A simple decision tree-based disturbance monitoring system for VSC-based HVDC transmission link integrating a DFIG wind farm

Rajesh Babu Damala1,2, Rajesh Kumar Patnaik2* and Ashish Ranjan Dash3

Abstract
Fault detection and classification is a key challenge for the protection of High Voltage DC (HVDC) transmission lines. In this paper, the Teager–Kaiser Energy Operator (TKEO) algorithm associated with a decision tree-based fault classifier is proposed to detect and classify various DC faults. The Change Identification Filter is applied to the average and differential current components, to detect the first instant of fault occurrence (above threshold) and register a Change Identified Point (CIP). Further, if a CIP is registered for a positive or negative line, only three samples of currents (i.e., CIP and each side of CIP) are sent to the proposed TKEO algorithm, which produces their respective 8 indices through which the fault can be detected along with its classification. The new approach enables quicker detection allowing utility grids to be restored as soon as possible. This novel approach also reduces computing complexity and the time required to identify faults with classification. The importance and accuracy of the proposed scheme are also thoroughly tested and compared with other methods for various faults on HVDC transmission lines.

Keywords: Change Identification Filter, Differential current, DC faults, Simple Decision Tree, Fault classifier, HVDC transmission link, Renewable Energy, TKEO algorithm

1 Introduction
Almost all industrial processes, as well as various aspects of daily life, rely on electrical energy or electricity [1, 2]. Electricity consumption is ever increasing, and in particular energy demands are even higher during peak hours, making it difficult to guarantee supply to consumers [1]. The adoption of distributed energy resources (DER), such as wind, solar, and fuel cells, has been proved to be a realistic alternative given several concerns, including rising energy consumption, exhaustion of conventional energy resources (such as fossil fuels and coal) and pollution [3]. Addressing the above problem, integrated wind farms have been proposed. HVDC transmission systems outperform HVAC transmission systems in high power ratings [4]. The losses in HVAC transmission lines increase as transmission distance increases because of increased resistance, inductance and capacitance [5], and thus the transmission efficiency is reduced for long transmission length while the skin effect and corona loss are also observed in HVAC [6, 7].

The use of an HVDC transmission system addresses the above mentioned loss issue. One of the most critical issues in an HVDC transmission system is fault identification and classification [3]. The entire power system could fail if the fault current on the HVDC transmission link is not interrupted for an extended period. It is challenging to distinguish the faulty system from the healthy components if proper methodology is not used [8, 9]. To restore system stability and limit economic losses, the type of fault and its classification on the transmission line should be determined as soon as possible. The purpose of this study is to detect and examine the four types of faults that can occur on HVDC
transmission lines and to evaluate the robustness of the Teager–Kaiser Energy Operator (TKEO) algorithm with a Simple Decision Tree-based mechanism for accurate results with low computing complexity and a reduced time for fault identification, and classification.

2 Literature review
For HVDC transmission systems, many fault detection and classification approaches have been proposed. However, because of the aforementioned problems, techniques for protecting HVDC transmission lines are more limited than methods for conventional transmission systems [10]. A detailed literature review is provided here to have a better understanding of the proposed fault detection and classification methods for HVDC transmission lines. From the survey, a gap in the available fault detection and classification systems for HVDC transmission lines is identified. As fault detection and classification methods in HVDC transmission lines are influenced by a variety of parameters, these factors are investigated from several perspectives, with each one being examined separately. To ensure a fair review, the methods are divided into two categories, i.e. model-based and data-driven-based strategies.

2.1 Group A: data-driven-based techniques
Examining data pertinent to a system or determining the relationship between input and output state variables are the roots of data-driven approaches [11]. Because of the complexity of and necessity for a large quantity of data, real-time protections based on these technologies are not commonly used in HVDC transmission lines [12]. However, because of a lack of deep knowledge of the system, these methods are sometimes adopted to detect abnormalities that model-based methods may not be able to detect.

2.1.1 Fuzzy-based techniques
For fault identification in an HVDC transmission line, a combination of wavelet singular entropy and fuzzy logic is described in [13]. Similarly, in [14, 15], differential protection techniques based on fuzzy inference processors are proposed. However, the following are some of the challenges associated with the fuzzy method in fault detection and classification:
(i) Finding accurate membership functions and fuzzy rules is difficult;
(ii) To evaluate and validate the fuzzy-based system, extensive hardware testing is required.

2.1.2 Decision tree and ANN-based techniques
To detect the faults, reference [16] employs local current measurements with wavelet transform and a decision tree. In addition, for fault classification, a sequence analyzer is employed to extract negative and zero sequence components. For HVDC transmission lines, a data-mining-based technique on two decision trees is described in [17], and artificial neural networks (ANNs) are used to detect and classify faults in HVDC transmission lines in [18]. The current signal is sent to two distinct ANNs that have been trained to detect and classify faults in [19]. However, the following are some of the potential drawbacks of using these methods:
(i) Extensive data are necessary for the training stage;
(ii) There may be inadequate (or missing) training data to derive estimates in the majority of cases.

2.2 Group B: model-based techniques
Model-based approaches aim to determine if the evaluated variables are consistent with the model, in [20]. These methods can be further categorized based on the detection method used to identify the fault in various modes of operation [21].

2.2.1 Differential-based techniques
In [22], a differential protection strategy is proposed, in which two time–frequency transformations, i.e., Hilbert_Huang and the S-Transform, are compared to calculate the difference in the spectral energy content of modified contours on two sides of a feeder. Using the average cumulative sum and transient estimation methods, a differential transient current-based fault detection method for HVDC transmission lines is proposed in [23]. The following are the primary issues that differential-based techniques face:
(i) In the event of a communication breakdown, backup protection is needed to protect the HVDC transmission line;
(ii) The system cost increase as a result of the communication systems.

2.2.2 Local variable-based methods
In [24], the loop type HVDC transmission lines are protected using the inherent characteristics of the local variable current and its derivation. The inverter output current is used as a local variable to calculate the recursive least squares and mathematical morphology (MM). However, the local variable-based approaches, in general, have some drawbacks:
(i) Failure of the protection used in the HVDC transmission line can cause changes in local factors;
(ii) They are highly dependent on the HVDC transmission line design, where most techniques are developed and proven for specific HVDC transmission lines;
(iii) Performance may suffer the consequences of time delays, as the time required for fault identification is dependent on the fault type and variable magnitude at the defective feeder.

