Stockwell-transform and random-forest based double-terminal fault diagnosis method for offshore wind farm transmission line

Xiaodong Wang | Xing Gao | Yingming Liu | Yonghao Wang

Abstract
Due to the difficulty and time-consumption in locating short-distance transmission lines for deep-sea offshore wind farm (DOWF), this paper proposes a novel double-terminal fault location method by using Stockwell-transform (ST) and random-forest (RF). After the fault type and branch are accurately determined, the accurate transmission line fault location is located. Firstly, Stockwell-transform is employed to extract fault eigenvalues from the collected wind turbine (WT) current signals, which will reduce the sensitivity of eigenvalues to noise. And the Pearson correlation coefficient (PCC) is introduced to remove duplicate eigenvalues. Secondly, the reserved fault eigenvalues are taken as inputs to the different random-forest to classify fault types and identify the fault branch, respectively. Finally, the double-terminal fault location principle is established in fault negative sequence network (only ABCG uses positive sequence components). Newton-Raphson method (NRM) is used to eliminate the influence of asynchrony data, which implies an accurate transmission line fault location for deep-sea offshore wind farm. More than 4000 fault cases data obtained by Simulink simulation verify the feasibility and performance of the proposed method. The results show that the proposed location method has a high fault recognition rate and is immune to fault inception angle, resistance, location and noise.

1 | INTRODUCTION

Deep-sea wind energy is a rapidly growing renewable energy source due to the abundant reserves [1–3]. A wind farm is a multi-branch transmission system with complex topology and short-distance transmission lines. The transmission lines of onshore wind farms are mainly composed of overhead lines [4]. Manual inspection is of low difficulty and is easy to operate. However, deep-sea offshore wind farm (DOWF) is mainly composed of submarine cables [5]. Due to the bad sea conditions and geological environment, submarine cables are vulnerable to the damage of current and seabed organisms [6]. The manual detection cost of submarine cable is higher and time-consuming than that of the overhead line because of the bad maintenance environment. The traditional way is to replace the whole fault branch, which makes the maintenance cost high. When the accurate fault location is determined, only the transmission line near the fault location needs to be repaired, which is more conducive to reduce the cost of manual detection and shorten the operation and maintenance time. Frequent transmission line faults reduce the wind power generation and limit the development and utilization of deep-sea wind energy [7–9]. Therefore, the fast and accurate fault location of deep-sea offshore wind farm transmission line (DWTL) plays a very important role in the safe operation of the power grid.

In recent research, the main methods for diagnosing transmission line fault can be mainly classified into three classes: the impedance [10–12], traveling wave [13–15] and learning algorithm-based methods [16–21]. The impedance method generally uses the voltage and current information of single-terminal or double-terminal, which is used to extract fault impedance and establish fault criterion. However, the accuracy of location largely depends on fault type, fault inception angle (\( \Delta f \)) and fault resistance (\( R_f \)). Although the relationship between fault distance and extreme point of system impedance spectrum is proposed in [12] which solves the influence of \( R_f \), the analysis of \( A_f \) and multi-branch system is still not considered. The

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Renewable Power Generation* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.
traveling wave methods use the time difference between the first traveling wave and the first reflected wave or the first wave front of double-terminals to determine the fault location. Although traveling wave methods based on wavelet coefficients have been proposed to improve the accuracy of fault location in [13], a method based on homogeneous transmission lines is proposed in [14]. However, $A_f$, $R_f$ and measurement noise (MN) on the performance of the method is not considered, and the judgment basis of fault type and fault branch is not given. The fault branch identification and fault location are carried out successfully in [15]. However, the above traveling wave based methods can only locate the fault in long-distance transmission line effectively. Smart grid is the development trend in the future. Learning algorithm-based methods has been widely adopted in the field of fault diagnosis. The intelligent algorithms do not affect the action of relay protection methods, such as the circuit breaker. On the contrary, the combination with relay protection greatly improves the intelligence of power grid. Wavelet transform and different machine learning are often used to locate faults in transmission lines, including ANN, FLS, SVM, and ELM [16–19]. However, wavelet transform usually loses the phase information when it decomposes the fault signal, which leads to the lack of accuracy. Combining adaptive convolution neural network and double-terminal fault location is proposed in [20]. But $A_f$ is not considered. And the original method is unable to adapt to the fault location of DWTLs. Reference [21] combines ST and particle swarm optimization with different machine learning to identify fault branch and locate fault. However, the accuracy for fault branch identification of DOWFs is limited by short-distance transmission lines. Although the reference [22] has successfully located the fault of the transmission line of onshore wind farm, the bus is ignored and regarded as non-impedance. The DWTLs are deeply buried in the seabed, and the working environment is worse. As mentioned earlier, its failure rate is high. Therefore, the bus is also studied by the proposed method.

It can be seen from the above analysis that the main factors restricting diagnostic performance include fault inception angle ($A_f$), fault resistance ($R_f$), fault location ($L_f$), measurement noise (MN) and length of transmission lines. All of the above methods have only concentrated on transmission lines of the distribution grid with little attention to wind farms. Meanwhile, most of them cannot be directly applied to DOWF due to the following reasons:

- Comparing the length of the distribution network and wind farm lines, it can be seen that most wind farms are composed of short-distance transmission lines. The shortest transmission line length in distribution network research is not less than 6.7 km [21, 23, 24]. In the study of the wind farm, the length of transmission lines is about 1 km [22]. At the same time, the closer the fault point is, the more similar the fault signal is. When the fault occurs in two adjacent branches, the closer the fault location is to the common node, the more similar the fault signal characteristics are [25, 26]. Due to the short distance of transmission lines in wind farms, the fault signals are more similar, so the accuracy of the existing methods will be slightly reduced when they are directly applied to DOWF.
- The distribution network is mainly composed of power grid, transmission line, transformer and load. Wind farm is mainly composed of power grid, transmission line, transformer and doubly-fed wind turbine. Comparing the two systems, we can see that the biggest difference lies in the load and wind turbine. When the system fails, the crowbar and internal control strategy of the wind turbine can restrain the change of voltage and current. Therefore, the existing fault methods cannot be directly applied to the DOWF fault research [27, 28].

Therefore, a novel double-terminal fault location method based on STRF is proposed to classify fault type, identify fault branch and locate fault point aiming at short-distance DWTLs. And considerable consideration is given by the proposed method to the above shortcomings. The proposed method has no conflict with the traditional relay protection method, and will not affect the circuit breaker and other line protection methods. On the contrary, the combination of the proposed method and relay protection is more advantageous to the operation and maintenance of offshore wind farms, which is beneficial to increase the reliability of line fault protection and reduce the operation and maintenance time. The proposed method starts with the extraction of fault eigenvalues from faulty current signals, collected from different branches of the DOWF by using ST. Then fault eigenvalues are taken as the input of random-forest (RF) to classify DWTL fault types. After the fault type is determined, the fault eigenvalues are sent to the corresponding RF to identify the fault branch. On this basis, combined with Kirchhoff voltage law (KVL) and Newton-Raphson method (NRM), the fault location is carried out by using double-terminal fault location principle.

