

On-line Fault Diagnosis Method of Large Transformer Winding based on High Frequency Excitation Injection at Load Side



Yong Kang¹, Guochao Qian^{2*}, Kun Yang², Hongwen Liu², Jin Hu², Weiju Dai², Liang Zhu³, Tao Guo⁴, Jun Shi⁵

1. Chuxiong Power Supply Bureau, Yunnan Power Grid Co., Ltd., Chuxiong, 675000, China
 2. Electric Power Research Institute of Yunnan Power Grid Co., Ltd., Kunming, 650214, China
 3. Qujing Power Supply Bureau, Yunnan Power Grid Co., Ltd., Qujing, 665000, China
 4. Honghe Power Supply Bureau, Yunnan Power Grid Co., Ltd., Honghe, 651400, China
 5. Puer Power Supply Bureau, Yunnan Power Grid Co., Ltd., Puer, 651400, China
- Corresponding Author's Email: qiangguochao_1981@163.com

Abstract: Winding defect is one of the important causes of transformer fault. Detecting the winding state of transformer is of great significance to ensure the safe operation of transformer ; In this paper, the on-line monitoring method of large transformer windings injected with high-frequency excitation at the load side is studied. Firstly, an experimental research platform for the deformation of transformer windings injected with high-frequency excitation at the load side is built, and the axial displacement, local warpage and inter cake short circuit faults of windings at different positions are simulated and tested respectively; Secondly, the gray difference statistical features and color features of the frequency response curve of the transformer winding are extracted, and the feature selection to characterize the fault type and region is carried out through the multimodal feature fusion strategy. Finally, the diagnosis method based on white shark optimization support vector machine is studied to diagnose the fault type and region of the transformer ; The results show that the online diagnosis of transformer winding deformation can be realized by using the method of high frequency excitation injection at the load side, and by using the proposed diagnosis model, the accuracy of fault type and region recognition can reach 93% and 91% respectively, which is about 10% higher than that of the traditional model.

Keywords: Transformer ; Winding Fault; Load Side Injection ; Feature Extraction; White Shark Optimization Algorithm; Support Vector Machine

1. Introduction

110kV and above large transformer is a vital power equipment in the power system, winding defects are one of the important causes of transformer failure, jeopardizing the safe and stable operation of the power system [1,2]. Therefore, timely and effective mastery of large transformer winding status is of great significance to ensure the safe operation of the power system [3-5].

At present, for the study of the detection method of transformer winding status, scholars at home and abroad have put forward a variety of winding fault detection methods, and the current detection methods are mainly divided into offline detection and online monitoring. Offline detection methods currently include: short-circuit impedance method, capacitance change method, vibration method, frequency response method, etc. [6-9]. Short-circuit impedance method is simple and convenient to test, but it is easy to be interfered by on-site equipment, and difficult to diagnose minor winding faults and other shortcomings [10]; the capacitance change method is convenient, but it is difficult to diagnose the deformation of winding faults; the vibration method requires multiple vibration sensors, and the entire signal transmission process is complex, and the signal is easily disturbed by the noise interference and aberrations [11]; the frequency response method is developed by the low-voltage impulse method, first proposed by Canadian scholars to be used in the field of testing and online monitoring. It was first proposed by Canadian scholars to be applied to transformer winding fault diagnosis [12], which has the advantages of convenient test operation, lightweight test equipment, high reliability and repeatability, so the frequency response method is used for offline detection of large transformers. Transformer winding state offline detection method exists inherent real-time poor limitations, and there are transformer outage difficult, time-consuming and laborious and other shortcomings. For this reason, scholars at home and abroad for large transformer winding state online monitoring methods have also carried out continuous exploration: scholars at Chongqing University used high-frequency pulse signals injected into the winding via the end screen of the high-voltage casing to monitor winding faults [13]; scholars in Thailand and other scholars will be complementary to the high-frequency current sensor and ultrasonic sensors to achieve on-line monitoring, which greatly improves the accuracy of localization of partial discharges [14]; Iranian scholars have proposed a method based on electromagnetic waves to detect the localization of partial discharge. Scholars proposed a method of online detection of axial displacement of high-voltage transformer windings based on ultra-wideband pulses of electromagnetic waves, which injects ultra-high-frequency short pulses into high-voltage windings and then uses a set of antennae to capture the reflection signals, and realizes the detection of axial displacement of windings by comparing and analyzing the reflection signals in different states of the windings [15]; scholars from Columbia University realized the online application of the low-pressure impulse method by means of the bushing The scholars at Columbia University realized the online application of low-

voltage pulse method, which injects high-frequency pulse signals into the transformer running on-line, thus realizing the online monitoring of the transformer windings [16]. The above research has laid a research foundation for online monitoring of transformer winding status, and the best feasibility is to inject the excitation signal outside the casing on the high-voltage side, but the impedance of the high-voltage side system is small, and the driving voltage power required for the excitation signal injection source is large, so the injection of the excitation signal from the high-voltage side has a poor effect of injection and there is a risk of levitation of the end screen of the casing, which restricts the practical application in the field.

In order to safely and real-time monitor the operation status of transformer windings, this paper investigates the online monitoring method of large transformer windings based on load-side high-frequency excitation injection: firstly, a load-side high-frequency excitation active injection test platform is built, and three winding fault types are simulated to test the test platform; secondly, the test curves are characterized by the multi-modal feature fusion strategy feature extraction method; and finally, the diagnosis method based on white shark optimization support vector machine is used to diagnose the fault types and fault areas of the transformer.

2. Load-side high-frequency excitation active injection test

2.1 Load-side high-frequency excitation active injection scheme

Load-side injection scheme as shown in Figure 1, first in the transformer load side through the capacitive voltage transformer injection of high-frequency excitation signals; then through the current transformer in the medium voltage side and high-voltage side of the capacitive voltage transformer ground line at the collection of high-frequency response signals, and through the Equation (1) calculated to get the amplitude-frequency characteristic curve of the windings; and finally on the frequency response curve to analyze the test to realize the transformer windings online Monitoring.

$$H(f) = 20 \lg[U_{out}(f)/U_{in}(f)] \quad (1)$$

Where: $H(f)$ is the transfer function value when the frequency is f ; $U_{out}(f)$ is the excitation end voltage RMS value when the frequency is f ; $U_{in}(f)$ is the response end voltage RMS value when the frequency is f . According to the power industry standard DL/T 911-2016 and related literature [17,18], the frequency of 100 Hz~1 Hz is selected as the test frequency range of this paper.