2.2.3 Adaptive methods

The updated mode of operation is checked at the relay point in adaptive-based approaches when the configuration changes. Current signals are obtained with current transformers (CTs) in [25] and compared using the cycle-by-cycle comparison method. The following are the key drawbacks for adaptive approaches:

(i) When the HVDC transmission system changes between different modes of operation, it is necessary to re-adjust the settings of the protection devices;
(ii) It is costly to use communication channels for setting updates and monitoring;
(iii) All feasible HVDC transmission configurations must be known prior to operation.

2.2.4 Traveling wave-based techniques

In [26], a traveling wave-based protection mechanism based on MM is proposed. Because MM technology only executes a few summations and subtractions, the introduced approach offers quick fault detection. However, high sampling-rate measuring equipment is required for traveling wave-based approaches. Although these procedures are quick and accurate, using high sampling-rate measurement instruments significantly reduces the benefits.

2.3 Aims and contribution

The aim of the proposed method is quicker detection and classification of the fault and to reduce computational complexity. A novel protection mechanism for HVDC transmission lines is proposed to detect and classify disturbances. When there is no fault, the differential current is zero, while it varies when there is a fault.

The method begins with the calculation of differential and average currents when CIP is identified. The numerical values at different faults are determined using a TKEO-based scheme in the next stage. The method is based on the “Teager Energy” tracked by the TKEO algorithm. The 8 indices are extracted from “Teager Energy”. In addition, TKEO calculates only three samples of data (at CIP and either side of CIP), resulting in a low computing burden and good time resolution. The next step is to generate eight separate indices based on “Teager energy” of differential and average currents to distinguish faulty from healthy sections, as well as the type of fault and the faulty line. Simulation and experimental systems are used to test the proposed method. The following are the main contributions of the proposed method:

(i) Processing only the current signal with simple rules;
(ii) Low computing burden and cost efficiency because no communication lines are required;
(iii) When evaluating the procedure, it takes into account a variety of challenging conditions.

3 Research gaps identified

Following a review of the literature, the following research gaps for fault analysis of an HVDC transmission system are identified:

• Distributed energy sources (DER) are the future of electricity production, but research on how to transmit this energy through HVDC lines with different design models is limited.
• A few studies used raw signal data processing units, which take much longer to analyse the fault classification and fault detection.
• According to several studies, the fault detection efficiency is limited to around 83%. Data is generated by a limited number of algorithms, and this may mislead the entire system.
• Several studies have concentrated on short distance HVDC lines only.
• The majority of studies fail to account for the computational burden and time required to diagnose the fault with classification.
• For fault analysis, very few little research has concentrated on the noise interface defective signal issue.

3.1 Novelties of the paper

The above-mentioned research gaps must be addressed to acquire the exact identification of faults as well as the classification of the HVDC transmission system. The gaps and limits described above must be addressed to conduct a realistic fault analysis of an HVDC transmission line. In contrast, the technique proposed can allow quick identification of the type of fault corresponding with classification in HVDC transmission link, as:
The required time to detect the fault and its classification is only about 10 ms.
The efficiency is improved to 98.75\% in terms of precise fault type and classification.
For the first time in an HVDC transmission system, the study uses the TKEO method in combination with a Simple Decision Tree-based fault classifier for fault analysis.
The proposed strategy overcomes the disadvantages in existing methods, such as computational complexity and requiring a long time to find the fault with classification.
When performing a fluctuating DC analysis, determining the magnitude of the threshold setting value is quite challenging.

4 Description of the designed model

Figure 1 depicts a single line diagram of a bipolar HVDC transmission system using a Voltage Source Converter (VSC) that is fed by a combination of offshore wind farms.

As can be seen 150 km HVDC transmission lines are presented. The fault analysis performed on the 150 km HVDC transmission line is the topic of this research. The HVDC system is fed by a wind farm of four units where each unit has a capacity of 9 MW. Each unit has 6 sub units of capacity 1.5 MW i.e. 6 \times 1.5 = 9 MW. So the total generating capacity is 36 MW i.e., 4 \times 9 = 36 MW. During simulation, the wind speed is kept constant at 15 m/s. A 30 km transmission line (TL1) with a 47 MVA step up transformer (T1) of 120 kV/25 kV transmits from the offshore to on shore. A 150 km transmission line (TL2) with capacity of 200 MVA ±100 kV connects the two converter stations with two 8 mH smoothing reactors. The AC voltage from the HVDC inverter is connected at bus B4 and to the utility via transformer T3.

4.1 Various faults on HVDC Transmission link

The four types of faults that can occur on the HVDC lines are:

1. Fault between the ground and the positive line (PG), shown in Fig. 2.

Fig. 1 Single line diagram of VSC based HVDC system fed by the integrated offshore wind farm

Fig. 2 Short circuit fault between positive and ground (PG) on an HVDC system
(2) Fault between the ground and the negative line (NG), shown in Fig. 3.
(3) Fault between the negative and positive lines (PN), shown in Fig. 4.
(4) Fault between the ground, the negative and positive lines (PNG), shown in Fig. 5.

In Figs. 6, 7, 8 and 9 depict the simulation results of the DC differential currents ($I_{dih}$) for the PG, PN, NG,
and PNG faults. During the simulation, the faults occur between 0.5 and 0.6 s at 25 km of the HVDC line.

In Figs. 6, 7, 8 and 9, simulation run time (T) is shown on the X-axis and the DC differential current (I_{diff}) on the Y-axis. As can be seen, the four faults result in different magnitudes of the DC differential currents at the same location at 25 km. The TKEO algorithm generates different “Teager Energies” from which 8 indices are generated. All four types of faults are simulated for each kilometer, and the resulting graphs and numerical data are saved. For simple presentation, only one graph of each fault at the same location 25 km is given.

The initialization of the proposed system is depicted in Fig. 10. The AC voltage from the wind generator is converted to a DC voltage to check that the designed model is in a stable condition. The converted wind farm DC voltage is not same as the DC voltage of transmission line. As can be seen, the DC output steady-state is reached after 1 s. Only once the designed system has reached a stable condition can fault analysis be performed. The designed system DC output voltage wave is not stable initially, because of the transient behavior which can last up to 1 s. The values of voltage and currents are in per unit.

