Electric Power Systems Research 216 (2023) 109025

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Machine learning based adaptive fault diagnosis considering hosting capacity amendment in active distribution network

Sourav Kumar Sahu [a], Millend Roy [b], Soham Dutta [c], Debomita Ghosh [a,*], Dusmanta Kumar Mohanta [a]

[a] EEE Department, Birla Institute of Technology, Ranchi, India
[b] Microsoft Research, Bengaluru, India
[c] Department of Electrical and Electronics engineering, Manipal Institute of Technology, Manipal academemy of Higher Education, Manipal, Karnataka, 576104, India

ARTICLE INFO

Keywords:
Machine learning (ML)
Fault diagnosis
Hosting capacity (HC)
Spectral kurtosis (SKS)
Histogram-based gradient boost (HGB)
Situational awareness (SA)

ABSTRACT

Augmentation of distributed energy resources (DERs) safely in distribution system termed as hosting capacity (HC) is one of the prominent needs to achieve energy sufficiency with minimum emission. However, any amendment in HC over premeditated injection sets up challenges in perspective of situational awareness (SA) of networks for precise decision-making related to fault prediction and location. In this work, authors propose histogram-based gradient boost (HGB) algorithm, an accurate machine learning (ML) technique for fault type detection and location. Due to the unique characteristic of noise cancelation, spectral-kurtosis is utilized for extraction of features of the faulted transient signals. For improved competence of the process, optimized feature importance values are considered. In order to study the efficacy of the proposed method, HC of the network is altered, leading to up-gradation of network parameters. These upgraded parameters are used for retraining the proposed ML algorithm for desired SA, with perception, comprehension, projection, and accurate decision making. The authors also considered other ML techniques to showcase a comparative study with the HGB. The entire analysis is tested on reconfigured IEEE-33 bus distribution system developed in Typhoon HIL real-time simulator. The proposed methodology is also meticulously compared with existing literature to establish its excellence.

1. Introduction

The advancement in technology contributed to the sprawl of the power system, enriching it towards a smart grid. The power distribution system being an integral part, is no exception to this development. With the paradigm shift of distribution system from passive to active, power flow dynamics changed to a great extent. This has led to incredible developments in the form of integration of intermittent source-based distributed generators. However, augmentation of DERs to enhance the sustenance of the distribution system termed as HC, without violating the network parameters is a key concern [1]. The maximum possible HC although is calculated in the planning stage, but exponential load growth results in a change of network parameters, i.e., voltage, current and THD, leading to alteration of power flow [2]. Simultaneously, estimation of upcoming renewable integration in the distribution grid is also uncertain, as dependent on the political and legal constraints of the concerned region and the economic benefits provided by the proposed states. This variable power integration into the distribution system may create several issues such as unwanted operation or failure in operation of the conventional protection system. This is due to the fact that integration of DER changes the line current in the network. With such variations in the active distribution system, the traditional protection scheme maloperates and may not provide desired SA [3]. In this context, the unavoidable dynamics of the distribution system need to be seen with utmost care by all the SA-based components of the power system, such as perception, comprehension, projection, and decision making [4]. Care should be taken with surgical precision for smooth and reliable operation of the network. Although various upgradation in protection philosophy such as providing additional time step in protection co-ordination, change in current setting and additional equipments can be added to counterbalance the degradation in fault detection due to addition of DERs, but correct fault detection could not be guaranteed at all the time [5,6].

* Corresponding author.
E-mail address: debomita.ghosh@bitmesra.ac.in (D. Ghosh).

https://doi.org/10.1016/j.epsr.2022.109025
Received 30 May 2022; Received in revised form 2 October 2022; Accepted 21 November 2022
Available online 1 December 2022

Nomenclature

DER	Distributed energy resources
HC	Hosting Capacity
SA	Situational awareness
HGB	Histogram-based gradient boost
ML	Machine learning
SKS	Spectral kurtosis
ADN	Active distribution network
PVDG	Photovoltaic distributed generation
SPV	Solar photovoltaic
STFT	Short time Fourier transform
DSO	Distribution system operator
PCC	Point of common coupling
CNS	Conditionally nonstationary
HMI	Human machine interface
DT	Decision Tree
NGB	Normal gradient boosting
XB	Xg Boost
RF	Random forest
CB	CatBoost
AB	AdaBoost
LGBM	Light GBM
V_{UG}	Voltage utility grid
R_{HC}	Equivalent resistance from considered bus
X_{HC}	Equivalent reactance from considered bus
P_{DER}	Active power injection by PVDG
Q_{DER}	Reactive power injection by PVDG
P_C	Active power demand at considered bus
Q_C	Reactive power demand at considered bus
V_C	Voltage at considered bus
$V_{C\ Max}$	Maximum permissible voltage at considered bus
C1	Constant amplitude
f	Frequency
δ	Constant initial phase
S	Training set
X	New observation
y	Category that the new observation belongs
Z	Set of class labels
F_0	Initial base model function
$\widehat{y}$	Predicted value
c	Weak learners
L	Categorical cross-entropy loss/ log loss

1.1. Motivation

To avoid various detrimental environmental issues and to deal with the limited amount of fossil fuel, renewable energy based DERs is one of the viable solutions. As the sustained growth is expected in DER integration into grid for upcoming decade, the importance of protection system to provide a reliable power is unassailable. With integration of DERs the existing protection system is susceptible to maloperation. Although various adjustments and inclusion of more devices can be made to improve the protection performance but these solutions are not sustainable. It may so happen that the integration of DERs can be happen abiding the HC of node, but the traditional protection system fails due to lack of ability to adopt the change or due to misoperation, which is not rare. It is impractical to minutely adjust the protection setting with each power injection upgradation of prosumer incase of very precise setting margin. Taking all the issues into consideration a sustainable, low-cost method which can work with the present as well as future network environment is essential.

1.2. Literature review

To improve the fault detection capabilities ML based approach seems to be a great solution, as depicted in [7–10]. The authors in [7], utilized a support vector machine to classifying the detected fault through the features extracted from the FFT analysis. High resolution synchrophasor data is vital for this proposed method to perform effective fault identification [7]. But, the exact faulted phase detection such as AN, AB, ABN etc., which may be crucial for some applications is not presented. In [8] the authors used a combination of continuous wavelet transform and convolutional neural network to identify the line-to-ground fault in the distribution network. However, the insight for measurement of lower sampling rates, which may cause difficulty in implementing this strategy with low-cost devices is not assessed. To detect and differentiate grid fault and anti-islanding in the presence of solar photovoltaic (SPV) in the distribution grid authors in [9,10] utilized SVM method, which is also effective in the presence of plugin electric vehicle. In [9,10] the author demonstrated this technique considering real example of a single grid connected SPV system. Although this fault detection technique providing promising result for single SPV system to the utility, the performance of the system needs to be evaluated in the presence of multiple DERs at different locations and with distributed network parameters. Even though ML is applied to enhance the capabilities to accurately detect various faults in power system network, effect of HC on protection is not discussed in the available literature thoroughly. The colossal growth in SPV integration into the grid across all the major counties, HC amendment imperative to all the essential entities of the utility such as its prosumers, consumers and the utility itself [11]. Researchers applied various techniques for amendment of HC in the distribution system as in [12–14]. The authors in [12] showcased a real-time simulation based iterative method to improve the HC of the distribution network. The authors considered various power system parameters such as voltage, current and THD as indicating parameter for HC analysis. Protection aspect is not considered in this work, while upgrading the HC of the network. In [13], the author assessed the HC of the active distribution network (ADN) in the presence of signal distortion in the distribution network. The authors also proposed a technique to design harmonic mitigating device to further improve the HC of the network. To design the filter, authors ensured the voltages and the currents should not violate the predefined limits in the network but, there is no insight about the protection due to addition of filters in the network. Not only with additional equipments, the authors in [14] proposed a reconfiguration based HC improvement solution by changing topology of the distribution network with the help of tie line breakers but by doing so the protection system may mal-operate and can lead to catastrophic event. Implementation of advanced techniques to improve the HC may cause various protection issues, as the line current of the upstream as well as the downstream can be contributed by the connected DERs. To tackle this protection concern, various protective measures can be considered such as re-estimated time grading of the protection settings, addition of additional protecting equipment, directional protection and adaptive protection [15]. Adoption of latest adaptive settings may also cause false tripping depending on setting margin in case of improper hosting capacity estimation and if the setting margin is very less the false tripping is always a concern irrespective of hosting capacity calculation [15]. To invalidate such protection issues, authors in [16–23] adopted some of the effective fault detection techniques.

Commercial protection has the disadvantage of variable accuracy with a change in line length. To counter this problem, in [16], the authors proposed ratio-based indices calculation for different types of fault

identification. Although detection is possible for LG and LL faults through the proposed technique, but for LLG faults, the accuracy is found to be sharply decreasing. For LLL fault identification, the scheme doesn't provide any insight. With the decrease in sampling rate below 40 kHz, the described method failed to work at its full potential. Additionally, a precise design of the filter is a must to implement this fault detection scheme. But the inclusion of auxiliary devices for fault analysis in the present day is not an effective techno-economical solution. Authors in [17]demonstrated an impedance-based fault analysis method for the distribution system. Although the method proved to be promising, but it lacks in performance when the photovoltaic distributed generation (PVDG) penetration is high. In addition to it, the performance is found to degrade with change in the network topology. The topological issue incurred in [17] is resolved in [18], where the authors discussed on high impedance fault in the distribution system. Although the topological concern is addressed, but this method did not give any insight into the integration of DERs and the requirement of high sampling rate. In [19], fault location is estimated in the distribution system with the highest accuracy, whereas no insight is given for the type of fault. The authors used capacitive current, considering the distribution system to have only underground cables, which is not feasible for all locations. Realising all these concerns, data-driven fault analysis schemes are presented in [20–23]. In [20] authors demonstrated a differential protection-based data mining procedure for fault identification. Although the proposed scheme provides efficient results, but it is highly unlikely to use differential protection in the distribution system except for the transformer. Secondly, for differential feature extraction, communication is required, which adds to the overall installation and maintenance cost. In [21], the authors utilized a wavelet-based data mining approach, where ten types of fault conditions are created for the decision tree classifier, to detect the fault or no-fault condition. The simulated ten types of faults are broadly classified into four types, namely LG, LL, LLG, and LLLG, respectively by the classifier, with additional input from the sequence analyzer. In the distribution system, it is essential to detect the faulted phase for further action required by the distribution system operator (DSO). In this context authors in [21] did not provide precise information regarding the faulted phase. It is eminent that due to the use of a sequence analyser, some time-delay may also occur for fault detection, which may not be appropriate for instantaneous applications. Authors in [22] proposed a wavelet optimization-based fault classification algorithm, which showcased efficient results. However, the authors did not give any insight into the location of the fault. Identification of faulted line section is showcased in [23], with the use of three ML classifiers. The authors discretized the features for the classifier to minimize the input. A detailed analysis due to variable sampling frequency on computational time as well as on accuracy is presented in this paper.

