01 Nov 2022

A Novel Microgrid Fault Detection and Classification Method using Maximal Overlap Discrete Wavelet Packet Transform and an Augmented Lagrangian Particle Swarm Optimization-Support Vector Machine

Masoud Ahmadipour

Muhammad Murtadha Othman

Rui Bo
Missouri University of Science and Technology, rbo@mst.edu

Zainal Salam

et. al. For a complete list of authors, see https://scholarsmine.mst.edu/ele_comeng_facwork/4567

Recommended Citation

M. Ahmadipour et al., "A Novel Microgrid Fault Detection and Classification Method using Maximal Overlap Discrete Wavelet Packet Transform and an Augmented Lagrangian Particle Swarm Optimization-Support Vector Machine," *Energy Reports*, vol. 8, pp. 4854 - 4870, Elsevier, Nov 2022.
The definitive version is available at https://doi.org/10.1016/j.egyr.2022.03.174

Energy Reports 8 (2022) 4854–4870

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Research paper

A novel microgrid fault detection and classification method using maximal overlap discrete wavelet packet transform and an augmented Lagrangian particle swarm optimization-support vector machine

Masoud Ahmadipour [a,*], Muhammad Murtadha Othman [a,*], Rui Bo [b], Zainal Salam [c], Hussein Mohammed Ridha [d], Kamrul Hasan [a]

[a] School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[b] Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
[c] Centre of Electrical Energy Systems, School of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia
[d] Advanced Lightning, Power & Energy Research (ALPER), Faculty of Engineering, Universiti Putra Malaysia, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 10 January 2022
Received in revised form 17 March 2022
Accepted 23 March 2022
Available online xxxx

Keywords:
Microgrids
Fault detection
Maximal overlap discrete wavelet packet transform
Augmented Lagrangian particle swarm optimization
Support vector machine
Real time digital simulator

ABSTRACT

In this paper, an intelligent method for fault detection and classification for a microgrid (MG) was proposed. The idea was based on the combination of three computational tools: signal processing using the maximal overlap discrete wavelet packet transform (MODWPT), parameter optimization by the augmented Lagrangian particle swarm optimization (ALPSO), and machine learning using the support vector machine (SVM). The MODWPT was applied to preprocess half cycle of the post-fault current samples measured at both ends of feeders. The wavelet coefficients derived from the MODWPT were statistically evaluated using the mean, standard deviation, energy, skewness, kurtosis, logarithmic energy entropy, max, min, and Shannon entropy. These were the input feature datasets and were used to train the SVM classifier. The ALPSO was utilized to reduce the feature subsets and select the sensitive parameters of the SVM (i.e., penalty factor and the slack variable) to further improve the performance of the SVM. The intelligent relaying scheme was executed on a real-time digital simulator (RTDS) which is integrated with Matlab. The performance of SVM-based protection method is compared to several different protection models in terms of signal processing tools, optimization techniques used for selecting datasets and sensitive parameters, and classifiers under different operating conditions. Numerous operating conditions, including islanded or non-islanded operation modes and radial and or loop topologies introducing different characteristics of fault were included as the case studies for the proposed technique. A comprehensive evaluation study of the consortium for electric reliability technology solutions (CERTS) MG system and IEEE 34-bus confirms that the proposed protection scheme is accurate, fast, and robust to noisy measurements. In addition, the obtained results illustrate that the proposed method is superior to the recently published works in the literature.

1. Introduction

Microgrids (MG) distributed generation and renewable energy (RE) will play important and distinctive roles in future energy concepts (Hirsch et al., 2018). These combinations are becoming an integral part of modern electrical power distributed systems (EPDS) that were designed to reduce carbon emissions, increase reliability, diversify the energy resources, and reduce cost (Hirsch et al., 2018; Lopez-Garcia et al., 2020). MG are controllable and operate in either grid connected mode or islanded mode. They are envisaged to offer significant benefits to customers and power generation companies. From the customer's point of view, MG enhance system efficiency, improve power quality, and increase reliability with reduced carbon emissions. From the perspectives of utility companies, they can be deployed to decrease consumption demands and to eliminate electricity consumption during peak hours, thus resulting in increased network losses (Gabbar, 2016). Despite the rapid development of MG in recent years, there are still technical issues pertaining to the MG design, control, and operation. One significant operational challenge is the protection of MG, especially when considering the amalgamation of distributed energy resources. The failure of protection systems can result different fault current levels, particularly in non-grid or islanded operations (Chakravorti et al., 2019; Chandra et al., 2021). This may cause maloperation of the protection relays. The

* Corresponding authors.
E-mail addresses: masoud@uitm.edu.my (M. Ahmadipour), mamat505my@yahoo.com (M.M. Othman).

https://doi.org/10.1016/j.egyr.2022.03.174

key to resolve this issue is to enhance the fault classification and detection mechanisms.

Various intelligent schemes were proposed in the literature regarding fault classification and detection in MG (Casagrande et al., 2013a,b; Hooshyar et al., 2015; Ahmadipour et al., 2019b; Dehghani et al., 2016; Bakar et al., 2014). In Casagrande et al. (2013a), a combination of Naïve Bayes (NB) and decision tree (DT) classifiers was utilized for fault detection in an islanded MG. The detection of symmetrical and unsymmetrical faults as well as high impedance faults in MG using a differential energy-based protection scheme were represented in Casagrande et al. (2013b). In another work, the differential sequence component protection method to protect the microgrids was investigated in Hooshyar et al. (2015). Ahmadipour et al. (2019b) relied on modified Slantlet transform and Ridgelet probabilistic neural network for grid fault detection in MG regarding grid-connected and islanded modes. The method proposed in Ahmadipour et al. (2019b) was based on a combination of wavelet singular entropy theory and fuzzy logic for detecting fault events. In Dehghani et al. (2016), Shannon entropy was extracted as a feature by wavelet transform and applied as input dataset feature vectors for fuzzy logic classifiers to detect and classify the fault events. In Bakar et al. (2014), a combined directional overcurrent and an earth fault protections were utilized to protect the MG. Nonetheless, it did not react quickly to trip fault signals. The approach proposed in Shafiullah and Abido (2018) relied on the extraction of the suitable features using S-transform from the measured current signals and conducted them to multilayer perceptron feedforward neural networks for grid faults detection and classification. James et al. (2017) presented wavelet transform in combination with deep neural networks to classify grid faults in MGs. A wavelet in combination with data-mining based fault detection technique was presented for MG in Mishra et al. (2015). The matching pursuit technique was proposed to find the perfect match for the signal's structure in Abdelgayed et al. (2017b) and Misiti et al. (1997-2015). The primary demerit of these works was the scalability limitation, so the identified set of wavelets for each signal grew with increased signals. To identify faults events, authors in Ahmadipour et al. (2018b) combined the Slantlet transform with different machine learning tools, including Ridgelet probabilistic neural network (RPNN), and the probabilistic neural network (PNN) was utilized by Ahmadipour and Hizam (2019). A combination of wavelet multi-resolution singular spectrum entropy and support vector machine was applied to classify different fault types (Ahmadipour et al., 2019a). In Singh et al. (2021), to classify different faults events in a 9-bus MG system, a fault detection method based on wavelet transform analysis and wavelet entropy approaches was utilized. A combination of the maximum overlap discrete wavelet transform (MODWT) and boosted tree method have been utilized for microgrid fault detection and classification (Patnaik et al., 2021). The feature extraction process was done by MODWT and differential energy of three-phase current and its zero-sequence current component at each of the decomposition levels of MODWT were calculated as input to an Extreme Gradient Boost (XGBoost) based machine learning to detect and classify the fault events in the microgrid. In Baloch and Muhammad (2021), the authors employed a Hilbert Transform and data mining approach to protecting the microgrid. Other intelligent techniques have been presented for microgrid fault detection and classification such as a combination of data mining and wavelet multiresolution analysis (Baloch et al., 2021), a combination of the wavelet transform, and Taguchi-based artificial neural network (Hong and Cabatac, 2019), and Fourier transform with machine learning algorithm (Ezzat et al., 2021).

Notwithstanding the numerous efforts to provide secure and intelligent fault detection methods, there is always room for improvement for more effective protection of MGs. The main interest is to increase the classification accuracy and to decrease the computational time that will improve the performance of technique in solving the MG fault detection and classification. In this paper, a novel method based on machine leaning technique was proposed for an improved fault classification and enhanced detection for AC microgrid protection systems. The idea was to combine a powerful signal processing tool, known as the maximal overlap discrete wavelet packet transform, (MODWPT) and the support vector machine (SVM) classifier. The MODWPT is more immune to noise and power disturbances compared to discrete wavelet transform (DWT) and has been successfully used in other fault detection works but not in MG, for example (Alves et al., 2016; Bagheri et al., 2017). It provides a uniform frequency band and processes the time-invariance properties that is ideal for real time estimation. The wavelet coefficients of MODWPT are analyzed using nine statistical instruments, namely the mean, standard deviation, energy, skewness, kurtosis, logarithmic energy entropy, max, min, and Shannon entropy. These statistical features were used as the input feature datasets for the fault detections and classifications by the SVM. However, it is known that the classifier efficiency is influenced by dataset size. Larger data sizes are preferable, but at the expense of computing times it will deprive the performance of protection method. Furthermore, the SVM is highly sensitive to control parameters' values, namely the penalty factor (C) and the slack variable (ξ). The augmented Lagrangian particle swarm optimization (ALPSO) is also used to select the best values of C and ξ and to optimize the sizes of the input feature datasets. An electric reliability technology solutions (CERTS) MG system (Lasseter et al., 2010) and a IEEE −34 bus test system is modeled in RSCAD that is a simulation software designed for interfacing with the RTDS platform. To validate the performance of the proposed MODWPT-ALPSO-SVM classifier, its performance is compared with the four well-known classifiers, i.e., decision tree (DT), k-nearest neighbor (K-NN), probabilistic neural network (PNN), and Naïve Bayes (NB). Moreover, in order to evaluate the effectiveness of the MODWPT and ALSPO used in the proposed algorithm, they are replaced with the signal processing methods (i.e.; discrete wavelet transform (DWT), Hilbert Huang (HH), fast discrete s-transform (FDST), and Variational Mode Decomposition (VMD)) and meta-heuristic techniques (i.e., the particle swarm optimization (PSO), grasshopper optimization algorithm (GOA), gray wolf algorithm (GWA), cuckoo search (CS), political optimizer (PO), Aquila optimizer (AO), and African vultures' optimization algorithm (AVOA)) separately in order to obtain the results required for comparison. Additionally, the investigation was extended using the following criterion: (1) grid-connected mode and islanding mode, (2) radial and loop topology, (3) all symmetrical and asymmetrical faults and high impedance fault (HIF), (4) various fault resistances, (5) different inception angles, (6) different fault locations for four different power lines of the MG, (7) pre-fault loading conditions, and (8) measurement noise presences.

The remaining work is organized as follows: Section 2 described a case study of MG system to analyze the proposed technique; Section 3 presents the theoretical background, and Section 4 focus on the proposed ALPSO-SVM scheme. Section 5 describes the obtained results and outlined the discussion. The conclusion is in Section 6.

2. AC microgrid under study

Fig. 1 shows the system used for the case study that is the consortium for electric reliability technology solutions (CERTS) MG and the details of which can be found in Lasseter et al. (2010). The MG operated at 0.48 kV, and 60 Hz; it consisted of two

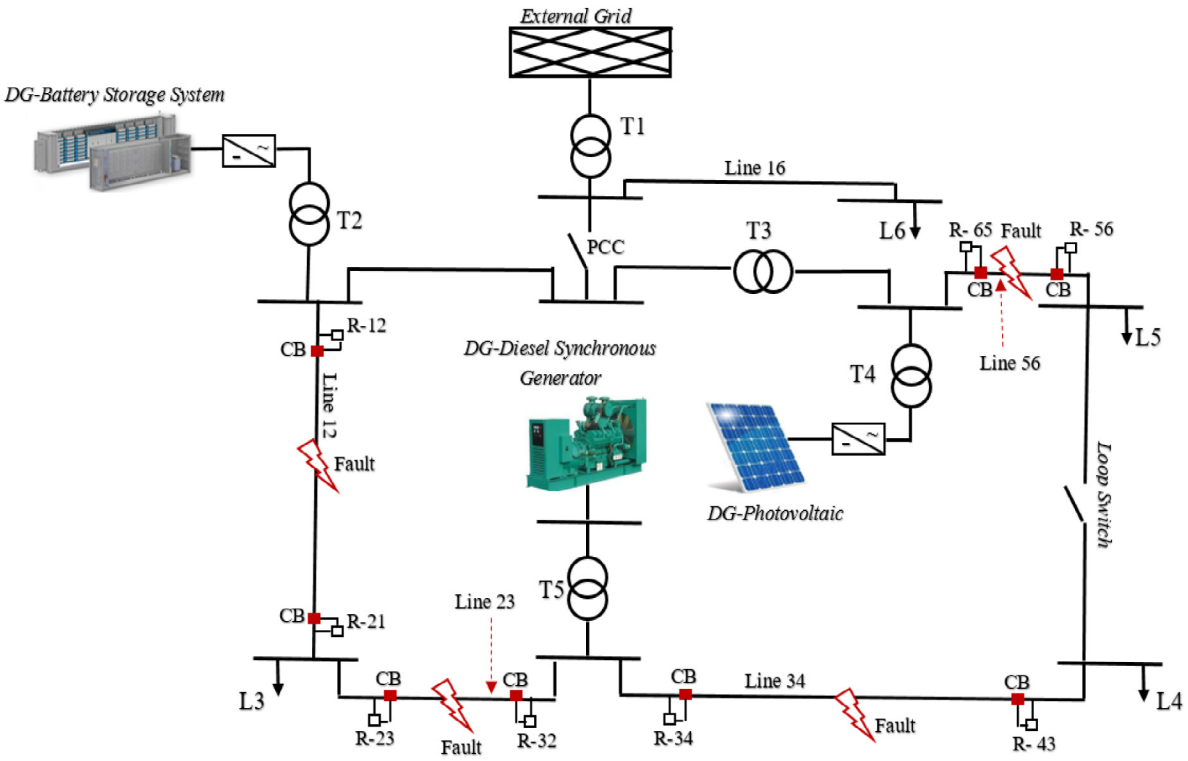

Fig. 1. The studied microgrid system structure.

operating modes: the islanded and non-islanded. It supported loads in both modes, and it was controlled through the common coupling (PCC) state point switch. During grid-connected mode, the system was supplied by the step-down 3-phase distribution transformer (rated 13.8/0.48 kV); during islanding mode, the MG was powered by the distributed generators (DGs). Furthermore, there was a loop switch that enabled the system to be operated in either the loop or radial topology. To investigate the viability of the proposed protection scheme, the system was simulated in RSCAD/RTDS.