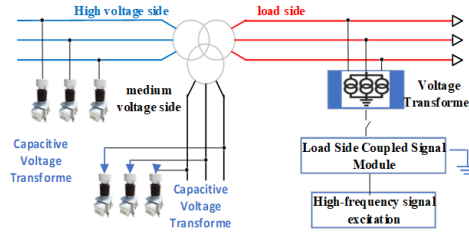


Figure. 1 Load side injection scheme

2.2 Test platform construction

In this paper, an online monitoring platform for transformer windings with load-side high-frequency excitation active injection is built, and the online monitoring platform is shown in Figure 2. The test transformer is a three-phase double-winding transformer of 10 kV voltage level, with delta wiring on the low-voltage side and star wiring on the high-voltage side, and the detailed parameters of the test transformer are shown in Table 1.

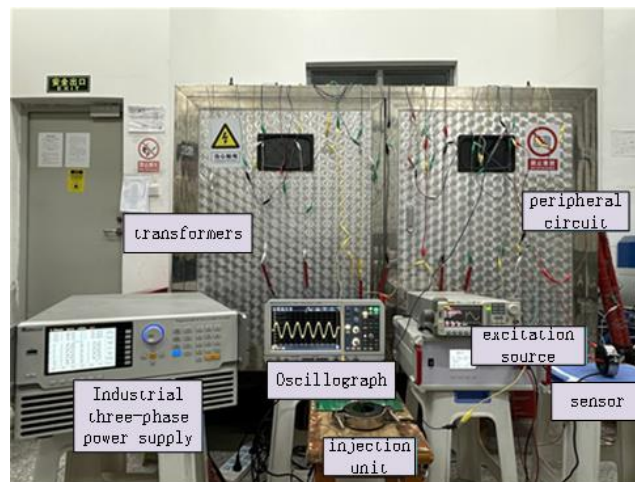


Figure. 2 Schematic diagram of online monitoring platform

Table 1 Main parameters of the test transformer

Transformer components	parameters	Value
Iron core	height	530mm
Double-layered flat bread	Individual height	20mm
	Total height	474mm
Low-voltage winding	Number of Turns per Cake	16 turns
	Outermost radius	89mm
	Width of oil channel	10mm
High-voltage winding	Number of Turns per Cake	32 turns
	Outermost radius	162mm
	Width of oil channel	14mm

In this paper, the resistance and inductance are connected in series to simulate a 10kM transmission line, the selected resistance $R = 2.5\Omega$, inductance $L = 13\text{mH}$; the se-lected resistance $R = 10\text{k}\Omega$ to simulate the low voltage side of the transformer load; through the programmable three-phase power supply of the industrial frequency through the high-voltage side of the simulated transmission line to the transformer high-voltage winding injected into the industrial frequency signal to simulate the transformer on-line operation of the actual operating conditions; wherein the high-voltage side A 10nF capacitor is used to simulate a capacitive voltage transform-er; the grounding wire of the three-phase AC power supply is jointly grounded with the transformer case. In the low-voltage side access to a model CHG-50VA high-precision voltage transformer to carry out the test, the measurement range of 0 ~ 60 V. The peak value of the high-frequency signal in the test is 20 V, using specifica-tions for the 85 * 55 * 25 high-frequency nickel-zinc ferrite ring made of current trans-former to measure the high-frequency current signal.

2.3 Winding fault simulation

2.3.1 Axial shift faults

As shown in Figure 3, the axial shift fault is simulated by adding pads between the cakes, and the thickness of a single pad is about 1.5 mm, simulating 1% axial shift fault requires adding three pads between the cakes. Simulation tests were carried out in the upper part of the winding (1-2 cakes, 3-4 cakes), the middle part (7-8 cakes, 9-10 cakes), the lower part (13-14 cakes, 15-16 cakes), and the axial shift faults with a fault degree of 1%, 3%, and 5% were simulated in the above six positions respectively, with 150 samples, and the tests were repeated three times for each fault, and a total of 54 groups of tests were carried out. The degree of axial displacement fault was calculated ac-cording to Equation (2) [19].

$$\Delta = \frac{\Delta h}{h} \times 100\% \quad (2)$$

Here: Δh is the axial displacement height; h is the total height of the winding.

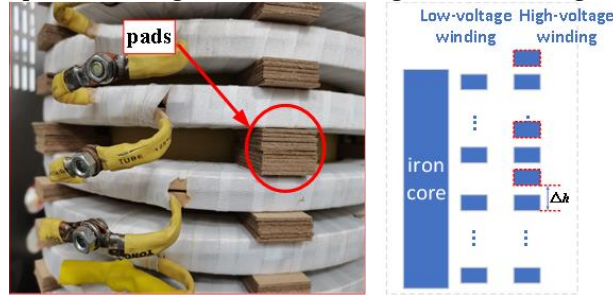


Figure. 3 Simulation diagram of axial displacement fault

2.3.2 Localized Warpage Failure

As shown in Figure 4, local warping faults are simulated by parallel longitudinal capacitance between different cakes, and the normal value of capacitance between cakes is calculated according to the relevant analytical equation [20,21], and the ca-pacitance values of 220pF, 470pF, and 680pF are selected to simulate different degrees of variation of faults such as the spacing between cakes, and the degree of each fault is simulated respectively in the upper part of the windings (1-2 cakes, 3-4 cakes), the middle part (7-8 cakes, 9-10 cakes), lower (13-14 cakes, 15-16 cakes) simulation, 150 samples, each type of fault repeat test 5 times, a total of 90 groups of tests.