To address any problem with ML, specific parameters related to an event need to be given as input to the ML model for learning, termed as the feature of the event. Spectral-kurtosis (SKS) is one of the promising tools for feature extraction, and these extracted features can be utilized as an input for the ML model to obtain an efficient result. Initially, SKS was limited to identifying the transients present in a signal along with their location in the frequency domain [24]. SKS is also used as an alternative to power spectral density, with the advantage of having a noise resilient property, i.e., SKS was able to detect transients, even in noisy signals or in case of various transients such as capacitor switching, load switching [25]. In [26], the concept of SKS is re-established on the basis of normalized fourth-order moment of the co-efficient for short time Fourier transform (STFT). This development enabled the multi-dimensional application of SKS that also includes power system protection.

To improve the SA of active distribution network, the features should be calculated with the timely up-gradation of renewable integration in the network [2,27]. This change in HC of the network contributes to a change in the amount of power flow in the network, which affects the extracted features from signals of the network. In [28], the authors explained the effect of change in the power flow and the need for timely up-gradation of various power system parameters. The authors also clearly explained the need for timely up-gradation of protection coordination parameters for power system networks. Although in the available literature, analysis is done for improvement of HC using various methods, but analysis on the effect of a wide range of fault in the presence of DERs are not extensively analyzed. Additionally, improvement of SA with the variation in the integration of additional DERs in real-time, with consideration of both HC and ML is one of the major areas of concern, which needs to be focused on.

1.3. Contribution

Based on reviewed literature in subsection B, context, the major contributions in context to state of the art are as follows:

i Histogram-based gradient boosting (HGB), a machine learning technique, is employed for fault classification as well as fault location identification. The classifier is trained with many scenarios that include wide variation in fault resistance, fault types, and locations. This model is tested with data that is generated in the real-time simulator with variation in above said parameters; thus, the robustness of the classifier is ensured.
ii Spectral-Kurtosis (SKS) based feature extraction is considered in this study, which has high precision in analyzing the wide range of fault conditions.
iii The algorithm is tested on real-time data with a wide variation of sampling rates for low-cost, user-friendly applications in the industry.
iv The developed algorithm is tested to have high accuracy and less classification time. It has the potential to accommodate upcoming dynamic changes in the network parameters, owing to the alteration in HC. The algorithm can also be updated to broadcast the fault classification information to the nearby DSOs based on which network SA can be enhanced.
v The presented work is compared with other ML techniques such as K-Nearest neighbour, Logistic regression, Gaussian RBF Kernel SVC, Gaussian naïve Bayers, Voting classifier, Gaussian RBF Kernel SVC, auto encoder, Decision Tree (DT), NGB, Xg Boost (XB),Random forest (RF), CatBoost (CB), HGB, AdaBoost (AB), and Light GBM (LGBM) to check its suitability over other techniques. The work is also compared with the available reported recent literature to justify its potency.
vi The algorithm is experimented to examine its potential to discriminate faults from other power system transients scenarios.

The rest of the paper is organized as follows. Section 2 briefs the overview of HC assessment. Section 3 details the modeling of active distribution network on Typhoon HIL real-time environment. Spectral-Kurtosis-based feature extraction for a variety of fault signals, extracted from the modeled reconfigured IEEE 33 bus distribution system, is resulted in Section 4. Section 5 depicts the significance of HGB technique for accurate fault diagnosis. Proposed methodology for SA based adaptive fault diagnosis with amendment of HC is depicted in Section 6. Section 7 deals with the result and discussion based on the extracted features required by the classifier for training. Also, its performance is compared with other methods and with existing literatures. Lastly, the conclusion is drawn in Section 8.

2. Brief review on hosting capacity assessment and its impact on fault

The possibility of maximum allowable DER integration i.e., HC of the network is subjected to various power system parametric fluctuations such as voltage, current, and THD. These subjectives are also due to the

uncertainty in DERs like wind speed and solar irradiation. Additionally, economic constraints can also be considered as limiting factors for HC. However, the major restrictive indicators considered for HC enhancement are voltage, current and THD. Hence, for network modeling and HC assessment IEEE and ANSI standards are taken into consideration. The voltage limit for HC is set according to IEEE Std 1250, which is ± 5 % and ampacity for HC is set according to ANSI C84.1-1989.THD, one of the significant indicators for HC assessment, is set according to IEEE-519-2014, which is 5% for the DER integrated nodes in the distribution side. For calculation of the power to be injected at the point of common coupling (PCC) as in Eq. (1), Kirchhoff's voltage law (KVL) is applied on the effective equivalent network of the active distribution system as in Fig. 1.

Considering the real part of the Eq. (1), and assuming Ω≈0 (as in distribution system $X/R \ll 1$), Eq. (2) can be deduced. From Eq. (3), it is evident that the maximum DER integration depends on parameters such as allowable voltage limit, line parameters, and load at PCC. It is evident from [1] that the maximum allowable voltage level is achieved prior to the thermal limit; hence taking care of the maximum voltage limit, seldom takes care of the current limit. In Eq. (3), if the reference voltage is considered as 1 p.u., and the maximum allowable voltage is taken as 1.05 p.u., then for a specific load, the maximum power injection obtained is found to be inversely proportional to the effective impedance at PCC. So, it can be deduced that with the decrease in the value of the effective impedance, the HC of the network increases and vice-versa [29]. Hence, at different nodes of the network, the HC obtained is different, as represented in Eq. (3).

$$V_C = V_{UG} + \frac{(P_{DER} - P_C) - j(Q_{DER} - Q_C)}{V_C \ < -\ \Omega} \times (R_{HC} + jX_{HC})\mathrm{V_C}$$
$$= \mathrm{V_{UG}} + \frac{(\mathrm{P_{DER}} - \mathrm{P_C}) - \mathrm{j}(\mathrm{Q_{DER}} - \mathrm{Q_C})}{\mathrm{V_C} \ < -\ \Omega} \times (\mathrm{R_{HC}} + \mathrm{jX_{HC}}) \quad (1)$$

$$\Rightarrow V_C = V_{UG} + \frac{(P_{DER} - P_C) - j(Q_{DER} - Q_C)}{V_C} \times \{(- X_{HC} Sin\ \Omega) + j\,(R_{HC} Sin\ \Omega) + (R_{HC} Cos\ \Omega) + j(X_{HC} Cos\ \Omega)\}$$

$$\Rightarrow V_C = V_{UG} + \frac{(P_{DER} - P_C)}{V_C} \times \{(R_{HC}\mathrm{Cos}\ \Omega - X_{HC}\ \mathrm{Sin}\ \Omega) + j(R_{HC}\mathrm{Sin}\ \Omega + X_{HC}\ \mathrm{Cos}\ \Omega)\} - j\frac{(Q_{DER} - Q_C)}{V_C} \times \{(R_{HC}\mathrm{Cos}\ \Omega - X_{HC}\ \mathrm{Sin}\ \Omega) + j(R_{HC}\mathrm{Sin}\ \Omega + X_{HC}\ \mathrm{Cos}\ \Omega)\} \quad (2)$$

$$\Rightarrow V_C = V_{UG} + \frac{(P_{DER} - P_C)}{V_C} \times R_{HC} + \frac{(Q_{DER} - Q_C)}{V_C} \times X_{HC}$$

$$\Rightarrow V_C(V_C - V_{UG}) = (P_{DER} - P_C) + (R_{HC} + \tan c \times X_{HC})$$

$$\Rightarrow (P_{DER} - P_C) = \frac{V_C(V_C - V_{UG})}{(R_{HC} + tan\ c \times X_{HC})}$$

$$\Rightarrow P_{DER} = \frac{V_C(V_C - V_{UG})}{(R_{HC} + tan\ c \times X_{HC})} + P_C$$

$$P_{HC} \ = \ V_C \times \frac{(V_{C\ Max} - V_{UG})}{(R_{HC} + X_{HC} * tan\ c)} + P_C \quad (3)$$

Eq. (3), depicts the HC of a node in the active distribution network. Due to addition of such DERs in the distribution network nodes, the current demand from the grid reduces and this reduction of current also depends on the number of DERs connected in the network as well as on the size of DERs on the distribution network. This may influence the protection setting of the protective relay. This phenomena may be concerning considering the growth rate of SPV integration into the utility grid. This generic issue for all the utility grid may degrade the SA of the system engineer to a very low level. To improve such degradation in SA, ML approach may be beneficial. Although various protection philosophy can be implemented as described in [15], but capital investment on equipment and on trained manpower may be economically as well as technically not feasible at all conditions. In this context ML algorithms can be utilised to detect fault and to improve SA in the complex distribution network.

3. Modelling of active distribution network in typhoon HIL real-time environment

For accurate analysis, real-time distribution system modelling is essential. Although the proposed methodology is applicable for any

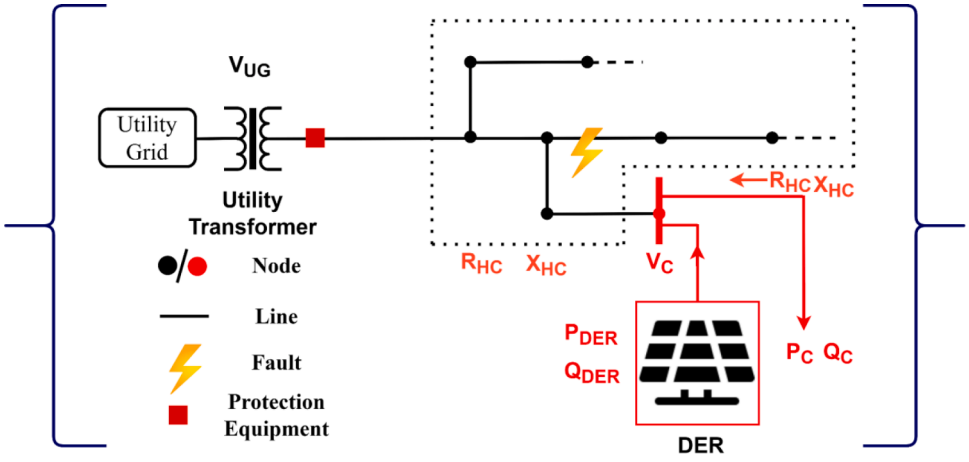

Fig. 1. Topology for hosting capacity analysis.