Three types of DG sources, namely, solar photovoltaic (DG-photovoltaic), DG-battery storage, and the DG-diesel synchronous generator were considered. The DG-photovoltaic and DG-battery were based on voltage source converters. While in the in grid connected mode, the former was controlled by the modified current-mode control with DC link voltage controller in the *d-q* frame. When operated in the islanding mode, the frequency-mode control in *d-q* frame was utilized. By contrast, the DG-battery storage was controlled by current-mode control. It was based on the active/reactive power controller in the *d-q* frames in both operation modes. The details about these schemes can be found in Yazdani and Iravani (2010) and Yazdani and Dash (2009).

In addition to these sources, the system included loads and transformers. The details about the components models can be found in James et al. (2017) and Abdelgayed et al. (2017b). Apart from that, at both ends of the transmission power lines (i.e., Line 12, 23, 34, and 56) digital protective relays were installed. Lines 12, 34, 56 were AWG2 type with lengths equal to 68.58 m, while Line 23 was AWG00 type with length equal to 22.86 m. Similar to previous works (James et al., 2017; Abdelgayed et al., 2017b), the focus was on these four mentioned lines, as denoted in Fig. 1. Using current transformers, these relays sampled the current magnitudes at 3.84 kHz, which was in accordance with Guillén et al. (2016), Jiang et al. (2002). The configuration in Guillén et al. (2016) utilized current transformers to obtain the incremental differential currents. These values were normalized and used as input signals to signal processing tool. The MODWPT was applied to the sampled current signals for half cycle after the fault happened.

3. Theoretical background

This section explains the overview of various tools and methods that are utilized to develop the proposed microgrid fault detection and classification scheme. The first stage of the proposed method involves wavelet tools that are applied to preprocess half cycle of the post-fault current samples measured at both ends of feeders. The wavelet coefficients derived from the wavelet tool are statistically evaluated using the mean, standard deviation, energy, skewness, kurtosis, logarithmic energy entropy, max, min, and Shannon entropy. These are the input feature datasets and are used to train the SVM classifier. The main concept behind the SVM is given in Section 3.2. The ALPSO is utilized to reduce the feature subsets and select the sensitive parameters of the SVM (i.e., penalty factor and the slack variable) to further improve the performance of the SVM classifier for fault detection and classification. Meanwhile, the main steps of ALPSO are summarized in Section 3.3.

3.1. Wavelet transform analysis

To extract the significant features of the MGs faults (from the obtained data), suitable signals for processing and conditioning were pre-requisites. The discrete wavelet packet transform (DWPT) is a digital signal processing technique that offers a wide range of possibilities for signal analysis. It has the capabilities to decompose both the scaling and wavelet coefficients at a specific decomposition level (Lasseter et al., 2010). In addition, the constant frequency band makes it suitable for power estimation applications. The DWPT is primarily used to extract the important features as input data for the Support Vector Machine. However, DWPT is a time-variant transformation; this leads to problems

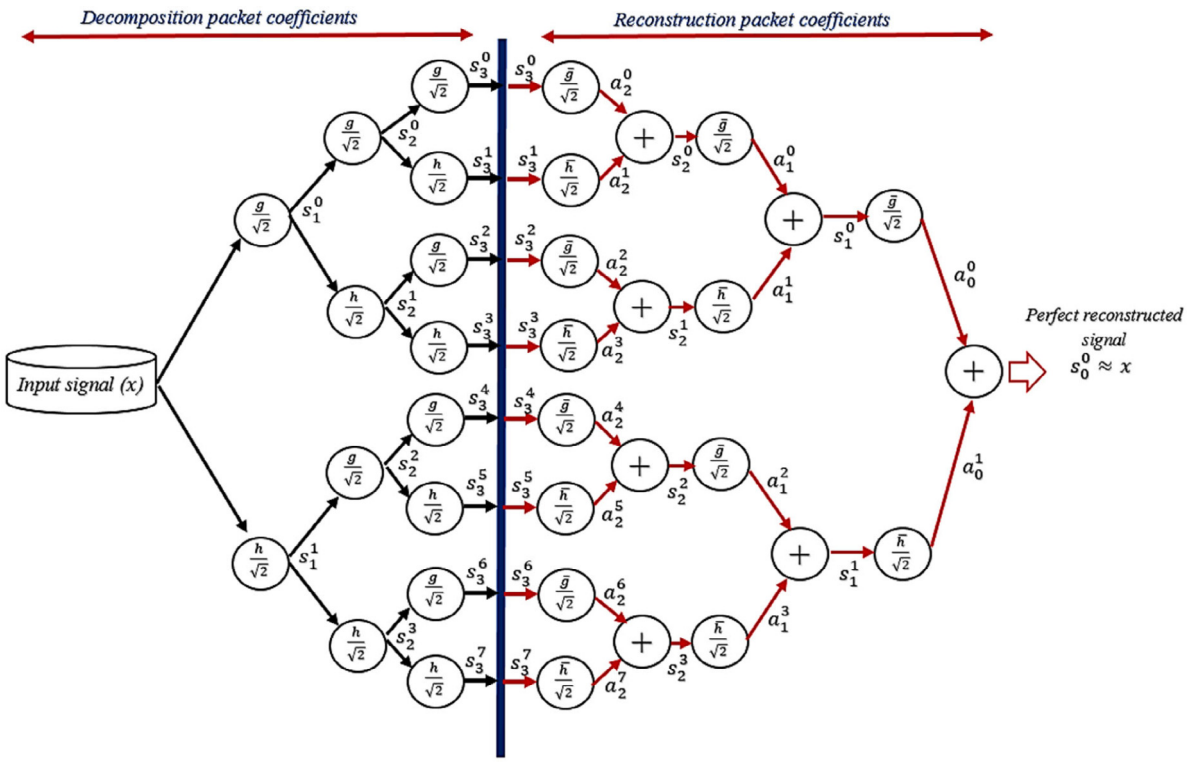

Fig. 2. The structure of MODWPT with three decomposition and reconstruction levels.

in the real-time detection of non-stationary signals (Alves et al., 2016). For this reason, in this paper, the maximal overlap discrete wavelet packet transform (MODWPT) was used to overcome the drawback of the conventional DWPT (Zhao and Ye, 2010; He, 2013).

3.1.1. Design of the MODWPT

In principle, the MODWPT design is similar to DWPT. The low pass and high pass filters are applied to the input signal at each level, thereby presenting uniform frequency output bands. This transformation can exert any sample size n. However, in contrast to the DWPT which is a time variant transformation, there is no down-sampling by factor of two in MODWPT (i.e., time-invariant transform) (Alves et al., 2016; Costa, 2014). In the reconstruction, the decomposition coefficients were convolved to the reverse low- and high-filters to reconstruct the original signal. Fig. 2 illustrates the MODWPT decomposition and reconstruction tree with three decomposition levels.

The MODWPT decomposition coefficients can be defined as follows (Alves et al., 2016):

$$S_{2z}(t) = \frac{1}{\sqrt{2}} \sum_{n=0}^{l-1} g[n] S_z(t+n-l+1) \tag{1}$$

$$S_{2z+1}(t) = \frac{1}{\sqrt{2}} \sum_{n=0}^{l-1} h[n] S_z(t+n-l+1) \tag{2}$$

in which the length of filters is l. t is always the current sampling, i.e., there are no samples regarding an index above t, and the computational burden must be less than $1/f_s$ seconds, where the sampling rate is f_s.

The reconstruction coefficient of MODWPT is denoted as:

$$a_{2z}(t) = \frac{1}{\sqrt{2}} \sum_{n=0}^{l-1} \overline{g}[n] S_{2z}(t-n) \tag{3}$$

$$a_{2z+1}(t) = \frac{1}{\sqrt{2}} \sum_{n=0}^{l-1} \overline{h}[n] S_{2z+1}(t-n) \tag{4}$$

Reconstructing the original signal x is formulated as follows:

$$S_Z(t) = a_{2z}(t) + a_{2z+1}(t) \tag{5}$$

in which $x = S_0$.

A carefully choosing wavelet functions is helpful to accurately extract the vector of the suitable training classification techniques, to precisely detect the events in the studied MG, and to enhance the reliability of system. Various wavelet functions were applied in microgrid fault detection such as Daubechies (db), symlets (sym), coiflets (coif), biorthogonal (bior) and haar (James et al., 2017; Mishra et al., 2015). These can affect the feature extraction capability of MODWPT due to having its own unique time–frequency domain characteristics (Alves et al., 2016). Hence, wavelet functions should be strategically selected based on the properties of the analyzed data (Alves et al., 2016). In this paper, ten wavelet members in the two wavelet functions Daubechies (db) and biorthogonal (bior), as the mother wavelets, were employed to transform the input signal. Moreover, considered MODWPT decomposition and reconstruction coefficients at level 4 were given to get the smooth data.

3.1.2. Feature extraction stage

The purpose of feature extraction is to specify the inimitable characteristics of the current signals that can be utilized to discriminate between different fault situations. To extract these features, the MODWPT was applied. The chronological order for the extraction was described in the following steps:

Step 1. The line current signals are acquired from current transformers (CTs) to obtain the incremental differential currents

Table 1
Statistical features.

Parameters	Label	Corresponding equations
Energy	E	$E = \sum S^2$
Shannon entropy	SE	$SE = -\sum \lvert S\rvert^2 \log \lvert S\rvert^2$
Mean	μ	$\mu_s = \frac{1}{N}\sum S$
Standard deviation	σ	$\sigma_s = \left[\frac{1}{N}\sum (S-\mu_S)^2\right]^{1/2}$
Minimum	Min	$Min_s = Min\{S\}$
Maximum	Max	$Max_s = Max\{S\}$
Skewness	$Skew$	$Skew_s = \frac{1}{N-1}\sum \left(\frac{(S-\mu_S)}{\sigma_s}\right)^3$
Kurtosis	$Kurt$	$Kurt_s = \frac{1}{N-1}\sum \left(\frac{(S-\mu_S)}{\sigma_s}\right)^4$
Logarithmic energy entropy	LEE	$LEE_s = \sum \log[s]^2$

N is the number of sampling point and S is the MODWPT decomposed coefficients.

(Guillén et al., 2016). Then, the incremental differential currents were normalized and used as input signals to MODWPT.

Step 2. The obtained current signals were decomposed by MODWPT into a series of coefficients using each of the mentioned mother wavelets in all decomposition levels.

Step 3. The statistical features of the MODWPT coefficients that included the essential data of the considered fault was calculated. These datasets were used as the input feature vector for machine learning algorithm. The statistical features of the decomposition coefficients were mean, standard deviation, energy, skewness, kurtosis, Logarithmic energy entropy, max, min, and Shannon entropy. They were computed using the equations shown in Table 1.

Step 4. Ten wavelet members in the two wavelet functions db and bior, as the mother wavelets, were employed to convert the input signal. Moreover, the considered MODWPT decomposition and reconstruction coefficients at level 4 were given to obtain the smooth data. Hence, MODWPT was applied to the sampled line current signals for half cycle after the fault. Next, the equations in Table 1 are used to determine several statistical features for wavelet coefficients derived from MODWPT coefficients. The feature descriptions were as follows:

- $F1$- energy of phase A.
- $F2$- energy of phase B.
- $F3$- energy of phase C.
- $F4$- Shannon entropy of phase A.
- $F5$- Shannon entropy of phase B.
- $F6$- Shannon entropy of phase C.
- $F7$- standard deviation of phase A.
- $F8$- standard deviation of phase B.
- $F9$- standard deviation of phase C.
- $F10$- mean of phase A.
- $F11$- mean of phase B.
- $F12$- mean of phase C.
- $F13$- skewness of phase A.
- $F14$- skewness of phase B.
- $F15$- skewness of phase C.
- $F16$- kurtosis of phase A.
- $F17$- kurtosis of phase B.
- $F18$- kurtosis of phase C.
- $F19$- max of phase A.
- $F20$- max of phase B.
- $F21$- max of phase C.
- $F22$- min of phase A.
- $F23$- min of phase B.
- $F24$- min of phase C.
- $F25$- Logarithmic energy entropy of phase A.
- $F26$- Logarithmic energy entropy of phase B.
- $F27$- Logarithmic energy entropy of phase C.