Figure. 4 Simulation diagram of local warping fault

2.3.3 short circuit fault between pancakes

As shown in Figure 5, the short-circuit fault between pies is realized by shorting wires to connect the copper nose connectors of the corresponding pies, respectively in the upper part of the winding (1-2 cakes,... , 1-6 cakes), the middle (7-8 cakes, ... , 7-12 cakes), lower (16-13 cakes, ..., 17-18 cakes) for the simulation of short-circuit faults between the cakes, 90 samples, each type of fault repeat test 5 times, a total of 75 groups of test

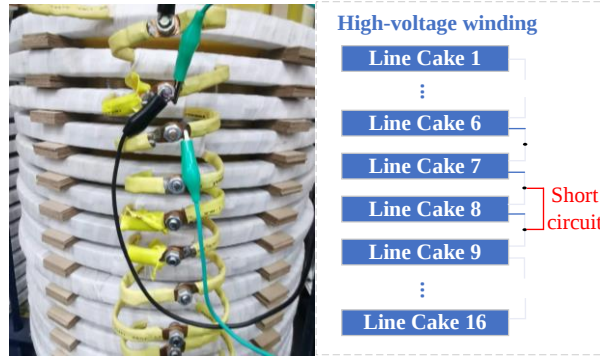


Figure. 5 Simulation diagram of inter cake short circuit fault

2.4 Pilot test results

As shown in Figure 6, this section analyzes the results of the winding axial displacement faults under the injection mode of capacitive voltage transformer as an example. As can be seen from Fig. 6, the high frequency band presents a certain degree of differentiation of the fault region, but still in the phenomenon of mutual interlacing; the frequency band 400~600kHz curve overall shift to the left, with the aggravation of the degree of fault, the curve to the left shift more and more, the fault region shows a certain degree of differentiation, but there are also mutual interlacing, so it is still necessary to further extract the characterization of the winding state.

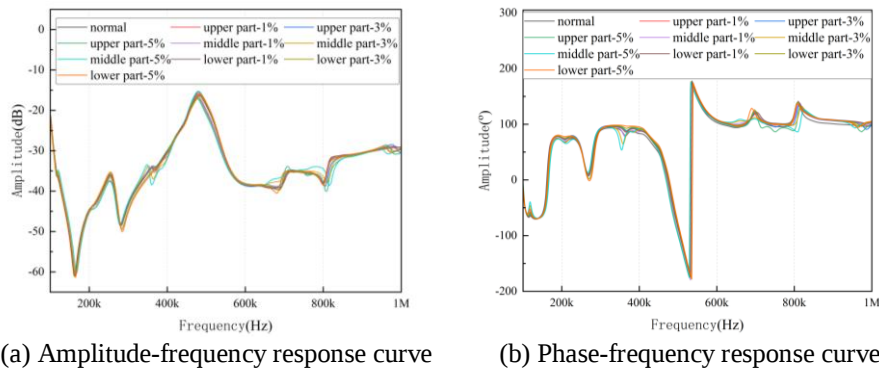


Figure. 6 Frequency response curve under winding axial displacement fault

3. Feature Extraction Methods

3.1 Gray scale difference statistical feature extraction

In this paper, a method of constructing gray-scale feature maps of frequency re-sponse curves is proposed by combining amplitude and phase frequency curves, and the flow is shown in Figure 7.

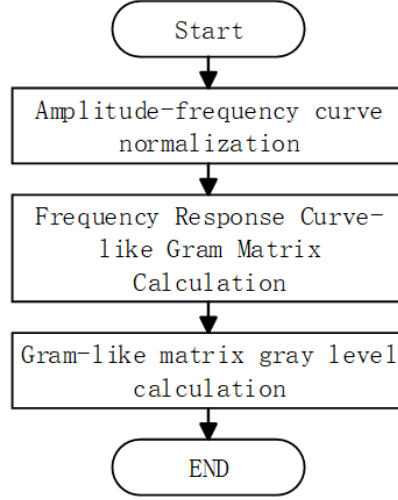


Figure. 7 Construction flow chart of frequency response curve gray characteristic spectrum

(1) Normalization. Avoiding too large an order of magnitude gap in the operation, the amplitude-frequency curve is normalized to between $[-1,1]$, Equation:

$$h(f) = \frac{H(f) - \min(H(f))}{\max(H(f)) - \min(H(f))} \quad (3)$$

(2) Class inner product operation. In order to better quantify the curve variation, the inner product operation is defined as the product of the frequency re-sponse curve difference, as shown in Equation (4):

$$x(f) \otimes y(f) = (h_x(f) - h_y(f)) * \cos(\eta_x(f) - \eta_y(f)) \quad (4)$$

Here $x(f)$ is the reference frequency response curve, $y(f)$ is the test frequency response curve, $h_x(f)$ is the reference normalized amplitude-frequency curve, $\eta_x(f)$ is the reference phase-frequency curve, $h_y(f)$ is the test normalized amplitude-frequency curve, and $\eta_y(f)$ is the test phase-frequency curve.

(3) Class Gram matrix generation. The class inner product operation is performed on the discrete points of the reference and test frequency response curves to generate the frequency response curve class Gram matrix G , as shown in Equation (5).

$$G = \begin{Bmatrix} x(1) \otimes y(1) & x(2) \otimes y(1) & \cdots & x(1) \otimes y(1000) \\ x(2) \otimes y(1) & x(2) \otimes y(2) & \cdots & x(2) \otimes y(1000) \\ \cdots & \cdots & \ddots & \vdots \\ x(1000) \otimes y(1) & x(2) \otimes y(2) & \cdots & x(1000) \otimes y(1000) \end{Bmatrix} \quad (5)$$

where $x(1) \otimes y(1)$ indicates the inner product operation of the frequency response curve class at 100 kHz, and G is the Gram matrix of the frequency response curve class.

(4) Gray scale mapping generation. In order to make the degree of black and white changes in the grayscale map obvious, 16-bit grayscale levels are used, i.e., all black to all white is divided into 16 levels, and according to Equation (6) the G matrix is normalized to be between $[0,15]$, and the generated grayscale map is shown in Figure 8.

$$g = 15 * \frac{G - \min(G)}{\max(G) - \min(G)} \quad (6)$$

Figure. 8 Gray characteristic spectrum of frequency response curve

In the grayscale mapping, when the winding is normal, according to Equation (4), it can be known that the whole of the diagonal of the Gram-like matrix is 0, i.e., the left diagonal of the grayscale feature mapping is all-

black, which can reflect the changes in the frequency response curve of the winding. The gray-scale feature map of the frequency response curve consists of pixel gray-scale values, and its gray-scale statistical information reflects the gray-scale distribution and change characteristics among image pixels, revealing spatial distribution and structural information [22]. Therefore, in this paper, Grey-scale Differential Statistics (GSDS) is used to quantitatively describe the frequency response curve gray-scale feature maps.