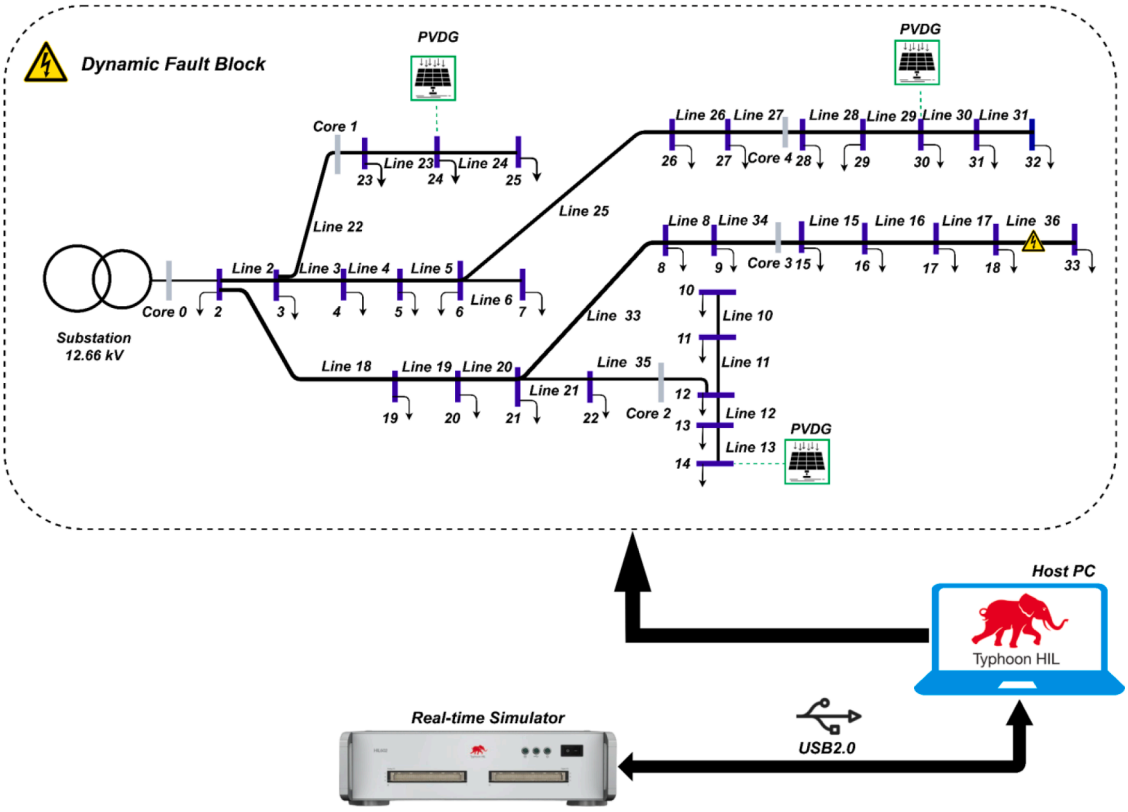

Fig. 2. Real-time simulation model of reconfigured IEEE-33 bus distribution system.

distribution network, but a reconfigured IEEE-33 bus system is considered for analysis of the proposed methodology. The consideration of reconfigured IEEE-33 bus system is due to various multi-dimensional positive aspects, which are detailed in [12]. It is to be noted that the objective of this work is to analyze the fault type and location in the active distribution network in the presence of DERs under variable fault resistance, type, and location. Hence the network with the location of PVDG as in [30] is experimented, as the losses obtained are low for the reconfigured case. The plant capacity is chosen according to the HC of the respective nodes as in [30]. Dynamic faults are created across the length of the network to collect the data for processing.

To test the proposed methodology for fault analysis in the case of PVDG penetrated distribution system, accurate data is required. To ensure the accuracy of data, the data are collected from the model executed in the real-time digital simulator as in Fig. 2. The test system developed is in Typhoon HIL's real-time environment, where six cores are used in parallel for a high sampling rate. As the proposed methodology deals with transient analysis, the sampling rate is further increased by optimizing the core partition, ranging from "core 0" to "core 4". Due to the involvement of the distribution system in this partitioning process of FPGA cores, core couplings are used. To simulate this network in a real-time platform, "configuration 2" in Typhoon HIL602+ device is considered, so that maximum number of FPGA cores can be utilised. The high-resolution data from the real-time simulation is recorded, which is as high as 2 million samples per second from the internal signal scope of the simulator. The recorded high-resolution data from the modeled distribution system although contains vital power system information; but it is impossible to process all frequency domain data into useful information. Thus to handle this issue, features are extracted from the available data using spectral kurtosis technique.

4. Brief review on spectral kurtosis for feature extraction and its application for fault signals

The fault signals in the power system are inherently nonstationary in nature. Thus, the SKS technique can be employed to indicate the impulsiveness of these fault signals efficiently. Authors in [31] detected the fault signal using SKS for a 5-bus system. Although detection is done, but it is impossible to draw all the required information about the event manually from the displayed kurtosis plots, when the data is huge. Due to this drawback, the method adopted in [31] cannot alone be effectively applied to the real distribution system, where inherently the fault data set will be high due to variation in location, type and resistance of faults. Proper calculation and selection of the SKS parameters needs to be done for better efficiency, which is detailed in the current section and Section 6, respectively. The detailed procedure for calculation of SKS parameters for a highly nonstationary signal is explained as follows.

The World-Cramer decomposition theory states that any nonstationary stochastic signal $L(t)$ can be expressed as output of a linear, causal, and time-varying system as expressed in Eq. (4), where $P(t,f)$ can be inferred as a complex envelope of $L(t)$ at a frequency f and $dU(f)$ as a unit variance orthogonal process [24,32]. Though $P(t,f)$ is considered as a deterministic function, there can be many instances when it is stochastic due to the time datum being unknown. To accommodate this, a more rugged approach is developed in view of a stochastic complex envelope $P(t,f,a)$. The shape of the envelope is governed by the random variable 'a'. This transforms the $L(t)$ process as a CNS process i.e., the process is generally stationary except at certain outcomes 'a' where it is nonstationary. The CNS processes have flatter tail for its probability density function in comparison to that of its generating system. Hence, with time-domain randomization, CNS can be constructed for any nonstationary signal. The spectral moments of order "$2n$" of the CNS process can be expressed as per Eq. (5) by ensemble averaging of many

outcomes for random variable 'a' [33]. The assumptions made in the equation includes $P(t,f)$ as a time stationary random field, independent of $dU(f)$ and $U(t)$ being a white process [34]. It is worth noting that taking the value of n as 1 in Eq. (5) yields the classical power density spectrum. The SKS is finally derived from the normalized fourth-order spectral cumulant of a CNS process as represented in Eq. (6). Factor 2 is used in Eq. (6) instead of factor 3 (which is generally used in classical cumulants definition) because $dU(f)$ is a circular random variable.

The SKS can be computed in many ways. The most widely used method is the STFT-based SKS [32,35]. The STFT of a signal $L(n)$ having an analysis window $w(n)$ with size N_W and a defined temporal step Q can be written as per Eq. (7). The length of the window is so chosen such that the analysis window is able to cover the quasi-stationary part of the signal and it is able to sample the complex envelope in a faster manner, lest some information is lost [32]. Thus, the spectral moment of order "$2n$" of $L_w(kQ, f)$ can be written as per Eq. (8), where $\overset{\Delta}{=}\langle.\rangle_k$ is called a time aggregator operator over an index 'k'. Hence, Eq. (9) shows the STFT based spectral kurtosis for a signal [26,36] .

$$L(t) = \int_{-\infty}^{+\infty} e^{2\pi jft} P(t,f)dU(f) \tag{4}$$

$$M_{2nL}(f) \overset{\Delta}{=} E\left\{|P(t,f)dU(f)|^{2n}\right\} \Big/ df = E\left\{|P(t,f)|^{2n}\right\}.M_{2nU} \tag{5}$$

$$SKS_L(f) \overset{\Delta}{=} \frac{M_{4L}(f)}{M_{2L}^2(f)} - 2,\ f \neq 0 \tag{6}$$

$$L_w(kQ, f) \overset{\Delta}{=} \sum_{n=-\infty}^{\infty} L(n)w(n-kQ)\exp(-2\pi jnf) \tag{7}$$

$$\widehat{M}_{2nL}(f) \overset{\Delta}{=} \left\langle |L_w(kQ,f)|^{2n} \right\rangle_k \tag{8}$$

$$\widehat{SKS}_L(f) \overset{\Delta}{=} \frac{\widehat{M}_{4L}(f)}{\widehat{M}_{2L}^2(f)} - 2,\ |f-|2|| > N_W^{-1} \tag{9}$$

SKS, being a fourth-order spectrum, offers resistance to noise, which helps in depicting the impulsiveness of the signal more vividly. Thus, the electrical anomalies in a power system can be detected more accurately with SKS. Additionally, SKS is computationally faster and less expensive tool [35–37]. Thus, SKS takes low values for non-transient signals while high values for transient signals. Therefore, SKS is competent in capturing signal transients effectively.

$$y(t) = C_1 sin(2\pi ft + \delta) \tag{10}$$

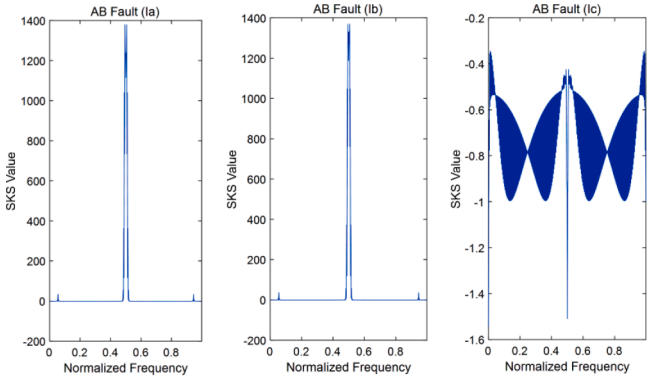

Fig. 3. SKS graphs for phase currents during AB fault for reconfigured IEEE-33 bus distribution system.

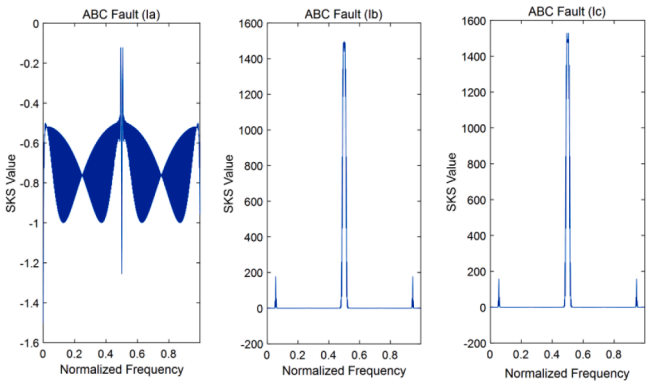

Fig. 4. SKS graphs for phase currents during ABC fault for reconfigured IEEE-33 bus distribution system.

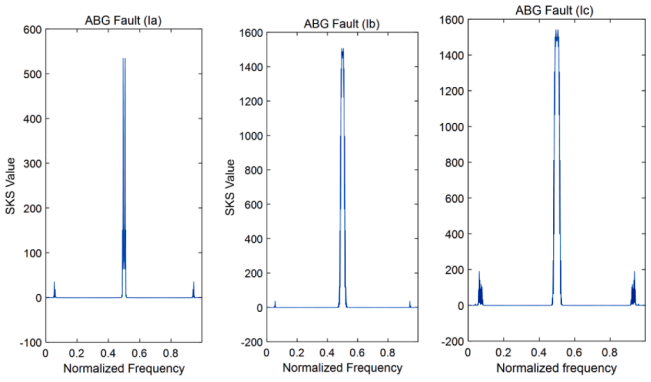

Fig. 5. SKS graphs for phase currents during ABG fault for reconfigured IEEE-33 bus distribution system.