Figure 11 shows the three-phase current and voltage waveforms before connecting the proposed system to the HVDC rectifier. The values of voltage and currents are in per unit. As seen, the voltage waveform becomes stable after 0.15 s while the current is stable after 0.2 s. The HVDC rectifier is only enabled after the voltage and current become stable.

Figure 12 shows the three-phase current and voltage before enabling the inverter. The voltage is stable from the beginning while the current waveform becomes stable after 0.15 s. The values of voltage and currents are in per unit.
Figure 13 shows the three-phase currents and voltages at the rectifier when the HVDC line is subjected to a fault. During the period of 0.5–0.6 s, a fault is created on the HVDC transmission line. For the duration of the fault, the differential current spikes can be seen at around 0.5–0.55. After a transient period, the current is stable at 0.05 s. The values of voltage and currents are in per unit.

Figure 14 shows the three-phase currents and voltages at the faulty inverter. For the duration of the fault the differential current spikes can be seen. Due to the transient effect (threshold value), the waveform is not stable at first, i.e., the current wave settles down after 0.05 s. The values of voltage and currents are in per unit.

5 Proposed methodology
The whole working methodology can be understood simply from Fig. 15.

The whole working method can be understood from Fig. 15. The DC differential current \( (I_1 - I_2) \) magnitude is almost zero at unfaulty conditions throughout the DC line. However, the magnitude increases \([I_1 - (-I_2)]\) rapidly when there is a fault on the HVDC line at any distance. The methodology contains 5 steps as follows:

1. Step 1 Whenever a fault occurs, the CIF technique is applied to HVDC transmission lines to detect the change in current wave form and registers CIP.
2. Step 2 Extracting “Teager Energy” either from differential or average currents at CIP at any distance on HVDC transmission line is the primary task for fault analysis of non-linear and non-stationary signals.
3. Step 3 TKEO, which tracks the “Teager Energy” of the respective signal at CIP with high time resolution is very efficient in terms of processing time as only three current samples (CIP and either side of CIP) are used for fault analysis.
4. Step 4 By using the extracted “Teager Energy”, the individual signal processing units of the 8 indices directly generate their numerical data, that is P1 (energy), P2 (amplified energy), P3 (mean), P4 (standard deviation), P5 (kurtosis), P6 (entropy), P7 (variance), and P8 from these differential and aver-

Fig. 12 Voltage and current before enabling the inverter under unfaulty condition

Fig. 13 Voltage and current at rectifier in faulty condition
current signals (maximum amplitude). All 8 indices have their own signal processing units.

**Step 5** In the final step, a Simple Decision Tree-based fault classifier is used. This allows the numerical data sets of the 8 indices to pass through it for fault detection and classification.

### 5.1 Computation of average and differential current

The averaged and differential current inputs are the two important components that differential current relays require for functioning and supervision in a differential protection scheme. The HVDC instant average and differential currents are given as:

\[
I_{\text{avg}}(f) = \frac{1}{2} \times (I_1(f) + I_2(f))
\]  

(1)

\[
I_{\text{diff}}(f) = (I_1(f) - I_2(f))
\]  

(2)

where \( f \) is the sampling instant.

### 5.2 Proposed algorithm

The proposed TKEO optimization technique avoids the shortcomings highlighted in the identified research gaps we cited previously, such as reducing computing complexity and the time required for fault identification and classification. This is because only three samples (CIP and either side of CIP) are being processed. TKEO is more sensitive to fluctuations in the signals under investigation. In this case, TKEO with the Simple Decision Tree-based classifier is a better alternative than other methods since it gives higher resolution and reduces the time taken to find faults. An "energy" tracing operator was invented by Teager and used by Kaiser to extract indices such as mean, energy, amplified energy, maximum amplitude, standard deviation, kurtosis, entropy, and variance from nonlinear signals. When compared to other commonly used algorithms, the TKEO algorithm outperforms them with high accuracy in fault detection and classification. The TKEO algorithm has formerly been employed in speech signal processing systems, but this is the first time it has shown promise in a non-stationary and nonlinear signal engineering application. Under any condition, TKEO is a simple method that is temporally localized, easy to compute, and capable of correctly monitoring the signal’s instantaneous changes in amplitude with respect to time. In terms of fault detection and categorization, the existing methods are quite lacking. In addition, the conventional methods process a large amount of data and this takes longer and increases the computational burden for fault analysis.
Change Identification Filter (CIF) working process:
Assume \( Z \) is the total number of samples for a signal \( x \) (either \( \text{avg} \) or \( \text{diff} \)) that has been sampled with \( S \) equal to 0 which is the initial sample. As a result, the CIF of the signal \( x \) may be described mathematically as:

\[
\text{CIF}_x(j) = \text{CIF}_x(j-1) + \sum_{m=0}^{Z} [x(m) - x(m - S)]
\]

(3)

where \( j \) is the iteration number, \( m \) is the sample number starting from the 1st sample, i.e., \( S + 1 \), while the initial \( \text{CIF}_x \) i.e., \( \text{CIF}_x(0) = 0 \). To identify the change on any HVDC line (positive or negative), the CIF formulation is applied to the current signal.

The TKEO algorithm procedure is explained as:

The numerical value of TKEO can be calculated using only 3 samples of the signal (CIP and either side of CIP).

The discrete energy of signal \( h(l) \) can be calculated as:

\[
\Psi[h(l)] = \left(\sum_{l=0}^{n} x(l)^2 \right) - h(l - 1) \times h(l + 1)
\]

(4)

where \( h(l-1) \) is the delayed sample and \( h(l+1) \) is the advanced sample of \( h(l) \).

\[
y(l) = h(l) - h(l - 1)
\]

(5)

\[
E(l) = 1 - \frac{\Psi[y(l)] - \Psi[y(l + 1)]}{4\Psi[h(l)]}
\]

(6)

The time can be calculated as:

\[
h_t(l) = \cos^{-1}[E(l)]
\]

(7)

The instantaneous amplitude is given by

\[
|A_i(l)| = \sqrt{\frac{\Psi[h(l)]}{1 - E^2(l)}}
\]

(8)

Indices Extraction from “Teager Energy”:
Consider a signal \( K(x) \), which comprises \( x \) samples and \( x = 1, 2 \ldots n \) then.

