IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 November 2023, accepted 30 December 2023, date of publication 3 January 2024,
date of current version 9 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3349410

APPLIED RESEARCH

A Novel Fault Diagnosis of GIS Partial Discharge Based on Improved Whale Optimization Algorithm

WEI SUN, HONGZHONG MA, AND SIHAN WANG
College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China
Corresponding author: Hongzhong Ma (18871716279@163.com)

ABSTRACT Partial discharge (PD) seriously affects the operational safety of power equipment. In order to effectively diagnose the PD in gas insulated switchgear (GIS), a GIS PD fault diagnosis method based on improved whale optimization algorithm (IWOA) is proposed, which optimizes variational mode decomposition (VMD) and support vector machine (SVM) to adaptively determine the appropriate parameters and further enhance performance. A laboratory GIS PD platform is built to collect four types of PD fault signals (point discharge, particle discharge, floating discharge, and air-gap discharge). Firstly, a nonlinear arctangent convergence factor and adaptive weight are proposed to address the issue of local optimization in the WOA optimization process. Then, IWOA is used to optimize parameters of VMD (mode parameter K and penalty factor α). Next, effective intrinsic mode functions (IMFs) are screened through correlation coefficients which are greater than 0.2. Because a single scale cannot fully reflect all signal information, and more important information is distributed in other scales, multiscale permutation entropy (MPE) is introduced for feature extraction. Furthermore, the principal component analysis (PCA) method is employed for dimension reduction of initial feature vectors, which reduces the dimension of 33 feature vectors to 7. Finally, SVM based on IWOA is applied to train and test the experimental data to identify different types of PD faults, and achieve diagnosis of GIS PD. Through experimental analysis and comparison with other methods such as EMD-MPE, WOA-VMD-MSE, etc., the proposed method has good diagnostic effects. Also, it proves the robustness and feasibility of the presented solution. The optimization model provides a reference for solving fault diagnosis of GIS PD problems.

INDEX TERMS Partial discharge, GIS, improved whale optimization algorithm, VMD, fault diagnosis.

I. INTRODUCTION

The gas insulated switchgear (GIS) is widely used in power system because of its good insulation performance, low electromagnetic pollution, compact structure, and easy installation [1]. However, with the expansion of GIS construction scale and the increase of operating time in power system, the number of partial discharge (PD) faults in GIS is also increasing, seriously shaking the stable operation of GIS [2], [3].

Effective extraction of feature information that can characterize different types of PD is the foundation for achieving PD identification. At present, the commonly used PD feature information extraction methods are mainly phase distribution method, including phase resolved partial discharge (PRPD) image and phase resolved pulse train (PRPS) image, which mainly extracts the spectral features of PD signals, including discharge phase, frequency, and discharge quantity to achieve the characterization of PD signals [4], [5], [6]. However, the method is easy to cause aliasing problem because of many sources of PD, which affects identification accuracy. In, addition, as the most widely applied artificial intelligence algorithm, the application of neural network has increased

The associate editor coordinating the review of this manuscript and approving it for publication was Mehrdad Saif.

in PD fault diagnosis in recent years [7], [8], [9], [10], [11], [12], [13], [14]. But massive PD experiments are expensive, and the fault sample data is small and limited, making it difficult to meet the large data volume requirements of neural network.

In the application of fault diagnosis, there are many practical data-driven fault diagnosis methods. Wavelet transform and MKELM have showed effective in [15]. In addition, cross-wavelet transform and variational Bayesian matrix factorization [16] and voltage difference analysis [17] have good performance in fault diagnosis. In [18], lower-dimensional features are produced through the kernel entropy component analysis as samples for training and testing a one-against-one least squares support vector machine, which revealed that the proposed approach has higher diagnosis accuracy. The above methods show better accuracy of fault diagnosis. However, there are still some problems, such as the limitation of wavelet basis function, high computational complexity and long computational time.

Empirical mode decomposition (EMD), as an adaptive signal processing method that decomposes a time series into some limited intrinsic mode functions (IMFs). It has already operated in the areas of fault detection, signal processing and data compression [19], [20], [21]. However, due to the problems of endpoint effects and mode mixing in non-stationary signal decomposition, EMD is limited in practical applications. Reference [22] proposed the local mean decomposition (LMD) method. Then, Huang et al. [23] proposed the ensemble empirical mode decomposition (EEMD) method, Yang et al. [24] proposed the ensemble local mean decomposition (ELMD) method and some improved EMD algorithms have also appeared [25], [26]. However, the decomposition error of these algorithms is larger and signal decomposition results are greatly affected by the sampling frequency. Also, they appear in different degrees of mode aliasing phenomenon, which makes it difficult to separate the components with similar frequencies, and the endpoint effect cannot be solved fundamentally.

To overcome the shortcomings of the analysis method mentioned above, variational mode decomposition (VMD) was proposed in 2014 [27]. Compared with EMD, VMD introduces a completely non recursive method to achieve effective frequency domain signal decomposition and component separation of multi-component signals, and VMD converts signal decomposition into non-recursive VMD modes, which has a solid theoretical basis [28]. In [29], VMD algorithm was used in electrocardiogram (ECG), and it has stronger denoising ability and better retains morphological characteristics of the original ECG signals. Reference [30] used VMD in the condition monitoring and fault diagnosis of rotary machinery. A singular kurtosis difference spectrum method was proposed to accurately determine the effective reconstructed order for signal noise reduction. The results demonstrate that the proposed method has superior ability to extract the early weak fault characteristics of the planetary gearbox. Sharma et al. [31] proposed a fast and effective VMD-based fault detection technique for the LVdc distribution system with penetration of renewable sources using the local end current measurements only. The proposed scheme detects all types of fault with wide variation in fault and operating conditions of the microgrid with faster response time. Practical applications have shown that VMD has the advantages of strong signal decomposition ability, good noise resistance, and fast signal processing speed.

Some scholars have also applied VMD into the field of PD. Two methods were proposed for PD signal denoising in [32], namely the VMD based on Shannon entropy and the VMD based on kurtosis-approximation entropy (APEN). Output signal-to-noise ratio (SNR), root mean squared error (RMSE), *etc,* are used as indexes to evaluate the denoising performance of the algorithms, and VMD obviously performs better. Shang et al. [33] proposed a new feature extraction method based on VMD. The original PD signal is decomposed with VMD to obtain IMFs and MDE values in each IMF are calculated. Experiment results demonstrate that, PD feature extraction method based on VMD can extract effective characteristic parameters. In [34], a denoising method combined with singular value decomposition (SVD) and VMD was proposed to eliminate noise in on-site PD signals from high-voltage electrical equipment. The PD signals are decomposed into K-IMFs by VMD, and the K value of white noise's IMF components is determined by EMD. The proposed method can eliminate periodic narrowband interference and white noise in different PD signals effectively. Reference [35] proposed feature extraction of PD in low-temperature composite insulation. Original and noisy signals of three typical PD defects are obtained and decomposed by VMD. Relative moments and grayscale co-occurrence matrix are employed for feature extraction by K-modal component diagram. Then the effectiveness of different feature extraction methods is evaluated by dimensionality reduction and pattern recognition, which shows the advantages of VMD decomposition. The literature [36] combined the VMD and Group-Sparse Total Variation, and applied the method to extract the simulated PD signal buried in white noise, discrete spectral interference and color noise. Simulation results show that the performance of the proposed VMD-GSTV method is superior to that of Wavelet and the recently introduced Wavelet Total Variation method.

Although the above research confirms the superiority of VMD, the mode parameter and penalty factor of some examples are manually set by experience, and cannot adapt to adjust parameters. The settings of K and α are key parameters that affect the result of signal decomposition. For example, if the value of K is too large, it will lead to over-decomposition, and vice versa, it will cause insufficient decomposition. Adopting intelligent optimization algorithm is a good solution, which can adaptively determine the suitable parameters and enhance performance. However, some existing swarm-based optimization algorithms have some disadvantages, such as slow convergence and easy to fall into local optimization. Aiming at these problems, this paper

introduces nonlinear arctangent convergence factor and adaptive weight in whale optimization algorithm (WOA) to solve these problems. In addition, more of the above research is focused on the denoising of local discharge signals and has limited involvement in PD fault diagnosis. Diagnosis of GIS PD has important engineering significance in power system. It is an innovative attempt to introduce VMD into PD fault diagnosis. Due to the experimental environment and expensive experimental equipment, the PD data obtained is limited, SVM is used to identify PD faults. The parameters (α, K) of VMD and (c, g) of SVM seriously affect the accuracy of fault diagnosis. To sum up, VMD and SVM are optimized automatically by improved whale optimization algorithm to realize diagnosis of GIS PD faults.

