



# A fault locating method for multi-branch hybrid transmission lines in wind farm based on redundancy parameter estimation



Ke ZHANG<sup>1,2</sup> , Yongli ZHU<sup>1</sup>, Xuechun LIU<sup>1</sup>

**Abstract** In order to solve the problem of “abandoned” wind caused by short circuit faults in a wind farm, a wind farm fault locating method based on redundancy parameter estimation is proposed. Using the characteristics of the traveling wave, transmission equations containing the position of the fault point are constructed. Parameter estimation from statistical theory is used to solve the redundant transmission equations formed by multiple measuring points to locate the faults. In addition, the bad data error detection capability of the parameter estimation is used to determine bad data and remove them. This improves locating accuracy. A length coefficient is introduced to solve the error enlargement problem caused by a transmission line sag. The proposed fault locating method can solve the fault branch misjudgment problem caused by the short circuit faults near the data measuring nodes of the wind farm based on the proposed fault interval criterion. It also avoids the requirements to the traveling wave speed of traditional methods, thus its fault location is more accurate.

Its effectiveness is verified through simulations in PSCAD/EMTDC, and the results shows that it can be used in the fault locating of hybrid transmission lines.

**Keywords** Transmission line, Fault location, Parameter estimation, Wind power, Least squares

## 1 Introduction

In recent years, wind power generation is increasing year by year, and the wind power is getting much attention as a form of energy power generation. However, due to the complex structure and the bad operational condition of any transmission line in the wind farm, its faults occur frequently, which reduces the wind power generation. Therefore, it is very important to quickly find the fault point after the line failure. Because there are many branches in a wind farm transmission line, conventional fault locating methods are not practical.

The existing fault locating methods can be mainly classified into two classes: the fault analysis methods [1–3] and traveling wave methods [4, 5]. The former ones can be further subdivided into single-ended and dual-ended. The single-ended method calculates the fault distance based on voltage and current phasors at one end of the faulty line. However, its accuracy largely depends on the fault type, ground impedance and the line model parameters etc. The dual-ended method uses voltage and current phasors at both ends of the line and is usually more accurate than the single-ended method.

The traveling wave methods can also be further classified into single-ended and dual-ended. The single-ended methods use the first traveling wave and the first reflected wave from the fault point measured by oneline terminal to

CrossCheck date: 18 October 2018

Received: 8 June 2018 / Accepted: 18 October 2018 / Published online: 1 April 2019

© The Author(s) 2019

✉ Yongli ZHU  
yonglipw@163.com

Ke ZHANG  
m18330220931@163.com

Xuechun LIU  
lxc13722958050@163.com

<sup>1</sup> State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Baoding 071003, China

<sup>2</sup> State Grid Wuxi Power Supply Company, Wuxi 214000, China



determine the fault location [6]. They can be further classified into A, C, E, and F fault type locators according to the generation mechanisms of their traveling waves. The dual-ended methods use the time differences between the arrivals of the first wavefronts at both terminals to achieve the fault location [7, 8]. They are classified as B and D fault type locators according to the synchronization mechanism of the line terminals. In general, the locating accuracy of the traveling wave methods is better than that of fault analysis methods.

The above methods are commonly used in end-to-end or simple structured transmission lines. Some new algorithms have been proposed for multi-branch hybrid transmission lines: ① improved methods based on traditional methods [9, 10]; ② fault locating methods based on pattern recognition techniques such as Artificial Neural Networks and K-means clustering algorithms [11]; and ③ a method based on dynamic state estimation is proposed [12, 13]. The principle of the first type of methods is first determining the fault branch and then locating the fault point based on the determined fault branch by the traditional method. The second type of methods often require a large quantity of fault data to constitute a database, which contains historical fault data of a real system or those derived from an accurate simulation system. The third type of methods has reference value in fault locating study. However, it did not consider the hybridity of overhead lines and cables. Although the cables in a wind farm transmission line are short, they can not be ignored, which can be demonstrated by Table A1 in Appendix A.

Aiming at the multiple branches and hybrid characteristics of the wind farm transmission lines, a redundancy parameter estimation method to locate fault points is proposed. Firstly, different from the traditional traveling wave method, this proposed method constructs the transmission equations based on the grid synchronization information. In addition, it comprehensively considers the hybridity of overhead lines and cables, and uses statistical methods to estimate the parameters to locate the fault point. Secondly, the length coefficient, which is the ratio of the fault distance and the total length of the fault branch, is used to indirectly find the position of the fault point on the fault branch. It can avoid the amplification of the locating error caused by the line length changes caused by the seasonal sag variations. Furthermore, as the bad measured data may cause large fault locating error, a bad data detecting algorithm is used. Finally, PSCAD/EMTDC simulation results verify the feasibility of the proposed method.

The proposed method has the following characteristics. ① It realizes the fault location on hybrid transmission lines including overhead and cable segments; ② Compared with the traveling wave method, it does not require the calculation of traveling wave velocity which easily causes a

locating error; ③ It takes the line sag effect into the fault locating model; ④ It constructs a fault locating model based on multiple measuring points which are adopted to do parameter estimation to minimize measurement errors; ⑤ The bad data detecting algorithm is used to ensure the correctness of the original data and locating results. The method for the parameter estimation in this paper is completely different with Reference [12] which uses the SCPQDM dynamic model to simulate the power grid and uses a dynamic state estimation to achieve fault location. Reference [12] is suitable for the power grids of lines of two ends and three ends, the proposed method can be used to locate the faults in the complex grids of multi-branch lines and transmission lines containing overhead and cable segments. It can consider the line sag effect and find the bad data to improve the fault locating results.

## 2 Fault locating method

### 2.1 Wind farm electric system model

The schematic electric system diagram of a large-scale wind farm is shown in Fig. 1. In general, current large-scale wind farm structure has the following characteristics.

- 1) Overhead and cable segments are hybridly connected in a wind farm transmission line.
- 2) The distance between adjacent wind turbines is short, ranging from a few hundred meters to a few kilometers.

