Article citation info:
Chen Z, Li D, Su H, Liu W. Fault diagnosis of rolling bearings in water pumps based on data dimensionality reduction and TWSVM. Diagnostyka. 2025;26(2):2025205. https://doi.org/10.29354/diag/203629.

DIAGNOSTYKA, 2025, Vol. 26, No. 2

e-ISSN 2449-5220
DOI: 10.29354/diag/203629

FAULT DIAGNOSIS OF ROLLING BEARINGS IN WATER PUMPS BASED ON DATA DIMENSIONALITY REDUCTION AND TWSVM

Zimeng CHEN * **, Diliang LI , Hua SU , Wei LIU**

Research Institute of Nuclear Power Operation, Wuhan, 430223, China
* Corresponding author, e-mail: chenzimeng2018@126.com

Abstracts

An algorithm combining particle swarm algorithm and twin support vector machines is proposed. Secondly, in an attempt to suppress the vibration signal noise and enhance the signal features, a rolling bearing signal feature extraction model based on improved variational mode decomposition is proposed. Then, in an attempt to reduce the data dimensionality to improve the computational speed, the study introduces the kernel principal component analysis to reduce the dimensionality of the data, and at the same time, the excavator distance algorithm is introduced so as to construct a condition monitoring model. Finally, the proposed algorithm is applied to the monitoring model to construct a rolling bearing fault diagnosis model based on the algorithm and data dimensionality reduction. The suggested approach outperformed the comparison algorithm in terms of average accuracy rate and loss value, with 97.2% and 1.39, respectively, according to a comparative performance analysis. The bearing defect diagnostic model underwent performance comparison study as well, and outcomes confirmed that the model's average diagnosis accuracy rate was 94.7%. The identification accuracy of inner ring pitting fault, outer ring pitting fault, outer ring fault and rolling element pitting fault are 96.2%, 94.7%, 94.5% and 84.9%, respectively, much higher than the comparison model.

Keywords: data dimensionality reduction; TWSVM; nuclear power plant; circulating water pump; rolling bearings

List of Symbols/Acronyms

FD: fault diagnosis;
CWP: Circulating Water pump;
RB: rolling bearings;
TWSVM: twin support vector machines;
SVM: support vector machine;
PSO: Particle swarm optimization;
KPCA: kernel principal component analysis;
HD: high-dimensional;
PCA: principal component analysis;
VMD: variational mode decomposition;
MPE: multi scale permutation entropy;
WHNR: weighted harmonic to noise ratio;
VS: vibration signals;
NPP: nuclear circulating pump;
KKT: Karush-Kuhn-Tucker;
FV: fitness value;
EV: extreme value;
IMF: intrinsic mode function;
FDF: frequency-domain features;
TDF: time-domain features.

1. INTRODUCTION

Circulating Water pump (CWP) is a key component of the operation of mechanical equipment, and the rolling bearing, as the core part of the water pump, is particularly important in ensuring the smooth operation of the CWP (1). Rolling bearings (RB) are prone to failure, and if the fault is not diagnosed in time, it may cause major safety accidents (2). The fault sources of RB usually come from four aspects. First, long-term operation causes cracks caused by material fatigue and failure under the action of contact stress on the working surface of bearings. Second, the relative sliding friction of bearings makes the metal constantly wear, especially when hard particles are added and lubrication conditions deteriorate, resulting in surface deformation and parts damage. Third, when the work is overloaded, the parts break and the materials reach the limit. Fourth, long-term contact with corrodes, such as current corrosion, expands the corrosion area, resulting in failure. Choosing the right method can effectively improve the accuracy rate of rolling bearings, timely maintenance of faults, so as to avoid safety accidents. However, Because the working environment of circulating pump is very bad, and the vibration signal has the characteristics of high frequency and non-stationarity, many methods cannot effectively process and extract the signal. At present, the fault diagnosis (FD) methods of water pump rolling bearing often have some problems such as inaccurate diagnosis. Although many scholars have done research in related fields,

Received 2024-09-25; Accepted 2025-04-01; Available online 2025-04-03

the effect is still not good (3). Although many scholars have done research in related fields, the results are still poor (4). For example, For example, Cheng et al. proposed a bearing residual service life model based on dynamic interactive hybrid model in order to solve the problem that it is difficult to accurately predict the residual service life of the shaft journal bearing of the nuclear circulating pump (NPP). Due to the inability to extract fault characteristics well, the prediction effect is poor (5). In addition, for the problem of pump fault diagnosis, Dutta et al. used the literature review method to study relevant literatures, but did not propose how to enhance fault characteristics to improve the diagnostic accuracy (6). Meanwhile, slightly effective studies also exist. Pu et al. presented a FD method based on depth separable convolutional neural network in an attempt to enhance the accuracy of FD of RB of water pumps. The method was experimentally validated and the results showed effectiveness (7).

