International Journal of Applied Power Engineering (IJAPE)
Vol. 15, No. 1, March 2026, pp. 98~109
ISSN: 2252-8792, DOI: 10.11591/ijape.v15.i1.pp98-109

High impedance fault discrimination in microgrid power system using stacking ensemble approach

Arangarajan Vinayagam[1], Raman Mohandas[1], Meyyappan Chindamani[2], Bhadravathi Gavirangapa Sujatha[3], Soumya Mishra[4], Arivoli Sundaramurthy[5]
[1]Department of Electrical and Electronics Engineering, New Horizon College of Engineering, Bangaluru, India
[2]Department of Electrical and Electronics Engineering, Sri Ramakrishna Engineering College, Coimbatore, India
[3]Department of Electrical and Electronics Engineering, Cambridge Institute of Technology, Bangaluru, India
[4]School of Electrical Engineering, Kalinga Institute of Industrial Technology, Bhubaneswar, India
[5]Department of Electrical and Electronics Engineering, PSG Institute of Technology and Applied Research, Coimbatore, India

Article Info

Article history:

Received Apr 18, 2025
Revised Sep 17, 2025
Accepted Oct 16, 2025

Keywords:

Decision tree
Discrete wavelet transform
High impedance fault
Microgrid
Photovoltaic
Stacking ensemble classifier

ABSTRACT

High impedance (HI) faults in microgrid (MG) power systems are non-linear, intermittent, and have low fault current magnitudes, making them challenging to detect by typical protective systems. Consequently, it is imperative to implement a sophisticated protection system that is dependent on the precision of fault detection. In this study, a stacking ensemble classifier (SEC) is proposed to discriminate HI fault from other transients within a photovoltaic (PV) generated MG power system. The MG model is simulated with the introduction of faults and transients. The features of data set from event signals are generated using the discrete wavelet transform (DWT) technique. The dataset is used to train the individual classifiers (Naïve Bayes (NB), decision tree J48 (DTJ), and K-nearest neighbors (KNN)) at initial and meta learner in the final stage of SEC. The SEC outperforms other classification methods with respect to accuracy of classification, rate of success in detecting HI fault, and performance measures. The outcomes of the classification study conducted under standard test conditions (STC) of solar PV and the noisy environment of event signals clearly demonstrate that the SEC is more dependable and performs better than the individual base classification approaches.

Corresponding Author:

Arangarajan Vinayagam
Department of Electrical and Electronics Engineering, New Horizon College of Engineering
Bangaluru, India
Email: arajanin@gmail.com

1. INTRODUCTION

Microgrids (MGs) are used to deliver dependable, cost-effective, and durable energy to remote areas [1]. However, integrating multiple distributed generation (DG) sources, such as conventional, non-linear, and intermittent renewable energy (RE) sources, and occurrence of abnormal events brought on by switching transients and faults (low and high impedance (HI) faults) can negatively impact the MG's security and reliability [2]. HI faults are prevalent in MG networks when conductors contact high-resistance surfaces like wet sand, asphalt, tree limbs, and gravel [3]. These faults can cause electrical shock, fire, and cascading failure, affecting healthy parts of the MG network [4], [5]. The HI fault exhibits a low fault current amplitude, complicating detection and isolation with conventional protective relays. In this instance, a sophisticated protection system is needed to accurately detect and differentiate HI faults in MG networks to isolate the faulty area more quickly and precisely [6]. Therefore, to find the problematic network segment faster and more accurately, an enhanced protection system that depends on precise detection and identification of HI faults is essential. To accomplish more precise detection and discrimination of HI faults in MG networks, advanced machine learning (ML) classification techniques can be implemented.

Researchers have used various ML algorithms to discover and categorize HI faults in various power system models and MG networks, ensuring a more efficient and accurate isolation of the problematic area. This sophisticated protection system is important to ensure the safe and reliable operation of MG networks. Multi-layer perceptron (MLP) neural networks [7], multi-class support vector machines (SVM) [8], [9], fuzzy and ANFIS techniques [10], Elman neural networks [11], Kalman filter-based techniques [12], and convolutional neural networks (CNN) [6] have been used to discriminate HI faults in power networks. HI fault occurrences have been analysed in MG networks using the Naive Bayes classifier [13] and in photovoltaic (PV) integrated power networks using mathematical morphology [14]. As per the literature, single-base classifiers are widely used, and they are generally effective for specific tasks but may struggle with generalisation due to their susceptibility to noise or overfitting. An ensemble model enhances the accuracy and consistency of single-base classifiers by training multiple classifiers on the same problem [15]. Researchers have used ensemble approaches like voting probability [16], bagged tree [17], and random forest [18] to classify electrical faults in power systems. Ensemble classifiers based on extended Kalman filters [19] and KNN-based random subspace approaches [3] have been used to discriminate high intensity current (HI) fault in PV-integrated power networks. A boosting ensemble method [20] and voting ensemble method [21] have been proposed for detecting and classifying electrical faults. The research suggests that ensemble classifiers are more reliable and efficient than single-base classifiers. The bagging and boosting ensemble algorithms use homogenous weak classifiers, however an advanced stacking ensemble classifier uses a collection of heterogeneous weak classifiers for superior generalization accuracy [22], [23]. Therefore, this investigation suggests a stacking methodology for classifying HI defects from other transients in a PV-generated MG power system. The present research addresses a gap in the literature by concentrating on the evaluation of HI faults using a stacking ensemble approach. This topic has been the subject of limited investigation in PV-based MG power systems. The method is designed to accurately classify faults in a noisy environment of event signals and standard test conditions of solar PV. Addressing these research gaps is crucial for improving fault classification methodologies in RE-integrated MG power networks and developing robust ensemble-based approaches for real-world applications. The study's primary contributions are:

- A stacking technique of an ensemble classifier is proposed to differentiate HI fault from other faults and transients in a PV-generated MG power system.
- Using the discrete wavelet transform (DWT) method, features of the dataset from faults and transient signals are generated to train the suggested stacking and individual classifiers.
- To verify the efficacy of the suggested stacking technique, a classification analysis is performed in terms of accuracy and performance metrics (PM) while MG is operating under standard test condition.
- To verify the resilience of the suggested ensemble model, a classification study is carried out under a noisy environment of event signals.

