Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 11, No. 1, 2024

Fault Distance Measurement Method Based on Wavelet Energy Spectrum and BWO Algorithm Optimized CNN-GRU Hybrid Neural Network

Haipeng Tian[1, 2, *], Hao Wu[1, 2]

[1] Automation and Information Engineering, Sichuan University of Science & Engineering, Zigong, Sichuan 643000, China
[2] Artificial Intelligence Key Laboratory of Sichuan Province, Zigong, Sichuan 643000, China
[*] Corresponding author: Haipeng Tian (Email: 3396001722@qq.com)

Abstract: Aiming at the problems such as noise interference in the current fault ranging methods for flexible DC transmission lines, a single-ended intelligent fault ranging method is proposed as a hybrid neural network based on wavelet energy spectrum and BWO algorithm to optimize the multi-head attention mechanism of CNN combined with gated recurrent unit GRU. First, the correlation between wavelet spectrum energy and fault distance is analyzed, and wavelet packet decomposition is used to extract the wavelet packet energy spectrum feature vector as the model input of the neural network. Second, the hybrid neural network model is constructed and trained to mine the deep fault information in the time series, and the parameters of the hybrid neural network model with added multi-head attention are optimized using the iterative optimization search of the beluga algorithm to achieve fast and accurate positioning of the fault distance. Finally, the network is trained using the constructed four-terminal Zhangbei flexible DC transmission system model to collect data, and the experimental results prove that the method has high distance measurement accuracy, strong anti-interference ability and generalization ability, and a certain degree of resistance to transition resistance.

Keywords: Wavelet packet decomposition, Wavelet energy spectrum, Optimization algorithm, GRU, Fault ranging.

1. Introduction

As China's power demand continues to grow, the pressure on the power system is increasing day by day. Transmission lines are crucial components of the power system, serving as the connecting link between the generation and load ends. However, transmission lines are susceptible to various external factors during operation, making them components with relatively high failure rates in the power system [1-3]. In recent years, China's power system construction has continuously expanded, leading to an increasing number of large-scale power outage incidents caused by transmission line faults, which have seriously disrupted people's normal production and life. Therefore, research on fault distance measurement of transmission lines has attracted widespread attention from various sectors of society [4-5]. To ensure the stable and reliable operation of the power grid, reduce the workload of scheduling, promptly eliminate faults to minimize economic losses, rapid and accurate fault localization after a fault occurs on the line is crucial.

Currently, there are mainly four methods for fault localization in DC transmission lines, including analytical calculation method [7], traveling wave distance measurement method [8], natural frequency method [9], and intelligent algorithm [9].

(1) The analytical calculation method is based on a detailed analysis of the topology and electrical parameters of the power grid. It utilizes mathematical models to transform the physical characteristics and operational parameters of the grid into mathematical expressions. Then, by solving these equations, the fault location is determined. The advantage of this method is its high precision, but it requires a large amount of grid information and complex computational processes [10].

(2) The traveling wave fault location method utilizes the propagation characteristics of traveling waves in power systems to determine the location of faults. This method is theoretically unaffected by transmission line parameters, fault types, and the initial occurrence time of faults. It demonstrates good performance in accuracy and speed. However, due to the difficulty in accurately identifying the traveling wave head, identification errors can significantly affect the effectiveness of fault distance measurement [11].

(3) The method of inherent frequency for distance measurement requires the accurate extraction of the inherent dominant frequency. When there are interference signals, the measurement results are often affected [12]. The energy of the traveling wave spectrum is concentrated near the inherent dominant frequency, making fault information easier to extract [13]. Moreover, the energy spectrum of wavelet packets can accurately extract signals from different frequency bands, so using the energy spectrum of wavelet packets for positioning is simpler and more reliable [14].

(4) With the expansion of the power grid, the stochastic nonlinearity of faults is enhanced. By utilizing the strong data mining and self-learning capabilities of artificial intelligence methods [15], more accurate fault information can be distilled, thus achieving fault localization. The literature [16] utilizes impedance matrices and admittance matrices to obtain the phasor matrix of the power system. It then employs phasor transformation and Fourier compensation algorithms to eliminate various factors affecting distance measurement accuracy, finally constructing a fault distance measurement model through the optimization of ant colony algorithm paths. Literature [17] proposes a hybrid system based on the optimal parameter adaptive neuro-fuzzy inference system (ANFIS) and Hilbert-Huang (HH) transform. It utilizes chaotic dynamic weight particle swarm optimization (CDWPSO)

algorithm to optimize network parameters, addressing wavefront identification issues and enhancing fault distance measurement accuracy. Literature [18] optimizes variational mode decomposition (VMD) parameters using sparrow search algorithm, extracts high-frequency modal components containing fault characteristics, and then obtains the time-frequency domain full waveform of fault traveling waves through transient extraction transformation, achieving fault distance measurement. Through formula derivation, literature [19] discovers a nonlinear relationship between the linear mode component of initial transient voltage traveling waves, the wavelet energy ratio of zero mode components, and fault distance. It proposes using Elman-Adaboost integrated artificial neural networks for fitting, constructing an integrated artificial neural network fault distance measurement model. Literature [20] converts voltage and current into relative offsets, then employs support vector regression (SVR) for deep learning, while effectively improving distance measurement accuracy through improved grey wolf algorithm optimization parameters.

Based on the above analysis, a hybrid neural network distance measurement method is proposed based on wavelet energy spectrum and whale optimization algorithm to optimize the multi-head attention mechanism. The wavelet energy spectrum of the linear mode voltage is taken as the fault feature vector. The whale optimization algorithm is used to optimize the multi-head attention mechanism to obtain attention weights. The attention weights are multiplied by the fault feature vector to obtain weighted data, which is then input into the convolutional neural network to extract features. Finally, it is input into the gated recurrent unit to extract temporal features, and the output prediction result serves as the distance measurement result. With the powerful nonlinear fitting capability of the gated recurrent unit, we can quickly extract effective information and store key information, which performs excellently in flexible DC transmission line distance measurement algorithms. Moreover, using the wavelet energy spectrum as the feature vector is not limited by system boundaries, thus solving problems such as unreliable wavefront identification and inaccurate wave arrival time recognition. Finally, the effectiveness of the proposed approach is verified through simulation experiments.

