots, q_n$) are obtained using the recursive formulae [21, 22]

$$q_2(\omega_i) = \frac{\omega_i - \omega_1}{q_1(\omega_i) - q_1(\omega_1)} \text{ for } i = 2,3,\ldots,n$$

$$q_3(\omega_i) = \frac{\omega_i - \omega_2}{q_2(\omega_i) - q_2(\omega_2)} \text{ for } i = 3,4,\ldots,n$$

$$\vdots$$

$$q_n(\omega_i) = \frac{\omega_i - \omega_{n-1}}{q_{n-1}(\omega_i) - q_{n-1}(\omega_{n-1})} \text{ for } i = n$$

Using Eqn. (4), the AIO models of f-o differentiators and integrators are derived and analyzed in the coming subsections.

The analysis of f-o differentiators based on Matsuda method is performed for all orders of α with FB [10^{-2}, 10^{2}] and n = 9. The frequency points within this band are set as [0.01, 0.0316, 0.1, 0.3162, 1, 3.1623, 10, 31.6228, 100]. Using this data, the AIO transfer functions for the FOD s^{α},($\alpha \in$([0.1:0.1:0.9] *and* 1/4, 3/4)) are derived and given in Table 1.

Table 1. AIO transfer function of FOD s^{α}, ($\alpha \in$ ([0.1:0.1:0.9] and 1/4, 3/4))using Matsuda method in FB

s^{α}	G_m atsuda(s)
$s^{0.1}$	1.828((s+52.78)(s+3.143)(s+0.2456)(s+0.01342)/ (s+74.5)(s+4.071)(s+0.3181)(s+0.01894))
$s^{0.2}$	3.3572((s+44.96)(s+2.766)(s+0.2155)(s+0.01111)/ (s+89.98)(s+4.64)(s+0.3615)(s+0.02224))
$s^{0.3}$	6.2275((s+38.54)(s+2.435)(s+0.1887)(s+0.009063)/ (s+110.3)(s+5.298)(s+0.4106)(s+0.02594))
$s^{0.4}$	11.7439((s+33.2)(s+2.145)(s+0.165)(s+0.007245)/ (s+138)(s+6.06)(s+0.4661)(s+0.03011))
$s^{0.5}$	22.7203((s+28.72)(s+1.89)(s+0.1439)(s+0.005634)/ (s+177.5)(s+6.948)(s+0.5291)(s+0.03481))
$s^{0.6}$	47.7342((s+24.93)(s+1.665)(s+0.1252)(s+0.004207)/ (s+237.7)(s+7.987)(s+0.6006)(s+0.0401))
$s^{0.7}$	98.224((s+21.7)(s+1.467)(s+0.1085)(s+0.002945)/ (s+339.5)(s+9.211)(s+0.6817)(s+0.04607))
$s^{0.8}$	237.755((s+18.94)(s+1.292)(s+0.09376)(s+0.001833)/ (s+545.5)(s+10.66)(s+0.7741)(s+0.05278))
$s^{0.9}$	769.99((s+16.56)(s+1.137)(s+0.08061)(s+0.0008552)/ (s+1169)(s+12.4)(s+0.8795)(s+0.06036))
$s^{0.25}$	4.565((s+41.6)(s+2.595)(s+0.2017)(s+0.01006)/ (s+99.42)(s+4.957)(s+0.3853)(s+0.02403))
$s^{0.75}$	149.6819((s+20.27)(s+1.376)(s+0.1009)(s+0.002371)/ (s+421.6)(s+9.906)(s+0.7264)(s+0.04932))

3. SIMULATION RESULTS

This section presents frequency response analysis of the approximated models. As examples magnitude and phase error plots of $s^{0.1}$, $s^{0.7}$ and $s^{0.25}$ are shown in Section 3.1. In section 3.1.1 the effect of variation of n and FB on FOD response is discussed and results analyzed. Similarly, magnitude and phase error plots of $1/s^{0.4}$, $1/s^{0.8}$ and $1/s^{0.75}$ are shown in Section 3.2 and the effect of variation of n and FB on FOI response is discussed in Section 3.2.1.

The outcome of the analysis is summarized highlighting the limitations and advantages of the proposed method. The Matsuda method based FOD models developed in this paper are compared with models developed using four other rational approximation techniques. The results of the comparison are tabulated in Section 3.3.

The comparison has three components: one is the maximum magnitude error in dB and second is the maximum phase error in degree with respect to their ideal values, and third is the order of the approximated model.

3.1. MATSUDA METHOD BASED FOD

The magnitude and phase plots of one-tenth differentiator $s^{0.1}$ (model given in Table 1) is shown in Fig. 1. Also shown is the ideal continuous bode of f-o differentiator $s^{0.1}$ for comparison. The magnitude and phase responses of Matsuda based one-tenth differentiator closely matches the continuous bode in the defined FB $[10^{-2},10^{2}]$.

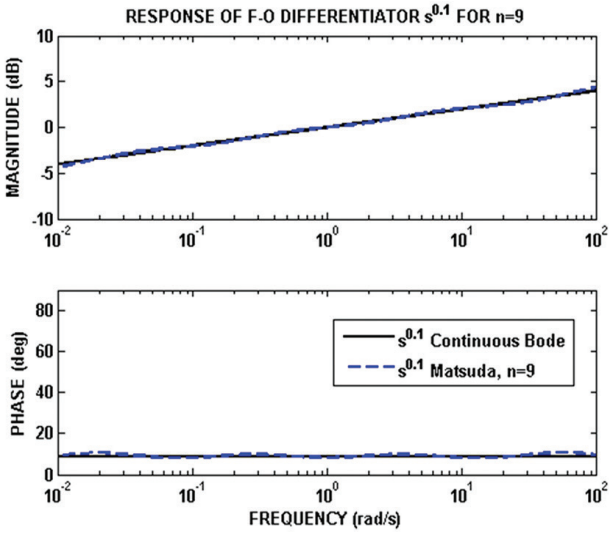


Fig. 1. Frequency Response of Matsuda based one-tenth differentiator $s^{0.1}$ in FB $[10^{-2}, 10^{2}]$,$n = 9$

Figs 2 and 3 show the magnitude and phase plots of Matsuda based f-o differentiators $s^{0.7}$ and $s^{0.25}$ (models given in Table 1) compared with their respective ideal continuous bode counterparts. It is seen that the responses closely match the ideal responses.

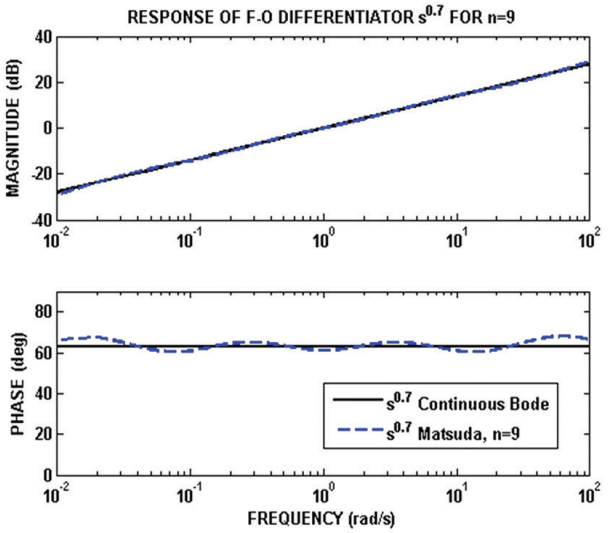


Fig. 2. Frequency Response of Matsuda based seven-tenth differentiator $s^{0.7}$ in FB $[10^{-2}, 10^{2}]$, $n = 9$

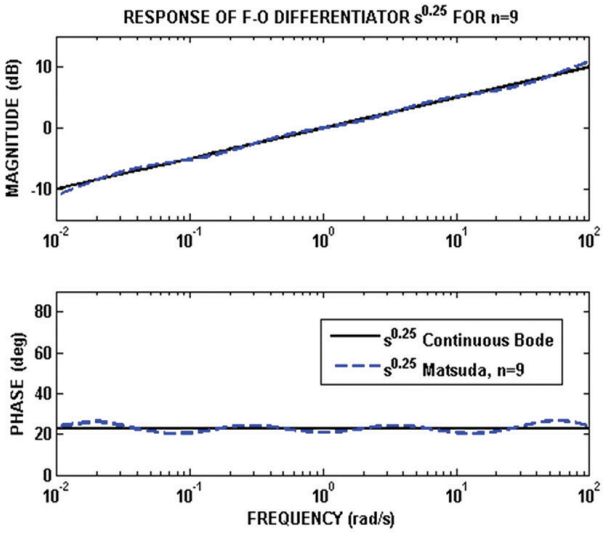


Fig. 3. Frequency Response of Matsuda based one-fourth differentiator $s^{0.25}$ in FB $[10^{-2}, 10^{2}]$, $n = 9$

3.1.1. Effect of Variation of n and FB on FOD Response

Matsuda method, the effect of variation of the two parameters - the specific FB and the number of frequency points within it, on the response of the AIO transfer function of FOD is studied and results analyzed.