Gray scale difference statistics are characterized by calculating the gray scale difference $gray_{\Delta}(x, y)$, as shown in Equation (7).

$$gray_{\Delta}(x, y) = gray(x, y) - gray(x + \Delta x, y + \Delta y) \quad (7)$$

Where (x, y) is the pixel point coordinates, $gray_{\Delta}(x, y)$ is the pixel point magnitude, $\Delta x, \Delta y$ is the gray scale difference shift direction.

The pixel point (x, y) is shifted over the whole image and the histogram of the gray level difference is counted, i.e., the probability of each value of the gray level difference, $P_{\Delta}(i)$ ($i=0, 1, 2, \dots, 15$), and the image texture features are extracted by the probability $P_{\Delta}(i)$:

1) average value:

$$Mea = \sum_{i=0}^{15} iP_{\Delta}(i) \quad (8)$$

2) contrast:

$$Con = \sum_{i=0}^{15} i^2 P_{\Delta}(i) \quad (9)$$

3) Second-order moments in the angular direction

$$Asm = \sum_{i=0}^{15} (P_{\Delta}(i))^2 \quad (10)$$

4) Entropy

$$Ent = -\sum_{i=0}^{15} P_{\Delta}(i) \log_2 P_{\Delta}(i) \quad (11)$$

3.2 Color feature extraction

The pseudo-color technique enhances the image by converting the grayscale map into color map to make the details clearer and to facilitate the extraction of features. In this study, pseudo-color image processing is carried out using the density layering method, where the grayscale map is divided into 15 intervals, each corresponding to a different color, to enhance the image features. Specifically, 16 levels of gray scale are uniformly divided to form 15 colors, which transition from dark blue to dark red, as shown in Equation 9.

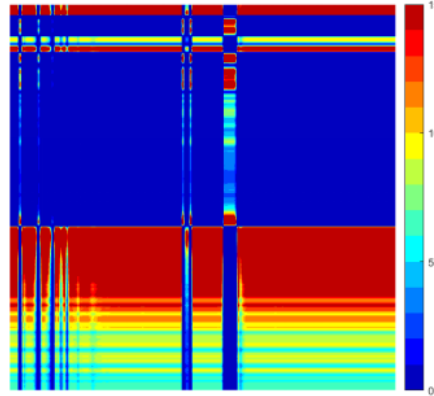


Figure. 9 Pseudo color characteristic spectrum of frequency response curve

The pseudo-color feature mapping in Figure 9 enriches the image information by converting the white region of the grayscale feature mapping in Figure 8 to a region with distinct colors through pseudo-color processing. The pseudo-color feature mapping is mainly reflected in the color depth and distribution characteristics, and the color features of the image can effectively extract relevant information. The advantage of color features lies in the weak dependence on image size, direction and viewing angle, and high stability, and the commonly used features include color moments and color aggregation vector features. Color moments can comprehensively describe the overall information of the image, and the image color information is mainly concentrated in the first, second and third order moments, which describe the average degree of color, variance and skewness, respectively. The color moments of the three color channels of the integrated image define the feature parameters as shown in Equation (12) to (15):

$$Col_1 = \frac{1}{3} \sum_{j=1}^3 E_j \quad (12)$$

$$Col_2 = \frac{1}{3} \sum_{j=1}^3 \left(\frac{1}{N} \sum_{i=1}^N (p_{ij} - E_j)^2 \right)^{1/2} \quad (13)$$

$$Col_3 = \frac{1}{3} \sum_{j=1}^3 \left(\frac{1}{N} \sum_{i=1}^N (p_{ij} - E_j)^3 \right)^{1/3} \quad (14)$$

$$E_j = \frac{1}{N} \sum_{i=1}^N p_{ij} \quad (15)$$

Where: p_{ij} is the i^{th} pixel value of the j^{th} color component; N is the total number of pixels in the image; E_j is the average pixel value of the j^{th} color component.

Color Coherence Vector (CCV) is constructed by dividing the connected regions by pixel connectivity and classifying each region as either cohesive or non-cohesive based on the total number of pixels in each region compared to a threshold value. Its calculation steps are as follows:

1) Quantization: each color component is uniformly quantized into n color intervals to generate a matrix of n -level color pixel values;

2) Division of connected regions: the matrix of quantized color pixel values is divided into a number of connected regions based on pixel connectivity;

3) Judgment of aggregation: count the number of pixels N_c in each connected region C and determine whether C is aggregated or not with a threshold ε (usually 1% of the total number of pixels N), as shown in Equation (16).

$$\begin{cases} N_c \geq \varepsilon & \text{Aggregated Pixel D} \\ N_c < \varepsilon & \text{Non-aggregated pixels W} \end{cases} \quad (16)$$

4) Aggregation vector: count the aggregated pixels D and non-aggregated pixels W in each connected region of level n according to the above judgment basis;

5) Feature parameter: the feature distance parameter Dw quantifies the difference between the CCVs of different images:

$$Dw = \frac{\sum_{j=1}^3 \sum_{i=1}^n (|D_{ij} - d_{ij}| + |W_{ij} - w_{ij}|)}{3 \sum_{j=1}^3 \frac{1}{N} \sum_{i=1}^n (|D_{ij} + W_{ij}|)} \quad (17)$$

Where: d_{ij} is an aggregated pixel at level i of the j^{th} color component of the reference image; D_{ij} is an aggregated pixel at level i of the j^{th} color component of the comparison image; w_{ij} is a non-aggregated pixel at level i of the j^{th} color component of the reference image; W_{ij} is a non-aggregated pixel at level i of the j^{th} color component of the comparison image.

3.3 feature selection

In this study, the fault types of all samples are screened by Random Forest feature selection method, and the fault type labels are defined as 1 for short-circuit between pancakes, 2 for bulge buckling, and 3 for axial displacement. The importance of each feature for fault classification is shown in Figure 10. Figure 10 shows that all gray differential statistics features have higher importance for winding fault types, and other features such as Col_2 , Col_3 and Dw have more prominent importance.