The main contribution is that a novel double-terminal fault location method based on STRF is proposed aiming at short-distance transmission lines for DOWF. The DWTL fault classification, branch identification and location accuracy are obviously improved. The influence of uncertainty $A_f$, $R_f$, $L_f$ is solved by combining ST with RF. ST itself has good noise resistance, which eliminates the influence of MN on fault eigenvalues. NRM is used to eliminate the asynchrony between the two terminals. The satisfied performance is kept even under different length of transmission line.

This paper is organized as follows: A novel fault location method based on STRF is proposed in Section 2. In Section 3, four DWTL fault characteristic matrices and the selection process of fault eigenvalues are given. In Section 4, the fault classification and branch identification method based on STRF the principle of double-terminal fault location by positive and negative sequence components are proposed. The effectiveness and competitiveness of the method is demonstrated by simulation results in Section 5. Section 6 concludes this paper.
PROPOSED DOUBLE-TERMINAL FAULT LOCATION METHOD BASED ON STRF

Flow chart of double-terminal fault location based on STRF for deep-sea offshore wind turbine transmission line is shown in Figure 1. The accuracy of fault classification is determined by the excellence of fault characteristics. And the high precision fault classification is the fundamental requirement of fault branch identification. Meanwhile, the necessary information for double-terminal fault location is provided by accurate fault branch identification. In this paper, a double-terminal fault location method based on STRF is proposed aiming at short-distance DWTLs for DOWF. The process of the proposed method can be divided into three parts: eigenvalues selection, fault classification and branch identification, and double-terminal fault location.

In the first part (eigenvalues selection): the WT current ($i_a$, $i_b$, and $i_c$) is dynamically measured by the proposed method. To select the useful eigenvalues, $i_a$, $i_b$, and $i_c$ are transformed by ST. And four fault eigenvalue matrices ($C_{stock}$, $S_{row}$, $S_{column}$ and $P_{stock}$) are produced after ST. The different statistical methods were applied to $C_{stock}$, $S_{row}$, $S_{column}$ and $P_{stock}$. And the fault eigenvalues are generated. The correlation between eigenvalues is obtained by Pearson correlation coefficient (PCC). The changing trend of fault current will be seriously damaged by the MN. The problem of MN can be solved by ST. Therefore, the accuracy under noise is guaranteed. The number and complexity of the fault data are reduced while the accuracy is not influenced by retaining the eigenvalues with weak correlation.

In the second part (fault classification and branch identification): the preserved fault classification eigenvalues are then fetched to the fault classification RF model to get the fault type of DWTLs. After getting the accurate fault type of DWTLs, 10 branch identification RF models corresponding to fault types are generated according to 10 different fault types. The different fault eigenvalues are then fetched to the fault branch identification RF model to identify the fault area based on the result of the fault type. High precision fault classification is the basis of branch identification, which effectively improves the accuracy of branch identification for DOWF.

In the third part (double-terminal fault location), after the first two parts, the fault types in offshore wind farms are accurately classified and the branches are accurately identified. Although fault classification, fault branch identification and fault location can be regarded as three independent parts, the result of fault classification is the basis of the high accuracy of fault branch identification. And the results of fault classification and branch identification serve the double-terminal fault location principle. The fault voltage and current from the double-terminal of the fault DWTL are extracted. The voltage and current components are converted by voltage and current. And then sequence components are fed to the fault location algorithm. The fault voltage equation is determined to locate the fault accurately by applying KVL to DWTLs. However, the asynchronous data will seriously affect the accuracy of fault location, so the synchronization angle $\xi$ is introduced based on the NRM.

3 | FAULT EIGENVALUES SELECTION BASED ON STOCKWELL-TRANSFORM AND PEARSON CORRELATION COEFFICIENT

3.1 | Stockwell-transform based fault matrices generation

Finding excellent fault type eigenvalues is the key to successful fault classification, resulting in high-accuracy. The change law of signal strength with time is the signal time-domain characteristic. And which single frequency signals are synthesized is the frequency domain characteristic. Time-domain signals can be transformed into the frequency domain by Fourier transform (FT) to analyse harmonics [29]. However, frequency resolution and time resolution cannot be taken into account simultaneously by both discrete, continuous and short-time Fourier transform [30, 31]. Better time-frequency resolution
cannot be provided by a fixed-width sampling window. The wavelet transform resolves the resolution problem. However, MN is inevitable in practice. The shortcoming of wavelet transform is the sensitivity to MN [32]. Meanwhile, wavelet transform will lose the phase information of non-stationary signals [32].

The main reason that affects the accuracy of the model is the extraction of fault eigenvalues. In machine learning, feature extraction is a very important part. In order to extract effective fault eigenvalues, the ST is introduced. The ST is a combination of STFT and wavelet transform, which retains the advantages of both STFT and wavelet transform [33]. The phase information is effectively maintained and the robustness to MN is improved by ST [27]. In the proposed method, the ST can be defined by

\[ S(\alpha, f) = \mathcal{S} \{ g(t) \} = \int_{-\infty}^{\infty} g(t) \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\alpha)^2}{2}} e^{-j2\pi ft} dt \]  

(1)

where \( f \) is frequency; \( \alpha \) and \( t \) are time variables; \( g(t) \) is input signal.

The Fourier spectrum is obtained by calculating the average value of the local spectrum with time [33]. \( G(f) \) is the Fourier transform of \( g(t) \), which can be written as:

\[ G(f) = \int_{-\infty}^{\infty} S(\alpha, f) d\alpha \]  

(2)

\( g(t) \) can be extracted from Equation (1), \( g(t) \) can be written as

\[ g(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} S(\alpha, f) d\alpha \right\} e^{j2\pi ft} df \]  

(3)

Then Equation (1) can be written as operations on the Fourier spectrum \( G(f) \) of \( g(t) \).

\[ S(\alpha, f) = \int_{-\infty}^{\infty} G(\mu + f) e^{-\frac{2\pi^2 \mu^2}{2}} e^{j2\pi \mu \alpha} d\mu \]  

(4)

\[ G(\mu + f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi (\mu + f)t} dt \]  

(5)

where \( f \neq 0 \). The discrete analogue of Equation (4) is used to compute the discrete ST.

In this paper, the fault current signal \( g[nT] \) is a discrete-time series, then the discrete Fourier transform is defined as [33, 34].

\[ G \left[ \frac{m}{MT} \right] = \frac{1}{M} \sum_{n=0}^{M-1} g[nT] e^{-j2\pi \frac{nm}{M}} \]  

(6)

where \( n, m = 0, 1, 2, \ldots, M-1, M \) is the number of samples; \( M \) is the number of samples; \( T \) is the sampling period, which is 0.0001.