Structure of the manuscript: section 2 describes MG model, HI fault model, and process steps of classification model; section 3 defines DWT method; section 4 presents the details of materials and methodology; section 5 discusses classification results; and section 6 summarizes research findings and future steps.

2. DESCRIPTION OF PV INTEGRATED MG MODEL

This study uses the ensemble classifier to distinguish HI faults from other transients in a PV-generated MG model. MATLAB-Simulink is used to simulate and analyse MG network model. Figure 1 shows the simple diagram of an islanded PV generated MG power system with integration of following elements: solar PV system: 3 units rated 300 kWp (100 kWp/unit); DC-DC powert converter (290 V/500 V DC) with maximum power point tracking control; voltage source inverter interface PV source into AC network through transformer T1 (0.260 kV/11 kV, 300 kVA, 50 Hz); diesel engine generator (DEG): 3.25 MVA DEG is interconnected to AC network through transformer T2; AC load maximum capacity of 2.2 MW at 11 kV; capacitor bank maximum capacity of 700 kVAR at 11 kV; and HI fault model with anti-parallel diodes, resistors (R1 and R2), and voltage sources (V1 and V2).

This study uses a specific procedure to generate 800 current signal samples (100 samples per event), with an MG model simulation time of 0.5 seconds: i) The HI fault model (Figure 2(a)) generates non-linear voltage-current samples by varying resistor values (RS1 and RS2) between 0.10 kΩ and 5.2 kΩ, and voltage sources (VS1 and VS2) between 0.5 kV and 10.2 kV in 0.3 s to 0.35 s. ii) Current signal samples for low impedance faults (LGF, LLGF, LLLGF, and LLF) are generated by adjusting resistance values (10-120% Ω) in 0.3-0.35 s steps. iii) Heavy loads (0.5 MW-2.4 MW) and capacitor banks (250 kVAR-700 kVAR) are turned on in stages to generate LST and CST signals (0.3 s switching).

2.1. High impedance fault transient model

The high impedance (HI) fault model is a circuit that mimics the properties of the Emanuel model [7], consisting of anti-parallel diodes, adjustable resistors, and DC source voltages [19]. It generates distinct non-linear voltage-current curves by varying resistance values and voltage levels. The HI fault voltage and current patterns within the microgrid network exhibit a nonlinear, asymmetrical, low-amplitude current waveform dominated by second- and third-order harmonics. The model's configuration and V-I characteristics of fault signal are illustrated in Figures 2(a) and 2(b), respectively.

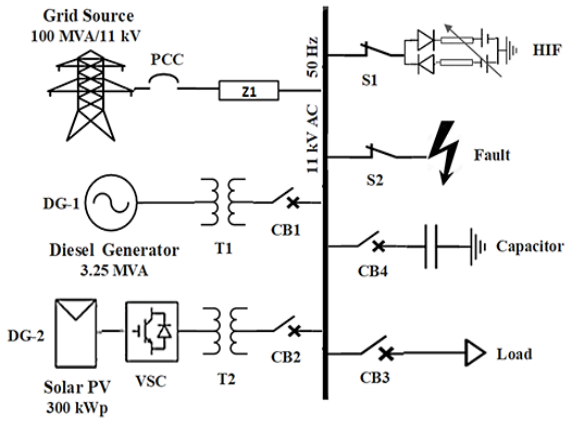
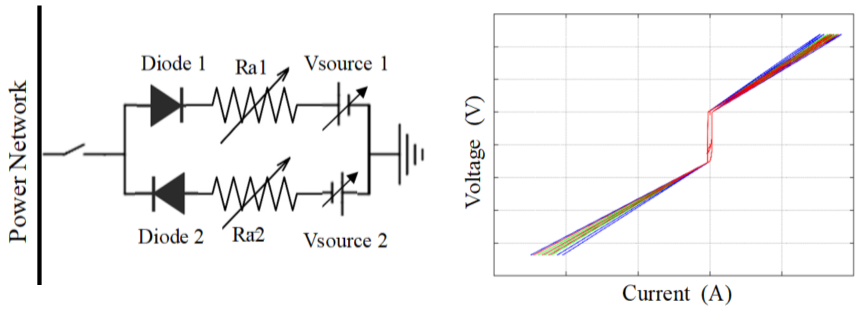


Figure 1. PV integrated MG model



(a) (b)

Figure 2. HI fault model: (a) basic configuration and (b) V-I curves

2.2. Signal processing with discrete wavelet transform (DWT)

DWT method is utilised in this work to decompose the HI fault and other transient signals to retrieve features of data set for learning the classifiers. The expression of DWT is written as (1) [24].

$$DWT(p,k) = \frac{1}{\sqrt{b_o^p}} \sum_m x(m) \times f \times \left(\frac{k - mab_o^p}{b_o^p}\right) \quad (1)$$

Where b_o^p is the scaling parameter, ab_o^p is the translation parameter, f is the mother wavelet function, p and m are integer variables, x(m) is time signal, and k is the number of samples of an input signal. (2)-(4) express co-efficient and energy value (EV) [22]. This study considers mother wavelet Daubechies 4 (db4) function and decomposition level (5th) for signal decomposition. The earlier study [23] discusses DWT approach.

$$A_i(m) = \sum_k x(m) \times LF1 \times (2m - k) \quad (2)$$

$$D_J(m) = \sum_k x(m) \times HF1 \times (2m - k) \quad (3)$$

$$EV = \sum_{J=1}^{N} [|D_J|^2] + |A_N|^2 \quad (4)$$

Where LF1 and HF1 stand for low and high pass filters, N denotes number of decomposition level, Ai and DJ represent the approximate and detailed coefficients, respectively.