Case 1: For this purpose, n is set as 9 and the AIO transfer function for one-tenth differentiator $s^{0.1}$ is

$$G_{0.1\_matsuda1}(s) = \frac{1.828s^4+102.7s^3+329.8s^2+78.9s+0.9996}{s^4+78.91s^3+329.8s^2+102.7s+1.827} \quad (6)$$

Case 2: If the FB is changed to $[10^{-2}, 10^{2}]$, the frequency points change [1, 1.7783, 3.1623, 5.6234, 10, 17.7828, 31.6228, 56.2341, 100] and keeping n same, the AIO for $s^{0.1}$ becomes

$$G_{0.1\_matsuda2}(s) = \frac{1.6s^4+33.29s^3+77.75s^2+27.43s+1}{s^4+27.43s^3+77.75s^2+33.29s+1.6} \quad (7)$$

We observe that the order of the transfer function remains same but the position of poles and zeros vary.

Case 3: Now, keeping FB as in Case 1 and varying n = 13 with frequency points set as [0.01, 0.0215, 0.0464, 0.1, 0.2154, 0.4642, 1, 2.1544, 4.6416, 10, 21.5443, 46.4159, 100], the approximated transfer function for $s^{0.1}$ is

$$G_{0.1\_matsuda3}(s) = \left(\frac{1.946s^6+268.2s^5+3779s^4+8363s^3 \\ +3162s^2+183.8s+0.9988}{s^6+183.9s^5+3163s^4+8363s^3 \\ +3778s^2+268s+1.944} \right) \quad (8)$$

From Eqn. (8) it is clear that when n is increased, the order of the transfer function increases.

Fig. 4 shows the position of poles and zeros for the three cases. Generalizing, it can be inferred that for different FBs, if n remains unchanged, only the position of poles and zeros vary; but the order remains same. And if n is varied, the order of the approximated transfer function also changes. Further, if n is increased, the plots exhibit better matching with the ideal continuous time (c-t) domain f-o differentiator. Analysis can also be performed for different FBs and different n.

The set of frequencies generated in the FB [10^0, 10^2] and various n are listed in Table 2. Table 3 shows the set of frequency points for n= 9 in different FBs.

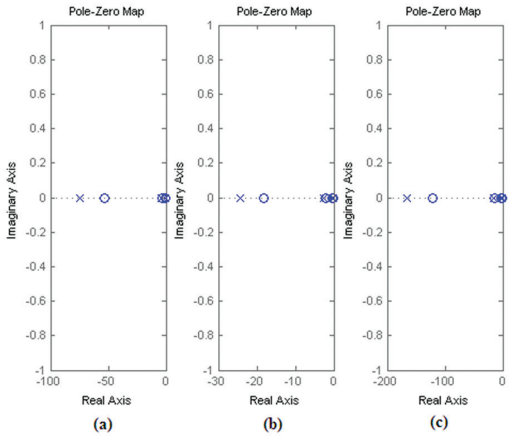


Fig. 4. Pole-zero maps of Matsuda based $s^{0.1}$ for different FBs & n (a) [10^{-2},10^2], n = 9 (b) [$10^{0.1}$, 10^2], n = 9 (c) [10^{-2},10^2], n =13

Table 2. Matsuda method: Set of frequency points for different n in the FB [10^{-2}, 10^2]

n	Frequency points
3	0.01, 1, 100
5	0.01, 0.1, 1, 10, 100
7	0.01, 0.0464, 0.2154, 1, 4.6416,21.5443, 100
9	0.01, 0.0316, 0.1, 0.3162, 1, 3.1623, 10, 31.6228, 100
11	0.01, 0.0251, 0.0631, 0.1585, 0.3981, 1, 2.5119, 6.3096, 15.8489, 39.8107, 100
13	0.01, 0.0215, 0.0464, 0.1, 0.2154, 0.4642, 1, 2.1544, 4.6416, 10, 21.5443, 46.4159, 100
15	0.01, 0.0193, 0.0373, 0.0720, 0.1389, 0.2683, 0.5179, 1, 1.9307, 3.7276, 7.1969, 13.8950, 26.8270, 51.7947, 100
17	0.01, 0.0178, 0.0316, 0.0562, 0.1, 0.1778, 0.3162, 0.5623, 1, 1.7783, 3.1623, 5.6234, 10, 17.7828, 31.6228, 56.2341, 100

FB	Group of frequencies for n=9
[10^0, 10^2]	1, 1.7783, 3.1623, 5.6234, 10, 17.7828, 31.6228, 56.2341, 100
[10^{-2}, 10^2]	0.01, 0.0316, 0.1, 0.3162, 1, 3.1623, 10, 31.6228, 100
[10^{-3}, 10^3]	0.001, 0.0056, 0.0316, 0.1778, 1, 5.6234, 31.6228, 177.8279, 1000
[10^{-1}, 10^5]	0.1, 0.5623, 3.1622, 17.7828, 100, 562.3413, 3162.2776, 17782.7941, 100000
[10^{-1}, 10^1]	0.1, 0.1778, 0.3162, 0.5623, 1, 1.7783, 3.1623, 5.6234, 10
[10^3, 10^5]	1000, 1778.2794, 3162.2776, 5623.4132, 10000, 17782.7941, 31622.7766, 56234.1325, 100000

Table 3. Matsuda method: Group of frequencies for different FBs, for n=9

To study the effect of variation of n, the response of one-tenth differentiator s^0.1 has been plotted with FB [10^{-2}, 10^2] with n = 7, 9, 11, 13 and is shown in Fig. 5. The corresponding magnitude and phase errors are shown in Fig. 6.

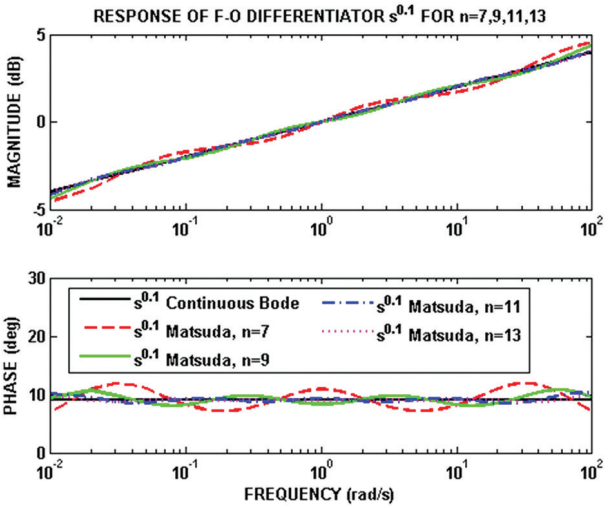


Fig. 5. Frequency Response of Matsuda based one-tenth differentiator $s^{0.1}$ in FB [10^{-2}, 10^2], n = 7,9,11,13

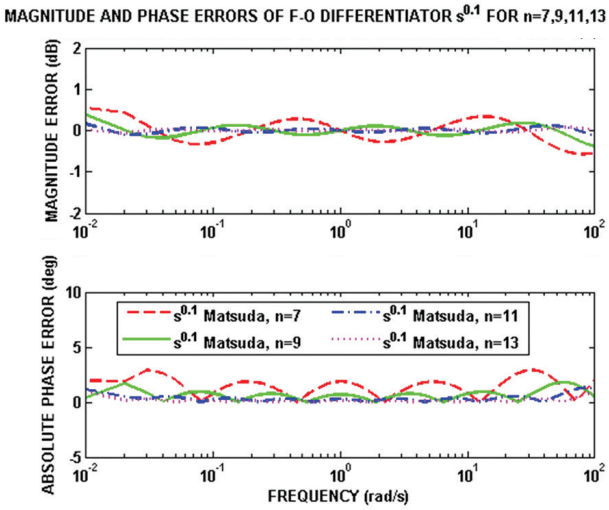


Fig. 6. Magnitude and phase errors of Matsuda based one-tenth differentiator $s^{0.1}$ in FB [10^{-2}, 10^2], n = 7,9,11,13

From the frequency response and error plots, we see that as the choice of number of frequency points increases the response matches more closely to the ideal c-t domain one-tenth differentiator. For the entire FB [10^{-2}, 10^{2}], the maximum magnitude error is less than 0.5 dB "*for*" $n \geq 9$ and the flatness of the phase diagram is obtained for n=13 which finds application in linear phase circuits. It is also observed that the magnitude error of the approximant is least at the chosen frequency points.