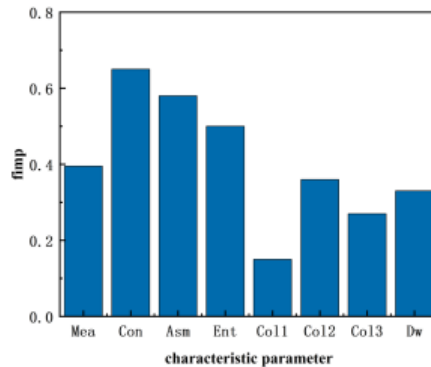


Figure. 10 Importance degree of winding fault type characteristics

In the winding fault region feature selection, the fault region labels are defined as top-1, middle-2 and bottom-3. The importance of features for fault regions under each fault type is shown in Figure 11. The figure shows that all the color features are more important for all types of winding fault regions and other feature parameters are more important for short circuit fault regions between pancake.

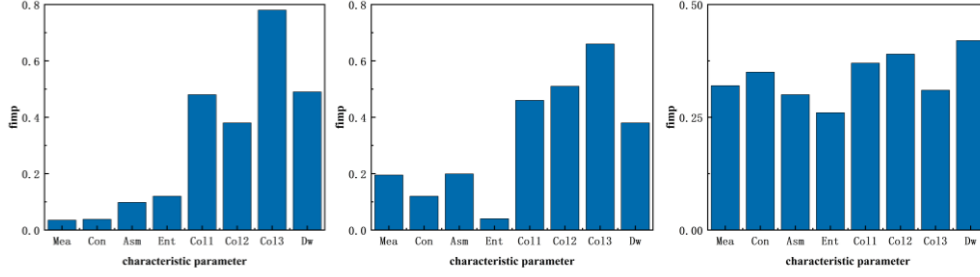


Figure. 11 Importance degree of winding fault area characteristics

According to the above random tree-based feature selection, the top-ranked features in terms of importance are selected as the effective characterization feature parameters of the winding state, as shown in Table 2.

Table 2 Effective characteristic parameters of winding state

Failure Properties		Grayscale Difference Statistical Features	Color features
Fault classification		Mea、Con、Asm、Ent	\
Fault area	Bugle warpage	\	Col_1 、 Col_2 、 Col_3 、 Dw
	Axial displacement	\	Col_1 、 Col_2 、 Col_3 、 Dw
	Inter cake short circuit	Mea、Con	Col_1 、 Col_2 、 Col_3 、 Dw

4. White Shark-Optimized SVM Diagnosis

4.1 White Shark Optimization Algorithm

White Shark Optimization Algorithm (WSO) is a bionic intelligent algorithm that simulates the hunting behavior of white sharks, and balances the global search and local exploitation through olfactory tracking, auditory localization and group collaboration mechanisms, effectively avoids the local optimum and achieves the search for the global optimal solution of the parameter optimization problem, and the White Shark Optimization Algorithm is mainly divided into the following stages:

1) Initialization: d is the total number of decision variables to be solved in the search space, then the white shark population w of n populations is shown below:

$$w = \begin{Bmatrix} w_1^1 & w_2^1 & \cdots & w_d^1 \\ w_1^2 & w_2^2 & \cdots & w_d^2 \\ \cdots & \cdots & \ddots & \vdots \\ w_1^n & w_2^n & \cdots & w_d^n \end{Bmatrix} \quad (18)$$

$$w_j^i = l_j + r \cdot (u_j - l_j) \quad (19)$$

Where: w_j^i is the location of white sharks in the search space, n is the total population size, u_j and l_j are the upper and lower bounds of the search space.

2) Rapid movement towards prey: When the white shark senses the position of its prey based on the fluctuations of the waves it hears when the prey is moving, it will move towards the prey in a fluctuating motion, as shown in Equation (20):

$$v_{k+1}^i = \mu[v_k^i + p_1(w_{gbestk} - w_k^i) \cdot c_1 + p_2(w_{gbest}^{v^i} - w_k^i) \cdot c_2] \quad (20)$$

where v_{k+1}^i represents the new velocity of the i^{th} white shark at $k+1$ steps, v_k^i represents the current velocity of the i^{th} white shark at k steps, w_{gbestk} represents the global best position vector obtained by the white shark in k iterations, w_k^i represents the position of the i^{th} white shark at k steps, v^i represents the i^{th} index vector at which the white shark arrives at the optimal position of the shark as defined in Equation (21), $w_{gbest}^{v^i}$ represents the position vector of the index vector at k steps, c_1 and c_2 are random numbers in the range of $[0,1]$, p_1 and p_2 are parameters controlling and influencing the variation of the white shark as shown in Equation (22) and (23), and p_1 and p_2 are parameters controlling and influencing the variation of the white shark as shown in Equation (22) and (23), respectively. c_1 and c_2 are random numbers in the range of $[0,1]$; p_1 and p_2 are parameters controlling and influencing the changes of the white shark, as shown in Equation (22) and (23).

$$v = \lfloor n \cdot rand(1, n) \rfloor + 1 \quad (21)$$

$$p_1 = p_{\max} + (p_{\max} - p_{\min}) \cdot e^{-(4k/K)^2} \quad (22)$$

$$p_2 = p_{\min} + (p_{\max} - p_{\min}) \cdot e^{-(4k/K)^2} \quad (23)$$

where $\text{rand}(1,n)$ is a vector of random numbers evenly distributed between $[0,1]$, k and K represent the current iteration number and the maximum iteration number, and p_{min} and p_{max} denote the initial and following velocities of the white shark's movement, which usually take the values of 0.5 and 1.5, respectively.

μ controls the search behavior of the white sand algorithm, which also helps to avoid early convergence and precludes the solution from falling into a local optimum, as shown in Equation (24).

$$\mu = \frac{2}{\left| 2 - \tau - \sqrt{\tau^2 - 4\tau} \right|} \quad (24)$$

where τ denotes the acceleration coefficient, which takes the value of 4.125.