3. MATERIALS AND METHOD OF CLASSIFICATION

The complete classification process, shown in Figure 3, includes various steps of analysis. First, the PV-integrated MG model is created, and simulation analysis is carried out with the introduction of different faults and transient events in the MATLAB/Simulink software environment. Also, using MATLAB/Simulink, using the DWT technique, the features of the dataset are extracted from the faults and transient events. The features are utilised to train the ML classifiers (NB, DTJ, KNN, and SEC) using an open-source tool, WEKA. The WEKA provides supervised and unsupervised ML methods, including grouping, visualisation, regression, and classification [23]. From the results of the confusion matrix, the classification accuracy of each classification method is estimated. The events like normal condition, LGF, LLGF, LLLGF, LLF, HI fault, LST, and CST in the network are classified in the class names CS1 to CS8. While training the classifiers, a k-fold cross validation technique is used to learn classifiers, which is more effective than the hold out data set method and can overcome overfitting issues with limited datasets [3], [23]. A 10-fold cross-validation strategy has been successful in evaluating classifier effectiveness, offering accurate approximations for classification accuracy across various tasks [3], [23]. This study uses a 10-fold cross-validation methodology to train classification models, focusing on proposed stacking (SEC) and individual classifiers (NB, DTJ, and KNN). Finally, the classification models effectively identify fault types based on these predictions. In case of a fault, the procedure is completed with a trip signal and passes to the protective system, or repeated in case of normal conditions. In addition, the materials, methods, and concepts of each ML algorithm employed in this study are delineated as follows:

3.1. Materials

The materials and tools used in this study are summarized as follows:
- MATLAB/Simulink (R2019b) software tool: used for MG model development/simulation and signal processing (wavelet tool box) with DWT approach.
- WEKA (V 3.9.6) open-source tool: machine learning models (for classification of events).
- Personal computer: Intel i5 CPU running at 2.4 GHz and 16 GB of RAM.
- Dataset: extracted features from the simulated current signals (faults/transients).

3.2. Methods

The overall methodology adopted in this study consists of the following stages:
- MG model simulation analysis: with introduction of faults and transients switching.
- Signal processing analysis: apply DWT method (decompose the signals to extract features from the evaluated wavelet coefficients).
- Classification analysis: train the classifier models (NB, DTJ, KNN, and SEC) and evaluate the results.
- System operation: generates the trip signal on fault detection and repeats the process in case of no fault conditions.

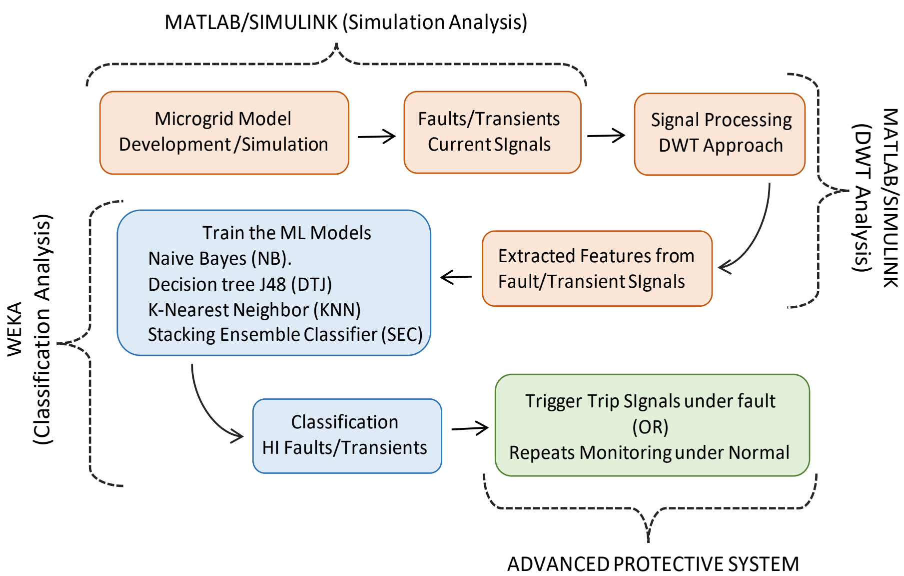


Figure 3. Block diagram of classification process

3.3. Classification method: Naïve bayes (NB)

The proficient NB classifier uses Bayes' probability theory for categorization [25], [26], renowned for its compute power, simplicity, and low variation [26]. It treats each characteristic independently during training, allowing for the computation of posterior probability during classification [26]. The output probability is the maximum argument for the given classification problem. The classifier's output probability can be regarded as the maximum argument (arg. max) for the given classification problem, written as (5) [27], [28].

$$\text{Max.} - Arg\left[\langle P_R(C_L|X_L)\rangle = \frac{P_R(C_L)\times\langle P_R(X_L|C_L)\rangle}{P_R(X_L)}\right] \quad (5)$$

Since the value of the PR(XL) attribute value rarely changes with class, (5) can be expressed as (6).

$$\text{Max.} - Arg[\langle P_R(C_L|X_L)\rangle = P_R(C_L) \times \langle P_R(X_L|C_L)\rangle] \quad (6)$$

When considering the likelihood PR(XL|CL) of variables in a data set (XL) is defined as [XL1, XL2, XL3, … XLN] for the class (CL) as [CL1, CL2, CL3, … CLK]. In the final assessment phase, the NB classifier labels the class of events based on the most probable output of each class value.

3.4. Classification method: decision tree J48 (DTJ)

The DTJ method with WEKA is a C4.5 decision tree approach that can handle large scale data sets with numerical and missing values [23], [29]. It consists of roots, branches, nodes, and leaves, with classification starting from the root node. The algorithm's working function is determined by entropy reduction. Each node receives test attribute values from learning instances, and a subset of these instances is used to build new nodes. The output attribute is added with the leaf generation, and nodes are created until all instances are used up [21]. The method uses a minimal confidence factor of 0.25 to prune the tree, considering 10 leaves and 18 tree sizes for the pruned tree.

3.5. Classification method: K-nearest neighbor (KNN)

This is a straightforward method, and learning from instances using the training set's closest to k values to categorize unknown sample distributions [30]. It uses distance metric functions like Euclidean, triangular, Gaussian, and Cosine to calculate distances between samples and determine nearest neighbors [31]. Instances are allocated to the majority class of the nearest k when there are multiple nearest neighbors. The maximum probability of XT1 being associated with class CS1 is expressed as (7) [32].

$$KNN(X_{T1}) = max\, P\,(C_{S1}, X_{T1}) \quad (7)$$

Where the probability of XT1 in class CS1 is denoted by P (CS1, XT1). The study uses a KNN classifier with multiple scanning techniques to find nearest neighbours. The optimal value of K is around 6, as it has been proven to produce the maximum classification accuracy.

3.6. Proposed classification method: stacking ensemble classifier (SEC)

The stacking ensemble model (Figure 4) uses DWT analysis to retrieve features from faults and transient signals, which are then used as input training datasets (DSET) for classifiers. In the first phase, single base classifiers like NB, DTJ, and KNN are learned using a 10 folds cross validation technique [30]. The KNN method is used at a meta level, as it outperforms NB and DTJ at the first level. In the second phase, the meta classifier KNN is trained using single base classifier predictions, with a K value of 6 and an Euclidean distance metric function. The meta-learner predicts class labels (CS1 to CS8) for various events. Table 1 illustrates the process steps of SEC.