### Table 1: The implication and size of fault eigenvalue matrices

| Name          | Implication                                                                 | Size    |
|---------------|----------------------------------------------------------------------------|---------|
| \( C_{stock} \) | The largest absolute value of each column of \( C \) matrix                | 1 \times 335 |
| \( S_{row} \)  | The largest absolute value of each row of \( ST \) matrix                  | 168 \times 1 |
| \( S_{column} \) | The largest absolute value of each column of \( ST \) matrix               | 1 \times 335 |
| \( P_{stock} \) | The phase angles corresponding to each element of \( S_{column} \) matrix   | 168 \times 1 |

By considering the Equations (4), (5) and (6), Stockwell transform of the fault current signal \( g[nT] \) is written as [33, 34]:

\[ S \left[ \frac{qT}{MT} \right] = \sum_{j=0}^{M-1} G \left[ \frac{j + \frac{m}{MT}}{MT} \right] e^{-\frac{2\pi q}{M} \frac{m}{MT}} \]  

(7)

where \( \alpha = qT; f = \frac{m}{MT}; q, m, y = 0, 1, 2, \ldots, M-1 \). Then the discrete ST is inverted, \( g[nT] \) can be written as:

\[ g[nT] = \sum_{j=0}^{M-1} \left\{ \frac{1}{M} \sum_{m=0}^{M-1} G \left[ \frac{j + \frac{m}{MT}}{MT} \right] e^{j2\pi \frac{nm}{M}} \right\} e^{j2\pi \frac{q}{M} \frac{m}{MT}} \]  

(8)

By considering the Equations (7) and (8), a fault matrix is generated by DWTL fault current through Stockwell transformation. The frequency and time information of the fault current is contained in the row and column of the fault matrix. The fault matrix is usually known as \( ST \) matrix which is a 168 \times 335 matrix.

The \( C \) matrix is indirect fault matrix, which is used to generate the fault characteristic matrix. The \( C \) matrix can be calculated by [27]:

\[ C_{168 \times 335} = |ST|_{168 \times 335}^2 \]  

(9)

In this paper, four fault eigenvalue matrices (\( C_{stock}, S_{row}, S_{column} \) and \( P_{stock} \)) are generated by \( C \) matrix or \( ST \) matrix. The implication and size of fault eigenvalue matrices are shown in Table 1.

### 3.2 Pearson correlation coefficient based fault eigenvalues selection

Different statistical methods are applied to the four fault eigenvalue matrices in the proposed method. The standard statistical methods such as standard deviations, maximum values, minimum value, and kurtosis are widely used to detect the distribution network ground fault [34, 35]. Kurtosis, mean values, amplitude and entropy are used for the analysis of power quality transients [36]. Seven statistical methods in the above method are applied to \( C_{stock} \), \( S_{row} \), \( S_{column} \) matrix and \( P_{stock} \) matrix in order to find the obvious fault...
TABLE 2  Unscreened eigenvalues by $C_{stock}$ matrix of DWTL faults

| Applied matrix | Statistical method | Eigenvalue |
|----------------|-------------------|------------|
| $C_{stock}$    | Kurtosis          | T1         |
| $C_{stock}$    | Mean value        | T2         |
| $C_{stock}$    | Minimum value     | T3         |
| $C_{stock}$    | Maximum value     | T4         |
| $C_{stock}$    | Standard deviation| T5         |
| $C_{stock}$    | Deviation         | T6         |
| $C_{stock}$    | Entropy           | T7         |

eigenvalues. The methods used include kurtosis, mean value, minimum value, maximum value, standard deviation, deviation and entropy. The generating process of unscreened eigenvalues by $C_{stock}$ matrix is shown in Table 2.

The generating process of unscreened eigenvalues by $S_{row}$ matrix, $S_{column}$ matrix and $P_{stock}$ matrix is the same as the $C_{stock}$ matrix. The fault eigenvalues of $S_{row}$, $S_{column}$ and $P_{stock}$ are numbered T8–T14, T15–T21 and T22–T28 respectively. Thus, 28 eigenvalues are generated from each phase after a DWTL fault occurs at the deep-sea wind farm. The number of eigenvalues is not the more the better, it will lead to useless calculation, and then affect the operation time of the model. So PCC is introduced in order to reduce the number of eigenvalues and maintain the accuracy. PCC is defined as:

$$\text{PCC} = \frac{\sum_{j=1}^{c} (T_{a_j} - \bar{T}_a)(T_{b_j} - \bar{T}_b)}{\sqrt{\sum_{j=1}^{c} (T_{a_j} - \bar{T}_a)^2 \sum_{j=1}^{c} (T_{b_j} - \bar{T}_b)^2}}$$

(10)

where $c$ is the length of the eigenvalue; $T_a$ and $T_b$ are two different fault eigenvalues of DWTL for DOWF; $\bar{T}_a = \sum_{j=1}^{c} T_{a_j} / c$; $\bar{T}_b = \sum_{j=1}^{c} T_{b_j} / c$.

PCC is the most commonly used linear correlation coefficient which can well reflect the correlation between the two fault eigenvalues [37]. -1 to +1 is the range of PCC. -1 indicates the complete negative correlation between the two fault eigenvalues, and +1 indicates the complete positive correlation between the two fault eigenvalues. 0 means there is no correlation between the two fault eigenvalues.

100 A-phase ground (AG) fault cases are randomly generated, resulting in a $28 \times 100$ fault matrix. The fault eigenvalues are fetched into Equation (12) in columns to calculate PCC. It is worth mentioning that some fault eigenvalues are constant (T1, T15 and T22) or zero (T7, T14, T21 and T28) in all DWTL fault cases. These unchanged fault eigenvalues are removed by the proposed method since they are meaningless for DWTL fault classification. PCC between all remaining eigenvalues is shown in Figure 2. The trend of partial fault eigenvalues in 100 fault cases is shown in Figure 3.

As shown in Figures 2 and 3, the performance of some fault eigenvalues is similar when a DWTL fault occurs at DOWF. The number of eigenvalues can be reduced without reducing the accuracy of DWTL fault classification by deleting these eigenvalues with high PCC coefficients. Table 3 summarizes the basis for the above Fault eigenvalues selection. The eigenvalues with PCC greater than 0.95 are deleted and only one of them is retained in the proposed method.
TABLE 3 The basis for the above fault eigenvalues selection

| Eigenvales | Action basis | Action |
|------------|--------------|--------|
| T1, T15, T22 | Constant value = 1.5 for all cases | Removed all |
| T7, T14, T21, T28 | Constant value = 0 for all cases | Removed all |
| T2, T3 | PCC = 0.9999 | Removed T3 |
| T2, T4 | PCC = 0.9999 | Removed T4 |
| T3, T19 | PCC = 0.9925 | Removed T19 |
| T6 | / | Reserved T6 |
| T8, T13 | PCC = 0.9929 | Removed T13 |
| T9 | / | Reserved T9 |
| T10 | / | Reserved T10 |
| T2, T11 | PCC = 0.9995 | Removed T11 |
| T12 | Correlation coefficient = 1.0000 | Reserved T12 |
| T2, T16 | PCC = 0.9995 | Removed T16 |
| T2, T17 | PCC = 0.9993 | Removed T17 |
| T2, T18 | PCC = 0.9995 | Removed T18 |
| T27 | / | Reserved T27 |

As can be concluded from Table 3, the final number of preserved eigenvalues is 11, including T2, T5, T6, T8, T9, T10, T12, T20, T23, T26 and T27. In this paper, the recorded two cycle current data of DWTL fault time, including the previous cycle and the latter cycle are sent into ST after a fault on DWTL. The ST matrix and four fault eigenvalue matrices (Cstock, Srow, Scolumn and Pstock) are generated by ST. The fault eigenvalues are calculated by Cstock, Srow, Scolumn and Pstock of each phase fault current data. A total of 3 × 11 = 33 eigenvalues are obtained at each deep-sea wind turbine side. By combining ST and PCC, not only the effective eigenvalues which can distinguish faults obviously are extracted, but also a large number of useless eigenvalues are reduced, which reduces unnecessary calculation without affecting the accuracy. Then, the fault eigenvalues are sent to the RF model for DWTL fault classification.