\( \text{Energy (P1)} \): The energy of the aforementioned samples, which is defined as the sum of the square of the sample, can be computed as:

\[
P_1 = \sum_{x=1}^{n} K_{(x)}^2
\]

(9)

\( \text{Amplified energy (P2)} \): It is defined as the sum of the sample’s product \( x \) and its square \( K(x) \), and can be calculated as:

\[
P_2 = \sum_{x=1}^{n} x \times K_{(x)}^2
\]

(10)

\( \text{Mean (P3)} \): The ratio of the sum of observations to the number of current signal samples \( x \), as:

\[
P_3 = \frac{\sum_{x=1}^{n} K_{(x)}}{x}
\]

(11)

\( \text{Standard Deviation (P4)} \): It is the difference between readings acquired from repeated measurements. It is also a way to quantify the variance or scatter of data set values, which may be determined by:

\[
P_4 = \left(\frac{1}{x-1} \sum_{x=1}^{n} |K_{(x)} - P_3|^2 \right)
\]

(12)

\( \text{Kurtosis (P5)} \) It is defined as a measure of the random variable’s tiredness. The central moment is defined as the moment of the mean of a random variable \( P_3 \), which can be calculated as:

\[
P_5 = \left(\frac{P_3^4}{P_4^4}\right)
\]

(13)

where \( P_4 \) is standard deviation, \( P_3 \) is mean.

\( \text{Entropy (P6)} \): It is a measure of a random variable’s randomness that can be calculated as:

\[
P_6 = \sum_{x=1}^{n} K_{(x)} \log_i K_{(x)}
\]

(14)

\( \text{Variance (P7)} \): It is defined as a measure as to how far a set of random numbers deviates from their mean value, which can be calculated as.

\[
P_7 = \sum_{x=1}^{n} \frac{(K_{(x)} - P_3)^2}{x}
\]

(15)

\( \text{Maximum Amplitude (P8)} \): It is described as a measure of the maximum amplitude value as:

\[
P_8 = \max(\text{signal})
\]

(16)

### 5.3 Simple Decision Tree-based fault classifier

Equations (1) and (2) are used to determine the differential and average current signals at CIP. The performance indices are calculated using the TKEO energy from the differential current signals. For fault type and identification, the proposed approach allows the indices described above to pass through an event classifier. The proposed method, as shown in the flow chart in Fig. 16, is used to handle the eight indices stated above. On a VSC-based HVDC bipolar transmission system, the Decision Tree-based fault classifier can quickly identify the fault type and its classification. The fault classifier’s decision-making process is described in detail in the tree diagram.
6 Results and discussion
MATLAB/Simulink is employed in this investigation and the simulation runtime for the system is one second. The following faults are specified on an HVDC line: PG fault, NG fault, PN fault, and PNG fault. The HVDC transmission link is 150 km long made up of two pi sections. At CIP, the energy is calculated using the Teager–Kaiser algorithm with three samples.

The targetted 8 indices P1–P8 are calculated using "Teager Energy", and are processed for fault detection and classification using a Simple Decision Tree-based fault classifier. The numerical data sets of all 8 indices are calculated and tabulated in Tables 1, 2, 3 and 4, which illustrate the ranges (maximum and minimum) of P1–P8 for all four fault types.

Tables 1, 2, 3 and 4 illustrate the ranges (maximum and minimum) of the indices: energy (P1), amplified energy (P2), mean (P3), standard deviation (P4), kurtosis (P5),...
entropy (P6), variance (P7), and maximum amplitude (P8) for all four fault types under faulty conditions. Tables 5, 6, 7 and 8 show the indices of the data from PG, NG, PN, and PNG faults, respectively. Each table contains 298 (149×2) data sets, one for each kilometer of the 150 km HVDC link. The two currents (differential and average currents) are used to tabulate each index data set at the faulty condition, each comprising 596×8 feature data sets. For the system, a total of (596×8)×2 feature data sets have been produced. The indices data sets for the four faults at fault positions at 30 km, 60 km, 90 km, 120 km, and 149 km are provided in the tables for ease of understanding. Tables 5, 6, 7 and 8 clearly indicate that the ranges of numerical values produced at various distances are associated with various types of indices, allowing for accurate fault detection and fault categorization.

Tables 3 and 4 show the ranges of the data from PN and PNG faults, respectively. Each table contains 298 (149×2) data sets, one for each kilometer of the 150 km HVDC link. The two currents (differential and average currents) are used to tabulate each index data set at the faulty condition, each comprising 596×8 feature data sets. For the system, a total of (596×8)×2 feature data sets have been produced. The indices data sets for the four faults at fault positions at 30 km, 60 km, 90 km, 120 km, and 149 km are provided in the tables for ease of understanding. Tables 3 and 4 clearly indicate that the ranges of numerical values produced at various distances are associated with various types of indices, allowing for accurate fault detection and fault categorization.

For a better understanding of the proposed strategy, a few example distances and their respective values for all the eight indices (P1–P8) are provided. The fault detection and its classification can be accurately determined since different attributes have unique values. Every individual index has its own unique number that does not overlap with other indices, indicating that the fault detection and classification can be determined accurately more quickly.

6.1 Performance evaluation of the proposed strategy
The performance of the proposed classifier is evaluated using indices generated from the “Teager Energy” using the TKEO method. The Simple Decision Tree-based fault classifier framework has a training and testing mechanism. During the training phase, initial parameters are optimized, and these values are then tested. From a total of 1192 (298×4) differential current data samples (I_{diff}), 477 (40%) are chosen at random throughout the testing phase. The fault-finding efficiency of the proposed approach is 98.75%. The efficiencies of the proposed approach are displayed in Table 9, which are calculated as:

$$\eta\% = \frac{\text{Number of rightly classified data samples}}{\text{Randomly pickdup samples from total group set}} \times 100$$  (17)

As previously described, the total fault samples of differential current (I_{diff}) are 298×4=1192, of which 60% (715) are trained for system training purposes and the rest 40% (477) are examined to check the efficiency and accuracy of the detection of the fault and fault type. The remaining 1192 average samples are ignored to reduce computational burden for fault analysis. 1192 + 1192 = 2384 are the total data sets generated for the 8 indices using the differential and average currents at CIP at each km.