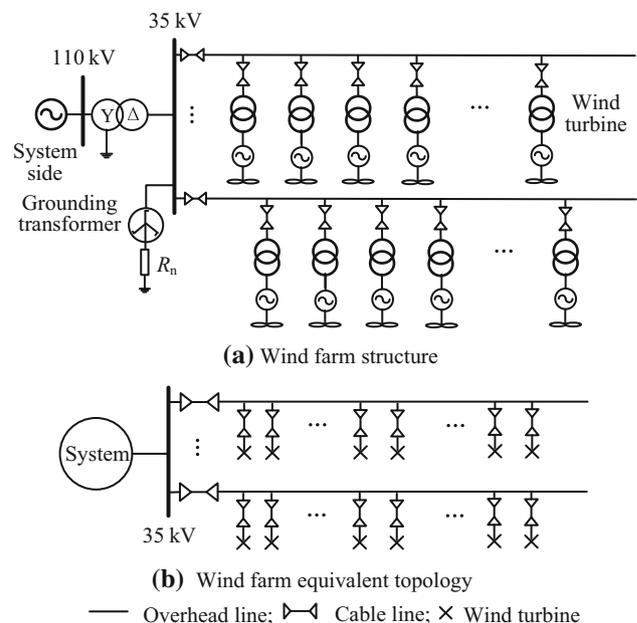


Fig. 1 Wind farm structure and equivalent diagram

**Table 1** Capability comparison of fault locating methods

Method	Multi-branch	Hybrid lines	Sag effect	Misjudgement near nodes
M1	○●	○	○	○
M2	○●	○	○	○●
M3	●	○	○	○●
M4	●	●	●	●

Note: ○ denotes cannot handle the characteristic; ● denotes can handle the characteristic; ○● denotes can handle the characteristic within a certain range

3) From Fig. 1b, it can be seen that the wind farm is a typical radial network and a large number of wind turbines are connected. In addition, the total length of a collecting power line is usually no more than 20 km.

The above characteristics of the wind farm lead to difficulty in finding fault points in a wind farm line by the conventional locating algorithms. In addition, the failure near the node are also difficult to locate by the traditional methods.

The fundamental frequency phasor-based methods are named M1 and the traveling wave-based methods are named M2. The method in [12] is named M3, and the proposed method is named M4. The capability of the fault locating methods to deal with the above characteristics of transmission lines is listed in Table 1.

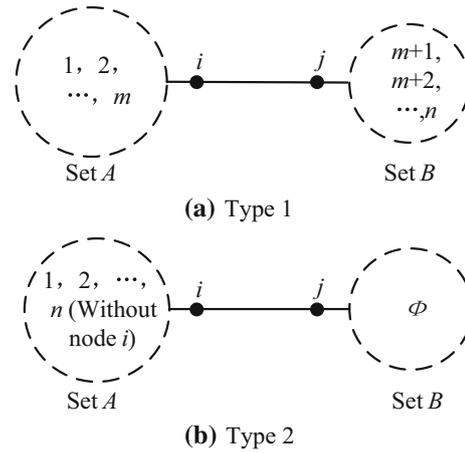
### 2.2 Construction of transmission line equations

When a fault occurs in a wind farm transmission line, the traveling wave generated by the fault is transmitted along the shortest path to each terminal of the line. Therefore, the time for the wave to reach any terminal of the line is a function of the shortest traveling path, which is defined as the transmission equation in the paper. There are two types of faulty line branches, as shown in Fig. 2, assuming that the fault occurs at branch  $i$ - $j$ .

Type 1 represents a fault occurs at the branch consisting only of middle nodes. Type 2 represents a fault occurring at the branch with a terminal node,  $j$  node is a line terminal in Fig. 2b. Assuming that there are  $n$  terminal nodes, the set  $A$  represents all the terminal nodes locating at the left of the node  $i$ ; the set  $B$  represents all the terminal nodes locating at the right of the node  $j$ . The corresponding transmission equations of the terminal nodes are shown in (1) and (2) respectively.

$$t_{Ap} = l_{1pi}/v_1 + l_{2pi}/v_2 + kl_{ij}/v + t_0 \quad p \in A \quad (1)$$

$$t_{Bq} = l_{1qj}/v_1 + l_{2qj}/v_2 + (1 - k)l_{ij}/v + t_0 \quad q \in B \quad (2)$$



**Fig. 2** Schematic diagram for faulty branch classification

where  $t_{Ap}$  and  $t_{Bq}$  correspond to the first wavefront arrival time moments measured by the  $p$  and  $q$  measuring points, respectively; all subscripts  $A$  and  $B$  in the above equations indicates their corresponding variables are relating to their terminal sets  $A$  and  $B$ ; subscripts 1 and 2 are related to overhead sections and cable sections respectively;  $l_{1pi}$  and  $l_{1qj}$  correspond to the length sum of the overhead sections' through which the first wavefront reaches the measuring points  $p$  and  $q$  from the node  $i$  and  $j$  of the fault branch, respectively; The representation  $l_{2pi}$  and  $l_{2qj}$  are similar to  $l_{1pi}$  and  $l_{1qj}$ .  $l_{ij}$  is the length of the fault branch;  $k$  is the ratio of the distance from the fault point to the node  $i$  to the length of the fault branch;  $v_1$  and  $v_2$  correspond to the wave speeds of traveling waves on overhead lines and cables, respectively;  $v$  is the wave speed on the fault line;  $t_0$  is the fault moment.

Because the total length of cable sections of a collecting power line are very short in a wind farm, a short-circuit fault is assumed to be an overhead type and the corresponding wave speed  $v = v_1$ . The transmission equations of all terminal measuring points are established in (3). The meaning of all the variables and their subscripts in (3) can be found according to that of (1) and (2).

$$\begin{cases} \begin{bmatrix} t_{A1} \\ \vdots \\ t_{Am} \end{bmatrix} = \begin{bmatrix} l_{11i} & l_{21i} & l_{ij} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ l_{1mi} & l_{2mi} & l_{ij} & 1 \end{bmatrix} \begin{bmatrix} 1/v_1 \\ 1/v_2 \\ k/v_1 \\ t_0 \end{bmatrix} \\ \begin{bmatrix} t_{B(m+1)} \\ \vdots \\ t_{Bn} \end{bmatrix} = \begin{bmatrix} l_{1(m+1)i} & l_{2(m+1)i} & l_{ij} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ l_{1ni} & l_{2ni} & l_{ij} & 1 \end{bmatrix} \begin{bmatrix} 1/v_1 \\ 1/v_2 \\ (1-k)/v_1 \\ t_0 \end{bmatrix} \end{cases} \quad (3)$$

Equation (3) can be rewritten into (4):



$$\begin{bmatrix} t_{A1} \\ \vdots \\ t_{Am} \\ t_{B(m+1)} \\ \vdots \\ t_{Bn} \end{bmatrix} = \begin{bmatrix} l_{11i} & l_{21i} & l_{ij} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ l_{1mi} & l_{2mi} & l_{ij} & 1 \\ l_{1(m+1)j} + l_{ij} & l_{2(m+1)j} & -l_{ij} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ l_{1nj} + l_{ij} & l_{2nj} & -l_{ij} & 1 \end{bmatrix} \begin{bmatrix} 1/v_1 \\ 1/v_2 \\ k/v_1 \\ t_0 \end{bmatrix} \quad (4)$$

Equation (4) can be abbreviated into the form of (5):

$$\mathbf{t} = \mathbf{H}\mathbf{x} \quad (5)$$

where  $\mathbf{t}$  is a column vector consisting of the arrival time moment of the first wavefront at each measuring point;  $\mathbf{H}$  is a matrix composed of coefficients of (4), which is called a coefficient matrix;  $\mathbf{x}$  is the independent variable vector.

The ratio of the number of independent measurements to the number of independent variables of  $\mathbf{x}$  is defined as redundancy, then (4) is a set of equations with a high degree of redundancy which are overdetermined equations. The traveling wave velocity and the position of the fault can be obtained through solving (5).

### 2.3 Solving transmission equations through parameter estimation

Although the equations in (4) or (5) can be used to obtain the independent variables, the solution is not as accurate as solving the whole overdetermined equations in (5). Because making full use of all the transmission equations can eliminate measured data errors, a statistical parameter estimation method based on redundancy is proposed to cope with the situation where the measured data have a high degree of redundancy. It can reduce the locating error and be realized through the parameter estimation algorithm [14]. The formula (5) has sufficient redundancy, which is discussed Section 2.2, to satisfy the preconditions of the statistical parameter estimation method.

Considering errors such as measurement errors, (5) should be rewritten to (6):

$$\mathbf{t} = \mathbf{H}\mathbf{x} + \boldsymbol{\delta} \quad (6)$$

where  $\boldsymbol{\delta}$  is a set of errors.

Using the weighted least squares criterion, an objective function with minimum error is established in (7) for (6).

$$J(\mathbf{x}) = [\mathbf{t} - \mathbf{H}\mathbf{x}]^T \mathbf{R}^{-1} [\mathbf{t} - \mathbf{H}\mathbf{x}] \quad (7)$$

where  $\mathbf{R}$  is the measurement error variance matrix, a diagonal matrix, which represents the weighting of the measured quantities. The diagonal elements are relating to the measurement equipment error of the quantities.

Perform the derivation to (7) and let  $\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = 0$ , i.e.:

$$\frac{\partial J(\mathbf{x})}{\partial x_k} = -2 \sum_{i=1}^m \frac{\left[ t_i - \sum_{j=1}^n h_{ij} x_j \right] h_{ik}}{\sigma_i^2} = 0 \quad (8)$$

$$k = 1, 2, \dots, n$$

where  $x_k$  represents the  $k$ -th variable;  $m$  is the number of transmission equations;  $n$  is the number of independent variables;  $\sigma_i^2$  is the  $i$ -th measurement error variance,  $1/\sigma_i^2$  is the  $i$ -th main diagonal element of matrix  $\mathbf{R}$ ;  $h_{ij}$  is the corresponding element of the  $i$ -th row and  $j$ -th column of the coefficient matrix  $\mathbf{H}$ ;  $h_{ik}$  is the corresponding element of the  $i$ -th row and  $k$ -th column of the coefficient matrix  $\mathbf{H}$ .

Equation (8) is an  $n$ -dimensional equations that contains  $n$  independent variables. we can obtain the best estimate  $\mathbf{x}'$  of  $\mathbf{x}$  without iterative calculation by solving (8). In the case of measured data without error, the result of the above process is the final locating result. However, it is unavoidable that the measured data in a real system is erroneous. In this case, the fault locating result is deviated or even mistaken. Therefore, the bad data of the sample need to be detected and removed to ensure the reliability of the data after the above model is solved. This paper uses the confidence interval of the sample residual as the judgment of the bad data. The specific judging operations are stated as follows. Firstly, the sample residual confidence interval graph is drawn. Secondly, if the confidence interval for all sample residuals contains the zero value, the conclusion is that there are no bad data in the measured data, indicating that the model's locating results are reliable. Thirdly, otherwise, the data will be removed as an outlier, thus the corresponding equation in (4) or (5) should be deleted, and then use the parameter estimation method to solve the updated (5) again.

### 2.4 Bad data detection

In order to ensure the correctness of the data, the transmission equations are solved to find and eliminate the bad data in the measured data by the parameter estimation equations. In the theory of parameter estimation, the measured data are only in a certain interval centering on the estimated values of the measured data at a certain confidence level. The intervals are called the confidence intervals. They can measure the credibility of the measured data. The difference between the measured data and their estimated values are called the residuals. The confidence intervals of the residuals are used to judge the bad data. Bad data detection can be easily achieved by calling the Regress function in MATLAB software. By plotting the residual confidence intervals, of all measuring points, the measuring points out of the intervals are regarded as the bad data.

### 2.5 Parameter estimation error analysis

Rewrite (8) into a matrix form (9):

$$(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{x}' = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{t} \tag{9}$$

where  $\mathbf{x}'$  is the estimated solution of the unknown variables, i.e. the estimated values.

1) The mean of the estimated error of the unknown  $\mathbf{x}'$  is:

$$\begin{aligned} E(\Delta \mathbf{x}) &= E(\mathbf{x} - \mathbf{x}') \\ &= E(\mathbf{x} - (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{t}) \\ &= E((\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{t})) \\ &= -(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} E(\delta) \end{aligned} \tag{10}$$

Since the mean of the measurement error  $\delta$  is usually zero (called unbiased, i.e.  $E(\delta) = 0$ ), the mean value of the estimation error of the unknown variables is zero, i.e.  $E(\Delta \mathbf{x}) = 0$ , this equation shows that the estimation of the unknown variables is unbiased .

2) In engineering, the covariance matrix of the estimated error is often used to measure the difference between an estimated value and its true value. The covariance matrix is shown in (11):

$$\begin{aligned} \mathbf{c} &= E(\Delta \mathbf{x} (\Delta \mathbf{x})^T) \\ &= (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} E(\delta \delta^T) \mathbf{R}^{-1} \mathbf{H} (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \end{aligned} \tag{11}$$

To simplify (11), let  $\mathbf{A} = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})$  and name it the information matrix. Consider  $E(\delta \delta^T) = \mathbf{R}$  ( $\mathbf{R}$  is the measurement error variance matrix). Then (11) is simplified to (12):

$$\begin{aligned} \mathbf{c} &= \mathbf{A}^{-1} \mathbf{H}^T \mathbf{R}^{-1} E(\delta \delta^T) \mathbf{R}^{-1} \mathbf{H} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{R} \mathbf{R}^{-1} \mathbf{H} \mathbf{A}^{-1} \\ &= \mathbf{A}^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \end{aligned} \tag{12}$$

The diagonal elements of  $\mathbf{A}^{-1}$  decrease with the increase of the number of the measurement quantity, which means that the more the measurement quantity, the more accurate the estimated values. This proves that the premise of better parameter estimation is under a high enough redundancy.

### 2.6 Feasibility analysis of length coefficient $k$

In the conventional methods, the length coefficient  $k$  is introduced to reduce the fault locating error caused by the sag change and other factors, which requires that the theoretical line length should not appear in the calculating equation. Compared with the conventional methods, the transmission equations in this paper still contains the theoretical line length, which seems irrationality. The following is used to show the rationality of this methodology.

When a short-circuit fault occurs in a transmission line  $i$ - $j$ , ignoring the sag effect, the transmission equation of the measuring point  $p$  is (13). However, considering the sag effect, the true transmission equation of the measuring point  $p$  is (14).

$$t_{Ap} = \frac{l_{1pi}}{v_1} + \frac{l_{2pj}}{v_2} + \frac{kl_{ij}}{v_1} + t_0 \tag{13}$$

$$\begin{aligned} t'_{Ap} &= \frac{\alpha l_{1pi}}{v_1} + \frac{\beta l_{2pj}}{v_2} + \frac{k(\alpha l_{ij})}{v_1} + t_0 \\ &= \frac{l_{1pi}}{v_1/\alpha} + \frac{l_{2pj}}{v_2/\beta} + \frac{kl_{ij}}{v_1/\alpha} + t_0 \\ &= \frac{l_{1pi}}{v'_1} + \frac{l_{2pj}}{v'_2} + \frac{kl_{ij}}{v'_1} + t_0 \end{aligned} \tag{14}$$

where  $\alpha$  and  $\beta$  are the telescopic factors of overhead lines and cables under the sag effect.

The sag effect causes  $t_{Ap}$  and  $t'_{Ap}$  to be unequal, but the parameter estimation in this research still uses the theoretical length of lines, which results in calculation bias. Comparing (13) and (14), it can be seen that the sag effect changes the calculated value of the traveling wave velocity without affecting the length factor  $k$  and the fault time moment  $t_0$ . In other words, this method obtains the exact  $k$  at the expense of imprecise wave velocity results.

### 2.7 Measuring point configuration in a line

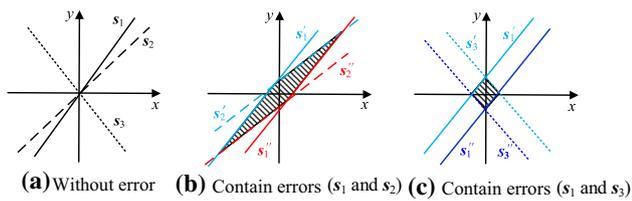
Traveling wave measuring points are generally installed at the terminal. For a wind farm transmission line, each wind turbine unit can be regarded as a terminal. However, the distance between adjacent terminals is extremely short in the line. If each terminal is equipped with a measuring point, it is not only not feasible, but also generates a large amount of similar data and amplification errors. Therefore, we must rationally configure the measuring points according to the structure of the line. That means we need to establish which measuring points are most favorable for improving the estimation accuracy or which measuring points can be added to achieve the best estimation effect based on the existing measuring points. Firstly, taking the two-point traveling wave transmission equations as an example, the analysis is as follows.

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \alpha & \beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \tag{15}$$

where  $t_1$  and  $t_2$  are the first wavefront arrival time moments of any two measuring points,  $\alpha$  and  $\beta$  are the elements in the coefficient matrix  $\mathbf{H}$ , specifying  $\alpha \neq \beta$ ,  $x_1$  and  $x_2$  are unknown variables, and  $\delta_1$  and  $\delta_2$  are measurement errors of the two points:

Solving (15), we can get (16):





**Fig. 3** Measuring point error amplification vector diagram

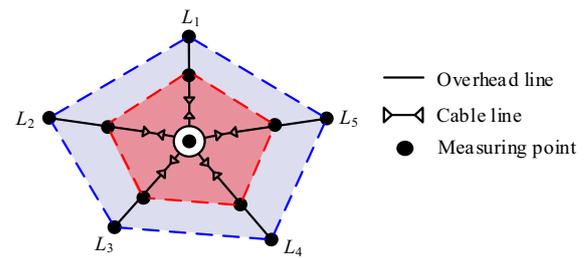
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\beta - \alpha} \begin{bmatrix} \beta & -1 \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} t_1 - \delta_1 \\ t_2 - \delta_2 \end{bmatrix} \quad (16)$$

From (16), it can be seen that the difference between  $\alpha$  and  $\beta$  is small when  $\beta$  is close to  $\alpha$ , causing the reciprocal of the difference to be large. This leads to an error amplification of the measuring point information  $t$ , which in turn increases the error of  $x$ . If we analyze (16) from a vector perspective, let  $s_1 = [\beta \ -1]$ ,  $s_2 = [-\alpha \ 1]$ . When  $\beta \approx \alpha$ , then  $s_1 \approx -s_2$ , so  $s_1$  and  $s_2$  are approximate vectors, nonorthogonal vectors, which increases the error. In order to avoid this situation, it is necessary to ensure the orthogonalization of  $s_1$  and  $s_2$ . When ignoring the error  $\delta$ ,  $s_1$  and  $s_2$  are plotted in Fig. 3a. When  $\beta = -\alpha = 1$  and  $s_1 \cdot s_2^T = 1 - 1 = 0$ , although the orthogonalization of  $s_1$  and  $s_2$  is also guaranteed, there is an infinite number of solutions for (15). Therefore, it is necessary to discard one measuring point.

When the error  $\delta$  is taken into account, the cross solution of the system of equations about  $s_1$  and  $s_2$  is shown in the shaded portion shown in Fig. 3b assuming  $t = 0$ . If one of the vectors is replaced by the vector  $s_3$  in Fig. 3a, that is, the orthogonal vector, the shadow portion shown in Fig. 3c is the smallest. This shows that the points that can constitute the most orthogonal vector are preferable to be selected as the final measuring points.

Although the orthogonal vector is ideal, the actual line terminals of measuring points are not constructed according to the orthogonal vector form. Therefore, the terminal which has the vector closest to the orthogonal vector is used as a measuring point. The method to measure the similarity of two vectors can be found in [15].

According to the equivalent topological graph of the wind farm line in Fig. 1b, the wind farm can be seen as a typical radial network as the center node. The measuring point configuration scheme is trying to select the wind turbine group of each overhead line terminal and the wind turbine group closest to the middle point of the line as the traveling wave measuring point under the premise of the above mentioned orthogonal vector. Taking the wind farm formed by five collecting power lines as an example, the measuring point network shown in Fig. 4 is constructed.



**Fig. 4** Wind farm measuring point configuration diagram

### 3 Fault interval determination

The fault locating method based on the parameter estimation needs to first determine the interval where the fault is located in the red area or in the blue area of the network shown in Fig. 4.

#### 3.1 Fault interval criterion

First, the theoretical time required for the first wavefront transmission in each section can be obtained and a theoretical time difference matrix  $\Delta t$  is constructed according to Fig. 4.  $\Delta t$  is a matrix of  $3 \times r$  order, where  $r$  is the number of collecting power lines in the wind farm. The matrix consists of the following parts: ① The elements of the first row are the theoretical traveling wave time differences between the terminal measuring points and the central measuring points; ② The elements of the second row are the theoretical traveling wave time differences between the middle measuring points and the central measuring points; ③ The elements of the third row are the time difference between the terminal measuring point and the middle measuring point.

Then, the actual time difference matrix  $\Delta t'$  is established according to the actual measured time at each measuring point and the difference matrix  $A$  is obtained as the difference between  $\Delta t$  and  $\Delta t'$ . Since  $\Delta t$  uses the theoretical wave velocity, the value of the differential matrix  $A$  in the non-faulty interval is not zero and there will be a slight deviation. Therefore, it is necessary to set a threshold (the value of which is determined based on the maximum locating error based on the theoretical wave velocity) so that the element is set to zero when the value of the difference matrix element is lower than the threshold. The first row of the differential matrix  $A$  detects on which collecting power line a fault has occurred, the second row detects whether a fault occurs in the first half of the collecting power line, and the third row detects whether a fault occurs in the second half of the collecting power line.

We use (17) as an example to illustrate the fault interval criterion, where  $L_1$ – $L_5$  represent the five collecting power lines in Fig. 4. Assuming a short-circuit fault occurs in the

second half of line  $L_4$ , the resulting differential matrix is as follows:

$$A = \begin{bmatrix} L_1 & L_2 & L_3 & L_4 & L_5 \\ 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 \end{bmatrix} \quad (17)$$

In the first row of  $A$ , the nonzero element is only the element corresponding to  $L_4$ , which indicates that short-circuit fault occurs in  $L_4$  and the corresponding element in the second row  $L_4$  is zero, indicating that no short-circuit fault occurs in the first half of the  $L_4$ . The  $L_4$  corresponding element in the third row is not zero, indicating that a short-circuit fault occurred in the second half of the  $L_4$ .

The above is the general case of the line fault. However, when a fault occurs near a measuring point, its differential matrix will become (18).

$$A = \begin{bmatrix} L_1 & L_2 & L_3 & L_4 & L_5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (18)$$

At this point, it is not possible to determine where the failure occurred based on (18).

### 3.2 Criterion for failure near nodes

The measuring points installed at the nodes can be divided into three types: terminal measuring point, middle measuring point, and central measuring point which is at bus node. Among them, since the terminal measuring point can be intuitively judged by the first wavefront arrival time, this kind of fault situation is easy to judge and need not to be considered in this research. For the middle measuring points (non-terminal points), considering of the simplicity of fault interval criterion, the transmission equation of the measuring point is artificially deleted to eliminate the possibility of misjudgment when a short-circuit fault occurs near a middle node. When a fault occurs near the center node, the criterion cannot be directly obtained because it connects a plurality of the collecting power lines. However, the criterion can be obtained by combining this with the parameter estimation algorithm. Assuming that the short-circuit fault occurs on each of the collecting power lines of wind farm bus in turn, the parameter estimation method is applied to the analyzed line to solve the corresponding locating results.

The specific operation is to use this algorithm to solve the corresponding locating results for the five collecting power lines connected to the center node in Fig. 4,

**Table 2** Fault interval criterion for fault around central node

No	Range of $k$				
	$L_1$	$L_2$	$L_3$	$L_4$	$L_5$
1	$(0, r/l_1)$	$\times$	$\times$	$\times$	$\times$
2	$\times$	$(0, r/l_2)$	$\times$	$\times$	$\times$
3	$\times$	$\times$	$(0, r/l_3)$	$\times$	$\times$
4	$\times$	$\times$	$\times$	$(0, r/l_4)$	$\times$
5	$\times$	$\times$	$\times$	$\times$	$(0, r/l_5)$

Note:  $\times$  denotes an abnormal value (locating error),  $k < 0$  or  $k > 1$

respectively. We assume that the radius of the fault-tolerant interval is  $r$ . The 1, 2, 3, 4, and 5 in Table 2 represent the locating results using the parameter estimation algorithm when assuming that lines  $L_1, L_2, L_3, L_4$ , and  $L_5$  are faulty lines, respectively. The  $k$  of the locating result should meet the fault-tolerant interval in Table 2, where each column represents the criterion for assuming the corresponding line fault.

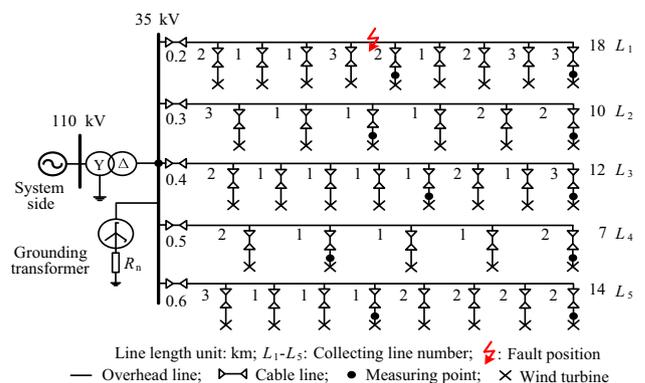
Summarizing the contents of Section 2 and Section 3, we can get the algorithm flowchart of the proposed method. See Fig. A1 of Appendix A.

## 4 Simulation result and discussion

### 4.1 Simulation model

A wind farm with five collecting power lines was used as a simulation example and its specific topology is shown in Fig. 5, where the position of each traveling wave measuring point has been marked.

We install two measuring points on each collecting power line. In addition, these measuring points are numbered from left to right and from top to bottom according to the transmission line number. Furthermore, a measuring



**Fig. 5** Wind farm simulation model

point is installed on the 35 kV bus. In summary, the entire simulation system contains 11 measuring points. We verify the feasibility of the algorithm in the following aspects: ① Demonstration of fault locating algorithm; ② Identification of bad data in measured data; ③ Fault locating results in different fault positions; ④ Identify the fault interval when a fault occurs near the node (or traveling wave measuring point).

Given that most faults in the grid are LG faults, LG faults are more representative. However, the fault locating calculation process of other faults, such as LL faults, is the same as for the LG fault. In addition, results for other fault types are similar. Due to space limitations, we here only present results for the LG fault.

### 4.2 Fault locating simulation and results analysis

#### 4.2.1 Algorithm demonstration

##### 1) Interval identification

In the simulation, a single-phase earth fault occurred on the  $L_1$ . The fault position was 8 km from the head end (cable end) and the fault occurred at 0.03 s. The first wavefront arrival time to each measuring point is shown in Table 3.

The theoretical time difference matrix  $\Delta t$  (in  $\mu s$ ) is :

$$\Delta t = \begin{bmatrix} 63.48 & 37.30 & 45.19 & 29.24 & 54.16 \\ 32.82 & 20.27 & 24.75 & 16.69 & 26.91 \\ 30.66 & 17.03 & 20.44 & 13.63 & 27.25 \end{bmatrix}$$

The actual measured time difference matrix  $\Delta t'$  (in  $\mu s$ ) for each measuring point is:

$$\Delta t' = \begin{bmatrix} 4.00 & 38.00 & 46.00 & 30.00 & 54.00 \\ 25.00 & 20.00 & 24.00 & 16.00 & 28.00 \\ 29.00 & 18.00 & 22.00 & 14.00 & 26.00 \end{bmatrix}$$

Given the difference between the two formulas and after setting the threshold, the difference matrix  $\Delta$  (in  $\mu s$ ) is:

$$\Delta = \begin{bmatrix} 59.42 & 0 & 0 & 0 & 0 \\ 7.82 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It can be seen from the first row elements of the differential matrix that the fault occurs on the  $L_1$ . From the second row and the third-row elements, it can be seen that the fault occurs in the first half of the  $L_1$ .

##### 2) Fault location

We write the transmission equation of the measuring point according to the measuring data. Taking the No. 3 measuring point (the left measuring point of the  $L_2$ ) as an example, the transmission equation is:

$$0.030048 = 5 \times \frac{1}{v_1} + 0.5 \times \frac{1}{v_2} + 9 \times \frac{k}{v} + t_0$$

The transmission equations for all the measuring points are combined and the result of solving (6) using the parameter estimation theory is:  $x = [3.359, 1.224, 2.969, 29999] \times 10^{-6}$ ,  $k = x_{(3)}/x_{(1)} = 0.88374$ . From the confidence interval of sample residuals in Fig. 6, no bad data is found in all measuring points. The final locating result is 7.9537 km and the error is 46.3 m.

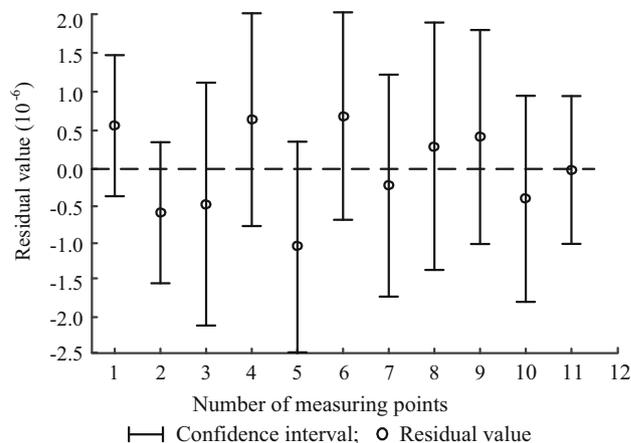
#### 4.2.2 Identification of bad data in measured data

We use the simulation data (i.e. the data in Table 1) in Section 4.2.1 to verify the identification of bad data. Among them, the value (30.028 ms) of the measuring point No. 11 is changed to 30.036 ms as a bad data, and its corresponding parameter estimation result is  $l_{ij-f} = 9.4875$  km. However, the true value is  $l_{ij-f} = 8$  km. It indicates that the locating result is wrong when there is bad data. We plot the residual confidence intervals for each measuring point in Fig. 7. It can be seen from Fig. 7 that the data of the point No. 11 is bad data. After removing the

**Table 3** Traveling wave point information

MP	$t$ (ms)	MP	$t$ (ms)
1	30.003	7	30.044
2	30.032	8	30.058
3	30.048	9	30.056
4	30.066	10	30.082
5	30.052	11	30.028
6	30.074		

Note: MP stands for the number of measuring points



**Fig. 6** Confidence intervals of measuring points residuals

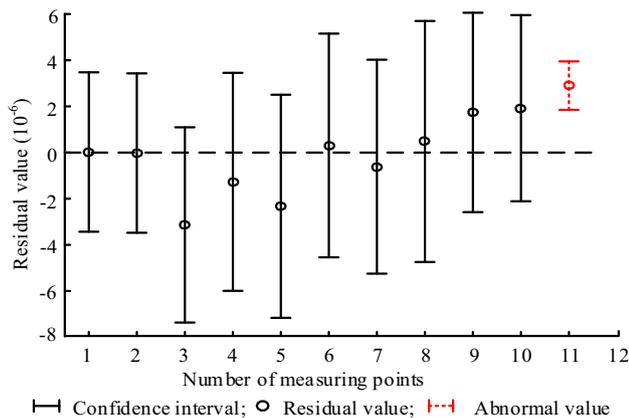


Fig. 7 Confidence intervals for measuring points residuals with bad data

bad data, the estimated result of the second parameter estimation is:  $l_{ij-f} = 7.9531$  km and the error is 46.1 m.

The above results show that the algorithm can identify bad data in the measurement data. In addition, it can still correctly locate the fault point and have high locating accuracy after rejecting bad data.

4.2.3 Fault locating results for different fault positions

Short-circuit faults are set at different positions on the  $L_1$ . The locating results obtained by the parameter estimation algorithm are shown in Fig. 8. In addition, in order to further verify the locating accuracy of this algorithm, the locating results of the dual-ended traveling wave method with higher accuracy are also displayed together in Fig. 8. It can be seen that the locating accuracy of this algorithm for other fault positions is higher than that of the dual-

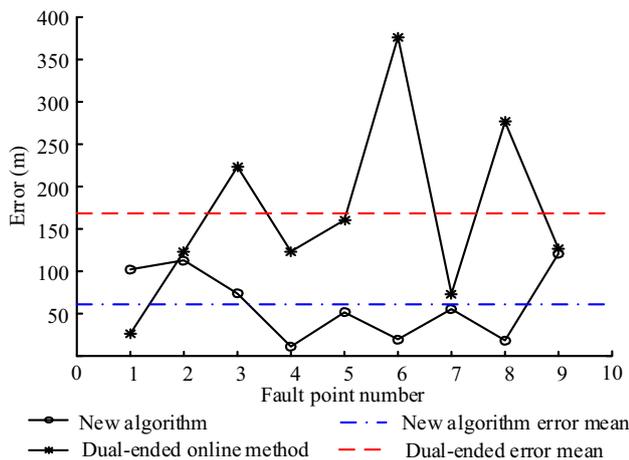


Fig. 8 Error comparison of fault locating results of different algorithms

Table 4 The  $k$  of the fault interval criterion in  $L_1-L_5$  and the locating result when a fault occurs on the  $L_1$

No	Range of $k$					Calculated $k$ of $L_1-L_5$
	$L_1$	$L_2$	$L_3$	$L_4$	$L_5$	
1	(0, 0.11)	X	X	X	X	0.070
2	X	(0, 0.20)	X	X	X	-0.689
3	X	X	(0, 0.17)	X	X	-0.499
4	XX	X	X	(0, 0.33)	X	-1.222
5	X	X	XX	X	(0, 0.17)	-0.819

Note:  $\hat{u}$  denotes an abnormal value (wrong location),  $k < 0$  or  $k > 1$

ended method in most cases. The overall locating accuracy is also superior to the dual-ended traveling wave method.

4.2.4 Fault section identification when a fault occurs near the center node

Suppose a short-circuit fault occurred on the  $L_1$  at a distance of 0.8 km from the side of the 35 kV bus, and the fault type was a single-phase earth fault. The calculation results of the “interval criterion for failure near the node” are shown in Table 4. In addition, the last column in Table 4 shows the calculated  $k$  of assuming faults on all lines  $L_1-L_5$  in turn under the real fault occurring on the  $L_1$ . According to the displayed results in Table 4, the calculated  $k$  (0.070) given fault on  $L_1$  conforms to the fault interval criterion of  $L_1$  because 0.07 is within (0,0.11), which indicates that the real fault branch is  $L_1$ . This example shows the correctness of the fault interval criterion algorithm.

4.3 Fault locating under sag effect

Reference [16] pointed out that the length of the overhead transmission line varies from 90% to 110% in different seasons. Therefore, the length variation of the overhead line in the simulation setting is  $\pm 10\%$  and the cable line does not change. In addition, when the length of the line changes, the position of the fault point also needs to be changed. The fault locating results of the algorithm of this paper are compared with those of the dual-ended traveling wave method in Table 5.

Table 5 shows that if the line sag effect is considered, the locating errors are not substantially affected, whereas the conventional dual-ended traveling wave method has a large error.



**Table 5** Fault locating results under sag effect of lines

Sag factor (%)	Real value $k$	Failure point (km)	$L_1$ (km)	Method of this paper $k$	Method error in this paper (km)	Dual-ended method (km)	Dual-ended method error (km)
+10	0.25	5.5	19.8	0.244	0.117	5.031	0.531
	0.50	11.0		0.503	0.056	9.960	0.964
	0.75	16.5		0.751	0.011	12.189	1.311
-10	0.25	4.5	16.2	0.252	0.038	5.093	0.593
	0.50	9.0		0.506	0.106	7.858	1.142
	0.75	13.5		0.756	0.108	14.539	1.039

## 5 Conclusion

This paper presents a fault locating method for hybrid transmission lines in a wind farm based on parameter estimation theory. Through a large number of simulations, the following conclusions have been reached:

- 1) The proposed transmission equations and the optimization of the redundancy parameter estimation can ensure good fault locating accuracy for short lines and hybrid lines consisting of overhead wires and cables in a wind farm. In addition, the bad data detecting capability of the method enhances the reliability of the measured data. Through PSCAD/EMTDC simulations, the feasibility of this method in wind farm fault locating has been verified.
- 2) The method does not require the traveling wave velocity. This avoids locating errors caused by the calculating errors of the traveling wave velocity.
- 3) In order to solve the problem of misjudgment caused by faults near the node, a criterion for effectively determining the faulty interval is proposed. The feasibility of the algorithm is verified by PSCAD/EMTDC simulations.
- 4) The introduction of the length coefficient  $k$  solves the problem of poor locating accuracy or misjudgment caused by line sag effects and the correctness of the algorithm has also been verified through simulations.
- 5) The comparison of fault locating results between this method and the dual-ended traveling wave method show that this method is more feasible in a multi-branch line.
- 6) Although this paper studies the fault locating of transmission lines in a wind farm, this method is also applicable to a distribution grids with lines of multi-branch that may be a mix of overhead wires and cables. Because of paper space limitation, the relative simulations and their locating results are not described.