4 DOUBLE-TERMINAL FAULT LOCATION BASED ON RANDOM-FOREST

4.1 Random-forest based fault classification

Machine learning and statistics are closely related fields [38]. Statistical inference is an important foundation of machine learning. From methodological principles to theoretical tools, statistics has been throughout the development of machine learning [39]. The selection of a classification algorithm is the key point of fault classification model construction. Artificial neural network (ANN) and support vector machines (SVM) are the main methods in the field of classification [40]. ANN is to adjust the input and output weights and bias, so as to determine the mapping relationship between input and type [35]. However, many parameters need to be adjusted by ANN. And ANN is also easy to fall into local minima and cannot guarantee that the result is the global optimal solution. SVM is based on statistical theory and structural risk minimization principle [21]. The local minima and over fitting are solved by SVM. However, the adjustment space of the core is limited.

RF is one of the algorithms based on statistical thought summary in the development of machine learning [41]. RF is a parallel ensemble learning model, which uses decision tree (DT) as classifier [42]. RF is widely used in classification because of its strong generalization ability, difficult to over fit, simple model and good classification effect [42]. In the proposed method, the CART algorithm is adopted in the DT construction algorithm. CART algorithm is an improved algorithm based on ID3 and C4.5. CART has excellent anti-noise performance and is easy to understand and use, which is very suitable for real-time fault classification [27]. The Gini coefficient is the basis for CART algorithm to select fault eigenvalues. The Gini coefficient is defined by:

\[
Gini = 1 - \sum_{d=1}^{D} \frac{P(d)}{I} \frac{P(\bar{d})}{I} \tag{11}
\]

where \(D\) is the total number of samples; \(P(d)\) is the probability of randomly extracting one sample and belonging to class \(d\) when the test variable is \(I\).

The fault current signals of six wind turbines are extracted by the RF based fault classification method. The fault current signal of each wind turbine generates 33 fault eigenvalues, as discussed in Chapter 2. Therefore, a total of 33 × 6 = 198 eigenvalues are generated for fault classification of DWTLs in this paper, which is \(D = 198\) in Equation (11). Ten common types of DWTL faults are considered, including A-phase ground fault (AG), B-phase ground fault (BG), C-phase ground fault (CG), AB-phase ground fault (ABG), AC-phase ground fault (ACG), BC-phase ground fault (BCG), ABC-phase ground fault (ABC), A-B phase fault (AB), A-C phase fault (AC), B-C phase fault (BC). Therefore, \(I\) has 10 categories, which is \(I = 10\) in Equation (11).

The construction process of RF is to extract \(Z\) fault training sets with the same number from the original fault eigenvalue samples by bootstrap method. The RF model is composed of \(Z\) DTs model which generated by CART algorithm. The final classification result is obtained by voting on \(Z\) results. The number of \(Z\) is one of the factors that directly influence the RF accuracy. After a large number of experiments, RF can maintain good performance when \(Z\) is set to 10 in this paper.

In order to illustrate the effectiveness and robustness of the proposed fault eigenvalues in fault classification. Taking T2 and T6 as examples, the performance of the two fault eigenvalues under partial fault types is shown in Figure 4. As is shown in Figure 4:

- The three fault types (AG, BG and CG) can be clearly divided into two groups by fault eigenvalues T2. The first group includes AG and BG. And the second group includes CG.
The two fault types (BG and CG) can be clearly divided into two groups by fault eigenvalues T12. The first group includes BG. And the second group includes CG. Therefore, the proposed fault eigenvalues can be used for accurate fault classification through RF.

- As can be seen comparing the performance of eigenvalues in Figure 4 under different noise levels (0 dB and 20 dB). Although some eigenvalues will change significantly under the noise level is 20 dB, different fault types can still be clearly distinguished. Therefore, the influences of MN can be completely eliminated by the proposed RF based fault classification method.

- It is worth noting that because S-transform has good noise immunity, all fault eigenvalues perform well in most cases. However, as shown in Figure 4(b), due to the influence of noise, the fault characteristic value T6 inevitably changes slightly. The judgment of RF is a top-down decision-making, so once the fault eigenvalue T6 exceeds the interval value, it will lead to misjudgement.

4.2 Random-forest based fault branch identification

The DWTL of DOWF includes collecting lines and multiple wind turbine branch lines, which is a typical multi-branch structure. The accuracy of conventional fault branch identification method cannot meet the requirements for DOWF due to the multi-branch structure. Besides, the difficulty of fault branch identification is increased because the DWTLs are generally short.

Fault branch identification can also be seen as a classification problem. The fault current date of six wind turbines are recorded for fault branch identification for DOWF. 33 fault eigenvalues are generated by a wind turbine. Then $33 \times 6 = 198$ eigenvalues are generated for fault branch identification of DWTLs which is $D = 198$ in Equation (11).

The fault type of DWTLs is determined. One RF fault branch identification model is established for each fault type in this paper. Then, the fault eigenvalues are sent to the corresponding RF to identify the fault branch of DWTLs according to different fault types. In order to illustrate the advantages of the proposed method, a total of 10 branches are selected to distinguish the fault branch, which is $I = 10$ in Equation (11). After a large number of experiments, RF in fault branch identification can maintain good performance when $Z$ is set to 50 in this paper.

In order to illustrate the effectiveness and robustness of the proposed fault eigenvalues in fault branch identification. Taking T4 and T20 as examples, the performance of the two fault eigenvalues under partial fault branches is shown in Figure 5.
is shown in Figure 5:

- The four fault branches (L1, L5, L6 and L7) can be clearly divided into two groups by fault eigenvalues T4. The first group includes L1 and L7. And the second group includes L5 and L6. The two fault branches (L2 and L6) can be clearly divided into two groups by fault eigenvalues T11. The first group is L6. And the second group is L2. Therefore, the proposed fault eigenvalues can be used for accurate fault branch identification through RF.
- As can be seen comparing the performance of eigenvalues in Figure 5 under different noise levels (0 dB and 20 dB). Although some eigenvalues will change significantly under the noise level is 20 dB, different fault branches can still be clearly distinguished. Therefore, the influences of MN can be completely eliminated by the proposed RF based fault branch identification method.
- Like the problem of fault classification, because S-transform has good noise immunity, all fault eigenvalues perform well in most cases. But once the fault eigenvalue T6 exceeds the interval value, it will lead to misjudgement of fault branch.

4.3 Double-terminal fault location

Rapid and accurate fault location after DWTLs fault can reduce the occurrence of wind power abandonment, which greatly improves the development and utilization of wind power. To further locate the fault quickly and accurately, the double-terminal fault location is introduced. The schematic diagram of the double-terminal location is shown in Figure 6.