2. Four-terminal flexible DC transmission system

According to the system parameters and topology of the Zhangbei Flexible HVDC Transmission Project, a frequency-dependent phase model is used to simulate overhead lines on the PSCAD/EMTDC simulation platform. A flexible HVDC grid model with four terminals as shown in Figure 1 is constructed. The rated DC voltage of the system is $\pm 500\,\mathrm{kV}$. Using line L_1 as a reference, fault F_1 on L_1 is an intra-area fault, including three types: positive and negative pole grounding faults, and bipolar short-circuit faults. DC faults F_2, F_3, and F_4 occurring on other lines are all inter-area faults, also including three types of faults. F_5 is an inter-area AC-side fault, including single-phase grounding, two-phase short-circuit, and three-phase grounding, etc. Lines L_1, L_2, L_3, and L_4 are 207km, 192km, 217km, and 50km long, respectively. Other main parameters of the system are shown in Table 1.

Table 1. The main parameters of the system

Parameters	MMC1	MMC2	MMC3	MMC4
Control parameter	Constant DC Voltage	Constant DC Voltage	certainly meritorious	certainly meritorious
Rated power（MW）	1500	3000	1500	3000
AC Voltage（kV）	220	220	220	220
Bridge Arm Inductors（mH）	29	19	116	29
Number of submodules	101	101	101	101
Submodule Capacitance（μF）	2500	2500	2500	2500

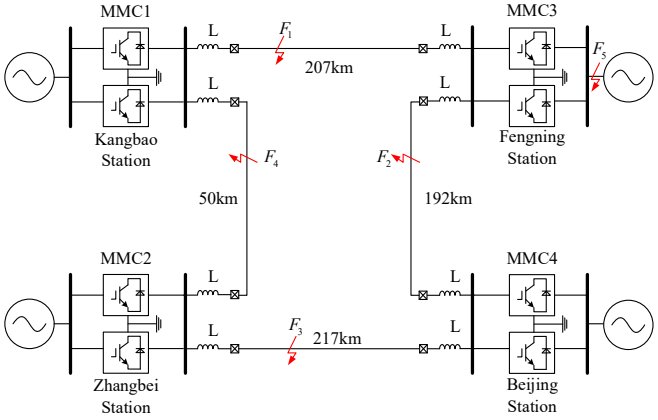


Figure 1. Four-terminal flexible DC grid topology

3. Wavelet Energy Spectrum Extraction

3.1. Principle of Wavelet Transform

Wavelet packet decomposition (WPD) is the process of decomposing a signal by a wavelet packet basis function. Compared with the traditional wavelet decomposition, wavelet packet decomposition provides a more flexible decomposition and can depict the frequency characteristics of the signal in more detail. Figure 2 shows a schematic diagram of a 3-layer wavelet packet decomposition of a signal.

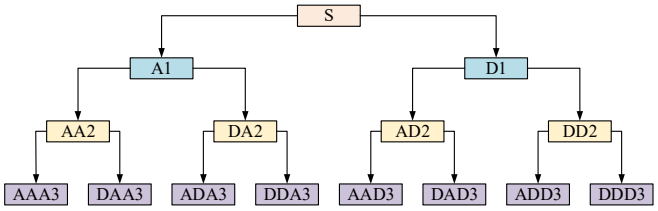


Figure 2. Schematic diagram of 3-layer wavelet packet decomposition

The wavelet packet decomposition expression is:

$$X(t)=\sum_{j=0}^{j}\sum_{k=0}^{2^j-1}d_{j,k}\phi_{j,k}(t)+\sum_{j=0}^{j}\sum_{k=0}^{2^j-1}w_{j,k}\psi_{j,k}(t) \quad (1)$$

Where $X(t)$ is the original signal, $d_{j,k}$ are the low frequency coefficients at scale j and location k, and $w_{j,k}$ are the low frequency coefficients at scale j and location k. In this expression, the first term represents the low-frequency

portion of the signal, while the second term represents the high-frequency portion of the signal, and by adjusting the value of the scale j, the degree of decomposition of the signal at different scales can be controlled.

The coefficients obtained from wavelet packet decomposition are utilized for hierarchical reconstruction. This process helps us to better understand the frequency characteristics of the signal and helps in signal processing, analysis and feature extraction. The expression for wavelet reconstruction is:

$$d_l^{j,n} = \sum_k \left(\tilde{h}_{2l-k} d_k^{j+1,2n} + \tilde{g}_{2l-k} d_k^{j+1,2n+1} \right) \tag{2}$$

3.2. Wavelet energy spectrum

The wavelet packet decomposition transform is performed to obtain the wavelet packet coefficients and then the energy values at each scale and frequency are calculated based on the wavelet packet coefficients, as shown in equation (3) we sum the modulus squares of the wavelet packet coefficients to calculate the wavelet packet energy.

$$E_j = \int \left| S_j(t) \right|^2 dt = \sum_{k=1}^{n} x_j(k)^2 \tag{3}$$

Where $x_j(k)$ is the wavelet packet coefficient at different scales.

The energy magnitude of different frequency components can be demonstrated by the energy spectrum, and through the difference of the energy spectrum under different frequency components, we are able to realize accurate fault ranging. According to equation (3) to get the wavelet packet energy under different scales, the combination of wavelet energy spectrum is:

$$E = \left[E_1, E_2, \cdots, E_j, \cdots \right] \tag{4}$$

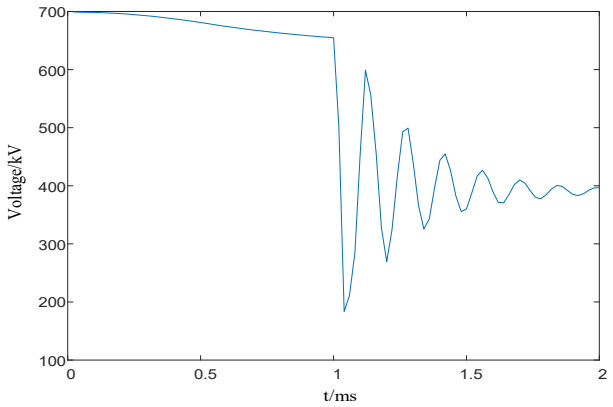
This wavelet energy spectrum integrates the wavelet packet energy of each frequency band, showing the energy distribution of the whole transmission line, which can reflect the overall fault information of the transmission line and provide help for fault ranging.