Because the circulating water pump is in a high load working state for a long time, its rolling bearing is prone to failure, which affects the normal work of the circulating water pump. How to accurately diagnose the rolling bearing fault has become an urgent problem to be solved. How to accurately diagnose the faults of RB has become an urgent problem. By resolving two related SVM-type difficulties, twin support vector machines (TWSVM), a binary support vector machine (SVM) classifier, identify two non-parallel planes (8). It has advantages such as better handling of multi-classification problems and is widely used in fields such as FD. However, TWSVM has the disadvantage that the parameters are difficult to specify. Particle swarm optimization (PSO) is a population-based stochastic optimization technique, which seeks the optimal solution of a problem by simulating the behavior of a group such as a flock of birds, and has the advantages of finding the global optimal objective, etc., and is able to make up for the defects of the TWSVM algorithm better (9). In addition, because the circulating water pump is in a complex environment and the data scale is large, it is prone to digital disaster. The fundamental concept of the kernel principal component analysis (KPCA) nonlinear dimensionality reduction technique is to use nonlinear transformations to map the sample data from the input space to the high-dimensional (HD) feature space, and then principal component analysis (PCA) to extract features (10). Many scholars have done related research. For example, Huang et al. suggested an automatic TSVM parameter selection technique based on chaotic mapping dung beetle optimization algorithm to address the issue of challenging TWSVM parameter selection. After experimental validation, it turned out that the method had high precision (11).

The above research shows that the current method used in the fault diagnosis (FD) of water pump rolling bearing is too low in FD due to the lack of effective feature extraction. In this study, the variational mode decomposition is improved and combined with multi-scale arrangement entropy, so as to enhance the extraction of fault vibration signal features of rolling bearing, so as to improve the diagnostic accuracy. Therefore, the study combines PSO and TWSVM to construct PSO-TWSVM algorithms. In addition, in order to better extract vibration signal features, a rolling bearing vibration signal feature extraction model based on improved variational mode decomposition (VMD) is constructed. Then, in order to improve the calculation speed of the algorithm and monitor the running state of the rolling bearing of the circulating water pump, the Earth Mover's Distance (EMD) algorithm is combined with KPCA algorithm and applied to the signal feature extraction model, so as to build the state monitoring model based on KPCA-improved VMD-EMD. Finally, the PSO-TWSVM algorithm is applied to the condition monitoring model to construct the FD model of circulating water pump rolling bearing based on data reduction and TWSVM. The innovation of this study is the combination of PSO algorithm and TWSVM, aiming to improve the FD accuracy of circulating water pump rolling bearing. It is expected that this method can contribute to enriching bearing FD theory.

2. METHODS AND MATERIALS

2.1. Fusion of PSO and TWSVM algorithm construction

One of the essential parts of CWP in NPP is RB, which is crucial to the proper operation of CWP in NPP. It is essential to the CWP in NPP operating safely. It is crucial to develop a technique for precisely identifying RB issues in order to guarantee the NPP's CWP operates safely TWSVM, developed from SVM, optimizes the classification effect by constructing two parallel decision boundaries to improve the prediction efficiency and accuracy (12). It has advantages such as better handling of multi-classification problems, and is widely used in fields such as FD (13). TWSVM is categorized into linear and nonlinear, and an example of linear TWSVM is shown in Figure 1 (14).

In Figure 1, TWSVM has two classification hyperplanes, whose two hyperplanes are not parallel to each other. Moreover, one of the hyperplanes is very close to the same class of training points, but very far from the other class of training points. The operation steps of linear TWSVM are as follows. First, assume that there is a training sample set $\{(x_j^i, y_j), i = 1, 2; j = 1, 2, \ldots, m\}$, $x_j^i \in R$. Among them, x_j^i is the j samples in sample i . i is the sample type. j is the samples. y_j is the sample corresponding label and $y_j \in \{+1, -1\}$. m is the total samples and $m = m_1 + m_2$. m_1 is the total sample type 1. m_2 is the total sample type 2.

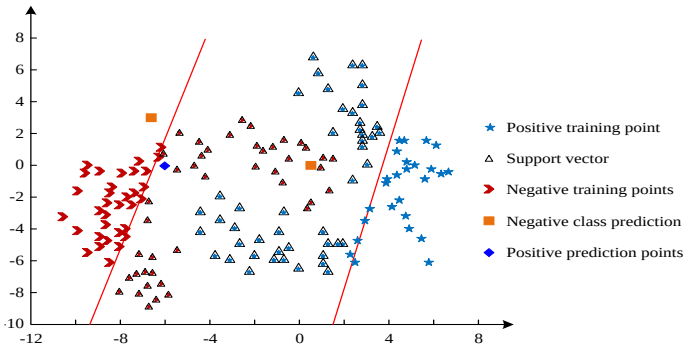


Fig. 1. Example diagram of a linear programmable model of TWSVM

Moreover, the set of sample type 1 and the set of sample type 2 are displayed by matrix $A \in R^{m_1 \times n}$ and matrix $B \in R^{m_2 \times n}$, respectively. Secondly, TWSVM is viewed as a quadratic programming problem and two hyperplanes TWSVM1 and TWSVM2 are constructed simultaneously. The computational expression is shown in Equation (1).

$$\begin{aligned} &\text{TWSVM1}: \min_{\omega^1,b^1,\xi} \frac{1}{2}(A\omega^1 + e_1 b^1)^T (A\omega^1 + e_1 b^1) + c_1 e_2^T \varsigma \\ &s.t. -(B\omega^1 + e_2 b^1) + \xi \geq e_2, \xi \geq 0 \\ &\text{TWSVM2}: \min_{\omega^2,b^2,\eta} \frac{1}{2}(B\omega^2 + e_2 b^2)^T (B\omega^2 + e_2 b^2) + c_2 e_1^T \eta \\ &s.t. -(A\omega^2 + e_1 b^2) + \eta \geq e_1, \eta \geq 0 \end{aligned} \tag{1}$$