The procedure for calculating efficiency is as follows:

- **Step 1**: Assign A = 0 if the test sample size is B and the number of correctly classified data sets is A.
- **Step 2**: A random number x, along with an initial guess, should be created. If x falls between data sets 1 and 298, the fault type is PG, and then u is set to 1. The fault is a PN type fault if x is between 299 and 596, and u is set to 2. The fault type is NG and u is set to 3, if x is between 597 and 894. Otherwise, the fault is of PNG type if x is between 895 and 1192, and u is set to 4.
- **Step 3**: The data sets ([P1, P2... P8]×8) are taken from the original data sets ([P1, P2... P8]596×8) for the testing data sets and are allowed to pass through the decision tree-based fault classifier to construct v, and the method works according to the following criteria for fault detection and classification:
Table 5  Generated values for all indices for a proposed system on the HVDC line for a PG fault

| Wind speed (m/s) | Fault resistance (Ω) | Fault location (km) | P1 Energy (per unit) | P2 Amplified energy (per unit) | P3 Mean (per unit) | P4 Standard deviation (per unit) | P5 Kurtosis (per unit) | P6 Entropy (per unit) | P7 Variance (per unit) | P8 Maximum amplitude (per unit) |
|-----------------|----------------------|---------------------|----------------------|-------------------------------|-------------------|--------------------------------|----------------------|---------------------|----------------------|-------------------------------|
| 15              | 0.001                | 30                  | 8.08E−07             | 5.986E−07                    | 0.166             | 0.00089                        | 0.508                | 0.402               | 3.67E−07             | 0.00089                        |
| 60              | 1.74E−11             | 3809E−11            | 0.095                | 4.172E−06                    | 0.621             | 0.392                          | 7.95E−12             | 5.14E−06            | 3.186E−05            | 0.0006                        |
| 90              | 1.533E−09            | 22.26E−09           | 0.0051               | 3.91E−05                     | 0.534             | 0.599                          | 1.05E−09             | 0.001               | 3.186E−05            | 0.0006                        |
| 120             | 4.508E−07            | 6.846E−07           | 0.1065               | 0.0006                        | 0.516             | 0.534                          | 2.967E−07            | 0.0006             | 3.186E−05            | 0.0006                        |
| 149             | 1.499E−06            | 1.499E−06           | 0.9148               | 0.0012                        | 0.516             | 0.659                          | 1.49E−06             | 0.001              | 3.186E−05            | 0.0006                        |
Table 6  Generated values for all indices for a proposed system on the HVDC line for an NG fault

| Wind speed (m/s) | Fault resistance (Ω) | Fault location (km) | P1 Energy (per unit) | P2 Amplified energy (per unit) | P3 Mean (per unit) | P4 Standard deviation (per unit) | P5 Kurtosis (per unit) | P6 Entropy (per unit) | P7 Variance (per unit) | P8 Maximum amplitude (per unit) |
|------------------|----------------------|---------------------|----------------------|-------------------------------|-------------------|----------------------------------|-----------------------|----------------------|-----------------------|--------------------------------|
| 15               | 0.001                | 30                  | 4.76E−13             | 1.03E−12                      | 0.0057            | 6.89E−07                         | 0.504                 | 0.436                | 2.24E−13              | 7.01E−07                      |
| 60               | 7.34E−11             | 30                  | 1.24E−11             | 2.46E−11                      | 0.016             | 8.56E−06                         | 0.515                 | 0.406                | 2.22E−10              | 8.59E−06                      |
| 90               | 1.2E−13              | 30                  | 3.62E−14             | 3.46E−07                      | 0.0228            | 4.27E−07                         | 0.301                 | 0.427                | 9.35E−13              | 3.07E−07                      |
| 120              | 1.6E−14              | 30                  | 2.94E−14             | 1.26E−07                      | 0.021             | 1.59E−05                         | 0.28                  | 0.568                | 1.06E−14              | 1.41E−07                      |
| 149              | 2.55E−10             | 30                  | 4.72E−11             | 0.285                         | 1.59E−05          | 0.508                            | 0.474                 | 1.4E−09               | 1.63E−05               | 7.01E−07                      |
| Wind speed (m/s) | Fault resistance (Ω) | Fault location (km) | P1 Energy (Per unit) | P2 Amplified energy (Per unit) | P3 Mean (Per unit) | P4 Standard deviation (Per unit) | P5 Kurtosis (Per unit) | P6 Entropy (Per unit) | P7 Variance (Per unit) | P8 Maximum amplitude (Per unit) |
|-----------------|----------------------|---------------------|----------------------|-------------------------------|-------------------|-------------------------------|---------------------|----------------------|----------------------|-------------------------------|
| 15              | 0.001                | 30                  | 8.85E−07             | 2.064E−06                     | 0.164             | 0.00094                       | 0.515               | 0.426                | 3.83E−07             | 0.00094                       |
| 60              | 1.81E−11             | 30                  | 3.97E−11             | 0.92                          | 4.26E−06          | 0.624                         | 0.457               | 8.31E−06             | 1.63E−09             | 5.26E−06                       |
| 90              | 2.37E−09             | 60                  | 3.44E−09             | 0.0051                        | 4.87E−05          | 0.528                         | 0.594               | 1.63E−09             | 4.17E−05             | 3.22E−07                       |
| 120             | 4.90E−07             | 60                  | 7.44E−07             | 0.1003                        | 0.0007            | 0.516                         | 0.504               | 3.22E−07             | 0.0007              |                               |
| 149             | 1.428E−06            | 90                  | 1.322E−06            | 0.389                         | 0.0011            | 0.516                         | 0.740               | 4.14E−06             | 0.0011              |                               |
Table 8  Generated values for all indices for a proposed system on the HVDC line for a PNG fault