Table 1. Process steps of stacking ensemble classifier

Step	Description
Input	Training dataset (D_{SET}) = {(x_{ia}, y_{io})}, where $x_{ia} \in \mathbb{R}^n$ and $y_{io} \in Y$
Output	Target output classes from the ensemble classifier E_P
(a)	Apply k-fold cross-validation (Kf = 10) and partition DSET into K equal subsets: DSET = {$DSET_1$, $DSET_2$, …, $DSET_k$}
(b)	Train individual base classifiers ACL_1, ACL_2, …, ACL_n using each subset. In this work: ACL_1 = NB, ACL_2 = DT, and ACL_3 = KNN (Stage 1)
(c)	Construct new training data from base classifier outputs: for each subset $X_{ja} \in DSET_k$, create instances (X'_{ja}, y_{jo}) where X'_{ja} = {$ACL_1(X_{ja})$, $ACL_2(X_{ja})$, …, $ACL_n(X_{ja})$}
(d)	Train the meta-learner (KNN) using the constructed dataset (Stage 2)
(e)	Generate final predicted class labels: $E_a(x) = L_a'(CS_1(x), CS_2(x), \ldots, CS_m(x))$, where M = 8

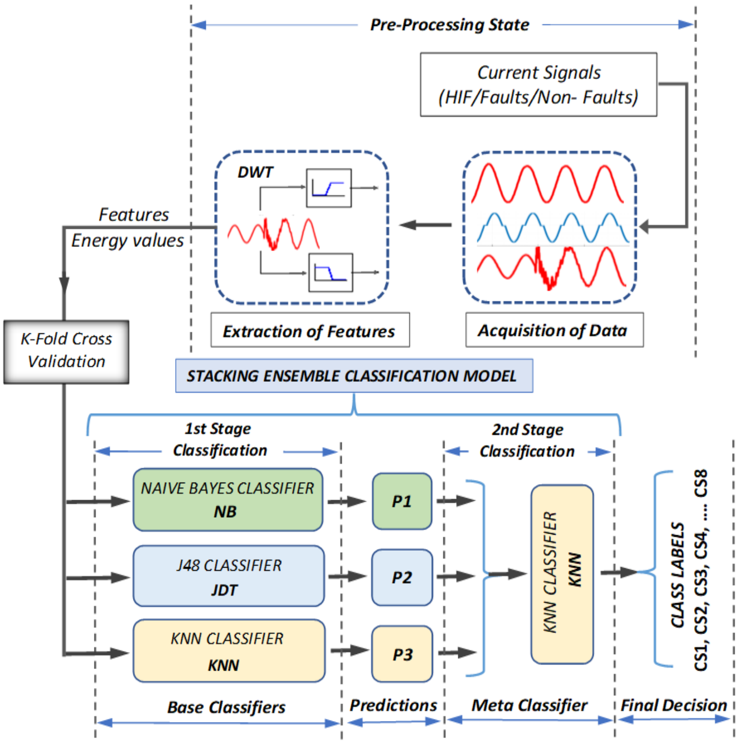


Figure 4. Proposed stacking ensemble classification model

4. RESULTS ANALYSIS AND DISCUSSION

The MG model was simulated with various class events, including normal, HIF, LGF, and LLGF (Figures 5(a) to 5(d)) and LLLGF, LLF, LST, and CST (Figures 6(a) to 6(d)). The results of decomposing signals of various events using the DWT approach are explained in section 4.1. Additionally, in section 4.2, the results of the classification analysis for the cases of PV-connected MG under standard test conditions and event signals exposed to noisy environments are discussed.

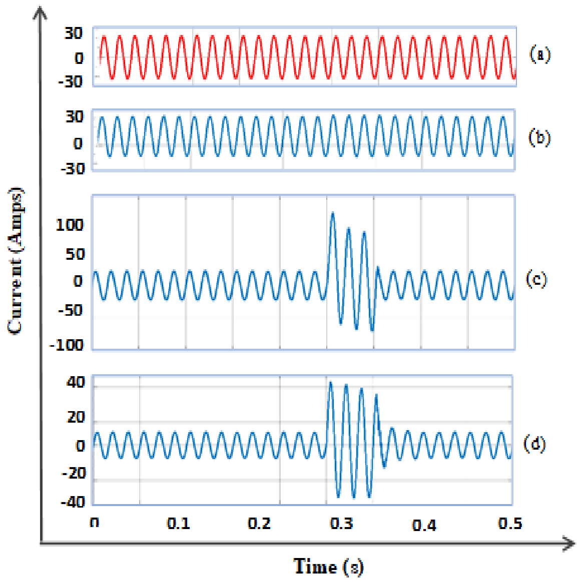


Figure 5. Signals of phase current: (a) normal, (b) HI fault, (c) LGF, and (d) LLGF

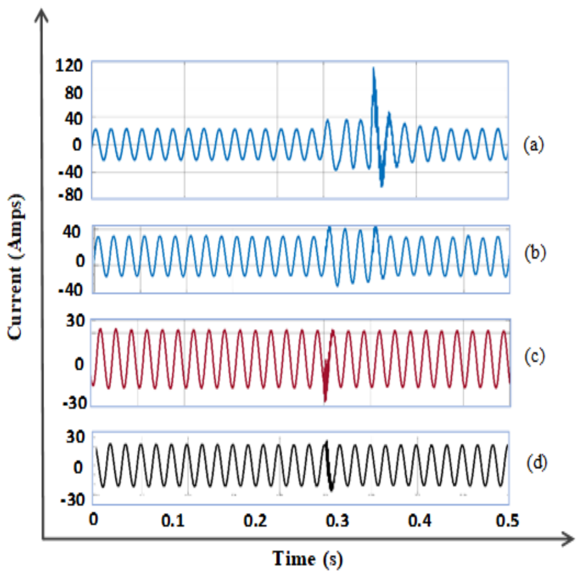


Figure 6. Signals of phase current: (a) LLLGF, (b) LLF, (c) CST, and (d) LST

4.1. Results of decomposed signals

This analysis generated wavelet coefficients by decomposing HI fault and other transient current signals at several levels. (4) assessed feature energy using recovered coefficients. SEC and other single base classifiers (NB, DTJ, and KNN) were trained utilising these features for classification. For decomposing signals, decomposition at 5th level, the mother wavelet function db4, and a sampling frequency of 24 kHz were taken into account. Figures 7(a) to 7(d) depict decomposed waveforms for normal, HI fault, LGF, and CST events in an islanded MG network, along with wavelet coefficient representations. These decompositions helped identify distinguishing features of each class event in the network. No spikes were observed in the waveform coefficients of the normal state event, but spikes were observed for LGF and CST events. LGF current magnitude was larger than the HI fault's small current. Similar decomposition procedures were applied to extract wavelet coefficients from all other fault and transient events.