Negative sequence network is often used for transmission line fault analysis. However, ABCG is different from other fault types and belongs to the symmetrical fault. It is well known that ABCG only produces positive sequence components. Therefore, only in the analysis of ABCG, the positive sequence component is used to replace the negative sequence component. The fault voltage and current obtained from the fault two terminals are converted into negative or positive sequence. According to KVL, the voltage component \( V_{1y} \) and \( V_{2y} \) measured by nodes B1 and B2 can be determined as follows [43]:

\[
V_{1y} = I_{1y}Z_y + V_f
\]

\[
V_{2y} = (L-\ell)I_{2y}Z_y + V_f
\]

where \( L \) is the total length of DWTL; \( \ell \) is fault distance from bus1 on DWTL; \( Z_y \) is the sequence impedance; \( I_{1y} \) is the current component measured at node B1; \( I_{2y} \) is the current component measured at node B2; \( y = 1, 2; 1 \) represents positive sequence (ABCG), 2 represents negative sequence (other fault types); \( V_f \) is fault voltage at fault point. Based on Equations (12) and (13), \( l \) can be calculated by:

\[
l = \frac{V_{1y} - V_{2y} + L \times I_{2y} \times Z_y}{Z_y \times (I_{1y} + I_{2y})}
\]

According to Equation (14), it can be seen that the proposed method can ignore the fault point voltage and calculate the fault location preliminarily [43, 44]. Based on the NRM, the synchronization angle \( \xi \) is introduced to achieve data synchronization. Then Equation (14) can be written as

\[
l = \frac{V_{1y} \times e^{j\xi} - V_{2y} + L \times I_{2y} \times Z_y}{Z_y \times (I_{1y} \times e^{j\xi} + I_{2y})}
\]

Synchronization angle \( \xi \) can be expressed in a polar form as:

\[
\xi = \cos(\xi) + j \times \sin(\xi)
\]

Based on Equation (15), Equation (16) can be rewritten as:

\[
l = \frac{\text{Re}(V_{1y}) \times \sin(\xi) + \text{Im}(V_{1y}) \times \cos(\xi) - \text{Im}(V_{2y}) + W_1}{W_2 \times \sin(\xi) + W_3 \times \cos(\xi) + W_4}
\]

where \( \text{Re}(\cdot) \) and \( \text{Im}(\cdot) \) are the real and imaginary of the fault signal. The coefficients \( W_1, W_2 \) and \( W_3 \) in Equation (17) are defined as [43]:

\[
W_1 = R \times \text{Im}(I_{2y}) + X \times \text{Re}(I_{2y})
\]

\[
W_2 = R \times \text{Re}(I_{1y}) - X \times \text{Im}(I_{1y})
\]

\[
W_3 = R \times \text{Im}(I_{1y}) + X \times \text{Re}(I_{1y})
\]

5 SIMULATION RESULTS AND DISCUSSIONS

5.1 Simulation environment

A DOWF transmission line model includes a power source, a large number of wind turbines and several short-distance transmission lines [22, 45]. A DOWF is taken as an example to construct a DWTL simulation model for offshore doubly-fed WT’s in real-time simulation software Matlab/Simulink, and its system wiring diagram is shown in Figure 7.

As mentioned earlier, the bus is also studied by the proposed method. 10 fault branches (6 fault transmission lines and 4 fault bus) and 10 fault types are considered in order to verify the
feasibility and overall accuracy of the proposed method for DWTLs. The 10 faulty branches are L1–L10. The 10 fault types include AG, BG, CG, ABG, ACG, BCG, ABCG, AB, AC and BC. Meanwhile, in order to verify the proposed method is not affected by $\alpha_f$, $R_f$ and $L_f$. The fault cases are generated by randomly changing $A_f$, $R_f$ and $L_f$. The parameter setting of simulation model and different fault conditions are summarized in Table 4.

The value range of $A_f$ must be greater than one simulation cycle (0–360°) in order to fully consider each fault inception angle. $A_f$ interval is set to 0.1 s to meet the above requirement. $R_f$ is in Ω and $L_f$ is in km. The range of $R_f$ and $L_f$ are 0.001–15 Ω and 0.1–0.9 km. $A_f$, $R_f$ and $L_f$ are randomly changed to generate a large number of random fault cases.

5.2 | Performance for fault classification

In order to test the efficacy of the proposed fault classification method, different DWTL faults are studied. 10 DWTL fault types are considered, including AG, BG, CG, ABG, ACG, BCG, ABCG, AB, AC and BC. On the fault branches L1, 50 fault cases are generated for each fault type. $10 \times 50$ fault cases are randomly generated on L1. Similarly, like L1 faulty cases, different fault types are applied to other branches including L2–L10. Thus, 500 fault cases are generated for each DWTL fault type. A total of 5000 fault cases are used to verify the overall effectiveness of the proposed classification method. 70% fault cases (3500 fault cases) are used for training and the remaining 30% (1500 fault cases) are used for testing. Table 5 shows a 9 × 9 confusion matrix obtained from the proposed classification method under a noise-free condition.

Under the same fault conditions ($\alpha_f$, $R_f$ and $L_f$), different levels of Gaussian white noise are added to the fault current signal. The SNR of MN includes 20, 30 and 40 dB. The fault eigenvalues generated by noisy current signal is used as the input of classification method to classify DWTLs fault. Different noises are added to verify that the proposed method is not influenced by MN. Results and accuracy of fault classification under different noise conditions are shown in Table 6.

As can be concluded from Tables 5 and 6, at each noise level, 1500 cases are used to test the effectiveness of the proposed fault classification method. The proposed fault classification method performs best without the influence of MN. The accuracy of fault classification can also reach 100%. The accuracy is slightly reduced when the noise is added to the fault current signal. But the accuracy of fault classification is still as high as 99.53% even under the influence of 20 dB noise. Meanwhile, the accuracy of the proposed method is 99.73% and 99.80% at 40 dB and 30 dB, respectively.

Consequently, the effectiveness of the proposed fault classification method in DWTL fault classification is verified by the above results. Meanwhile, the method is not influenced by fault inception angle, fault resistance, fault location and measurement noise, and the overall accuracy is as high as 99.77%.

5.3 | Performance for fault branch identification

After accurate classification of fault types, the fault types of DWTLs in DOWF are determined. According to different fault types, 10 corresponding fault branch identification models are generated. Then, the fault eigenvalues generated by three-phase fault current data recorded in the previous branch are fed to the corresponding fault branch identification model. 400 AG fault cases are generated on each DWTL fault branch. Thus, a total of 4000 AG fault cases are randomly generated to verify the effectiveness of the proposed classification method when AG occurs in DWTL. 70% fault cases (2800 AG fault cases) are used for training and the remaining 30% (1200 AG fault cases) are used for testing. A 10 × 10 AG confusion matrix obtained
TABLE 5  Results and accuracy of fault classification for noise-free measurement

|       | AG | BG | CG | ABG | ACG | BCG | ABCG | AB | AC | BC |
|-------|----|----|----|-----|-----|-----|------|----|----|----|
| AG    | 150| 0  | 0  | 0   | 0   | 0   | 0    | 0  | 0  | 0  |
| BG    | 0  | 150| 0  | 0   | 0   | 0   | 0    | 0  | 0  | 0  |
| CG    | 0  | 0  | 150| 0   | 0   | 0   | 0    | 0  | 0  | 0  |
| ABG   | 0  | 0  | 0  | 150 | 0   | 0   | 0    | 0  | 0  | 0  |
| ACG   | 0  | 0  | 0  | 0   | 150 | 0   | 0    | 0  | 0  | 0  |
| BCG   | 0  | 0  | 0  | 0   | 0   | 150 | 0    | 0  | 0  | 0  |
| ABCG  | 0  | 0  | 0  | 0   | 0   | 0   | 150  | 0  | 0  | 0  |
| AB    | 0  | 0  | 0  | 0   | 0   | 0   | 0    | 150| 0  | 0  |
| AC    | 0  | 0  | 0  | 0   | 0   | 0   | 0    | 0  | 150| 0  |
| BC    | 0  | 0  | 0  | 0   | 0   | 0   | 0    | 0  | 0  | 150|
| Accuracy (%) | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Overall accuracy(%) is 100