The remaining parts are organized as follows: Section II introduces the basic theory of VMD, MPE and WOA. A detailed introduction was given to the IWOA-VMD method and fault diagnosis process of GIS PD. Section III introduces the process of GIS PD experiments, including platform construction, simulation of typical faults, data collection, and other processes. Section IV is the analysis of the decomposition results of IWOA-VMD, including the optimized decomposition results of the PD signals, MPE calculation, screening of IMFs and PCA-based dimension reduction. Then achieve PD fault identification through IWOA-SVM. Finally, Section V concludes this paper.

II. A DESCRIPTION OF THEORETICAL BACKGROUND

A. VARIATIONAL MODE DECOMPOSITION

VMD decomposes one real signal into K independent subsignal u_k, which has specific sparsity. This procedure gets the minimum bandwidth estimation of each modal [37]. The procedure of signal decomposition is to solve the variational problem. The variational model with constraint condition is:

$$\begin{cases} \min \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^{K} f(t) \end{cases} \tag{1}$$

where $\{\mu_k\} = \{\mu_1, \mu_2, \ldots, \mu_k\}$ demonstrates the modal components, $\{\omega_k\} = \{\omega_1, \omega_2, \ldots, \omega_k\}$ is the center frequency of each modal component, $\delta(t)$ represents impulse function, $*$ is a convolutional symbol, and f is the original signal.

In order to obtain the optimal solution of such constrained variational problem, Lagrangian multiplier $\lambda(t)$ is introduced. The constrained variational problem is transformed into non-constrained problem:

$$\begin{aligned} &L(\{\mu_k\}, \{\omega_k\}, \lambda) \\ &= \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ &\quad + \left\langle \lambda(t), x(t) - \sum_{k=1}^{K} \mu_k(t) \right\rangle \end{aligned} \tag{2}$$

where α is the quadratic penalty factor. Alternate direction method of multipliers (ADMM) is introduced to obtain the saddle point of such Lagrangian function, which is the optimal solution.

The procedure of VMD can be summarized in the following steps:

1) Initialize each modal component $\{\mu_k^1\}$, center frequency $\{\omega_k^1\}$, and operators λ^1. Set $n = 0$.
2) Update $\hat{\mu}_k$ in non-negative frequency intervals:

$$\hat{\mu}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{\mu}_k(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)} \tag{3}$$

3) Update ω_k.

$$\omega_k^{n+1} = \frac{\int_0^{\infty} \omega \left| \hat{\mu}_k(\omega) \right|^2 d\omega}{\int_0^{\infty} \left| \hat{\mu}_k(\omega) \right|^2 d\omega} \tag{4}$$

4) Update λ in non-negative frequency intervals:

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left[\hat{f}(\omega) - \sum_k \hat{\mu}_k^{n+1}(\omega) \right] \tag{5}$$

For a given precision $\varepsilon > 0$, if $\frac{\sum_k \left\| \hat{\mu}_k^{n+1} - \hat{\mu}_k^n \right\|_2^2}{\left\| \hat{\mu}_k^n \right\|_2^2} < \varepsilon$, then stop iteration. Otherwise, return to 2).

B. THEORY OF MULTISCALE PERMUTATION ENTROPY (MPE)

The basic idea of MPE is to coarsen time series according to multi-scale factors s, which reflects the arrangement entropy of "average length" [38]. Set the one-dimensional time series $X = \{x_i, i = 1, 2, \ldots, N\}$ with a length of N, and coarsen it to obtain the time series:

$$\begin{cases} y_j^{(s)} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_i \\ j = 1, 2, \ldots, [N/s] \end{cases} \tag{6}$$

where s is the scale factor, $[N/s]$ is the rounding of N/s. When $s = 1$, the coarse-grained sequence is the original sequence. Reconstruct the time of $y_j^{(s)}$ to obtain:

$$\gamma_l^{(s)} = \left\{ y_l^{(s)}, y_{l+\tau}^{(s)}, \ldots, y_{l+(m-1)\tau}^{(s)} \right\} \tag{7}$$

where l is the l-th reconstructed component, $l = 1, 2, \ldots, N\text{-}(m\text{-}1)\tau$, m is the embedding dimension, τ is the delay time.

Represent reconstructed components $\gamma_l^{(s)}$ with $l_1, l_2, \ldots, l_m$, and arrange the reconstructed components in ascending order as follows:

$$y_{l+(l_1-1)\tau}^{(s)} \leq y_{l+(l_2-1)\tau}^{(s)} \leq \cdots \leq y_{l+(l_m-1)\tau}^{(s)} \tag{8}$$

where $r = 1, 2, \ldots, \mathrm{R}$ and $\mathrm{R} \leq m!$, there are $m!$ permutations in the time reconstruction sequence with embedding dimension m, the symbol sequence $S(r)$ is one of the permutations. After calculating the probability $P_r(r = 1, 2, \ldots, \mathrm{R})$ of each symbol sequence, define the arrangement entropy $Hp(m)$

of different symbol sequences in the form of information entropy as:

$$H_p(m) = -\sum_{r=1}^{R} P_r \ln P_r \tag{9}$$

when $P_r = \frac{1}{m!}$, normalize $Hp\ (m)$:

$$H_p = \frac{H_p\ (m)}{\ln\ (m!)} \tag{10}$$

where H_p is the normalized arrangement entropy value, and $0 \leq\ H_p\ \leq 1$. The smaller value of H, the more regular and smooth the time series; on the contrary, the more complex the time series, the more random it is. The use of multi-scale arrangement entropy can effectively reflect the randomness of PD signals of GIS at multiple scales.

C. INTRODUCTION TO THE WHALE OPTIMIZATION ALGORITHM (WOA)

The Whale Optimization Algorithm (WOA) is a meta heuristic optimization algorithm to simulate the hunting behavior of humpback whale populations [39]. What sets it apart is that this method uses a random search agent to simulate hunting behavior, achieving global search. This special hunting method is called bubble net predation, where humpback whales swim up in a circle or ''9'' shape and continuously release bubbles to surround their prey. This behavior can be summarized into the following three parts: surround prey, bubble net attack prey, and random search for prey.

1) SURROUND PREY

WOA assumes in advance that the hunting target or location close to the optimal search agent is the current best candidate location for the prey. The other experience updates their position relative to the best agent in real-time, with the mathematical expression as follows:

$$X(t+1) = X^*(t) - A \cdot D \tag{11}$$

$$D = \left|C \cdot X^*(t) - X(t)\right| \tag{12}$$

where t represents the current number of iterations, $X(t)$ represents the position vector of the current search agent, $X^*(t)$ represents the position vector of the current best search agent, A and C represent the coefficient matrix, and the calculation formula is:

$$\begin{cases} A = 2ar_1 - a \\ C = 2r_2 \end{cases} \tag{13}$$

where r_1 and r_2 are random vectors of [0, 1], with a linearly decreasing at [0, 2].

2) BUBBLE NET ATTACK PREY (DEVELOPMENT PHASE)

The bubble net attack behavior of humpback whales can be divided into two mechanisms: contraction encirclement and spiral position update. These two behaviors occur simultaneously, and whales randomly choose to shrink their circle or swim along a spiral path, with a 50% probability of each. During the optimization process, the whale's position is updated in real-time based on this, and the mathematical expression is as follows:

$$X(t+1) = \begin{cases} X^*(t) - A \cdot D, & p < 0.5 \\ X^*(t) + D' \cdot e^{bl} \cos(2\pi l) & p \geq 0.5 \end{cases} \tag{14}$$

$$D' = \left|X^*(t) - X(t)\right| \tag{15}$$

where D' represents the distance between the current search agent and the optimal search agent location, b determines the constant of the spiral path shape, l is a uniformly distributed random number within $[-1, 1]$, and p is a random number of [0, 1].

3) SEARCH FOR PREY (SEARCH STAGE)

Unlike the development phase, with each iteration, the remaining search agents update their location in real-time based on randomly selected search agents rather than the current best search agent. The iterative process is as follows:

$$D = |C \cdot X_{rand}(t) - X(t)| \tag{16}$$

$$X(t+1) = X_{rand}(t) - A \cdot D \tag{17}$$

where $X_{rand}(t)$ represents the location of the random search agent.