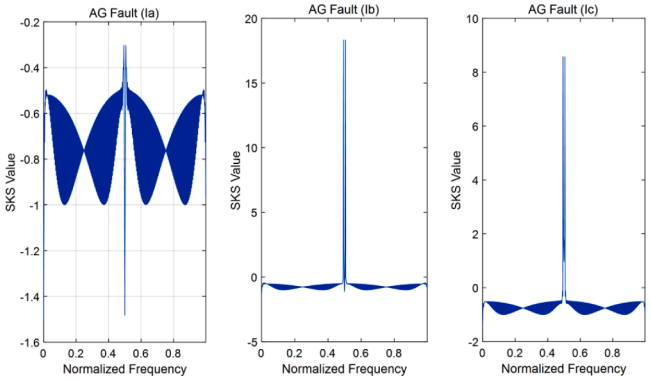

Fig. 6. SKS graphs for phase currents during AG fault for reconfigured IEEE-33 bus distribution system.

$$\widehat{SKS}_L(f) \overset{\Delta}{=} \frac{\widehat{M}_{4L}(f)}{\widehat{M}_{2L}^2(f)} - 2 = \frac{E\left\{|c_1|^4\right\}}{E\left\{|c_1|^2\right\}} - 2 = 1 - 2 = -1 \tag{11}$$

To depict the applicability of SKS in fault detection and classification, several types of faults are created between nodes 18 and 33 of the modelled IEEE-33 bus reconfigured distribution system as shown in Fig. 2. The graphs of SKS for these different fault condition is presented in Figs. 3–6, respectively. For obtaining the SKS, a window of Hanning type with size 256 having 75% overlap is employed for SKS as advocated in [34,23]. In the SKS graph, a stationary component is indicated by a value of −1. This can also be proved analytically with an example of a

sinusoid. A constant amplitude (C_1) and frequency (f) sinusoid signal with a constant initial phase (δ) can be represented as in Eq. (10). The SKS can then be calculated as per Eq. (11). The nonstationary (transients) components of the signal are indicated by positive peaks [37,38]. Thus, as seen from the graphs, the nature of the SKS is different for various faults. This indicates that the nature of the transients varies with the fault types, and the differences are effectively captured in the SKS graphs.

The graphs of SKS as in Fig. 3–6, shows different values for different types of faults, is utilized as an input to the multiple ML classifiers to analyse the fault location and fault type.

5. Application of HGB technique for accurate fault prediction

Ensemble method uses a conglomeration of models to achieve a better performance result. The elementary apprehension for the success of ensemble method are statistical, computational, representational, bias-variance decomposition, and strength correlation. Given the theoretical justifications behind ensemble methods, a vast number of models under ensemble method are well equipped for classification, regression, and optimization fields. Since the problems addressed in fault classification and fault location determination are supervised learning classification task, with multiple labels, methods are explained in this section with a perspective to address classification difficulties. Classification aims to identify a discrete category of new observation by studying a training set (S) of data [with 'm' rows/ data samples and 'n' features] as represented by Eqs. (12)–(13), where 'X' is the new observation as a feature vector with features [$F_1, F_2, F_3, \ldots, F_v$], '$y$' is the category that the new observation belongs to, $f(\cdot)$ is the trained classification function, 'α' is the classification function's parameter set, 'Z' is the set of class labels.

$$y = f(X, \alpha),\ y \in Z \tag{12}$$

$$X = [x_1, x_2, x_3, \ldots, x_v] \tag{13}$$

Boosting follows a sequential mechanism where weak learners are sequentially produced during the training phase. The weak learners are hence trained on different distributions of the training set to create weak rules. The predictions from the multiple generated weak learners are combined to form one strong decision rule that can correctly classify the labels. The parameters required as input are the number of classifier models or weak learners (say 't') and the training dataset (having size of 'm' rows/records). The steps included in the normal gradient boosting (NGB) algorithm are as follows:

1 The first base learners take the complete training dataset and assigns similar weights to each of the records or observations in the dataset.
2 When some error occurs in the prediction given by the first base learners, greater attention is given to those records by assigning more weights to them. The second base learners by then are applied again to the total training dataset with 'm' rows.
3 Step 2 iterates for all the base learners until a better accuracy is achieved by aggregating all the base learners' decision rules.

In the NGB method, decision tree-based weak learners are kept on adding sequentially and the model is developed. The main disadvantage of gradient boosting models is that the time complexity of the models are high because of which the model takes a longer time to train. This is because the trees have to be added sequentially and cannot be trained in parallel, exploiting multiple CPU cores. Hence, to have a better computing efficiency alongside fast training, HGB model evolved.

Since most of the algorithms under ensemble method requires the use of decision trees, hence the speed of the algorithms gets constricted by the number of rows and features in the training dataset resulting in slow construction of the trees. Therefore, the building of decision trees is speeded up in HGB by reducing the quantity of values of the continuous set of features through discretization or binning into a fixed number of cumulative buckets. Tailored implementation of histograms act as an efficient data structure for binning of the input data during the construction of the decision trees. The coarse approximation of putting into integer brackets does impact the performance of the HGB models to a greater extent, but it helps dramatically of accelerating the algorithm in terms of building of the decision trees.

The inputs of a HGB model comprises a training dataset (with 'm' rows/records and 'n' columns/features), a number of weak learners (c) and a loss function which is differentiable as represented by Eq. (14), where, 'L' is categorical cross-entropy loss/ log loss since there are multiple class labels that has to be predicted. The steps followed in HGB are as follows:

1. A base model is computed which provides a constant value as its prediction output presented in Eq. (15).

   Where,
   F_0 is the initial base model function,
   $\widehat{y}$ is the predicted value which is its constant for all 'm' observations; Finding its minimum value is a 1-dimensional optimization problem where first order derivative of the loss function with respect to $\widehat{y}$, gives the output of minimum error.

2. Iterate for all the weak learners i.e. $k = 1$ to 'c_2':
   a Pseudo-residuals are computed for every observation taking the actual value and the predicted value using the formula in Eq. (16), where, $i = 1, 2, 3, \ldots, m$ for all rows/ records/ observations.
   b A weak learner $h_b(x)$ is trained on the pseudo-residuals, using the same dataset feature columns, with the residuals as the predictor column label. Therefore, the residual column produced by $R_{i,k}$ acts as the actual label, which helps in training. The new column produced namely $h_b(x) = \widehat{R_{i,\,b}}$ is the predicted value after training.
   c The multiplier/ learning rate α_b is computed by solving the following optimization problem as per Eq. (17).
   d The model is updated as per Eq. (18).
3. Finally, the output $F_c(x)$ is obtained, which can be represented by Eq. (19):

$$L(y, F(x)) \tag{14}$$

$$F_0\langle x\rangle = \underset{\widehat{y}}{\operatorname{argmin}} \sum_{i=1}^{m} L\langle y_i, \widehat{y}\rangle \tag{15}$$

$$R_{i,b} = -\left[\frac{\partial\, L\langle y_i, F_{b-1}(x_i)\rangle}{\partial\, F_{b-1}(x_i)}\right] \tag{16}$$

$$\alpha_b = \underset{\alpha}{\operatorname{argmin}} \sum_{i=1}^{m} L(y_i,\ F_{b-1}(x) + \alpha h_b(x)) \tag{17}$$

$$F_b(x) = F_{b-1}(x) + \alpha_b h_b(x) \tag{18}$$

$$F_t(x) = F_0\langle x\rangle + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \ldots + \alpha_t h_t(x)$$

$$= F_0\langle x\rangle + \sum_{i=1}^{t} \alpha_i h_i(x) \tag{19}$$

The classifiers need data for the preparation of its training and testing dataset. While the accuracy of the classifiers increases with the increase in the number of data features, but it comes at the cost of increased computational complicacy, time, and involved budget. Thus, the number of data features needs to be optimized so that a balance is achieved between accuracy, computational complicacy, time, and cost. Fig. 7, shows the overall flow of the work.

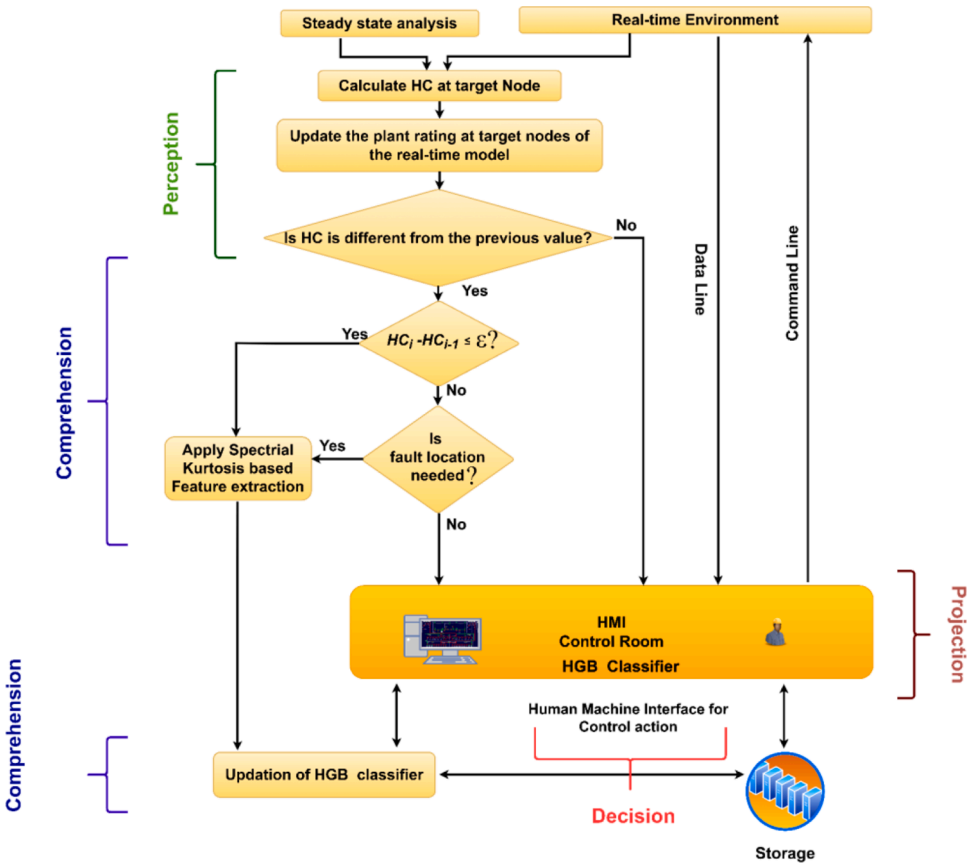

Fig. 7. Proposed methodology for situational awareness based adaptive fault diagnosis with hosting capacity alteration.

