WILEY | Hindawi

Research Article

Islanding Detection and Power Quality Diagnosis of Wind Power Integrated Microgrid with Reduced Feature Trained Novel Optimized Random Decision Forest

Sairam Mishra [1], **Ranjan K. Mallick** [2], **Debadatta A. Gadanayak** [2], **Pravati Nayak** [1]
Aymen Flah [3,4,5,6,7,8], **Claude Ziad El-Bayeh** [9], **Habib Kraiem** [10], and **Lukas Prokop** [7]

[1]Department of Electrical Engineering, Siksha 'O' Anushandhan Deemed to Be University, Bhubaneswar, India
[2]Department of Electrical and Electronics Engineering, Siksha 'O' Anushandhan Deemed to Be University, Bhubaneswar, India
[3]Energy Processes Environment and Electrical Systems Unit, National Engineering School of Gabès, University of Gabes, Gabès 6029, Tunisia
[4]University of Business and Technology (UBT), College of Engineering, Jeddah 21448, Saudi Arabia
[5]MEU Research Unit, Middle East University, Amman 11831, Jordan
[6]University of Gabès, Private Higher School of Applied Sciences and Technology of Gabès, Gabès 6029, Tunisia
[7]ENET Centre, VSB-Technical University of Ostrava, Ostrava, Czech Republic
[8]Applied Science Research Center, Applied Science Private University, Amman 11931, Jordan
[9]Department of Electrical Engineering, Bayeh Institute, 55 Kfar Saleh-Hay El Arbe Street, Amchit, Mount Lebanon, Lebanon
[10]Department of Electrical Engineering, College of Engineering, Northern Border University, Arar, Saudi Arabia

Correspondence should be addressed to Habib Kraiem; habib.kraiem@yahoo.fr

Distributed generations (DGs) have been increasingly addressing the ongoing power deficit in the electricity market. However, a significant concern in DG-integrated microgrids is the detection of accidental islanding. To tackle this issue, this article proposes a cost-friendly, novel data-driven passive islanding detection scheme named EEMD-HOBRC, combining noise-assisted ensemble empirical mode decomposition (EEMD) and a hybrid optimization-based random forest classifier (HOBRFC). The detection scheme employs a diverse set of features extracted from both raw and EEMD decomposed signals. Essential features are selected using the binary grey wolf optimizer (BGWO) to reduce computational burden. To further improve classification accuracy, the parameters of the random forest classifier are optimized through a hybrid particle swarm and reformed grey wolf optimization (PSRGWO) technique with Cohen's kappa index as the cost function. The proposed technique is rigorously validated in two different multi-DG environments, encompassing islanding and various nonislanding events. The results demonstrate the effectiveness of the approach in terms of enhanced accuracy, detection time, and performance under both noisy and noise-free conditions. The accuracy of detection under ideal and high noise scenarios is found to be 99.88% and 99.2%, respectively, with maximum detection time of 34.27 ms. Comparative analysis with other algorithms also supports the superiority of the proposed technique. Finally, the method is successfully applied to shrink the nondetection zone (NDZ) with minimal power mismatch, further enhancing its utility in practical applications.

1. Introduction

Distributed generators (DGs) are proving to be the best solution to match the ongoing demand for electricity. Since every solution is equipped with new problem(s), the DGs are equipped with the challenge of unintentional islanding. In simple terms, islanding is such a scenario where the DG(s) of an area is/are kept on supplying power to the nearby loads even though the central grid is no longer connected. This occurrence should be detected within 2 sec as

per IEEE 1547 [1] norms. The conventional strategies of islanding detection [2–5] have a common concern that is reducing the nondetection zone (NDZ). By definition, NDZ is a power mismatch region where the islanding is remained undetected due to supply-demand equilibrium. Several research works have already proposed to conquer the NDZ issue by applying signal processing approach (SPA) to the conventional passive detection techniques [6, 7]. But still, these have issues like computational complexity, affected by noise interference, and unapplicable to nonlinear and nonstationary signals. Hence, the necessity of further research is essential in this particular field. Eventually, mode decomposition- (MD-) oriented SPAs are gaining popularity that brings forward empirical MD (EMD) [8], variational MD (VMD) [9], and empirical wavelet transform (EWT) [10] into the limelight. But these MD techniques basically suffer from mode mixing issues. This major drawback is wonderfully eliminated by a recently proposed noise-assisted technique called ensemble EMD (EEMD) [11] by taking the mean of an aggregate of trials that comprises of signal and finite amplitude white noise. It has also displayed significant results in detecting differential protection [12], load forecasting [13], wind speed prediction [14], etc. Thus, EEMD is chosen as the base SPA in this work.

Disturbance detection only with SPA brings thresholding challenges in front. Setting such critical thresholds is a tough job. A suitable alternative is to go for machine learning classifiers (MLCs) that detect the class of event by training-testing approach [15]. Such an application utilizes the inherent features of a signal. Since there are lot of simple as well as complex features can be obtained from any signal, it is important to look for an effective feature selection technique (FST) so as to reduce the time complexity and thereby increasing the efficacy of the MLC with least number of features. Many researchers have already suggested several FSTs where the feature selection process is considered as a combinatorial problem [16]. But the classical FSTs have some flaws which can be overcame by evolutionary optimization techniques [17–20] those can be of two types, namely, continuous and discrete (binary). As the decision variable of feature selection varies in binary space, a binary optimization algorithm can be chosen to deal with this [21–23]. But it is essential to look for an algorithm that maintains a proper balance between exploration and exploitation. One of the recent bioinspired techniques is grey wolf optimization (GWO) that is gaining popularity due to its smooth transition between exploration and exploitation [24, 25]. A binary variant of GWO has been proposed by Ahmadipour et al.[26] which has shown significant result for feature selection problems. That is why BGWO has been chosen for feature selection. Once the features are obtained, then it is required to look for a classifier. A number of MLCs are already implemented in several research works starting from consensus-based intelligent control [27], machine faults [28–30], and islanding detection [31–35], which have their own uniqueness and limitations. Eventually, ensemble machine learning paradigms are stepping ahead with the concept of combining more than one learner together to bolster the classification rate. One of such learners is random

forest (RF) which ensembles the decision trees (DTs). RF possesses the advantages like well performance with less parameters, hardly affected by fitting issues, and has better generalizing capability [36]. Furthermore, the efficacy of RF is tested under public data [37], and it has observed that RF is displaying significant result among the MLCs. Hence, RF is considered as the base classifier in this study. But before implementing RF classifier, it is necessary to look through two major factors; those are DTs inside forest and cardinality of feature subset. These two deciding factors also called hyperparameters can be evaluated by implementing any optimization algorithm. A hybrid optimization algorithm combining the social behaviour like swarm etiquettes and grey wolf hunting is proposed in this study to extremize the major hyperparameters of RF which can enhance the robustness and classification capability of the classifier. Moreover, for solving any optimization problem, it is required to define an objective function. Here, the objective function can be the classification accuracy or some other criterion that might consider the best trade-off between attribute extraction, computational burden, and efficiency. For instant, evaluating the performance of any classifier only considering the number of misclassifications may be deceitful [38], if there is prediction failure obtained for one or more features. Hence, a different statistical measure is suggested in this study for assessing the classification accuracy called Cohen's kappa. The kappa index is generally a measure of inter-rater agreement that indemnifies the possibility of classification by chance [39]. Further, it is unrealistic to rate all the error as identical. For example, an islanding event if misclassified as capacitor switching is much more destructive than a transformer switching misclassified as capacitor switching. Thus, for cost sensitive classifications, the preliminary job is to ensure that the errors are free of constraints, assumptions, and biasing which can be achieved by estimating Cohen's kappa.

Furthermore, integration of multiple DG units with the exiting grid is creating added complexity, which is to be addressed. The generated voltage out of any DG unit can be AC or DC depending on the nature of the source. It has to be converted first to the nominal voltage and frequency with the nominal parameters before feeding to the grid. For that reason, transformers and/or inverters are usually used while DG-GRID integration. Hence, DGs can be classified into two major categories such as inverter-based DG and rotating machine DG which have separate impact individually on the associated system. For an instant, when a wind power generation unit is disconnected from grid, the voltage of the system will swell up for few cycles due to the sudden drop in reactive power of the system. On the contrary, the outage of a PV unit certainly brings some fluctuation in frequency along with voltage amplitude that resembles a flicker. Hence, it is highly desired to detect the PQ events arising because of these DGs along with other PQ events and islanding [40].

Briefly, the work focuses on islanding and power quality disturbance classification from the BGWO reduced feature dataset that fed into the novel particle swarm and reformed grey wolf optimization (PSRGWO) random forest classifier.

Further, the proposed approach is validated with two different multi-DG systems under ideal and noisy conditions. The descriptive information of all the concepts is discussed in the Section 2. Section 3 is dedicated to microgrid system under consideration whereas Section 4 explains about feature extraction and data collection. The methodology part is covered in Section 5, and the results are analyzed in Section 6. At last, the study is concluded in Section 7.

The major highlights of this paper can be as follows:

(i) Application of EEMD on extracted signals to decompose into different IMFs and most informative IMF is chosen to extract multiple features, and some of the features are extracted from raw signal

(ii) Maidan application of binary grey wolf optimization to select optimal informative features

(iii) A novel hybrid PSRGWO optimization algorithm has been formulated to optimize the parameters of the proposed random forest classifier, i.e., termed as hybrid optimization-based random forest classifier (HOBRFC)

(iv) Kappa statics is used as objective function for the optimization problem to eradicate the possibility of classification by chance

(v) The proposed method is used for classifying the islanding and other power quality disturbances

(vi) The efficacy of the proposed classifier is tested for two different test systems for ideal and noisy scenarios

(vii) Finally, the reduction of nondetection zone is validated with a separate dataset of very small power mismatch

2. Theoretical Background

2.1. Binary Grey Wolf Optimization. The grey wolves are preferring to stay in a pack by maintaining a social hierarchy where the alpha wolves lead the pack followed by beta and delta wolves. The movement of the wolves in GWO algorithm took place stochastically in the search space. There are problems, whose solutions are restricted to binary 0/1 or true/false type that inspires to go for a special variant of optimization technique, i.e., binary optimization algorithms. Thus, BGWO comes into picture as an FST. In BGWO, the conventional grey wolf optimization is carried out initially and the position updating equation is forced to be binary.