Step 5. The obtained feature vectors were normalized and used as input data for the proposed classifier for fault detection and classification. Features $F1$ to $F18$ are given as input along with the class to build a ALPSO-SVM model for fault detection. The classes used to build detection SVM are fault and non-fault.

Step 6. To train a ALPSO-SVM model to classify a fault, all features F1 to F27 are given as inputs along with the class of fault.

The construction of the proposed protection scheme is shown in Fig. 3.

3.2. Support vector machine design

The support vector machine (SVM) is a powerful tool to solve regression, classification, and pattern recognition problems. It is a supervised learning machine that was introduced by Vapnik in 1995 (Cortes and Vapnik, 1995). The SMV is a binary classifier. It was developed by specifying a decision boundary to separate the training cases into their related classes. To distinguish the fault and non-fault cases, SVM produces a hyperplane to divide information into their separate classes in a d-dimensional feature space using nonlinear decision boundary.

Consider the dataset as follows:

$$\{x_i, y_i | i = 1, 2, \ldots, N\}, x_i \epsilon R^d$$
$$y_i \epsilon \{-1, 1\} \tag{6}$$

in which x_i denotes the independent variables, y_i denotes the dependent variables, and N is the number of sampling. A class decision function linked with a suitable hyperplane is assumed to be:

$$f(x) = w^T x + b = \sum_{i=}^{N} w.x_i + b = 0 \tag{7}$$

in which w and b represent the weight vector and the bias term, respectively. These variables are utilized to determine the position of the splitting hyperplane, that ought to fulfill the following constrains:

$$y_i f(x) = y_i \left(w^T x_i + b\right) \geq +1 \quad for \quad i = 1, 2, \ldots .N \tag{8}$$

Mathematically, the optimal separating hyperplane can be calculated as follows:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i, \quad i = 1, 2, 3, \ldots, N \tag{9}$$

which is subject to

$$\begin{cases} y_i (w.x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \tag{10}$$

where x_i represents the input feature vector with the class label $y \epsilon \{-1, 1\}$. Parameters of C and ξ_i are the error's penalty factor and the slack variable, respectively.

To solve the non-linear classification problems, kernel functions are utilized. Using the non-linear vector function can map input vector x from n-dimensional to m-dimensional:

$$\varphi(x) = \varphi_1(x), \varphi_2(x), \varphi_3(x) \ldots, \varphi_m(x) \quad m \gg n. \tag{11}$$

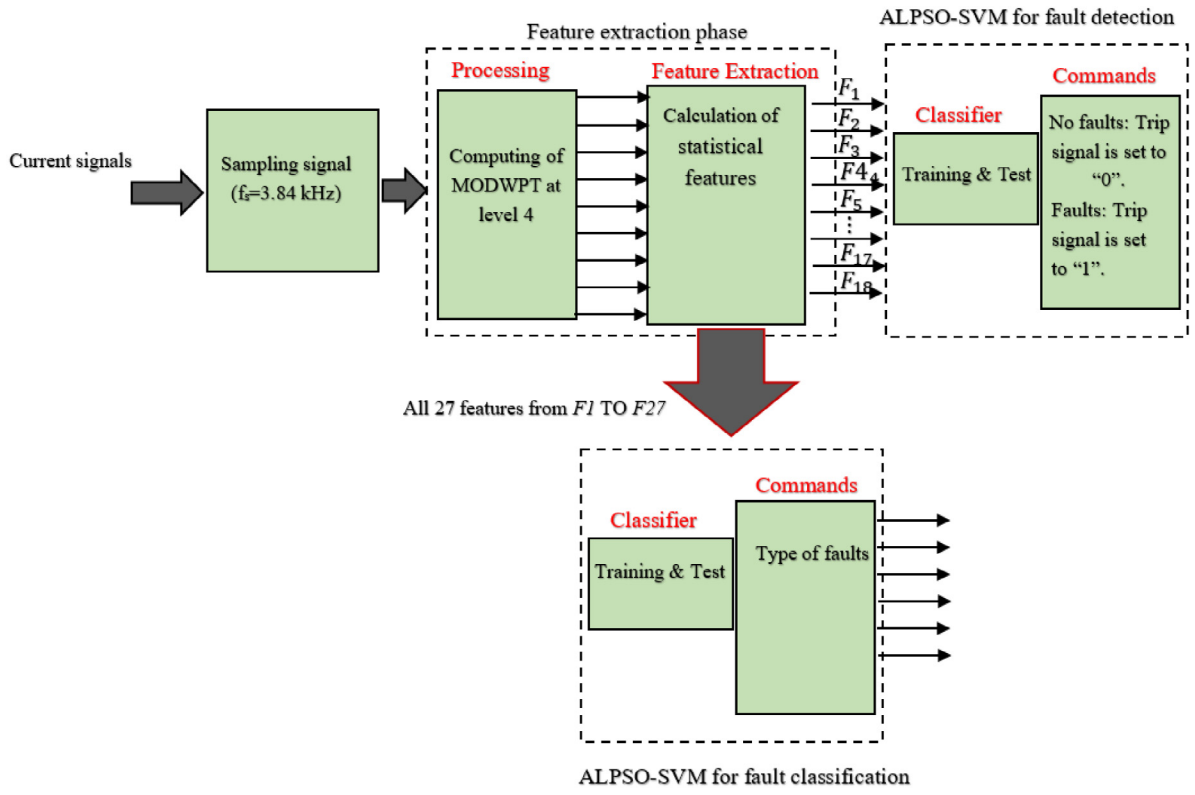

Fig. 3. The construction of the proposed protection scheme.

Thus, the decision function of the SVM is obtained using the following equation:

$$f(x) = sign\left(\sum_{i=1}^{N} \lambda_i y_i K(x, x_i) + b\right) \tag{12}$$

where λ_i denotes the Lagrangian multiplier ($0 \le \lambda_i \le C$). $K(x, x_i) = \varphi^T(x)\varphi(x)$ represents a kernel function. The linear, polynomial, Gaussian radial basis, and sigmoid functions are the most commonly kernel functions that are utilized in the SVM (Motlagh and Foroud, 2021). Each of them has inimitable characteristics. Gaussian radial basis function (RBF) was considered the kernel function in this paper. It has an outstanding performance in practice in comparison with other kernels used and can calibrate easily (Motlagh and Foroud, 2021). The RBF is defined as the following equation:

$$K(x, x_i) = e^{(-\gamma \|x - x_i\|^2)} \tag{13}$$

in which γ is the kernel parameter with the following value:

$$\gamma = \frac{1}{2\sigma^2} \tag{14}$$

where σ denotes the width parameter of the radial basis function.

The feature selection plays an essential role in pattern recognition. However, the SVM classifiers were faced with restrictions regarding feature dataset and optimizing parameter selections. Datasets include irrelevant and unnecessary components that can affect the classifier performances and increase the response times. Thus, it was preferable to find the subset of features that were deemed to be more important than the others. Merits of feature subset selection included (1) decreased size and storage requirements, (2) simplified data grasp and conception, (3) decreased computational cost, (4) enhanced the classification accuracy (due to the removal of unrelated data), and (5) improved generalization by decreasing over-fitting. Moreover, there were sensitive parameters in SVM that can affect performance: the penalty factor C, slack variable ξ_i, the type of kernel function. Hence, choosing a suitable combination among them was vital to improve the SVM performance. In this study, the augmented Lagrangian particle swarm optimization (ALPSO) was applied to address the aforementioned problems.

3.3. Augmented Lagrangian particle swarm optimization

The particle swarm optimization (PSO) is one of the most interesting global optimal algorithms owed to its search abilities and flexibility in various solution spaces. It was first was introduced by Eberhart in Kennedy and Eberhart (1995). The algorithm was inspired by the swarm behavior, where the elements coordinated motion throughout maneuvers for actions, such as searching food and staging defense. The PSO procedure can be summarized as follows:

Step 1. Initialized PSO particle position and velocity and randomly initialized the values of particle position $p.p$ and velocity vel for each decision variable, as the following equations:

$$\begin{cases} p.p_{i,j} = X_{jmin} + rand \in (0,1) \times (X_{jmax} - X_{jmin}) & 0 < i < N_p \\ vel_{i,j} = X_{jmin} + rand \in (0,1) \times (X_{jmax} - X_{jmin}) & 0 < j < N_{vr} \end{cases} \tag{15}$$

in which N_p and N_{vr} represent the number of particles and the number of decision variables respectively. Furthermore, the maximum and the minimum values of each decision variable are denoted by X_{max} and X_{min}, respectively. $rand \epsilon (0, 1)$ is a random number that uniformly distributes in the interval [0,1].

Step 2. Determined the best position of particle $p.p_{i,j}^{best,k=0}$ and the best global position $p.p_{swarm}^{best,k=0}$ that had the lowest objective

functions for value of each particle and in the whole swarm, respectively.

Step 3. Checked termination criteria. If fulfilled, the algorithm terminated with the solution $p.p^* = p.p_{swarm}^{best,k}$, otherwise next step was applied.

Step 4. In each iteration, the velocity and particle position were updated and evaluated regarding the corresponding objective function in each position. At iteration k, the velocity and each new particle position were determined as the follows:

$$vel_{i,j}^{k+1} = \mathcal{W}.vel_{i,j}^{k} + c_1 \times rand\ \epsilon\ (0.1) \times \left(p.p_{i,j}^{best,k} - p.p_{i,j}^{k}\right) + c_2 \times rand\ \epsilon\ (0.1) \times (p.p_{swarm}^{best,k} - p.p_{i,j}^{k}) \quad (16)$$

$$p.p_{i,j}^{k+1} = p.p_{i,j}^{k} + vel_{i,j}^{k+1} \quad (17)$$

where $p.p_{i,j}^{best,k}$ and $p.p_{swarm}^{best,k}$ are the best particle position and the best global position at iteration k, respectively. The term $c_1 \times rand\epsilon\ (0.1) \times \left(p.p_{i,j}^{best,k} - p.p_{i,j}^{k}\right)$ is based on cognition, as it considers only the best location of the particle's own experience. The term $c_2 \times rand\epsilon\ (0.1) \times (p.p_{swarm}^{best,k} - p.p_{i,j}^{k})$ denotes the social particles interaction. Thus, c_1 and c_2 are referred to as cognitive and social knowledge values, respectively. $\mathcal{W}$ is the inertia factor.

Step 5. Stopped condition. If the iteration number $k < k_{max}$ (maximum number of iterations), increment and go back to step 3; otherwise, update the best and global best position of the particles with the minimum value of objective function as the final solution.

Despite of the merits of PSO, such as efficiency, robustness, and simplicity, it was faced with a velocity control mechanism issue (Eberhard and Sedlaczek, 2009). If the values of velocity are small, the particles investigate only their local areas, thereby increasing the possibility of being trapped in local minima. By contrast, the large velocity values tend to cause particles to leave the defined boundary constraints of the issue and to the divergence from the swarm. Hence, to overcome the mentioned problem, an augmented Lagrangian particle swarm optimization (ALPSO) for constrained optimization problems was proposed (Eberhard and Sedlaczek, 2009). Mathematically, ALPSO was formulated as follows:

$$\min_x f(x), \quad x\epsilon\mathbb{D}\bigcap\mathbb{F}, \quad \mathbb{D} \subseteq \mathcal{R}^n \quad (18)$$

which is subject to:

$$\begin{cases} g\,(x) = 0, & g : \mathcal{R}^n \to \mathcal{R}^{me} \\ h\,(x) \le 0, & h : \mathcal{R}^n \to \mathcal{R}^{mi} \end{cases} \quad (19)$$

in which $f(x)$ is the nonlinear objective function, that is to minimize based on the design variable x. Furthermore, it is subject to the nonlinear equality $g(x)$ and unfairness constrains $h(x)$. $\mathbb{F}$ represents the achievable region, and $\mathbb{D}$ denotes the search space that is also bounded by the simple bounds $x_{low} \le x \le x_{upper}$.