6. Proposed methodology for situational awareness based adaptive fault diagnosis with hosting capacity amendment

Precise SA about unforeseen events to the control engineers and proactive legitimate decision that is most suitable to an event, is discussed in this proposed strategy. The proposed methodology considered can be represented as in Fig. 7. The methodology used can be summarized in the following points:

a The steady-state analysis is performed considering the PVDG specification that the DSO plans to integrate in the distribution system.
b The analysis of the HC with the integrated PVDG in the distribution grid is performed.
c The real-time modeling of the analyzed network as in (b) is performed in Typhoon HIL, and all possible fault scenarios are generated.
d The features of SKS are obtained from these generated scenarios that are fed to the ML classifier for fault classification and fault location detection. For the initial training of the classifier as well as for retraining, the procedure is followed as discussed in Section 5.
e Generated fault signal is fed to ML classifier, checked if classifier model is able to distinguish and locate the faults correctly.
f Update the HC in the network. If the altered HC is less than and equal to 'ε' no up-dation is required in the classifier i.e, retraining of the classifier is not needed.
g If the altered HC is greater than 'ε', training of the classifier is again executed with the new features obtained after additional DER integration.

These steps are performed for every additional PVDG integration. The extend of PVDG integration whose value is greater than 'ε', may need retraining of the model, which will ensure the highest degree of SA with negligible investment, for accurate fault location and type prediction. It is worth noting that the value of 'ε' and its selection totally depends on the utility.

To clearly understand the work flow, the process of data generation, feature extraction, and choosing the suitable features for training the classifier is detailed in the Section 7.

7. Result and analysis: data generation, feature extraction, feature selection and performance evaluation

7.1. Real-time data generation for different fault location and type

To mimic the actual power system distribution network, the datasets required for the classifier are taken from real-time Typhoon HIL simulator for maximum reliability and accuracy. The data required for the training as well as the testing purpose of the targeted classifier are considered with variable fault resistance, location, and type as in Table 1 for the considered IEEE 33 bus reconfigured distribution system as in Fig. 2. The fault data are collected from various locations throughout the network. For classification of fault type, the predicted output is considered as LG, LLG, LLL, and LL fault. For fault location detection, the predicted output is considered as the location between different nodes.

Table 1
Real-time data generation for different fault scenarios in reconfigured IEEE 33 bus distribution system.

Classification task	Description	Dataset size		
		Training	Testing	Total
Fault type	Different types of fault with variable fault resistance at different location	1402	468	1870
Fault location		1447	483	1930

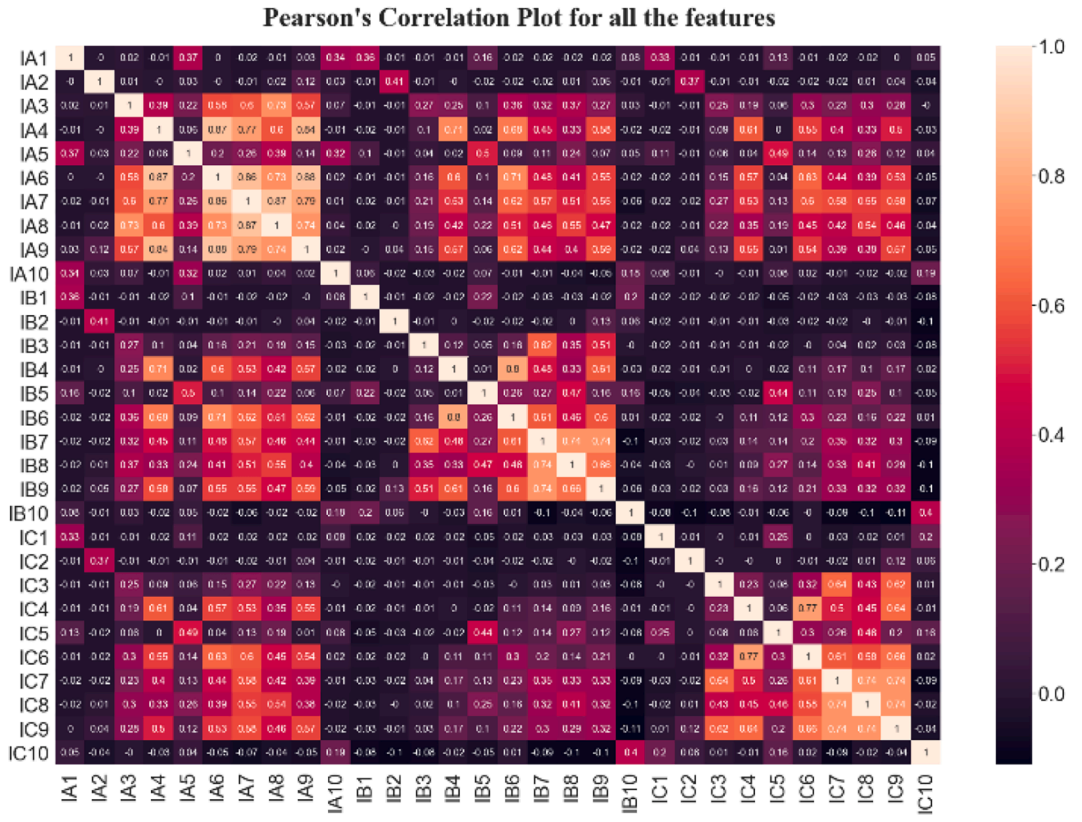

Fig. 8. Pearson's correlation plot for all the features to identify the fault type tested on reconfigured IEEE-33 bus distribution system.

Table 2
Details of fault type dataset for reconfigured IEEE-33 bus distribution system.

Fault type	AB	AC	BC	AN	BN	CN	ABN	ACN	BCN	ABC
Number of faults	187 per fault									
Training Dataset	1402									
Testing Dataset	468									

To proceed with the development of the classifier, thirty features are generated (IA_n, IB_n, IC_n), $n \in N, 0 \leq n \leq 10$ from the faulted current signals. These thirty features are the SKS values normalized at frequencies ranging from 0.1 to 1 for each phase current signal. For the determination of the efficient features to be utilized in the classifier, a clear picture of correlation mapping among all the possible features are needed to be performed. Fig. 8, presents a heat map of Pearson's correlation matrix plot among all the possible features. The diagonal values in Fig. 8 i.e., the value 1 shows the maximum correlation between the features and the value −1 shows the minimum co-relation between the features. For better clarity, the numerical values of the heat map plot are rounded off upto two places. Table 2 shows the details of the dataset utilized for the fault classification problem. During the development of the classifier, 75% of the dataset are used for the training purpose and 25% of the data are used for the testing purpose.

7.2. *Feature selection and accuracy optimization for identification of fault type*

The co-relation plot in Fig. 8 shows thirty features considered for each set of data, which is time-consuming. This problem can be solved by optimally choosing the set of least co-related features contributing to

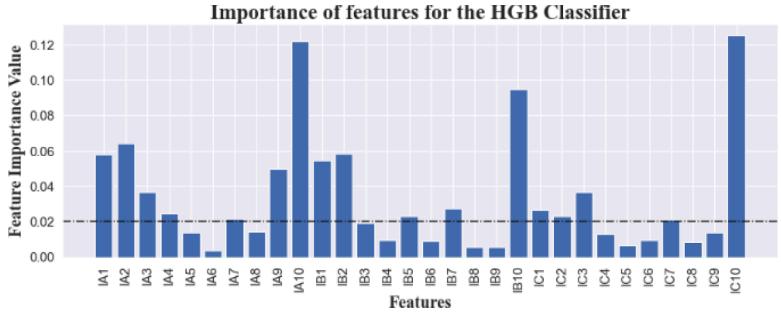

Fig. 9. Feature importance value for fault classification in reconfigured IEEE-33 bus distribution system.

Table 3
Accuracy optimization with reduced features for fault type determination in reconfigured IEEE-33 bus distribution system.

DT	NGB	XB	RF	CB	HGB	AB	LGBM
0.89	0.93	0.94	0.95	0.95	0.95	0.95	0.96
(30)	(30)	(30)	(30)	(30)	(30)	(30)	(30)
0.87	0.94	0.94	0.96	0.94	0.95	0.96	0.96
(26)	(26)	(26)	(25)	(26)	(26)	(26)	(26)
0.89	0.93	0.95	0.96	0.95	0.96	0.94	0.95
(21)	(21)	(21)	(21)	(21)	(21)	(21)	(21)
0.88	0.92	0.95	0.93	0.95	0.94	0.94	0.94
(17)	(16)	(16)	(17)	(16)	(18)	(16)	(16)
0.86	0.91	0.94	0.95	0.94	0.96	0.93	0.93
(16)	(14)	(15)	(16)	(15)	(16)	(14)	(14)
0.85	0.91	0.93	0.95	0.93	0.94	0.93	0.95
(14)	(11)	(14)	(14)	(14)	(14)	(11)	(12)
0.84	0.89	0.91	0.92	0.92	0.93	0.89	0.91
(10)	(9)	(9)	(13)	(9)	(9)	(9)	(7)
0.80	0.84	0.89	0.90	0.88	0.89	0.88	0.89
(7)	(6)	(6)	(7)	(6)	(6)	(8)	(6)
0.76	0.76	0.79	0.82	0.79	0.82	0.81	0.80
(3)	(3)	(3)	(3)	(3)	(3)	(3)	(3)

the highest accuracy from the co-relation plot. However, it is very tedious to determine the least co-related features directly from the plot. To get a clearer picture of this, the histogram in Fig. 9 is constructed, which represents the feature importance value for the optimal selection of features in order to gain maximum accuracy with minimal features. With numerous data analysis collected from the real-time simulator, the threshold for feature importance value is chosen as 0.02. To verify the claimed threshold value, detailed analysis, with different number of input features along with the multiple ensemble modes are presented in Table 3 namely Decision Tree (DT), NGB, Xg Boost (XB),Random forest (RF), CatBoost (CB), HGB, AdaBoost (AB), Light GBM (LGBM). It can be observed from Table 3 that different number of features are considered for the same row. This is due to the fact that the features exceeding the threshold feature value are distinct for different types of algorithms. From the results produced in Table 3 for reduced features, it can be observed that HGB with 16 features result in the highest accuracy of 96% for detection of fault type, as compared to the other seven methods with various input features.