In this paper, taking the fault occurring at F_3 of the four-terminal flexible DC network in Figure 1 as an example, the sampling frequency is set to 50kHz, the total running time of the simulation is set to 1.8s, and the moment of fault occurrence is set to 1.6s. The fault voltage and current data are collected from the transmission line L_1 with a fault distance of 100km and a transition resistance of 50Ω. Due to the coupling effect of positive and negative lines in flexible DC transmission, again we need to decouple first. The decoupling is performed using the Kellenberger transform matrix shown in equation (5) to obtain the line mode voltage and line mode current we need.

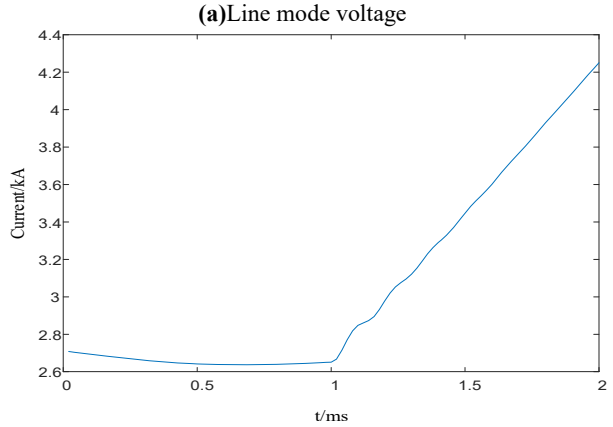
$$\begin{bmatrix} u_0 \\ u_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_P \\ u_N \end{bmatrix} \tag{5}$$

Where u_P and u_N are the collected voltages on the positive and negative lines, u_0 is the zero mode voltage, and u_1 is the line mode voltage.

The waveforms of line mode voltage and line mode current for 2ms before and after the fault occurrence are obtained by decoupling equation (5) as shown in Figure 3(a) and Figure 3(b).



(a) Line mode voltage



(b) Line mode current

Figure 3. Line mode voltage and line mode current

After obtaining the line mode voltage within 2ms before and after the fault, it is decomposed into wavelet packets according to the equations from equation (1) to (4), and then the wavelet energy spectrum is calculated.

The db6 wavelet is a tightly supported orthogonal real wavelet with excellent regularity and vanishing moments. We choose to use the db6 wavelet as the wavelet basis function to perform the wavelet packet decomposition and extract the wavelet energy of the line-mode voltage for 2ms before and after the occurrence of the fault. Because the line mode voltage is mainly concentrated in the lower frequency, in order to make the signal band after wavelet packet decomposition can completely contain the low-frequency band, the line mode voltage is decomposed to 8 layers, there will be a total of 2^8=256 wavelet packet sub-bands, i.e., the obtained wavelet energy spectrum is a 1×256 vector.

In flexible DC transmission lines, lightning faults generate a large number of electromagnetic waves with high-frequency components, which affect the high-frequency region (d_0,d_1,d_2), resulting in abnormally high energy levels, which are different from normal short-circuit faults. This interference affects the ranging system and reduces the ranging accuracy. In order to eliminate the lightning fault, the wavelet energy in the high frequency region needs to be eliminated, so in this paper, the wavelet energy spectrum in the region of d_3-d_{255} is retained, and the fault feature vector obtained as shown in equation (6) is a 1×253 vector.

$$E=\left[E_3,E_4,\cdots,E_{255}\right] \tag{6}$$

This paper takes the DC transmission line L1 with a length of 207km as an example, carries out the simulation experiments of positive-pole grounding faults with fault distances of 10, 50, 100, 150, and 200km, respectively, and the transition resistances are set to 50Ω, and obtains positive and negative voltage currents under different fault distances, calculates the line mode voltage, and then calculates the wavelet energy spectrum. The wavelet energy spectrum of different fault distances is shown in Figure 4, from which it can be seen that different fault distances lead to different wavelet energy spectra of the line mode voltage, and fault ranging can be realized by combining the different wavelet energy spectra with the fault distance.

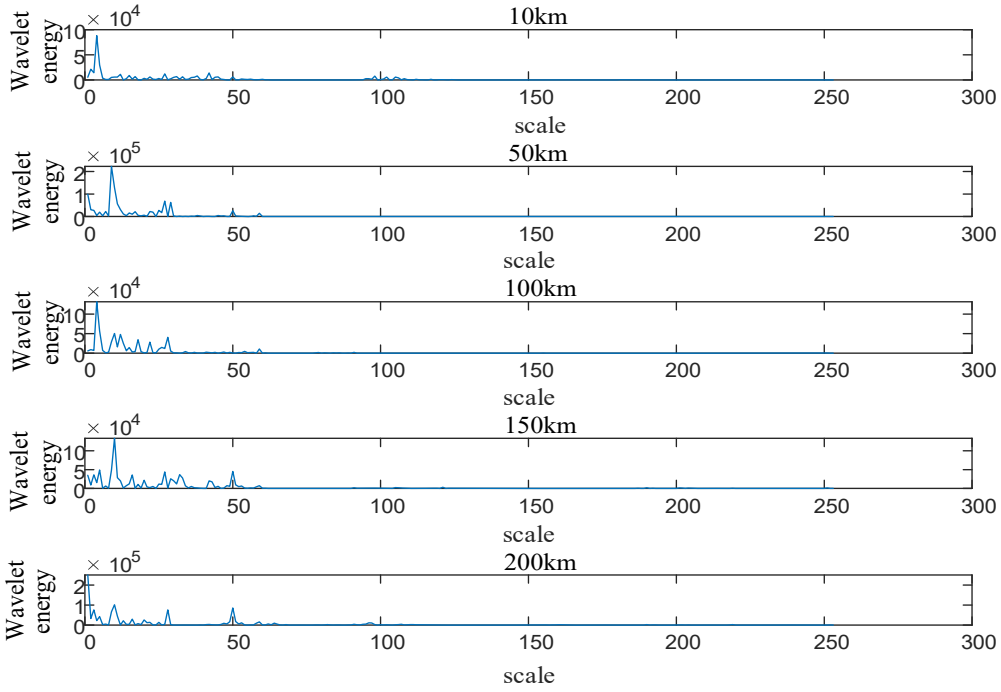


Figure 4. Relationship between fault distance and wavelet energy

4. Improved 1D-CNN-GRU Convolutional Neural Network

In this paper, the 1D-CNN-GRU hybrid neural network is improved by introducing the beluga algorithm into the multi-head attention mechanism in order to optimize the allocation of attention weights and improve the model's ranging accuracy. The model also combines a convolutional neural network and a gated recurrent unit to extract local and temporal features in the data to further improve the ranging performance of the model.