TABLE 6  Results and accuracy of fault classification under different noise conditions

| Fault type | Number of successful cases classified |
|------------|--------------------------------------|
|            | Free | 40 dB  | 30 dB  | 20 dB  |
| AG         | 150  | 150    | 150    | 150    |
| BG         | 150  | 150    | 150    | 150    |
| CG         | 150  | 149    | 149    | 149    |
| ABG        | 150  | 150    | 149    | 149    |
| ACG        | 150  | 150    | 149    | 149    |
| BCG        | 150  | 150    | 149    | 149    |
| ABCG       | 150  | 149    | 149    | 149    |
| AB          | 150  | 150    | 149    | 149    |
| AC          | 150  | 150    | 149    | 149    |
| BC          | 150  | 149    | 150    | 149    |
| Accuracy (%) | 100  | 99.80  | 99.73  | 99.53  |

Overall accuracy(%) is 99.77

from the proposed branch identification method under a noise-free condition is shown in Table 7.

As with the fault classification problem, \( A_f, R_f, L_f \) and MN are randomly selected when all AG fault cases are generated. Different white Gaussian noise is also added to the fault current signal. The SNR of MN includes 20, 30 and 40 dB. Results and accuracy of fault branch identification under different noise conditions are shown in Table 8.

As can be concluded from Tables 7 and 8, in case of AG fault on the DWTLs, the effectiveness of the fault branch identification method is verified by 1200 fault cases. The accuracy of the fault branch identification method for DWTL can reach 99.83% without the influence of MN. Even under the influence of noise, the accuracy of the proposed method is 99.75%, 99.58% and 99.17% at 40 dB, 30 dB and 20 dB respectively. Consequently, the proposed fault branch identification method is not influenced by \( A_f, R_f, L_f \) and MN. The overall accuracy is as high as 99.58%, which fully verifies the effectiveness of the fault branch identification method when AG occurs on DWTLs.

As with AG fault cases, 4000 different fault cases are generated under the other fault types respectively. Results and accuracy of fault branch identification under different noise conditions in all fault type cases are shown in Table 9.

As can be concluded from Table 9, the proposed fault branch identification method is verified by 4000 fault cases under different DWTL fault types. When the noise level is 20 dB, the accuracy of fault branch identification is as low as 99.17% after AG fault occurs. When the fault condition is noise-free, the accuracy of fault branch identification is as high as 100% after BG fault occurs. The overall accuracy can reach 99.58% (AG), 97.69% (BG), 99.63% (CG), 99.58% (ABG), 99.58% (ACG), 99.65% (BCG), 99.50% (ABCG), 99.48% (AB), 99.58% (AC) and 99.58% (BC) in the cases of considering noise. The accuracy of fault branch identification decreases with the increase of noise level no matter what kind of fault occurs on DWTLs. But the overall classification accuracy is approximately 99.5%.

5.4  Performance for fault location

The accurate fault type and fault branch of DWTLs are determined by the above proposed method. The fault location of DWTL can be located by sending the fault dates recorded at double-terminal of the fault branch to Equation (17). 200 fault cases are randomly generated on the fault branch \( L_1 \), where \( A_f, R_f, L_f \) are randomly set as the fault classification and fault branch identification problems. Location results and errors of 50 fault cases on \( L_1 \) are shown in Figure 8.

As can be seen from Figure 8, when the DWTL fault occurs in branch \( L_1 \), the fault location can be accurately located by the proposed method on the premise of determining the fault line and fault type. The maximum error can be reached 0.25%, and the minimum error can be reached 0.01%. In the detection of 1 km transmission line, the location error can be controlled
TABLE 7  Results and accuracy of fault branch identification for noise-free in the AG cases

| L1 | L2 | L3 | L4 | L5 | L6 | L7 | L8 | L9 | L10 |
|----|----|----|----|----|----|----|----|----|-----|
| 120 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 120| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 0  | 120| 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 120| 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 1  | 0  | 0  | 0  | 119| 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 120| 0  | 0  | 0  | 0  | 119| 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   |

Accuracy (%) 100 100 100 100 100 99.17 100 100 100 99.17

Overall accuracy(%) is 99.83

TABLE 8  Results and accuracy of branch identification under different noise conditions in AG cases

| Number of successful cases classified     | Free | 40 dB | 30 dB | 20 dB |
|------------------------------------------|------|-------|-------|-------|
| Fault section                            | 120  | 120   | 120   | 120   |
| L1                                       | 120  | 120   | 120   | 120   |
| L2                                       | 120  | 120   | 120   | 120   |
| L3                                       | 120  | 120   | 119   | 118   |
| L4                                       | 120  | 120   | 120   | 120   |
| L5                                       | 120  | 120   | 120   | 120   |
| L6                                       | 119  | 119   | 118   | 118   |
| L7                                       | 120  | 119   | 119   | 119   |
| L8                                       | 120  | 120   | 120   | 120   |
| L9                                       | 120  | 120   | 120   | 120   |
| L10                                      | 119  | 119   | 119   | 119   |

Accuracy (%) 99.83 99.75 99.58 99.17

Overall accuracy(%) is 99.58

TABLE 9  Overall accuracy of fault branch identification under different noise conditions

| Fault type | Free | 40 dB | 30 dB | 20 dB | Overall accuracy |
|------------|------|-------|-------|-------|------------------|
| AG         | 99.83| 99.75 | 99.58 | 99.17 | 99.58            |
| BG         | 100  | 99.83 | 99.67 | 99.25 | 99.69            |
| CG         | 99.92| 99.67 | 99.50 | 99.42 | 99.69            |
| ABG        | 99.83| 99.67 | 99.58 | 99.25 | 99.58            |
| ACG        | 99.92| 99.75 | 99.50 | 99.17 | 99.58            |
| BCG        | 99.83| 99.75 | 99.67 | 99.33 | 99.65            |
| ABCG       | 99.75| 99.58 | 99.50 | 99.17 | 99.50            |
| AB          | 99.75| 99.58 | 99.42 | 99.17 | 99.48            |
| AC          | 99.83| 99.75 | 99.50 | 99.25 | 99.58            |
| BC          | 99.83| 99.75 | 99.58 | 99.17 | 99.58            |

5.5  Influence of system parameters on accuracy

The length of DWTL and the sampling frequency of the system are important factors that influence the accuracy of fault classification, fault branch identification and the error of double-terminal fault location. In the research of line fault, the setting of sampling frequency is not uniform. Different reference sets different sampling frequency. Most of the sampling frequencies are 3.2, 6.4, 10, 12.6 and 20 kHz [23, 46–48]. In order to illustrate the selection basis of the sampling frequency used in the proposed method, the variations of sampling frequencies (3.2, 6.4, 10, 12.6 and 20 kHz) are studied. Only the sampling frequency of the system is changed, the accuracy or error of the proposed method is shown in Table 11 when other fault conditions are the same.