| Wind speed (m/s) | Fault resistance (Ω) | Fault location (km) | P1 Energy (per unit) | P2 Amplified energy (per unit) | P3 Mean (per unit) | P4 Standard deviation (per unit) | P5 Kurtosis (per unit) | P6 Entropy (per unit) | P7 Variance (per unit) | P8 Maximum amplitude (per unit) |
|-----------------|----------------------|---------------------|----------------------|-------------------------------|-------------------|---------------------------------|----------------------|----------------------|----------------------|-----------------------------|
| 15              | 0.001                | 30                  | 8.51E−07             | 1.98E−06                     | 0.164             | 0.00092                         | 0.5152               | 0.426                | 3.68E−07             | 0.0009                      |
| 60              | 1.746E−11            | 30                  | 3.82E−11             | 0.094                         | 4.178E−06         | 0.622                           | 0.433                | 7.981E−12            | 5.153E−06             |                             |
| 90              | 1.568E−09            | 30                  | 2.77E−09             | 0.006                         | 3.96E−05          | 0.534                           | 0.606                | 1.08E−09             | 3.23E−05             |                             |
| 120             | 4.711E−07            | 30                  | 7.15E−07             | 0.114                         | 0.00068           | 0.516                           | 0.511                | 3.100E−07            | 0.00068              |                             |
| 149             | 1.50163E−06          | 30                  | 1.3905E−06           | 0.393                         | 0.0012            | 0.516                           | 0.697                | 1.62E−06             | 0.0012               |                             |
• If the algorithm determines that the fault is caused by a PG short circuit, set v to 1.
• If the program detects a short circuit between the negative pole and the positive pole (PN), v is set to 2.
• v is set to 3 if the program checks for a short circuit fault between ground and the negative pole.
• v is set to 4 if the program checks for short circuit between grounds, negative pole, and positive pole (PNG).

Step4 If u = v, then A = A + 1.
Step5 To detect the fault, the efficiency is calculated as: efficiency (percentage) = [(A B) 100].

The proposed approach is compared to existing methods in terms of efficiency to confirm that it provides improved protection efficiency, as illustrated in Table 10. The following conventional methods are considered:

1. An approach based on Park theory and a wavelet transform [28]. Converting line voltage or current signals into dq0 components and analyzing their behavior during faults to find patterns that signal the starting of a fault is part of the process. By filtering one of the dq0 components using the wavelet transformation and isolating band frequencies of interest, the finite difference between samples of the filtered signal can be used to detect faults.
2. The method based on mathematical morphology [26]. This method detects and classifies faults by applying the MM concept’s dilation and erosion median filters on current signal.
3. The method based on correlation concept [29]. To detect and classify the faults, this technique combines synchronized measured line currents and the correlation notion. The statistical cross-alienation coefficients for the measured current signals at the transmitting and receiving ends of each feeder are determined using this method. Changes in the synchronized and discretized waveforms of current signals inside a moveable window are taken into account in the fault detection and classification procedure (one-fourth cycle).
4. The method depended on reactive energy [30] to calculate the superimposed reactive energy (SRE). This technique employs the Hilbert transform. SRE is the integral of superimposed reactive power over a given time period. To identify faults in HVDC lines, several ratios are defined based on SRE.

The performance of the proposed strategy is compared in Table 11 to the aforementioned methods using the four criteria of accuracy, required average time, computational complexity and robustness to operate. Accuracy is defined as:

\[
\text{Accuracy \%} = \left[ 1 - \frac{\text{Number of incorrect discrimination}}{\text{Number of whole cases from total group set}} \right] \times 100
\]

(18)

Based on the results from Table 11, it is clear that the effectiveness of the proposed system has increased, and so has its computational complexity and classification efficiency.

Table 12 shows that the average time taken to identify the fault with different fault resistances is 10 ms highlighting the superiority over other methods.

7 Conclusions
To detect and classify power system faults on an HVDC transmission link, a novel “Teager–Kaiser Energy Operator” (TKEO) method which is combined with a Simple Decision Tree-based fault classifier has been investigated. The differential and average current components are subjected to a Change Identification Filter (CIF), which detects the first instant of fault (greater than the threshold value) incidence

| Serial Number | Fault Type | Randomly collected total data set | Number of random samples of data sets picked up for testing the proposed technique | Fault finding efficiency% (In percentage) |
|---------------|------------|----------------------------------|----------------------------------------------------------------------------------|------------------------------------------|
| 1             | PG         | 149 × 2 = 298                    | 477 (40% of total) random samples data sets from 1192 are taken for the testing purpose | 98.75                                    |
| 2             | PN         | 149 × 2 = 298                    |                                                                                  |                                          |
| 3             | NG         | 149 × 2 = 298                    |                                                                                  |                                          |
| 4             | PNG        | 149 × 2 = 298                    |                                                                                  |                                          |
| Total data sets |           | 298 × 4 = 1192                  |                                                                                  |                                          |
and registers a Change Identified Point (CIP). Further, if the CIP is registered either for a positive or negative line, only three samples of differential and average currents at CIP and each side of the CIP are sent through the proposed TKEO algorithm, which produces their respective 8 indices for fault detection and classification. The new method allows for quick detection of the fault type and classification. This cutting-edge technology increases fault identification efficiency while improving fault classification with greater accuracy. This approach also reduces the computing complexity and the average time required to identify faults is 10 ms as only three samples are required. The importance and significance of the proposed scheme have also been thoroughly tested and compared with some conventional methods for various faults on HVDC transmission lines. The outputs are satisfactory, demonstrating the real-time applicability of the proposed scheme. This can be useful for broad area protection.