To explain the learning of HGB in a more descriptive manner, a pictorial view of the learning procedure for predicting fault type in the considered reconfigured IEEE -33 bus system with three features are shown in Fig. 10. The reason for showcasing the pictorial representation only for three features is due to the constraint in space as the total

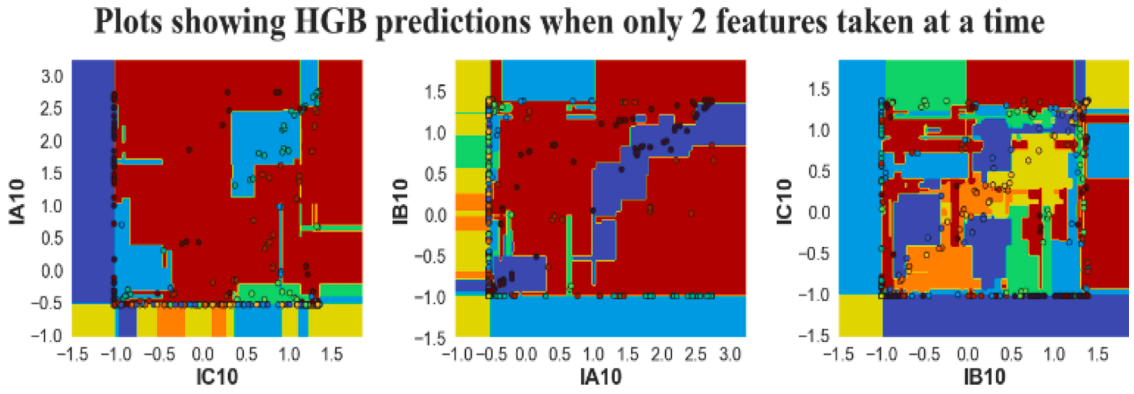

Fig. 10. Predictions of fault type using HGB method.

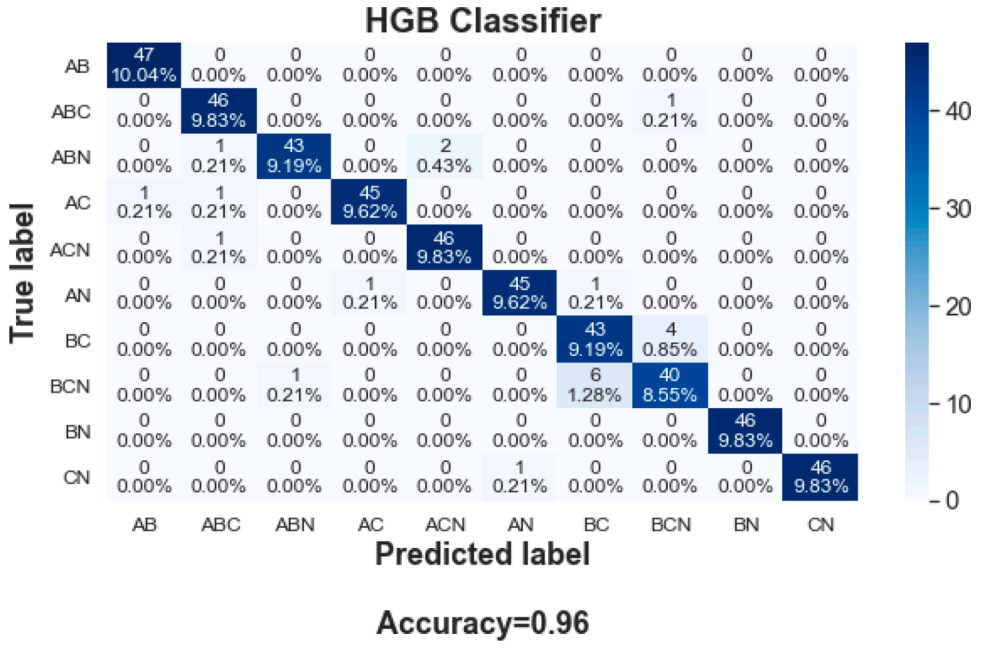

Fig. 11. Confusion matrix considering reduced features for determination of fault types tested for reconfigured IEEE-33 bus distribution system.

number of figures generated will be ${}^{n}C_{r}$. Where 'n' represents a number of features and 'r' represents 2 as it is a 2D figure. It shows the training as well as the prediction of the HGB. The colored surfaces behind the dots showcase the training procedure, and the dots show the prediction values. If the background surface color and the dot color matches, the condition is explained that the prediction is correct with respect to its training. If dot and the background have different colors, then the prediction is different from the training, which is not a favorable condition.

From the above results as in Table 3 and Fig. 10, it is evident that HGB with sixteen features have the maximum efficacy with minimum number of features. This unique characteristic makes HGB suitable for fault detection in the active distribution system. Thus, in the present work, HGB with 16 features are considered for fault classification. These 16 features include IA1, IA2, IA3, IA4, IA7, IA9, IA10, IB1, IB2, IB5, IB7, IB10, IC1, IC2, IC3 and IC10, respectively. The fault detection results considering only 16 features at a time with the accuracy of 96 % is shown in confusion matrix in Fig. 11. The "Y axis" represents the true value and "X axis" represents the predicted value by the classifier. The diagonal elements represent the correct prediction by the classifier and the off-diagonal element represents the flawed prediction by the classier. The dark color shows that the predicted values are the same as the true values. With the fade in the color depth, the predicted value diverges from the true value as per the color scale showcased in Fig. 11.

7.3. Feature selection and accuracy optimization for identification of fault location

For fault location estimation, a similar approach is adopted, i.e., the Person's correlation plot is utilized to find out the correlation among the input features. The total data set selected for fault location identification is 1930. From the data set of 1930, 1448 data sets are used for the training purpose. The remaining 482 data sets are used for testing purpose for fault location. All the features shown in the Person's correlation plot for training as well as testing can be considered as in Fig. 8. Consideration of all the features for training as well as testing may not result in the highest accuracy, although it is time-consuming. In such a case, the input feature reduction according to the feature importance value can provide better results. Fig. 12 represents the selection of important features for the HGB algorithm, where with minimum of 21 features, the attained accuracy is 83%.

Table 4 result verifies the importance of claimed feature reduction for improving the accuracy with fewer features, as shown in Fig. 12. With rigorous real-time data analysis, the feature importance value is selected as '0.03'. With a specific feature importance value, it is observed that different algorithms have a different number of input features and accuracy. For example, with a feature importance value of 0.03 the DT, NGB, XB and RF algorithm shows the accuracy of 66%, 74%, 81%, and 79%, respectively. Thus, in the present work, HGB with 21 features is considered for fault location determination. These 21 features include IA2, IA3, IA7, IB1, IB2, IB4, IB5, IB6, IB7, IB8, IB9, IB10, IC1, IC2, IC3, IC4, IC5, IC6, IC7, IC8, and IC10 respectively.

Table 4
Accuracy optimization with reduced features for fault location identification in reconfigured IEEE-33 bus distribution system.

DT	NGB	XB	RF	CB	HGB	AB	LGBM
0.68	0.77	0.82	0.79	0.78	0.82	0.83	0.82
(30)	(30)	(30)	(30)	(30)	(30)	(30)	(30)
0.66	0.76	0.82	0.81	0.78	0.83	0.83	0.83
(26)	(26)	(26)	(25)	(26)	(26)	(26)	(26)
0.66	0.74	0.81	0.79	0.79	0.83	0.80	0.82
(21)	(21)	(21)	(21)	(21)	(21)	(21)	(21)
0.65	0.73	0.80	0.80	0.76	0.79	0.80	0.79
(17)	(16)	(16)	(17)	(16)	(18)	(16)	(16)
0.66	0.72	0.77	0.79	0.77	0.78	0.79	0.79
(16)	(14)	(15)	(16)	(15)	(16)	(14)	(14)
0.66	0.71	0.76	0.77	0.77	0.78	0.79	0.77
(14)	(11)	(14)	(14)	(14)	(14)	(11)	(12)
0.63	0.68	0.76	0.77	0.70	0.73	0.79	0.76
(10)	(9)	(9)	(13)	(9)	(9)	(9)	(7)
0.62	0.59	0.72	0.72	0.69	0.70	0.78	0.71
(7)	(6)	(6)	(7)	(6)	(6)	(8)	(6)
0.54	0.55	0.60	0.65	0.52	0.62	0.65	0.63(3)
(3)	(3)	(3)	(3)	(3)	(3)	(3)	

Considering the reduced 21 features as in Table 4 for the HGB algorithm, the confusion matrix is built for understanding of correctness regarding the location identification as in Fig. 13. To verify HGB algorithm's performance for identifying fault type and location in the active distribution system, the performance of other methods such as K-Nearest neighbour, Logistic regression, Gaussian RBF Kernel SVC, Gaussian naïve Bayers, Voting classifier, Gaussian RBF Kernel SVC, auto encoder, DT, NGB, XB, RF, CB, AB and LGBM needs to be checked under the identical condition as that of the HGB algorithm. Fig. 14 shows a detailed comparison of these mentioned methods under identical conditions. It is to be observed that the HGB algorithm has the highest accuracy, with a minimum number of features. A time-based performance analysis is also performed in details for both fault type as well as location identification. In this context, utilization of selected optimal values as in Figs. 9 and 12 is very crucial, as this decides the execution time efficiency. A detailed time efficiency performance of the projected framework is displayed in Table 5. From Table 5; it can be inferred that with optimized feature importance values, 39.39% time is reduced for fault type detection, which is 0.20 ms. Similarly, the time required for fault location identification is reduced by 13.09%, which is to 0.73 ms from 0.84 ms.

7.4. Performance comparison of the proposed method with existing reported methods in the literature

A detailed comparative analysis of the proposed method existing literature are tabulated in Table 6. In [21], the accuracy varies between

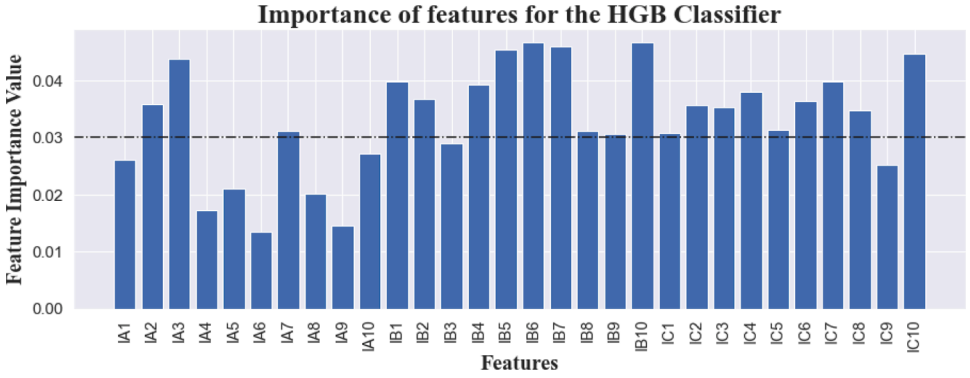

Fig. 12. Feature importance value for fault location in reconfigured IEEE-33 bus distribution system.