III. THE PROPOSED METHODS

A. IMPROVED WHALE OPTIMIZATION ALGORITHM (IWOA)

In traditional WOA, there is a significant imbalance between the global search ability and local development ability, resulting in poor algorithm accuracy and a possible decrease in population diversity during continuous evolution, making the algorithm easily fall into local optimization and slow convergence. To solve the above problems, this section comprehensively improved WOA though a nonlinear arctangent convergence factor and adaptive weight.

In WOA, the convergence factor a decays linearly from 2 to 0, causing the algorithm to lose balance between global search and local development capabilities throughout the entire iteration process. To enhance the global search and local development capabilities, it is desired to maintain a high value of the convergence factor during the first half of the iteration process, enhance the global search ability of the algorithm, and explore the solution space as much as possible with a larger step size. In the second half of the iteration process, it is best to perform local search with a smaller step size for the convergence factor, which can improve the solution accuracy. In this regard, this paper proposes a nonlinear arctangent convergence factor.

$$a = 2 - \frac{4}{\pi} \arctan\left(\frac{10t}{Max_{iter}}\right) \tag{18}$$

where Max_{iter} is the maximum number of iteration.

Compared with the linear iteration of the convergence factor of standard WOA, the arctangent convergence factor proposed in this paper decreases nonlinearly

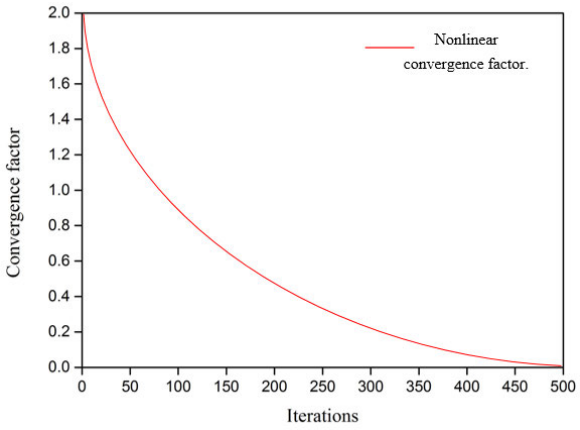


FIGURE 1. Nonlinear convergence factor change curve.

from 2 to 0 faster, reducing the imbalance between global and local search capabilities caused by the convergence factor in the original algorithm, and is more suitable for handling high-dimensional complex optimization problems.

In addition, based on the idea of weight, adaptive weight is introduced to realize location update, which further enhances the global optimization ability of the algorithm in the early stage of population evolution, and avoids the defect that the population is easy to fall into the local optimal solution to a certain extent. The mathematical model of adaptive weight is:

$$\omega = \sin(\frac{\pi t}{2Max_{iter}} + \pi) + 1 \tag{19}$$

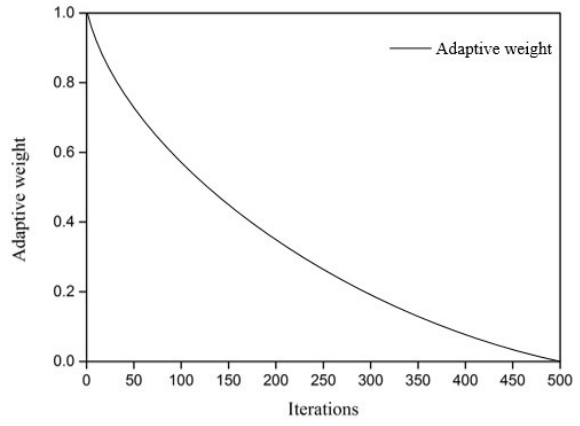


FIGURE 2. Adaptive weight change curve.

Adaptive weight maintains relatively large weight in the early stages of population evolution, which is beneficial for global exploration. As the number of iterations increases, the weight begins to decrease. Smaller weight can enhance the local expansion ability of the algorithm in the late iteration, so as to achieve the balance of global and local optimization ability, and improve the convergence speed and solution accuracy of the algorithm.. After introducing sine inertia weight and improving convergence factor, the formula (11), (14), (17) is updated to(20), (21), (22):

$$X(t+1) = \omega X^*(t) - A \bullet D \tag{20}$$

$$X(t+1) = \omega D^* \bullet e^{bl} \bullet \cos(2\pi l) + X^*(t) \tag{21}$$

$$X(t+1) = \omega X_{rand} - A \bullet D \tag{22}$$

B. PD FAULT DIAGNOSIS BASED ON IWOA-VMD

The optimization process of VMD parameters (mode parameter K and penalty factor α) based on IWOA is as follows:

Step 1: Collect PD signals.
Step 2: Set relevant parameters, including the population size and the maximum number of iteration. Initialize the position of the whale group and use a random method to set the initial position of the whale group.
Step 3: According to the objective function (α, K). Calculate the fitness value of each individual whale.
Step 4: Calculate and update the optimal position and value.
Step 5: Determine whether the algorithm has ended and skip to *Step* 9. Otherwise, proceed to *Step* 6.
Step 6: Update parameters (a, A, C, l).
Step 7: Generate a random number p between [0, 1].
Step 8: If p is not less than 0.5, update the whale position according to formula (21) and jump to *Step* 3. Otherwise jump to *Step* 9.
Step 9: If A is not less than 1, update the whale position according to formula (22). Otherwise, update the position according to formula (20) and skip to *Step* 3.
Step 10: Output optimal parameters (a, K).

C. FAULT DIAGNOSIS MODEL

In this paper, the fault diagnosis of GIS PD method combines feature extraction and pattern identification. Firstly, the original PD signals are decomposed using IWOA-VMD to obtain the IMFs. Secondly, MPE of each IMF is calculated. Then principal component analysis (PCA) [10] is introduced to select principal components of MPE as PD feature vectors. Finally, the extracted vectors are sent to IWOA-SVM to identify different PD faults. The fault diagnosis process is as follows:

Step 1: Extract different types of PD signals in the experimental environment, including point discharge, particle discharge, floating discharge and air-gap discharge.
Step 2: Though IWOA-VMD, we select the optimal parameters [α, K] of signals of different fault types, and decomposes the signals into several IMFs.
Step 3: Calculate the correlation coefficient between each IMF and the original PD signals. If it is greater than 0.2, it will be retained as a valid IMF, otherwise it will be abandoned.
Step 4: Choose the appropriate scale factor for IMFs and calculate the MPE of extracted IMFs as original PD feature vectors.