Table 10  Comparison in terms of efficiency of the proposed method with existing methods

| Serial number | Methods                                      | Fault finding efficiency (%) |
|---------------|----------------------------------------------|------------------------------|
| 1             | Hilbert–Huang transform (HHT) method [27]    | 90.41                        |
| 2             | Artificial neural network (ANN) method [19]   | 95.90                        |
| 3             | The proposed method (TKEO with Decision Tree) | 98.75                        |

Table 11  Comparison of proposed with conventional methods in various factors

| parameter              | Method (Conventional) | Proposed method |
|------------------------|-----------------------|-----------------|
|                        | 1 2 3 4               |                 |
| Accuracy %             | 64.27%                | 69.96%          |
|                        | 82.27%                | 74.29%          |
| Required average       | 18                    | 23              |
| operating time(in ms)  | 23                    | 24              |
|                        | 15                    | 11              |
| Robust in operation    | No                    | No              |
|                        | No                    | No              |
| Computational burden   | High                  | High            |
|                        | High                  | Yes             |
|                        | High                  | Very low        |

Table 12  The time taken to detect the faults when all wind farms are turned ON

| Sl no | Fault type | Fault resistance (in ohms) | Relay trip timing (milliseconds) | Wind farm operation mode |
|-------|------------|----------------------------|---------------------------------|--------------------------|
| 1     | PG         | 0.1                        | 9.17                            | When all 4 wind farms are turned ON |
| PN    | 0.1        | 10.11                      |                                 |
| NG    | 0.1        | 9.15                       |                                 |
| PNG   | 0.1        | 10.11                      |                                 |
| 2     | PG         | 0.2                        | 10.24                           | When all 4 wind farms are turned ON |
| PN    | 0.2        | 9.25                       |                                 |
| NG    | 0.2        | 9.22                       |                                 |
| PNG   | 0.2        | 10.87                      |                                 |
| 3     | PG         | 0.5                        | 10.85                           | When all 4 wind farms are turned ON |
| PN    | 0.5        | 10.65                      |                                 |
| NG    | 0.5        | 10.37                      |                                 |
| PNG   | 0.5        | 9.52                       |                                 |
| 4     | PG         | 1                          | 10.99                           | When all 4 wind farms are turned ON |
| PN    | 1          | 9.57                       |                                 |
| NG    | 1          | 10.55                      |                                 |
| PNG   | 1          | 9.16                       |                                 |
### Index:

| S. No | Component       | Specific details of components                        |
|-------|----------------|-------------------------------------------------------|
| 1     | Wind turbine    | DFIG: 9 MW of unit of frequency 60 Hz and average speed of wind 15 m/sec [2] |
| 2     | PV Arrays       | Sun Power SPR 415-E WHT-M model with 7 series strings and 88 parallel strings |
| 3     | AC T/m link     | Length 30 km, frequency 60 Hz, capacitance $C_1$, $C_0$, $L_1$, $L_0$, resistance $R_1$, $R_0$, inductance $L_1$, $L_0$, 0.0851Ω per km, inductance $L_1$, $L_0$, 1.05e−3 3.32e−3 H/km |
| 4     | DC T/m link     | Length 150 km, PI Sects. 2, inductances/Unit length 1.5900e−004 Henry per km, capacitances/Unit length 2.3100e−007 Farad per km and Resistances/Unit length 1.3900e−003 Q/km |
| 5     | Phase reactor   | 0.15 p.u., inductance 0.15 x (100^2/200)/(Ω x pi x 60) mH and resistance 0.0015 x (100^2/200) mΩ |
| 6     | AC filter       | Frequency 60 Hz, Quality factor 15 and reactive component 40 MVAr |
| 7     | DC filter       | 12 micro farads |
| 8     | DC capacitor    | 500 micro farads |
| 9     | Smoothing reactor | Resistance 0.0851Ω, Inductances 25.0e−3 H |
| 10    | DC fault        | $R_0$, which is resistance of ground is 0.022(Ω) and Time of Switching 0.5 s [2] |
| 11    | Three level bridge IGBT/Diodes | Resistance of internal 1 millihm (Ω), Snubber circuit resistance 5000 Ω (Ω) and capacitance 1micro. F |
| 12    | VSC             | Voltage source converter |

### Abbreviations

DFIG: Double Fed Induction Generator; VSC: Voltage Source Converter; TKEO: Teager–Kaiser Energy Operator; IDC: Differential current; IAVRG: Average current; PG: Fault between ground and positive lines; NG: Fault between ground and negative lines; PN: Fault between negative and positive lines; PNG: Fault between ground, negative line and positive lines; T11: HVAC transmission line1; T12: HVDC transmission line2; T1: Transformer1; T2: Transformer2; T3: Transformer3; GT: Grounding transform; CIP: Change Identification Point; CIF: Change Identification Filter.

### Acknowledgements

We acknowledge the management of Centurion University of Technology and Management and GMR Institute of Technology for providing the opportunity to carry out the experiment and frame this paper.

### Author contributions

RBD carried out the overall work for the paper and drafted the whole paper. ARD has helped in drafting the paper. All authors read and approved the final manuscript.

### Funding

Availability of data and materials Data will not be shared as it is a Ph.D. research work.

### Declarations

#### Competing interests

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.

#### Author details

1. Department of EEE, Centurion University of Technology and Management, Paralakhemundi, Odisha, India & GMR Institute of Technology, Rajam, Andhra Pradesh, India. 2. Department of EEE, GMR Institute of Technology, Rajam, Srikakulam, Andhra Pradesh, India. 3. Department of EEE, Centurion University of Technology and Management, Paralakhemundi, Odisha, India.