4.1. Beluga whale optimization

Beluga whale optimization (BWO) is a heuristic optimization algorithm inspired by the feeding behavior of whales. The algorithm simulates the behavior of individuals in a population of whales, including beluga swimming, feeding, and "whale-falling" behaviors, and BWO optimizes the problem-solving process by simulating these behaviors.

The beluga optimization algorithm is based on the mechanism of population, where each beluga whale is assumed to be a candidate solution and is continuously updated during the optimization process, and the population consisting of n beluga whales can be represented by a matrix as follows:

$$X=\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,1} & \cdots & x_{n,d} \end{bmatrix} \tag{7}$$

Then, the fitness function for all beluga whales can be expressed as:

$$F_X=\begin{bmatrix} f\left(x_{1,1},x_{1,2},\cdots,x_{1,d}\right) \\ f\left(x_{2,1},x_{2,2},\cdots,x_{2,d}\right) \\ \vdots \\ f\left(x_{n,1},x_{n,2},\cdots,x_{n,d}\right) \end{bmatrix} \tag{8}$$

In equation (7) and (8), d is the dimension of the variable to be optimized, n is the population size, and f is the fitness relationship.

The exploration phase is mathematically modeled by the swimming behavior of beluga whales, and the location of the search agent is determined by the swimming of a pair of beluga whales, which are updated as follows:

$$\begin{cases} X_{i,j}^{t+1}=X_{i,p_j}^{t}+\left(X_{r,p_1}^{t}-X_{i,p_j}^{t}\right)\left(1+r_1\right)\sin\left(2\pi r_2\right), j=even \\ X_{i,j}^{t+1}=X_{i,p_j}^{t}+\left(X_{r,p_1}^{t}-X_{i,p_j}^{t}\right)\left(1+r_1\right)\cos\left(2\pi r_2\right), j=odd \end{cases} \tag{9}$$

Where t is the current number of iterations.

The development phase was inspired by the feeding behavior of beluga whales, where individuals make jumps to catch prey via Levy flight, and obtain food based on the cooperation and positional movement of the beluga whales around them, with the formula:

$$X_i^{t+1} = r_3 X_{best}^t - r_4 X_i^t + C_1 \bullet L_F \bullet \left(X_r^t - X_i^t \right) \tag{10}$$

Where r_3 and r_4 are random numbers between (0,1), $C_1 = 2r_4\left(1 - t/T_{\max}\right)$ denotes the measurement of the random jump strength of the Levy flight, and L_F is calculated as:

$$\begin{cases} L_F = 0.05 \times \dfrac{\mu \times \sigma}{|v|^{1/\beta}} \\ \sigma = \left(\dfrac{\Gamma(1+\beta) \times \sin(\pi\beta/2)}{\Gamma\left((1+\beta)/2\right) \times \beta \times 2^{(\beta-1)/2}} \right) \end{cases} \tag{11}$$

Where u and v are normally distributed random numbers and β is a constant that defaults to 1.5.

The whale fall phase simulates the change of individuals in the group, and depending on a certain probability, an individual may be selected for a fall operation. If an individual is selected, its position may change, which represents a situation where the individual may move to another location, or the individual may be shot down and fall into the deep sea. In order to keep the population size constant, the position of the individual after the fall needs to be updated. The method of updating is to utilize the beluga whale position and the whale fall step size to update the whale position:

$$X_i^{t+1} = r_5 X_i^t - r_6 X_i^t + r_7 X_{step} \tag{12}$$

where r_5, r_6 and r_7 are random numbers between (0,1), $X_{step} = \left(u_b - l_b\right)\exp\left(-C_2 t/T\right)$ denotes the step size of the whale fall, $C_2 = 2W_f \times n$ is the step factor, u_b and l_b are the upper and lower bounds of the variables. The probability W_f of a whale fall is designed as a linear function:

$$W_f = 0.1 - 0.05t/T \tag{13}$$

Where the optimization process proceeds, the probability of a whale falling gradually decreases from 0.1 in the initial iteration to 0.05 in the last iteration, which reflects when beluga whales are closer to the food source during the optimization process. When beluga whales are close to the food source, they are also less at risk because they are more likely to find enough food without having to continue searching.

4.2. Multiple Attention Mechanisms

Multiple attention mechanisms have been developed from self-attention mechanisms to improve the model's ability to understand and represent input sequences. In the traditional self-attention mechanism, the model correlates the representation of each position in the input sequence with the representation of other positions so that global information can be taken into account when generating the output. In contrast, the multi-head attention mechanism allows the model to pay attention under multiple representation spaces by introducing multiple attention heads (i.e., multiple different weight matrices), each of which can learn different attention weights.