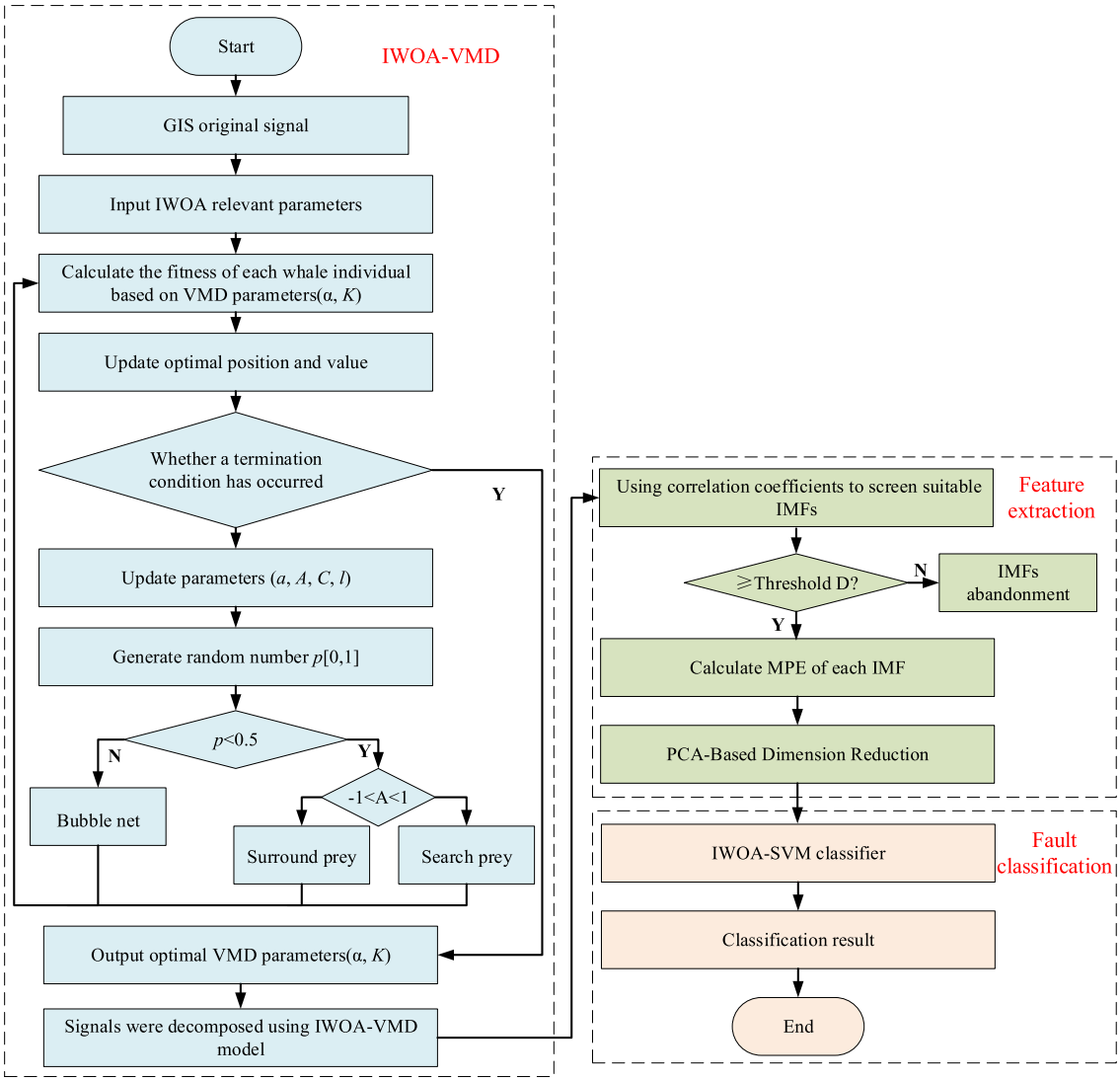


FIGURE 3. The flowchart of GIS fault diagnosis model.

Step 5: By using the PCA method, fewer representative principal components are extracted as PD feature vectors.
Step 6: Inputs the multi-dimensional fault feature vector into IWOA-SVM to classify and recognize different fault signals.

The optimization process of VMD parameters [α, K] and fault diagnosis model are shown in Fig. 3.

IV. GIS PD EXPERIMENT

A. EXPERIMENTAL SETUPS

To be more representative, four typical faults in GIS PD are selected in this paper [41]: point discharge, particle discharge, floating discharge, and air-gap discharge. The experiment sets up four defect models as shown in Fig. 4. As shown in Fig. 4(a), the point discharge simulated the protrusion of an electrode through a needle with a tip radius of 15 μm and a diameter of 1 mm, while the distance between the needle and the ground electrode is 10 mm. To simulate the free particle discharge, small spheres with a diameter of 2.0 mm are placed on a concave ground electrode, and the high-voltage electrode is secured 10mm away from the ground electrode, as represented in Fig. 4(b). To simulate floating discharge, the epoxy with a thickness of 10mm is placed, and the suspended material is directly placed between the high-voltage electrode and epoxy. The aluminum is about 4mm high, and there is a certain gap between the high-voltage electrode and the aluminum. For the air-gap discharge, there is a small gap between the epoxy and the high voltage electrode, as shown in Fig. 4(d). We use UHF sensors to collect PD signals of four different defects.

Fig. 5 shows the experimental GIS cavity and related equipment, and experimental wiring circuit is shown in Fig. 6. The power supply part of the experimental device mainly consists of a 220V AC power supply, an autotransformer,

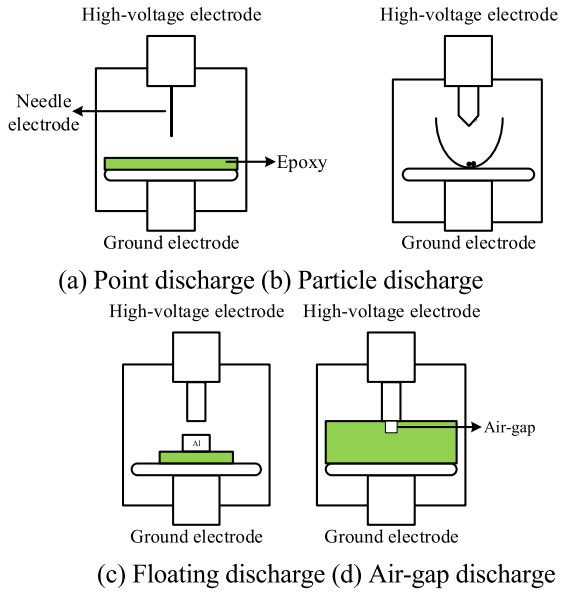


FIGURE 4. Four defect models.

and an isolation transformer. The PD test part mainly consists of protection resistance, detection impedance, and coupling capacitance. The insulation medium filled in the experimental device is SF_6 gas.

The oscilloscope used in this experiment has a maximum sampling rate of 2.5GSa/s. Adjust the test voltage, and the step size of each voltage increase is 1kV. When the GIS bus voltage starts to appear discharge pulse, it is considered that the defect starts to occur PD, and the corresponding voltage is the initial discharge voltage. After that, the bus voltage is reduced to 0, and the second test is carried out after SF6 recovers the insulation ability. This is done five times, and the median voltage of the five experiments is taken as the initial discharge voltage of the defect.

Due to the defects in GIS, the rated or higher voltage may lead to rapid breakdown of GIS. To ensure that as much as possible discharge data is collected under each defect model, the test voltage is uniformly increased to 1kV above the initial discharge voltage by the regulator, which is taken as the test voltage. Initial discharge voltage and experimental voltage under various defects are shown in the Table 1.

TABLE 1. PD experimental voltage.

Defect type	Initial discharge voltage	Testing voltage
Point discharge	15 kV	16 kV
Particle discharge	23 kV	24 kV
Floating discharge	25 kV	26 kV
Air-gap discharge	32 kV	33 kV

Set the sampling rate to 50MSa/s, 100 discharge points are selected randomly for each type of PD. 80 points are used as training sets, and 20 points are used as test sets, with each discharge point containing 500 sampling points.

B. PD SIGNALS EXTRACTION

The UHF sensors are adsorbed on the inner wall of the GIS shell, and can detect the UHF electromagnetic wave signal generated during the simulated PD in the GIS, and then transmit the signal to the oscilloscope and computer through wired mode. To verify the effectiveness of the method proposed in this article, set the sampling rate to 50MSa/s, the fault pulse signals of four PD types are collected as Fig. 7. Due to the four defects, the electric field distribution is non-uniform, the electric field intensity near the defects is large, and some areas reach the breakdown limit of SF6, resulting in PD phenomenon. PD is a kind of pulse discharge, which produces strong pulse crest when discharge occurs.

V. EXPERIMENTAL RESULTS ANALYSIS

A. DECOMPOSITION OF PD SIGNALS

Perform IWOA-VMD decomposition on PD signals, and the parameter settings of IWOA are listed in Table 2. The population size and maximum iteration number are determined after experimental comparison and verification. The optimization results of VMD parameters are shown in Table 3.

TABLE 2. Parameter settings.

Population Size	Maximum iteration number	Boundary(α, K)	
		Minimum value	Maximum value
20	50	(100，3)	(2000，8)

TABLE 3. IWOA-VMD optimization results.

Defect type	Optimized value of α	Optimized value of K
Point discharge	268.6	4
Particle discharge	231	7
Floating discharge	1557.4	3
Air-gap discharge	1615.7	7

Taking the point discharge and particle discharge as examples, the IWOA-VMD decomposition diagram and spectrum diagram are shown as Fig. 8 and Fig. 9.

As a comparison, EMD is used to decompose the point discharge signals, as shown in Figure 10. The graph shows the EMD decomposition results containing 10 IMFs components, resulting in 9 IMFs components and one residual component. EMD is the recursive modal decomposition, modal aliasing is existed which make it difficult to separate the components with similar frequency, and the endpoint effect has also appeared. Through the above comparison, the effectiveness and superiority of VMD have been verified, and VMD is more suitable for the decomposition of PD signals.

B. IMFS SELECTION

To obtain effective IMFs, calculate the correlation coefficient between each IMF and the original PD signals. Given the threshold D, if correlation coefficient is greater than D, the

(a) Overall experimental platform (b) SF6 inflation device (c) Voltage control platform

FIGURE 5. GIS cavity and related equipments.