In Equation (1), ξ and η are slack variables. e_1 and e_2 are unit vectors. b_1 and b_2 are biases. ω^1 and ω^2 are weight vectors. Then, the optimization problem for the two inequalities is solved by Lagrangian method. Moreover, the pairwise problems of TWSVM1 and TWSVM2 are obtained by combining the Karush-Kuhn-Tucker conditions (KKT). Among them, the TWSVM1 formula is shown in Equation (2).

$$\begin{aligned} &\min_{\alpha} \quad \frac{1}{2}\alpha^T G(H^T H)^{-1} G^T \alpha - e_2^T \alpha \\ &s.t. \quad 0 \leq \alpha \leq c_1 \end{aligned} \tag{2}$$

In Equation (2), α is the Lagrange multiplier, H and G are matrices and $H = [A^T \ \ e_1^T]$, $G = [B^T \ \ e_2^T]$. The TWSVM2 formula is shown in Equation (3).

$$\begin{aligned} &\min_{\alpha} \quad \frac{1}{2}\gamma^T P(Q^T Q)^{-1} P^T \gamma - e_1^T \gamma \\ &s.t. \quad 0 \leq \gamma \leq c_2 \end{aligned} \tag{3}$$

In Equation (3), γ is the Lagrange multiplier. P and Q are matrices and $P = [A \ \ e_1]$, $Q = [B \ \ e_2]$. Finally, the final discriminant formula $Label(x)$ computational expression is obtained as shown in Equation (4).

$$\begin{aligned} Label(x) &= \underset{i=1,2}{argmin}\{d_i\} = \underset{i=1,2}{argmin}\left|x^T \omega^i + b^i\right| \\ &= \begin{cases} d_1 \Rightarrow x \in class \ \ 1 \\ d_2 \Rightarrow x \in class \ \ 2 \end{cases} \end{aligned} \tag{4}$$

In Equation (4), d_1 and d_2 are dimensions. $class$ is the type. However, TWSVM has the defect of parameter uncertainty. PSO algorithm has advantages such as finding the global optimal objective, which can better compensate for the defects of TWSVM algorithm (15). The flow of PSO algorithm is shown in Figure 2 (16).

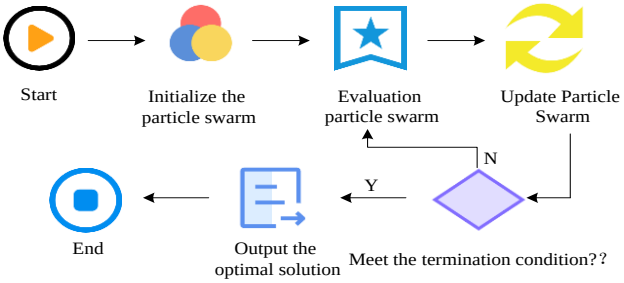


Fig. 2. Flow chart of PSO algorithm

The PSO method in Figure 2 initializes each and every particle in the population first. Second, the fitness value (FV) of every particle is computed for every generation of evolution. Subsequently, the FV of every particle is contrasted with its unique extreme value (EV) of the optimal place in history. If it is better than the past value, then it is taken as the individual EV of the current position. Each particle's FV is concurrently compared to its previous global EV. The global EV's index number is reset if it is superior. The particles' position and velocity are then updated to see if the iteration has reached the maximum of generations or if the FV meets the predetermined criteria. Provide the best outcome if the condition is met. Instead, go back to the second step. Among them, after each iteration, the particle's velocity update formula is shown in Equation (5).

$$v_{id} = wv_{id} + q_1.rand().(p_{id} - \beta_{id}) + q_2.Rand().(p_{gd} - \beta_{id}) \tag{5}$$

In Equation (5), v_{id} is the particle velocity. q_1 and q_2 are the weights. p_{id} and p_{gd} are the d th dimension components. w is the parameter. $rand()$ and $Rand()$ are random numbers. The particle position update formula is shown in Equation (6).

$$\beta_{id} = \beta_{id} + v_{id} \tag{6}$$

Based on the above, PSO is combined with TWSVM to construct PSO-TWSVM algorithm. The flow of this algorithm is shown in Figure 3.

Figure 3 shows the specific operation steps of PSO-TWSVM algorithm. First, the particle swarm's size and maximum iterations are determined. The particles' speed and direction are also initialized simultaneously. Secondly, the particles are substituted into the TWSVM and the training set data are classified and the accuracy of the classification is used as the fitness. The particles are evaluated using the fitness to obtain the initial fitness of each particle. Then iterative optimization search is performed, utilizing the position update formula and velocity update formula for the velocity and direction of the particles. Followed by the number of iterations are updated, while the adaptation degree of the corresponding particle is calculated. If the updated fitness is superior, the value is updated. The value of the global optimal fitness is updated if the particle's optimal fitness exceeds the current value. Then, the current number of iterations is judged and stopped if the maximum iteration is reached. Instead, continue the iteration. Finally, the optimal value obtained is extracted and the operation is terminated after obtaining the optimal parameters.

2.2. Design of bearing FD model based on data dimensionality reduction and TWSVM

The intricate operating environment of CWP in NPP results in a high level of noise in the VS, making it challenging to identify the characteristics of the RB fault. To increase the accuracy of FD of CWP in NPP, it is crucial to figure out how to process the signal and add features. Before extracting the features, the signal needs to be processed, so the first study introduces the VMD model. Moreover, to address the shortcomings of the VMD model such as low identification accuracy of intrinsic mode function (IMF), penalty factor and multilevel decomposition are introduced for decomposition. In this way, the optimal IMF is obtained to denoise the signal and enhance the features. Secondly, in an effort to control the number of decomposition layers, safeguard the integrity of fault features, eliminate redundant modes, and control the IMF bandwidth as well as to prevent the problem of insufficient or insufficient noise reduction, the study introduces the energy ratio index, 1.5-dimensional Weighted Harmonic to Noise Ratio (WHNR index), and constructs a method based on the improved VMD VS decomposition. Figure 4 displays the method's flow chart.

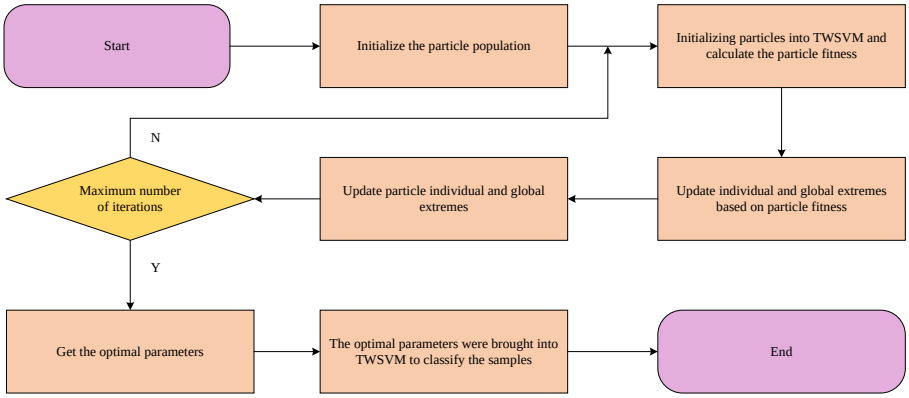


Fig. 3. Flow chart of PSO-TWSVM algorithm

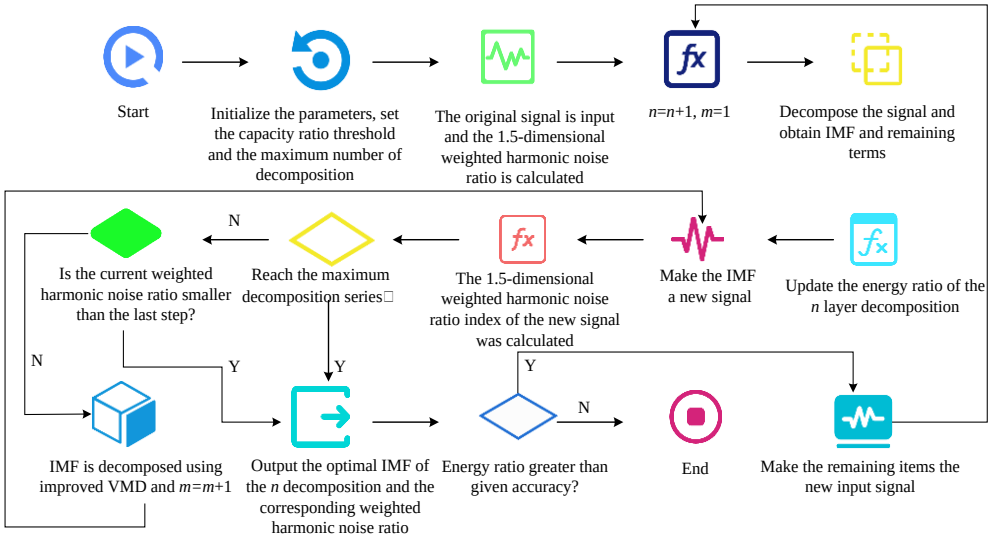


Fig. 4. The improved VMD vibration signal decomposition method

Figure 4 shows the specific operation flow of the method. In the first step, each parameter is initialized to operate, and the energy ratio threshold and the maximum number of graded levels are set. Meanwhile, the original VS is input and then its 1.5 dimensional energy spectrum analysis WHNR index is calculated, so that the times n is added to 1 and the decomposition levels m is equal to 1. In the second step, the signal is decomposed by using the VMD to obtain the IMF and the residual term. It also updates and calculates the energy ratio threshold for the n level of decomposition, while making the IMF for the new signal to calculate the 1.5 dimensional energy spectrum analysis WHNR for its decomposition level m. In the third step, if the decomposition level reaches the maximum decomposition level, the optimal IMF for the decomposition of the n level as well as the corresponding 1.5-dimensional energy spectrum analysis WHNR are outputted. Determine whether its threshold is greater than the given accuracy, if yes then the remaining terms are used as new input signals and return to the first step to recalculate. Otherwise output the result and the decomposition is finished. In the fourth step, if the decomposition level has not reached the maximum number of decomposition levels then determine whether the current level 1.5 dimensional energy spectrum analysis WHNR is smaller than the previous level. Proceed to the following stage if the answer is affirmative. Rather, it goes back to step two. In the first step, the energy ratio needs to be calculated, the energy ratio calculation formula is shown in Equation (7).

$$R_E = \frac{\dfrac{\sum_{f=CF-\tau}^{CF} \left|E_{imf}(f)\right|^2}{2\tau}}{\dfrac{\sum_{f=1}^{f_s} \left|E_{sig}(f)\right|^2}{f_s}} \tag{7}$$

In Equation (7), R_E is the energy ratio. CF is the center frequency. $E_{sig}(f)$ and $E_{imf}(f)$ are the spectral amplitude of the original signal and the frequency in the IMF, respectively. f is the signal sampling frequency. τ is the calculation range. 1.5-dimensional energy spectrum analysis $WHNR$ calculation formula is shown in Equation (8).

$$WHNR(f_o) = \sum_{f_o=f_o+T}^{f_o+T} 10log \frac{r_Z(f_o)}{r_Z(0)-r_Z(f_o)} \tag{8}$$

In Equation (8), f_o is the fault frequency. r_Z is the 1.5-dimensional energy spectrum analysis total energy. T is the weighting factor. Secondly, MultiScale Permutation Entropy (MPE) is used to extract the time-domain features (TDFs) of the vibration signals (VS) and 1.5-dimensional energy spectrum analysis S-index to extract the frequency-domain features (FDFs) of the VS. Construct the fault feature extraction model of roller bearing signal of CWP in NPP. The specific process of the model is as follows: firstly, the sensors are arranged in two orthogonal directions, respectively +X-axis and +Y-axis, and the vertical direction is +Z-axis in the shell part of the CWP, with a total of three positions. Then, MPE and 1.5-Dimensional Energy Spectrum $WHNR$ are used to enhance the VS features in the three directions, and the three optimal IMFs are obtained at the same time. Finally, the TDF MPE and FDF $WHNR$ of the optimal IMFs of each sensor are extracted, and the time-frequency feature matrix of the VSs is constructed. Among them, the expression for calculating the MPE value $\overline{J_\delta(\phi)}$ after normalizing the alignment entropy of each coarse-grained sequence is shown in Equation (9).

$$\overline{J_\delta(\phi)} = \frac{-\sum_{g=1}^{k} I_g ln I_g}{ln(\phi!)} \tag{9}$$

In Equation (9), ϕ is the embedding dimension and $g = 1,2,\ldots,k, k < \phi!$, I are the elements. Here, the RB fault feature extraction of CWP in NPP is completed. Next, in order to solve the problem that CWP in NPP in the monitoring state of HD features will have dimensional catastrophe, the study introduces the KPCA algorithm to reduce the dimensionality of the data. Its calculation method steps first set the input matrix as $L_j \in W^{\varphi}(j = 1,2,\cdots,N)$. Among them, N is the number of samples and φ is the variables. The data in the original space is mapped to the HD feature space F by the nonlinear mapping function ϑ. The vector L_j is mapped to $\varphi(L_j)$ and the de-meaned mapping $\varphi(l_j)$ is assumed to exist in the HD space. The resulting expression for the covariance matrix calculation is shown in Equation (10).

$$C^F = \frac{1}{N}(\sum_{j=1}^{N} \varphi(l_j)\varphi(l_j)^T \tag{10}$$

In Equation (10), C^F is the covariance matrix. The covariance matrix's eigenvalues and eigenvectors must then be calculated, and the computational expression must be conformed as shown by Equation (11).

$$C^F \partial = \partial \upsilon \tag{11}$$

In Equation (11), ∂ is the eigenvalue. υ is the eigenvector. Next, υ is regarded as a linear data set. The coefficients $\psi_r(r = 1,2,\cdots,N)$ exist, which leads to the eigenvector computational expression as shown in Equation (12).

$$\upsilon = \sum_{r-1}^{N} \psi_r \varphi(l_r) \tag{12}$$

As a result, Equation (2) can be simplified to obtain the computational formula $N\partial r = Kr$. Among them, K is the kernel matrix of $N \times N$. and K_{rj} is the dot product of $\varphi(l_j)$ and $\varphi(l_r)$. As a result, the expression for the computation of t_k

, the projection of $\varphi(l)$ in the eigenvector υ_k, is obtained as shown in Equation (13).

$$t_k = \langle \upsilon_k, \varphi(l) \rangle = \sum_{r=1}^{N} \psi_j^k \langle \varphi(l_r), \varphi(l) \rangle \quad (13)$$

In addition, the calculation of the kernel matrix requires a kernel function, and the study combines the practical situation with the Gaussian kernel function. Equation (14) provides the expression for its computation.

$$K(l_r, l_j) = exp\left(-(l_r - l_j)/2\sigma^2\right) \quad (14)$$

In Equation (14), σ is the nuclear parameter. Immediately after that, to better detect the operation state of the RB of the CWP. The study introduces the EMD combined with KPCA under the Lajda criterion to construct the EMD-KPCA algorithm. Moreover, it is applied to the feature extraction model of RB VS based on improved VMD, so as to construct the condition monitoring model based on KPCA-improved VMD-EMD. The specific flow of the model is shown in Figure 5.

Figure 5 shows the specific flow of the model. First of all, the initialization operation is carried out on the parameters of the model. Moreover, the features of the VS are extracted using the kernel cycle based on the improved VMD VS decomposition method as well as the pump roller bearing signal fault feature extraction model, and the feature data are preprocessed after the feature matrix at the same time. Secondly, the processed feature matrix is set to two-dimensional data, and KPCA is utilized to downsize the data. Then, the dimensionality reduced data is input to the distance of each principal component in the EMD and the distance threshold is calculated using the Lajda criterion to determine the KPCA. Finally, the data to be measured is input to the KPCA and the EMD distance is calculated by repeating the previous step. The RB condition monitoring is realized by comparing with the threshold value. Finally, to provide accurate diagnosis of RB faults, the study applies the PSO-TWSVM algorithm to the KPCA-improved VMD-EMD-based condition monitoring model. The FD model of RB of CWP in NPP is constructed based on the reduced dimensional data and TWSVM. Figure 6 displays the model's flow.

Figure 6 shows the specific flow of the proposed FD model. First, basic parameters such as the population size and the maximum number of iterations of the PSO algorithm are set, and the speed and direction of particles in the particle swarm are initialized. Secondly, the particles obtained after initialization are substituted into TWSVM, the training data set is trained and classified, and the classification accuracy is used as the fitness to evaluate the particles, thus obtaining the initial fitness of all particles. Then the particle position and velocity are updated, and the corresponding particle fitness is calculated. If the new particle fitness is better than its own fitness, the value is updated; if the fitness of a particle is better than the global fitness, the value is updated to the global optimal value. Then, the number of iterations of the algorithm is judged. If the maximum number of iterations is reached, the algorithm is stopped, and vice versa, the iteration continues. Finally, the obtained optimal value is put into TWSVM for training and input test data, so as to obtain fault classification results and classification accuracy.

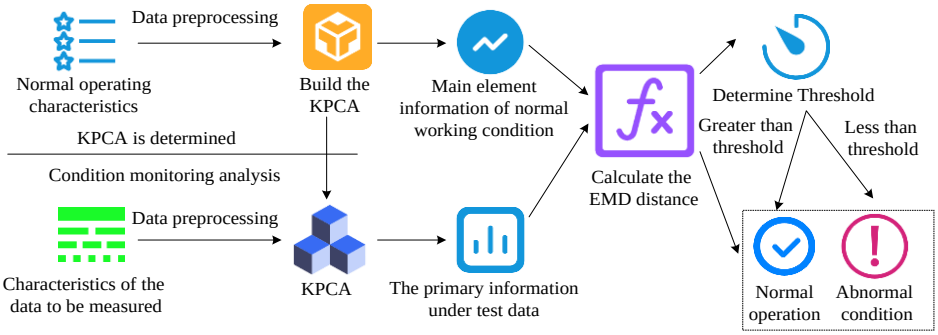


Fig. 5. State monitoring model based on KPCA-improved VMD-EMD

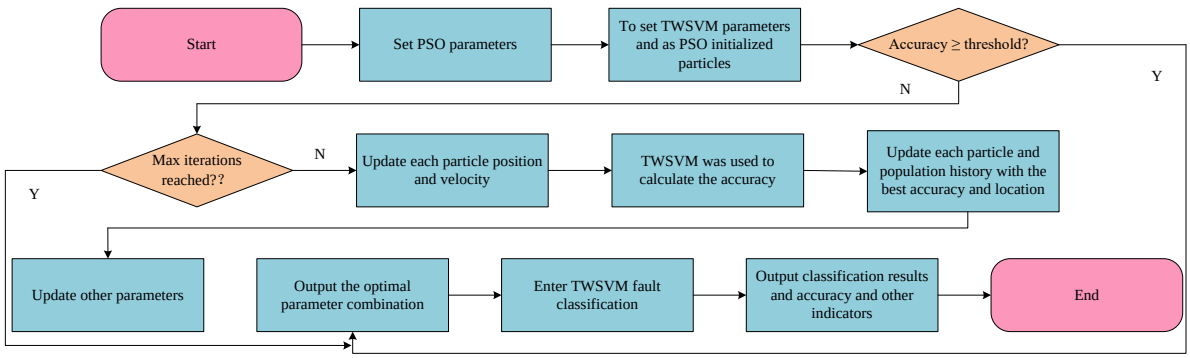


Fig. 6. FD model of rolling bearing of CWP in NPP based on dimensionality reduction data and TWSVM

3. RESULTS

3.1. PSO-TWSVM performance comparison test

For verifying the superior performance of the PSO-TWSVM algorithm (Algorithm 1) proposed in the study, it is compared with the GA-SVM algorithm (Algorithm 2), the PSO-CNN algorithm (Algorithm 3) and the multi-scale collaborative fusion residual neural network (MCFRNN) algorithm (Algorithm 4) in the Matlab simulation software for experimental comparison. The experimental metrics include adaptation and accuracy, etc. In order to ensure the effectiveness of the experiment, a data acquisition system was designed, as shown in Figure 7.

Signal acquisition by setting up the experimental platform and using the data acquisition system to collect data. As shown in Figure. 7 (a), the combined water tank is composed of a heating water tank and a storage water tank with a vertical circulating water pump as the driving pressure head. The data is collected by a data acquisition system composed of electromagnetic flowmeter and temperature sensor arranged around the pump. As can be seen in Figure 7 (b) and Figure 7 (c), the data acquisition system is composed of a performance parameter acquisition module and a vibration signal acquisition module. By imitating the faults such as pitting and breaking of the inner and outer rings of the bearing ring and the rolling body, the bearing signals are collected using sensors and electromagnetic currents, and transmitted to the data acquisition and signal analysis system using the network intelligent acquisition instrument, and finally transmitted to the terminal to complete the collection. The sensors used to extract the experimental signal are temperature sensor and pressure sensor respectively. The temperature measuring element of the temperature sensor is Pt100 thermal resistance with a measuring range of 0 ~ 100℃ and the output current is 4 ~ 20mA standard current. The pressure sensor adopts the static pressure gauge produced by Guangzhou Sennuo Instrument Co., LTD., with a measuring range of 0-2.5MPa, the output signal is 4-20mA, and the pressure measurement accuracy is ±0.1%. The pump is a vertical centrifugal pump, the model is GDS50-250, the flow is 35m3, the head is 70m, the speed is 3950r/min, the driving power is 18.5kW. Network intelligent acquisition instrument, the model is INV3062, 8-channel / 24-bit, each channel sampling frequency is 6.25-51.2kHz. Data acquisition and signal analysis system, model DASP-V11, oscillographic sampling, INV high precision frequency meter, time domain analysis, format conversion. Acceleration sensor, model INV9822, installed resonant frequency >25kHz, resolution is 0.002m/s2. The frequency domain legend of normal vibration signal, bearing inner ring, outer ring and rolling element common pitting failure in the data is shown in Figure 8.

It can be seen from Figure 8 that the time domain distribution of VS of normal bearing VS, inner ring of bearing, outer ring of bearing and rolling body are different when pitting corrosion fault occurs, which have their own unique characteristics. The frequency domain characteristic legend of normal bearing vibration signal, bearing inner ring, outer ring and rolling element when pitting fault occurs is shown in Figure 9. It can be seen from Figure 9 that when pitting corrosion fault occurs, there are obvious differences in the amplitude distribution characteristics of the frequency domain of the inner, outer, rolling body and normal VS of the bearing ring. By extracting and classifying the features, the fault location and fault type can be accurately distinguished effectively. When the bearing fails, the peak value will change dramatically. The peak value of vibration signal in frequency domain and time domain is reflected in the data set, and the fault is inferred by the peak value. Through the data acquisition system, a total of 27621 pieces of data were collected as the vibration signal data of rolling bearings, 80% of which were used as the training set and 20% as the test set. The specific experimental environment of this research is shown in Table 1.

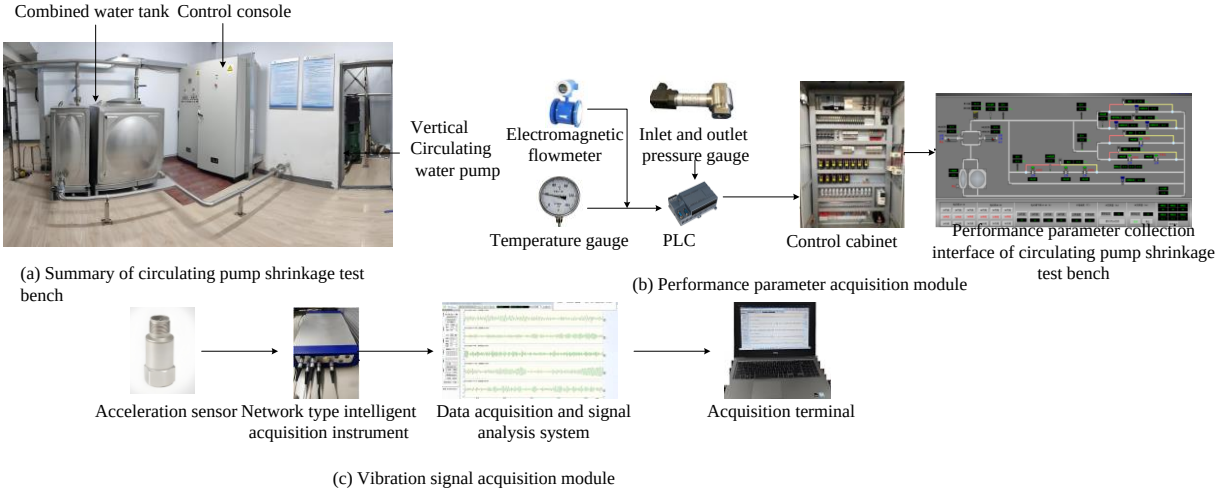


Fig. 7. Data acquisition system

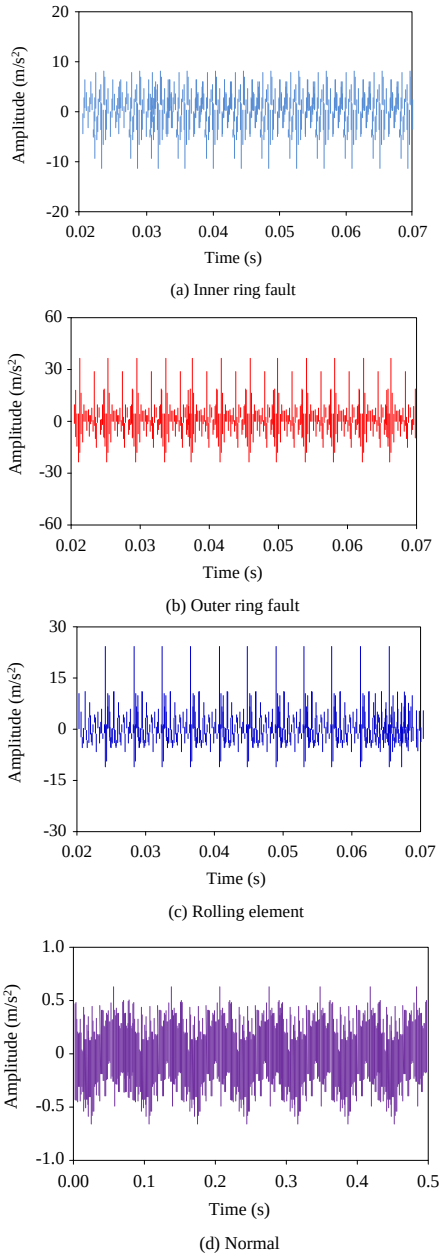


Fig. 8. Four kinds of fault time domain images

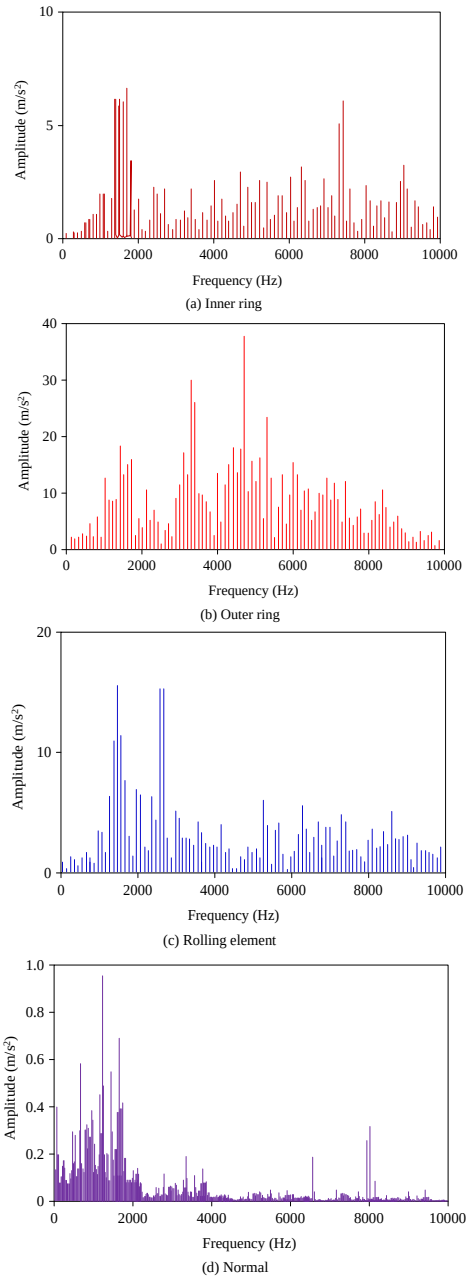