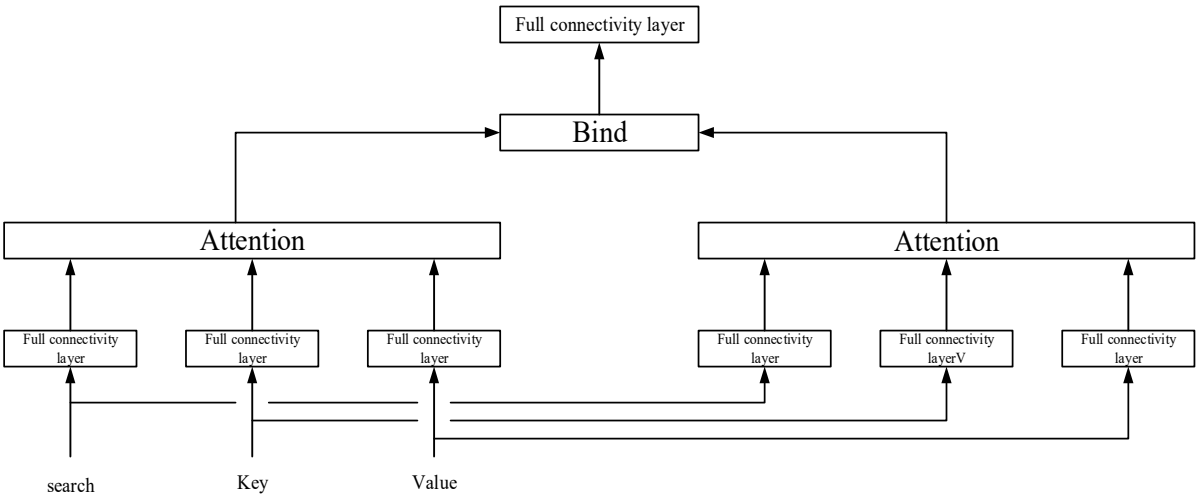


Figure 5. Structural diagram of the multi-attention mechanism

4.3. GRU algorithm

When a fault occurs on a transmission line of a flexible DC grid, the fault waveform is complex and random, and it is difficult to intuitively find the nonlinear relationship between the fault distance and the wavelet energy spectrum of the line mode voltage, and it is very difficult to want to output the ranging results. Therefore, a hybrid CNN-GRU neural network is created, which combines the advantages of CNN for feature extraction and GRU for sequence modeling, and at the same time, a multi-attention mechanism is added to enhance the ability of the model to capture feature data, and the relationship between the wavelet energy spectrum of the line-mode voltage and the fault distance is explored in depth, so as to achieve high-precision fault ranging.

Figure 6 shows the structure of the GRU unit, which is a variant of recurrent neural network that realizes selective processing of input data and memorization of historical

information through a gating mechanism. Compared with traditional recurrent neural networks, GRU has certain advantages in dealing with long-term dependence and gradient explosion problems, and has high computational efficiency.

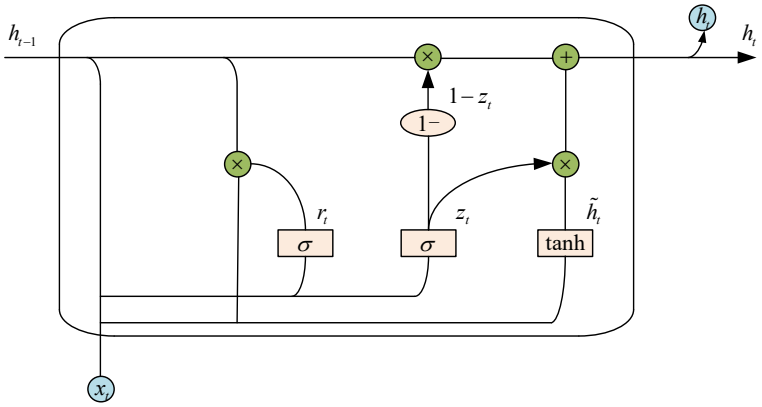


Figure 6. Structure diagram of GRU unit

The arithmetic process of GRU is denoted as:

$$\begin{cases} z_t = \sigma\left(W_z x_t + U_z h_{t-1} + b_z\right) \\ r_t = \sigma\left(W_r x_t + U_r h_{t-1} + b_r\right) \\ \tilde{h}_t = \tanh\left(W_h x_t + U_h\left(r_t \otimes h_{t-1}\right) + b_h\right) \\ h_t = \left(1 - z_t\right) \otimes h_{t-1} + z_t \otimes \tilde{h}_t \end{cases} \tag{14}$$

In the formula, the reset gate r_t controls how the hidden state of the previous moment should be mixed with the information input at the current moment, and its value between 0 and 1 determines how much of the previous moment's memory should be retained. The update gate z_t integrates the current input with the previous moment's memory, and through the update gate, the model can selectively update the memory state to fit the current input information. The candidate memory $\tilde{h}_t$ is obtained by the reset gate, which contains the input of the current moment and part of the memory information of the previous moment. The updated memory cell h_t is an updated memory cell obtained by the update gate and candidate memory computation. It combines the input of the current moment and the memory information of the previous moment and is updated to a new hidden state.

4.4. Improved 1D-CNN-GRU fault ranging model based on wavelet energy spectrum

4.4.1. Basic Processes

The improved 1D-CNN-GRU fault ranging model based on wavelet energy spectrum designed in this paper is shown in Figure 7.

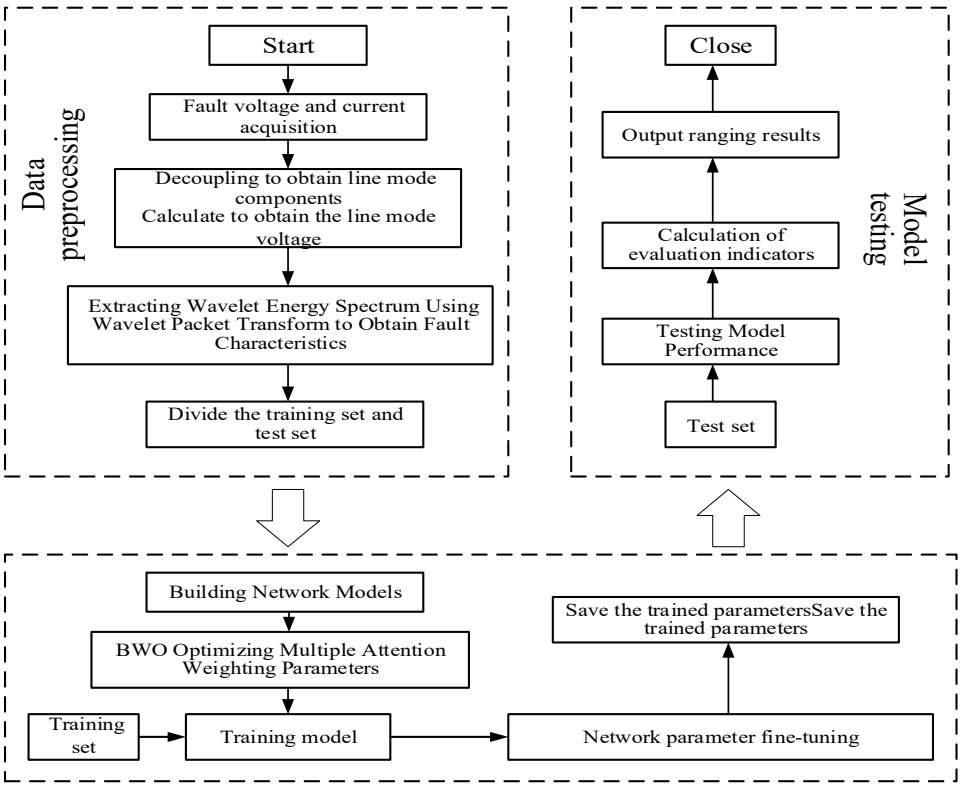


Figure 7. Wavelet energy spectrum with improved 1D-CNN-GRU fault ranging models

The basic flow of the algorithm is as follows:

(1) Collect the original data of fault voltage and fault current of multi-terminal flexible DC transmission line through the simulation of PSCAD platform;

(2) Decouple the voltage and current data 2ms before and after the occurrence of the fault using the Kellenberger transform, and obtain the line mode voltage 2ms before and after the occurrence of the fault;

(3) Carry out wavelet packet transform on the line mode voltage before and after the fault occurs for 2ms to obtain the wavelet energy spectra of different frequency bands, and after eliminating the energy of the high-frequency region, the obtained wavelet energy spectra are used as the feature vectors;

(4) Convert all the collected raw data into feature vectors according to steps (1) (2) (3), construct a dataset, and divide the training set and test set, and the ratio of the training set to the test set is 9:1;

(5) Train the 1D-CNN-GRU hybrid neural network using the training set, and use the Adam optimizer to update the model parameters to speed up the convergence of the network;

(6) Fine-tune the parameters of the model, such as epochs and batchsize, according to the training effect, while the learning rate decay strategy uses segmented constant decay, the learning rate decline period is set to 20 iterations, and the learning rate decline factor is set to 0.8, and finally save the model structure and weight parameters with the best training effect;

(7) Test the trained network model using the test set, calculate the relevant evaluation indexes according to the output results, and output the ranging results.

4.4.2. Indicators for evaluating the effectiveness of ranging

In order to have a specific measure of the ranging effect of the network model and reflect the effectiveness of the ranging algorithm, the ranging effect of the CNN-GRU hybrid neural network is evaluated by using two metrics, the mean absolute percentage error (MAPE) and the decidability coefficient, which are computed by the formulas of the two metrics, respectively:

$$e_{MAPE}\left(Y,h\right)=\frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{h\left(x_i\right)-y_i}{y_i}\right| \tag{15}$$

$$R^2\left(Y,h\right)=1-\frac{\sum_{i=1}^{m}\left(h\left(x_i\right)-y_i\right)^2}{\sum_{i=1}^{m}\left(\overline{y}_i-y_i\right)^2} \tag{16}$$

Where Y is the sample set of the test set, y_i is the true fault distance of each sample in the test set, $h\left(x_i\right)$ is the ranging result outputted by each sample in the test set after the network model, and $\overline{y}_i$ is the mean value of the true fault distance of each sample in the test set. Where the smaller the value of e_{MAPE}, the larger the value of the decidable coefficient R^2, it means the better the fitting effect of the fault distance and the higher the accuracy of the fault ranging.

5. Simulation Verification

In this chapter, a large number of simulations are carried out under different fault conditions based on the four-terminal Zhangbei flexible straight transmission model in Fig. 1, the sampling frequency is set to 50 kHZ, the fault occurrence moment of the Zhangbei flexible straight model is 1.6 s, and the duration of the fault is 0.2 s. Voltage and current data on the rectifier side before and after the fault occurs for 2 ms are collected, and after the phase-mode transformation, the line-mode voltage inverse travelling wave is computed, and the wavelet energy spectrum is extracted by wavelet packet transform for the line-mode voltage inverse travelling wave. The wavelet energy spectrum is extracted by wavelet packet transform, and the wavelet energy spectrum is used as a feature vector, which is inputted into the CNN-GRU hybrid neural network optimized for multi-attention mechanism by the beluga algorithm for training to realize fault ranging.

The length of line L_1 is 207km, and on this line, starting from 7km on the rectifier side, three types of faults are set every 1km, namely, positive fault, negative fault and bipolar fault, until the distance from the rectifier side stops at 200km, and the transition resistances are set to 10Ω, 50Ω and 90Ω, obtaining a total of 1719 sets of sample data.

5.1. Convergence validation of CNN-GRU hybrid neural network models

The dataset is divided into training set and test set according to the ratio of 9:1 and input into the CNN-GRU hybrid neural network model for training. Judge whether the network model is overfitting by observing the change of the loss value of the training set with the number of training rounds, and if overfitting occurs, stop the network training, adjust the parameters of the network model, and then re-train. Finally, the MiniBatchSize of the network model is set to 24, the epoch is set to 60, and the initial learning rate is set to 0.01. The loss value curve of the network training is shown in Figure 8, from which it can be seen that the network model tends to converge in the 11th round of training, and the network basically converges in the 60th round, and the loss value of the network tends to be close to 0.00001. It can be seen that the range network model in this paper has good convergence, and the convergence speed is very fast.

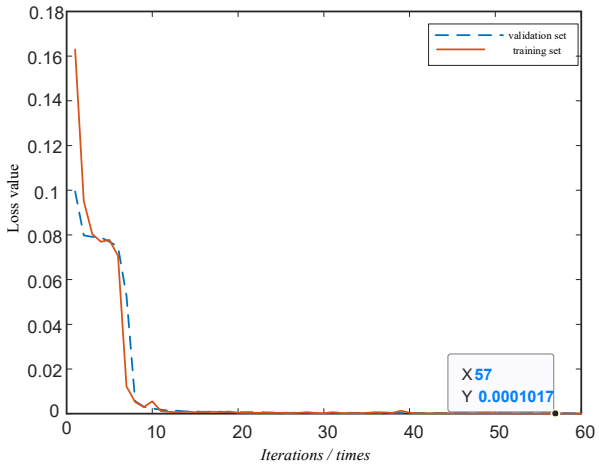


Figure 8. Loss curve

5.2. Fault ranging results analysis

The ranging effect of the CNN-GRU hybrid neural network model is verified using the test set, and the average absolute percentage error and the decidability coefficient of the

ranging results of the test set are calculated according to Eqs. (15) and (16), and the average absolute percentage error of the ranging results of the test set, MAPE, is 0.2182% and the decidability coefficient R^2 is 99.9983%. In order to visualize the effect of fault ranging, the inputs and outputs of the model are normalized to between 0 and 1, indicating the percentage of fault distance to the total length of the transmission line. The inputs and outputs of all the samples in the test set are drawn as a series of points with horizontal and vertical coordinates respectively, and the vertical distance from each point to y=x is the error between the ranging result and the real fault distance, and the smaller the distance is, the better the ranging effect is and the higher the accuracy is. From Figure 9, it can be seen that these points form a straight line approximating y=x, which indicates that the ranging results are almost completely overlapped with the real fault distance, indicating that the ranging effect of the CNN-GRU hybrid neural network model in this paper is better.