As can be concluded from Table 11, the overall accuracy is close to the peak value at 10, 12.6 and 20 kHz, which is about
99.8%. The maximum error is close to the peak value at 6.4, 10, 12.6 and 20 kHz, which is about 0.6%. The accuracy is improved and the error is reduced with the increase of the sampling frequency. The results show that the proposed method is suitable for sampling frequency above 10 kHz.

To evaluate the proposed method, the variations of DWTL length (1, 2, 5 and 10 km) are studied. Only the length of DWTL is changed, the accuracy or error of the proposed method is shown in Table 12 when other fault conditions are the same. As can be concluded from Table 12, the accuracy of fault classification is improved and the error of fault location is reduced with the increase of the DWTL length. The overall accuracy of fault classification is close to the peak value at 5 and 10 km, which is 100%. The overall accuracy of fault branch identification is basically unchanged. The maximum error of fault location is close to the peak value at 10 km, which is 0.55. The results show that the proposed method is not only suitable for short-distance transmission fault detection, but also suitable for long-distance transmission line detection.

### 5.6 Comparison of performance under different methods

In order to verify the superiority of the proposed STRF-based double-terminal fault location method, different reference methods are applied to fault classification, branch identification and location of DWTLs in DOWF. Different reference methods are compared with the proposed method in terms of overall accuracy and maximum error. The comparisons of the fault classification, branch identification and location methods are shown in Figure 9.
 FIGURE 9  The comparisons of the fault classification, branch identification and location methods (a) The comparisons of fault classification, (b) the comparisons of the fault branch identification, (c) the comparisons of the fault location

As can be concluded from Figure 9, after all the methods mentioned above are applied to DOWF fault diagnosis, some methods still maintain high performance, but the overall accuracy is low in fault branch identification problem. The proposed method outperforms the others in fault classification, branch identification and location aspect. Meanwhile, the influence of $A_f$, $R_f$ and $L_f$ on accuracy can be ignored by the proposed method. Thus, the proposed method can effectively locate faults with better or a competitive accuracy over the referenced methods.

The speed of fault classification, branch identification and location is briefly demonstrated in this section. The speed determines whether it can meet the real-time requirements, and then determines whether it can be implemented online. For any complex structure line fault monitoring problem, the problem can be roughly divided into four steps: fault detection, classification, branch identification and location. The starting point of this paper is to study the remaining steps after the fault detection, so the proposed method does not need to be implemented online. But in order to demonstrate the feasibility and reliability of the proposed method, the speed of fault classification, fault branch identification and fault location is shown in Table 13.

As can be seen from Table 13, the proposed fault classification method requires an average time of 7.8541 s to generate the final model. After receiving the three-phase fault current signal, the average time required for fault classification is 0.0156 s, which is less than one cycle of 60 Hz system. The proposed fault branch identification method requires an average time of 8.1523 s to generate the final model. The average time required for fault branch identification is 0.0177 s. The proposed fault location method needs 0.0248 s to estimate the fault location. As mentioned earlier, the proposed method is to operate after the fault occurs, so it does not need to be implemented online. However, the proposed method can still complete different steps in about a cycle of 60 Hz system.

6  CONCLUSIONS

This paper proposes an innovative double-terminal fault location method based on Stockwell-transform (ST), Random-forest (RF) and Newton-Raphson (NRM) for deep-sea offshore wind farm (DOWF). Through a large number of simulations, the feasibility and effectiveness of this method in fault types classifying, branches identifying and fault locating have been verified for short-distance transmission lines. The following conclusions have been reached:

1. ST is used to decompose the fault current signals, and a large number of fault eigenvalues are generated by using the generated fault eigenvalue matrices. The introduction of PCC removes a lot of duplicate eigenvalues, while it does not affect the performance of eigenvalues.
2. After filtering eigenvalues, according to the representation of eigenvalues in different fault types and fault branches, the selection basis of RF is given. And the correctness of RF has also been verified through a lot of simulations.

3. After the fault type and branch are accurately determined, the double-terminal fault location equation is established. In order to solve the problem of calculating errors caused by asynchrony data, NRM is introduced. Synchronization angle $\xi$ greatly improves the accuracy of fault location.

4. More than 4000 fault cases show that the accuracy of fault classification and branch identification for DWTL has been significantly improved by this method. The overall accuracy of fault classification and branch identification be as high as 99.77% and 99.58% respectively with almost 100%. The location error can be controlled within 6.1 m for short-distance transmission fault location of DOWF.

5. The proposed method overcomes the influence of transmission line length, fault conditions uncertainty (fault inception angle, fault resistance, fault location, measurement noise) and complex wind farm topology, which ensure good accuracy of fault classification, branch identification and location for short-distance transmission lines in DOWF.

6. As a part of the future extension, considering the integration of renewable energy, the sensitivity of the developed method be investigated by applying in a microgrid. Besides, the feasibility of the method is studied by considering wind speed change, load change and different line parameters, so as to solve the difficulty of the short distance submarine transmission line location and reduce the operation and maintenance time of smart grid.

ORCID

Xiaodong Wang https://orcid.org/0000-0002-0911-9214
Xing Gao https://orcid.org/0000-0002-0312-4020
Yingming Liu https://orcid.org/0000-0003-1204-6199
Yonghao Wang https://orcid.org/0000-0003-0368-7357

REFERENCES

1. Wang, X.D., Wang, Y.W., Liu, Y.M.: Dynamic load frequency control for high-penetration wind power considering wind turbine fatigue load. Int. J. Electr. Power 117(1), 1–11 (2020)

2. Chatzipetros, D., Pilgrim, J.A.: Review of the accuracy of single core equivalent thermal model for offshore wind farm cables. IEEE Trans. Power Delivery 33(4), 1913–1921 (2018)

3. Colm, M.A.H., Pilgrim, J.A.: Cable thermal risk estimation for over-planted wind farms. IEEE Trans. Power Delivery 35(2), 609–617 (2020)

4. Zheng, L.M., et al.: A novel structural similarity based pilot protection for renewable power transmission line. IEEE Trans. Power Delivery 35(6), 2672–2681 (2020)

5. Maragan, A.P., Marquez, F.P.G.: Advanced analytics for detection and diagnosis of false alarms and faults: A real case study. Wind Energy 22(11), 1622–1635 (2019)

6. Agarwal, S., et al.: Fault analysis method of integrated high voltage direct current transmission lines for onshore wind farm. J. Mod. Power Syst. Clean Energy 7(3), 621–632 (2019)

7. Catmull, S., et al.: Cyclic load profiles for offshore wind farm cable rating. IEEE Trans. Power Delivery 31(3), 1242–1250 (2016)

8. Rizk, M.E.M., et al.: Investigation of lightning electromagnetic fields on underground cables in wind farms. IEEE Trans. Electromagn. Compat. 58(1), 143–152 (2016)

9. Guo, G., Zha, K., Zhang, J., et al.: Grounding fault in series-connection-based offshore wind farms: fault clearance. IEEE Trans. Power Electron. 35(9), 9357–9367 (2020)

10. Herlender, J., Izykowski, J., Solak, K.: Compensation of the current transformer saturation effects for transmission line fault location with impedance-differential relay. Electr. Power Syst. Res. 182, 106223 (2020)

11. Liang, Y., et al.: A novel fault impedance calculation method for distance protection against fault resistance. IEEE Trans. Power Delivery 35(1), 396–407 (2020)