As the choice of the number of frequency points increases, the order of the AIO transfer function also increases and accuracy towards the actual value is attained.

An important point worth mentioning here is that the order of the approximated transfer function does not depend on the order of the fractional operator.

Another important inference is that proper approximated models are obtained only for odd values of n and the system is non-causal for even values of n. To illustrate this, let the number of poles be n_p and the number of zeros be n_z.

For odd value of n, $n_p = n_z = (n\text{-}1)/2$ ie. the numerator and denominator polynomials have the same order. For even value of n, $n_p = (n/2)\text{-}1$ and $n_z = n/2$ the numerator polynomial has order one higher than denominator polynomial. Since system is non-causal, further analysis has not been pursued for FODs. Table 4 lists the order of the AIO models obtained for different values of n.

Table 4. Order of Approximated transfer function of FOD s^{α} using Matsuda method for different values of n

n	3	5	7	9	11	13	15	17
Order of approximated integer order model	1	2	3	4	5	6	7	8

3.2. MATSUDA METHOD BASED FOI

The AIO transfer functions of the FOI in Eqn. (3) are obtained by directly inverting the transfer function of the models obtained in Section 2. This is due to the fact that the order of numerator and denominator polynomials is same in the models of Table 1. All the models thus obtained are stable.

Figures 7, 8 and 9 show the frequency responses of Matsuda based f-o integrator $1/s^{0.4}$,$1/s^{0.8}$ and $1/s^{0.75}$ respectively compared with their ideal continuous bode responses in the FB [10^{-2}, 10^{2}] and n set as 9. It is observed that the responses of the Matsuda based f-o integrator models closely match the ideal responses.

In Section 2 the AIO models of FOD were obtained only for odd values of n; the models being non-causal for even values of n. An important observation is that by inverting the models with even values of n, we could obtain transfer functions with order of the numerator one less than that of denominator. The order of the model developed for even values of n is $n/2$. Figs. 10, 11 and 12 show the frequency responses of Matsuda based f-o integrator $1/s^{0.1}$,$1/s^{0.3}$ and $1/s^{0.8}$ respectively compared with their ideal continuous bode responses in the *FB* [10^{-2}, 10^{2}] and n set as 8. It is seen that the responses of these Matsuda based f-o integrator models for even values of n also closely match the ideal responses.

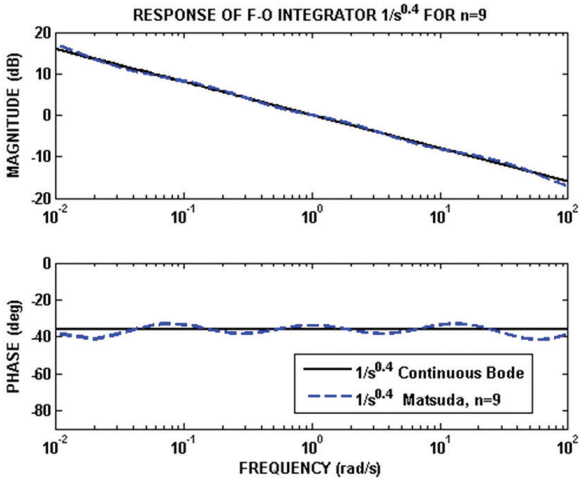


Fig. 7. Frequency Response of Matsuda based four-tenth integrator $1/s^{0.4}$ in FB [10^{-2}, 10^{2}], $n = 9$

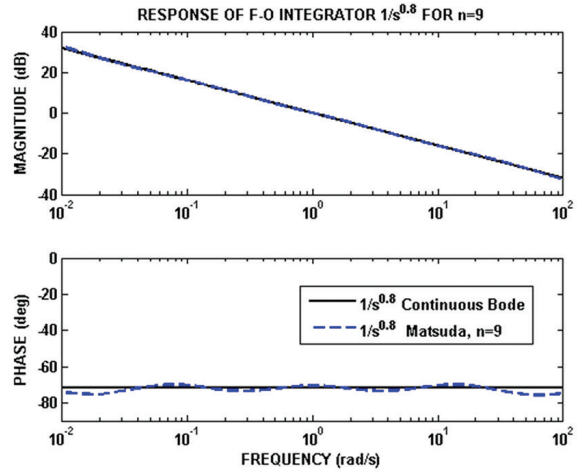


Fig. 8. Frequency Response of Matsuda based eight-tenth integrator $1/s^{0.8}$ in FB [10^{-2}, 10^{2}], $n = 9$

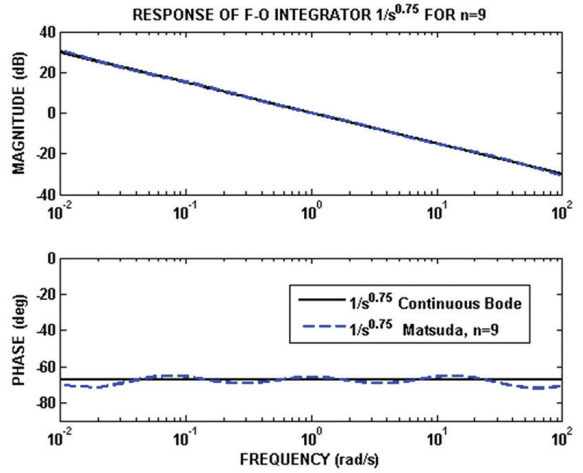


Fig. 9. Frequency Response of Matsuda based three-fourth integrator $1/s^{0.75}$ in FB [10^{-2}, 10^{2}], $n = 9$

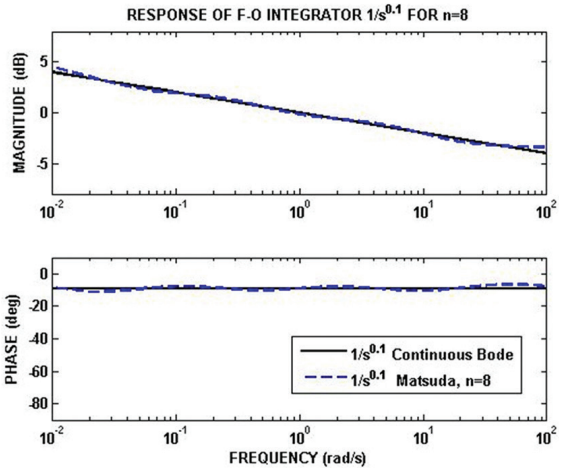


Fig. 10. Frequency Response of Matsuda based one-tenth integrator $1/s^{0.1}$ in FB [10^{-2}, 10^{2}], $n = 8$

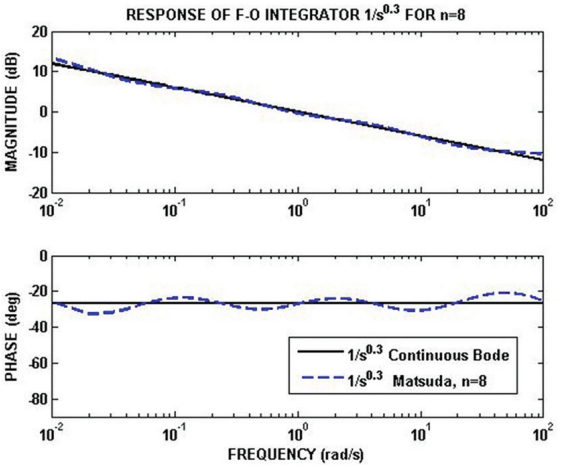


Fig. 11. Frequency Response of Matsuda based three-tenth integrator $1/s^{0.3}$ in FB [10^{-2}, 10^{2}], $n = 8$

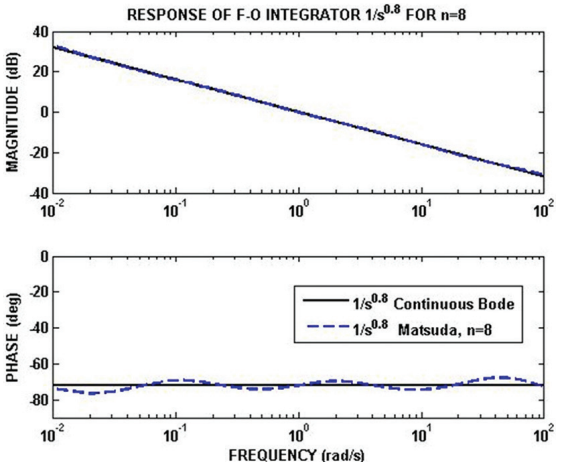


Fig. 12. Frequency Response of Matsuda based eight-tenth integrator $1/s^{0.8}$ in FB [10^{-2}, 10^{2}], $n = 8$

3.2.1. Effect of Variation of *n* and FB on FOI Response

Since FOI models are obtained directly by inverting the FOD models, the discussion presented in section 2 also holds true for FOI response.