### References

1. Ahmad, S., Ahmad, A., Naeem, M., Ejaz, W., & Kim, H. S. (2018). A compendium of performance metrics, pricing schemes, optimization objectives, and solution methodologies of demand side management for the smart grid. *Energies, 11*(10), 1–33. https://doi.org/10.3390/en11102801
2. Butt, I. Q., Kazmi, S. A. A., Ali Bukhari, S. B., & Ahmad, S. (2021). Passive islanding detection using square law method. In *2021 Int. Conf. Emerg. Power Technol.* (ICETP 2021), pp. 1–6. https://doi.org/10.1109/ICETPS706.2021.9435493
3. Dharmapandit, O., Patnaik, R. K., & Dash, P. K. (2017). A fast time-frequency response based differential spectral energy protection of AC microgrids including fault location. *Protection and Control of Modern Power Systems*. https://doi.org/10.1109/s41601-017-0062-0
4. Zou, G., Huang, Q., Song, S., Tong, B., & Gao, H. (2017). Novel transient-energy-based directional pilot protection method for HVDC line. *Protection and Control of Modern Power Systems*. https://doi.org/10.1186/s41601-017-0047-z
5. Xu, Y., Liu, J., & Fu, Y. (2018). Fault-line selection and fault-type recognition in DC systems based on graph theory. *Protection and Control of Modern Power Systems*. https://doi.org/10.1186/s41601-018-0098-9
6. Aziz, S., Peng, J., Wang, H., & Jiang, H. (2019). ADMM-based distributed optimization of hybrid MTDC-AC grid for determining smooth operation point. *IEEE Access*, 7, 74238–74247. https://doi.org/10.1109/ACCESS.2019.2919700
7. Muniappan, M. (2021). A comprehensive review of DC fault protection methods in HVDC transmission systems. *Protection and Control of Modern Power Systems*, 6(1), 1–20. https://doi.org/10.1186/s41601-020-00173-9
8. Jarrahi, M. A., Samet, H., & Ghanbari, T. (2020). Novel change detection and fault classification scheme for AC microgrids. *IEEE Systems Journal*, 14(3), 3987–3998. https://doi.org/10.1109/JSYST.2020.2966686
9. Liu, Z., Gao, H., Luo, S., Zhao, L., & Feng, Y. (2020). A fast boundary protection for an AC transmission line connected to an LCC-HVDC inverter station. *Protection and Control of Modern Power Systems*. https://doi.org/10.1186/s41601-020-00175-7
10. Jarrahi, M. A., & Samet, H. (2018). Modal current and cumulative sum based fault detection approach in transmission lines. *International Journal of Emerging Electric Power Systems*, 19(6), 1–24. https://doi.org/10.1515/ijeeps-2018-0037
11. Samet, H. (2016). Evaluation of digital metering methods used in protection and reactive power compensation of micro-grids. *Renewable and Sustainable Energy Reviews, 62*, 260–279. https://doi.org/10.1016/j.rser.2016.04.032
12. Damala, R. B., Misra, B. P., & Patnaik, R. K. (2019). High voltage DC transmission systems for offshore wind farms with different topologies.
13. 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. *International Journal of Electrical Power & Energy Systems*, 78, 455–462. https://doi.org/10.1016/j.ijepes.2015.11.048

14. Kar, S., & Ranjan-Samantaray, S. (2015). A fuzzy rule base approach for intelligent protection of microgrids. *Electric Power Components and Systems*, 43(18), 2082–2093. https://doi.org/10.1080/15325008.2015.1070584

15. Bhunia, S. B. A., Haider, R., Saeed-Uz-Zaman, M., Oh, Y. S., Cho, G. J., & Kim, C. H. (2018). An interval type-2 fuzzy logic based strategy for microgrid protection. *International Journal of Electrical Power & Energy Systems*, 98, 209–218. https://doi.org/10.1016/j.ijepes.2017.11.045

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

17. Kar, S., & Samantaray, S. R. (2016). Data-mining based comprehensive primary and backup protection scheme for micro-grid. In *2015 IEEE Power, Commun. Inf. Technol. Conf. PCITC 2015: Proc.*, pp. 505–510. doi: https://doi.org/10.1109/PCITC.2015.7438217

18. Panigrahi, B. K., Ray, P. K., Rout, P. K., Mohanty, A., & Pal, K. (2018). Detection and classification of faults in a microgrid using wavelet neural network. *Journal of Information and Optimization Sciences*, 39(1), 327–335. https://doi.org/10.1080/02522667.2017.1374738

19. Patnaik, R. K., Prasa, V., & Patnaik, R. K. (2019). An advanced signal processing based anfis classifier for fault analysis of dfig based hvdc transmission links. *Journal of Advanced Research in Dynamical and Control Systems*, 11(2), 1154–1167.

20. Hare, J., Shi, X., Gupta, S., & Bazzi, A. (2016). Fault diagnostics in smart micro-grids: A survey. *Renewable and Sustainable Energy Reviews*, 60, 1114–1124. https://doi.org/10.1016/j.rser.2016.01.122

21. Chakravorti, T., Patnaik, R. K., & Dash, P. K. (2017). Advanced signal processing techniques for multiclass disturbance detection and classification in microgrids. *IET Science, Measurement and Technology*, 11(4), 504–515. https://doi.org/10.1049/iet-smt.2016.0432

22. Gururani, A., Mohanty, S. R., & Mohanta, J. C. (2016). Microgrid protection using Hilbert-Huang transform based differential scheme. *IET Generation, Transmission and Distribution*, 10(15), 3707–3716. https://doi.org/10.1049/iet-gtd.2015.1563

23. Sadeghkhani, I., Golshan, M. E. H., Mehrizi-Sani, A., Guerrero, J. M., & Ketabi, A. (2018). Transient monitoring function-based fault detection for inverter-interfaced microgrids. *IEEE Transactions on Smart Grid*, 9(3), 2097–2107. https://doi.org/10.1109/TSG.2016.2606519

24. Gush, T., et al. (2018). Fault detection and location in a microgrid using mathematical morphology and recursive least square methods. *International Journal of Electrical Power & Energy Systems*, 102(February), 324–331. https://doi.org/10.1016/j.ijepes.2018.04.009

25. Muda, H., & Jena, P. (2017). Superimposed adaptive sequence current based microgrid protection: A new technique. *IEEE Transactions on Power Delivery*, 32(2), 757–767. https://doi.org/10.1109/TPWRD.2016.2601921

26. Li, X., Dy, A., & Burt, G. (2014). Traveling wave based protection scheme for inverter dominated microgrid using mathematical morphology. *IEEE Transactions on Smart Grid*, 5(5), 2211–2218.

27. J. Y., & Wang, H. (2018). A revised Hilbert-Huang transform and its application to fault diagnosis in a rotor system. *Sensors*, 18(12), 4239. https://doi.org/10.3390/s18124329

28. Escudero, R., Noel, J., Elizondo, J., & Kirtley, J. (2017). Microgrid fault detection based on wavelet transformation and Park's vector approach. *Electric Power Systems Research*, 152, 401–410. https://doi.org/10.1016/j.epsr.2017.07.028

29. Mahfouz, M. M. A., & El-sayed, M. A. H. (2016). Smart grid fault detection and classification with multi-distributed generation based on current signals approach. *IET Generation, Transmission & Distribution*, 10, 4040–4047. https://doi.org/10.1049/iet-gtd.2016.0364

30. Bukhari, S. B. A., Zaman, M. S. U., Haider, R., Yun-Sik, O., & Kim, C.-H. (2017). Electrical power and energy systems a protection scheme for microgrid with multiple distributed generations using superimposed reactive energy. *International Journal of Electrical Power & Energy Systems*, 92, 156–158.