To transform the mentioned constrained optimization into an unconstrained optimization problem, the augmented Lagrange multiplier technique is:

$$L\,(x, \lambda, \beta) = f\,(x) + \sum_{i=1}^{me+mi} \lambda_i \theta_i\,(x) + \sum_{i=1}^{me+mi} \beta_i \theta_i^2(x) \quad (20)$$

with

$$\theta_i = \begin{cases} g_i\,(x)\,, & 1 \le i \le me \\ \max\left[h_{i-me}\,(x)\,, \dfrac{-\lambda_i}{2\beta_i}\right], & me+1 \le i \le me+mi \end{cases} \quad (21)$$

and

$$\lambda_i = [\lambda_1, \lambda_2, \ldots, \lambda_{me+mi}]^T \epsilon \mathcal{R}^{me+me} \quad (22)$$

$$\beta_i = [\beta_1, \beta_2, \ldots, \beta_{me+mi}]^T \epsilon \mathcal{R}^{me+me} \quad (23)$$

where λ_i, and β_i denote the Lagrange multipliers and the penalty factors, respectively. Based on continuous and differentiable problems, the term $\frac{-\lambda_i}{2\beta_i}$ is selected for continuous derivatives $\frac{\partial L}{\partial x}$ at $\hat{x}$ in which $h_{i-me}\left(\hat{x}\right) = \frac{-\lambda_i}{2\beta_i}$. Moreover, the term $\sum_{i=1}^{me+mi} \beta_i \theta_i^2(x)$ guarantees that x^* is a stationary point of L for the correct Lagrange multiplier λ^*. Nevertheless, x^* in which the symbol $*$ shows the best value of variables is a minimum of L. In addition, the correct Lagrange multipliers λ^* and the suitable penalty factors β^* are problem dependent, therefore unfamiliar. The solution x^* is not able to be directly calculated by a signal unconstrained minimization of Eq. (20). Thus, an update scheme for Lagrange multipliers and the penalty factors is applied as follows:

$$\lambda_i^{j+1} = \lambda_i^j + 2\beta_i^j \theta_i(x^j) \quad (24)$$

$$\beta_i^{j+1} = \begin{cases} 2\beta_i^j & if \quad \left|g_i(x^j)\right| > \left|g_i(x^{j-1})\right| \wedge \left|g_i(x^j)\right| > \epsilon_g, \\ \dfrac{1}{2}\beta_i^j & if \quad \left|g_i(x^j)\right| \le \epsilon_g, \\ \beta_i^j & else, \end{cases} \quad (25)$$

$$1 \le i < me$$

$$\beta_{l+me}^{j+1} = \begin{cases} 2\beta_{l+me}^j & if \quad \left|h_l(x^j)\right| > \left|h_l(x^{j-1})\right| \wedge \left|h_l(x^j)\right| > \epsilon_h, \\ \dfrac{1}{2}\beta_{l+me}^j & if \quad \left|h_l(x^j)\right| \le \epsilon_h, \\ \beta_{l+me}^j & else, \end{cases}$$

$$1 + me \le l < me + mi \quad (26)$$

in which ϵ_g and ϵ_h denote the user-defined tolerance for acceptable constraint violations. Moreover, x^j mentioned in Eqs. (24)–(26) is updated as follows:

$$vel_{i,j}^{k+1} = \mathcal{W}.vel_{i,j}^{k} + c_{i,j}^{k} \times rand\epsilon\ (0.1) \times \left(p.p_{i,j}^{best,k} - p.p_{i,j}^{k}\right) + d_{i,j}^{k} \times rand\epsilon\ (0.1) \times (p.p_{swarm}^{best,k} - p.p_{i,j}^{k}) \quad (27)$$

$$p.p_{i,j}^{k+1} = p.p_{i,j}^{k} + vel_{i,j}^{k+1} \quad (28)$$

Hence,

$$p.p_{i,j}^{best,k} := arg \min_{x_i^p}\{L\left(x_i^p, \lambda^j, \beta^j\right), 0 \le p \le k\} \quad (29)$$

$$p.p_{swarm}^{best,k} := arg \min_{x_i^k}\{L\left(x_i^k, \lambda^j, \beta^j\right), \forall i\} \quad (30)$$

The $p.p_{i,j}^{best,k}$ is the best obtained position of the *ith* particle in optimization process. The $p.p_{swarm}^{best,k}$ represents the best position in the whole swarm at the current *th* iteration. $\mathcal{W}$ is the inertia weight.

Regarding Eqs. (24), (25) and (26), $x^j = p.p_{swarm}^{best,kmax}$, in which k_{max} is a predefined iteration number. The flowchart of the ALPSO implementation is represented in Fig. 4.

4. The proposed ALPSO-SVM scheme

To improve the fault detection and classification performance of the SVM classifier, the ALPSO was applied. In addition, the ALPSO can also be used to decrease the redundancy of the input feature dataset. The detailed procedure of the proposed scheme was as follows.

Step 1. The feature data set matrix was obtained, and the kernel function of SVM classifier was decided. In this paper, RBF kernel function was used. It examined higher dimensional data and needed only two parameters, C and ξ that are the penalty factor and the slack variable, respectively. Then, the features used as input attributed and the parameters (C and ξ) were optimized

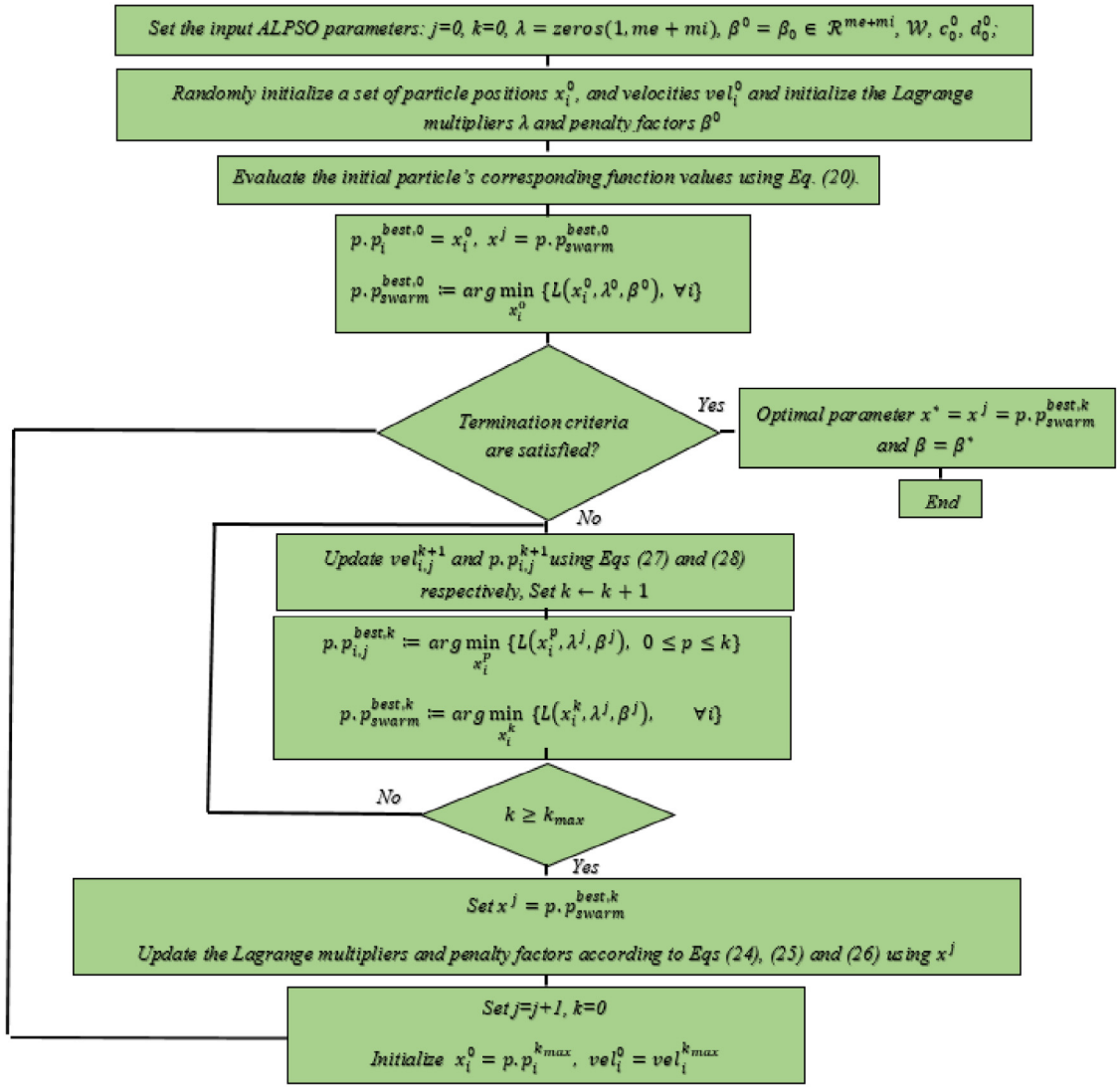

Fig. 4. The flowchart of the ALPSO.

Table 2
The particle representation.

	Particle type							
	Input feature mask						C	ξ
Representation	$x_{i,1}$	$x_{i,2}$	...	$x_{i,d}$	...	x_{i,n_f}	x_{i,n_f+1}	x_{i,n_f+2}
n_f denotes the number features that changes from different datasets.								

by the ALPSO-SVM system. Therefore, particles consisted of three parts, including input features mask, C and ξ, as shown in Table 2.

Step 2. Particle initialization and ALPSO parameters setting:

Initial particles composed of the feature mask, C, and ξ were generated. Furthermore, ALPSO parameters including the swarm size, the number of iterations, the number of particles, particle dimension, the velocity limitation, and inertia weight were set. Set $j = 0$, iteration ($k = 0$), Lagrange multipliers $\lambda^0 = 0$, penalty factor $\beta^0 = [1, 1, \ldots, 1]$, $vel_i^0 = 0$, $c_i^0 = 2$, and $d_i^0 = 2$.

Step 3. Evaluated the initiate particles' corresponding function values using Eq. (20).

Step 4. Checked the satisfying stopping criterion in which the number of fitness assessments. If the criterion were met, set $x^* = x^j = p.p_{swarm}^{best,k}$ and $\beta = \beta^*$; and do the training procedure from Steps 5–10.

Step 5. Set iteration $k \leftarrow k + 1$

Step 6. The training of SVM classifier

- The input features, according to feature mask, were selected to train and validate data sets.
- The calculation of the SVM classifier accuracy: for the training data set, a 10-fold cross validation was considered on the training data set, and the average of cross validation accuracy based on the (C, ξ) was calculated.
- The classification accuracy on the validation data-set was assessed based on trained SVM that was dependent on the (C, ξ) and the entire training data set.

Table 3
Dataset generation parameters for simulating the fault cases.

Parameter	Possible condition	Number of conditions
Grid fault type	abc, abc-g, ab, bc, ac, a-g, b-g, c-g, ab-g, bc-g, ac-g, HIF	12
Fault location	10%, 20%, 30%, …, 90% on fault line	9
Fault resistance	0.01, 2, 10, and 100 ohm	4
Operational mode	Islanded and non-islanded	2
Topology	Radial and loop	2
Fault line	Line 12, 23, 34, and 56	4
Fault inception angles	0^0, 45^0, 90^0 , and 180^0	4
L3 and L4 load	(90 kW, 45 kVAr) or (45 kW, 25 kVAr)	2
L5 load	(90 kW, −40 kVAr) or (45 kW, −20 kVAr)	2
L6 load	(90 kW, −20 kVAr) or (45 kW, −10 kVAr)	2
Total		221,184

Step 7. Fitness evaluation: the fitness function was evaluated for each particle according to the following formula:

$$Fit_i = \mathcal{W}_{acc} \times acc_i + \mathcal{W}_f \times \left[1 - \frac{\left(\sum_{j=1}^{n_f} f_i\right)}{n_f}\right] \tag{31}$$

As seen in Eq. (31), the fitness function had two predefined weights, called $\mathcal{W}_{acc}$ (classification accuracy weight) and $\mathcal{W}_f$ (number of selected feature weights). The f_i is the value of the feature mask and n_f denotes the entire number of features. Furthermore, the acc_i is the SVM classification accuracy which was evaluated as follows:

$$acc = \frac{cc}{cc + ic} \times 100\% \tag{32}$$

in which cc and ic are the numbers of examples that were classified correctly and incorrectly, respectively.

Note that accuracy in Eq. (30) was set as average cross-validation accuracy, as obtained in the previous step.

Step 8. Updated the velocities and positions according to Eqs. (27) and (28).

Step 9. According to fitness evaluation results, the global and individual best were updated. Moreover, their average training cross-validation accuracies and validation accuracies were recorded.

Step 10. Updated the Lagrange multipliers and penalty factors according to Eqs. (24), (25) and (26) used $x^j = p.p_{swarm}^{best.kmax}$, and set $j = j + 1$, and $k = 0$.

Step 11. Initialized $x_i^0 = p.p_i^{kmax}$, $vel_i^0 = vel_i^{kmax}$.

Step 12. Stopped condition checking: The global and individual best based on the fitness evaluation results were discovered. A check was performed on the maximum iteration limit. If the limit was reached, the best position emerged as the optimal parameter values. Otherwise, proceed with Step 5.

The overall procedure of the ALPSO-optimized SVM is demonstrated by the flowchart in Fig. 5.