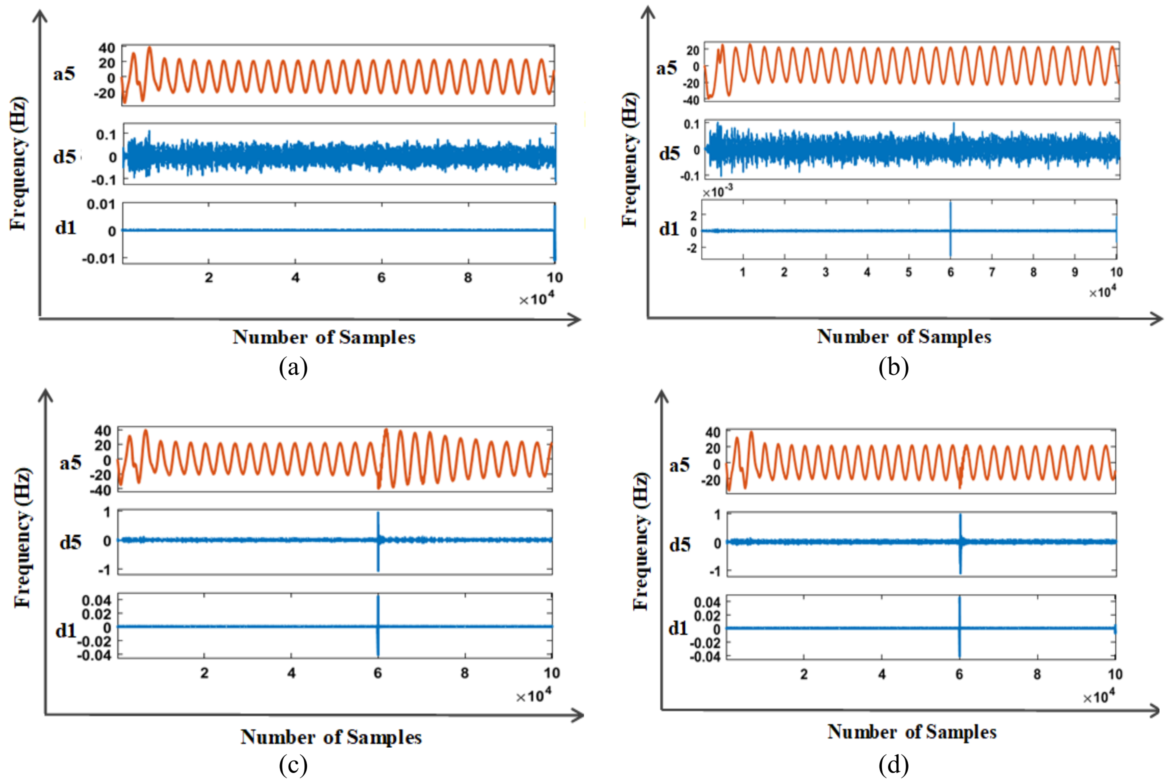


Figure 7. Results of decomposed signals: (a) normal state, (b) HI fault, (c) LGF, and (d) CST

4.2. Results of classification

The study used attributes from faults and transient signals to train classifiers using a 10-fold cross-validation technique [23]. The discrimination of HI fault from other faults and transients in PV generated MG networks was considered. The accuracy and HI fault success rate of ensemble (SEC) and individual (NB, DTJ, and KNN) classifiers were evaluated based on confusion matrix outcomes. The definition of accuracy and success rate of HI fault is expressed as (8) and (9) [3].

$$\text{Classification accuracy} = \frac{\text{Correctly classified instances}}{\text{Total number of instances}} \times 100\ \% \tag{8}$$

$$\text{HI fault succesive rate} = \frac{\text{Correctly classified instances of event}}{\text{Total number of instances of event}} \times 100\ \% \tag{9}$$

4.2.1. Classification analysis: in PV connected MG network (under standard test condition)

This analysis distinguished between faults and transients in PV-generated MG power system. From the results of confusion matrix (Tables 2 to 5), classification accuracy, HI fault success rate, and performance indices were assessed. According to the results, the NB classifier incorrectly classified 140 instances, compared to 80 and 70 for DTJ (90% and 82%) and KNN, respectively. This resulted in lower accuracy and HI fault success rate compared to DTJ (91.25% and 78%) and KNN (91.25% and 88%) classifiers. Stacking classifier had higher accuracy (96.25%) and HI fault success rate (92%) than base classifiers. The stacking classifier had 30 misclassified instances, the fewest. Overall, the proposed stacking model surpasses base classification methods with respect to accuracy and rate of success in detecting HI Fault in PV-connected MG networks.

Table 2. Classification results: NB classifier

Events	Classes	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	Incorrectly classified	Correctly classified
Normal	CS1	100	0	0	0	0	0	0	0	0	100
LGF	CS2	12	88	0	0	0	0	0	0	12	88
LLGF	CS3	0	0	82	10	0	0	8	0	18	82
LLLGF	CS4	4	12	4	80	0	0	0	0	20	80
LLF	CS5	10	10	0	0	80	0	0	0	20	80
HI fault	CS6	0	0	0	0	0	80	0	20	20	80
LST	CS7	0	0	12	0	0	0	62	26	38	62
CST	CS8	0	12	0	0	0	0	0	88	12	88
Overall accuracy = 82.5% and success rate of HI fault = 80%											