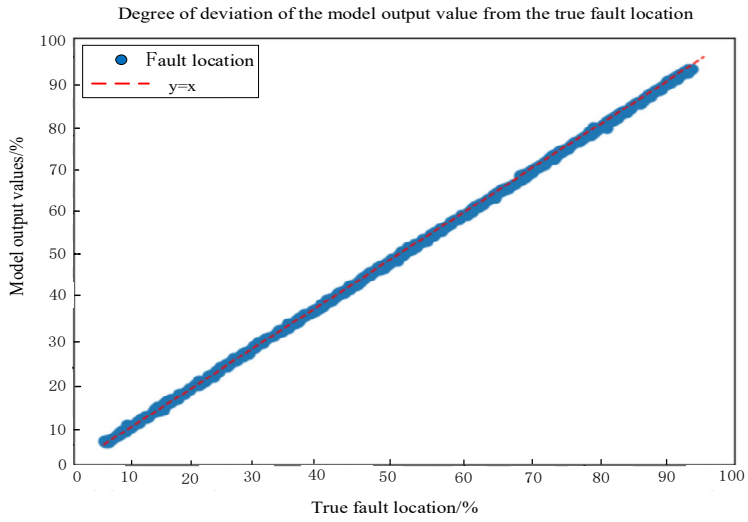


Figure 9. Fault ranging regression curve

The simulation of three types of faults, namely, positive ground fault, negative ground fault and bipolar short-circuit fault, is carried out on the L_1 line, and the fault distances are set to 10km, 40km, 80km and 120km, and the transition resistances are set to 30Ω, 60Ω, 90Ω, 120Ω and 150Ω, respectively.A total of 3×4×5=60 sets of fault sample data are obtained, which are feature extracted and input into the The fault ranging model is tested, and the ranging results are shown in Table 2, with an average absolute percentage error MAPE of 0.21%.

Table 2. Fault ranging results

Fault distance (km)	transition resistance(Ω)	Positioning results (%)	absolute error(km)	relative error(%)
10	30	9.947	0.053	0.265
	60	10.134	0.134	0.67
	90	9.949	0.051	0.255
	120	9.812	0.188	0.94
	150	9.874	0.126	0.63
40	30	40.018	0.018	0.09
	60	40.058	0.058	0.29
	90	39.872	0.128	0.64
	120	39.858	0.142	0.71
	150	39.804	0.194	0.97
80	30	80.107	0.107	0.535
	60	79.923	0.077	0.385
	90	80.048	0.048	0.24
	120	79.950	0.05	0.25
	150	80.181	0.181	0.905
120	30	119.882	0.118	0.59
	60	119.842	0.158	0.79
	90	119.923	0.077	0.385
	120	120.011	0.011	0.055
	150	119.802	0.198	0.99

6. Performance Analysis

6.1. Anti-noise performance analysis

In practical engineering, data are often subject to various disturbances that introduce noise, and the presence of noise reduces the ranging accuracy of the network model, so it is necessary to analyze the network model's ability to resist noise. In the field of data processing, the signal-to-noise ratio SNR is often used to measure the level of noise in the data. A higher SNR means that the signal is stronger relative to the

noise, whereas a lower SNR indicates a relatively large amount of noise, which may negatively affect data analysis and model prediction. In this paper, we add different intensities of noise to the dataset to analyze the effect of noise on the network model, and the relationship between the mean absolute percentage error MAPE and the decidability coefficient and SNR for the dataset with different intensities of noise are shown in Figure 10 and Figure 11, respectively.

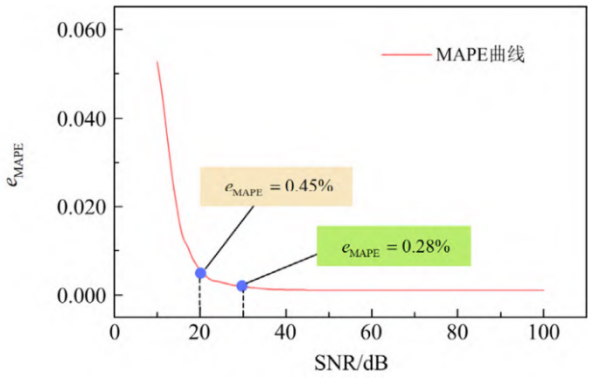


Figure 10. MAPE variation diagram under different levels of noise

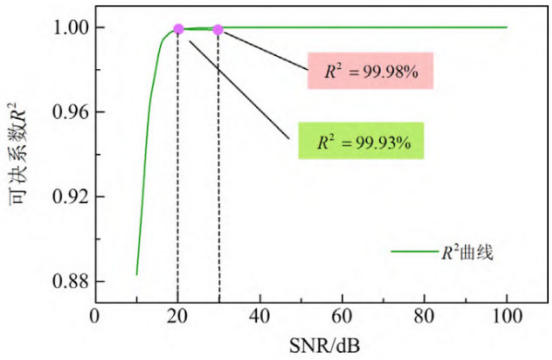


Figure 11. R^2 variation diagram under different levels of noise

As can be seen from Figure 10 and Figure 11, the mean absolute percentage error MAPE decreases with the weakening of the noise, and the coefficient of determination R^2, on the contrary, increases with the weakening of the noise. When the noise is 30db, the average absolute percentage error MAPE is 0.28%, and the coefficient of determination R^2 is 99.98%; when the noise is 20db, the average absolute percentage error MAPE is still 0.45%, and the coefficient of determination R^2 is 99.93%; when the noise is further strengthened, the average absolute percentage error MAPE rises more, and the coefficient of determination R^2 decreases more quickly. It shows that the maximum noise intensity of the fault ranging model in this paper is 20db, which is strong in noise resistance and can meet the requirements of actual ranging.

6.2. Transition resistance performance analysis

The amplitude of the wavelet energy spectrum will be affected by the transition resistance, the data is normalized before inputting into the network training, but it is still necessary to test the transition resistance on the fault ranging accuracy will be affected, so the high resistance test is carried out to verify that the CNN-GRU hybrid neural network model in this paper can withstand the transition resistance ability. For line L_1, four fault distances are selected, and the fault data of 10 different transition resistances (150Ω to 600Ω, with a step size of 50Ω) are collected for each fault distance, and the features are extracted and input into the network model for testing. The obtained ranging results are shown in Table 5-2, the localization results are the average of the test results of multiple trials, the average error is the average error of the test results of multiple trials, and the maximum error is the maximum value of the error in multiple trials.