5. Results and discussion

To assess the performance of the proposed ALPSO-SVM method for fault detection and classification of the MG, a series of simulations were performed. The efficiency of the ALPSO-SVM was compared to four well-known classifiers, i.e., decision tree (DT), k-nearest neighbor (K-NN), probabilistic neural network (PNN), and Naïve Bayes (NB). For a fair comparison, all the aforesaid classifiers were subjected to equal conditions, such as the input data set and the feature selection methods. In addition, the same training, validation, and testing samples were used for all classifiers. The impacts of noise on the fault classification performance were investigated. The measured metrics, namely, precision, recall, and F-measure were calculated to evaluate the effectiveness of the proposed method. Moreover, in order to evaluate the effectiveness of the MODWPT and ALSPO used in the proposed algorithm, they are replaced with famous signal processing methods and meta-heuristic techniques separately and their obtained results are compared. Finally, the obtained results from the proposed scheme were compared with the advanced fault detection and classification techniques in the recently published works.

All time-series simulations and numerical calculations were performed on a computer with an 8-core Intel Zeon E5-2630V3 processor with an Intel C612 chipset that can process data with a frequency of 2.4 to 3.2 GHz with 256 GB RAM. It had a maximum of 256 GB of supported RAM. Furthermore, this model computer supported an internal HDD hard drive with a SATA interface and a maximum capacity of 1TB.

5.1. Dataset generation

After simulating the CERTS MG system, a RSCAD script is written for generating fault and non-fault scenarios under various operating conditions including: changing the fault resistances, types of symmetrical and asymmetrical faults and HIF, different operating modes, different network topologies, changing fault current inception angles, fault lines, locations on the lines, and loads The operation of simulator is controlled by the script file. The script file analyzes the data without user interaction. As shown in Fig. 1, the faults were simulated at the designated location between the two bus. Moreover, the three phase current signals were measured at the relying point. The current signals were sampled at 64 samples/cycle of the grid frequency 60 Hz, and the sampling frequency was 3.84 kHz. Fault condition details are outlined in Table 3. A summary of conditions for no-fault cases is outlined in Table 4.

In total, 221,184 fault cases and 1088 no-fault cases were generated. The simulated current signals were used as measurement samples in each case through the protective relays at arbitrary times. The sampled signals were processed by ten wavelet members in the two wavelet functions, namely db and bior, as the mother wavelets of the MODWPT. The analyzed window was half cycle after the fault occurred. The sampled current signals were decomposed with 4 levels to ensure total features were extracted in different frequency bands. Table 5 summarizes the frequency bands of each level (Alves et al., 2016). The sampling frequency was 3.84 kHz, while the sampling number per cycle was 64 samples (considering $f = 60$ Hz).

In total, 16 MODWPT coefficient nodes, i.e., $s_4^0, s_4^1, s_4^2, \ldots, s_4^{15}$ were obtained for each current signal. The MODWPT coefficient of statistical features, i.e., the MODWPT coefficient for energies (i.e. $E_{S_4^0}, E_{S_4^1}, E_{S_4^2}, \ldots, E_{S_4^{15}}$), Shannon entropies ($SE_{S_4^0}, SE_{S_4^1}, SE_{S_4^2}, \ldots, SE_{S_4^{15}}$), means ($\mu_{S_4^0}, \mu_{S_4^1}, \mu_{S_4^2}, \ldots, \mu_{S_4^{15}}$), standard deviations ($\sigma_{S_4^0}, \sigma_{S_4^1}, \sigma_{S_4^2}, \ldots, \sigma_{S_4^{15}}$), minimums ($Min_{S_4^0}, Min_{S_4^1}, Min_{S_4^2},$

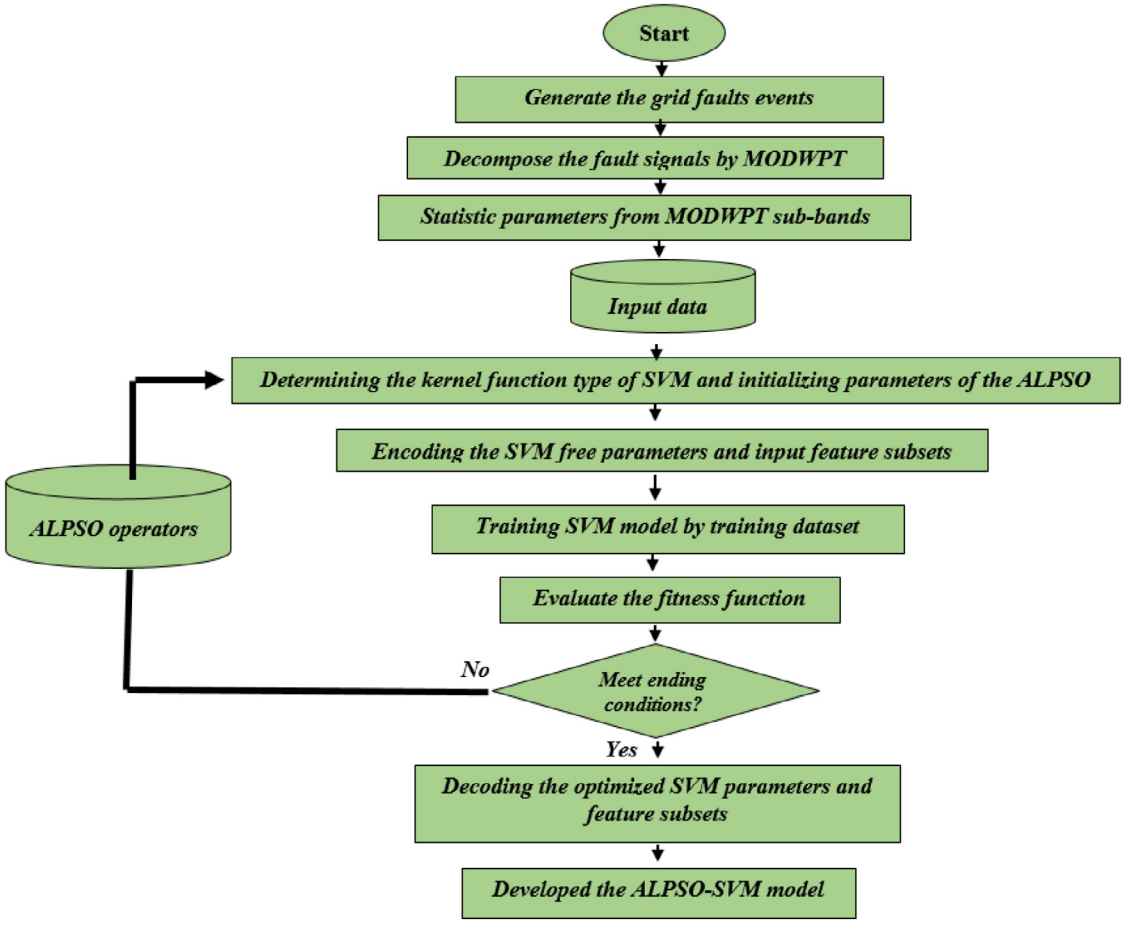

Fig. 5. Flowchart of the ALPSO-SVM method.

Table 4
Dataset generation parameters for simulating the no-fault cases.

Parameter	Possible condition	Number of conditions
Operational modes	Islanded and non-islanded	2
Topology	Radial and loop	2
L3 and L4 load	(90 kW, 45 kVAr) or (45 kW, 25 kVAr)	2
L5 load	(90 kW, −40 kVAr) or (45 kW, −20 kVAr)	2
L6 load	(90 kW, −20 kVAr) or (45 kW, −10 kVAr)	2
Event	Mode of operation change, topology change, L3, L4, L5, and L6 load change by ±5%, ±10%, ±15%, and ±20%	34
Total		1088

Table 5
Frequency bands of each level of MODWPT.

MODWPT level	Frequency bands (Hz)
1	0–1920; 1920–3840
2	0–960; 960–1920; 1920–2880; 2880–3840
3	0–480; 480–960; 960–1440; 1440–1920; 1920–2400; 2400–2880; 2880–3360; 3360–3840
4	0–240; 240–480; 480–720; 720–960; 960–1200; 1200–1440; 1440–1680; 1680–1920; 1920–2160; 2160–2400, 2400–2640; 2640–2880; 2880–3120; 3120–3360; 3360–3600; 3600–3840

$\ldots, Min_{S_4^{15}}$), maximums (i.e. $Max_{S_4^0}, Max_{S_4^1}, Max_{S_4^2}, \ldots, Max_{S_4^{15}}$), skewness's (i.e. $Skew_{S_4^0}, Skew_{S_4^1}, Skew_{S_4^2}, \ldots, Skew_{S_4^{15}}$), Kurtosis's (i.e. $Kurt_{S_4^0}, Kurt_{S_4^1}, Kurt_{S_4^2}, \ldots, Kurt_{S_4^{15}}$), and logarithmic energy entropies (i.e. $LEE_{S_4^0}, LEE_{S_4^1}, LEE_{S_4^2}, \ldots, LEE_{S_4^{15}}$) for the 3-phase current signals were calculated. The vector of each statistical feature contained 48 elements (16 values for MODWPT coefficient statistical features of the current signal multiplied by 3-phase signals = 48). Hence, the whole features vector contained 432 elements (48 multiplied by 9 statistical parameters). Therefore, the dimension of the feature matrix was 221,184 fault measurements multiplied by 432 elements (input data to the classifier).

5.2. ALPSO-SVM modeling and parameter optimization for fault detection and classification

To set up the ALPSO-SVM model, the radial basis function (RBF) was applied as the kernel function. The selection of RBF was due to its superior performance and fewer parameters to be tuned. Only two sensitive parameters, namely the penalty factor

Table 6
ALPSO based combination of selecting feature and optimizing parameter technique for SVM.

Iteration	Fitness value (%)	Fault and non-fault accuracy (%)	Fault type accuracy (%)	Number of selected features	Optimized C	Optimized ξ
1	95.5566	97.6982	96.5985	090	128.0	0.50000
2	95.9441	98.1061	97.0018	090	16.00	0.12500
3	95.1690	97.6982	96.5985	105	256.0	0.03125
4	95.9543	97.2903	96.1951	060	64.00	0.12500
5	95.7605	97.4942	96.3968	075	32.00	0.50000
6	95.3628	97.9021	96.8001	105	8.000	0.50000
7	96.1379	98.7180	97.6068	105	2.000	0.06250
8	95.7605	97.4942	96.3968	075	0.500	0.50000
9	95.9441	98.5140	97.4051	105	0.125	0.12500
10	95.3628	97.9021	96.8001	105	0.125	0.50000
11	96.1481	97.9021	96.8001	075	2.000	0.01563
12	95.9543	97.6982	96.5985	075	0.125	0.25000
13	95.7503	97.9021	96.8001	090	0.500	0.25000
14	***97.3107***	***99.5338***	***98.4135***	***090***	***16.00***	***0.03125***
15	96.3316	98.5140	97.4051	090	0.250	0.50000
16	96.1379	98.3101	97.2035	090	16.00	1.00000
17	95.7401	98.7180	97.6068	090	8.000	0.50000
18	95.9543	97.6982	96.5985	120	2.000	0.25000
19	96.1690	98.2903	97.1951	075	0.250	0.06250
20	96.1379	98.3101	97.2035	090	128.0	0.50000

(C) and the slack variable (ξ). had to be optimized. The search ranges for C and ξ were selected as $C\epsilon\left\{2^{-24}, 2^{-23}, \ldots, 2^{24}, 2^{25}\right\}$ and $\xi\epsilon\left\{2^{4}, 2^{3}, \ldots, 2^{-9}, 2^{-10}\right\}$, respectively. The details on parameter setting for the ALPSO are summarized as follows. The swarm size was equal to 27; the maximum iteration ($k_{max} = 20$); Lagrange multipliers $\lambda^0 = 0$; penalty factor $\beta^0 = [1, 1, \ldots, 1]$; $vel_i^0 = 0$; $c_i^0 = 1.1$; $d_i^0 = 1.1$; and $\epsilon_{g=}10^4$ in Eq. (28). Furthermore, $\mathcal{W}_{acc}$ was adjusted to 95% and $\mathcal{W}_f$ was set to 5% in Eq. (34). The SVM classifier was trained with 70% of the total dataset and then tested with the remaining 30%. The training data included the prerequisite to tune the parameters for the RBF kernel and the optimization of the SVM parameters. Note that the training and testing data were selected randomly from the feature dataset. To generalize the performance of SVM, 10-fold cross-validation was considered.

Table 6 denotes the iteration process of ALPSO to achieve the best solutions. These included the optimal parameters, the number of selected features, fitness values, fault, and non-fault cases accuracies, fault type classification accuracies using the SVM method. As observed, the best performance occurred at the 14th iteration. At this condition, the proposed method exhibited the highest fitness value (97.3107%), distinguished fault and non-fault accuracy (99.5338%), and yielded fault type classification accuracy (98.4135%). The recorded values of C and ξ were 16 and 0.03125, respectively.

To assess the efficacy of the proposed technique, accuracy and computational time evaluations were made. For the accuracy assessment, comparisons were made between three variations of the algorithm: (1) the baseline MODWPT-SVM algorithm (2) the ALPSO-SVM with the reduced selected feature but without optimized C and ξ and (3) ALPSO-SVM with the reduced selected feature and optimized C and ξ . For the baseline MODWPT-SVM, all generated datasets for the features were fed as inputs to the SVM classifier. The ALPSO was not involved in the process of reducing the features. Furthermore, the two critical parameters of SVM were obtained based on trial and error ($C = 2048$ and $\xi = 2.0$).