To study the effect of variation of n, simulations have been performed for four-tenth integrator $1/s^{0.4}$ with different odd values of n and one-tenth integrator $1/s^{0.1}$ with different even values of n.

Fig. 13 shows the frequency response of approximated Matsuda based four-tenth integrator $1/s^{0.4}$ as compared to the c-t domain FOI for the FB [10^{-2}, 10^{2}], n = 9, 11, 13, 15 and their corresponding magnitude and phase error plots are shown in Fig. 14. It is seen that if n is increased, the response matches very closely to ideal c-t domain FOI $1/s^{0.4}$. The maximum magnitude error is less than 0.5 dB and the maximum phase error is less than 5° for $n \geq 11$. The maximum magnitude error is less than 0.1 dB for $n \geq 13$ and the phase response is flat for $n = 15$ in the FB [10^{-1}, 10^{1}]. The frequency response of Matsuda based one-tenth integrator $1/s^{0.1}$ in the FB [10^{-2}, 10^{2}]," for " n =6, 8, 10, 12 is shown in Fig. 15.

The plot is compared with the ideal response of one-tenth integrator and it is seen that if n is increased, the response matches closely to ideal c-t domain FOI $1/s^{0.1}$. The corresponding magnitude and phase error plots are shown in Fig 16. It is seen that the maximum magnitude error is less than 0.3 dB for $n = 12$ and maximum phase error is less than 5° for n =6, 8, 10, 12.

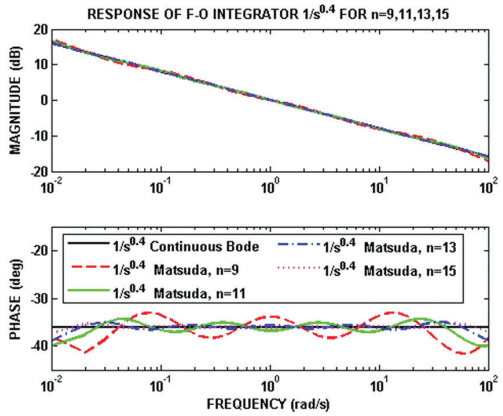


Fig. 13. Frequency Response of Matsuda based four-tenth integrator $1/s^{0.4}$ in FB [10^{-2}, 10^{2}], $n = 9, 11, 13, 15$

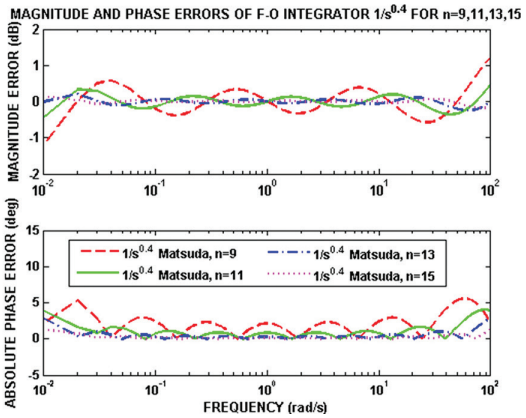


Fig. 14. Magnitude and phase errors of Matsuda based four-tenth integrator $1/s^{0.4}$ in FB [10^{-2}, 10^{2}], n = 9, 11,13,15

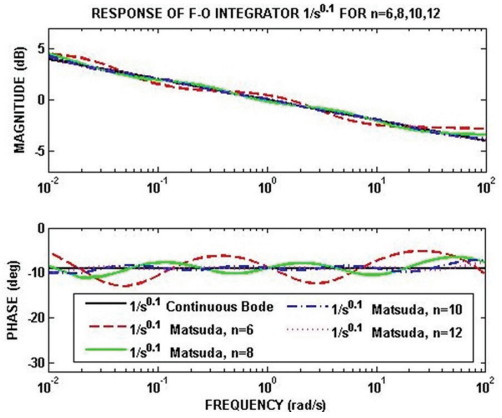


Fig. 15. Frequency Response of Matsuda based one-tenth integrator $1/s^{0.1}$ in FB [10^{-2}, 10^{2}], $n = 6, 8, 10, 12$

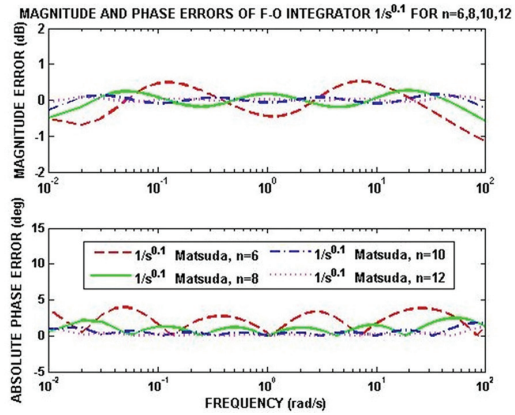


Fig. 16. Magnitude and phase errors of Matsuda based one-tenth integrator $1/s^{0.1}$ in FB [10^{-2}, 10^{2}], $n = 6, 8, 10, 12$

Simulations have been performed for FOD and FOIs for order α ($\alpha \in$ ([±0.1:±0.1:±0.9] and ±1/4, ±3/4)) in the FB [10^{-2}, 10^{2}] for different odd and even values of n using Matsuda method of approximation. Similar observations were noted. It is seen that accurate approximation is achieved for both magnitude and phase for higher values of n, specifically for $n \geq 13$.

3.3. COMPARATIVE STUDY

In this section, a comparative analysis of the FOD models developed in this paper is done with models developed using four other rational approximation techniques. These are Charef method, Carlson method, Oustaloup method and Modified Oustaloup method. The comparison is based on the frequency response of the models i.e. maximum magnitude and phase errors are compared as shown in Tables 5 and 6. Also order of the developed approximated models are mentioned in the comparison table as it is an important factor for realization purposes. Since the number of active/passive elements required for hardware implementation depends upon the order of the approximated model, a compact hardware can be made only with lesser number of components.

In Carlson method and Charef method there is no provision in the formula to select the desired order for which the approximation is to be developed [20, 39]. Only the desired frequency band can be chosen a priori. The order of the approximated model depends on the fractional order α. But in the Oustaloup method and Modified Oustaloup method both desired order and desired frequency band can be selected before applying the formula [16, 17]. This can also be seen from the Tables 5 and 6 that all the models developed by applying Matsuda, Oustaloup and modified Oustaloup methods have fixed orders, 4, 5 and 5 respectively. Whereas for the models based on Charef and Carlson methods the integer orders are not fixed. It depends on the value of α. Although we can the assign a parameter for desired order in Oustaloup and modified Oustaloup methods, but the order of the generated transfer function will always be an odd number [16, 17]. For comparison, the order of these models are chosen as 5, because the models having order 3 gave very poor results.

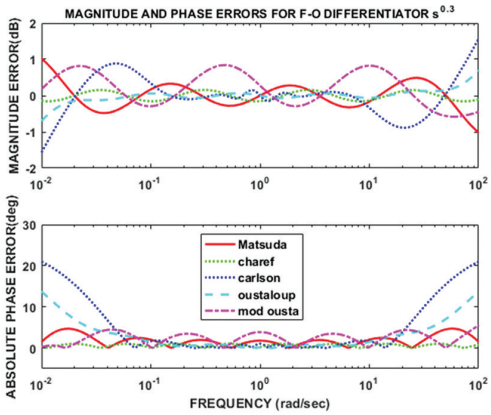


Fig. 17. Magnitude and phase errors of Matsuda, Charef, Carlson, Oustaloup and Modified Oustaloup based three-tenth differentiator $s^{0.3}$ in FB [10^{-2}, 10^{2}]

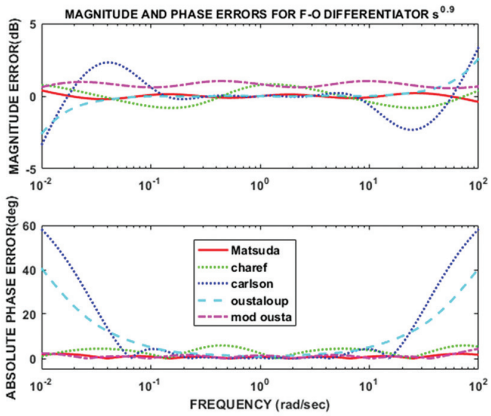


Fig. 18. Magnitude and phase errors of Matsuda, Charef, Carlson, Oustaloup and Modified Oustaloup based nine-tenth differentiator $s^{0.9}$ in FB [10^{-2}, 10^{2}]