Table 3. Classification results: DTJ classifier

Events	Classes	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	Incorrectly classified	Correctly classified
Normal	CS1	100	0	0	0	0	0	0	0	0	100
LGF	CS2	0	96	4	0	0	0	0	0	4	96
LLGF	CS3	0	10	90	0	0	0	0	0	10	90
LLLGF	CS4	0	0	4	96	0	0	0	0	4	96
LLF	CS5	0	12	0	0	88	0	0	0	12	88
HI fault	CS6	0	0	0	0	0	82	8	10	18	82
LST	CS7	0	0	0	0	0	10	80	10	20	80
CST	CS8	0	0	0	0	0	0	12	88	12	88
Overall accuracy = 90% and success rate of HI fault = 82%											

Table 4. Classification results: KNN classifier

Events	Classes	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	Incorrectly classified	Correctly classified
Normal	CS1	100	0	0	0	0	0	0	0	0	100
LGF	CS2	0	96	4	0	0	0	0	0	4	96
LLGF	CS3	0	12	88	0	0	0	0	0	12	88
LLLGF	CS4	0	12	0	88	0	0	0	0	12	88
LLF	CS5	0	0	12	0	88	0	0	0	12	88
HI Fault	CS6	0	0	0	0	4	88	0	8	12	88
LST	CS7	0	0	0	0	0	0	90	10	10	90
CST	CS8	0	0	0	0	4	0	4	92	8	92
Overall accuracy = 91.25% and success rate of HI fault = 88%											

Table 5. Classification results: SEC classifier

Events	Classes	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	Incorrectly classified	Correctly classified
Normal	CS1	100	0	0	0	0	0	0	0	0	100
LGF	CS2	0	100	0	0	0	0	0	0	0	100
LLGF	CS3	0	0	100	0	0	0	0	0	0	100
LLLGF	CS4	0	0	0	100	0	0	0	0	0	100
LLF	CS5	0	0	10	0	90	0	0	0	10	90
HI Fault	CS6	0	0	0	0	0	92	4	4	8	92
LST	CS7	0	0	0	0	0	8	92	0	8	92
CST	CS8	0	0	0	0	4	0	0	96	4	96
Overall accuracy = 96.25% and success rate of HI fault = 92%											

4.2.2. Results of performance analysis

To assess the classifier effectiveness in further level, performance measures (PM) like precision (PR), recall (RC), F-measure (FMS), and receiver operating characteristics (ROCS) were assessed. Performance measures are detailed in earlier works [22]. Table 6 illustrate performance parameters (PR, RC, FMS, and ROCS) results of all classifiers. Individual base classifier PR and RC values vary from 0.886 to 0.929 and 0.896 to 0.913. Individual base classifier FMS and ROCS scores range from 0.898 to 0.907 and 0.9 to 0.976. In comparison to NB and DTJ, the KNN exhibits promising results and outperforms. Additionally, the proposed stacking classifier surpasses individual base classifiers in PR, RC, FMS, and ROCS, with values above 0.96 and 0.98, respectively.

4.2.3. Classification analysis with noisy event signals

The power system's electrical signals are often noisy. To realize disruption in noisy environments, power system event signals are analysed with the addition of white Gaussian noise (WGN) [33]. The study used noise levels with signal-to-noise ratio (SNR) values between 20 and 40 dB. The SNR is computed as (10) [33].

$$SNR(db) = 10\,log\left(\frac{FS}{FN}\right) \tag{10}$$

The proposed ensemble classifier's classification accuracy at different noise levels is shown in Table 7. The classifier's accuracy and HI Fault success rate ranged from 91% to 94.75% at 20 dB to 40 dB, and 84% to 90% at 20 dB of SNR. However, at higher noise levels, both accuracy and HI Fault success rate were lower. Despite the noisier environment, the ensemble classifier still showed good results with respect to accuracy (91%) and rate of success (84%) in detecting HI fault. The proposed ensemble classifier demonstrates superior precision and robustness in the presence of elevated noise levels with event signals.

Table 6. Results of performance analysis (under STC)

Classifiers	PR	RC	FMS	ROCS
Weighted average metrics				
NB	0.886	0.896	0.898	0.900
DTJ	0.909	0.9	0.9	0.962
KNN	0.929	0.913	0.907	0.976
SEC	0.966	0.963	0.962	0.984
Results of metrics with HI fault event				
NB	0.875	0.612	0.778	0.841
DTJ	0.878	0.622	0.786	0.982
KNN	1	0.7	0.867	1
SEC	0.909	1	0.952	1

Table 7. Results of classification under noisy environment

Class events	No Noise	20 dB		30 dB		40 dB	
	Accuracy (%)	Mis-classified instances	Accuracy (%)	Mis-classified instances	Accuracy (%)	Mis-classified instances	Accuracy (%)
CS1	100	5	95	3	97	1	99
CS2	100	4	96	2	98	2	98
CS3	100	4	96	3	97	1	99
CS4	100	6	94	3	97	1	99
CS5	90	15	85	12	88	11	89
CS6	92	16	84	12	88	10	90
CS7	92	13	87	11	89	10	90
CS8	96	9	91	8	92	6	94
Overall accuracy	96.25%	91%		93.25%		94.75%	
Success rate of HI fault	92%	84%		88%		90%	

5. CONCLUSION

This study proposes a stacking ensemble classifier to differentiate between HI faults and transients in a PV-generated Microgrid power system. The DWT approach is employed in the pre-stage of data analysis to extract training set features from transients and fault events. The proposed stacking model has two steps: learning individual classifiers (NB, DTJ, and KNN) using a 10-fold cross-validation method and learning the meta-learner (KNN) to obtain targeted class labels. The outcomes from this study are as follows: i) According to the outcomes of the analysis in the MG network (at STC), the suggested model outperforms NB, DTJ, and KNN with respect to accuracy (96.25%) and rate of success (92%) in detecting HI faults. ii) The suggested stacking model achieves 91% accuracy and 84% success rate in detecting HI faults, even in noisy event data. iii) The performance measures (PR: 0.966; RC: 0.963; FMS: 0.962; and ROCS: 0.984) prove that the suggested stacking model performs well in identifying faults in MG power systems.

Thus, the ensemble classifier's consistent results demonstrate its adaptability and usefulness in practical situations. Adopting suggested ensemble classifier could improve fault detection systems in noisy and dynamic environments, enhancing dependability and operational efficiency. Future study will investigate the application of sophisticated deep learning techniques to differentiate between HI faults and other transients in multi-MG power systems.