As seen in Table 7, the performance of the proposed ALPSO-SVM method with reduced feature selection and optimized C and ξ provided the highest fault and non-fault accuracies. In addition, it also exhibited the highest fault type detection. Its fault type classification accuracy was 98.4135%, and the fault cases identification accuracies from the non-fault types were 99.5338%. In terms of computational time, the baseline algorithm was the slowest. This was expected because the processing involved all feature sets. In contrast, the ALPSO-SVM method with reduced feature selection eliminated approximately 80% of the redundant features. Thus, it increased the computational speed by almost one order of magnitude. The significant reduction in the computation speed suggested that the method could be implemented in real-time.

The performance of the ALPSO-SVM method in identifying of the fault and non-fault cases and in assessing accuracy on fault type classification for each relay in the CERTS MG is shown Table 8. As seen, the proposed method effectively yielded the correct information in training and testing cases. More than 99% accuracy can be achieved to distinguish typical fault and non-fault cases in each relay. The average accuracy for training cases and testing cases were 99.7584% and 99.5338%, respectively. For the fault type classification, the performance of the proposed scheme was also acceptable. It provided considerable classification accuracy with an overall 98.6179% for training cases and 98.4135% for testing cases. From the obtained results, the selection of suitable features and sensitive parameters of SVM by ALPSO were enhanced regarding performance of the proposed fault detection method, and classification accuracy was rendered.

5.3. Performance of ALPSO with different classifiers

The performance of ALPSO with SVM classifier was compared with four well-known classifiers, i.e., decision tree (DT) (Ranjbar et al., 2020), k-nearest neighbor (K-NN) (Cepeda et al., 2020), probabilistic neural network (PNN) (Ahmadipour et al., 2018a), and Naïve Bayes (NB) (Mishra and Rout, 2017). For a fair comparison, all of the aforementioned classifiers had equal conditions, such as the input data set, the feature selection method, and the same training, validation, and testing samples. These classifiers had certain parameters that influenced the classification performances. Hence, the ALPSO was employed to find optimal values of classifiers' parameters. As seen, among all models of Table 9, ALPSO-SVM had the highest accuracy in classifying and detecting faults. To distinguish typical fault and non-fault cases, average accuracy in training cases and testing cases were 99.7584% and 99.5338%, respectively. The average fault type classification accuracies for training cases and testing cases were 98.6179% and 98.4135%, respectively.

5.4. Performance in a noisy environment

To assess the ALPSO-SVM fault detection method in terms of robustness, its performance was explored in a noisy environment. Based on the procedures outlined by previous researchers

Table 7
Fault detection accuracy with and without ALPSO optimal selection.

Algorithm	Number of selected features	Fault and non-fault accuracy (%)	Fault type accuracy (%)	Computational detection time (sec.)
Baseline MODWPT-SVM algorithm	Whole feature set	96.2703	95.1868	0.15336
ALPSO-SVM with reduced feature selection	90	99.1258	98.0101	0.02372
ALPSO-SVM with reduced feature selection and optimized SVM parameter	90	99.5338	98.4135	0.02372

Table 8
Fault detection accuracy on the CERTS microgrid system.

Relay	Fault and non-fault accuracy (%)		Fault type accuracy (%)	
	Training	Testing	Training	Testing
R-12	99.7229	99.7831	99.1898	98.6970
R-21	100.000	99.5325	100.000	98.2735
R-23	99.4824	99.4623	98.7663	99.1004
R-32	99.7630	99.9234	98.7361	97.1442
R-34	99.4323	99.2819	97.2861	98.1324
R-43	99.6828	99.3320	98.0707	98.5458
R-56	100.000	99.4223	98.4135	97.9206
R-65	99.9835	99.5325	98.4135	99.4936
Average	99.7584	99.5338	98.6179	98.4135

Table 9
Comparison of overall detection accuracy for different classifiers.

Model	Fault and non-fault accuracy (%)		Fault type accuracy (%)	
	Training	Testing	Training	Testing
ALPSO-DT	97.9039	97.1574	96.7846	96.4638
ALPSO-KNN	99.0316	98.2650	97.8994	97.1589
ALPSO-PNN	97.2093	96.4752	96.0979	95.3893
ALPSO-NB	98.9788	98.2131	97.8472	97.1076
ALPSO-SVM	99.7584	99.5338	98.6179	98.4135

Table 10
The performance of the proposed method under no-noise and noisy conditions.

Condition	Number of selected features	Accuracy		Computational detection time (S)
		Fault	Type	
No noise	90	99.5338	98.4135	0.02372
40 dB SNR	84	99.5137	98.4538	0.03630
30 dB SNR	84	99.4636	98.3630	0.05497
20 dB SNR	90	99.4235	98.3832	0.05816
10 dB SNR	90	99.032	97.8636	0.08058

(Casagrande et al., 2013b; James et al., 2017), the current measured data was distorted with white Gaussian noise. Hence, the performance of the proposed method was tested by adding consistently distributed Gaussian noise to the fault signals. A white Gaussian noise ratio of different values—10 dB single-to-noise ratio (SNR), 20 dB SNR, 30 dB SNR and 40 dB SNR—was uniformly applied to all fault signals (Casagrande et al., 2013b). Table 10 denotes the obtained results of classification accuracy with different levels of noise ratios. Despite the insignificant impact of noise on the offered technique's performance accuracy, this method's performance was acceptable. In the worst-case scenario (10 dB), the accuracy was decreased approximately 0.5% in comparison with the perfect measurements.

Moreover, the different types of colored noise were tested to validate the performance of the proposed method in terms of fault detection and classification. Colored noise refers to the noise with non-uniform distribution of power spectral density (PSD) in the frequency domain Xiangyu et al. (2021). Here, four kinds of common colored noises are tested, which are pink noise, blue noise, purple noise, and brown noise which are applied to all fault signals (Xiangyu et al., 2021). The setting of colored noises can be found in Xiangyu et al. (2021). Fig. 6 denotes the performance of the proposed method. Despite the insignificant impact of noise on the offered technique's performance accuracy, this method's performance was acceptable.

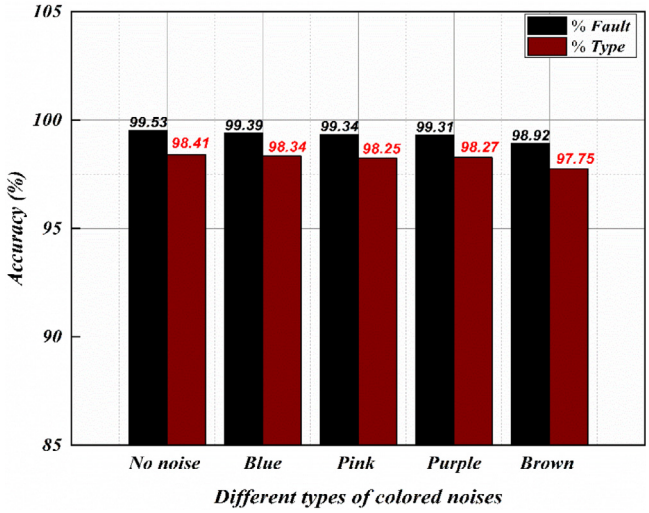

Fig. 6. The performance of the proposed method during no-noise and colored noise conditions. . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.5. The performance of the proposed method under high impedance faults (HIFs)

Detection of high impedance fault (HIF) in distributed power systems poses a serious challenge owing to low fault current amplitude. This is because the current amplitude is almost similar to the load current magnitude. Thus, detection using conventional overcurrent protection is very difficult. To evaluate the performance of the proposed MODWPT ALPSO-SVM method under this condition, the arc-associated HIF model is considered. The suggested model is shown in Fig. 7 which is based on the work by Michalik et al. (2006). Fig. 8 shows the measured voltage during a high impedance fault. The model includes a non-linear resistance that is connected to the feeder at the fault location. To test the effectiveness of the proposed scheme, 160 high-impedance fault cases are imposed. These comprise 85 and 75 scenarios in the grid-connected and islanded modes, respectively. Furthermore, the system is operated under different operating conditions as mentioned in Table 3. The performance of the proposed method for HIF detection during no-noise and noisy conditions is shown in Fig. 9. It is clear that the proposed method can detect the HIFs effectively under the grid-connected and islanded modes. Furthermore, it is found that the noisy conditions have little impact on the accuracy of the detection.

5.6. The performance of the proposed method with considering of several measures test

The classification efficiency may not denote the statistical significance of accurate classification of the SVM over the aforesaid

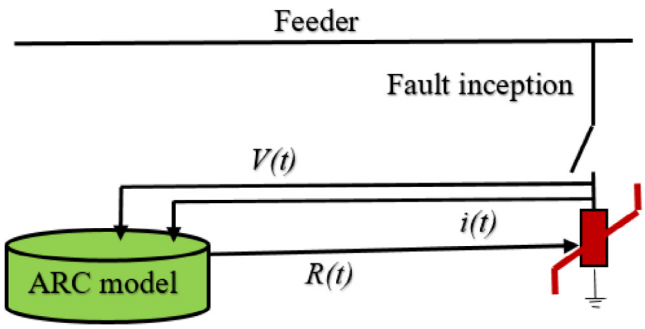

Fig. 7. The model of high impedance fault (HIF).

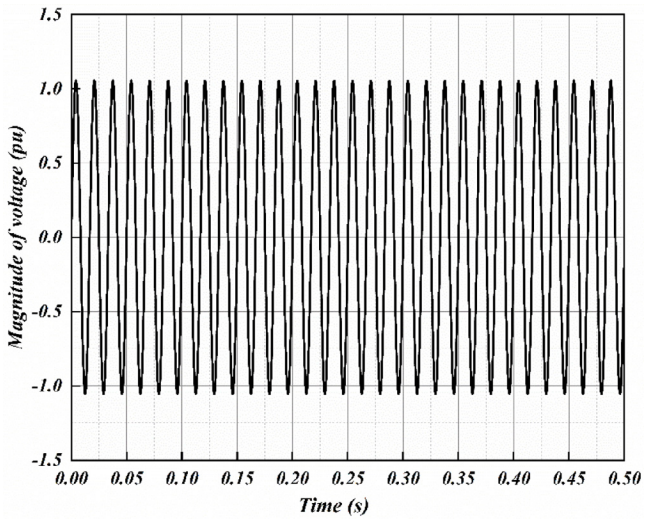

Fig. 8. Measured voltage during a HIF.

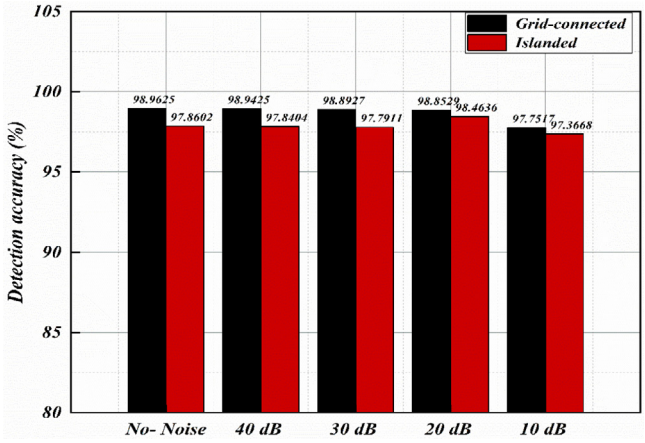

Fig. 9. The performance of the proposed method for HIF detection during no-noise and noisy conditions.

classifiers. Hence, measurements, such as precision, recall, and F-measure were calculated from the confusion matrix between the proposed method and other mentioned models in Table 9. The equations for these measured metrics are as follows:

$$precision = \frac{TP}{TP + FP} \times 100\% \tag{33}$$

$$recal = \frac{TP}{TP + TN} \times 100\% \tag{34}$$

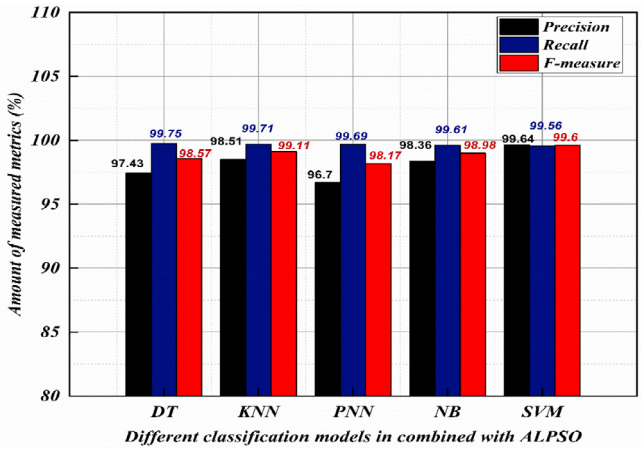

Fig. 10. Comparison of different classifiers combined with ALPSO in terms of precision, recall, and F-measure.

$$F - measure = 2 \times \frac{precision \times recall}{precision + recall} \times 100\% \tag{35}$$

in which *TP* denotes the whole number of fault cases correctly detected by the models, and *TN* is the whole number of other non-fault cases that are correctly predicted. *FP* and *FN* are the total numbers of fault cases incorrectly detected, as other non-fault cases and the total number of other non-fault cases wrongly predicted the fault events.