As an example, the maximum magnitude and phase error plots is shown in Fig 17 for FOD $s^{0.3}$. The corresponding error values obtained after simulation is presented in Tables 5 and 6. It can be seen that maximum

magnitude error is least for Charef method which is 0.15 dB, and is 0.69 dB in Oustaloup method, 0.84 dB in Modified Oustaloup method, 1.02 dB in Matsuda method and 1.56 dB in Carlson method. The maximum phase error is 1.04°, 4.72°, 5.54°, 13.72° and 21.01° for models based on Charef, Matsuda, Modified Oustaloup, Oustaloup and Carlson methods respectively. Fig. 18 shows the maximum magnitude and phase error plots of FOD $s^{0.9}$ and its corresponding error values are tabulated in Tables 5 and 6. It can be seen that Matsuda based model has least maximum magnitude error of 0.39 dB. The maximum magnitude error for Charef, Modified Outaloup, Oustaloup and Carlson methods are 0.81 dB, 1.03 dB, 2.58 db and 3.36 dB respectively. Also Matsuda based model has least maximum phase error of 2.08° and it is 4.39° and 5.85° for Modified Oustaloup and Charef based models respectively and is very high for Oustaloup (40.57°) and Carlson (58.21°) method approximated models. Similarly, simulations have been performed for all FOD of order α (α∈([0.1:0.1:0.9] and 1/4, 3/4)) in the FB [10^{-2}, 10^{2}] and results tabulated. It can be seen that order of Matsuda based models is 4, which is least among all the other methods. Also it can be seen that the accuracy of the proposed Matsuda based models is very good in comparison to Carlson, Oustaloup and Modified Oustaloup methods. Although Charef based models have lesser magnitude and phase errors but the order of the approximated models are very high which is not suitable for implementation purposes. Therefore Matsuda based models are both accurate and have least order among all other methods and are therefore most suitable for applications in linear phase circuits.

Table 5. Maximum Magnitude Error comparison between proposed FOD s^{α}, (α∈([0.1:0.1:0.9] and 1/4, 3/4)) in the FB [10^{-2}, 10^{2}] with four different methods

	Maximum Magnitude Error (in dB)				
s^{α}	Matsuda Method (ALL order 4)	Charef Method	Carlson Method	Oustaloup Method (ALL order 5)	Modified Oustaloup Method (ALL order 5)
$s^{0.1}$	0.38	0.81 (Order 4)	0.54 (Order 12)	0.22	0.30
$s^{0.2}$	0.73	0.33 (Order 7)	1.02 (Order 7)	0.45	0.59
$s^{0.3}$	1.02	0.15 (Order 9)	1.56 (Order 19)	0.69	0.84
$s^{0.4}$	1.20	0.09 (Order 10)	2.04 (Order 14)	0.95	1.03
$s^{0.5}$	1.27	0.08 (Order 11)	1.32 (Order 4)	1.22	1.15
$s^{0.6}$	1.21	0.09 (Order 11)	1.86 (Order 16)	1.52	1.21
$s^{0.7}$	1.03	0.15 (Order 11)	2.34 (Order 11)	1.84	1.20
$s^{0.8}$	0.74	0.33 (Order 8)	2.88 (Order 23)	2.20	1.14
$s^{0.9}$	0.39	0.81 (Order 5)	3.36 (Order 18)	2.58	1.03
$s^{0.25}$	0.88	0.22 (Order 8)	1.21 (Order 6)	0.57	0.72
$s^{0.75}$	0.89	0.22 (Order 9)	2.53 (Order 10)	2.02	1.18

Table 6. Maximum Phase Error comparison between proposed FOD
FOD s^{α}, (α∈([0.1:0.1:0.9] and 1/4, 3/4)) in the FB [10^{-2}, 10^{2}] with four different methods

	Maximum Phase Error (in degrees)				
s^{α}	Matsuda Method (ALL order 4)	Charef Method	Carlson Method	Oustaloup Method (ALL order 5)	Modified Oustaloup Method (ALL order 5)
$s^{0.1}$	1.79	5.85 (Order 4)	7.06 (Order 12)	4.57	2.22
$s^{0.2}$	3.40	2.24 (Order 7)	13.95 (Order 7)	9.15	4.06
$s^{0.3}$	4.72	1.04 (Order 9)	21.01 (Order 19)	13.72	5.54
$s^{0.4}$	5.58	0.64 (Order 10)	27.90 (Order 14)	18.25	6.54
$s^{0.5}$	5.94	0.56 (Order 11)	30.31 (Order 4)	22.77	7.03
$s^{0.6}$	5.75	0.65 (Order 11)	37.38 (Order 16)	27.26	7.01
$s^{0.7}$	5.02	1.05 (Order 11)	44.26 (Order 11)	31.72	6.52
$s^{0.8}$	3.78	2.24 (Order 8)	51.32 (Order 23)	36.16	5.61
$s^{0.9}$	2.08	5.85 (Order 5)	58.21 (Order 18)	40.57	4.39
$s^{0.25}$	4.11	1.48 (Order 8)	17.25 (Order 6)	11.44	4.86
$s^{0.75}$	4.46	1.49 (Order 9)	47.56 (Order 10)	33.94	6.11

4. IMPLEMENTATION OF FOD

The proposed Matsuda method based fractional order differentiators listed in Table 1 are developed using the CFE formula given in Eqn. 4. The frequency response plots and the maximum magnitude and phase error plots of these s-domain models are generated using MATLAB software. In this section RC circuits of these s-domain models are created using partial fraction expansion formula given in Eqn. 9 and the results verified using OrCAD Capture CIS circuit simulator.

4.1. RC (RESISTOR CAPACITOR) CIRCUIT MODEL

The integer order function of FOD for the FB $[\omega_L\ \omega_H]$ rad/s is converted into partial fraction expansion form as follows [39]:

$$G(s) = A_P + \sum_{i=1}^{n}\left(\frac{sA_i}{(s+p_i)}\right) \tag{9}$$

The objective here is to develop an RC model of the FOD, analogous to the circuit of Fig. 19. The circuit is having one resistor and cascaded RC sets connected in parallel [39]. The RC sets count is dependent upon the number of poles of the integer order function. The resultant admittance of the circuit is [39]

$$Y(j\omega) = \frac{1}{R_p} + \sum_{i=1}^{n}\left(\frac{\frac{j\omega}{R_i}}{\left(j\omega+\frac{1}{R_iC_i}\right)}\right) \tag{10}$$

Comparing (9) and (10),

$$\frac{1}{R_p} = A_p\ ,\ \frac{1}{R_i} = A_i\ ,\ \frac{1}{R_iC_i} = p_i \tag{11}$$

Fig. 19 becomes the equivalent circuit of FOD.

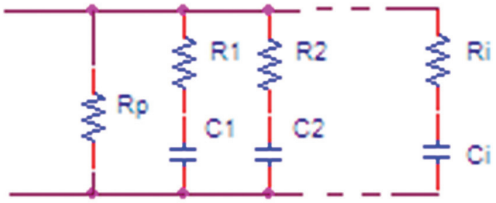


Fig. 19. RC circuit model

4.2. SCALING

Once an RC model of FOD is generated in one FB, the conventional method used to generate the RC model for the same FOD in another FB was to start from scratch i.e. first develop the integer order approximation in the desired FB and then find the values of R & C. But scaling is a technique in which this repetition of procedure is not required. Directly from the RC model of FOD first developed in the chosen FB, the models in other bands can be obtained as explained in this section. In case if it is desired to change only the magnitude of FOD and not the FB, magnitude scaling is also presented here. There are three types of scaling, namely: magnitude scaling, frequency scaling and magnitude-frequency scaling as given in Table 7.

Table 7. Scaling parameters

Scaling	Magnitude Scaling	Frequency Scaling	Magnitude-Frequency Scaling
Scale factor	m	k	a, b
New Scaled values	$[R'=R/m,\ C'=mC]$	$[R'=R,\ C'=C/k]$	$[R'=aR,\ C'=C/(ab)]$
Remark	To shift magnitude, plot up/down	To shift magnitude and phase plots right/left	To simultaneously shift magnitude plot up/down and phase plot right/left

4.3. ONE-TENTH DIFFERENTIATOR $s^{0.1}$

The first example considered is a one-tenth differentiator $s^{0.1}$. Matsuda method is used to obtain the approximated integer order model given as

$$G_{0.1_{Matsuda}}(s) = \frac{3.647s^4+2.049e5s^3+6.58e8s^2+1.575e11s+1.995e12}{s^4+7.891e4s^3+3.298e8s^2+1.027e11s+1.828e12} \tag{12}$$

The frequency range is set as 10^1 to 10^5 rad/s. For Matsuda based AIO transfer function nine frequency points are chosen within this range including the end frequencies. Hence the set of frequencies are [(10, 31.62, 100, 316.22, 1000, 3162.28, 10000, 31622.8, 100000)]. Eqn. 12 is transformed into partial fraction expansion form first and then the passive elements of the circuit are obtained using Eqn. 11.