From Table 3, it can be seen that the CNN-GRU hybrid neural network model proposed in this paper still has high ranging accuracy when the maximum transition resistance is 600Ω. The maximum ranging error distances under several different fault distances are not more than 1 km, and the maximum ranging error is not more than 0.2%. Therefore, the CNN-GRU hybrid neural network model proposed in this paper has a certain ability to resist transition resistance.

Table 3. High resistance fault ranging results

Fault distance（km）	Ranging results（km）	average error（%）	maximum error（%）
50	50.465	0.093	0.127
69	68.17	0.166	0.182
118	117.455	0.109	0.161
200	200.57	0.114	0.127

7. Conclusions

For the fault ranging of transmission lines in flexible DC grid, a CNN-GRU hybrid neural network single-ended intelligent fault ranging method based on wavelet energy spectrum is proposed. The wavelet energy spectrum of the line mode voltage is used as the fault feature vector, the attention weight is obtained by optimizing the multi-attention mechanism using the beluga algorithm, the attention weight is multiplied with the fault feature vector to obtain the weighted data and input into the convolutional general network to extract the features, and finally input into the gated recurrent unit to extract the timing features, and the output prediction results are used as the ranging results. Using the strong nonlinear fitting ability of the gated loop unit, we are able to quickly extract effective information and store key information, which is excellent in the flexible DC transmission line ranging algorithm, and at the same time, the use of the wavelet energy spectrum as the feature vector is not subject to the limitations of the system boundaries, which is able to solve the problems of unreliable identification of the traveling waveheads and inaccurate arrival time of the traveling waves.

Comprehensive experimental results show that the CNN-GRU hybrid neural network model proposed in this paper has better robustness compared with the shallow neural network used in traditional intelligent fault ranging methods. The ranging model has less influence on the fault type, has stronger resistance to noise interference and transition resistance, and can still accurately realize fault ranging in the face of high resistance faults of 600Ω.

Acknowledgment

Sichuan University of Science and Engineering Graduate Student Innovation Fund(Y2022114)

References

[1] W Y Fang .Research on power transformer fault diagnosis based on hybrid intelligent algorithm. Northesat electric power technology, 2020, 41(01): 12-15.

[2] X Chen, L H Zhang, X R Zhang, et al. An asynchronous fault ranging algorithm for T-type transmission lines considering the characteristics of ranging results[J]. Automation of Electric Power Systems, 2020, 44(21): 132-138.

[3] F Xiao, B D Zhang, W C Sun, et al. A power gird fault diagnosis method based on Petri net combining the analogue and digital information[J]. Journal of Electric Power Science and Technology, 2022, 37(1): 113-121.

[4] H L Zhang, W Liu, D W Kong, et al. Troubleshooting and treatment of FF and Profibus fieldbus in thermal power plant applications[J]. Process Automation Instrumentation, 2020, 41(07): 102-106.

[5] G Y Zhang, P Y Zhang, X Fang. An overview of fault localization in high-voltage transmission lines[J]. Technology Wind, 2023(03): 86-89+99.

[6] G K Zhao, K Jia, J F Chen, et al. A single terminal fault location method for a DC transmission line based on circuit breaker reclosing[J]. Power System Protection and Control, 2021, 49(7): 48-56.

[7] K Duan, Y F Fan, Y G Wang, et al. A traveling wave ranging method for a DC line based on wave velocity compensation and fault distance approaching its real value gradually[J]. Power System Protection and Control, 2021, 49(11): 70-78.

[8] K Liao, Z Y He, X P Li. Fault location of HVDC transmission line based on the natural frequency of traveling wave[J]. Automation of Electric Power Systems, 2013, 37(3): 104-109.

[9] Desai J P, Makwana V H. A novel out of step relaying algorithm based on wavelet transform and a deep learning machine model[J]. Protection and Control of Modern Power Systems, 2021, 6(4): 500-511.

[10] Ghorbani A, Mehrjerd H. Negative-sequence network based fault location scheme for double-circuit multi-terminal transmission lines[J]. IEEE Transactions on Power Delivery, 2019, 34(3): 1109-1117.

[11] B H Zhu, Y Chen,J J Ma. Wave front steepness based single-ended traveling wave fault location for transmission lines[J]. Automation of Electric Power Systems, 2021, 45(9): 130-135..

[12] S Lin, Z Y He, X P Li, et al. Single terminal fault location by natural frequencies of travelling wave considering its time-domain characteristics[J]. Power System Technology, 2012, 36(7): 243-248.

[13] K Z Liu, H C Shu, J L Yu, et al. Fault location based on wavelet energy spectrum and neural network for ±800 kV HVDC transmission line[J]. Electric Power Automation Equipment, 2014, 34(4): 141-147.

[14] X G Zeng, X L Zhang, H J Ma, et al. Traveling wave fault location method for power grids based on wavelet packet energy spectra[J]. High Voltage Engineering, 2008, 34(11): 2311-2316.

[15] J H He, G M Luo, M X Chen, et al. A research review on application of artificial intelligence in power system fault analysis and location[J]. Proceedings of the CSEE, 2020, 40(17): 5506-5516.

[16] B T Gao, C X Zheng, H Li, et al. Research on fault ranging model of transmission line based on improved ant colony algorithm[J]. Automation of Electric Power Systems, 2022, 43(02): 61-64.

[17] Rohani R, Koochaki A. A Hybrid Method Based on Optimized Neuro-Fuzz System and Effective Features for Fault Location in VSC-HVDC Systems[J]. IEEE access 2020(8): 70861-70869.

[18] Z B Chi, C J Xia, M J Yang. Single-end fault ranging method for flexible straight line based on parameter optimized VMD and TET[J]. Power System Protection and Control, 2024, 52(04): 1-11.

[19] Y Q Wang, S L Chen, R Z Wei, et al. Elman-Adaboost Integrated Artificial Neural Network for High Voltage AC Transmission Line Single-phase Ground Initial Voltage Traveling Wave Modulus Amplitude Ratio for Single-Ended Fault Ranging[J]. Electric Power Science and Engineering, 2023,39(09): 10-19.

[20] T Li, P F Tian, Z T Sun, et al. Double-ended fault ranging method for transmission lines based on relative fault volume offset and improved GWO-SVR[J]. Engineering Journal of Wuhan University, 2024, 57(01): 89-95.