Fig. 10 shows the performance of the proposed method in comparison with other mentioned models in Table 9. As observed, the F-measure and precision of the proposed method were highest, and the recall measure was lowest in comparison with other models. It proved that the number of the fault and non-fault cases, that were predicted by the proposed algorithm, were more than other methods. Meaning, the performance of the proposed method is better than other models.

5.7. Validation through IEEE 34-bus microgrid system

To show that the proposed MODWPT-ALPSO-SVM is capable of real-time implementation, the modified IEEE-34 bus system (Faqhruldin et al., 2014) is utilized for the analysis fault detection and classification. The protective relays are installed on transmission lines 808–812, 816–824, 834–842, and 846–848. Using a similar technique as we conducted on our case study system, we develop 150,292 cases for training and testing the proposed scheme. The results corresponding to this additional system validation are shown in Table 11. It can be clearly analyzed from Table 11 that the proposed method performs well with precision 98.1124%, recall 99.3641%, and F-measure 98.7343% for fault detection in IEEE-34 bus test system. The computational time that has been recorded by the proposed scheme is 0.01123 s. Hence, it can be concluded that the proposed method can be generalized and applied to microgrid systems with large sizes in real-time.

5.8. Comparative analysis

5.8.1. Comparison of the performance of the MODWPT

To evaluate the effectiveness of the proposed feature selection, the MODWPT it is replaced by discrete wavelet transform (DWT) (Abdullah, 2017), Hilbert Huang (HH) (Azizi and Seker, 2021), fast discrete s-transform (FDST) (Mondal et al., 2020), and Variational Mode Decomposition (VMD) (Wang et al., 2019). For a fair comparison, all aforesaid methods are subjected to the same classifier (i.e., SVM), and the measurements are made for

Table 11
Performance of fault detection in IEEE-34 bus test system.

Models	Precision (%)	Recall (%)	F-measure (%)	Computational time (S)
The proposed method	98.1124	99.3641	98.7343	0.01123

Table 12
Overall fault detection accuracy for different signal processing techniques.

Model	Fault and non-fault accuracy (%)		Fault type accuracy (%)	
	Training	Testing	Training	Testing
DWT-ALPSO-SVM	97.5710	96.8270	96.4555	96.1358
FDST-ALPSO-SVM	98.6948	97.9309	97.5665	96.8285
HH-ALPSO-SVM	98.8229	98.0766	97.6931	96.9727
VMD-ALPSO-SVM	99.0678	98.3014	97.9352	97.1950
The proposed method	99.7584	99.5338	98.6179	98.4135

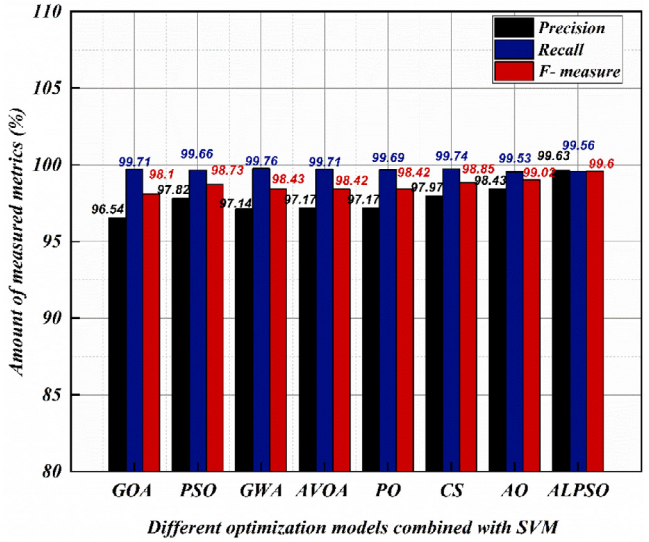

Fig. 11. Comparison of ALPSO and other meta-heuristic methods in terms of precision, recall and F-measure.

the three-phase current. As can be observed from the simulation results in Table 12, the proposed MODWPT exhibits the highest detection accuracy among all competing methods. These results proved the superiority of the MODWPT to other feature selection methods.

5.8.2. Comparison of the performance of ALPSO

The ALPSO method is incorporated to reduce the feature subsets and to select the sensitive parameters of the SVM (i.e., the penalty factor and the slack variable). To demonstrate the effectiveness of these functions, the ALPSO is compared with other meta-heuristic optimizers, namely the particle swarm optimization (PSO) (Kennedy and Eberhart, 1995), grasshopper optimization algorithm (GOA) (Zeng et al., 2021), gray wolf algorithm (GWA) (Chen et al., 2021), cuckoo search (CS)) (Mehedi et al., 2021), Political optimizer (Askari et al., 2020), Aquila optimizer (Abualigah et al., 2021), and African vultures' optimization algorithm (Abdollahzadeh et al., 2021). For a fair comparison, all competing optimizers have the same set of candidate input, preprocessing techniques, training and validation, and forecast samples. Only the optimizer type is changed (i.e., instead of ALPSO-SVM, the GA-SVM is used, etc.). The results are shown in *Fig.* 11. Based on the tests, the ALPSO exhibits the highest precision (99.6350%) and F-measure (99.6011%). Thus, it can be concluded that the ALPSO has superior fault detection capability compared to other methods".

5.8.3. Comparison of performance of the performance method with related fault detection techniques

The performance of the proposed ALPSO-SVM fault detection method was compared to existing state-of-the-art schemes. The evaluations were based on the overall accuracy for microgrid fault detections and classifications. As seen in Table 13, except for work in Mishra et al. (2015), most techniques exhibited satisfactory accuracies for classification and detection of faults. The overall accuracy of the ALPSO-SVM exceeded 99% for fault detection and classification. Furthermore, the proposed method can outperform advanced fault detection methods for MGs. However, results in Table 13 were not essentially assessed using a similar microgrid system; thus, the numbers were not absolute. The purpose of the exercise was only to provide an overview regarding the accuracy of various protection methods for MG.

6. Conclusion

In this paper, a large fault dataset was produced using the simulated CERT MG system in RSCAD/RTDS under several scenarios, such as different operating modes (i.e., grid-connected mode, and islanding mode), radial and loop topology, different fault types (i.e., symmetrical and asymmetrical faults), High impedance fault, different fault resistances, varied inception angles, different fault locations in four different power lines of the MG, pre-fault loading conditions, and measured noises (10, 20, 30, and 40 dB SNR). The three-phase current signals measurements sampled by the relaying points were input into the method. A MODWPT was applied to decompose the measurement data and the most effective statistical features that contained mean, standard deviation, energy, skewness, kurtosis, logarithmic energy entropy, max, min, and Shannon entropy were extracted from the results. Then, the mentioned facets were used as input data for the SVM classifier to yield the final relaying decision. The primary innovation of this work is in the proposed ALPSO-based method that aimed at optimizing the SVM classifier performance in terms of classification accuracies and computational times through detecting the best subset of available features and optimizing the sensitive parameters of SVM (i.e., C and ξ which are the error's penalty factor and the slack variable respectively). For most accurate fault detection and classification, the performance of the SVM classifier was compared with four well-known classifiers, i.e. decision tree (DT), k-nearest neighbor (K-NN), probabilistic neural network (PNN), and Naïve Bayes (NB). Furthermore, the impacts of noise measurements on fault detection and classification performance were assessed. Moreover, the performance of the proposed method is tested on a modified IEEE 34-bus system and their obtained results remain satisfactory. Finally, in order to evaluate the effectiveness of the MODWPT and ALSPO used in the proposed algorithm, they are replaced with famous

Table 13
Comparison with other state-of-the-art fault detection techniques.

Ref	Name of the technique	Overall accuracy (%)	
		Fault	Type
James et al. (2017)	Wavelet based deep neural networks	99.31	97.60
Mishra et al. (2015)	DWT+DT	97	85
Mishra et al. (2015)	DWT+RF	99.00	94.00
Mishra et al. (2015)	Overcurrent relay	56	–
Mishra et al. (2015)	Current differential relay	96	–
Abdelgayed et al. (2017b)	OWFMP+DT	More than 90.40	90.40
Abdelgayed et al. (2017b)	OWFMP+KNN	More than 95.63	95.63
Abdelgayed et al. (2017b)	OWFMP+SVM	More than 93.30	93.30
Abdelgayed et al. (2017b)	OWFMP+NB	More than 94.24	94.24
Mishra and Rout (2017)	HHT+NB	Less than 96.75	Less than 91.41
Mishra and Rout (2017)	HHT+SVM	Less than 96.15	Less than 91.91
Mishra and Rout (2017)	HHT+ELM	Less than 96.99	Less than 93.93
Mishra and Rout (2017)	Overcurrent relay	Less than 61.25	–
Mishra and Rout (2017)	Current differential relay	Less than 88.53	–
Kar et al. (2015)	S-transform+DT	Less than 99.475	–
Kar et al. (2015)	S-transform+SVM	Less than 99.26	–
Gashteroodkhani et al. (2020)	TT-DBN	99.80	99.32
Kar and Ranjan Samantaray (2015)	A fuzzy rule base	Less than 99.22	98.32
Abdelgayed et al. (2017a)	SSML+DT	More than 97.81	97.81
Abdelgayed et al. (2017a)	SSML+KNN	More than 96.70	96.70
The proposed method	MODWPT+ALPSO-SVM	99.53	98.41

signal processing methods and meta-heuristic techniques separately and their obtained results are compared. Furthermore, the proposed scheme was compared with state-of-the-art methods in the literature. The simulation results showed that the proposed method is accurate, fast, robust in noisy environments and able to detect and classify the fault events in microgrids; it superior to existing methods regarding accuracies and computation times.

Nomenclatures

MODWPT	Maximal Overlap Discrete Wavelet Packet Transform
RBF	Radial basis function
PNN	Probabilistic neural network
RF	Random forest
DWT	Discrete wavelet transform
OWFMP	Optimal wavelet functions matching pursuit
ELM	Extreme learning machine
HIF	High impedance fault
TT	Time-time transform
DBN	Deep belief network
SSML	Semisupervised machine learning
SVM	Support Vector Machine
MG	Microgrid
PO	Political optimizer
AO	Aquila optimizer
AVOA	African vultures' optimization algorithm
ALPSO	Augmented Lagrangian Particle Swarm Optimization
CERTS	Consortium for Electric Reliability Technology Solutions
RTDS	*Real time digital simulator*
EPDS	Electrical power distributed system
DER	Distributed energy resource
PV	Photovoltaic
NB	Naïve Bayes
DT	Decision tree
db	Daubechies
sym	Symlets
PCC	Point of common coupling
DWPT	Discrete wavelet packet transform
$g[n]$	Low pass filter
$h[n]$	High pass filter
$S_0(t)$	Scaling
$S_1(t)$	Wavelet function
$d_j^{2m}[n]$	Wavelet coefficients
l	Length of filters
f_s	Sampling frequency
Coif	Coiflet
bior	Biorthogonal
E	Energy
SE	Shannon entropy
μ	Mean
σ	Standard deviation
Min	Minimum
Max	Maximum
$Skew$	Skewness
$Kurt$	Kurtosis
LEE	Logarithmic energy entropy
N	Number of sampling point
S	Decomposed coefficients
x_i	Independent variables
y_i	Dependent variables
w	Weight vector
b	Bias term
C	Penalty factor
ξ_i	Slack variable
$\varphi(x)$	Non-linear vector function
λ_i	Lagrangian multiplier
$K(x, x_i)$	kernel function
$p.p$	Value of particle position
vel	Velocity
N_p	Number of particles
N_{vr}	Number of decision variables
X_{max}	Maximum value of each decision variables
X_{min}	Minimum value of each decision variables
$p.p_{i,j}^{best,k=0}$	Best position of particle
$p.p_{swarm}^{best,k=0}$	Best global position
$p.p_{i,j}^{best,k}$	Best position of particle at iteration k
$p.p_{swarm}^{best,k}$	Best global position at iteration k
K	Number of iteration
c_1	Cognitive values
c_2	Social knowledge values.
$\mathcal{W}$	Inertia factor.

$g(x)$	Nonlinear equality
$h(x)$	Inequality constrains
$\mathbb{D}$	Search space
$L(\ldots)$	Augmented Lagrange multiplier
λ_i	Lagrange multipliers
β_i	Penalty factors
λ^*	Correct Lagrange multiplier
ϵ_g	User-defined tolerance
ϵ_h	User-defined tolerance
$\mathcal{W}_{acc}$	Weight for SVM classification accuracy
$\mathcal{W}_f$	Weight for the number of selected features
f_i	Value of feature mask
n_f	Total number of features
acc_i	SVM classification accuracy
Fit_i	Fitness function

CRediT authorship contribution statement

Masoud Ahmadipour: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Visualization, Supervision, Formal analysis, Resources. **Muhammad Murtadha Othman:** Visualization, Investigation, Validation, Project administration, Supervision. **Rui Bo:** Writing – review & editing, Formal analysis, Validation, Software. **Zainal Salam:** Writing – review & editing, Formal analysis, Validation, Conceptualization. **Hussein Mohammed Ridha:** Software, Validation, Writing – review & editing. **Kamrul Hasan:** Software, Formal analysis, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the Long-Term Research Grant (LRGS), Ministry of Education Malaysia for the program titled "Decarbonization of Grid with an Optimal Controller and Energy Management for Energy Storage System in Microgrid Applications" with project code LRGS/1/2018/UNITEN/01/1/3; and also by the Research Management Centre (RMC), Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia with project code 100-RMC 5/3/SRP (019/2021). The authors would also like to acknowledge RMC, UiTM for the facilities provided to support on this research.