The values for Matsuda approximated models are:

R_p=0.9163Ω;

R_1=0.9273 Ω; R_2=1.7050 Ω; R_3=2.2232 Ω; R_4=2.2676 Ω;

C_1=0.0144 mF; C_2=0.1440 mF; C_3=1.4000 mF; C_4=23.300 mF

All the values of R and C are positive.

These RC values are used to plot the frequency response of one-tenth differentiator $s^{0.1}$. The results obtained using OrCAD Capture CIS simulator is shown in Fig. 20. It is seen that the plot exhibits 2 dB/dec rise and the phase is approximately 9 deg as desired in the FB $[10^1\ \ 10^4]$ rad/s.

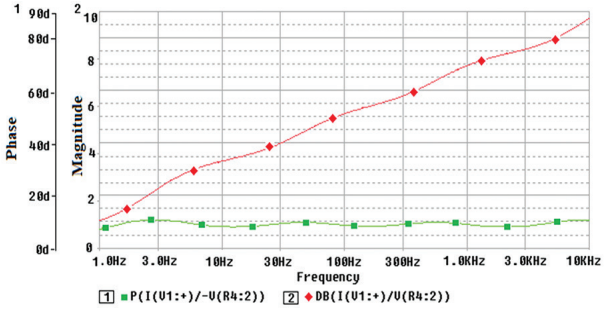


Fig. 20. Bode plot of one-tenth differentiator $s^{0.1}$ in the FB $[10^1\ \ 10^4]$ obtained using RC values based on Matsuda method

In this example, frequency scaling is demonstrated on Eqn. (12) by changing the FBs. It has been mentioned earlier in Section 4.2 that, in case of frequency scaling only the capacitance values change. The scaling

factor 'k' is the frequency value by which the FB is to be changed. All the capacitances in the circuit are divided by 'k' to obtain the desired response.

Thus the new values of capacitances for only frequency shift are given as

$$C' = C \times \frac{1}{k}\ ;\ C' is\ new\ scaled\ capacitance \quad (13)$$

Now, choosing k=10, the scaled capacitances obtained for one-tenth differentiator $s^{0.1}$ in the FB [10^2 10^5] are C_1'=1.4475 µF, C_2'=0.0144 mF, C_3'=0.14 mF and C_4'=2.33mF.

Similarly choosing $k = 10^3$, 10^6, 10^9, 10^{12} and 10^{15}, the scaled parameter is derived for the FBs [10^4 10^7],[10^7 10^{10}], [10^{10} 10^{13}], [10^{13} 10^{16}] and [10^{16} 10^{19}] respectively.

Simulations with scaled capacitances for one-tenth differentiator $s^{0.1}$ are performed and it was observed that the frequency response plots were in correspondence to the frequency response plots of ideal one-tenth differentiator in the subsequent FBs of three-decade widths.

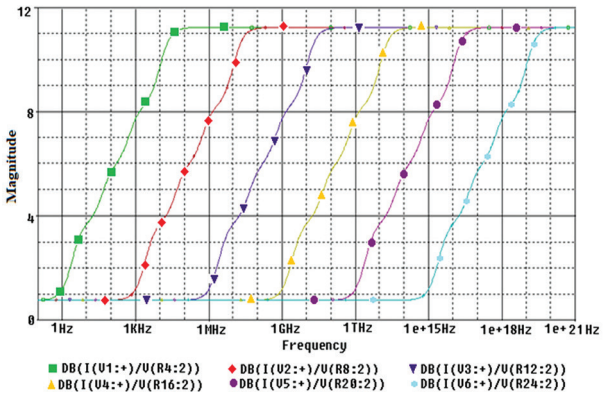


Fig. 21. Magnitude plots of Matsuda method based one-tenth differentiator $s^{0.1}$ for different FBs obtained by scaling of capacitances in the FB of interest

Figures 21 and 22 show magnitude and phase plots of one-tenth differentiator $s^{0.1}$ for different FBs [10^1 10^4], [10^4 10^7], [10^7 10^{10}], [10^{10} 10^{13}], [10^{13} 10^{16}] and [10^{16} 10^{19}] of 3 decades each. The total span of frequency is from 10^1 to 10^{19}, obtained by scaling of capacitances.

From Fig. 21, it is observed that the responses of $s^{0.1}$ differentiator obtained by scaling the capacitances in different FBs, match with their ideal responses in the FB of interest (i.e. for three decades). In each case the slope of the magnitude plot is 2dB/dec (as desired for a $s^{0.1}$ differentiator). It is further observed that there is no change in the magnitude when the FBs change. For each FB, the phase plots are shown in Fig. 22 (a-f).

It is again clear that the phase is approximately 9° (as desired for a $s^{0.1}$ differentiator) in the region of interest for the whole range with a maximum error of 1.8° and 0.8° in the ranges [10^1 10^4] and [10^2 10^4] rad/s respectively in each plot. Separate subplots have been purposefully shown so as to validate the results.

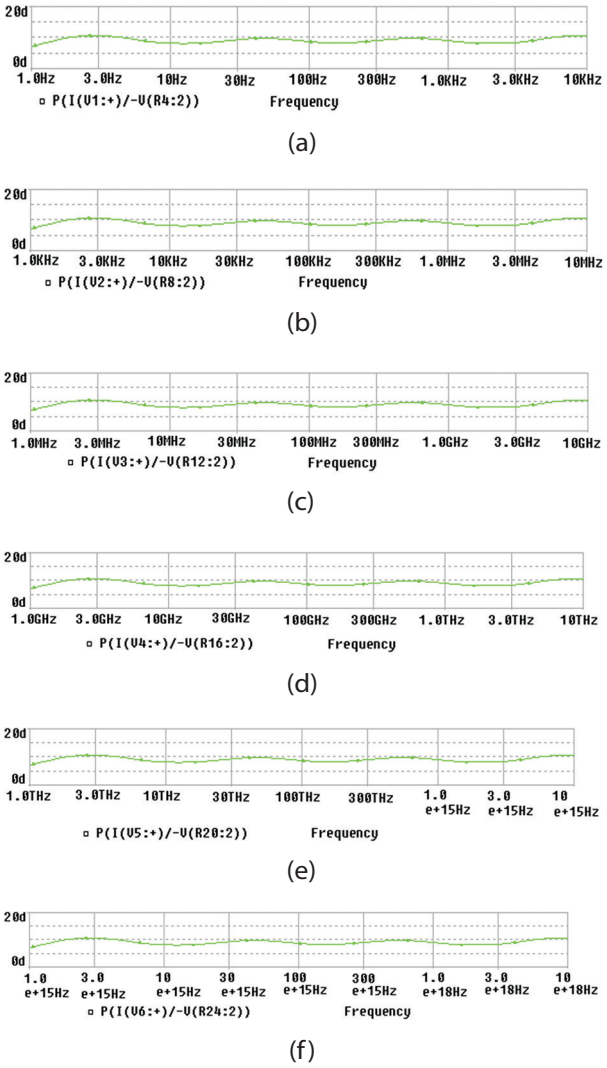


Fig. 22. (a)-(f). Phase plots of Matsuda method based one-tenth differentiator s^{α} (α=0.1) for different FBs [10^1 10^4], [10^4 10^7], [10^7 10^{10}], [10^{10} 10^{13}], [10^{13} 10^{16}] and [10^{16} 10^{19}] obtained by scaling of capacitances

4.4. SEVEN-TENTH DIFFERENTIATOR $s^{0.7}$

The Matsuda method is used to develop integer order approximations of $s^{0.7}$ in the FB [10^{-2} 10^2]. The approximation is given in Eqn. 14.

$$G_{0.7\_Matsuda}(s) = \frac{98.22s^4+2287s^3+3380s^2+349.3s+0.9996}{s^4+349.4s^3+3381s^2+2287s+98.2} \quad (14)$$

Eqn (14) is then transformed into partial fraction expansion form. Since the order of the approximation is 4, the number of circuit elements required as per Fig. 19 is nine (five resistances and four capacitances).