3) Surrounding the best prey: The position of the white shark is constantly changing during the search for prey, and this hunting behavior of the white shark is simulated by Equation (25).

$$w_{k+1}^i = \begin{cases} w_k^i \cdot \neg \oplus w_o + u \cdot a + l \cdot b & \text{rand} < mv \\ w_k^i + v_k^i / f & \text{rand} \geq mv \end{cases} \quad (25)$$

Where w_{k+1}^i is the new position vector of the i^{th} white shark in the $k+1^{st}$ step; $\neg$ is the negation operator, $\oplus$ is the bit-by-bit dissimilarity operation; a and b are one-dimensional binary vectors as defined in Equation (26) and (27); w_o is the logic vector defined in Equation (28); f represents the frequency of the white shark's wave motion, as shown in Equation (29); mv represents the increase in the movement force of a white shark approaching its prey, which expresses the white shark's auditory and olfactory intensity, as an iterative increasing function, as shown in Equation (30):

$$a = \text{sgn}(w_k^i - u) > 0 \quad (26)$$

$$b = \text{sgn}(w_k^i - l) < 0 \quad (27)$$

$$w_o = \oplus(a, b) \quad (28)$$

$$f = f_{\min} + \frac{f_{\max} - f_{\min}}{f_{\max} + f_{\min}} \quad (29)$$

$$mv = \frac{1}{a_0 + e^{(K/2-k)a_1}} \quad (30)$$

In the Equation, $f_{\min}$ and $f_{\max}$ are the minimum and maximum frequencies of the white shark's fluctuating movement, which take the values of 0.07 and 0.075, respectively; a_0 and a_1 denote the white shark's sense of hearing and olfaction, which take the values of 6.25 and 100. The white sharks search locally in the smallest time, and globally in the opposite way. Therefore, the values of a_0 and a_1 can control the search efficiency of white sharks to locate prey effectively.

When $\text{rand} < mv$, it mainly indicates that white sharks randomly change their position in many directions and areas when hunting for the best prey.

When $\text{rand} > mv$, it mainly indicates that white sharks can update their position in the search space based on the sounds they hear from their prey.

4) Approach to the best white sharks: Once the white sharks have found the prey and surrounded the prey, they approach to their best attack position to kill the prey as shown in Equation (31):

$$w_{k+1}^i = w_{gbest} + r_1 D_w \text{sgn}(r_2 - 0.5) \quad r_3 < s_s \quad (31)$$

$$D_w = \left| \text{rand} \cdot (w_{gbest} - w_k^i) \right| \quad (32)$$

$$s_s = \left| 1 - e^{(-a_2 \cdot k/K)} \right| \quad (33)$$

Where w_{k+1}^i is the updated position of the i^{th} white shark relative to the location of the prey, and $\text{sgn}(r_2 - 0.5)$ is given as 1 or -1 to change the direction of search; the variables r_1 , r_2 , and r_3 are random numbers in the range of $[0,1]$; D_w is the distance between the prey and the white sharks; s_s is used to represent the olfactory and visual intensities of the white sharks as they follow other white sharks close to the best prey, and a_2 is a control for the normal number of white shark hunting intensity, which takes the value of 0.0005.

4) Fish School Behavior: The schooling behavior of white sharks is defined according to Equation (34):

$$w_{k+1}^i = \frac{w_k^i + w_{k+1}^i}{2 \cdot \text{rand}} \quad (34)$$

The final position of the white shark is one that will be ideally located around prey in the search space, very close to the best prey. Fish schooling behavior and movement toward the best white sharks suggests collective WSO behavior, which expands the range of E&P features.

4.2 White Shark Optimization Support Vector Machine

Support vector machine (SVM) uses kernel mapping to upgrade low-dimensional data to linearly differentiable space to construct an optimal classification hyperplane, and its performance is synergistically regulated by the kernel function parameter, which determines the complexity of the feature space, and the penalty factor C , which balances the model's fault-tolerance and generalization ability, and the joint optimization of the two can significantly improve the classification accuracy. The white shark optimization algorithm simulates the sensory and schooling behaviors of white sharks, and identifies the global optimal solution quickly and accurately by constantly searching for the best shark and target location, which has superior search capability and is suitable for SVM parameter optimization. The SVM flow based on the white shark algorithm optimization is shown in Figure 12.

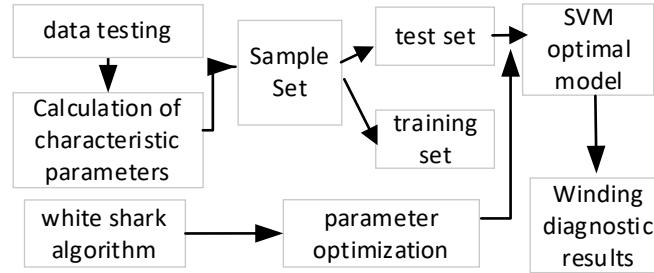


Figure. 12 Support vector machine optimized by white shark algorithm

4.3 Diagnostic process and methodology

In this paper, based on the load-side injection of high-frequency excitation signals, the frequency response test of the winding under different fault conditions is carried out, and the specific steps are as follows:

- (1) Carry out the frequency response test of the winding under normal conditions according to the test experiment platform. The data are collected once every 5 kHz, 181 times, and 2504 moments are collected at each frequency.
- (2) Set up different fault conditions of the winding and repeat step (1), the fault setting samples include: 150 samples of axial displacement faults, 150 samples of local warping faults, and 90 samples of short-circuit faults between cakes.
- (3) Mathematical statistical features, waveform features, gray scale differential statistical feature extraction and color feature extraction are performed on the frequency response curves under different fault conditions and feature selection is performed.
- (4) The optimally selected feature quantities are input into the white shark optimization support vector machine algorithm for fault diagnosis of transformer windings;
- (5) Analyze the test results.

4.4 Testbed-based results

4.4.1 Diagnosis of winding fault types

Based on the 390 samples obtained from the transformer winding experimental platform, 90 samples of short-circuit faults between pancake and set the label as 1, 150 samples of bulging and warping faults set the label as 2, and 150 samples of axial displacement faults set the label as 3, as shown in Table 3.

Table 3 Fault type sample label

Fault type	Sample size	Training set	Test set	Type label
Inter cake short circuit	90	273	117	1
Bulge warpage	150			2
Axial displacement	150			3

For the winding fault diagnosis, the data was divided into training and validation sets in a ratio of 7:3; specifically, the training set contained 273 samples while the test set contained 117 samples. In this study, the features *Mea*, *Con*, *Asm* and *Ent* screened in the previous section were used as input variables for the model. The SVM is trained using the White Shark optimization algorithm to improve the accuracy of fault type identification. The final result of the confusion matrix for fault prediction is displayed in Figure 13.