FUNDING INFORMATION

The authors state no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Arangarajan Vinayagam	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Raman Mohandas		✓				✓		✓	✓	✓	✓	✓		
Meyyappan Chindamani	✓		✓	✓			✓			✓	✓			✓
Bhadravathi Gavirangapa Sujatha	✓		✓			✓		✓		✓			✓	
Soumya Mishra					✓		✓			✓		✓		✓
Arivoli Sundaramurthy		✓		✓	✓		✓		✓		✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [AV], upon reasonable request.

REFERENCES

[1] N. Hatziargyriou, Ed., *Microgrids: architectures and control*. Hoboken, NJ, USA: John Wiley & Sons, 2014.
[2] M. Manohar, E. Koley, and S. Ghosh, "Reliable protection scheme for PV integrated microgrid using an ensemble classifier approach with real-time validation," *IET Science, Measurement & Technology*, vol. 12, no. 2, pp. 200–208, Mar. 2018, doi: 10.1049/iet-smt.2017.0270.
[3] K. S. V. Swarna, A. Vinayagam, M. B. J. Ananth, P. V. Kumar, V. Veerasamy, and P. Radhakrishnan, "A KNN based random subspace ensemble classifier for detection and discrimination of high impedance fault in PV integrated power network," *Measurement*, vol. 187, no. 1, p. 110333, Jan. 2022, doi: 10.1016/j.measurement.2021.110333.
[4] A. Ghaderi, H. A. Mohammadpour, H. L. Ginn, and Y.-J. Shin, "High-impedance fault detection in the distribution network using the time-frequency-based algorithm," *IEEE Transactions on Power Delivery*, vol. 30, no. 3, pp. 1260–1268, Jun. 2014, doi: 10.1109/TPWRD.2014.2361207.
[5] B. K. Chaitanya, A. Yadav, and M. Pazoki, "An intelligent detection of high-impedance faults for distribution lines integrated with distributed generators," *IEEE Systems Journal*, vol. 14, no. 1, pp. 870–879, Mar. 2020, doi: 10.1109/JSYST.2019.2911529.
[6] S. Wang and P. Dehghanian, "On the use of artificial intelligence for high impedance fault detection and electrical safety," *IEEE Transactions on Industry Applications*, vol. 56, no. 6, pp. 7208–7216, Nov. 2020, doi: 10.1109/TIA.2020.3017698.
[7] V. Veerasamy *et al.*, "A novel discrete wavelet transform-based graphical language classifier for identification of high-impedance fault in distribution power system," *International Transactions on Electrical Energy Systems*, vol. 30, no. 6, Jun. 2020, doi: 10.1002/2050-7038.12378.
[8] K. Sekar and N. K. Mohanty, "A fuzzy rule base approach for high impedance fault detection in distribution system using morphology gradient filter," *Journal of King Saud University - Engineering Sciences*, vol. 32, no. 3, pp. 177–185, Mar. 2020, doi: 10.1016/j.jksues.2018.12.001.
[9] M. Sarwar, F. Mehmood, M. Abid, A. Q. Khan, S. T. Gul, and A. S. Khan, "High impedance fault detection and isolation in power distribution networks using support vector machines," *Journal of King Saud University - Engineering Sciences*, vol. 32, no. 8, pp. 524–535, Dec. 2020, doi: 10.1016/j.jksues.2019.07.001.
[10] I. Baqui, I. Zamora, J. Mazón, and G. Buigues, "High impedance fault detection methodology using wavelet transform and artificial neural networks," *Electric Power Systems Research*, vol. 81, no. 7, pp. 1325–1333, Jul. 2011, doi: 10.1016/j.epsr.2011.01.022.
[11] V. Veerasamy *et al.*, "High-impedance fault detection in medium-voltage distribution network using computational intelligence-based classifiers," *Neural Computing and Applications*, vol. 31, no. 12, pp. 9127–9143, Dec. 2019, doi: 10.1007/s00521-019-04445-w.
[12] Q. Cui, K. El-Arroudi, and Y. Weng, "A feature selection method for high impedance fault detection," *IEEE Transactions on Power Delivery*, vol. 34, no. 3, pp. 1203–1215, Jun. 2019, doi: 10.1109/TPWRD.2019.2901634.