The wolves try to encircle their target first which can be mathematically represented as

$$\vec{S} = \left| \vec{Z} \vec{T}(j) - \vec{W}(j) \right|,$$

$$\vec{W}(j+1) = \vec{T} + \vec{Y}.\vec{S}, \tag{1}$$

where j is the iteration number, $\vec{S}$ is the distance vector, $\vec{T}$ is the position vector of target, $\vec{W}$ is the position vector of grey wolf, and $\vec{Y}$ and $\vec{Z}$ are the coefficient vectors with values.

$$\vec{Y} = 2g.\vec{r_1} - g,$$

$$\vec{Z} = 2\vec{r_2}. \tag{2}$$

These coefficients are aiding to the moment of the wolves around the search space whereas g is the factor that keep on decaying to shift the wolves closer to the prey (target) in each iteration as per (3), and $\vec{r_1}$ and $\vec{r_2}$ have random values between 0 and 1.

$$g = 2 - \frac{2}{\text{Total}_{\text{Iterations}}} * j. \tag{3}$$

Moreover, the hierarchy assumption takes three best solutions, and others try to update their positions according to the top three. The updating equation is

$$W(j+1) = \frac{W_1 + W_2 + W_3}{3}. \tag{4}$$

Here, $\vec{W_1} = \vec{W_\alpha} - \vec{Y_1}.(\vec{S_\alpha})$, $\vec{W_2} = \vec{W_\beta} - \vec{Y_2}.(\vec{S_\beta})$, and $\vec{W_3} = \vec{X_\delta} - \vec{Y_3}.(\vec{S_\delta})$.

Now to binaries this problem, the result obtained in (4) undergoes a sigmoidal function validation as follows:

$$W_b(j+1) = \begin{cases} 1, & \text{if sigmoid}(W(j+1)) \geq \text{rand}, \\ 0, & \text{or else}, \end{cases} \tag{5}$$

where sigmoid of any quantity x can be defined as

$$\text{Sig}(x) = \frac{1}{1 - e^{-10(x-(1/2))}}. \tag{6}$$

2.2. Ensemble Empirical Mode Decomposition (EEMD). Mode decomposition is eventually gaining enormous interest both in the community of researchers as well as industrialists. EMD is the start point of such signal preprocessors and has already used in wide range of applications. The descriptive explanation of EMD can be found in [14]. It simply divides the original signal into several intrinsic mode functions. But the major problem arises here is the mixing of modes which can be eliminated by implementing a noise-assisted method called EEMD that collaborate the different scale signals to their respective band of IMFs.

In general, every recorded data is impure and an amalgam mixture of signal and noise which arises question about the clarity of signal measurement. A noise-assisted method assumes that the collected data are separately measured observations by adding different series of noises to it. To realize this conceptualization, let us presume a recorded data
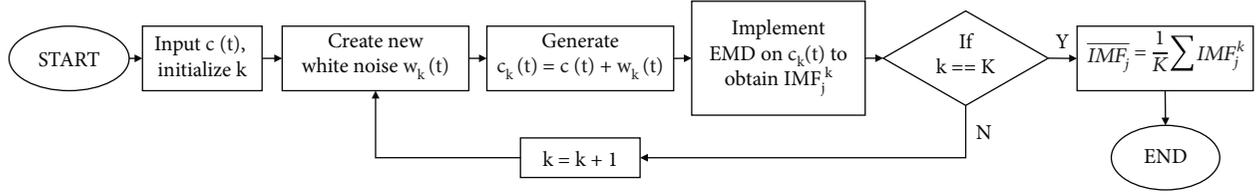
FIGURE 1: EEMD workflow.

as $c(t)$. A set of white noises are added to it which can be seen as a kth artificial observation.

$$c_k(t) = c(t) + w_k(t). \qquad (7)$$

Here, each of the observation is impersonated by appending different realization of white noise to a single observation $c(t)$. Even though the white noise inclusion brings about the drop in signal-to-noise ratio, but it comes up with relatively uniform distribution of reference scale which actually eradicates the problem of mode mixing. Hence, EEMD determines the true IMF by calculating the average on ensemble of trails where each trail consists of a signal and finite amplitude white noise. Figure 1 represents a flow diagram for the implementation of EEMD.

The critical abstractions of the technique are as follows:

(i) The cluster of white noises abandons each other in an ensemble mean that only left out the components of original signal

(ii) The finite amplitude noise actually facilitates the inhabitation of different scale signal into the corresponding IMF which makes the ensemble mean much more worthwhile

Moreover, EEMD introduces an element of randomness to the decomposition process by incorporating white noise into the disturbance signal. This results in the generation of multiple slightly varied realizations of the signal. The IMFs from these realizations are then averaged, providing a stronger illustration of each IMF. The ensemble averaging technique proves effective in mitigating the mode mixing problem. Because the added noise in each realization is random, the mixing effects tend to cancel out when combining results from multiple realizations. Additionally, EEMD follows statistical process to address mode mixing, leveraging the law of large numbers. With an increasing number of realizations, the signal's statistical properties become more evident, leading to a more reliable decomposition process [11, 14].

2.3. Random Decision Forest Classifier. RF is an ensemble classifier whose prediction is not established on a single classifier rather than on a bunch of classifiers whose decision-making is much more accurate and noise robust. A random forest generally consists of several decision trees that justifies the term "forest" and can be represented mathematically as

TABLE 1: Parameters of TS-I.

Parameters	Specification
Utility grid	120 kV, 60 Hz
DG1, DG2	9 MW wind system, 6 units of 1.5 MW, 60 Hz, 575 V
DG3	3.125 MVA synchronous diesel engine, 60 Hz, 575 V
Switching capacitor (C)	25 kV, 2 MVAR
Induction motor (M)	6 MVA
TF1	25 MVA, 120 kV/25 kV star-delta
TF2, TF3	12 MVA, 25 kV/575 V delta-star
TF4	5 MVA, 25 kV/575 V delta-star
L1, L2, L3	22 km, 60 Hz, 3-phase π-section line
LOAD3,4	8 MW, 1.75 MVAR
LD5	2.5 MW, 0.75 MVAR
LOAD1,2	6 MW, 1.4 MVAR

$$\text{Forest} = \{\text{TR}_1(x_{\text{rand}}, c_{\text{rand}}), \text{TR}_2(x_{\text{rand}}, c_{\text{rand}}), \\ \cdots .., \text{TR}_z(x_{\text{rand}}, c_{\text{rand}})\}. \qquad (8)$$

Here, x_{rand} is the random feature set at each node split, and c_{rand} is the subset of classes a tree can predict.

It is important to note that, out of the whole dataset, a certain amount of data vector is designated for training and rest which are kept for out-of-bag estimation. Once the model is trained, it can predict a new set of data with the majority vote of the trees. For any input data vector X, it can be defined as

$$P_{rf}^Z = \text{maxvote}\{P(X)_1, P(X)_2, \cdots, P(X)_z\}, \qquad (9)$$

where $P(X)_z$ symbolizes the prediction of zth tree.

Furthermore, the limitation of RF is high variance and biasing that arises fitting issues. That means a delta amount of change in training data is supposed to create a completely different variant of decision trees. However, the problem of underfitting and overfitting is effectively addresses through two primary strategies: bagging and random feature selection. Bagging entails creating multiple subsets of the training data by sampling with replacement. Each subset undergoes training with a decision tree. This approach diminishes the model's variance, reducing its sensitivity to noise or outliers in the data. Another important strategy is random feature selection, where instead of using all features at each split of the decision tree, a random subset of features is chosen. This
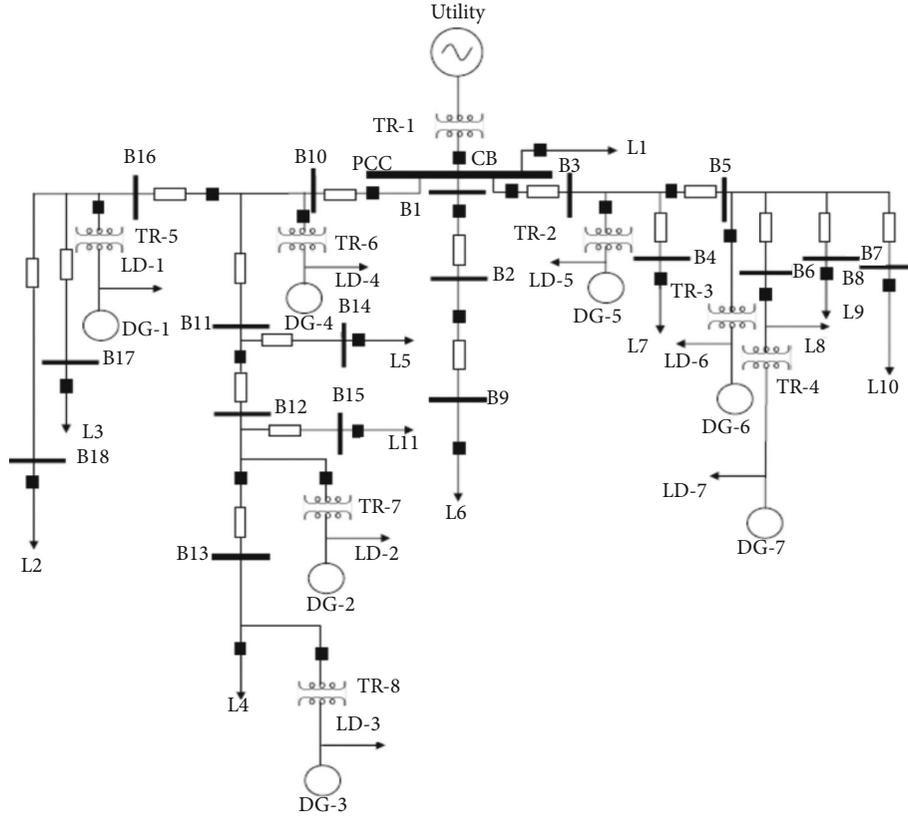
Figure 2: TS-I microgrid.

helps minimize the correlation among trees and boosts the model's resilience to irrelevant or redundant features. By combining these two techniques, the random decision forest achieves a well-balanced trade-off between bias and variance, effectively managing challenges related to underfitting and overfitting issues [41, 42].

3. Systems under Study

In this study, the validations are made with two separate test systems, namely, TS-I and TS-II, to ensure the dynamism of the proposed work. The first microgrid TS-I [40] is connected to a central grid at 120 kV bus which can be intentionally islanded by operating circuit breaker 1 (CB1). The related TS-I parameters are summarized in Table 1. The main transformer TF1 steps down the voltage from 120 kV to 25 kV and connected to the PCC at the low voltage end. Figure 2 shows the one-line diagram of TS-I that comprises of 3 DGs at separate location. On the other hand, the TS-II is consisting of 7 numbers of DGs where DG-1,2,7 are wind farms, DG-3,5,6 are diesel generators deployed in the time of emergency, and DG-4 is alone a PV unit. It is shown in Figure 3. The specifications of TS-II are presented in [43].