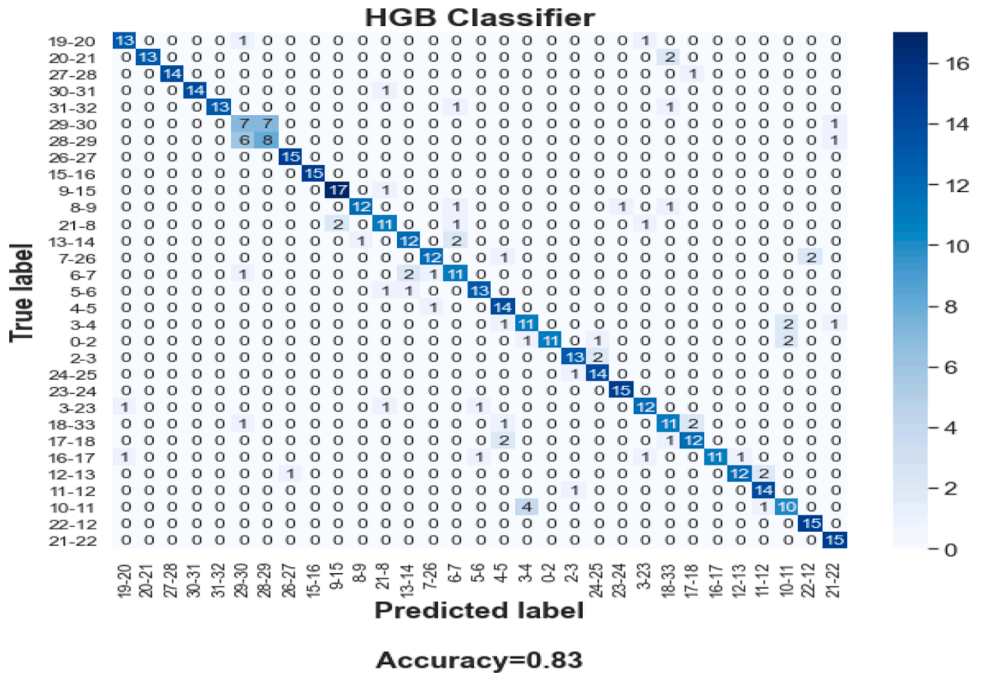

Fig. 13. Confusion matrix considering reduced features for fault location.

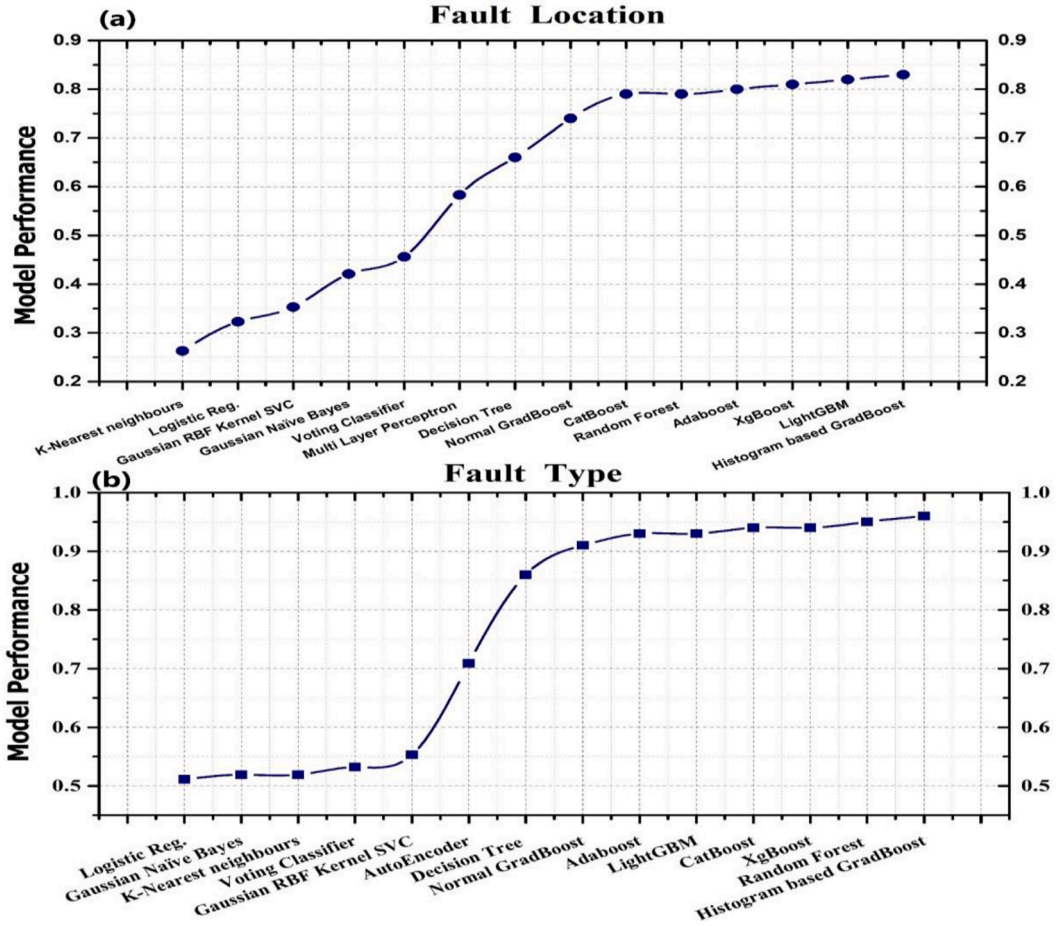

Fig. 14. Comparison of various model performance with respect to HGB for (a) Fault location, (b) Fault type.

Table 5
Time based performance analysis of HGB method.

	Fault type feature importance		Fault location feature importance	
	Without optimization	With optimization	Without optimization	With optimization
Time (in mSec)	0.33	0.20	0.84	0.73
% Improvement	-	39.39	-	13.09

Table 6
Comparison with existing literatures.

Technique Parameters	Proposed	[20]	[21]	[22]	[23]	[9,10]
Detection time	0.20 msec	NA	NA	NA	33.79 msec	40 msec
Hosting capacity considered	Yes	No	No	No	NA	NA
Fault detection accuracy	96%	99.7%	85–94%	95.63%	92.52%	100%
Fault detection accuracy dependence	Independent	On differential communication system	NA	NA	On sampling frequency	NA
Fault location detected	Yes	No	No	No	No	NA
Samples / Cycle	20-2K	NA	NA	64	NA	4
Real-time validation	Yes	No	Yes	No	No	Yes

Table 7
Testing HGB algorithm based on HC amendment.

	Cases	Differentiable as per previous trained classifier?	Retraining required?	Differentiable as per new trained classifier
Fault Type	Variation in HC beyond '$\boldsymbol{\varepsilon}$'	Partially	Yes	Yes
	Variation in HC within '$\boldsymbol{\varepsilon}$'	Yes	No	Yes
Fault Location	Variation in HC beyond '$\boldsymbol{\varepsilon}$'	Dependent on 'ε'	Yes	Yes
	Variation in HC within '$\boldsymbol{\varepsilon}$'	Yes	No	Yes

85 and 94% as compared to the proposed method. Although the detection accuracy is good in [20], the technique relies on a differential communication mechanism which may diminish the reliability of the overall system. The accuracy of the proposed method as per Table 6 is not only better over [22], but also when tested in wide range of sampling frequency, the accuracy is found to remain unchanged under wide variation in sampling frequency. Due to the capability of accurate fault detection under wide range of sampling frequency, micro - PMUs and improved IEDs at substation with variable sampling time can be used in this proposed method. The fault detection time in the proposed method is 0.20 msec which is very negligible as compared to the method explained in [23]; additionally, the accuracy of the proposed method is high as compared to the methods in [21–23]. The detection accuracy for [9,10] is found to be 100% and for the study the authors consider only one node of the utility, whereas the proposed method is applied to all the nodes of the reconfigured IEEE 33 bus distribution system. The authors in this present work applied SVM method as in [9,10], but the accuracy is found to be low, which is 55.3% and 35.3% for fault type and location detection respectively. Additionally, the authors set sampling time of 5 msec i.e, 4 samples per cycle as depicted in Table 6, which may not able to capture fast transients. Unlike other literature explained in Table 6, the proposed method also considers the HC of the network for the fault analysis.

With the amendment of HC in network the proposed HGB classifier is tested to accurately comprehend the SA of the network. The detailed analysis for various possible conditions are tested and tabulated in Table 7.

With the change in HC beyond 'ε', the training may be required depending on the application. For fault type detection, the training is not required as the proposed ML algorithm will be able to distinguish these irrespective of the variation in HC, but it is preferable to train the model to have maximum accuracy. In case of fault location, the model need to be retrained to give accurate fault location in the ADN.

7.5. Performance of the proposed method in other transient scenarios

In the aforementioned sections, the algorithm has been trained and tested with fault scenarios only. However, in the modern distribution grid, there are some scenarios present which generates transients [25]. A fault detection algorithm must be able to differentiate between these transients. Different scenarios of transients are created, i.e. transients due to islanding, sudden load addition and curtailment, switching capacitor branch. Islanding is a scenario which causes transients that appears similar to fault transients. Islanding of ADN is a scenario, when the ADN is disconnected from the main grid either advertently or inadvertently. The island is caused either due to a fault, equipment failures, human error or due to some natural calamity. The algorithm is tested for islanding caused due to fault and due to other non-fault reasons. In case of non-faulted scenarios, the islanding is done by opening of respective circuit breakers manually. The results are tabulated in Table 8. It is observed that the algorithm appropriately functions and identify fault when islanding occurs due to fault. However, it shows no fault in case when islanding occurs due to some other reasons. Other power system scenarios that generates transients are sudden load connection or load curtailment and capacitor bank switching. All these cases are tested for the proposed algorithm as detailed in Table 8. The load connection scenarios are simulated by connecting additional load

Table 8
HGB algorithm tested for other transient scenarios.

Scenarios	Description	Is fault detected?
Island	Caused due to fault	Yes
	Caused due to other reasons	No
Sudden additional load connection	Loads of 100 kW at node 21	No
	Loads of 100 kW, 50 kVAr at node 26	No
Sudden load curtailment	Loads of 60 kW, 30 kVAr at node 5	No
	Loads of 150 kW, 70 kVAr at node 31	No
Capacitor bank switching	100 kVAr	No
	400 kVAr	No

of 100 kW at node 21 and 100 kW, 50 kVAr at node 26 respectively. Similarly, load curtailment scenarios are simulated by disconnecting the rated loads at node 5 and 31 respectively. Two scenarios of capacitor switching of rated value of 100 kVAr and 400 kVAr are considered. As observed from Table 8, the algorithm does not baffle to detect fault with these transients. This showcases that the algorithm can safely be employed in practical power systems.

8. Conclusion

The integration of DERs, smart metering, and monitoring devices leading to escalated automation is reforming the conventional distribution system to smart distribution system (SDS). This has transformed the power distribution system from a data deficit to a data surplus system. Thus improvement of reliability of SDSs demands enhanced network SA with the amplified perception, comprehension, and projection for precise decision making with the boon of accessibility of real-time network data.