As can be seen from the graph, the support vector machine optimized by the White Shark algorithm improves the accuracy, precision, and sensitivity in the identification of winding fault types by about 10%, which proves the effectiveness of the White Shark algorithm in optimizing the parameters. Specifically, the accuracy of identification of winding faults reached 93.2%, with the lowest accuracy for bulging warpage faults, but also exceeding 88%, and the lowest sensitivity for axial displacement faults, which also exceeded 91%. These results show that the classification model built by the support vector machine optimized by the White Shark algorithm can effectively identify the type of winding faults.

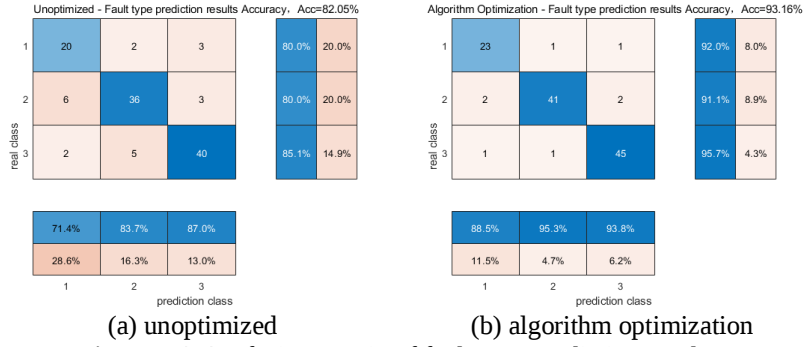


Figure. 13 Confusion matrix of fault type prediction results

4.4.2 Winding fault area diagnosis

The fault region is divided into three parts: top labeled -1, middle labeled -2, and bottom labeled -3. The training set, validation set, and test set are assigned according to the ratio of 7:3. For the short-circuit fault between pies, there are 63 samples in the training set and 27 samples in the test set, while for the bulge warping and axial displacement faults, the training set contains 105 samples and the test set contains 45 samples. The detailed allocation can be found in Table 4.

The confusion matrix for fault region prediction is displayed in Figure 14 to Figure 16. Without optimization, the recognition accuracy is about 80%, with a sensitivity of about 68% for the bulge-warp fault region and about 70% for the short-circuit fault region between pies. Applying the optimized support vector machine with the White Shark algorithm, the overall recognition accuracy is improved by about 10%, both exceeding 91%, while the accuracy and sensitivity are also improved. These results further confirm that the classification model built based on the support vector machine optimized by the White Shark algorithm can effectively distinguish the winding fault areas under different fault types.

Table 4 Winding area sample label

Fault type	Fault area	Sample size	Training set	Test set	Type label
Inter cake short circuit	Upper part	30	63	27	1
	Middle part	30			2
	Lower part	30			3
Bulge warpage	Upper part	50	90	45	1
	Middle part	50			2
	Lower part	50			3
Axial displacement	Upper part	50	90	45	1
	Middle part	50			2
	Lower part	50			3

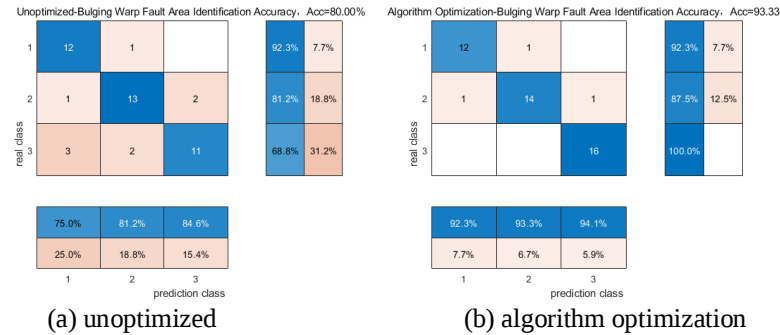


Figure. 14 Confusion matrix of prediction results of bulge warpage fault area

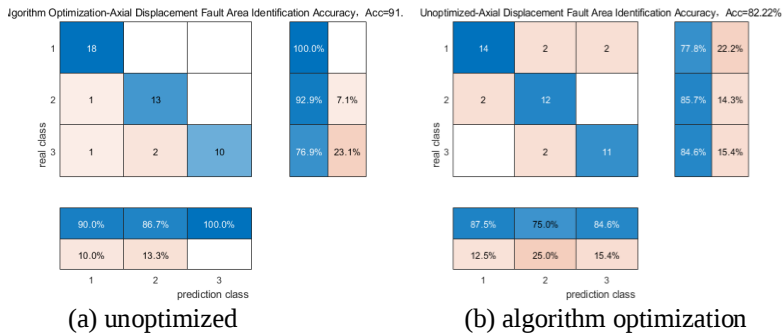


Figure. 15 Confusion matrix of prediction results of axial displacement fault area

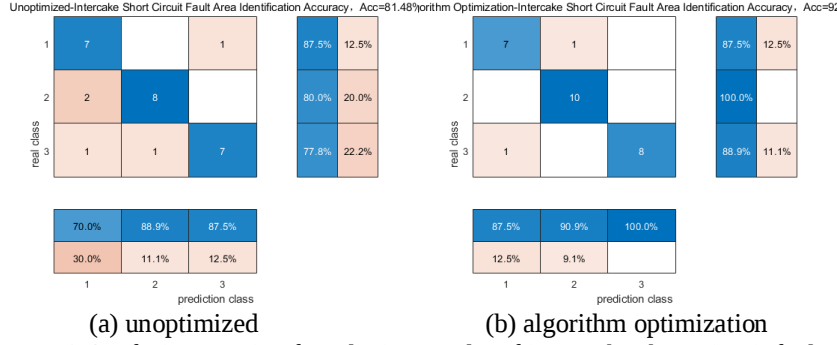


Figure. 16 Confusion matrix of prediction results of inter cake short circuit fault area

5. Conclusions

This paper carries out a research on the online diagnosis method of large transformer winding faults injected by load-side high-frequency excitation, and the following conclusions are obtained:

(1) Based on the extracted gray-scale differential statistical features and color features of the winding frequency response curve, the influence weights of each feature parameter on the winding fault types and fault regions are evaluated by the Random Forest feature selection algorithm, respectively, and in terms of fault types, the importance of all the gray-scale differential statistical features is higher, and the importance of the Col_2 , Col_3 , and Dw features is also more prominent; and in terms of fault regions, the importance of all the color features have higher importance for the three winding fault regions, and Mea and Con features have higher importance only for the short circuit fault region between cakes.