12. Jain, L., Jianshi, N., Yu, D.: A unified matrix algorithm for fault segment judgment and isolation in distribution network system. Autom. Electr. Power Syst. 1(1), 31–33 (1999)

13. Tang, J., Yin, X., Zhang, Z.: Traveling-wave-based fault location in electrical distribution systems with digital simulations. Telkomnika 12(2), 297–304 (2014)

14. Lopes, F.V., et al.: Practical methodology for two-terminal traveling wave-based fault location eliminating the need for line parameters and time synchronization. IEEE Trans. Power Delivery 34(6), 2123–2134 (2019)

15. Shu, H., et al.: Fault model and travelling wave matching based single terminal fault location algorithm for T-connection transmission line: A Yunnan power grid study. Energies 13(6), 1506 (2020)

16. Rafinia, A., Moshtagh, J.: A new approach to fault location in three-phase underground distribution system using combination of wavelet analysis with ANN and FLS. Int. J. Electr. Power Energy Syst. 55(6), 261–274 (2014)

17. Nakhlé, M.P., Safavi, A.A.: Path characteristic frequency-based fault locating in radial distribution systems using wavelets and neural networks. IEEE Trans. Power Delivery 26(2), 772–781 (2011)

18. Ray, P., Mishra, D.: Artificial intelligence based fault location in a distribution system. In: Proceedings of the International Conference on Information Technology. Bhubaneswar, India, pp. 22–24 (2014)

19. Shafiuullah, M., et al.: Optimized support vector machine & wavelet transform for distribution grid fault location. In: Proceedings of the International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG). Cadiz, Spain, pp. 4–6 (2017)

20. Liang, J., et al.: Two-terminal fault location method of distribution network based on adaptive convolution neural network. IEEE Access 8, 54035–54043 (2020)

21. Shafiuullah, M., Abido, M.A., Fattah, T.A.: Distribution grids fault location employing ST based optimized machine learning approach. Energies 11(9), 2328 (2018)

22. Zhang, K., Zhu, Y., Liu, X.: A fault locating method for multi-branch hybrid transmission lines in wind farm based on redundancy parameter estimation. J. Mod. Power Syst. Clean Energy 7, 1033–1043 (2019)

23. Zaki, M.I., et al.: An investigated reactive power measurements-based fault-identification scheme for teel transmission lines. Measurement 136, 185–200 (2019)

24. Wang, X.W., et al.: Single line to ground fault detection in a non-effectively grounded distribution network. IEEE Trans. Power Delivery 33(6), 3173–3186 (2018)

25. Wang, X.W., Gao, J., Chen, M.F., et al.: Faulty line detection method based on optimized bistable system for distribution network. IEEE Trans. Ind. Inf. 14(4), 1370–1381 (2018)

26. Barik, M.A., et al.: A decentralized fault detection technique for detecting single phase to ground faults in power distribution systems with resonant grounding. IEEE Trans. Power Delivery 33(5), 2462–2473 (2018)

27. Wang, X.D., et al.: WRC-SDT based on-line detection method for offshore wind farm transmission line. IEEE Access 8, 53547–53560 (2020)

28. Raghavendran, C.R., Roselyn, J.P., Devaraj, D.: Development and performance analysis of intelligent fault ride through control scheme in the dynamic behaviour of grid connected DFIG based wind systems. Energy Rep. 6, 2560–2576 (2020)

29. Bhandari, A., Eldar, Y.C.: Sampling and super-resolution of sparse signals beyond the Fourier domain. IEEE Trans. Signal Process. 67(6), 1508–1521 (2019)

30. Tyagi, T., Sumathi, P.: Comprehensive performance evaluation of computationally efficient discrete Fourier transforms for frequency estimation. IEEE Trans. Instrum. Meas. 69(5), 2155–2163 (2020)
31. Garrdio, M.: The feedforward short-time Fourier transform. IEEE Trans.
    Circuits Syst. II Express Briefs 63(9), 868–872 (2016)
32. Swetapadma, A., Yadav, A.: A novel decision tree regression based fault
    distance estimation scheme for transmission lines. IEEE Trans. Power
    Delivery 32(1), 234–245 (2017)
33. Stockwell, R.G., Mansinha, L., Lowe, R.P.: Localization of the complex
    spectrum: The S transform. IEEE Trans. Signal Process. 44(4), 998–1001
    (1996)
34. Perveen, R., Kishor, N., Mohanty, S.R.: Fault detection for offshore
    wind farm connected to onshore grid via voltage source converter-high
    voltage direct current. IET Gener. Transm. Distrib. 9(16), 2544–2554
    (2015)
35. Shafullah, M., Abido, M.A.: S-transform based FFNN approach for distribu-
    tion grids fault detection and classification. IEEE Access 6, 8080–8088
    (2018)
36. Mishra, S., Bhende, C.N., Panigrahi, B.K.: Detection and classification of
    power quality disturbances using S-transform and probabilistic neural network.
    IEEE Trans. Power Delivery 23(1), 280–287 (2008)
37. Lin, H., et al.: Adaptive protection combined with machine learning for
    microgrids. IET Gener. Transm. Distrib. 13(6), 770–779 (2019)
38. Webmaster: Statistics and Machine Learning. 2nd ed., p. 2013. MIT Press,
    Cambridge, MA (2014)
39. James, G., Witten, D., Hastie, T.: An Introduction to Statistical Learning.
    2nd ed., p. 2012. SNY Press, New York (2013)
40. Oyang, Y.J., et al.: Data classification with radial basis function networks
    based on a novel kernel density estimation algorithm. IEEE Trans. Neural
    Networks 16(1), 225–236 (2005)
41. Breiman, L.: Statistical modeling: The two cultures (with comments and a
    rejoinder by the author). Stat. Sci. 16(3), 199–215 (2001)
42. Zhong, T., et al.: Power-quality disturbance recognition based on time-
    frequency analysis and decision tree. IET Gener. Transm. Distrib. 12(18),
    4153–4162 (2018)
43. Novosel, D., Hart, D.G.: Unsynchronized two-terminal fault location esti-
    mation. IEEE Trans. Power Delivery 11(1), 130–138 (1996)
44. Ramseebhuck, N., Awodele, K.: Development, testing and comparison of
    smart fault location algorithms for smart transmission grids. In: Pro-
    ceedings of the International SAUPEC/RobMech/PRASA Conference
    (SRPC). Bloemfontein, South Africa, pp. 28–30 (2019)
45. Perez, J.A., et al.: Global optimization of offshore wind farm collection
    systems. IEEE Trans. Power Syst. 35(3), 2256–2267 (2020)
46. Zaki, M.I., et al.: Sensitive/stable complementary fault identification
    scheme for overhead transmission lines. IET Gener. Transm. Distrib.
    13(15), 3252–3263 (2019)
47. Wang, X.W., et al.: Fault feeder detection method utilized steady state
    and transient components based on FFT backstepping in distribution net-
    works. Int. J. Electr. Power Syst. 114, 105391 (2020)
48. Lima, E.M., et al.: High impedance fault detection based on Stockwell
    transform and third harmonic current phase angle. Electr. Power Syst. Res.
    175, 105931 (2019)

How to cite this article: Wang, X., et al.: Stockwell-transform and random-forest based
double-terminal fault diagnosis method for offshore wind farm transmission line. IET Renew.
Power Gener. 15, 2368–2382 (2021).
https://doi.org/10.1049/rpg2.12170