Taking aforesaid conditions, it is indispensable to apply ML to data-rich SDS. Although these available high-resolution data are capable of projecting a clear SA about the power system events, but most of the data could not be used efficiently. In context to this, the paper proposes an adaptive ensemble learning-based fault classification methodology with alterations in the HC for a PV penetrated active distribution network. In this work, the algorithm classifies fault and also determines the location of the fault by analyzing the fault signals with SKS. Several features are extracted with SKS which are fed to a HGB ML classifier. The algorithm proves to have high efficiency and low computational time when compared with other methods such as K-Nearest neighbour, Logistic regression, Gaussian RBF Kernel SVC, Gaussian naïve Bayers, Voting classifier, Gaussian RBF Kernel SVC, auto encoder, DT, NGB, XB, RF, CB, AB and LGBM. On comparison of the proposed method with the existing literature, the proposed method proved to be noteworthy based on tested parameters. It also depicts it capability in distinguishing faults from other power system transients.

Another advantage of the algorithm includes the facility to update the classifier when DSOs plan to integrate additional PV to enhance the HC of the network. The developed algorithm is tested for real-time compatibility using Typhoon HIL environment, which guarantees the accuracy and practical applicability of the method.

CRediT authorship contribution statement

Sourav Kumar Sahu: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Millend Roy:** Methodology, Software, Writing – original draft. **Soham Dutta:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Debomita Ghosh:** Conceptualization, Validation, Formal analysis, Investigation, Writing – review & editing, Supervision. **Dusmanta Kumar Mohanta:** Visualization, Supervision.

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.

Data availability

Data will be made available on request.

References

[1] S.M. Ismael, S.H.E. Abdel Aleem, A.Y. Abdelaziz, A.F. Zobaa, State-of-the-art of hosting capacity in modern power systems with distributed generation, Renew. Energy 130 (2019) 1002–1020, https://doi.org/10.1016/j.renene.2018.07.008.

[2] S.K. Sahu, D. Ghosh, Operational hosting capacity-based sustainable energy management and enhancement, Int. J. Energy Res. 46 (3) (2022) 2418–2437, https://doi.org/10.1002/er.7317. Mar.

[3] M.S. Ballal, A.R. Kulkarni, Improvements in existing system integrity protection schemes under stressed conditions by synchrophasor technology—case studies, IEEE Access 9 (2021) 20788–20807, https://doi.org/10.1109/ACCESS.2021.3054792.

[4] Y.-K. Wu, S.M. Chang, Y.-L. Hu, Literature review of power system blackouts, Energy Procedia 141 (2017) 428–431, https://doi.org/10.1016/j.egypro.2017.11.055. Dec.

[5] M. H. J. Bollen, "Literature search for reliability data of components in electric literature search for reliability data of components in electric distribution networks," 1993.

[6] L. Ndahepele, S. Chowdhury, Impact of distributed generation on traditional protection in distribution and transmission systems: a review, in: 2020 IEEE PES/IAS PowerAfrica, PowerAfrica 2020, 2020, https://doi.org/10.1109/PowerAfrica49420.2020.9219840.

[7] P. Gopakumar, M.J.B. Reddy, D.K. Mohanta, Adaptive fault identification and classification methodology for smart power grids using synchronous phasor angle measurements, IET Gener. Transm. Distrib. 9 (2) (2015) 133–145, https://doi.org/10.1049/iet-gtd.2014.0024.

[8] M.F. Guo, X.D. Zeng, D.Y. Chen, N.C. Yang, Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems, IEEE Sens. J. 18 (3) (2018) 1291–1300, https://doi.org/10.1109/JSEN.2017.2776238.

[9] H.R. Baghaee, D. Mlakic, S. Nikolovski, T. Dragicevic, Support vector machine-based islanding and grid fault detection in active distribution networks, IEEE J. Emerg. Sel. Top. Power Electron. 8 (3) (2020) 2385–2403, https://doi.org/10.1109/JESTPE.2019.2916621.

[10] H.R. Baghaee, D. Mlakic, S. Nikolovski, T. Dragicevic, Anti-islanding protection of PV-based microgrids consisting of PHEVs using SVMs, IEEE Trans. Smart Grid 11 (1) (2020) 483–500, https://doi.org/10.1109/TSG.2019.2924290.

[11] IEA, "Solar PV - analysis - IEA," *June 2020*. p. 12, 2020.

[12] S.K. Sahu, G. Debomita, Hosting capacity enhancement in distribution system in highly trenchant photo-voltaic environment: a hardware in loop approach, IEEE Access 8 (2020) 1440–14451, https://doi.org/10.1109/ACCESS.2019.2962263.

[13] S. Sakar, M.E. Balci, S.H.E.A. Aleem, A.F. Zobaa, Hosting capacity assessment and improvement for photovoltaic-based distributed generation in distorted distribution networks, in: 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), 2016, pp. 1–6, https://doi.org/10.1109/EEEIC.2016.7555515.

[14] E. Kazemi-Robati, M.S. Sepasian, H. Hafezi, H. Arasteh, PV-hosting-capacity enhancement and power-quality improvement through multiobjective reconfiguration of harmonic-polluted distribution systems, Int. J. Electr. Power Energy Syst. 140 (2022), 107972, https://doi.org/10.1016/j.ijepes.2022.107972. Sep.

[15] Math Bollen, Fainan Hassan, Integration Of Distributed Generation In The Power System, A John Wiley & Sons, INC., New Jersey, 2014.

[16] J. Wijekoon, A.D. Rajapakse, N.M. Haleem, Fast and reliable method for identifying fault type and faulted phases using band limited transient currents, IEEE Trans. Power Deliv. 36 (5) (2020) 2839–2850, https://doi.org/10.1109/tpwrd.2020.3027793.

[17] R. Dashti, M. Ghasemi, M. Daisy, Fault location in power distribution network with presence of distributed generation resources using impedance based method and applying π line model, Energy 159 (2018) 344–360, https://doi.org/10.1016/j.energy.2018.06.111. Sep.

[18] S.H. Mortazavi, Z. Moravej, S.M. Shahrtash, A searching based method for locating high impedance arcing fault in distribution networks, IEEE Trans. Power Deliv. 34 (2) (2019) 438–447, https://doi.org/10.1109/TPWRD.2018.2874879. Apr.

[19] J. Li, M. Gao, B. Liu, F. Gao, J. Chen, Fault location algorithm in distribution networks considering distributed capacitive current, IEEE Trans. Power Deliv. 36 (5) (2021) 2785–2793, https://doi.org/10.1109/TPWRD.2020.3026835.

[20] S. Kar, S.R. Samantaray, M.D. Zadeh, Data-mining model based intelligent differential microgrid protection scheme, IEEE Syst. J. 11 (2) (2017) 1161–1169, https://doi.org/10.1109/JSYST.2014.2380432.

[21] D.P. Mishra, S.R. Samantaray, G. Joos, A combined wavelet and data-mining based intelligent protection scheme for microgrid, IEEE Trans. Smart Grid 7 (5) (2016) 2295–2304, https://doi.org/10.1109/TSG.2015.2487501.

[22] T.S. Abdelgayed, W.G. Morsi, T.S. Sidhu, A new approach for fault classification in microgrids using optimal wavelet functions matching pursuit, IEEE Trans. Smart Grid 9 (5) (2018) 4838–4846, https://doi.org/10.1109/TSG.2017.2672881.

[23] S. Jamali, S. Ranjbar, A. Bahmanyar, Identification of faulted line section in microgrids using data mining method based on feature discretisation, Int. Trans. Electr. Energy Syst. 30 (6) (2020) 1–16, https://doi.org/10.1002/2050-7038.12353.

[24] R. Dwyer, Detection of non-Gaussian signals by frequency domain kurtosis estimation, in: ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing 8, 1983, pp. 607–610, https://doi.org/10.1109/ICASSP.1983.1172264.

[25] S. Dutta, P.K. Sadhu, M.J.B. Reddy, D.K. Mohanta, Smart inadvertent islanding detection employing p-type μPMU for an active distribution network, IET Gener.

Transm. Distrib. 12 (20) (2018) 4615–4625, https://doi.org/10.1049/iet-gtd.2018.5805. Nov.
[26] C. Ottonello, S. Pagnan, Modified frequency domain kurtosis for signal processing, Electron. Lett. 30 (14) (1994) 1117–1118, https://doi.org/10.1049/el:19940777. Jul.
[27] M. Bajaj, A.K. Singh, Design and analysis of optimal passive filters for increasing the harmonic-constrained hosting capacity of inverter-based DG systems in non-sinusoidal grids, Electr. Eng. (2021), https://doi.org/10.1007/s00202-021-01415-1.
[28] W. Sun, Maximising Renewable Hosting Capacity in Electricity Networks, The University of Edinburgh, Edinburgh, 2014.
[29] S. Jothibasu, A. Dubey, S. Santoso, Two-stage distribution circuit design framework for high levels of photovoltaic generation, IEEE Trans. Power Syst. 34 (6) (2019) 5217–5226, https://doi.org/10.1109/TPWRS.2018.2871640.
[30] S. Kansal, V. Kumar, B. Tyagi, Hybrid approach for optimal placement of multiple DGs of multiple types in distribution networks, Int. J. Electr. Power Energy Syst. 75 (2016) 226–235, https://doi.org/10.1016/j.ijepes.2015.09.002.
[31] S.K. Sahu, D. Ghosh, Spectral kurtosis-based fault detection for a highly penetrated distributed generation: a real-time analysis. Advances in Smart Grid Automation and Industry 4.0, Springer, Singapore, 2021, pp. 649–656, https://doi.org/10.1007/978-981-15-7675-1_65. M. M Jaya Bharata, Reddy, Dusmanta Kr.
[32] J. Antoni, The spectral kurtosis of nonstationary signals: formalisation, some properties, and application, in: 12th European Signal Processing Conference, 2004, pp. 1167–1170.
[33] R. Dwyer, Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals, IEEE J. Ocean. Eng. 9 (2) (1984) 85–92, https://doi.org/10.1109/JOE.1984.1145602. Apr.
[34] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, Mech. Syst. Signal Process. 20 (2) (2006) 282–307, https://doi.org/10.1016/j.ymssp.2004.09.001. Feb.
[35] G. Mousmoulis, C. Yiakopoulos, G. Aggidis, I. Antoniadis, I. Anagnostopoulos, Application of spectral kurtosis on vibration signals for the detection of cavitation in centrifugal pumps, Appl. Acoust. 182 (2021), 108289, https://doi.org/10.1016/J.APACOUST.2021.108289. Nov.
[36] N. Cotuk, A.H. Kayran, A. Aytun, K. Savci, Detection of marine noise radars with spectral kurtosis method, IEEE Aerosp. Electron. Syst. Mag. 35 (9) (2020) 22–31, https://doi.org/10.1109/MAES.2020.2990590. Sep.
[37] V. Vrabie, P. Granjon, and C. Serviere, “Spectral kurtosis: from definition to application,” p. xx, 2003.
[38] S. Dutta, S. Olla, P.K. Sadhu, A secured, reliable and accurate unplanned island detection method in a renewable energy based microgrid, Eng. Sci. Technol. Int. J. 24 (5) (2021) 1102–1115, https://doi.org/10.1016/j.jestch.2021.01.015. Oct.