(2) Based on the fact that the combination of the White Shark optimization algorithm and SVM can significantly improve the classification accuracy, intelligent diagnosis of winding fault categories and locations is achieved by using the screened important feature parameters of the windings as inputs to the support vector machine model optimized by the White Shark algorithm. The experimental results show that the proposed diagnostic model has an identification accuracy of 93% for the fault type and more than 91% for the fault area, and the accuracy is improved by about 10% compared with the traditional model.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

This research was funded by the Electric Power Research Institute of Yunnan Power Grid Co., Ltd. (Project No. YNKJXM20222300) and Chuxiong Power Supply Bureau of Yunnan Power Grid Co., Ltd. (Project No. YNKJXM20222330).

References

- [1] Asadi N; Kelk H M. Modeling, analysis, and detection of internal winding faults in power transformers. IEEE Transactions on Power Delivery, 2015, 30(6), 2419-2426.
- [2] Zhou Lijun; Yun Xiucheng; Wang Dongyang ; Zhou Meng; Zhang Jun. Transformer Fault Winding Identification Method Based on Voltage Damped Oscillation. Transactions of China Electrotechnical Society, 2024, 39, 3218-3231.
- [3] Ahmadi, SA; Sanaye-Pasand, M; Abedini, M; Samimi, MH. Online Sensitive Turn-to-Turn Fault Detection in Power Transformers. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 2022, 69(12), 13555-13564.
- [4] S. Bagheri;Z. Moravej and G. B. Gharehpetian. Effect of transformer winding mechanical defects, internal and external electrical faults and inrush currents on performance of differential protection. IET GENERATION TRANSMISSION & DISTRIBUTION, 2017, 11, 2508-2520.
- [5] Zhao, XZ; Yao, CG; Zhou, ZH; Li, CX; Wang, XM; Zhu, TY; Abu-Siada, A. Experimental Evaluation of Transformer Internal Fault Detection Based on V-I Characteristics. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 2020, 67, 4108-4119.
- [6] Dong, L; Wang, G; Qi, X. Power Transformer Winding Deformation Diagnosis Based on the Sum of Two-Way Shortest Distance. IEEE TRANSACTIONS ON ELECTRICAL AND ELECTRONIC ENGINEERING, 2023, 18, 1869-1875.

- [7] Yao, CG; Zhao, ZY; Mi, Y; Li, CX; Liao, YF; Qian, GC. Improved Online Monitoring Method for Transformer Winding Deformations Based on the Lissajous Graphical Analysis of Voltage and Current. IEEE TRANSACTIONS ON POWER DELIVERY, 2015, 30, 1965-1973.
- [8] Lu Fangcheng; Wang Xinyu; Wang Ping; Gao Shuguo; Geng Jianghai. Mechanical Deformation Identification of Transformer Winding under Multiple Short-Circuit Impacts Based on Vibration Deviation and Weighted Entropy. Transactions of China Electrotechnical Society, 2023, 38, 3022-3032.
- [9] Jiang Junfei; Zhou Lijun; Li Wei; Liu Huancheng; Wang Dongyang. FRA Modeling and Fault Winding Identification of Autotransformer With Split Windings. Proceedings of the Chinese Society of Electrical Engineering, 2018, 38, 3102-3108.
- [10] M. Wang, et al. Review of condition assessment power transformers in service. IEEE Electrical Insulation Magazine, 2002, 18(6), 12-25.
- [11] Zhang Fan; Ji Shengchang; Shi Yuhang; Lu Weifeng; Li Dajian; Chen Liangyuan. Research on Transformer Winding Vibration and Propagation Characteristics. Proceedings of the Chinese Society of Electrical Engineering, 2018, 30, 2790-2798.
- [12] E. P. Dick, et al. Transformer diagnostic testing by frequency response analysis. IEEE Transactions on Power Apparatus and Systems, 1978, 97(6), 2144-2153.
- [13] Li Chengxiang; Xia Qi; Zhu Tianyu; Zhao Zhongyong; Yao Chenguo; Mi Yan. Live transformer winding deformation detector based on impulse frequency response method. Electric Power Automation Equipment, 2018, 38, 110
- [14] Thungsuk, N; Mungkung, N; Songruk, A; Tunlasakun, K; Tikakosol, K; Nilawat, S; Boonsomchuae, K; Yuji, T; Arunrungrusmi, S; Kinoshita, H. The Investigation of Detect Position of Partial Discharge in Cast-Resin Transformer Using High-Frequency Current Transformer Sensor and Acoustic Emission Sensor. APPLIED SCIENCES-BASEL, 2022, 12(3), 1310.
- [15] Karami, H; Tabarsa, H; Gharehpetian, GB; Norouzi, Y; Hejazi, MA. Feasibility Study on Simultaneous Detection of Partial Discharge and Axial Displacement of HV Transformer Winding Using Electromagnetic Waves. IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, 2020, 16, 67-76.
- [16] De Rybel, T. Singh, A. Vandermaar, et al. Apparatus for Online Power Transformer Winding Monitoring Using Bushing Tap Injection. IEEE TRANSACTIONS ON POWER DELIVERY, 2009, 24(3), 996-1003.
- [17] DL/T 911-2016, Analysis method of frequency response for winding deformation of power transformers, China Energy Administration, Feb. 2016.
- [18] IEEE Guide for the Application and Interpretation of Frequency Response Analysis for oil Immersed Transformers, Mar. 2013.
- [19] Hashemnia N; Abu-Siada A; Islam S. Improved Power Transformer Winding Fault Detection using FRA Diagnostics—Part 1: Axial Displacement Simulation. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22(1), 556-563.
- [20] Usha, K, et al. Inter Disc Fault Location in Transformer Windings Using SFRA. IEEE TRANSACTIONS ON DIELECTRICS AND ELECTRICAL INSULATION, 2015, 22(6), 3567-3573.
- [21] Abeywickrama; Nilanga, et al. High-frequency modeling of power transformers for use in frequency response analysis (FRA). IEEE TRANSACTIONS ON POWER DELIVERY, 2008, 23(4), 2042-2049.
- [22] Liu Hui; Zhang YunSheng; Zhang YinHui; He ZiFen. Texture Feature Extraction of Flame Image Based on Gray-Scale Difference Statistics. Control Engineering of China, 2013, 20, 213-218.