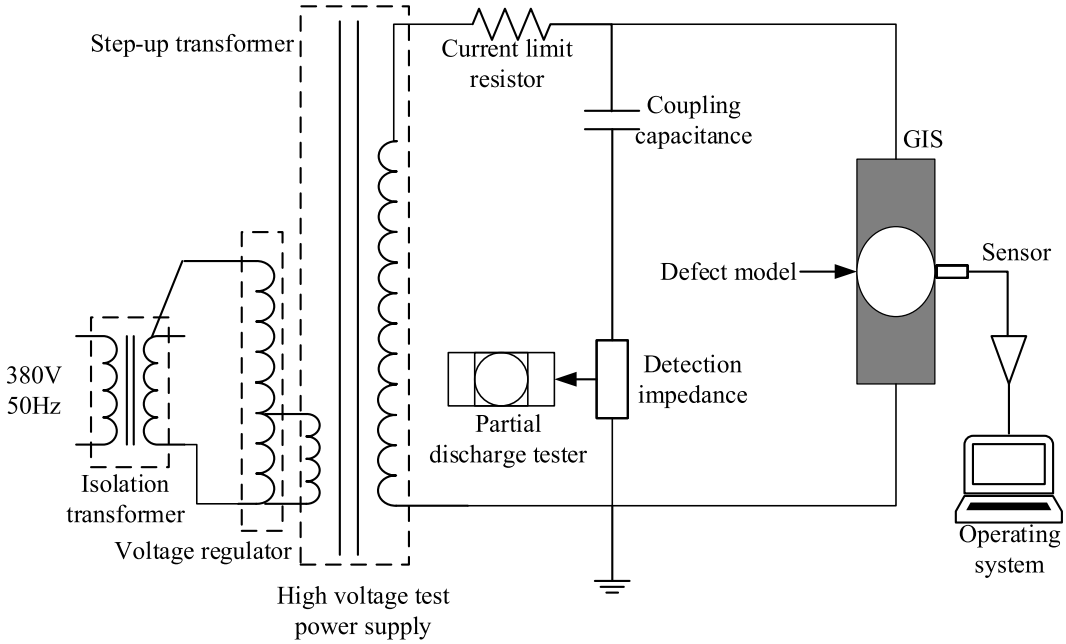


FIGURE 6. Experimental wiring circuit.

IMF is selected as the effective component, otherwise, it will be considered as a false ingredient and abandoned. In the paper, to collect as many fault signal components as possible, D is set to 0.2. The IMFs correlation coefficient values of IWOA-VMD are shown as Table 4.

TABLE 4. IMF correlation coefficient values.

Defect type	D1	D2	D3	D4	D5	D6	D7
Point discharge	0.752	0.524	0.315	0.065	-	-	-
Particle discharge	0.836	0.721	0.558	0.214	0.156	0.084	0.0045
Floating discharge	0.84	0.65	0.31	-	-	-	-
Air-gap discharge	0.89	0.765	0.51	0.31	0.089	0.012	0.0016

From Table 4, it is indicated that IMF1 to IMF3 are effective for point discharge and floating discharge. As for particle discharge and air-gap discharge, we select the IMF1 to IMF4 as components.

C. FEATURE EXTRACTION

Next, calculate the MPE of each IMF. During the MPE calculation process, some preset parameters need to be provided, including the scale factor s, embedding dimension m, and time delay δ. Generally speaking, the embedding dimension takes values of 3-7. If m is too small, the reconstructed sequence contains too few states and cannot reflect the dynamic changes of the time series, and if m is too large, it not only consumes time but also cannot reflect small changes in the time series. The impact of time delay δ on the time series can be ignored, and the value of the scale factor s is greater than 10. In order to better detect the dynamic changes of the signals, through some debugging and comparison in the early stage, m is set to 6, s is set to 18, and δ is set to 1.

Calculate the MPE values of four different PD signals extracted by the laboratory. For each type of PD, the MPE value is averaged using each effective IMF, as shown in Fig. 9.

From Fig. 11, it is illustrated that as the scaling factor changes, the MPE values of different types of PD signals also vary. The reason is that when a PD fault occurs, the

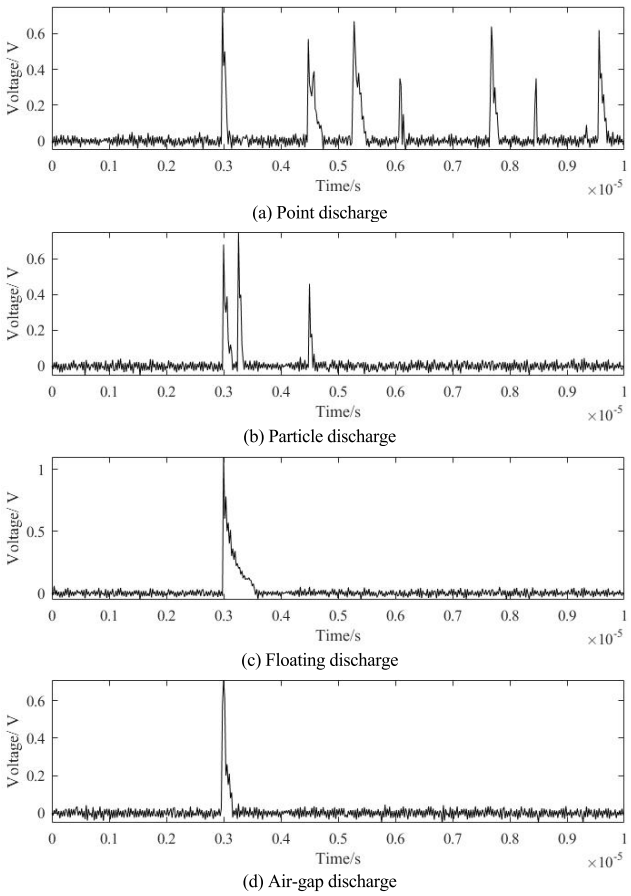


FIGURE 7. PD signals.

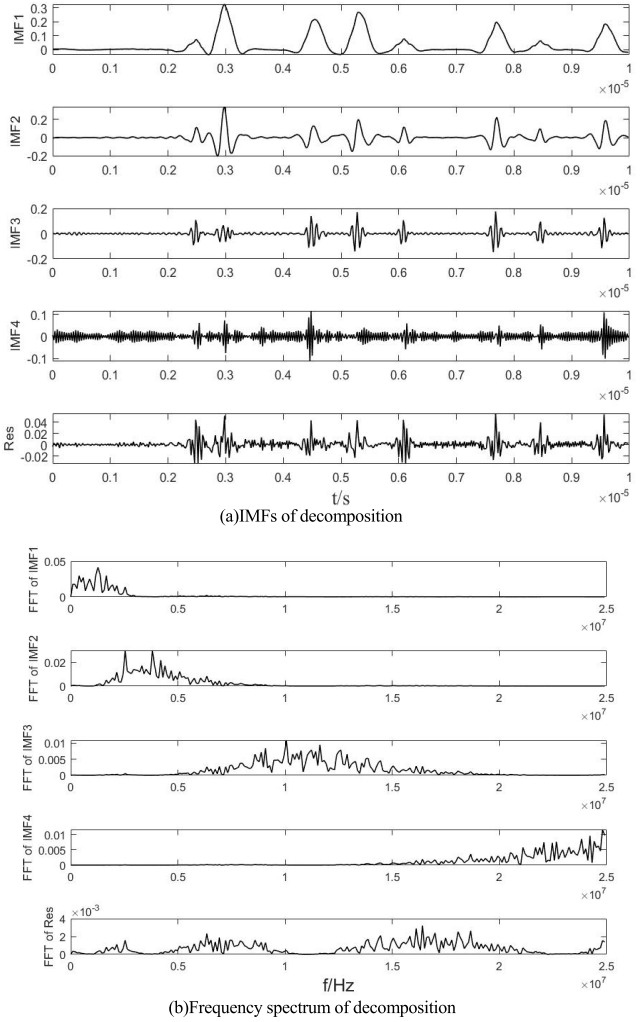


FIGURE 8. IWOA-VMD decomposition results of point discharge.

randomness of the PD signals is changing, which will change the MPE. This also indicates that a single scale cannot fully reflect all signal information, and more important information is distributed in other scales. MPE can effectively detect the dynamic changes of PD signals representing different scale fault features. In addition, the MDE values begin to stabilize after $s = 11$. Therefore, set the scale factor to 11. Taking the point discharge and particle discharge as examples, calculate the MPE of each effective IMF at a scale factor of 11.

According to Figure 12, as the scaling factor changes, the MPE values extracted from VMD of different fault signals are different obviously. Therefore, the MPE can be used as feature vectors.

D. PCA REDUCTION PROCESSING

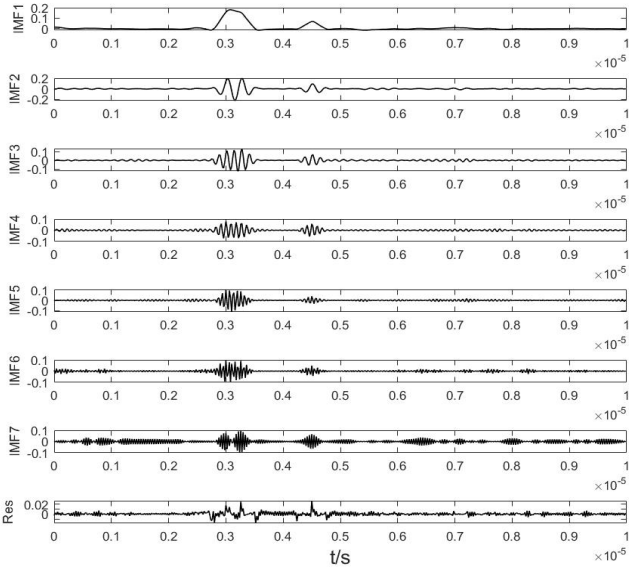
Taking the point discharge as an example, its feature vectors are shown in Table 5.

TABLE 5. Feature vectors of point discharge.

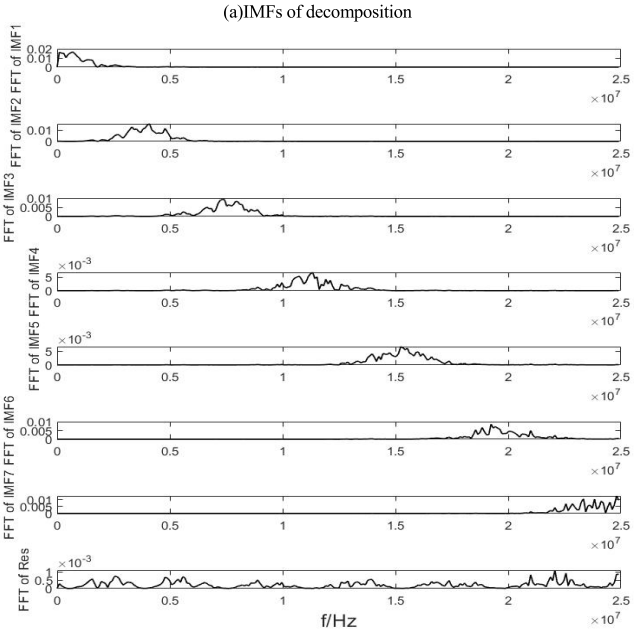
IMF	Vectors
IMF1	a1, a2, a3, …, a11
IMF2	d1, d2, d3, …, d11
IMF3	f1, f2, f3, …, f11

It is obvious that the dimension of the initially extracted feature vectors are too high, which can lead to a slow SVM classification process. The paper uses PCA to reduce the dimension of the initial feature vectors. For IMF1, construct a covariance matrix to obtain the principal components, perform a linear transformation on the original vector to obtain the eigenvalues and eigenvectors of the covariance matrix. In order to achieve dimension reduction, factors with eigenvalues greater than 1 are selected as principal components. The eigenvalues of the covariance matrix and their corresponding contribution rates are shown in Fig. 13.

The eigenvalue of feature vector $a2$ is 2.32, and the contribution rate decreases to 0.28. The eigenvalue of feature vector $a3$ is reduced to 0.852, with a contribution rate of 0.064. The subsequent values of both are reduced to negligible. The combined contribution rates of $a1$ and $a2$ reach 0.894, so they represent most of the vector information and are suitable for preservation. Therefore, through the PCA method, the original 11 feature vectors are



(a)IMFs of decomposition



(b)Frequency spectrum of decomposition

FIGURE 9. IWOA-VMD decomposition results of particle discharge.

reduced to 2, and similar methods can be used to obtain the principal components of IMF2 and IMF3, as shown in Table 6.

TABLE 6. Screening of principle component.

IMF	Contribution Rate	Principle Component
IMF1	0.894	a1, a2
IMF2	0.913	d1, d2, d3
IMF3	0.842	f1, f2

Seven principal component factors are extracted from 33 feature vectors, and the total contribution rate was close to 0.9, which can better contain the original information.

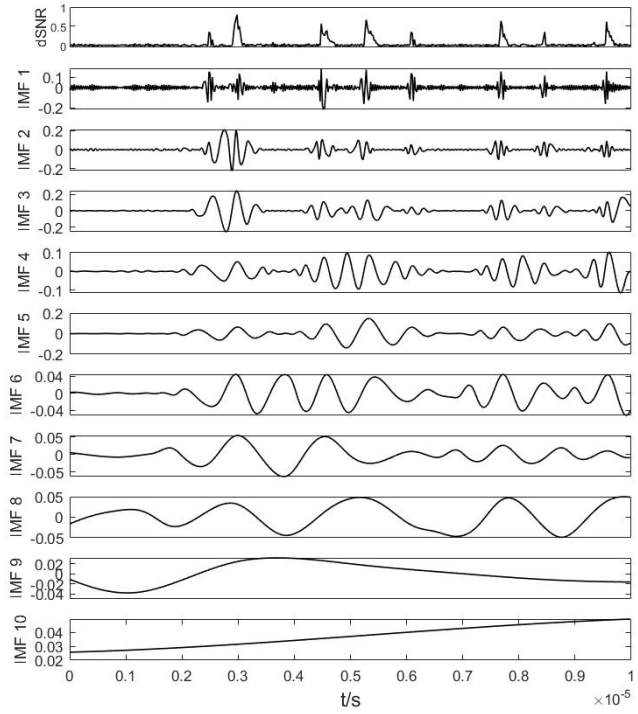


FIGURE 10. Results of EMD decomposition of point discharge.

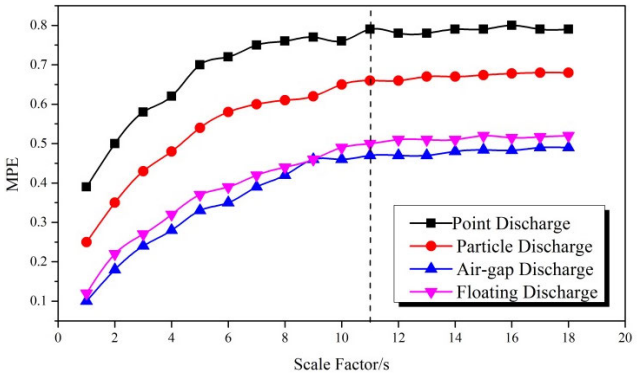


FIGURE 11. Values of MPE.

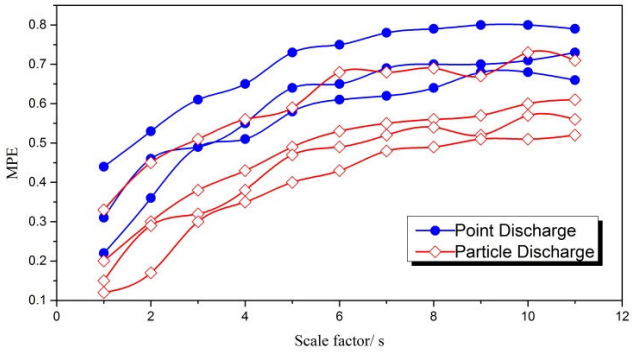


FIGURE 12. MPE of IMFs.

Similarly, through the above steps, the principal component factors for 4 PD types are calculated as shown in Table 7.

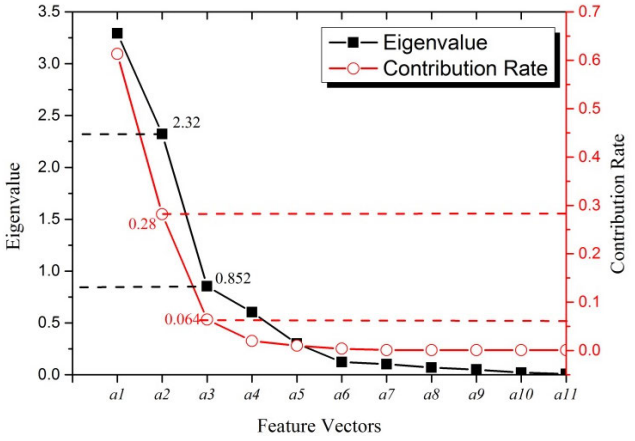


FIGURE 13. Eigenvalues and Contribution.

TABLE 7. Results of PCA dimensionality reduction.

PD type	IMF1	IMF2	IMF3	IMF4
Point discharge	a1, a2	d1, d2, d3	f1, f2	-
Particle discharge	a1, a2	d1, d2	f1, f2, f3	h1, h2
Floating discharge	a1, a2	d1, d2	f1, f2	-
Air-gap discharge	a1, a2, a3	d1, d2	f1, f2, f3	h1, h2, h3

E. GIS PD FAULT IDENTIFICATION

Calculate the MPE of the collected samples according to above steps and construct a feature vector matrix. Input samples into the SVM model, and use IWOA to optimize the parameter penalty factor c and kernel parameters g of SVM.

In this step, the population size is set to 15, the maximum number of iterations is set to 50, the range of penalty factor c is [1, 2000], and the range of kernel parameter g is [1, 200]. This paper uses accuracy as the fitness function for optimization, and takes the median accuracy of the five experiments as representative. The experimental accuracy and optimization results are shown in Table 8.

TABLE 8. Results of IWOA-SVM.

Identification Accuracy	(c, g)
98.5%	(312.481, 86.954)

To avoid the randomness of SVM, we conduct five experiments for each algorithm. The box diagram of experimental results is shown in Fig. 14. We take the median of experimental results of each group as reference and compared them in the Table 9.

From the perspective of running time, #1 is one third of #2, and #4 is much smaller than #5, which reflects the effectiveness of the improved whale optimization algorithm proposed in the paper. By comparing #1 and #4, it can be seen that MPE is used as an indicator with faster computing speed and higher recognition accuracy. EMD performs poorly

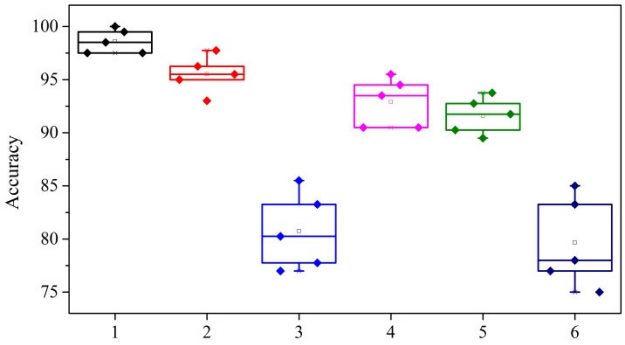


FIGURE 14. Experimental results.

TABLE 9. Identification result with different algorithm.

NO.	Algorithm types	Running Time/s	Identification Accuracy/%
#1	IWOA-VMD-MPE	1.453	98.5
#2	WOA-VMD-MPE	4.317	95.5
#3	EMD-MPE	9.724	80.25
#4	IWOA-VMD-MSE	3.426	93.5
#5	WOA-VMD-MSE	7.241	91.75
#6	EMD-MSE	16.235	78

in any category. Combining accuracy and running time, it is indicated that the method presented in the paper (#1) shows its robustness and superiority.

VI. CONCLUSION

In this paper, a novel PD fault diagnosis method is proposed. This method combines PD feature extraction based on IWOA-VMD-MPE with PD pattern identification based on IWOA-SVM. IWOA-VMD is employed for PD signal decomposition and determines the appropriate (α, K) automatically. Next, determine the effective IMFs through the correlation coefficient method and calculate its MPE. After that, PCA is introduced to select the effective principal components in MPE as the final feature vectors. Finally, IWOA-SVM is used for fault identification. The conclusions obtained are as follows:

1) Through the introduction of nonlinear arctangent convergence factor and adaptive weight, IWOA has a very good effect on parameter optimization of VMD and SVM.
2) The multi-scale calculation and dimensionality reduction of MPE can measure different fault characteristics well.
3) Compared with other PD feature extraction methods, IWOA-VMD-MPE has 98.5% accuracy and 1.453s of operation speed, showing the validity and superiority of the theory proposed in this paper.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and help to improve this manuscript.

REFERENCES

[1] U. Schichler, W. Koltunowicz, D. Gautschi, A. Girodet, H. Hama, K. Juhre, J. Lopez-Roldan, S. Okabe, S. Neuhold, C. Neumann, J. Pearson, R. Pietsch, U. Riechert, and S. Tenbohlen, "UHF partial discharge detection system for GIS: Application guide for sensitivity verification: CIGRE WG D1.25," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 3, pp. 1313–1321, Jun. 2016.

[2] C. Xia, M. Ren, R. Chen, J. Yu, C. Li, Y. Chen, K. Wang, S. Wang, and M. Dong, "Multispectral optical partial discharge detection, recognition, and assessment," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022, doi: 10.1109/TIM.2022.3162284.

[3] Z. Shu, W. Wang, C. Yang, Y. Guo, J. Ji, Y. Yang, T. Shi, Z. Zhao, and Y. Zheng, "External partial discharge detection of gas-insulated switchgears using a low-noise and enhanced-sensitivity UHF sensor module," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–10, 2023, doi: 10.1109/TIM.2023.3277980.

[4] K. Yin, Y. Wang, S. Liu, P. Li, Y. Xue, B. Li, and K. Dai, "GIS partial discharge pattern recognition based on multi-feature information fusion of PRPD image," *Symmetry*, vol. 14, no. 11, p. 2464, Nov. 2022, doi: 10.3390/SYM14112464.

[5] S. M. Song, Y. Qian, H. Wang, Y. M. Zang, G. H. Sheng, and X. C. Jiang, "Partial discharge pattern recognition based on 3D graphs of phase resolved pulse sequence," *Energies*, vol. 13, no. 16, p. 4103, Aug. 2020, doi: 10.3390/EN13164103.

[6] Z. Li, Y. Qian, H. Wang, X. Zhou, G. Sheng, and X. Jiang, "Partial discharge fault diagnosis based on Zernike moment and improved bacterial foraging optimization algorithm," *Electric Power Syst. Res.*, vol. 207, Jun. 2022, Art. no. 107854, doi: 10.1016/J.EPSR.2022.107854.

[7] F. Lv, G. Liu, Q. Wang, X. Lu, S. Lei, S. Wang, and K. Ma, "Pattern recognition of partial discharge in power transformer based on InfoGAN and CNN," *J. Electr. Eng. Technol.*, vol. 18, no. 2, pp. 829–841, Mar. 2023, doi: 10.1007/S42835-022-01260-7.

[8] Q. Jing, J. Yan, L. Lu, Y. Xu, and F. Yang, "A novel method for pattern recognition of GIS partial discharge via multi-information ensemble learning," *Entropy*, vol. 24, no. 7, p. 954, Jul. 2022, doi: 10.3390/E24070954.

[9] Q. Jing, J. Yan, Y. Wang, R. He, and L. Lu, "A novel differentiable neural network architecture automatic search method for GIS partial discharge pattern recognition," *Measurement*, vol. 195, May 2022, Art. no. 111154, doi: 10.1016/J.MEASUREMENT.2022.111154.

[10] Y. Wang, J. Yan, X. Ye, Z. Qi, J. Wang, and Y. Geng, "GIS partial discharge pattern recognition via a novel capsule deep graph convolutional network," *IET Gener., Transmiss. Distrib.*, vol. 16, no. 14, pp. 2903–2912, Jul. 2022, doi: 10.1049/GTD2.12508.

[11] Y. Wang, J. Yan, Z. Yang, J. Wang, and Y. Geng, "A novel 1DCNN and domain adversarial transfer strategy for small sample GIS partial discharge pattern recognition," *Meas. Sci. Technol.*, vol. 32, no. 12, Dec. 2021, Art. no. 125118, doi: 10.1088/1361-6501/AC27E8.

[12] T. Liu, J. Yan, Y. Wang, Y. Xu, and Y. Zhao, "GIS partial discharge pattern recognition based on a novel convolutional neural networks and long short-term memory," *Entropy*, vol. 23, no. 6, p. 774, Jun. 2021, doi: 10.3390/E23060774.

[13] X. Zhou, X. Wu, P. Ding, X. Li, N. He, G. Zhang, and X. Zhang, "Research on transformer partial discharge UHF pattern recognition based on CNN-LSTM," *Energies*, vol. 13, no. 1, p. 61, Dec. 2019, doi: 10.3390/EN13010061.

[14] X. Peng, F. Yang, G. Wang, Y. Wu, L. Li, Z. Li, A. A. Bhatti, C. Zhou, D. M. Hepburn, A. J. Reid, M. D. Judd, and W. H. Siew, "A convolutional neural network-based deep learning methodology for recognition of partial discharge patterns from high-voltage cables," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1460–1469, Aug. 2019, doi: 10.1109/TPWRD.2019.2906086.

[15] C. Zhang, Y. He, T. Yang, B. Zhang, and J. Wu, "An analog circuit fault diagnosis approach based on improved wavelet transform and MKELM," *Circuits, Syst., Signal Process.*, vol. 41, no. 3, pp. 1255–1286, Mar. 2022, doi: 10.1007/S00034-021-01842-2.

[16] W. He, Y. He, B. Li, and C. Zhang, "Feature extraction of analogue circuit fault signals via cross-wavelet transform and variational Bayesian matrix factorisation," *IET Sci., Meas. Technol.*, vol. 13, no. 2, pp. 318–327, Mar. 2019, doi: 10.1049/IET-SMT.2018.5432.

[17] C. Zhang, S. Zhao, Z. Yang, and Y. He, "A multi-fault diagnosis method for lithium-ion battery pack using curvilinear Manhattan distance evaluation and voltage difference analysis," *J. Energy Storage*, vol. 67, Sep. 2023, Art. no. 107575, doi: 10.1016/J.EST.2023.107575.

[18] C. Zhang, Y. He, L. Zuo, J. Wang, and W. He, "A novel approach to diagnosis of analog circuit incipient faults based on KECA and OAO LSSVM," *Metrol. Meas. Syst.*, vol. 22, no. 2, pp. 251–262, Jun. 2015, doi: 10.1515/MMS-2015-0025.

[19] Y. Li, M. Xu, X. Liang, and W. Huang, "Application of bandwidth EMD and adaptive multiscale morphology analysis for incipient fault diagnosis of rolling bearings," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6506–6517, Aug. 2017, doi: 10.1109/TIE.2017.2650873.

[20] A. Bustos, H. Rubio, C. Castejón, and J. García-Prada, "EMD-based methodology for the identification of a high-speed train running in a gear operating state," *Sensors*, vol. 18, no. 3, p. 793, Mar. 2018, doi: 10.3390/s18030793.

[21] Y. Zhou, J. Li, H. Yan, and X. Yan, "Low-frequency ultrasound thoracic signal processing based on music algorithm and EMD wavelet thresholding," *IEEE Access*, vol. 11, pp. 73912–73921, 2023, doi: 10.1109/ACCESS.2023.3296465.

[22] J. S. Smith, "The local mean decomposition and its application to EEG perception data," *J. Roy. Soc. Interface*, vol. 2, no. 5, pp. 443–454, Dec. 2005, doi: 10.1098/RSIF.2005.0058.

[23] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009, doi: 10.1142/S1793536909000047.

[24] Y. Yang, J. Cheng, and K. Zhang, "An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems," *Measurement*, vol. 45, no. 3, pp. 561–570, Apr. 2012, doi: 10.1016/J.MEASUREMENT.2011.10.010.

[25] J. Zheng, M. Su, W. Ying, J. Tong, and Z. Pan, "Improved uniform phase empirical mode decomposition and its application in machinery fault diagnosis," *Measurement*, vol. 179, Jul. 2021, Art. no. 109425, doi: 10.1016/J.MEASUREMENT.2021.109425.

[26] Y. Imaouchen, M. Kedadouche, R. Alkama, and M. Thomas, "A frequency-weighted energy operator and complementary ensemble empirical mode decomposition for bearing fault detection," *Mech. Syst. Signal Process.*, vol. 82, pp. 103–116, Jan. 2017, doi: 10.1016/J.YMSSP.2016.05.009.

[27] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 1, 2014, doi: 10.1109/TSP.2013.2288675.

[28] J. Yao, Y. Xiang, S. Qian, S. Wang, and S. Wu, "Noise source identification of diesel engine based on variational mode decomposition and robust independent component analysis," *Appl. Acoust.*, vol. 116, pp. 184–194, Jan. 2017, doi: 10.1016/J.APACOUST.2016.09.026.

[29] C. Li, Y. Wu, H. Lin, J. Li, F. Zhang, and Y. Yang, "ECG denoising method based on an improved VMD algorithm," *IEEE Sensors J.*, vol. 22, no. 23, pp. 22725–22733, Dec. 2022, doi: 10.1109/JSEN.2022.3214239.

[30] C. Wang, H. Li, G. Huang, and J. Ou, "Early fault diagnosis for planetary gearbox based on adaptive parameter optimized VMD and singular kurtosis difference spectrum," *IEEE Access*, vol. 7, pp. 31501–31516, 2019, doi: 10.1109/ACCESS.2019.2903204.

[31] N. K. Sharma, S. R. Samantaray, and C. N. Bhende, "VMD-enabled current-based fast fault detection scheme for DC microgrid," *IEEE Syst. J.*, vol. 16, no. 1, pp. 933–944, Mar. 2022, doi: 10.1109/JSYST.2021.3057334.

[32] S. Jiao, W. Shi, and Y. Yang, "Denoising of UHF partial discharge signals by using VMD based on Shannon entropy and kurtosis-approximation entropy," in *Proc. 14th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2019, pp. 1742–1747, doi: 10.1109/ICIEA.2019.8834081.

[33] H. Shang, F. Li, and Y. Wu, "Partial discharge fault diagnosis based on multi-scale dispersion entropy and a hypersphere multiclass support vector machine," *Entropy*, vol. 21, no. 1, p. 81, Jan. 2019, doi: 10.3390/e21010081.

[34] Z. Lei, F. Wang, and C. Li, "A denoising method of partial discharge signal based on improved SVD-VMD," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 30, no. 5, p. 1, Oct. 2023, doi: 10.1109/TDEI.2023.3269725.

[35] X. Chen, X. Shao, X. Pan, G. Luo, M. Bi, T. Jiang, and K. Wei, "Feature extraction of partial discharge in low-temperature composite insulation based on VMD-MSE-IF," *CAAI Trans. Intell. Technol.*, vol. 7, no. 2, pp. 301–312, Jun. 2022, doi: 10.1049/CIT2.12087.

[36] R. Dhandapani, I. Mitiche, V. S. Mallela, and G. Morison, "Partial discharge signal denoising based on VMD and group-sparse TV denoising," in *Proc. 2nd Int. Conf. Signal Process. Inf. Secur. (ICSPIS)*, Dubai, United Arab Emirates, Oct. 2019, pp. 1–4, doi: 10.1109/ICSPIS48135.2019.9045901.

[37] Z. Wang, L. Jia, and Y. Qin, "Adaptive diagnosis for rotating machineries using information geometrical kernel-ELM based on VMD-SVD," *Entropy*, vol. 20, no. 1, p. 73, Jan. 2018, doi: 10.3390/E20010073.
[38] M. Jabloun, P. Ravier, and O. Buttelli, "Improving of the interpretation of linear filtering preprocessing-based multiscale permutation entropy," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Hanoi, Vietnam, Jul. 2023, pp. 190–194, doi: 10.1109/SSP53291.2023.10208030.
[39] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016, doi: 10.1016/J.ADVENGSOFT.2016.01.008.
[40] X. Deng, X. Tian, S. Chen, and C. J. Harris, "Deep principal component analysis based on layerwise feature extraction and its application to nonlinear process monitoring," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 6, pp. 2526–2540, Nov. 2019, doi: 10.1109/TCST.2018.2865413.
[41] I. A. Metwally, "Status review on partial discharge measurement techniques in gas-insulated switchgear/lines," *Electr. Power Syst. Res.*, vol. 69, no. 1, pp. 25–36, Apr. 2004, doi: 10.1016/J.EPSR.2003.07.006.

WEI SUN was born in Huanggang, Hubei, China, in 1996. He received the B.Eng. and M.Sc. degrees from the College of Electrical Engineering and New Energy, China Three Gorges University, in 2018 and 2021, respectively. He is currently pursuing the Ph.D. degree with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. His current research interests include high-voltage insulation technology, the fault diagnosis of high voltage power equipment, and the online monitoring of power equipment.

HONGZHONG MA received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering, Southeast University, Nanjing, China, in 1989, 1994, and 2002, respectively.

He is currently a Professor in electrical engineering with Hohai University, Nanjing. His current research interests include high voltage insulation technology, condition monitoring and the fault diagnosis of electrical equipment, and electrical machines and drives.

SIHAN WANG was born in Lianyungang, Jiangsu, China, in 1998. He received the B.Eng. degree from the College of Electrical Engineering, Northeast Electric Power University, in 2021. He is currently pursuing the Graduate degree with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. His current research interests include high-voltage insulation technology and the fault diagnosis of high voltage power equipment.