The values of the circuit elements obtained for Matsuda based AIO transfer functions are:

R_p=98.0392 Ω;

R_1=10.6mΩ; R_2=298.5m Ω; R_3=1.9869 Ω; R_4=11.1859 Ω;

C_1=277.7 mF; C_2=363.7 mF; C_3=738.2 mF; C_4=1.9393 mF

Using these RC values, the circuit is simulated with input voltage as 1V ac. The magnitude and phase plots

of seven-tenth differentiator $s^{0.7}$ is shown in Fig. 23. It is observed that the magnitude and phase values as obtained from the plot for all the frequency points in the chosen FB are very close to their corresponding ideal values in the whole FB.

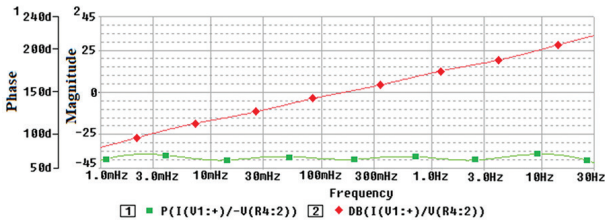


Fig. 23. Bode plot of seven-tenth differentiator $s^{0.7}$ in the FB [10^{-2} 10^{2}] obtained using RC values based on Matsuda method

The analysis for magnitude scaling is performed on the Matsuda model.

Now, if it is desired to change the magnitude (gain) of the FOD, magnitude scaling is applied. The scaling factor 'm' depends upon the desired change in magnitude. For an 'A' dB rise or fall in magnitude, the scaling factor is obtained as follows:

$$20 \log (m) = \pm \mathrm{A} \tag{15}$$

$$\therefore \ \mathrm{m} = 10^{\pm \frac{A}{20}} \tag{16}$$

Correspondingly all resistances are divided by factor 'm' and all capacitances are multiplied by factor 'm'.

Thus, new values of resistances and capacitances for magnitude shift are given as

$$\begin{aligned} R' &= R \times 1/m \,;\ R' \textit{is new scaled resistance} \\ C' &= C \times m \,;\ C' \textit{is new scaled capacitance} \end{aligned} \tag{17}$$

The following two cases are considered for simulations on Matsuda approximated seven-tenth differentiator:

Case 1: 14 dB rise in magnitude

The scaling factor 'm' is calculated as 5.01187. The scaled values of resistances and capacitances are

Rp=19.56 Ω;

R_1=2.1149 mΩ; R_2=59.5586 mΩ; R_3=396.439 mΩ; R_4=.2318 Ω;

C_1=1.3917 F; C_2=1.8228 F; C_3=3.6997 F; C_4=1.7195 F

Case 2: 25 dB fall in magnitude

The scaling factor 'm' is calculated as 0.056234. The scaled values of resistances and capacitances are

R_p=174.34 KΩ;

R_1=188.84 mΩ; R_2=5.3081 Ω; R_3=35.33 Ω; R_4=198.91 Ω;

C_1=15.616 mF; C_2=20.45 mF; C_3=41.51 mF; C_4=109.054 mF

It can be seen that, all the values of R and C are positive.

Figure 24 shows the frequency response plot of seven-tenth-tenth differentiator for different two different scaling factors. It can be seen that the magnitude plot is shifted up by 14 dB and down by 25 dB in the entire FB for Case 1 (red line with diamond shapes) and Case 2 (blue line with triangle shapes) respectively as desired. There is no change in the phase plots.

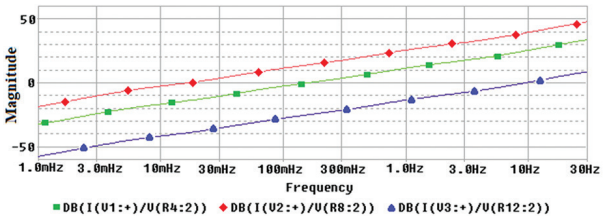


Fig. 24. Magnitude plots of Matsuda method based seven-tenth differentiator $s^{0.7}$ for different scaling factors

Similar simulations were performed for other FODs of order ($\alpha \in$([±0.1:±0.1:±0.9] and ±1/4, ±3/4)) (models presented in Table 1) with all RC elements having positive values.

5. CONCLUSION

In this paper, we have proposed s-domain stable models of f-o operator s^{α} developed using continued fraction expansion technique, popularly known as Matsuda method. The findings of the parameter sensitivity analysis are as follows: the parameter of importance is n, the number of frequency points within the FB of interest. It is observed that the selection of 'n' is directly proportional to the order of the approximated integer order transfer function and that the approximation is non-causal for odd values of n in case of FOD. A detailed frequency analysis is performed and it is seen that for the models proposed in our work, the error is less. Also, these models are most suitable for linear phase circuits as the phase response is almost flat in desired FB. In the Matsuda approximation method, if the number of frequency points chosen is $n \geq 13$, the phase response shows linear behaviour throughout the desired FB. An important point worth mentioning here is that the order of the approximated transfer function does not depend on the order of the fractional operatorα. For the proposed model's direct realization in hardware is possible for f-o control purposes as the order is less and accuracy is high. In order to validate the proposed models so that they can be used for realization purposes, a structure with passive elements is designed and results of one tenth and seven tenth differentiator presented. Similar results were found for other FODs of order α($\alpha \in$([±0.1:±0.1:±0.9] and ±1/4, ±3/4)) (models presented in Table 1) with all RC elements having positive values. This shows that negative impedance converter is not required and also inductors, which would actually make the system bulky have not been used in circuit realization. This reduces the complexity of hardware realization. Another advantage of this design procedure is that there is only one junction other than reference junction. The scaling relations of R and C have been developed for smooth transition from one FB to another. Simulations were performed with scaled

parameters and it was found that the magnitude and phase response matched *20α dB/decade and 90α degrees* respectively.

6. REFERENCES

[1] X. Cui, D. Xue, F. Pan, "A fractional SVIR-B epidemic model for Cholera with imperfect vaccination and saturated treatment", The European Physics Journal Plus, Vol. 137, No. 1361, 2022.

[2] H. Sun, A. Chang, Y. Zhang, W. Chen, "A Review on Variable-Order Fractional Differential Equations: Mathematical Foundations, Physical Models, Numerical Methods and Applications", Fractional Calculus and Applied Analysis, Vol. 22, 2019, pp. 27-59.

[3] Y. Wei, Y. Chen, Y. Wei, X. Zhao, "Lyapunov stability analysis for nonlinear nabla tempered fractional order systems", Asian Journal of Control, Vol. 25, No. 4, 2023, pp. 3057-3066.

[4] S. Yadav, D. Kumar, J. Singh, D. Baleanu, "Analysis and dynamics of fractional order Covid-19 model with memory effect", Results in Physics, Vol. 24, 2021, p. 104017.

[5] A. Boudaoui, Y. E. Moussa, Z. Hammouch, S. Ullah, "A fractional-order model describing the dynamics of the novel coronavirus (COVID-19) with nonsingular kernel", Chaos, Solitons & Fractals, Vol. 146, 2021, p. 110859.

[6] P. Veeresha, E. Ilhan, D. G. Prakasha, H. M. Baskonus, W. Gao, "Regarding on the fractional mathematical model of tumour invasion and metastasis", Computer Modeling in Engineering & Sciences, Vol. 127, No. 3, 2021, pp. 1013-1036.

[7] A. Atangana, D. Baleanu, "New fractional derivatives with nonlocal and non-singular kernel: theory and application to heat transfer model", Thermal Science, Vol. 20, No. 2, 2016, pp. 763-769.

[8] K. B. Oldham, J. Spanier, "The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order", 1st Edition, Dover Publications Mineola New York, USA, 2006.

[9] H. Sun, Y. Zhang, D. Baleanu, W. Chen, Y. Chen, "A new collection of real world applications of fractional calculus in science and engineering", Communications in Nonlinear Science and Numerical Simulation, Vol. 64, 2018, pp. 213-231.

[10] I. Petráš, "Fractional-Order Nonlinear Systems: Modeling, Analysis and Simulation", 1st Edition, Springer Berlin Heidelberg, 2011.

[11] W. Zheng, Y. Luo, X. Wang, Y. Pi, Y. Chen, "Fractional order PIλDμ controller design for satisfying time and frequency domain specifications simultaneously", ISA Transactions, Vol. 68, 2017, pp. 212-222.

[12] E. Yumuk, M. Güzelkaya, İ. Eksin, "A robust fractional-order controller design with gain and phase margin specifications based on delayed Bode's ideal transfer function", Journal of the Franklin Institute, Vol. 359, No. 11, 2022, pp. 5341-5353.

[13] R. Shalaby, M. El-Hossainy, B. Abo-Zalam, T. A. Mahmoud, "Optimal fractional-order PID controller based on fractional-order actor-critic algorithm", Neural Computing and Applications, Vol. 35, 2023, pp. 2347-2380.