References

Abdelgayed, T.S., Morsi, W.G., Sidhu, T.S., 2017a. Fault detection and classification based on co-training of semisupervised machine learning. IEEE Trans. Ind. Electron. 65 (2), 1595–1605.

Abdelgayed, T.S., Morsi, W.G., Sidhu, T.S., 2017b. A new approach for fault classification in microgrids using optimal wavelet functions matching pursuit. IEEE Trans. Smart Grid 9 (5), 4838–4846.

Abdollahzadeh, B., Gharehchopogh, F.S., Mirjalili, S., 2021. African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems. Comput. Ind. Eng. 158, 107408.

Abdullah, A., 2017. Ultrafast transmission line fault detection using a DWT-based ANN. IEEE Trans. Ind. Appl. 54 (2), 1182–1193.

Abualigah, L., Yousri, D., Abd Elaziz, M., Ewees, A.A., Al-Qaness, M.A., Gandomi, A.H., 2021. Aquila optimizer: a novel meta-heuristic optimization algorithm. Comput. Ind. Eng. 157, 107250.

Ahmadipour, M., Hizam, H., 2019. A new islanding detection scheme based on combination of Slantlet transform and probabilistic neural network for grid-tied photovoltaic system. In: 2019 International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE). IEEE, pp. 1–6.

Ahmadipour, M., Hizam, H., Lutfi Othman, M., Amran Mohd Radzi, M., 2018a. An anti-islanding protection technique using a wavelet packet transform and a probabilistic neural network. Energies 11 (10), 2701.

Ahmadipour, M., Hizam, H., Othman, M.L., Mohd Radzi, M.A., Chireh, N., 2019a. A fast fault identification in a grid-connected photovoltaic system using wavelet multi-resolution singular spectrum entropy and support vector machine. Energies 12 (13), 2508.

Ahmadipour, M., Hizam, H., Othman, M.L., Radzi, M.A.M., Chireh, N., 2019b. A novel islanding detection technique using modified Slantlet transform in multi-distributed generation. Int. J. Electr. Power Energy Syst. 112, 460–475.

Ahmadipour, M., Hizam, H., Othman, M.L., Radzi, M.A.M., Murthy, A.S., 2018b. Islanding detection technique using Slantlet transform and ridgelet probabilistic neural network in grid-connected photovoltaic system. Appl. Energy 231, 645–659.

Alves, D.K., Costa, F.B., de Araujo Ribeiro, R.L., de Sousa Neto, C.M., Rocha, T.D.O.A., 2016. Real-time power measurement using the maximal overlap discrete wavelet-packet transform. IEEE Trans. Ind. Electron. 64 (4), 3177–3187.

Askari, Q., Younas, I., Saeed, M., 2020. Political optimizer: A novel socio-inspired meta-heuristic for global optimization. Knowl.-Based Syst. 195, 05709.

Azizi, R., Seker, S., 2021. Microgrid fault detection and classification based on the boosting ensemble method with the Hilbert-Huang transform. IEEE Trans. Power Deliv..

Bagheri, S., Moravej, Z., Gharehpetian, G.B., 2017. Classification and discrimination among winding mechanical defects, internal and external electrical faults, and inrush current of transformer. IEEE Trans. Ind. Inf. 14 (2), 484–493.

Bakar, A.H.A., Ooi, B., Govindasamy, P., Tan, C., Illias, H.A., Mokhlis, H., 2014. Directional overcurrent and earth-fault protections for a biomass microgrid system in Malaysia. Int. J. Electr. Power Energy Syst. 55, 581–591.

Baloch, S., Muhammad, M.S., 2021. An intelligent data mining-based fault detection and classification strategy for microgrid. IEEE Access 9, 22470–22479.

Baloch, S., Samsani, S.S., Muhammad, M.S., 2021. Fault protection in microgrid using wavelet multiresolution analysis and data mining. IEEE Access.

Casagrande, E., Woon, W.L., Zeineldin, H.H., Kan'an, N.H., 2013a. Data mining approach to fault detection for isolated inverter-based microgrids. IET Gener. Transm. Distrib. 7 (7), 745–754.

Casagrande, E., Woon, W.L., Zeineldin, H.H., Svetinovic, D., 2013b. A differential sequence component protection scheme for microgrids with inverter-based distributed generators. IEEE Trans. Smart Grid 5 (1), 29–37.

Cepeda, C., Orozco-Henao, C., Percybrooks, W., Pulgarín-Rivera, J.D., Montoya, O.D., Gil-González, W., Vélez, J.C., 2020. Intelligent fault detection system for microgrids. Energies 13 (5), 1223.

Chakravorti, T., Priyadarshini, L., Dash, P.K., Sahu, B.N., 2019. Islanding and non-islanding disturbance detection in microgrid using optimized modes decomposition based robust random vector functional link network. Eng. Appl. Artif. Intell. 85, 122–136.

Chandra, A., Singh, G.K., Pant, V., 2021. Protection of AC microgrid integrated with renewable energy sources–A research review and future trends. Electr. Power Syst. Res. 193, 107036.

Chen, X., Yi, Z., Zhou, Y., Guo, P., Farkoush, S.G., Niroumandi, H., 2021. Artificial neural network modeling and optimization of the solid oxide fuel cell parameters using grey wolf optimizer. Energy Rep. 7, 3449–3459.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.

Costa, F.B., 2014. Fault induced transient detection based on real time analysis of the wavelet coefficient energy. IEEE Trans. Power Deliv. 29 (1), 140–153.

Dehghani, M., Khooban, M.H., Niknam, T., 2016. Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations. Int. J. Electr. Power Energy Syst. 78, 455–462.

Eberhard, P., Sedlaczek, K., 2009. Using augmented lagrangian particle swarm optimization for constrained problems in engineering. In: Advanced Design of Mechanical Systems: From Analysis to Optimization. Springer, Vienna, pp. 253–271.

Ezzat, A., Elnaghi, B.E., Abdelsalam, A.A., 2021. Microgrids islanding detection using Fourier transform and machine learning algorithm. Electr. Power Syst. Res. 196, 107224.

Faqhruldin, O.N., El-Saadany, E.F., Zeineldin, H.H., 2014. A universal islanding detection technique for distributed generation using pattern recognition. IEEE Trans. Smart Grid 5 (4), 1985–1992.

Gabbar, H., 2016. Smart Energy Grid Engineering. Academic Press, http://dx.doi.org/10.1016/B978-0-12-805343-0.00001-2.

Gashteroodkhani, O.A., Majidi, M., Etezadi-Amoli, M., 2020. A combined deep belief network and time-time transform based intelligent protection Scheme for microgrids. Electr. Power Syst. Res. 182, 106239.

Guillén, D., Esponda, H., Vázquez, E., Idárraga-Ospina, G., 2016. Algorithm for transformer differential protection based on wavelet correlation modes. IET Gener. Transm. Distrib. 10 (12), 2871–2879.

He, Q., 2013. Vibration signal classification by wavelet packet energy flow manifold learning. J. Sound Vib. 332 (7), 1881–1894.

Hirsch, A., Parag, Y., Guerrero, J., 2018. Microgrids: A review of technologies, key drivers, and outstanding issues. Renew. Sustain. Energy Rev. 90, 402–411.

Hong, Y.Y., Cabatac, M.T.A.M., 2019. Fault detection, classification, and location by static switch in microgrids using wavelet transform and taguchi-based artificial neural network. IEEE Syst. J. 14 (2), 2725–2735.
Hooshyar, A., El-Saadany, E.F., Sanaye-Pasand, M., 2015. Fault type classification in microgrids including photovoltaic DGs. IEEE Trans. Smart Grid 7 (5), 2218–2229.
James, J.Q., Hou, Y., Lam, A.Y., Li, V.O., 2017. Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks. IEEE Trans. Smart Grid 10 (2), 1694–1703.
Jiang, J.A., Liu, C.W., Chen, C.S., 2002. A novel adaptive PMU-based transmission-line relay-design and EMTP simulation results. IEEE Trans. Power Deliv. 17 (4), 930–937.
Kar, S., Ranjan Samantaray, S., 2015. A fuzzy rule base approach for intelligent protection of microgrids. Electr. Power Compon. Syst. 43 (18), 2082–2093.
Kar, S., Samantaray, S.R., Zadeh, M.D., 2015. Data-mining model based intelligent differential microgrid protection scheme. IEEE Syst. J. 11 (2), 1161–1169.
Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of ICNN'95-International Conference on Neural Networks, Vol. 4. IEEE, pp. 1942–1948.
Lasseter, R.H., Eto, J.H., Schenkman, B., Stevens, J., Vollkommer, H., Klapp …, D., Roy, J., 2010. CERTS microgrid laboratory test bed. IEEE Trans. Power Deliv. 26 (1), 325–332.
Lopez-Garcia, T.B., Coronado-Mendoza, A., Domínguez-Navarro, J.A., 2020. Artificial neural networks in microgrids: A review. Eng. Appl. Artif. Intell. 95, 103894.
Mehedi, I.M., Ahmadipour, M., Salam, Z., Ridha, H.M., Bassi, H., Rawa, M.J.H., … Abdullah, M.P., 2021. Optimal feature selection using modified cuckoo search for classification of power quality disturbances. Appl. Soft Comput. 113, 107897.
Michalik, M., Rebizant, W., Lukowicz, M., Lee, S.J., Kang, S.H., 2006. High-impedance fault detection in distribution networks with use of wavelet-based algorithm. IEEE Trans. Power Deliv. 21 (4), 1793–1802.
Mishra, M., Rout, P.K., 2017. Detection and classification of micro-grid faults based on HHT and machine learning techniques. IET Gener. Transm. Distrib. 12 (2), 388–397.
Mishra, D.P., Samantaray, S.R., Joos, G., 2015. A combined wavelet and data-mining based intelligent protection scheme for microgrid. IEEE Trans. Smart Grid 7 (5), 2295–2304.
Misiti, M., Misiti, Y., Oppenheim, G., Poggi, J.-M., 1997-2015. Wavelet Toolbox—User's Guide. The Math Works Inc., Natick, MA, USA.
Mondal, A., Das, S., Patel, B., 2020. Fault detection during power swing using fast discrete s-transform. In: Computational Advancement in Communication Circuits and Systems. Springer, Singapore, pp. 73–79.
Motlagh, S.Z., Foroud, A.A., 2021. Power quality disturbances recognition using adaptive chirp mode pursuit and grasshopper optimized support vector machines. Measurement 168, 108461.
Patnaik, B., Mishra, M., Bansal, R.C., Jena, R.K., 2021. MODWT-XGBoost based smart energy solution for fault detection and classification in a smart microgrid. Appl. Energy 285, 116457.
Ranjbar, S., Farsa, A.R., Jamali, S., 2020. Voltage-based protection of microgrids using decision tree algorithms. Int. Trans. Electr. Energy Syst. 30 (4), e12274.
Shafiullah, M., Abido, M.A., 2018. S-transform based FFNN approach for distribution grids fault detection and classification. IEEE Access 6, 8080–8088.
Singh, P., Singh, N., Choudhary, N.K., 2021. Fault detection and classification in microgrid using wavelet transform and artificial neural network. In: Advances in VLSI, Communication, and Signal Processing. Springer, Singapore, pp. 17–31.
Wang, X., Gao, J., Wei, X., Song, G., Wu, L., Liu, J., … Kheshti, M., 2019. High impedance fault detection method based on variational mode decomposition and Teager–Kaiser energy operators for distribution network. IEEE Trans. Smart Grid 10 (6), 6041–6054.
Xiangyu, L.U., Xudong, H.E., Huaihai, C.H.E.N., Zheng, R., 2021. Operational modal parameter identification with colored noise excitation. Chin. J. Aeronaut. 34 (2), 288–300.
Yazdani, A., Dash, P.P., 2009. A control methodology and characterization of dynamics for a photovoltaic (PV) system interfaced with a distribution network. IEEE Trans. Power Deliv. 24 (3), 1538–1551.
Yazdani, A., Iravani, R., 2010. Voltage-Sourced Converters in Power Systems, Vol. 39. John Wiley & Sons, Hoboken, NJ, USA.
Zeng, X., Hammid, A.T., Kumar, N.M., Subramaniam, U., Almakhles, D.J., 2021. A grasshopper optimization algorithm for optimal short-term hydrothermal scheduling. Energy Rep. 7, 314–323.
Zhao, X., Ye, B., 2010. Convolution wavelet packet transform and its applications to signal processing. Digit. Signal Process. 20 (5), 1352–1364.