4. Data Collection and Feature Extraction

The purpose of the study is classification of disturbance events with machine learning which requires labelled dataset. Thus, dataset formation is very essential while imple-

menting any MLC algorithm. But the preliminary step is to accumulate the skeletal parts of the dataset, i.e., features. This is done by the following steps:

(a) The voltage and current signals form the PCC are collected from few cycles and saved as time series data

(b) The signals are sampled uniformly at 2.4 kHz, and data samples are stored within a time frame of 0.6 to 0.8 seconds as disturbance is intentionally occurred within this interval. This can be seen in Figures 4(a) and 4(b)

(c) From the retrieved signals, 2 types of features are extracted, namely, raw signal features (RSFs) and decomposed signal features (DSFs)

The mathematical formulations of RSF and DSF are mentioned from Eq. (10)–(26). Here, (10)–(20) are for RSF and rest are applied on 3rd IMF of voltage signal to obtain DSF. Now for any signal $y(k)$ having K number of samples, the features are represented as follows:

$$\text{Voltage RMS,} \quad V_{\text{rms}}, \tag{10}$$

$$\text{Current RMS,} \quad I_{\text{rms}}, \tag{11}$$

$$\text{Mean of first derivative of voltage,} \quad \text{MEAN}\left(\frac{dv}{dt}\right), \tag{12}$$
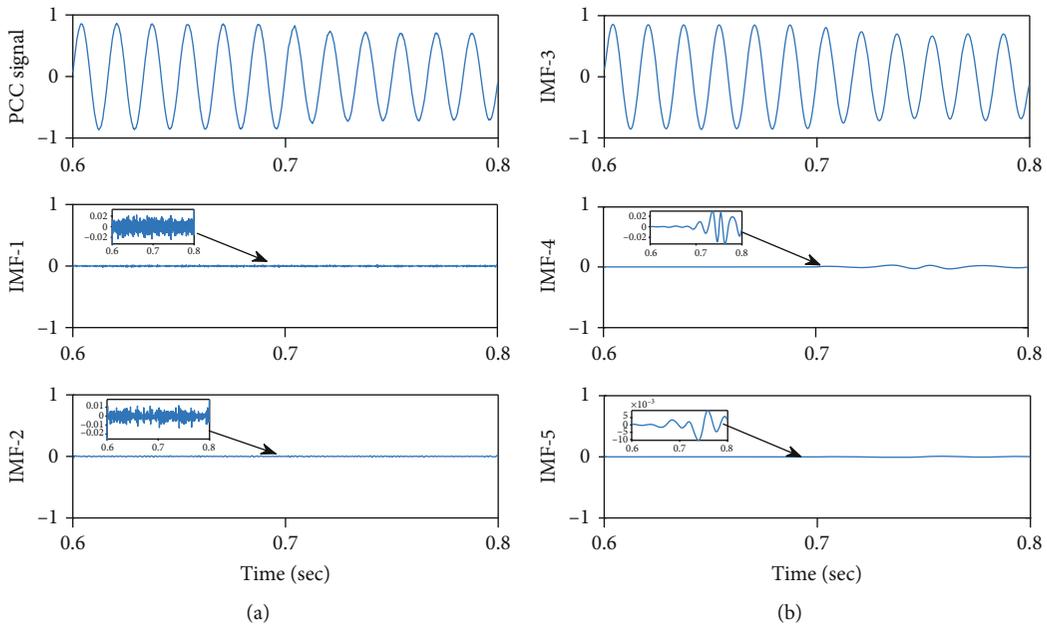
FIGURE 3: Studied microgrid system-II.

FIGURE 4: PCC voltage and its decomposed IMFs: (a) PCC voltage with 1st and 2nd IMFs; (b) 3rd to 5th IMF.

$$\text{Mean of second derivative of voltage,} \quad \text{MEAN}\left(\frac{d^2v}{dt^2}\right), \tag{13}$$

$$\text{Mean of first derivative of current,} \quad \text{MEAN}\left(\frac{di}{dt}\right), \tag{14}$$

FIGURE 5: Flow diagram of proposed method.

Mean of second derivative of current, $\quad \text{MEAN}\left(\dfrac{d^2 i}{dt^2}\right),$

$$(15)$$

Mean of first derivative of frequency, $\quad \text{MEAN}\left(\dfrac{df}{dt}\right),$

$$(16)$$

Mean of second derivative of frequency, $\quad \text{MEAN}\left(\dfrac{d^2 f}{dt^2}\right),$

$$(17)$$

Mean of first derivative of power, $\quad \text{MEAN}\left(\dfrac{dP}{dt}\right), \quad (18)$

Table 2: Class of events.

Event classes	Name of event	Description	Samples
EC-1	Islanding	CB1 opening	800
EC-2	Outage of DG	DG-1	800
EC-3		DG-2	800
EC-4	Capacitor switching	S1 switching	800
EC-5	Line switching	CB3 opening	800
EC-6	Induction motor switching	S3 switching	800
EC-7	Load switching	S2 switching	800
EC-8	Fault	LG, LLG, LLL, LLLG	200
EC-9	Voltage sag	At target DG	800
EC-10	Voltage swell	At target DG	800

$$\text{Voltage THD,} \quad \text{THD}_v = \frac{\sqrt{\sum_{n=2}^{\infty} V_{n\_\text{RMS}}^2}}{V_{fd\_\text{RMS}}}, \quad (19)$$

$$\text{Current THD,} \quad \text{THD}_i = \frac{\sqrt{\sum_{n=2}^{\infty} I_{n_{\text{RMS}}}^2}}{I_{fd_{\text{RMS}}}}, \quad (20)$$

$$\text{Mean,} \quad A_f = \sum_{k=1}^{K} y_k, \quad (21)$$

$$\text{Energy,} \quad \text{ENG}_f = \sum_{k=1}^{K} y_k^2, \quad (22)$$

$$\text{Variance,} \quad \text{VAR}_f = \frac{1}{(\text{K}-1)} \sum_{k=1}^{K} (y_k - A_f)^2, \quad (23)$$

$$\text{Standard deviation,} \quad \text{DEV}_f = \sqrt{\frac{1}{K} \sum_{k=1}^{K} (y_k - A_f)^2}, \quad (24)$$

$$\text{Skewness,} \quad \text{SKW}_f = \frac{1}{K \times \text{DEV}_f^3} \sum_{k=1}^{K} (y_k - A_f)^3, \quad (25)$$

$$\text{Kurtosis,} \quad \text{KUT}_f = \frac{1}{K \times \text{DEV}_f^4} \sum_{k=1}^{K} (y_k - A_f)^4, \quad (26)$$

where V_{n_RMS} and I_{n_RMS} are the nth harmonic voltage/current and $V_{fd_{\text{RMS}}}$ and $I_{fd_{\text{RMS}}}$ are the fundamental voltage/current.

5. Methodology

There are several events in power system such as transformer/capacitor switching, load/line switching, and operating of circuit breakers that are wrongly identified as islanding events by the traditional detection schemes. The working flow of the proposed techniques is undetected due to low power mismatching which should be avoided. A flow diagram of the proposed work is presented in Figure 5 for detection of islanding and nonislanding (I&NI) events correctly. Keeping in mind the actual grid scenarios, a wide range of I&NI events are deliberately simulated in MATLAB

Table 3: Correlation of IMFs with healthy signal.

IMF$_i$	IMF1	IMF2	IMF3	IMF4	IMF5
PERCC	0.0173	0.0392	0.9137	0.1733	0.0851

Simulink environment. In this process, 10 different classes of events are obtained as explained in Table 2. For each event occurrence, both voltage and current signals are collected from PCC. From the collected signal, a total of 17 features including RSF and DSF are extracted. As discussed before, RSFs are calculated directly from raw signal, but DSFs are extracted from the 3rd IMF of the EEMD decomposed voltage signal. The motive behind considering the 3rd IMF and discarding others is only because of the highest similarity of this with respect to the disturbance signal. This can be visualized by Pearson's correlation coefficient (PERCC). PERCC serves as a gauge for the linear relationship between two signals. Its value ranges from -1 to 1. This coefficient indicates a perfect negative correlation at -1, no correlation at 0, and a perfect positive correlation at 1. It can be defined as

$$\text{PERCC}(X, Y) = \frac{1}{M} \sum_{k=1}^{M} \left(\frac{x_k - X_\mu}{X_\sigma}\right) \left(\frac{y_k - Y_\mu}{Y_\sigma}\right), \quad (27)$$

where X_μ and Y_μ are the mean and X_σ and Y_σ are the standard deviation of the respective signals.

The PERCC of all the IMFs to that of the healthy signal is given in Table 3. It clearly displays that IMF-3 has the highest correlation of 0.9137 with the disturbance signal and hence chosen for DSF extraction. Thus, 800 data instances of each EC have been saved to the database. An 8000×17 feature vector has been extracted in this procedure and fed to a FST called binary grey wolf optimizer (BGWO) which will eliminate the redundant features and only hold 9 uncorrelated features. Those are V_{rms}, I_{THD}, MEAN(dv/dt), MEAN(d^2i/dt^2), MEAN(df/dt), MEAN(dP/dt), ENG$_f$, DEV$_f$, and SKW$_f$. The objective function of BGWO is taken as the overall accuracy error (OAE). It is defined as

$$\text{OAE} = 1 - \text{OA}, \quad (28)$$

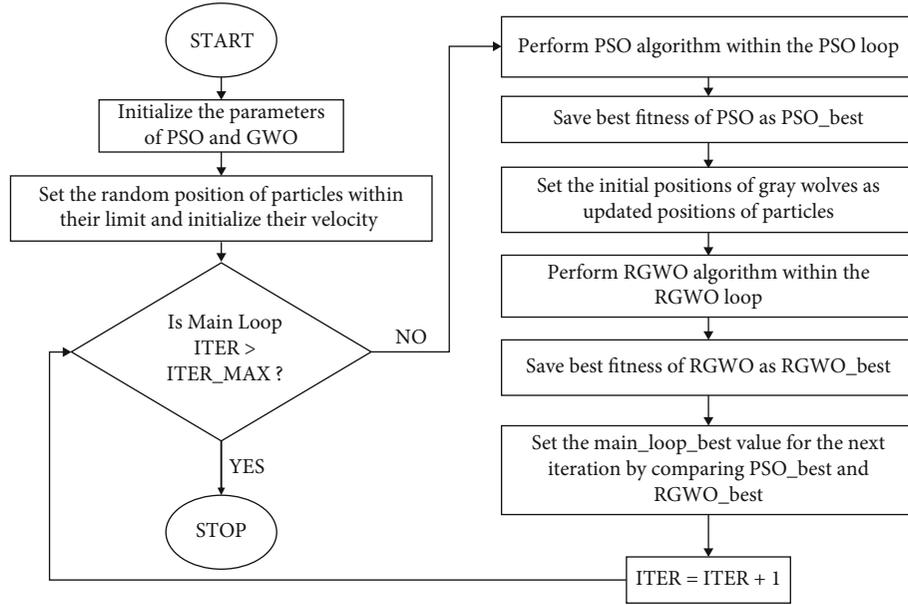where OA stands for overall accuracy.

Figure 6: Flow diagram of proposed PSRGWO.

Next, the compressed dataset undergoes data preparation stage where the data is get divided into training and testing data to fed into RFC. But with the focus on improving the efficacy of this classifier, a hybrid optimization technique, namely, PSO with reformed-GWO (PSRGWO), has been proposed in this paper. This hybrid optimizer is basically going to optimize two parameters. The detailed information regarding the application of PSRGWO is presented in the next subsection.

5.1. Hybrid PSO and Reformed GWO Algorithm (PSRGWO). In this section, a hybrid algorithm is suggested combining the traditional particle swarm optimization with a reformed version of grey wolf optimization so as to strengthen the exploration capabilities. Here, within a single iteration of main loop, there exist few iterations of PSO followed by a few iterations of GWO. At the end of every loop, the outcomes of PSO and GWO are compared to save the best value till that particular iteration. As a result, the modified GWO will ensure the reduction in probability of PSO to get trapped in local extrema. The workflow of the proposed algorithm is shown in Figure 6.

Furthermore, there are few modifications which are imposed to the conventional GWO algorithm to make the entire optimization process more efficient. In any population-based optimization technique, the agents will first get distributed among the entire search space to explore it and then exploit the collected knowledge to converge the solution to global extremum. Mathematically, if $|\vec{Y}| > 1$ (as of 3), then exploring is going on and it continues till half of the total GWO iterations and in the next half of iterations $|\vec{Y}| < 1$ that carry out the exploitation work. As higher exploration reduces the possibility of local stagnation, the iteration dedicated to exploration can be increased by modifying (3). It is defined as

$$\tilde{g} = 2 - \frac{2}{(\text{Total}_{\text{Iterations}})^2} * j^2. \tag{29}$$

This will devote more than two-thirds of the total iterations for exploration job, and the rest are committed to exploitation. Secondly, there is a mere possibility of the solution to get trapped other than the major three wolves called omega ($\grave{\omega}$) wolves. For this cause, the positions of omega wolves are also updated in GWO reformation. It can be represented as

$$\overrightarrow{W_4} = \overrightarrow{X_\omega} - \overrightarrow{Y_3} \cdot \left(\overrightarrow{S_\omega}\right). \tag{30}$$

Thirdly, even though it was said in conventional GWO that the wolves are following a specific hierarchy but while updating the position vector equation, all three major wolves are treated equally as shown in (4). Here, the updating equation is rewritten with a weighted valuation of wolves. Thus, the reformed updating equation is as follows:

$$W_R(j+1) = \frac{4W_1 + 3W_2 + 2W_3 + W_4}{10}. \tag{31}$$

Moreover, the cost function used in such type of problems is generally based upon accuracy index. But in this error minimization problem, Cohen's kappa error is minimized to ensure an optimum number of trees and especially least number of FPS. This will guarantee the highest accuracy with lowest correlations between trees. The kappa index can be mathematically represented as

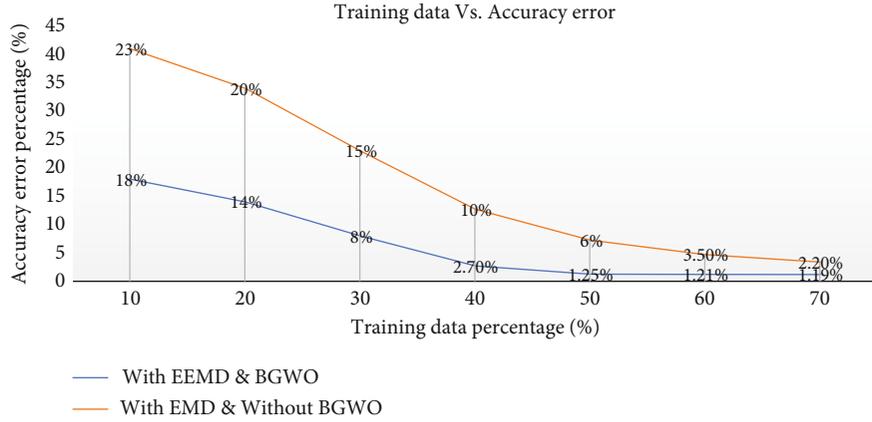$$\mathbb{K} = \frac{\text{PA} - \text{PA}_C}{1 - \text{PA}_C}, \tag{32}$$

FIGURE 7: Training data vs. accuracy error graph.

where PA is the probability of agreement and PA_C is the probability of agreement by chance. Now the objective function of the minimization problem can be defined as

$$\mathbb{K}_{error} = 1 - \mathbb{K}. \tag{33}$$

Here, $\mathbb{K}_{error}$ represents the kappa error, and $\mathbb{K}$ represents the actual kappa index. These are function of number of trees (TR) in the forest and feature involved in each split (FPS) of any tree construction.

6. Result Analysis

The proposed system is simulated in MATLAB (version 2018a) software running in a Lenovo personal computer with INTEL core-i5 8th generation 2.3 GHz processor and 8 GB RAM. The performance will be evaluated using the feature optimized dataset consisting of 9 distinct features (column wise) that holds 10 sets of event classes (row wise) to form an 8000×9 data vector. The collection of data samples is made under a wide range of power mismatch condition for both I&NI occurrences which results in a diversified dataset that have the seldom possibility of redundancy and biasing. Here, half of the data are taken for training and the remaining half for testing to verify the hybrid optimization-based random forest classifier (HOBRFC). Normally, a higher percentage of data (70 to 80%) are taken to train any classifier to ensure higher accuracy, but purity and clarity of data abstraction can play a vital role to amplify the performance indices of any classifier. A "training data percentage" versus "accuracy error" graph in Figure 7 can well describe the meaning of the previous statement. It can be observed that the reduced dataset can attain maximum prediction level even with 50% of training data. This signifies the unmatched data abstraction ability of EEMD as well as the unbiased feature selection capability of BGWO.

6.1. Optimal Parameter Selection for HOBRFC. RF classifier's performance is based on the decision of several DTs whose branch division is carried out with a random subset of features. Training such type of models requires the dimensionality definition of both of these two elements which can be
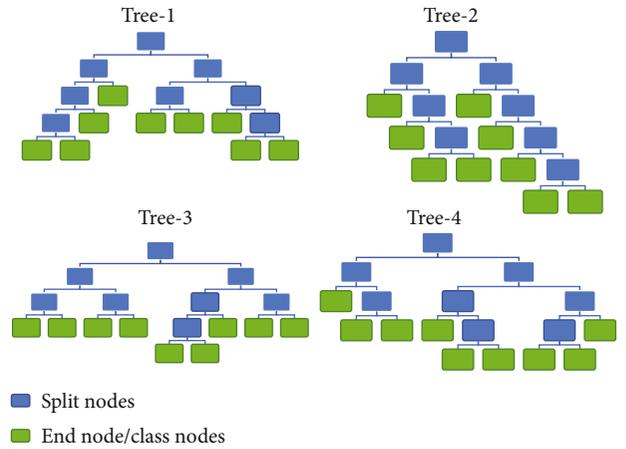

FIGURE 8: Tree structure in a RF model each having 9 classes with unique split.
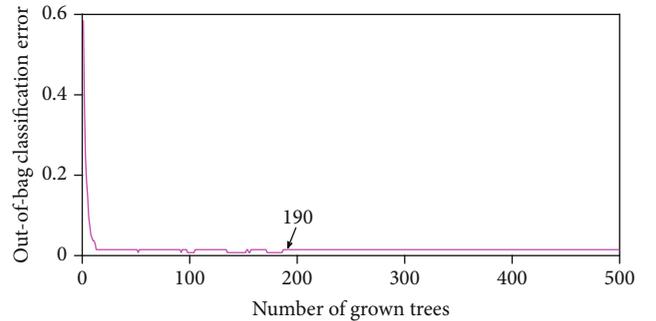

FIGURE 9: OBCE to determine maximum number of trees required for optimization.

achieved using optimization. As we are optimizing TR and FPS, the preliminary requirement is to set the lower and upper limits of these two. The default FPS [42] in an RFC is chosen as $\sqrt{p}$, where p is the maximum number of available features. In this case, $p = 9$ so the lower limit of FPS is

TABLE 4: Parameters of PSRGWO.

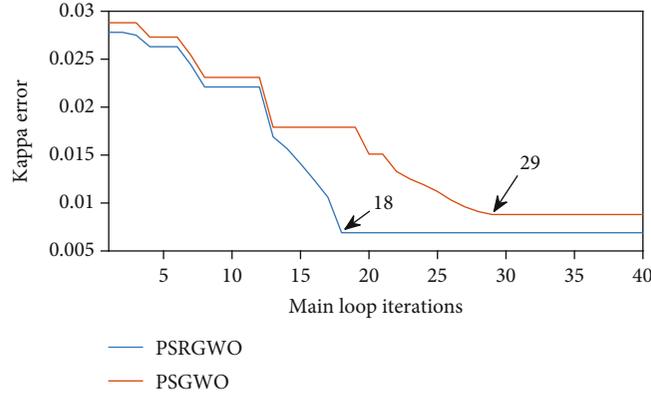| Parameters	Values	Parameters	Values
Max iteration of main loop	40	Search agents	20
Max iteration of PSO loop	15	Max iteration of GWO loop	15
Range of split features	[3 8]	Range of trees	[9 200]

FIGURE 10: Comparison of PSGWO and PSRGWO.

taken as 3. On the other side setting, the max value as the upper limit, certainly going against the core definition of the term "random" as this signifies any node of a DT inside the RF is subdivided into further branches using a random subset of features.

Moreover, a higher value of FPS can increase the correlation between the trees which in turn reduces their individuality. Hence, the upper limit of FPS is taken as 8 instead of 9. Further, setting the limits of TR is quite tricky. Choosing a much smaller lower limit can affect agreement to disagreement ratio as the decision-making of RF relies on collective voting approach. Since lower limit of FPS is 3 and maximum feature available is 9, a total of 84 combinations of feature subsets can be formed as per (33). If each subset is assumed to be used uniquely to split only a single node of a tree in the entire forest, it is required at least 9 trees to accommodate all. A glimpse of Figure 8 can well explain why it has been said. There are 4 example trees where blue-colored nodes are split nodes and green-colored nodes are end nodes or class nodes. It can be noted that, to reach 9 end nodes, there should be minimum 8 split nodes in a tree. As the end nodes of trees are the classes, at least "$m - 1$" node splits are needed to specify "m" classes in a tree. Since the number of classes is 10 over here and 9 node splits are required to make 10 end nodes, $84/10 = 8.4(\approx 9)$ trees are desired to accommodate 84 feature subsets. Hence, the lower limit of TR is set to be 9. Although RF never overfits while adding more and more numbers of trees [43], still it is important to find out the optimum trees for the classifier speed point of view. So, the upper limit can be decided by plotting, out-of-bag classification error (OBCE) vs. number of trees as shown in Figure 9. It can be seen that the OBCE remains stable

approximately after 190 trees. Therefore, with a small margin, the upper limit is placed at 200.

$$\text{Combinations} = C_r^n (n \geq r) = \frac{n!}{r!(n-r)!}. \quad (34)$$

The parameters used during the implementation of proposed optimization algorithm are presented in Table 4. It can be seen from Figure 10 that PSRGWO's performance is proving better both in terms of a number of iterations and error margin. While the PSGWO is taking 29 iterations to get the least converged error, on the other hand, the proposed PSRGWO is only taking 18 iterations to achieve an even low error margin. Finally, the optimum values of FPS (O_{FPS}) and TR (O_{TR}) are found to be 4 and 141, respectively, after taking the modes of 100 trails.

6.2. Performance Indices. Performance indices (PIs) are the measures to analyze the prediction capability of a classifier once it is trained. It may sound a little weird if the formulae of different indices are directly stated over here. But it will be much easier to understand the PIs once after understanding the concept of the confusion matrix (CM). Simply, CM is a square matrix datasheet that estimates the performance of any ML classification model.

The core of the random forest machine learning model is decision trees, and each decision tree will train with random subset of features. Furthermore, the end nodes of these trees are the event classes. Therefore, a model trained with 10 classes will one islanding, and 9 PQ events (case 1) have separate tree structures than that of a model trained with only 9 PQ event classes (case 2). For this reason, the proposed model is tested with both the conditions.
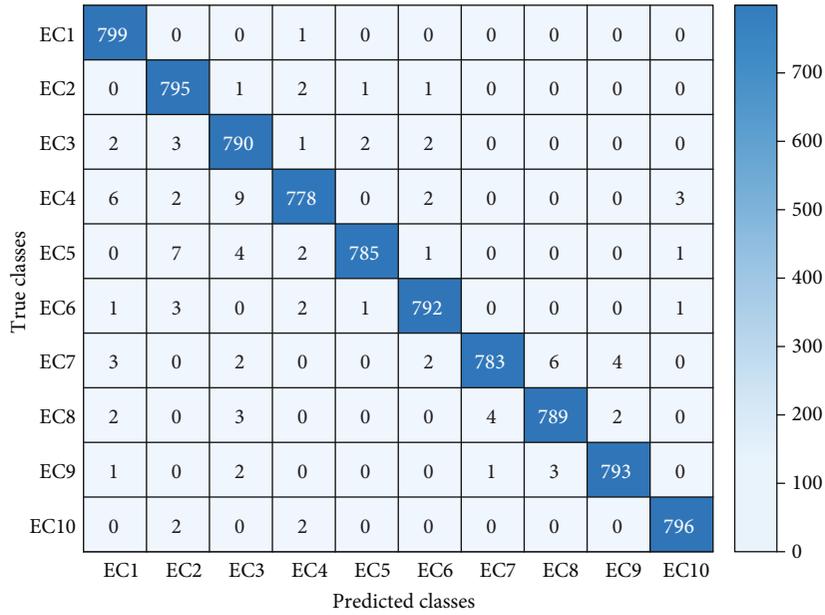
(a) CM with islanding and PQ scenarios
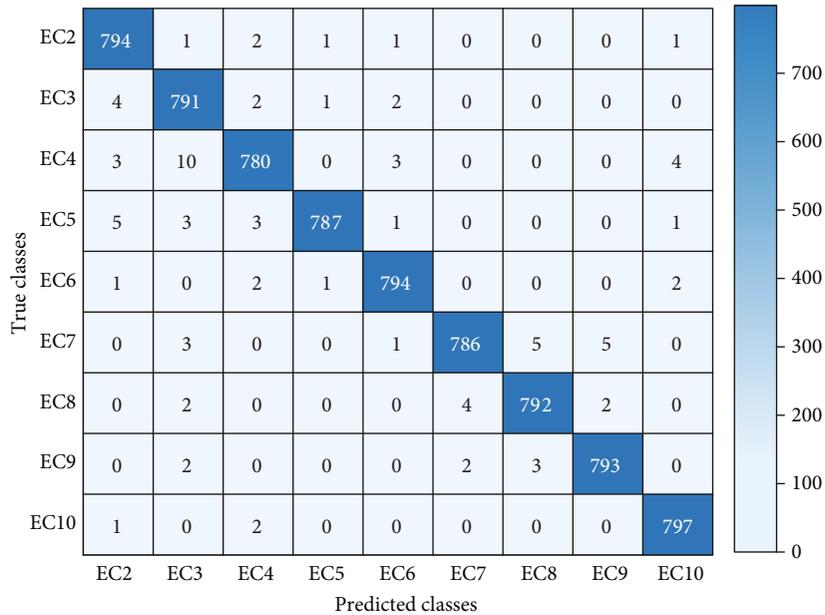
(b) CM under nonislanding scenario

FIGURE 11: CM with 800 samples/EC for TS-I.

In Figure 11(a), the confusing matrix is obtained for the model trained with case 1, where 800 events per class are predicted using the proposed detection technique. It is found that the method shows 99.88% accuracy in detecting the islanding scenarios as 799 islanding cases are truly detected out of 800 classes. Apart from that, other nonislanding events are also showing promising accuracy of detection. Here, classes EC-1 to EC-10 are mapped as 1 to 10. Similarly, in Figure 11(b), only the case 2 events are considered. The proposed classifier also displays much accuracy for nonislanding events alone where EC-10 has highest accuracy rate of 99.63% since 797 events have predicted truly out of

800. Here, classes EC-2 to EC-10 are mapped as 1 to 9. The figures are represented as heatmaps. The darker color of the diagonal elements reflecting the elegant performance of the classifier. Now the above statement can be verified mathematically with the help of several PIs.

The mathematical formulae of several PIs such as accuracy, specificity, sensitivity, and Matthew's correlation coefficient (MCC) are presented in the (35)–(38). Two bar chart distributions of all the PIs of every EC both under I&NI events are shown in Figures 12(a) and 12(b). Due to the randomness property of RF, all the indices are obtained by taking the mean of 100 trials. These results are ensuing

the efficacy of the proposed HOBRFC which is well capable of detecting islanding and PQ disturbance events.

$$\text{Accuracy, } A_{\text{EC}_x} = \frac{\text{truly predicted as EC}_X}{\text{total number of events from EC}_X}, \tag{35}$$

$$\begin{aligned} \text{Sensitivity, } SV_{\text{EC}_x} &= \frac{\text{truly EC}_X}{\text{truly EC}_X + \text{falsely other than EC}_X} \\ &= \frac{\text{Pos}(XX)}{\sum X\text{column}}, \end{aligned} \tag{36}$$

$$\begin{aligned} \text{Specificity, } SP_{\text{EC}_x} &= \frac{\text{truly not EC}_X}{\text{truly not EC}_X + \text{falsely EC}_X} \\ &= \frac{\sum \text{main\_diagonal} - \text{Pos}(XX)}{(\sum \text{main\_diagonal} - \text{Pos}(XX)) + (\sum X\text{row} - \text{Pos}(XX))}, \end{aligned} \tag{37}$$

$$MCC_{\text{EC}_x} = \sqrt{\frac{\text{Pos}(XX) * [\sum MD - \text{Pos}(XX)] - [\sum X\text{row} - \text{Pos}(XX)] * [\sum X\text{column} - \text{Pos}(XX)]}{\sum X\text{row} * \sum X\text{column} * [\sum MD + \sum X\text{row} - 2 * \text{Pos}(XX)] * [\sum MD + \sum X\text{column} - 2 * \text{Pos}(XX)]}}. \tag{38}$$

6.3. Noise Tolerance Evaluation. Noise is a nonavoidable part of any system, thus also visible in power system. The companionship of communication lines with power lines creates electromagnetic interference which is the main reason for noise existence in collected signal.

As the studied system is designed in MATLAB environment, to simulate the noisy condition, additive white Gaussian noise with 3 different signal-to-noise ratio is added to the collected signal. It can be expressed as

$$R_{S/N} = 10 * \log\left(\frac{\text{power of signal}}{\text{power of noise}}\right) (\text{in dB}). \tag{39}$$

The proficiency of the suggested approach under noisy conditions can be contemplated in Figures 13(a) and 13(b). It is observed that the noise-free accuracy of important islanding events (EC-1) is reaching 99.88% whereas it is hitting 99.2% even under the dense noise of 20 dB. Similarly, the mean accuracies of the other events under islanding conditions represented by EC-2 to EC-10 are maintained at 98.79%, 98.58%, 98.39%, and 98.18% for 40 dB, 30 dB, and 20 dB SNR, respectively.

Additionally, the proposed EEMD-HOBRFC also provides faster detection of all the events, and it is noted to be only 23.16 msec under the ideal situation. The exceptional ability of the technique makes it to elapse almost same time to detect any type of trained event. But the performance is stagnated with the introduction of noise. A series of 500 disturbance cases is taken for each SNR level, and the mean time of detection is illustrated in Figure 14.

6.4. Performance Assessment in TS-II. The extraordinary efficacy of the EEMD-HOBRFC is already seen in the previous subsections for test system 1. However, it is essential to verify whether it is performing in the same manner or not for a different system. So as to validate it, a separate test system is taken into consideration as shown in Figure 3. A 400 number of cases are extracted for each event class only having the BGWO optimized prediction variables (9 numbers of features) from test system 2. This extraction has taken place in both islanding and nonislanding scenarios. The confusion matrix in relation to both the scenarios can be seen in Figures 15(a) and 15(b). It can be seen from Figure 15(a) that the method shows 100% accuracy in detecting the islanding scenarios as all the 400 cases are truly detected. Here, classes EC-1 to EC-10 are mapped as 1 to 10. On the other hand, the proposed classifier also displays much accuracy for nonislanding events alone where EC-15 has highest accuracy rate of 99.5% since 398 events have predicted truly out of 400. Here, classes EC-2 to EC-10 are mapped as 1 to 9.

It can be noted that, under islanding condition (Figure 15(a)), the method of EEMD-HOBRFC successfully detected all the 400 islanding events. In addition, it is also displayed its detection capability for the other events with and without islanding occurrences. Hence, the proposed method is found to be effective even in other microgrid structures.

6.5. Comparative Study. This section covers up the comparative study of the proposed EEMD-HOBRFC method with several other methods. Firstly, both EMD and EEMD are combined with simple unoptimized RF and latterly clubbed with HOBRFC for 3 different SNR levels as presented in Table 5. The resulting accuracy data can well define the dominancy of the proposed method that reaching 99.2%

	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
■ Accuracy	99.88%	99.38%	98.75%	97.25%	98.13%	99%	97.88%	98.63%	99.13%	99.50%
■ Sensitivity	98.16%	97.91%	97.41%	98.73%	99.49%	99%	99.36%	98.87%	99.25%	99.37%
■ Specificity	99.99%	99.93%	99.86%	99.69%	99.79%	99.89%	99.76%	99.85%	99.90%	99.94%
■ MCC	98.90%	98.48%	97.86%	97.76%	98.67%	98.89%	98.46%	98.61%	99.10%	99.37%

(a) Under islanding situation

	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
■ Accuracy	99.25%	98.88%	97.50%	98.38%	99.25%	98.25%	99%	99.13%	99.63%
■ Sensitivity	98.27%	97.41%	98.61%	99.62%	99%	99.24%	99%	99.13%	99.01%
■ Specificity	99.91%	99.86%	99.69%	99.79%	99.91%	99.78%	99.87%	99.89%	99.95%
■ MCC	98.60%	97.91%	97.81%	98.87%	99.02%	98.59%	98.87%	99.01%	99.23%

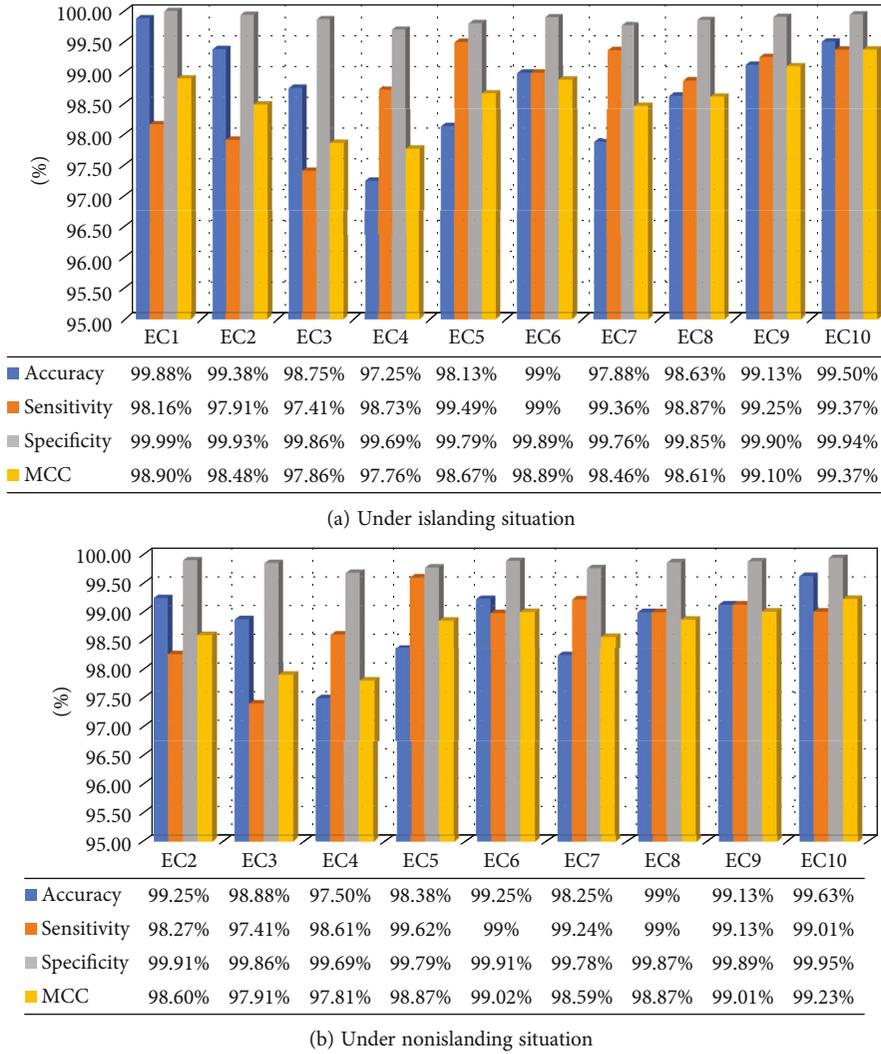(b) Under nonislanding situation
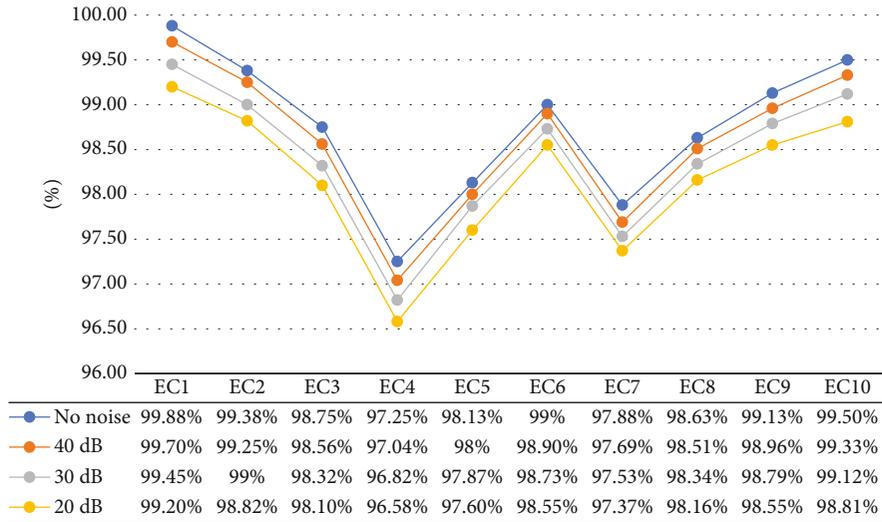
FIGURE 12: Performance indices of each event class (EC).

under substantial noise (20 dB) and 99.88% without any noise. Moreover, Table 6 unveils its supremacy over other ML tools. It can be noticed that the EEMD-HOBRFC is displaying appreciable accuracy with respect to all these MLTs with maintaining a mean accuracy of 98.75%.

Further, the analysis is extended in terms of detection time, detection accuracy, and performance under noisy conditions. The time required for islanding detection for different preproposed techniques with their references can be seen in Table 7. The maximum detection time taken is around 55 ms for a transient monitoring [45] based detection system which is more than twice of the detection time of the proposed one. Besides that, the performance of the EEMD-HOBRFC is evaluated against other prepublished work in Table 7. Furthermore, in Table 8, it can be noticed that, although [48, 52] delivering an accuracy level of 100%, still they are underperforming in the presence of noise. But the proposed method has built up with both noise-assisted signal processing technique and noise robust machine learning technique that is why performing at towering level of accuracy in all three noisy conditions.
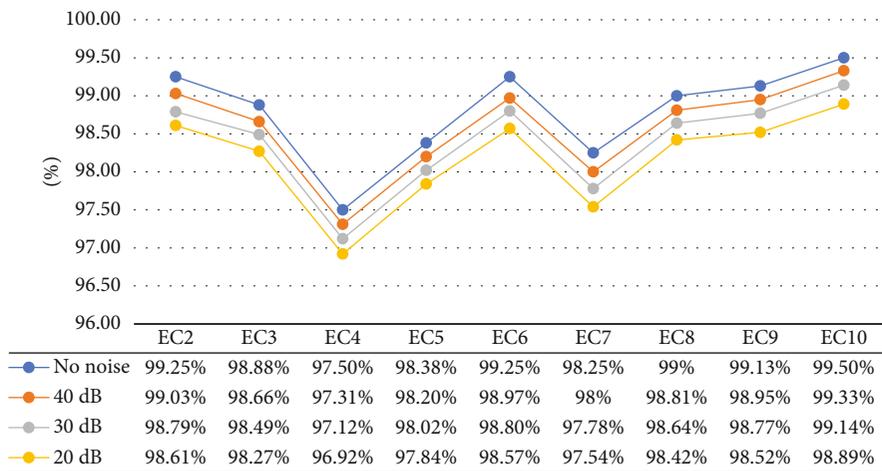
6.6. Nondetection Zone Validation. The contraction of NDZ can be validated by considering a mix of 500 diversified islanding data samples within a power mismatch range of 0.0005 pu to 0.03 pu with 30 dB SNR. Even in such a narrow range, the EEMD-HOBRFC is still capable of detecting 498 islanding event classes. The other two are detected falsely as DG outage and capacitor switching. The outcomes not only indicate the impact of the proposed method under low power mismatch conditions but also replicate the efficacy under the noisy situations. That means it can be well implemented in practical environment. The result of this test is given in Table 9, and the contracted region of nondetection is shown in grey color in Figure 16.

7. Conclusion

The manuscript presents an islanding detection scheme, namely, EEMD-HOBRFC, that can identify and classify the islanding as well as nonislanding power quality events. In the proposed approach, some of the features are retrieved from the raw signal to retain the original characteristics of

	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
No noise	99.88%	99.38%	98.75%	97.25%	98.13%	99%	97.88%	98.63%	99.13%	99.50%
40 dB	99.70%	99.25%	98.56%	97.04%	98%	98.90%	97.69%	98.51%	98.96%	99.33%
30 dB	99.45%	99%	98.32%	96.82%	97.87%	98.73%	97.53%	98.34%	98.79%	99.12%
20 dB	99.20%	98.82%	98.10%	96.58%	97.60%	98.55%	97.37%	98.16%	98.55%	98.81%

(a) Under islanding situation

	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10
No noise	99.25%	98.88%	97.50%	98.38%	99.25%	98.25%	99%	99.13%	99.50%
40 dB	99.03%	98.66%	97.31%	98.20%	98.97%	98%	98.81%	98.95%	99.33%
30 dB	98.79%	98.49%	97.12%	98.02%	98.80%	97.78%	98.64%	98.77%	99.14%
20 dB	98.61%	98.27%	96.92%	97.84%	98.57%	97.54%	98.42%	98.52%	98.89%

(b) Under nonislanding situation

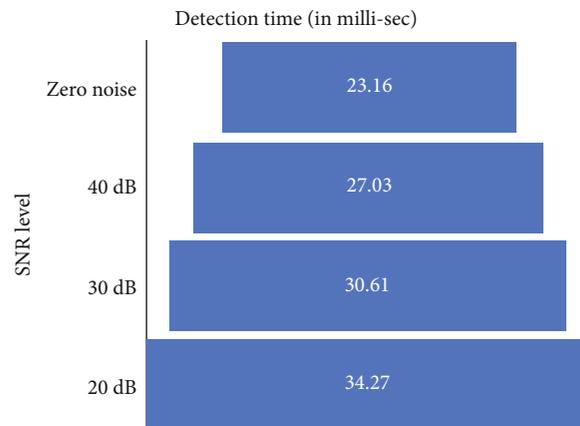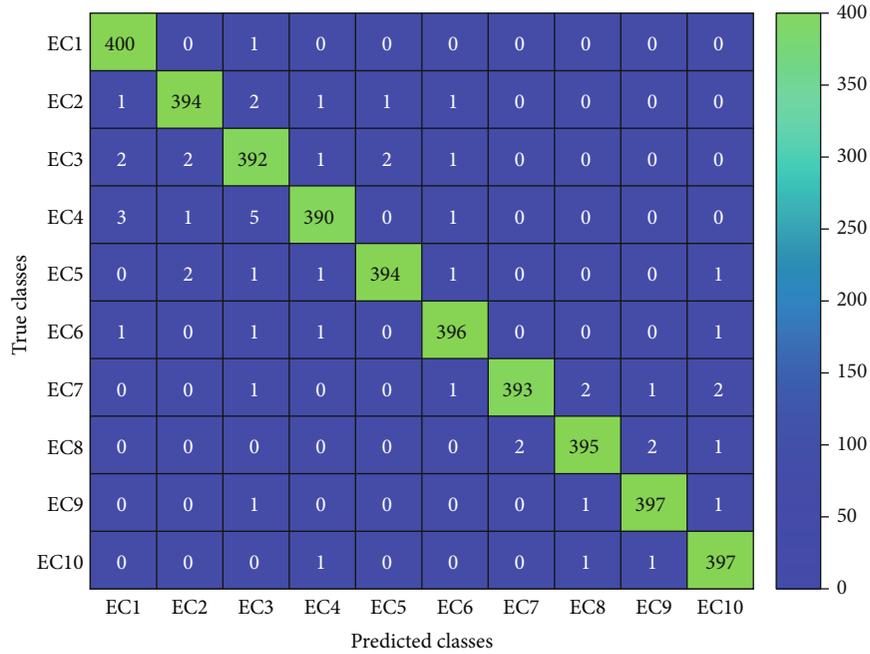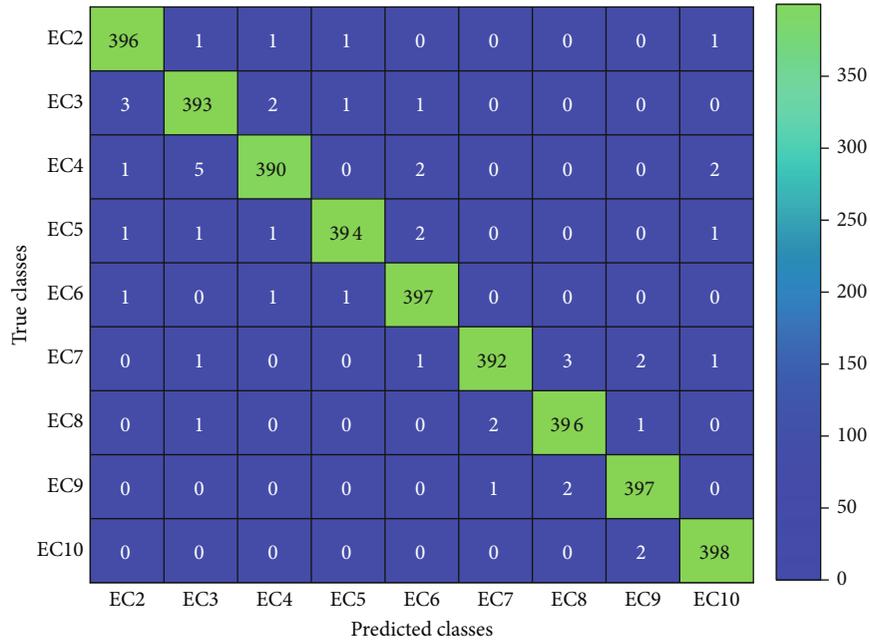FIGURE 13: Impact of noise on classification accuracy.

FIGURE 14: Impact of noise on detection time.

(a) CM with islanding and PQ scenarios

(b) Under nonislanding scenario

Figure 15: CM with 400 samples/EC for TS-2.

Table 5: Comparing with optimized and nonoptimized EMD and EEMD techniques.

Techniques	Zero noise	SNR level		
		40 dB	30 dB	20 dB
EMD-RF	97.6%	97.4%	97.1%	96.6%
EMD-HOBRFC	98.6%	98.42%	98.24%	97.97%
EEMD-RF	99.1%	98.87%	98.69%	98.4%
EEMD-HOBRFC	99.88%	99.7%	99.45%	99.2%

TABLE 6: Comparing with other ML techniques with EEMD in absence of any noise.

Classes	NB	ELM	KNN	SVM	DT	Proposed
EC-1	93.58%	98.65%	98.81%	99.02%	99.3%	99.88%
EC-2	92.95%	98.11%	98.25%	98.47%	98.81%	99.38%
EC-3	92.78%	97.6%	97.77%	98.01%	98.33%	98.75%
EC-4	91%	96.07%	96.23%	96.43%	96.78%	97.25%
EC-5	91.72%	96.95%	97.09%	97.31%	97.6%	98.13%
EC-6	92.53%	97.76%	97.93%	98.15%	98.48%	99%
EC-7	91.47%	96.63%	96.8%	97.03%	97.34%	97.88%
EC-8	92.39%	97.5%	97.64%	97.87%	98.17%	98.63%
EC-9	92.75%	97.89%	98.03%	98.26%	98.61%	99.13%
EC-10	93.04%	98.18%	98.36%	98.55%	98.98%	99.50%
Mean	92.421%	97.534%	97.691%	97.91%	98.24%	98.75%

TABLE 7: Detection time of the proposed method with respect to other research works.

Sl no.	Related works ref.	Detection time
1	[6]	Less than 28 ms
2	[44]	45 to 60 ms
3	[45]	Less than 34 ms
4	[33]	40 ms
5	[46]	50 ms
6	[47]	Around 55 ms
7	Proposed	23.16 ms

TABLE 8: Accuracy tally with some of the prepublished articles.

Sl no.	Published work	Accuracy			
		Noise free	20 dB	30 dB	40 dB
1	[26]	98.8%	98.3%	97.5%	97%
2	[34]	99.09%	91.51%	96.96%	—
3	[48]	100%	97.66%	—	—
4	[49]	98.19%	—	—	—
5	[50]	98%	—	—	—
6	[51]	—	—	98.3%	—
7	[52]	100%	90%	98.78%	—
8	[53]	99.33%	—	—	—
9	[54]	98%	94.2%	—	—
10	Proposed	99.88%	99.2%	99.45%	99.7%

TABLE 9: Event detection inside narrow range to validate NDZ.

True class	Predicted class									
	EC-1	EC-2	EC-3	EC-4	EC-5	EC-6	EC-7	EC-8	EC-9	EC-10
EC-1	498	1	0	1	0	0	0	0	0	0

the signal and the rest of the features are extracted from the decomposition of noise-assisted signal preprocessor called EEMD to capture the inherent characteristics.
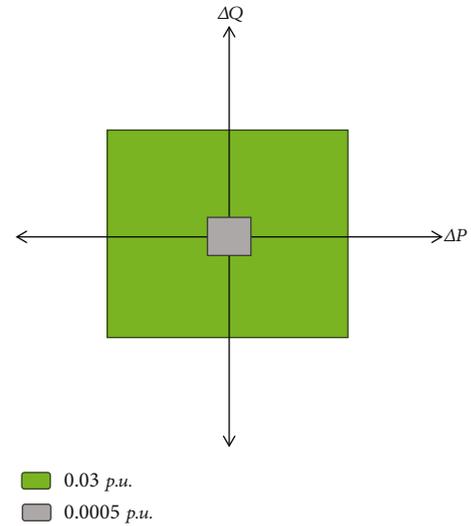

FIGURE 16: Nondetection zone validation.

Moreover, the least correlated features are filtered with BGWO and then fed to a PSRGWO optimized random forest classifier. The proposed technique is validated with a system consisting of multi-DG system with high wind energy penetration, and accuracy of islanding detection under ideal condition is found to be 99.88%.

Its performance is not worsening under noisy condition and maintained at 99.2% even in the dense noise of 20 dB because both EEMD and RF are noise robust in nature. Additionally, the detection time is found to be as low as 23.16 ms with no noise and 34.27 ms at heavier noise.

Lastly, the method is also able to detect islanding event in real scenario (30 dB SNR) alongside a power mismatch range of 0.03 pu to 0.005 pu with an appreciable accuracy of 99.6%. Hence, the method is found to be very competitive with reference to the related research works of recent past. From rigorous simulation study, it is found that the proposed EEMD-HOBRFC is able to detect and classify the islanding and nonislanding event classes efficiently.

Abbreviations

BGWO:	Binary grey wolf optimization
CM:	Confusion matrix
DG:	Distributed generator
DSF:	Decomposed signal feature
DT:	Decision tree
EC:	Event class
EEMD:	Ensemble empirical mode decomposition
EMD:	Empirical mode decomposition
EWT:	Empirical wavelet transform
FPS:	Feature per split
FST:	Feature selection tool
GWO:	Grey wolf optimization
HOBRFC:	Hybrid optimization-based random forest classifier
I&NI:	Islanding and nonislanding
IMF:	Intrinsic mode function
MCC:	Matthew's correlation coefficient
MLC:	Machine learning classifier
NDZ:	Nondetection zone
OAE:	Overall accuracy error
OBCE:	Out-of-bag classification error
PCC:	Point of common coupling
PERCC:	Pearson's correlation coefficient
PI:	Performance index
PSRGWO:	Particle swarm and reformed grey wolf optimization
RF:	Random forest
RSF:	Raw signal feature
SPA:	Signal processing approach
TR:	Number of tress in forest
TS:	Test system
VMD:	Variational mode decomposition.

Data Availability

There are no data to mention.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] I. E. E. E. Standards Board, *IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems: 1547-2003*, IEEE, 2003.

[2] M. E. Ropp, M. Begovic, and A. Rohatgi, "Analysis and performance assessment of the active frequency drift method of islanding prevention," *IEEE Transactions on Energy Conversion*, vol. 14, no. 3, pp. 810–816, 1999.