[14] A. Latif, S. M. S. Hussain, D. C. Das, T. S. Ustun, A. Iqbal, "A review on fractional order (FO) controllers' optimization for load frequency stabilization in power networks", Energy Reports, Vol. 7, 2021, pp. 4009-4021.

[15] Z. Wu, J. Viola, Y. Luo, Y. Chen, D. Li, "Robust fractional-order [proportional integral derivative] controller design with specification constraints: more flat phase idea", International Journal of Control, Vol. 97, 2024, pp. 111-129.

[16] A. Oustaloup, F. Levron, B. F. M. Nanot, "Frequency band complex noninteger differentiator: Characterization and synthesis", IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, Vol. 47, 2000, pp. 25-39.

[17] D. Y. Xue, C. Zhao, Y. Q. Chen, "A modified approximation method of fractional order system", Proceedings of the IEEE International Conference on Mechatronics and Automation, Louyang, China, 25-28 June 2006, pp. 1043-1048.

[18] A. Charef, H. H. Sun, Y. Y. Tsao, B. Onaral, "Fractal system as represented by singularity function", IEEE Transactions on Automatic Control, Vol. 37, 1992, pp. 1465-1470.

[19] G. Carlson, C. Halijak, "Approximation of fractional capacitors $(1/s)^{1/n}$ by a regular Newton process", IEEE Transactions on Circuit Theory, Vol. 11, 1964, pp. 210-213.

[20] N. Shrivastava, P. Varshney, "Rational Approximation of Fractional Order Systems Using Carlson Method", Proceedings of the International Conference on Soft Computing Techniques and Implementation, Faridabad, India, 8-10 October 2015, pp. 76-80.

[21] K. Matsuda, H. Fujii, "H(∞)-optimized wave-absorbing control -Analytical and experimental results", Journal of Guidance, Control & Dynamics, Vol. 16, 1993, pp. 1146-1153.

[22] B. M. Vinagre, I. Podlubny, A. Hernandez, V. Feliu, "Some approximations of fractional order operators used in control theory and applications", Journal of Fractional Calculus and Applied Physics, Vol. 3, 2000, pp. 231-248.

[23] G. Maione, "On the Laguerre Rational Approximation to Fractional Discrete Derivative and Integral Operators", IEEE Transactions on Automatic Control, Vol. 58, No. 6, 2013, pp. 1579-1585.

[24] G. Maione, R. Caponetto, A. Pisano, "Optimization of zero-pole interlacing for indirect discrete approximations of noninteger order operators", International Journal of Computers and Mathematics with Applications, Vol. 66, 2013, pp. 746-754.

[25] J. D. Colín-Cervantes, C. Sanchez-Lopez, R. Ochoa-Montiel, D. Torres-Muñoz, C. M. Hernández-Mejía, L. A. Sánchez-Gaspariano, H. G. González-Hernández, "Rational Approximations of Arbitrary Order: A Survey", Fractal and Fractional, Vol. 5, No. 4, 2021, p. 267.

[26] C. Sánchez-López, "An experimental synthesis methodology of fractional-order chaotic attractors", Nonlinear Dynamics, Vol. 100, 2020, pp. 3907-3923.

[27] C. Sánchez-López, V. H. Carbajal-Gómez, M. A. Carrasco-Aguilar, I. Carro-Pérez, "Fractional-Order Memristor Emulator Circuits", Complexity, Vol. 2018, No. 1, 2018, p. 2806976.

[28] I. Podlubny, "Fractional-order systems and PIλDμ-controllers", IEEE Transactions on Automatic Control, Vol. 44, 1999, pp. 208-214.

[29] C. Muñiz-Montero, L. V. García-Jiménez, L. A. Sánchez-Gaspariano, C. Sánchez-López, V. R. González-Díaz, E. Tlelo-Cuautle, "New alternatives for analog implementation of fractional-order integrators, differentiators and PID controllers based on integer-order integrators", Nonlinear Dynamics, Vol. 90, 2017, pp. 241-256.

[30] H. F. Raynaud, A. Zergaïnoh, "State-space representation for fractional order controllers", Automatica, Vol. 36, No. 4, 2000, pp. 1017-1021.

[31] C. A. Monje, B. M. Vinagre, A. J. Calderon, V. Feliu, Y. Q. Chen, "Auto-tuning of fractional lead-lag compensators", Proceedings of the 16th IFAC World Congress, Prague, Czech Republic, 4-8 July 2005, pp. 319-324.

[32] S. Kapoulea, C. Psychalinos, A. S. Elwakil, "Double exponent fractional-order filters: Approximation methods and realization", Circuits, Systems, and Signal Processing, Vol. 40, 2021, pp. 993-1004.

[33] S. Kapoulea, C. Psychalinos, A. S. Elwakil, "Power law filters: A new class of fractional-order filters without a fractional-order Laplacian operator", AEU - International Journal of Electronics and Communications, Vol. 129, 2021, p. 153537.

[34] D. Mondal, K. Biswas, "Performance study of fractional order integrator using single component fractional order element", IET Circuit, Devices and Systems, Vol. 5, No. 4, 2011, pp. 334-342.

[35] L. Dorcak, J. Valsa, E. Gonzalez, J. Terpak, I. Petras, "Analogue Realization of Fractional-Order Dynamical System", Entropy, Vol. 15, No. 10, 2013, pp. 4199-4214.

[36] I. Podlubny, I. Petras, B. M. Vinagre, P. O'Leary, L. Dorcak, "Analogue Realization of Fractional-Order Controllers", Nonlinear Dynamics, Vol. 29, 2002, pp. 281-296.

[37] A. S. Elwakil, "Fractional-order circuits and systems: An emerging interdisciplinary research area", IEEE Circuits and Systems Magazine, Vol. 10, No. 4, 2010, pp. 40-50.

[38] G. Radwan, A. Shamim, K. N. Salama, "Theory of fractional order elements based on impedance matching networks", IEEE Microwave and Wireless Components Letters, Vol. 21, No. 3, 2011, pp. 120-122.

[39] A. Charef, "Analogue realisation of fractional-order integrator, differentiator and fractional $PI^{\lambda}D^{\mu}$ con-

troller", IEE Proceedings - Control Theory and Applications, Vol. 153, No. 6, 2006, pp. 714-720.

[40] D. Sierociuk, I. Podlubny, I. Petras, "Experimental Evidence of Variable-Order Behaviour of Ladders and Nested Ladders", IEEE Transactions on Control Systems Technology, Vol. 21, No. 2, 2013, pp. 459-466.

[41] M. A. Valencia-Ponce, P. R. Castañeda-Aviña, E. Tlelo-Cuautle, V. H. Carbajal-Gómez, V. R. González-Díaz, Y. Sandoval-Ibarra, J. Nuñez-Perez, "CMOS OTA-Based Filters for Designing Fractional-Order Chaotic Oscillators", Fractal and Fractional, Vol. 5, No. 3, 2021, p. 122.

[42] D. Patrinos, G. Tsirmpas, P. Bertsias, C. Psychalinos, A. S. Elwakil, "Implementation and Experimental Verification of Resistorless Fractional-Order Basic Filters", Electronics, Vol. 11, No. 23, 2022, p. 3988.

[43] B. T. Krishna, "Recent Developments on the Realization of Fractance Device", Fractional Calculus and Applied Analysis, Vol. 24, 2021, pp. 1831-1852.

[44] G. Tsirimokou, A. Kartci, J. Koton, N. Herencsar, C. Psychalinos, "Comparative Study of Discrete Component Realizations of Fractional-Order Capacitor and Inductor Active Emulators", Journal of Circuits, Systems and Computers, Vol. 27, No. 11, 2018, p. 1850170.

[45] M. F. Tolba, A. M. AbdelAty, N. S. Soliman, L. A. Said, A. H. Madian, A. T. Azar, A. G. Radwan, "FPGA implementation of two fractional order chaotic systems", AEU - International Journal of Electronics and Communications, Vol. 78, 2017, pp. 162-172.

[46] S. M. Mohamed, W. S. Sayed, L. A. Said, A. G. Radwan, "Reconfigurable FPGA Realization of Fractional-Order Chaotic Systems", IEEE Access, Vol. 9, 2021, pp. 89376-89389.

[47] A. Silva-Juárez, E. Tlelo-Cuautle, L. G. Fraga, R. Li, "FPAA-based implementation of fractional-order chaotic oscillators using first-order active filter blocks", Journal of Advanced Research, Vol. 25, 2020, pp. 77-85.

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

Copyright

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

TITLE OF ARTICLE (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

Author/Authorized Agent **Date**

CONTACT
International Journal of Electrical and Computer Engineering Systems (IJECES)
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr