Precise Lightning Strike Detection in Overhead Lines Using KL-VMD and PE-SGMD Innovations

Dong Xinsheng *, Liu Jucheng , He Shan , Han Lu , Dong Zhongkai , Cai Minbo

Posted Date: 9 January 2024

doi: 10.20944/preprints202401.0706.v1

Keywords: Keywords: Traveling wave; KL-VMD; PE-SGMD; Lightning strike; Zero mode voltage

Article

Precise Lightning Strike Detection in Overhead Lines Using KL-VMD and PE-SGMD Innovations

Dong Xinsheng [1,*], Liu Jucheng [2], He Shan [2,3], Han Lu [2,3], Dong Zhongkai [4] and Cai Minbo [5]

[1] State Grid Xinjiang Electric Power Research Institute, Urumqi, Xinjiang Uygur Autonomous Region 830011, China; dongxinsheng56@foxmail.com
[2] Xinjiang University Autonomous Region Key Laboratory of Renewable Energy Power Generation and Grid Connected Technology, Urumqi, Xinjiang Uygur Autonomous Region 830017, China; heshanxju@sina.com; attentioncheng@163.com; xjdhl@xju.edu.cn
[3] Engineering Research Center of Ministry of Education for Renewable Energy Power Generation and Grid Connection Control, Urumqi, Xinjiang Uygur Autonomous Region 830017, China; heshanxju@sina.com; xjdhl@xju.edu.cn
[4] State Grid Turpan Power Supply Company, Turpan, Xinjiang Uygur Autonomous Region 838000, China; dzklndj@sina.com
[5] State Grid Altay Power Supply Company, Altay, Xinjiang Uygur Autonomous Region 836500, China; xjalthc@sina.com
* Correspondence: dongxinsheng56@foxmail.com

Abstract: When overhead lines are impacted by lightning, the traveling wave of the fault contains a wealth of fault information. The accurate extraction of feature quantities from transient components and their classification is fundamental to the identification of lightning faults. The extraction process may involve modal aliasing, optimal wavelet base issues, and inconsistencies between the lightning strike distance and the fault point, which have the potential to impact the effectiveness of recognition. This paper presents a method for identifying lightning strike faults by utilizing Kullback-Leibler (KL) divergence enhanced Variational Mode Decomposition (VMD) and Symmetric Geometry Mode Decomposition (SGMD) improved with Permutation Entropy (PE) to address the aforementioned issues. A model of a 220 kV overhead line is constructed using real faults to replicate scenarios of winding strike, counterstrike, and short-circuit. The three-phase voltage is chosen and then subjected to Karenbaren decoupling in order to transform it into zero mode, line mode 1, and line mode 2. The zero mode voltage is decomposed using KL-VMD and PE-SGMD methods, and lightning identification criteria are developed based on various transient energy ratios. The research findings demonstrate that the criteria effectively differentiate between winding strike, counterstrike, and short-circuit faults, thus confirming the accuracy and efficacy of the lightning fault identification criteria utilizing KL-VMD and PE-SGMD.

Keywords: traveling wave; KL-VMD; PE-SGMD; Lightning strike; Zero mode voltage

1. Introduction

High-voltage overhead lines serve as the primary conduits for the transmission of electric power and constitute a crucial component of the power grid [1,2]. These entities are situated in expansive regions with diverse natural and geographical features, rendering them particularly vulnerable to lightning strikes. More than 50% of power system accidents in Japan result from lightning strikes on transmission lines. In nations such as the former Soviet Union and the United States, 60% of incidents involving 275-500 kV overhead lines are attributed to lightning strikes. In China, lightning strikes account for 40% to 70% of all trip accidents [3]. The transient stage of the fault contains abundant fault information [4]. Accurately detecting signal singularities and extracting feature quantities from complex transient components, and promptly and precisely identifying the type of fault, are crucial for ensuring the stable operation of the power system.

Concurrently, a significant body of research has been dedicated to the development of feature extraction and identification methods for lightning strikes on transmission lines. Identification

 doi:10.20944/preprints202401.0706.v1

methods primarily rely on two aspects: the high-frequency transient components and the amplitude of lightning. Depending on the approach to feature extraction, the primary methods for identifying lightning strikes include the amplitude method, time-domain waveform method, modal extremum method, wavelet energy spectrum method, and high-frequency attenuation ratio method.

Zou et al. [5] suggests that the mutation points of transient traveling waves correspond one-to-one with the wavelet transform modulus maxima points. The strength of the traveling wave changes is characterized by the magnitude of the wavelet modulus, whereas the direction of signal change is indicated by the polarity of the wavelet changes. Consequently, the wavelet transform modulus maxima can be utilized for the identification of the type of lightning strike. Shu et al. [6] employs the ratio of the maximum initial modulus to the maximum second modulus of transient traveling waves to distinguish between non-fault lightning strike current waves that are not truncated by the fault point and fault traveling waves that are truncated by the grounding point. This approach aids in the identification of non-fault and fault lightning strikes. Sima et al. [7] utilizes wavelet transform modulus maxima for the identification of the polarity of initial transient voltage surges and the duration of surge arrester discharge. This is achieved by analyzing the polarity of wavelet modulus maxima and discharge time to detect winding strike and counterstrike faults. Guo et al. [8] differentiates between short-circuit and lightning faults by analyzing their high-frequency distribution. This is achieved by using the modulus maxima of the detail coefficients obtained through wavelet transform as criteria. These identification methods [5-8] exhibit high reliability and remain unaffected by the initial phase angle of the fault, the type of fault, and the transient fault resistance. However, the identification of modulus maxima presents a challenge and lacks a standardized approach.

Zhong et al. [9] utilizes mathematical morphology for the analysis of variations in fault waveforms. It compares the clustering effects of single-scale and multi-scale fractal curves to classify and identify lightning faults. Si et al. [10] employs a linear regression to model the disturbed traveling wave by analyzing the variations in the waveform of the transient traveling wave over time. The fit is satisfactory for cases of faults but inconsistent for non-fault lightning strikes. The waveform's consistency coefficient is utilized for assessing the occurrence of a fault on the line; however, the establishment of its threshold value lacks theoretical justification. Gao et al. [11] employs wavelet transform for multi-scale decomposition and utilizes the proportion of the transient signal energy frequency band characteristic values as identification criteria. The wavelet transform encounters challenges in selecting the optimal wavelet basis and is vulnerable to noise, which can result in potentially inaccurate outcomes. Chen et al. [12] employs the Hilbert-Huang Transform method to classify lightning types by analyzing the energy distribution across various frequency bands. This approach comprises two essential stages: Empirical Mode Decomposition (EMD) and Hilbert Transform. However, EMD decomposition process is subject to phenomena such as over-enveloping, under-enveloping, and modal mixing. Zhao et al. [13] integrates Fluorescence-Imaged MicroDeformation (FIMD) and Hilbert Transform for the detection of line traveling waves, however, the signal decomposition process is susceptible to modal mixing phenomena. Han et al. [14] differentiates non-fault lightning strikes from short-circuit faults by assessing whether the product of the time difference between the first two wavefronts at the reference side bus and the wave speed exceeds the total length of the line. However, this approach does not account for situations in which the short-circuit point or lightning strike point is located at the midpoint of the line.

In response to the identified limitations and the requirement for manual configuration of the modal components k and penalty factor α [15] in Variational Mode Decomposition (VMD) process, VMD has been enhanced through the utilization of Kullback-Leibler (KL) divergence. The optimized VMD algorithm not only effectively mitigates modal mixing phenomena but also exhibits improved resilience to noise.

This study examines the zero-mode voltage of 220 kV overhead lines under the influence of winding strike, counterstrike, and short-circuit faults following phase-mode transformation decoupling. The analysis employs KL-VMD and Permutation Entropy-Symmetric Geometry Mode Decomposition (PE-SGMD). In light of this, a proposed criterion is capable of distinguishing between

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

winding strike, counterstrike, and short-circuit faults. The criterion has been validated using MATLAB to extensively identify Power Systems Computer Aided Design (PSCAD) simulation data, demonstrating its effectiveness in distinguishing different types of line faults.

2. Lightning Simulation Model

2.1. 220 kV Overhead Line Model

A specific 220 kV overhead line encounters a lightning fault, spanning a total length of 156 km and comprising 480 towers, with an average spacing of 325 m between them. The simulation model is depicted in Figure 1. The system employs a Phase model, encountering a lightning fault at a distance C from the 50 km end busbar, with dual lightning lines installed along the line.

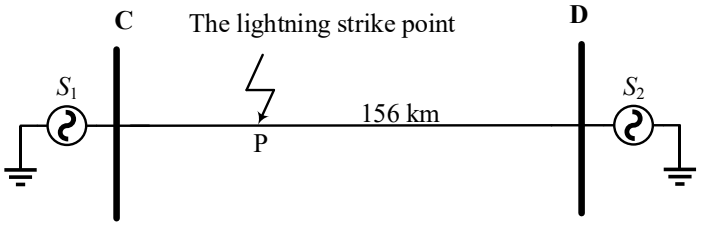

Figure 1. Simulation system model.

2.2. Lightning Current Model

Empirical data indicates that a substantial proportion, ranging from 70% to 90%, of natural lightning carries a negative charge. Therefore, a 1.2/50 μs negative polarity double exponential lightning current model is employed for analysis [16]. Equation (1) represents the calculation for lightning current, while its waveform is illustrated in Figure 2.

$$i(t) = AI_0(e^{-\beta t} - e^{-\delta t}) \tag{1}$$

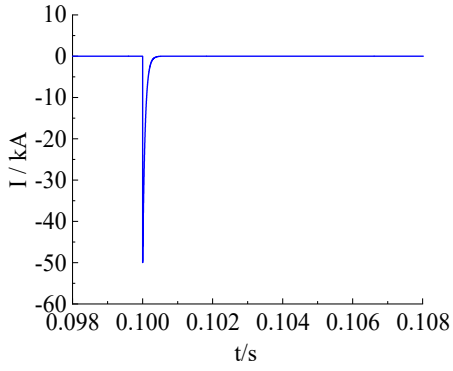

Figure 2. Lightning current waveform.

In Equation (1): A is the lightning current amplitude correction factor; I_0 is the maximum value of the lightning current; and β and δ are the corresponding waveform decay coefficients.

Figure 3 illustrates overhead lines experiencing winding strikes and counterstrikes. In this representation, I_L denotes a controlled current source, while Z_0 denotes the lightning channel's equivalent wave impedance, which is specified as 300 Ω [17].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

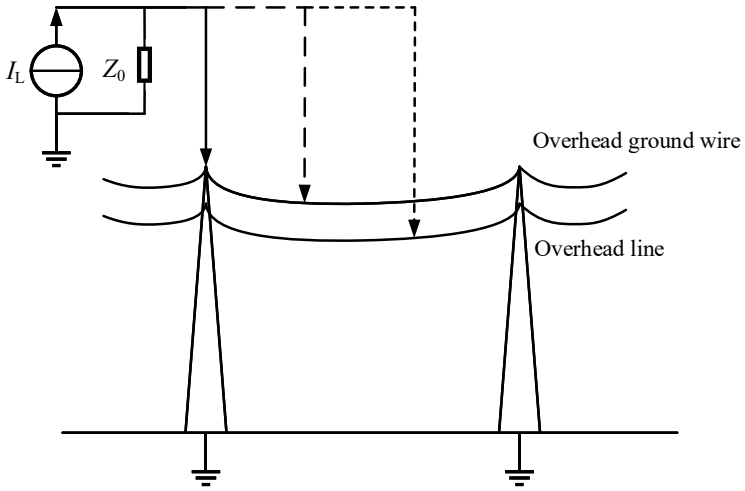

Figure 3. Schematic diagram of lightning struck line.

2.3. Tower Model

Various simulation models for transmission towers are commonly used, including the standard model, inductive model, and wave impedance model. The standard model does not account for the phenomenon of overvoltage change, while the inductive model overlooks the influence of the tower on the lightning current wave [18]. This study employs a multi-wave impedance model derived from Bergeron to replicate the behavior of 220 kV cup-type iron towers. The model accurately reproduces the refraction and reflection of lightning current waves on the tower, as depicted in Figure 4. In this context, Z_H denotes the transverse carrier wave impedance, Z_M denotes the main brace wave impedance of the tower, and R_g represents the grounding resistance, which is set at 10 Ω.

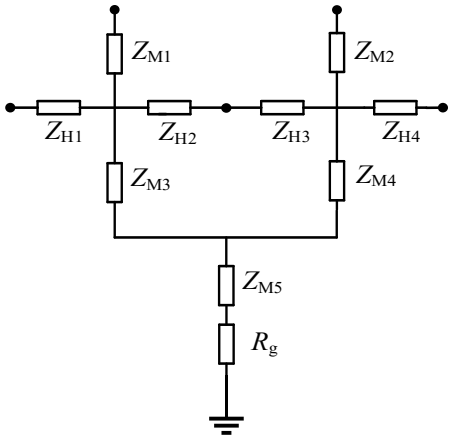

Figure 4. Multi-wave impedance model of tower.

2.4. Insulator Flashover Model

To minimize the disparity between simulation outcomes and real-world scenarios, the intersection method (involving the intersection of the overvoltage waveform curve with the volt-second characteristic curve) is selected for the characterization of insulator flashover. A capacitor paralleled with a voltage-controlled switch is used to simulate the flashover process [19], as depicted in Equation (2):

$$u(t) = 400L_X + \frac{710L_X}{t^{0.75}} \tag{2}$$

In Equation (2): L_X represents the length of the insulator string.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

3. VMD Algorithm

The parameters for VMD are frequently established using empirical values or the center frequency observation method. KL divergence is capable of indicating the similarity between two signals, and is therefore employed as the objective function for optimizing VMD algorithm. In accordance with modulation standards, IMFs are characterized as amplitude-frequency modulated signals, as depicted in Equation (3):

$$u_k(t) = A_k(t)cos[\varphi_k(t)] \tag{3}$$

In Equation (3): $\varphi_k(t)$ is the phase function and $A_k(t)$ is the instantaneous amplitude.

The original signal is decomposed into k intrinsic mode components known as Intrinsic Mode Functions (IMFs), along with the amplitude $u_k(t)$ and center frequency $\omega_k(t)$ of each IMF. The bandwidth of each modal signal is estimated through the application of Gaussian smoothness [20], and the constraint problem is expressed as follows:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \{\sum\limits_k ||\partial_t[(\delta(t)+j/\pi t)\times u_k(t)]e^{-j\omega_k t}||_2^2\} \\ s.t.\sum\limits_k u_k = f \end{cases} \tag{4}$$

In Equation (4): $\{u_k\}$ are the k decomposed modal components of the signal and $\{\omega_k\}$ are the central frequencies of each component.

To determine the optimal solution for constrained variational mode mingling, a quadratic penalty factor α and Lagrange multipliers λ are introduced, leading to the equation of an augmented Lagrangian function [21] as follows:

$$\begin{aligned} L(\{u_k\},\{\omega_k\},\lambda) = ||f(t)-\sum_k u_k(t)||_2^2 + <\lambda(t), f(t)-\sum_k u_k(t)> + \\ \alpha\sum_k ||\partial_t[(\delta(t)+j/\pi t)*u_k(t)]e^{-j\omega_k t}||_2^2 \end{aligned} \tag{5}$$

In Equation (5): * is the convolution operator. Alternation Direction Method of Multipliers (ADMM) is a technique that involves iteratively updating $\{u_k\}$ and $\{\omega_k\}$ in order to seek the Lagrange saddle point, until the iteration termination criteria are satisfied. This process is expressed as follows:

$$u_k^{n+1} = ||f(t)-\sum_k u_k(t)+\lambda(t)/2||_2^2 + \underset{u_k\in X}{argmin}\{\alpha||\partial_t[(\delta(t)+j/\pi t)\times u_k(t)]e^{-j\omega_k t}||_2^2\} \tag{6}$$

$$\omega_k^{n+1} = \underset{\omega_k\in X}{argmin}\{||\partial_t[(\delta(t)+j/\pi t)\times u_k(t)]e^{-j\omega_k t}||_2^2\} \tag{7}$$

KL divergence, also referred to as relative entropy, quantifies the difference between two probability distributions P and Q [22]. It is defined in Equation (8):

$$D_{KL}(P||Q) = \sum_i P(i)\,log\frac{P(i)}{Q(i)} \tag{8}$$

The probability density distribution of a signal is determined by applying the kernel density estimation method to the probability distribution and probability density distribution functions of two signals, as depicted in Equation (9):

$$P(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left[\frac{x_i - x}{h}\right], x\in R \tag{9}$$

$$k(u) = \frac{1}{\sqrt{2\pi}}e^{-u^2/2} \tag{10}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

In Equation (9): $P(x)$ is the probability density function and h is the window parameter, also referred to as the smoothing parameter. Equation (10) defines the Gaussian kernel function. The process by which KL divergence enhances the performance of VMD is depicted in Figure 5.

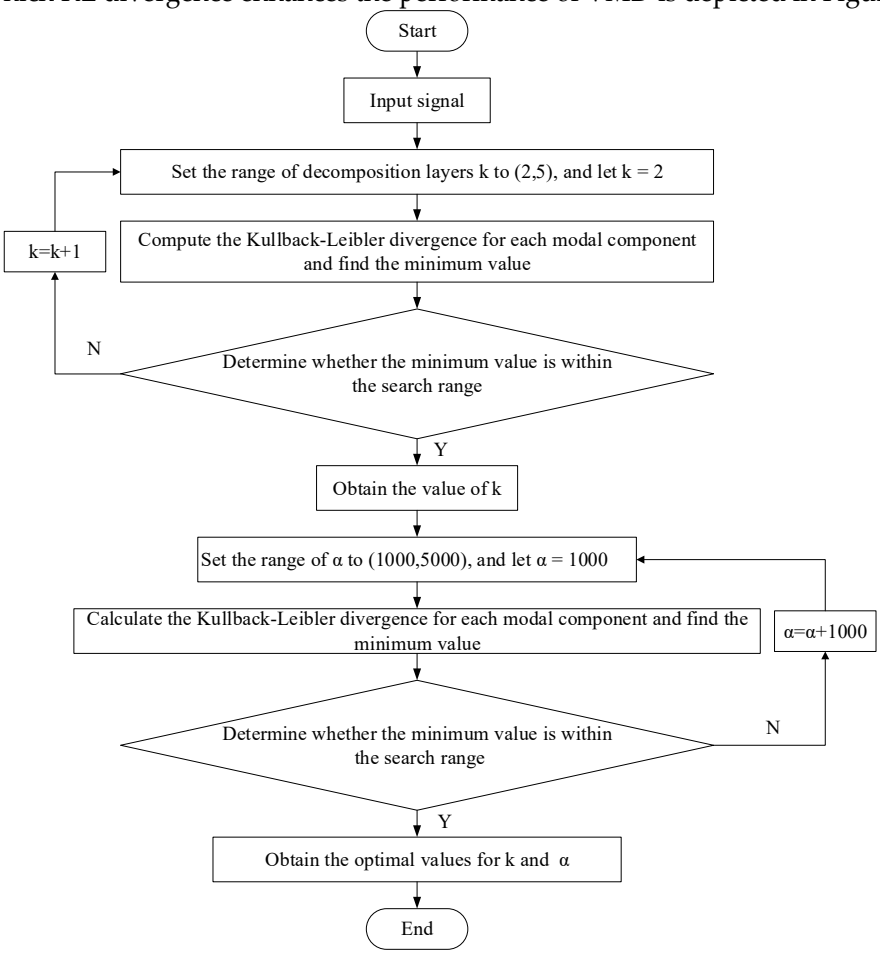

Figure 5. KL divergence optimized VMD parameter flowchart.

4. PE Algorithm

4.1. Introduction to PE

PE was introduced by Bandt et al. [23] as a technique for identifying randomness and dynamical variations in time series data. Principal Component Analysis (PCA) is recognized for its simplicity, rapid computational speed, and robust resistance to interference, rendering it especially well-suited for nonlinear data. The computational method can be outlined in Equation (11).

Upon conducting phase space reconstruction [24], the time series $\{x(i), i = 1, 2, ..., N\}$ of length N transforms into the following time series:

$$\left.\begin{aligned} X(1) &= \{x(1), x(1+\lambda), \cdots, x(1+(m-1)\lambda)\} \\ X(k) &= \{x(i), x(i+\lambda), \cdots, x(i+(m-1)\lambda)\} \\ &\vdots \\ X(N-(\text{m-1})\lambda) &= \{x(N-(m-1)\lambda), x(N-(m-2)\lambda), \cdots, x(N)\} \end{aligned}\right\} \quad (11)$$

In Equation (11): m is the embedding dimension and λ is the time delay.

The m vectors of $X(i)$, denoted as $X(i) = \{x(i), x(i+\lambda), ..., x(i+(m-1)\lambda)\}$, are rearranged in ascending order as follows:

$$X(i) = \{x(i+(j_1-1)\lambda) \le x(i+(j_2-1)\lambda) \le \cdots \le x(i+(j_m-1)\lambda\} \quad (12)$$

If $x(i+(j_{i1}-1)\lambda) = x(i+(j_{i2}-1)\lambda)$ exists, the values are arranged based on the magnitude of j, such that when $j_{k1} < j_{k2}$, it implies: $x(i+(j_a-1)\lambda \le x(i+(j_a-1)\lambda)$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

Therefore, any vector $X(i)$ has the potential to generate a set of symbolic sequences:

$$S(g) = [j_1, j_2, \cdots, j_m] \tag{13}$$

In Equation (13): g = 1, 2, ..., k, $k \leq m!$. The m symbols [j_1, j_2, ..., j_m] amount to m factorial permutations. According to various permutations [25], there are m! corresponding permutations. There are various symbolic sequences, with $S(g)$ representing just one of $m!$. The probability of occurrence for each type of symbolic sequence is calculated through $P_1, P_2, \cdots, P_k, \sum_{g=1}^{k} P_g = 1$.

PE of the time series $\{x(i), i = 1, 2, ..., N\}$ can be expressed in terms of Shannon entropy as follows [26]:

$$H_p(m) = -\sum_{g=1}^{k} P_g \, ln \, P_g \tag{14}$$

When P_g=1/$m!$, the function $H_p(m)$ reaches its maximum value at $In(m!)$. PE $H_p(m)$ can be normalized using the $In(m!)$ transformation.The expression is presented as follows:

$$H_p = H_p(m)/In(m!) \tag{15}$$

The range of H_p values falls within the interval $0 \leq H_p \leq 1$. The magnitude of H_p serves as an indicator of the level of randomness present in the time series. A higher H_p value indicates a greater degree of randomness in the time series, while a lower H_p value suggests a more predictable pattern in the time series.

4.2. Selection of PE Parameters

In the calculation of PE, it is necessary to determine three parameters: the length of the time series N, the embedding dimension m, and the time delay λ. The selection of the embedding dimension, denoted as m, typically falls within the range of 3 to 7. If m is too small, the resulting reconstructed vector will contain an insufficient number of states, thereby undermining the meaningfulness of the algorithm. Conversely, if m is too large, the reconstruction of the phase space will lead to a homogenization of the time series, resulting in increased computational demands and an inability to capture subtle changes in the series. As a result, this paper chooses m = 5. The time delay parameter λ exerts a negligible influence on the calculation of the time series. In this study, a value of λ = 1 is selected.

4.3. Symplectic Geometry Mode Decomposition

4.3.1. Phase Space Reconstruction

According to Takens' embedding theorem, a one-dimensional discrete original signal s = (x_1, x_2, ..., x_n) [27] can be topologically equivalent to the construction of a multidimensional time series matrix, denoted as the trajectory matrix X in Equation (16):

$$X = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(d-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(d-1)\tau} \\ \vdots & \vdots & & \vdots \\ x_m & x_{m+\tau} & \cdots & x_{m+(d-1)\tau} \end{bmatrix} \tag{16}$$

In Equation (16): d is the difference dimension and τ is the delay time. The trajectory matrix X is characterized by two primary parameters, d and τ. Varied values of d and τ can yield distinct matrices X, thereby exerting a substantial influence on the the analytical results.

4.3.2. Symplectic Orthogonal Matrix QR Decomposition

Let $A = X^T X$ construct M as a Hamilton matrix [28], namely:

$$M = \begin{bmatrix} A^T & 0 \\ 0 & -A \end{bmatrix} \tag{17}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

Let $N = M^2$, then both M and N are Hamilton matrices. A symplectic orthogonal matrix Q is to be constructed such that:

$$Q^T N Q = \begin{bmatrix} B & R \\ 0 & B^T \end{bmatrix} \tag{18}$$

In Equation (18): B is an upper triangular matrix, $b_{ij} = 0 (i > j+1)$. The transformation of N using the defined symplectic orthogonal type for 2 yields B, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_d$. According to the given properties, $\sigma_i = \sqrt{\lambda_i}$, where Q_i represents the eigenvectors of matrix A corresponding to σ_i. By constructing transformation coefficients $S_i = Q_i^T X^T$ using the matrix Q_i and the trajectory matrix X, and defining $Y = Q_i S_i$, the matrix $X_i = Y^T$ can be obtained. The original matrix X for phase space reconstruction comprises d components, which can be expressed as $X = X_1 + X_2 + \ldots + X_d$.

4.3.3. Diagonal Averaging

This step involves transforming each component X_k $(1 \le k \le d)$ into a time series of length n, resulting in d sets of time series, each with a length equivalent to the original time series n.

Let $X_{m \times d} = (x_{ij})_{m \times d}$, $d^* = \min(m,d)$, $m^* = \max(m,d)$, $n = m + (d-1)\tau$.

$$y_k = \begin{cases} \dfrac{1}{k} \displaystyle\sum_{p=1}^{k} y^*_{p,k-p+1}, & 1 \le k \le d^* \\ \dfrac{1}{d} \displaystyle\sum_{p=1}^{d^*} y^*_{p,k-p+1}, & d^* \le k \le m^* \\ \dfrac{1}{n-k+1} n \displaystyle\sum_{p=k-m^*+1}^{n-m^*+1} y^*_{p,k-p+1}, & m^* < k \le n \end{cases} \tag{19}$$

Equation (19) yields a time series $Y_k = (y_1, y_2, \ldots, y_n)$ based on X_i, leading to d groups of independent superimposed components in the time domain, specifically d of Symplectic Geometry Component (SGC) components.

4.4. PE Improved SGMD

Initially, an empty array z is initialized. The multiscale PE function is employed for computing PE of every row in SGC matrix, and the outcomes are saved in z array. Subsequently, the values in array z are sorted, and the resulting sorted indices are retained. The minimum and maximum values in array z are calculated and then assigned to d_1 and d_2, respectively. The value of d is determined by the equation $d = (d_2 - d_1)/4$. Four arrays, namely s_1, s_2, s_3, and s_4, are initialized and populated based on the conditions associated with the value of z. These arrays store the sums of rows of SGC matrix. As z values are iterated through, the corresponding rows of SGC are added to the respective arrays s based on the specified conditions. The arrays are stacked vertically and transposed to derive x array, which is subsequently assigned to the output variable.

5. Method for Identifying Lightning Faults

5.1. Identification of Short Circuit and Direct Strike Faults

The study primarily focuses on direct strikes as induced lightning poses minimal threat to 220 kV overhead lines. Utilizing PSCAD software for simulating the model depicted in Figure 1, single-phase short circuits, winding strikes, and counterstrike faults are configured to take place at a distance of 50 km from C end of the line. The voltage traveling waves in three phases, measured at C end, are depicted in Figure 6:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

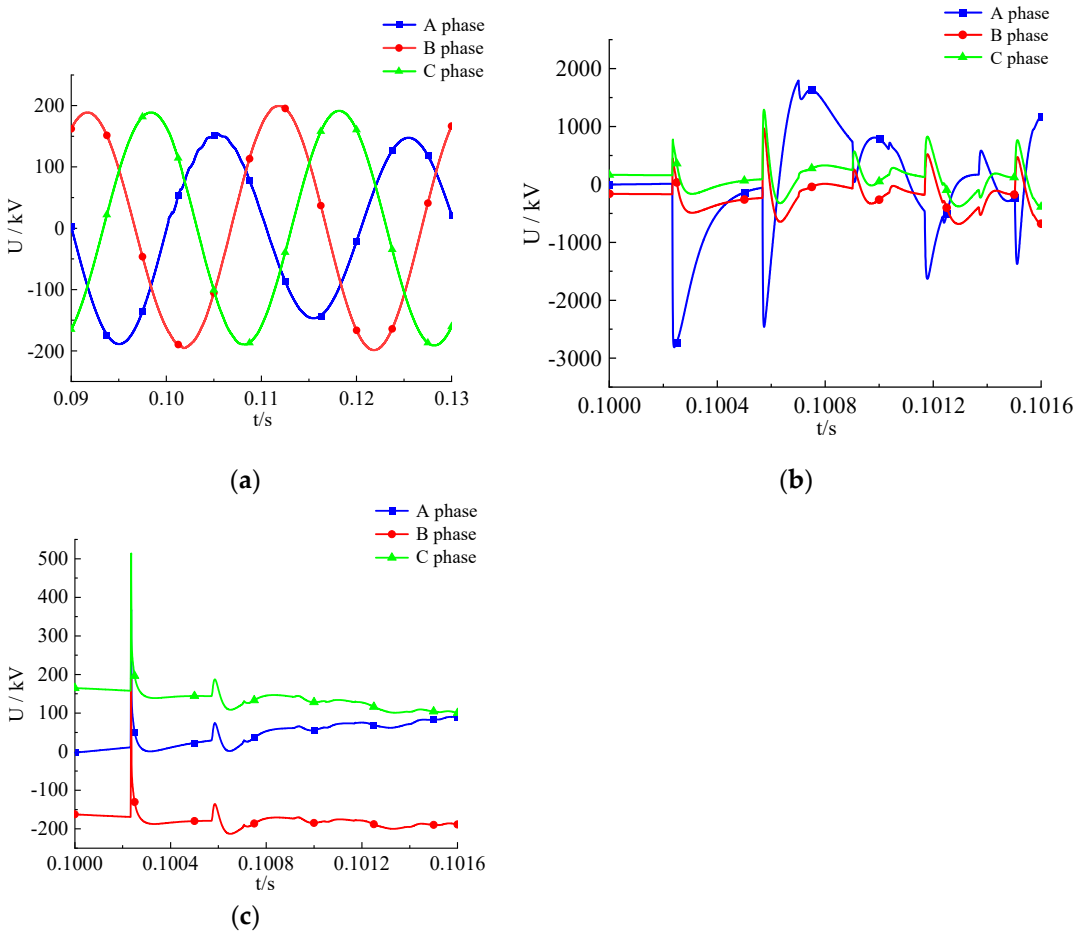

Figure 6. Three-phase voltage waveform during faults: (**a**) Short circuit fault; (**b**) Winding strike fault; (**c**) Counterstrike fault.

The simulated waveforms were observed, and the following differences were illustrated as follows:

(1) As depicted in Figure 6(a), in the event of a single-phase ground short circuit experienced by phase A, its voltage amplitude is lower than that of phases B and C, with a difference of approximately 100 kV. The voltage value of the short-circuit process fluctuates in accordance with a sine function.

(2) As depicted in Figure 6(b) and 6(c), the voltage amplitude of winding strike faults exceeds that of counterstrike faults for the lightning current in Figure 2. Specifically, the first wavefront voltage amplitudes reach approximately 2800 kV and 500 kV, respectively, and both exhibit steep shapes. When lightning bypasses lightning rods and poles and directly strikes the phase A conductor, the polarity of its voltage traveling wave is opposite to that of the other two phases. When the lightning rod or pole is directly struck by lightning, the transient voltage traveling wave polarities of all three phases are identical.

The conclusions drawn are based on the comparison of time-domain waveforms, and it is important to acknowledge that there are inherent limitations in practical applications. For instance, variations in fault distances may potentially affect the accuracy of identification results. This paper examines the transient energy ratio of high-frequency components of zero-mode voltage from a theoretical standpoint, and subsequently proposes a more comprehensive method for identification.

The electromagnetic coupling among the three-phase transmission lines necessitates the use of Karhunen-Loeve phase-mode transformation to mitigate its impact [29], as demonstrated in Equation (20).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

$$\begin{pmatrix} U_0 \\ U_1 \\ U_2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} U_a \\ U_b \\ U_c \end{pmatrix} \tag{20}$$

In Equation (20): U_0, U_1, and U_2 denote the zero-mode component, line mode 1 component, and line mode 2 component of the phase voltages U_a, U_b, and U_c, respectively, as per Karhunen-Loeve transformation.

When the lightning strike occurs near C end, there is a high detected transient energy ratio of the high-frequency component, whereas near D end, the detected transient energy ratio of the high-frequency component is low, and vice versa. In light of the unpredictable characteristics of lightning strikes on overhead lines, it has been determined that calculating the combined energy ratios from both ends is necessary to mitigate the influence of strike distance on the transient energy ratio of high-frequency components. The zero-mode voltage at both ends of the line is extracted and then subjected to KL-VMD decomposition, as depicted in Figure 7.

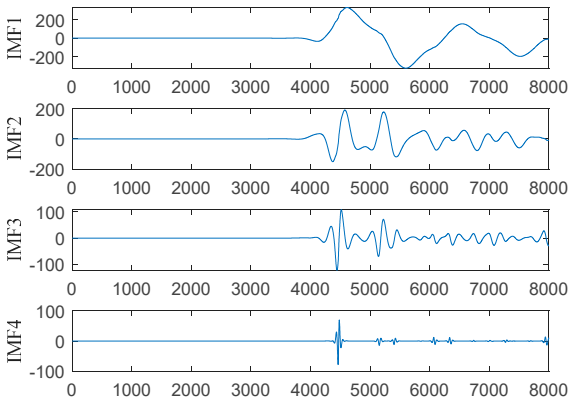

Figure 7. KL-VMD decomposition results.

Through optimization, m intrinsic mode components are decomposed at D end, while n intrinsic mode components are decomposed at C end. The absolute values are computed to derive the energy distribution of the IMF (I_1, I_2, ..., I_m and F_1, F_2, ..., F_n), extracting the high-frequency components of the highest layer I_m and F_n.

The calculation involves determining the sum of the transient energy ratios of the high-frequency components of the zero-mode voltage relative entropy at both ends of the line, denoted as E_{K1}. The definition is as follows:

$$E_{K1} = \frac{I_m}{\sum_{i=1}^{m} I_i} + \frac{F_n}{\sum_{i=1}^{n} F_i} \tag{21}$$

In Equation (21): m and n are the number of KL-VMD decomposed components IMF at C and D ends, respectively; and i is the starting index.

Table 1 presents the summation of the energy ratios of the high-frequency components of the relative entropy of the zero-mode voltage at both ends of the line, denoted as E_{K1}, for short circuit, winding strike, and counterstrike faults at varying distances from C end. The data presented in Table 1 indicates that E_{K1} value for lightning faults significantly exceeds that for single-phase ground short circuits. Consequently, a threshold of $E_1 = 0.001$ is established. If $E_1 > E_{K1}$, it corresponds to a short circuit fault; if E_1 is less than or equal to E_{K1}, it corresponds to a direct strike fault. The lightning current functions as a unipolar high-frequency impulse signal, whereas a short circuit fault can be likened to the application of a reverse voltage source. Consequently, E_{K1} value for short circuit faults is lower. Extensive simulation data calculations demonstrate that the method is capable of accurately identifying short circuit faults, but it does not differentiate between winding strike and counterstrike faults.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

Table 1. Summation of the energy ratio of high-frequency components of relative entropy in zero mode voltage at both ends of the line under different conditions E_{K1}.

Fault Type	PC/km	PD/km	E_{K1}
Short Circuit Fault	30	126	1.3326e-05
Short Circuit Fault	40	116	1.3516e-05
Short Circuit Fault	50	106	1.4201e-05
Short Circuit Fault	60	96	1.5051e-05
Winding Strike	30	126	0.0159
Winding Strike	40	116	0.0356
Winding Strike	50	106	0.5490
Winding Strike	60	96	0.2999
Counterstrike	30	126	0.0664
Counterstrike	40	116	0.0352
Counterstrike	50	106	0.0318
Counterstrike	60	96	0.0343

5.2. Identification of Winding Strike and Counterstrike Faults

To accurately differentiate between winding and counterstrike faults, it is essential to employ a suitable method that effectively identifies the transient differences between the two. The SGMD method demonstrates proficiency in signal decomposition by addressing the challenge of empirical parameter selection in EEMD methods and wavelet transforms, as well as circumventing the issue of forced component in LCD methods. Consequently, it achieves enhanced accuracy in fault identification.

At the outset, the zero-mode voltage at both terminals of the line is decomposed using PE-SGMD, with the objective of optimizing the decomposition into p intrinsic mode components at the C terminal and q intrinsic mode components at the D terminal. The absolute values are computed to derive the energy distribution of the IMF (M_1, M_2, ..., M_p and W_1, W_2, ..., W_q), by extracting the high-frequency components M_p and W_q from the highest layer. Finally, the calculation involves determining the sample entropy, symplectic geometry mode decomposition, and the high-frequency component energy ratio E_{K2} of the zero-mode voltage at both ends of the line, which is defined as follows:

$$E_{K2} = \frac{M_p}{\sum_{i=1}^{p} M_i} + \frac{W_q}{\sum_{i=1}^{q} W_i} \tag{22}$$

In Equation (22): p and q are the number of PE-SGMD decomposed components IMF at C and D ends, respectively; and i is the starting index.

The data presented in Table 2 indicates disparities in E_{K2} values between winding strike faults and counterstrike faults. E_{K2} values for winding strike faults are consistently below 0.08, whereas E_{K2} values for counterstrike faults consistently exceed 0.08. Consequently, a threshold of E_{K2} = 0.08 has been established for differentiation. The flowchart illustrating the lightning fault identification criterion can be observed in Figure 8.

Table 2. Summation of the energy ratio of the high-frequency component in PE symplectic geometric modal component of zero mode voltage at both ends of the line under different conditions E_{k2}.

Fault Type	PC/km	PD/km	E_{k2}
Winding Strike	30	126	0.0026
Winding Strike	40	116	0.0043
Winding Strike	50	106	0.0033
Winding Strike	60	96	0.0029
Winding Strike	70	86	0.0068

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

Counterstrike	30	126	0.1930
Counterstrike	40	116	0.1724
Counterstrike	50	106	0.1901
Counterstrike	60	96	0.1923
Counterstrike	70	86	0.1641

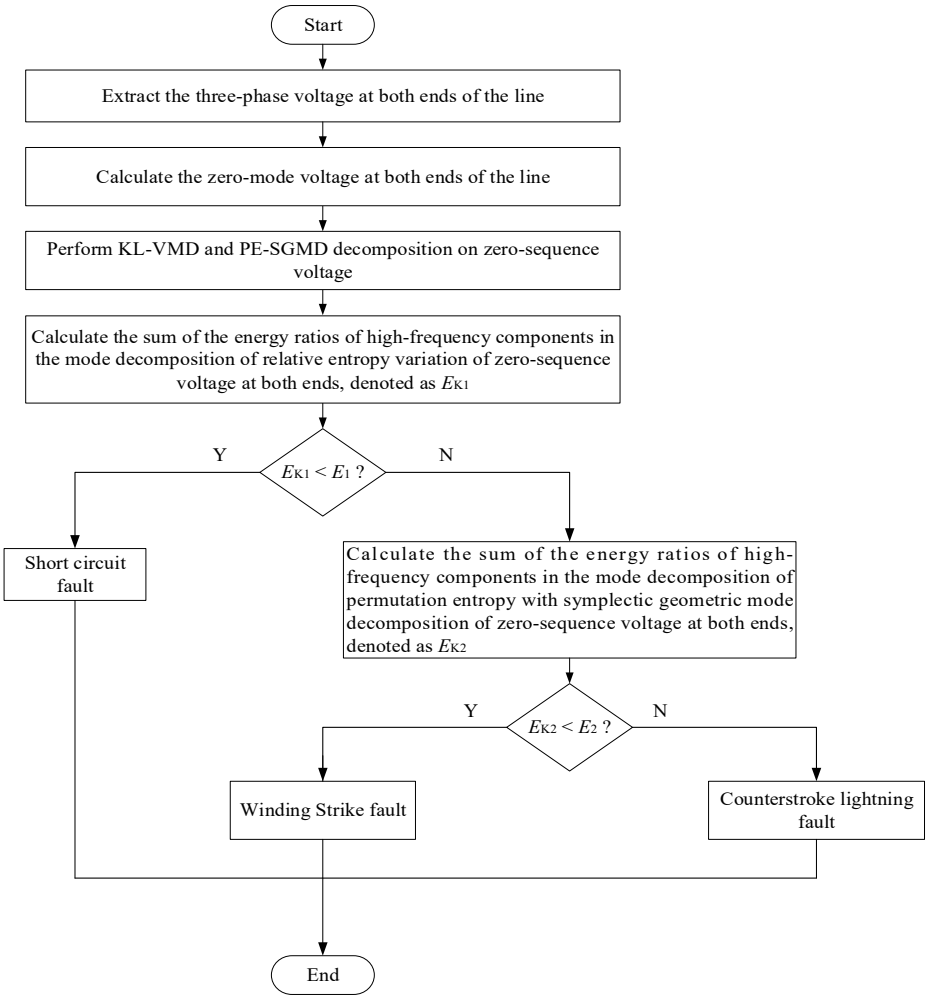

Figure 8. Fault type identification process using criterion method.

6. Simulation Validation

Utilizing the overhead line model depicted in Figure 1, simulations are performed for various operational conditions, such as short circuit, winding strike, and counterstrike faults. The double exponential model for lightning current is standardized with a 1.2/50 μs and 2.6/50 μs waveform, with initial phase angles set at 0° and 90°. The amplitude range of the winding lightning current is set between 30-70 kA, while the counterstrike lightning current amplitude range is set between 50-90 kA. The symmetrical line configuration places the lightning fault point P at a distance of 2080 km from the busbar. The fault range for single-phase ground short circuit and two-phase ground short circuit is set between 2070 km. Transition resistances are established at 10 Ω, 30 Ω, 50 Ω, 70 Ω, and 100 Ω, resulting in a total of 300 sample data sets. The data is imported into MATLAB for identification, and certain calculation results are presented in Table 3.

The analysis of the data suggests that the method yields high recognition results for short circuit, winding strike, and counterstrike faults across various operating conditions. The use of KL-VMD and SE-SGMD for extracting high-frequency components, in combination with specific criteria, has been shown to facilitate the efficient and accurate identification of overhead line fault types in diverse scenarios.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0706.v1

Table 3. Partial calculation results.

Distance/km	30km			50km			70km		
	E_{K1}	E_{K2}	Result	E_{K1}	E_{K2}	Result	E_{K1}	E_{K2}	Result
Phase A ground short circuit, R_g=30Ω, θ=90°	1.3326e-05	—	Short Circuit	1.4201e-05	—	Short Circuit	2.0681e-05	—	Short Circuit
Phase A ground short circuit, R_g=50Ω, θ=0°	1.6745e-05	—	Short Circuit	1.8363e-05	—	Short Circuit	2.7329e-05	—	Short Circuit
Phase AB ground short circuit, R_g=30Ω, θ=90°	2.9690e-05	—	Short Circuit	2.6474e-05	—	Short Circuit	3.0331e-05	—	Short Circuit
Phase AB ground short circuit, R_g=50Ω, θ=90°	3.1702e-05	—	Short Circuit	2.8562e-05	—	Short Circuit	3.0805e-05	—	Short Circuit
Winding Strike, I_{max}=30 kA, θ=0°, 1.2/50 μs	0.3145	0.0035	Winding Strike	0.5705	0.0088	Winding Strike	0.5166	0.0088	Winding Strike
Winding Strike, I_{max}=40 kA, θ=0°, 2.6/50 μs	0.3168	0.0029	Winding Strike	0.5781	0.0087	Winding Strike	0.5196	0.0091	Winding Strike
Winding Strike, I_{max}=50 kA, θ=90°, 1.2/50 μs	0.0159	0.0026	Winding Strike	0.5492	0.0033	Winding Strike	0.2793	0.0068	Winding Strike
Winding Strike, I_{max}=60 kA, θ=90°, 2.6/50 μs	0.0165	0.0026	Winding Strike	0.5526	0.0039	Winding Strike	0.2817	0.0063	Winding Strike
Counterstrike, Imax=50 kA, θ=90°, 1.2/50 μs	0.0664	0.1914	Counterstrike	0.0318	0.1896	Counterstrike	0.0324	0.1636	Counterstrike
Counterstrike, Imax=60 kA, θ=90°, 2.6/50 μs	0.0701	0.1707	Counterstrike	0.0318	0.1557	Counterstrike	0.0328	0.1443	Counterstrike
Counterstrike, Imax=70 kA, θ=0°, 1.2/50 μs	0.0783	0.1930	Counterstrike	0.0622	0.1901	Counterstrike	0.0404	0.1641	Counterstrike
Counterstrike, Imax=80 kA, θ=0°, 2.6/50 μs	0.0697	0.1688	Counterstrike	0.0669	0.1562	Counterstrike	0.0540	0.1447	Counterstrike

7. Conclusion

This study has developed a model for identifying overhead line lightning faults by employing KL-VMD and SE-SGMD to decompose the zero-mode voltage. The study also established lightning recognition criteria through extensive data calculations, resulting in the following conclusions:

1. (1) In situations where signal decomposition is hindered by challenges such as modal mingling, the utilization of KL-VMD can automatically optimize the decomposition layers and penalty factors. This approach effectively extracts transient characteristic quantities, demonstrating its strong adaptability in fault signal decomposition;
2. (2) A criterion is proposed for identifying winding strike, counterstrike, and short circuit faults by analyzing the fault stage traveling wave amplitude, wavefront polarity, rate of change, and modal energy distribution using KL-VMD and PE-SGMD. Following thorough data calculations, the validity and accuracy of this criterion have been confirmed;
3. (3) The criterion demonstrates high reliability in accurately distinguishing between short circuit faults and lightning conditions under various lightning current amplitudes, distances, and initial phase angles. It also provides a reference for line fault identification.

Funding: This research was funded by the key laboratory of Xinjiang Uygur Autonomous Region is open for research projects (2023D04029), the key research and development program projects in Xinjiang Uygur Autonomous Region (2022B01003-3) and the graduate innovation projects in Xinjiang Uygur Autonomous Region (XJ2023G051).

References

1. Dong,X.; Zhang,D. Across the Tianshan mountains of 750 kV transmission line lightning risk assessment. *Electric Power Science and Engineering* **2015**, *31*, 66-71.
2. Zhang,D.; Dong,X.; Tao,F.; Wang,S.; Research on reduction of the monsoon influence on the transmission line shielding failure. *Insulators and Surge Arresters* **2014**, *259*, 57-61.
3. Zhao,C.; Chen,J.; Gu,S.; Ruan,J.; Li,X.; Tong,X.; Hu,W. Research on differentiated lightning protection comprehensive management for the 500-kV power network in the area near the three gorges project. IEEE *transactions on power delivery* **2011**, *27*, 337-352.
4. Jiang,K.; Du, L.; Chen, H.; Yang, F.; Wang, Y. Non-contact measurement and polarity discrimination-based identification method for direct lightning strokes. *Energies* **2019**, *596*, 713-718.
5. Zou,G.; Gao,H.; Zhu,F.; Wang,H. Integral identification method of lightning stroke and fault for transmission line. *Power System Protection and Control* **2012**, *40*, 43-48.
6. Shu,H.; Wang,Y.; Cheng,C.; Sun,S. Analysis of electromagnetic transient and fault detection on ±800 kV UHVDC transmission lines under lightning stroke. *Proceedings of the CSEE* **2008**, *28*, 93-100.
7. Sima,W.; Xie,B.; Yang,Q.; Wang,J. Identification of lightning over-voltage about UHV transmission line.*High Voltage Engineering* **2010**, *36*, 306-312.
8. Guo,N.; Qin,J. Locating method of short-circuit point for transmission lines under lightning stroke fault. *Automation of Electric Power Systems* **2009**, *33*, 74-77+85.
9. Zhong,H.; Chen,J.; Fu Q.; Hua,M.; Lightning strike identification algorithm of an all-parallel auto-transformer traction power supply system based on morphological fractal theory, *IEEE Transactions on Power Delivery* **2023**, *38*, 2119-2132.
10. Si,D.; Shu,H.; Chen,X.; Yu,J. Study on characteristics and identification of transients on transmission lines caused by lightning stroke. *Proceedings of the CSEE* **2005**, *25*, 64-69.
11. Gao,Y.; Zhu,Y.; Yan,H.; Yan,H. Identification of lightning strike and short-circuit fault based on wavelet energy spectrum and transient waveform characteristics. *Mechatronics and Industrial Informatics* **2014**, *596*, 713-718.
12. Chen,Z.; Pei,H.; Teng,C.; Identification of transmission line lightning based on HHT. *International Journal of Smart Home* **2016**, *10*, 201-212.
13. Zhao,H.; Wang,X.; Gao, C. Traveling wave fault location method for high voltage transmission lines based on FIMD and Hilbert transform. *Electrical Measurement & Instrumentation* **2020**, *57*, 77-82.
14. Han,Z.; Rao,S.; Jiang,Y.; Wang,S.; Liu,J.; Lu,Y. Energy ratio function-based traveling-wave fault location for transmission lines. *Power System Technology* **2011**, *35*, 216-220.
15. Liu,J.; Su,Y.; Deng,R.;Sun,F. Performance degradation assessment of rolling bearing based on KL-VMD and comprehensive characteristic indexes. *Journal of Electronic Measurement and Instrumentation* **2022**, *36*, 78-88.

16. Xi,Y.; Cui,Y.; Tang,X.; Li,Z.; Zeng,X. Fault location of lightning strikes using residual analysis based on an adaptive kalman filter, *IEEE Access* **2019**, 7, 88126-88137
17. Gu,Y.; Song,G.; Guo,A.; Tao,R.; Liu,Y. A lightning recognition method for DC line travling-wave protection of HVDC. *Proceedings of the CSEE* **2018**, *38*, 3837-3845+4024.
18. Long,Y.; Yao,C.; Mi,Y.; Hu,D.; Yang,N.; Liao,Y Identification of direct lightning strike faults based on mahalanobis distance and S-transform *IEEE Transactions on Dielectrics and Electrical Insulation* **2015**, *22*, 2019-2030.
19. Song,X.; Gao,Y.; Ding,G.;Yan H. Lightning strike interference and fault identification of transmission system. *Insulators and Surge Arresters* **2021**, 299, 96-102+110.
20. Deng L, Zhang A, Zhao R. Intelligent identification of incipient rolling bearing faults based on VMD and PCA-SVM. *Advances in Mechanical Engineering* **2022**, *14*, 1-18.
21. Li,J.; Zhu,X.; Guo,J. Bridge modal identification based on successive variational mode decomposition using a moving test vehicle. *Advances in Structural Engineering* **2022**, *25*, 2284-2300.
22. Harmouche,J.; Delpha,C.; Diallo,D. Incipient fault amplitude estimation using KL divergence with a probabilistic approach. *Signal Processing* **2016**, *120*, 1-7.
23. Bandt,C.; Pompe,B. Permutation entropy: a natural complexity measure for time series. *Physical review letters*, **2002**, 88, 174102.
24. Li X, Ouyang G, Richards D A. Predictability analysis of absence seizures with permutation entropy. *Epilepsy research* **2007**, *77*, 70-74.
25. Ricci,L.; Politi,A. Permutation entropy of weakly noise-affected signals. *Entropy* **2022**, *24*, 54-64.
26. Ouyang,G.; Li,J.; Liu,X.; Li,X. Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis. *Epilepsy research* **2013**, *104*, 246-252.
27. Zheng, Z.; Xin, G. Fault feature extraction of hydraulic pumps based on symplectic geometry mode decomposition and power spectral entropy. *Entropy* **2019**, *21*, 476-500.
28. Pan,H.; Yang,Y.; Li,X.; Zheng,J.; Cheng,J. Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis. *Mechanical Systems and Signal Processing* **2019**, *114*, 189-211.
29. Dai,J.; Liu,Y.; Jiang,W.; Liu,Z.; Sheng,G.; Yan,Y.; Jiang,X. Identification of back striking and shielding failure on transmission line based on time domain characteristics of trav*eling wave. Transactions of China Electrotechnical Society* **2016**, *31*, 242-250.