[3] W. I. Bower and M. Ropp, *Evaluation of islanding detection methods for utility-interactive inverters in photovoltaic systems (no. SAND2002-3591)*, Sandia National lab.(SNL-NM), Albuquerque, NM (United States); Sandia National Lab.(SNL-CA), Livermore, CA (United States), 2002.

[4] W. Freitas, W. Xu, C. M. Affonso, and Z. Huang, "Comparative analysis between ROCOF and vector surge relays for distributed generation applications," *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 1315–1324, 2005.

[5] S. Liu, S. Zhuang, Q. Xu, and J. Xiao, "Improved voltage shift islanding detection method for multi-inverter grid-connected photovoltaic systems," *IET Generation, Transmission & Distribution*, vol. 10, no. 13, pp. 3163–3169, 2016.

[6] P. K. Ray, S. R. Mohanty, and N. Kishor, "Disturbance detection in grid-connected distributed generation system using wavelet and S-transform," *Electric Power Systems Research*, vol. 81, no. 3, pp. 805–819, 2011.

[7] S. Raza, H. Mokhlis, H. Arof, J. A. Laghari, and L. Wang, "Application of signal processing techniques for islanding detection of distributed generation in distribution network: a review," *Energy Conversion and Management*, vol. 96, pp. 613–624, 2015.

[8] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London," *Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[9] S. Admasie, S. B. A. Bukhari, R. Bengovik Haider, T. Gush, and C. H. Kim, "A passive islanding detection scheme using variational mode decomposition-based mode singular entropy for integrated microgrids," *Electric Power Systems Research*, vol. 177, article 105983, 2019.

[10] J. Gilles, "Empirical wavelet transform," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3999–4010, 2013.

[11] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[12] A. Anand and S. Affijulla, "Ensemble empirical mode decomposition-based differential protection scheme for islanded and grid-tied AC microgrid," *IET Generation, Transmission & Distribution*, vol. 14, no. 26, pp. 6674–6681, 2020.

[13] Z. Zhang and W. C. Hong, "Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm," *Nonlinear Dynamics*, vol. 98, no. 2, pp. 1107–1136, 2019.

[14] M. Santhosh, C. Venkaiah, and D. V. Kumar, "Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction," *Energy Conversion and Management*, vol. 168, pp. 482–493, 2018.

[15] J. A. Laghari, H. Mokhlis, M. Karimi, A. H. A. Bakar, and H. Mohamad, "Computational intelligence based techniques for islanding detection of distributed generation in

distribution network: a review," *Energy Conversion and Management*, vol. 88, pp. 139–152, 2014.

[16] R. O. Duda and P. E. Hart, *Pattern classification*, John Wiley & Sons, 2006.

[17] M. Dashtdar, A. Flah, S. M. S. Hosseinimoghadam et al., "Improving the power quality of island microgrid with voltage and frequency control based on a hybrid genetic algorithm and PSO," *IEEE Access*, vol. 10, pp. 105352–105365, 2022.

[18] G. Vashishtha, S. Chauhan, A. Kumar, and R. Kumar, "An ameliorated African vulture optimization algorithm to diagnose the rolling bearing defects," *Measurement Science and Technology*, vol. 33, no. 7, article 075013, 2022.

[19] M. Dashtdar, A. Flah, S. M. S. Hosseinimoghadam et al., "Optimal operation of microgrids with demand-side management based on a combination of genetic algorithm and artificial bee colony," *Sustainability*, vol. 14, no. 11, p. 6759, 2022.

[20] M. Babanezhad, S. Arabi Nowdeh, A. Y. Abdelaziz, K. M. AboRas, and H. Kotb, "Reactive power based capacitors allocation in distribution network using mathematical remora optimization algorithm considering operation cost and loading conditions," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 10511–10526, 2022.

[21] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "BGSA: binary gravitational search algorithm," *Natural Computing*, vol. 9, no. 3, pp. 727–745, 2010.

[22] R. Y. Nakamura, L. A. Pereira, K. A. Costa, D. Rodrigues, J. P. Papa, and X. S. Yang, "BBA: a binary bat algorithm for feature selection," in *2012 25th SIBGRAPI Conference on Graphics, Patterns and Images*, Ouro Preto, Brazil, 2012.

[23] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, pp. 54–65, 2016.

[24] G. Vashishtha and R. Kumar, "An amended grey wolf optimization with mutation strategy to diagnose bucket defects in Pelton wheel," *Measurement*, vol. 187, article 110272, 2022.

[25] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, pp. 371–381, 2016.

[26] M. Ahmadipour, H. Hizam, M. Lutfi Othman, and M. Amran Mohd Radzi, "An anti-islanding protection technique using a wavelet packet transform and a probabilistic neural network," *Energies*, vol. 11, no. 10, p. 2701, 2018.

[27] A. M. Jasim, B. H. Jasim, F. Aymen, H. Kotb, and A. Althobaiti, "Consensus-based intelligent distributed secondary control for multiagent islanded microgrid," *International Transactions on Electrical Energy Systems*, vol. 2023, article 6812351, pp. 1–20, 2023.

[28] M. E. Korkmaz, M. K. Gupta, M. Kuntoğlu et al., "Prediction and classification of tool wear and its state in sustainable machining of Bohler steel with different machine learning models," *Measurement*, vol. 223, article 113825, 2023.

[29] G. Vashishtha and R. Kumar, "Unsupervised learning model of sparse filtering enhanced using Wasserstein distance for intelligent fault diagnosis," *Journal of Vibration Engineering & Technologies*, vol. 11, no. 7, pp. 2985–3002, 2023.

[30] G. Vashishtha, S. Chauhan, N. Yadav, A. Kumar, and R. Kumar, "A two-level adaptive chirp mode decomposition and tangent entropy in estimation of single-valued neutrosophic cross-entropy for detecting impeller defects in centrifugal pump," *Applied Acoustics*, vol. 197, article 108905, 2022.

[31] E. Shahryari, M. Nooshyar, and B. Sobhani, "Combination of neural network and wavelet transform for islanding detection of distributed generation in a small-scale network," *International Journal of Ambient Energy*, vol. 40, no. 3, pp. 263–273, 2019.

[32] E. A. Gomes, J. C. Vieira, D. V. Coury, and A. C. Delbem, "Islanding detection of synchronous distributed generators using data mining complex correlations," *IET Generation, Transmission & Distribution*, vol. 12, no. 17, pp. 3935–3942, 2018.

[33] D. Mlakić, H. R. Baghaee, and S. Nikolovski, "A novel ANFIS-based islanding detection for inverter-interfaced microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4411–4424, 2019.

[34] M. Mishra, M. Sahani, and P. K. Rout, "An islanding detection algorithm for distributed generation based on Hilbert–Huang transform and extreme learning machine," *Sustainable Energy, Grids and Networks*, vol. 9, pp. 13–26, 2017.

[35] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, no. 5, p. 272, 2012.

[36] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3133–3181, 2014.

[37] A. Ben-David, "Comparison of classification accuracy using Cohen's weighted kappa," *Expert Systems with Applications*, vol. 34, no. 2, pp. 825–832, 2008.

[38] S. M. Vieira, U. Kaymak, and J. M. Sousa, "Cohen's kappa coefficient as a performance measure for feature selection," in *International Conference on Fuzzy Systems*, Barcelona, Spain, 2010.

[39] P. K. Ray, N. Kishor, and S. R. Mohanty, "Islanding and power quality disturbance detection in grid-connected hybrid power system using wavelet and S-transform," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1082–1094, 2012.

[40] M. A. Farhan and K. Shanti Swarup, "Mathematical morphology-based islanding detection for distributed generation," *IET Generation, Transmission & Distribution*, vol. 10, no. 2, pp. 518–525, 2016.

[41] A. T. Azar, H. I. Elshazly, A. E. Hassanien, and A. M. Elkorany, "A random forest classifier for lymph diseases," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 2, pp. 465–473, 2014.

[42] C. Zhang, Y. Li, Z. Yu, and F. Tian, "Feature selection of power system transient stability assessment based on random forest and recursive feature elimination," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 1264–1268, Xi'an, China, 2016.

[43] A. Samui and S. R. Samantaray, "Wavelet singular entropy-based islanding detection in distributed generation," *IEEE Transactions on Power Delivery*, vol. 28, no. 1, pp. 411–418, 2013.

[44] H. R. Baghaee, D. Mlakić, S. Nikolovski, and T. Dragicčvić, "Anti-islanding protection of PV-based microgrids consisting of PHEVs using SVMs," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 483–500, 2020.

[45] A. M. Niaki and S. Afsharnia, "A new passive islanding detection method and its performance evaluation for multi-DG systems," *Electric Power Systems Research*, vol. 110, pp. 180–187, 2014.

[46] B. Matic-Cuka and M. Kezunovic, "Islanding detection for inverter-based distributed generation using support vector machine method," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2676–2686, 2014.

[47] R. Dubey, M. Popov, and S. R. Samantaray, "Transient monitoring function-based islanding detection in power distribution network," *IET Generation, Transmission & Distribution*, vol. 13, no. 6, pp. 805–813, 2019.

[48] M. Ahmadipour, H. Hizam, M. L. Othman, M. A. M. Radzi, and A. S. Murthy, "Islanding detection technique using slantlet transform and ridgelet probabilistic neural network in grid-connected photovoltaic system," *Applied Energy*, vol. 231, pp. 645–659, 2018.

[49] A. Moeini, A. Darabi, S. M. R. Rafiei, and M. Karimi, "Intelligent islanding detection of a synchronous distributed generation using governor signal clustering," *Electric Power Systems Research*, vol. 81, no. 2, pp. 608–616, 2011.

[50] O. N. Faqhruldin, E. F. El-Saadany, and H. H. Zeineldin, "A universal islanding detection technique for distributed generation using pattern recognition," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1985–1992, 2014.

[51] X. Kong, X. Xu, Z. Yan, S. Chen, H. Yang, and D. Han, "Deep learning hybrid method for islanding detection in distributed generation," *Applied Energy*, vol. 210, pp. 776–785, 2018.

[52] M. Mishra and P. K. Rout, "Fast discrete s-transform and extreme learning machine based approach to islanding detection in grid-connected distributed generation," *Energy Systems*, vol. 10, no. 3, pp. 757–789, 2019.

[53] M. R. Alam, M. T. A. Begum, and B. Mather, "Islanding detection of distributed generation using electrical variables in space vector domain," *IEEE Transactions on Power Delivery*, vol. 35, no. 2, pp. 861–870, 2020.

[54] S. A. Kumar, M. S. P. Subathra, N. M. Kumar et al., "A novel islanding detection technique for a resilient photovoltaic-based distributed power generation system using a tunable-q wavelet transform and an artificial neural network," *Energies*, vol. 13, no. 16, p. 4238, 2020.