[13] M. Mishra and P. K. Rout, "Detection and classification of micro-grid faults based on HHT and machine learning techniques," *IET Generation, Transmission & Distribution*, vol. 12, no. 2, pp. 388–397, Jan. 2018, doi: 10.1049/iet-gtd.2017.0502.
[14] M. Kavi, Y. Mishra, and M. Vilathgamuwa, "Challenges in high impedance fault detection due to increasing penetration of photovoltaics in radial distribution feeder," in *2017 IEEE Power & Energy Society General Meeting*, Jul. 2017, pp. 1–5, doi: 10.1109/PESGM.2017.8274658.
[15] G. Niu, T. Han, B.-S. Yang, and A. C. C. Tan, "Multi-agent decision fusion for motor fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 3, pp. 1285–1299, Apr. 2007, doi: 10.1016/j.ymssp.2006.03.003.
[16] A. Eskandari, J. Milimonfared, and M. Aghaei, "Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on I-V characteristics," *Solar Energy*, vol. 211, pp. 354–365, 2020, doi: 10.1016/j.solener.2020.09.071.
[17] P. K. Mishra, A. Yadav, and M. Pazoki, "A novel fault classification scheme for series capacitor compensated transmission line based on bagged tree ensemble classifier," *IEEE Access*, vol. 6, pp. 27373–27382, 2018, doi: 10.1109/ACCESS.2018.2836401.
[18] P. Balakrishnan and S. Gopinath, "A new intelligent scheme for power system faults detection and classification: a hybrid technique," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 33, no. 5, Sep. 2020, doi: 10.1002/jnm.2728.
[19] S. R. Samantaray, "Ensemble decision trees for high impedance fault detection in power distribution network," *International Journal of Electrical Power & Energy Systems*, vol. 43, no. 1, pp. 1048–1055, Dec. 2012, doi: 10.1016/j.ijepes.2012.06.006.
[20] R. Azizi and S. Seker, "Microgrid fault detection and classification based on the boosting ensemble method with the Hilbert-Huang transform," *IEEE Transactions on Power Delivery*, vol. 37, no. 3, pp. 2289–2300, Jun. 2022, doi: 10.1109/TPWRD.2021.3109023.
[21] M. M. Islam, M. U. Usman, A. Newaz, and M. O. Faruque, "Ensemble voting-based fault classification and location identification for a distribution system with microgrids using smart meter measurements," *IET Smart Grid*, vol. 6, no. 3, pp. 219–232, 2023, doi: 10.1049/stg2.12091.
[22] A. Vinayagam, V. Veerasamy, M. Tariq, and A. Aziz, "Heterogeneous learning method of ensemble classifiers for identification and classification of power quality events and fault transients in wind power integrated microgrid," *Sustainable Energy, Grids and Networks*, vol. 31, 2022, doi: 10.1016/j.segan.2022.100752.
[23] P. Radhakrishnan, K. Ramaiyan, A. Vinayagam, and V. Veerasamy, "A stacking ensemble classification model for detection and classification of power quality disturbances in PV integrated power network," *Measurement: Journal of the International Measurement Confederation*, vol. 175, 2021, doi: 10.1016/j.measurement.2021.109025.
[24] H. Dehghani, B. Vahidi, R. A. Naghizadeh, and S. H. Hosseinian, "Power quality disturbance classification using a statistical and wavelet-based hidden Markov model with Dempster–Shafer algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 47, pp. 368–377, May 2013, doi: 10.1016/j.ijepes.2012.11.005.
[25] A. H. Aljemely, J. Xuan, L. Xu, F. K. J. Jawad, and O. Al-Azzawi, "Wise-local response convolutional neural network based on Naïve Bayes theorem for rotating machinery fault classification," *Applied Intelligence*, vol. 51, no. 10, pp. 6932–6950, Oct. 2021, doi: 10.1007/s10489-021-02252-2.
[26] S. Suganthi, A. Vinayagam, V. Veerasamy, A. Deepa, M. Abouhawwash, and M. Thirumeni, "Detection and classification of multiple power quality disturbances in microgrid network using probabilistic based intelligent classifier," *Sustainable Energy Technologies and Assessments*, vol. 47, p. 101470, Oct. 2021, doi: 10.1016/j.seta.2021.101470.
[27] D. Michaelson, H. Mahmood, and J. Jiang, "A predictive energy management system using pre-emptive load shedding for islanded photovoltaic microgrids," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 7, pp. 5440–5448, Jul. 2017, doi: 10.1109/TIE.2017.2677317.
[28] N. Zhang, L. Wu, J. Yang, and Y. Guan, "Naive Bayes bearing fault diagnosis based on enhanced independence of data," *Sensors (Switzerland)*, vol. 18, no. 2, 2018, doi: 10.3390/s18020463.
[29] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*. Hoboken, NJ, USA: John Wiley & Sons, 2011.
[30] T. Ngo, "Data mining: practical machine learning tools and technique, by ian h. witten, eibe frank, mark a. hell," *ACM SIGSOFT Software Engineering Notes*, vol. 36, no. 5, pp. 51–52, 2011.
[31] O. Kherif, Y. Benmahamed, M. Teguar, A. Boubakeur, and S. S. M. Ghoneim, "Accuracy improvement of power transformer faults diagnostic using KNN classifier with decision tree principle," *IEEE Access*, vol. 9, pp. 81693–81701, 2021, doi: 10.1109/ACCESS.2021.3086135.
[32] M. Manohar, E. Koley, Y. Kumar, and S. Ghosh, "Discrete wavelet transform and kNN-based fault detector and classifier for PV integrated microgrid," *Advances in Data and Information Sciences: Proceedings of ICDIS-2017, Volume 1*, 2018, pp. 19–28, doi: 10.1007/978-981-10-8360-0_2.
[33] O. Jeba Singh *et al.*, "Robust detection of real-time power quality disturbances under noisy condition using FTDD features," *Automatika*, vol. 60, no. 1, pp. 11–18, Jan. 2019, doi: 10.1080/00051144.2019.1565337.

BIOGRAPHIES OF AUTHORS

Arangarajan Vinayagam has over 15 years of industrial experience in the power plant and renewable energy industries. He also has academic experience (teaching and research) in smart grid and microgrid power systems, renewable energy technologies, power generation, and distribution control. His research interests include microgrids, electric vehicle, machine learning, power quality, forecasting, and fault analysis of solar PV and wind generation. He received Ph.D. (in the field of PV integrated microgrid power systems) from Deakin University, Australia, in 2017. He is currently working as a visiting professor and research fellow (remote) at the Institute of Power Engineering, UNITEN, Malaysia. He can be contacted at email: arajanin@gmail.com.

Raman Mohandas received his Ph.D. degree in 2020 from Anna University, Chennai, India. He is currently working as an associate professor in the Department of Electrical and Electronics Engineering at New Horizon College of Engineering, Bangaluru, India. His research interests include microgrids, PWM techniques, power electronics, and renewable energy sources. He can be contacted at email: mohandas@newhorizonindia.edu.

Meyyappan Chindamani is currently serving as an assistant professor in the Department of Electrical and Electronics Engineering at Sri Ramakrishna Engineering College, Coimbatore. She earned her Ph.D. in electrical engineering from Anna University, Chennai. Her research interests encompass power converters, electrical machines, power systems, data science, and artificial intelligence. She can be contacted at email: chindamani.meyyappan@srec.ac.in.

Bhadravathi Gavirangapa Sujatha completed Ph.D. degree in electrical engineering and electronics engineering from the Jain University, Bangalore, India in 2021. She is currently working as a professor and HoD in the Department of Electrical and Electronics Engineering at Cambridge Institute of Technology, Bengaluru, India. Her research interests include solar PV system, power system, electrical machines, and electric vehicles. She can be contacted at email: sujatha.eee@cambridge.edu.in.

Soumya Mishra received his Ph.D. degree in electrical engineering from the National Institute of Technology Rourkela, Odisha, India, in 2018. He is currently an associate professor with the School of Electrical Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Bhubaneswar. His research interests include microgrids, grid integration of renewable energy systems, and electric vehicles. He can be contacted at email: soumya.mishra@kiit.ac.in.

Arivoli Sundaramurthy received his Ph.D. degree in electrical engineering from Anna University, Chennai, India, in 2024. He is currently an assistant professor in the Department of Electrical and Electronics Engineering at PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India. His research interests include artificial intelligence, machine learning, microgrid optimization, stability analysis, and image processing. He can be contacted at email: arivoli@psgitech.ac.in.