**Acknowledgements** This work was supported in part by National Natural Science Foundation of China (No. 51677072).

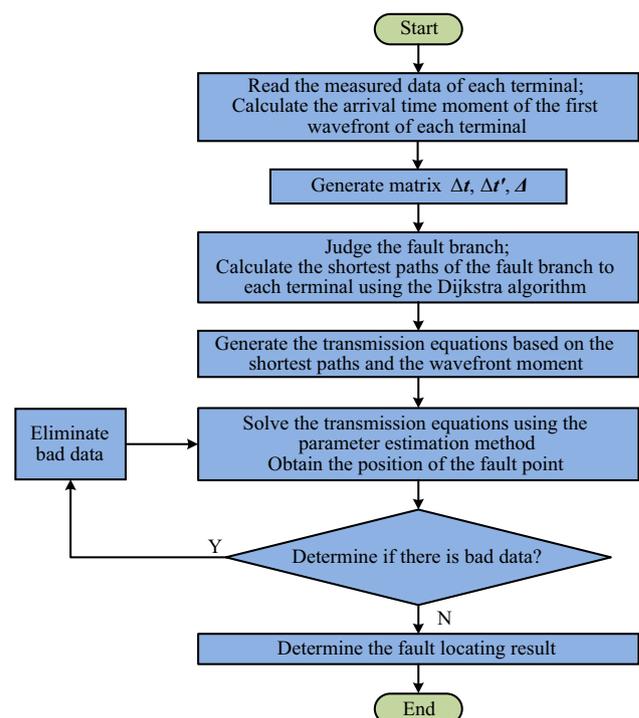
**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix A

**Table A1** Locating results comparison of ignoring and considering cables

Case	Ignoring cable	Cable equivalence	Considering cable
$x$ (km)	8.6877	8.4707	7.9537

Note: Actual fault point at 8.0 km

**Fig. A1** Flowchart of fault locating

## References

- [1] Li YW, Nejabatkhah F (2014) Overview of control, integration and energy management of microgrids. *J Mod Power Syst Clean Energy* 2(3):212–222
- [2] Xu HH, Hui ZB, Lai LZ (2002) A novel principle of single-ended fault location technique for EHV. *IEEE Power Eng Rev* 22(11):61
- [3] Weng J, Liu D, Luo N et al (2015) Distributed processing based fault location, isolation, and service restoration method for active distribution network. *J Mod Power Syst Clean Energy* 3(4):494–503
- [4] Xu F, Dong X, Wang B et al (2015) Self-adapted single-ended travelling wave fault location algorithm considering transfer characteristics of the secondary circuit. *IET Gener Transm Distrib* 9(14):1913–1921
- [5] Lee JW, Kim WK, Han J et al (2016) Fault area estimation using traveling wave for wide area protection. *J Mod Power Syst Clean Energy* 4(3):478–486
- [6] Azizi S, Sanaye-Pasand M, Abedini M et al (2014) A traveling wave-based methodology for wide-area fault location in multi-terminal DC systems. *IEEE Trans Power Deliv* 29(6):2552–2560
- [7] Esau Z, Jayaweera D (2014) Reliability assessment in active distribution networks with detailed effects of PV systems. *J Mod Power Syst Clean Energy* 2(1):59–68
- [8] Jafarian P, Sanaye-Pasand M (2010) A traveling wave based protection technique using wavelet/PCA analysis. *IEEE Trans Power Deliv* 25(2):588–599
- [9] Das S, Santoso S, Gaikwad A et al (2014) Impedance-based fault location in transmission networks: theory and application. *IEEE Access* 2:537–557
- [10] Phadke AG, Wall P, Ding L et al (2016) Improving the performance of power system protection using wide area monitoring systems. *J Mod Power Syst Clean Energy* 4(3):319–331
- [11] Mora-Flórez J, Cormane-Angarita J, Ordóñez-Plata G (2009)  $k$ -means algorithm and mixture distributions for locating faults in power systems. *Electr Power Syst Res* 79(5):714–721
- [12] Liu Y, Meliopoulos APS, Tan Z et al (2017) Dynamic state estimation-based fault locating on transmission lines. *IET Gener Transm Distrib* 11(17):4184–4192
- [13] Wu ZJ, Xu JJ, Yu XH et al (2017) Review on state estimation technique of active distribution network. *Autom Electr Power Syst* 41(13):182–191
- [14] Wang Y, Chao LU, Zhu L et al (2016) Comprehensive modeling and parameter identification of wind farms based on wide-area measurement systems. *J Mod Power Syst Clean Energy* 4(3):383–393
- [15] Zhang Y, Liu YD, Ji Z (2009) Vector similarity measurement method. *Techn Acoust* 28(4):532–536
- [16] Suonan JL, Zhang YN, Jun Q et al (2006) Time domain fault location method based on transmission line parameter identification using two terminals data. *Power Syst Technol* 30(8):65–70

**Ke ZHANG** received B.E. degree in 2016, from North China Electric Power University, China. He is now a graduate student studying at North China Electric Power University, China. His research interests include fault analysis and fault location of the power system.

**Yongli ZHU** received Ph.D. degree in 1992, from North China Electric Power University, China. He is now a professor of North China Electric Power University, China. His research interests include power system security defense and recovery control, power system network monitoring and intelligent information processing.

**Xuechun LIU** received B.E. degree in 2017, from Xi'an University of Technology, China. She is now a graduate student studying at North China Electric Power University, China. Her research interests include power system fault analysis and unit combination.