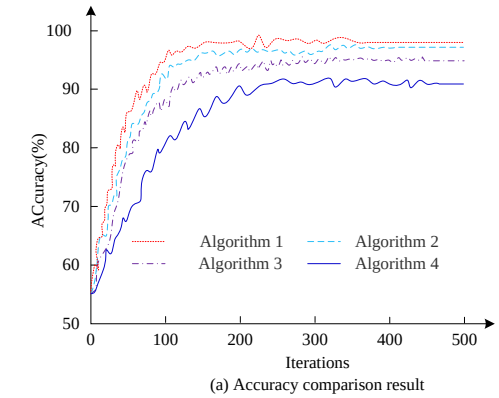
Fig. 9. Frequency domain diagram of four fault types

Table 1. Specific experimental environment of this study

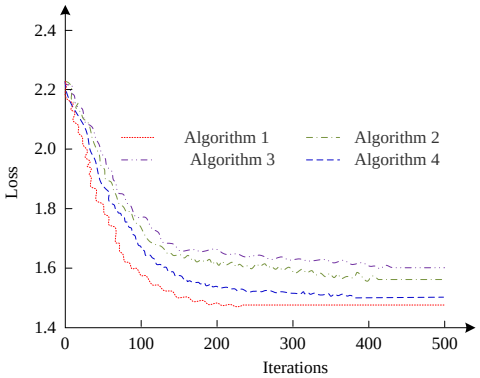
Parameter names	Parameter
Processor	Intel Core i9-13900K
Main frequency	5.8GHz
Internal memory	32GB
Hard disk capacity	500GB
Operating system	Windows 10 64
Matlab version	Matlab 2023a
Data analysis software	Spss24.0

In the above environment, the four algorithms are firstly experimented for accuracy and loss value comparison. Figure 10 displays the outcomes of each algorithm's accuracy and loss numbers.

In Figure 10(a), the average accuracy of Algorithm 1, 2, 3, and 4 are 97.8%, 95.1%, 94.2%, and 89.6%, respectively. Among them, Algorithm 1 proposed by the study has the highest average accuracy. In Figure 10(b), the loss value curve of Algorithm 1 starts to converge at iteration 227 generations. Its average loss value is 1.39, which is lower than 1.62 for Algorithm 2, 1.69 for Algorithm 3, and 1.52 for Algorithm 4. The above results illustrate that Algorithm 1 outperforms the comparative algorithms in terms of both quasi-average accuracy and loss value dimensions. The results of the fitness curves and runtime comparisons of the algorithms are shown in Figure 11.

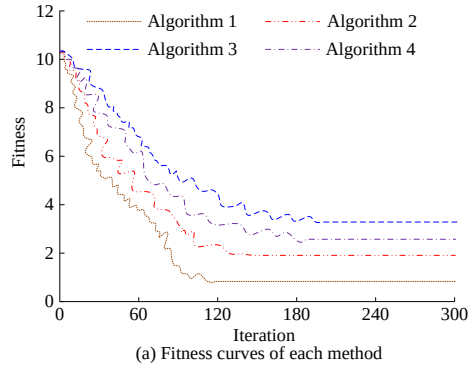


(a) Accuracy comparison result

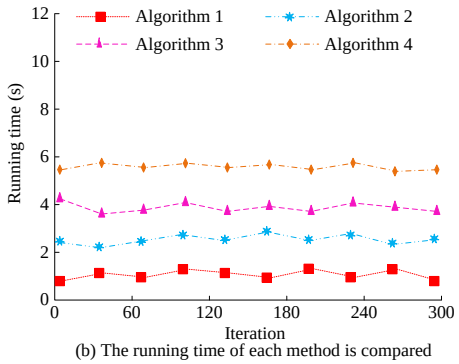


(b) Loss comparison result

Fig. 10. Loss and accuracy curves for each algorithm



(a) Fitness curves of each method

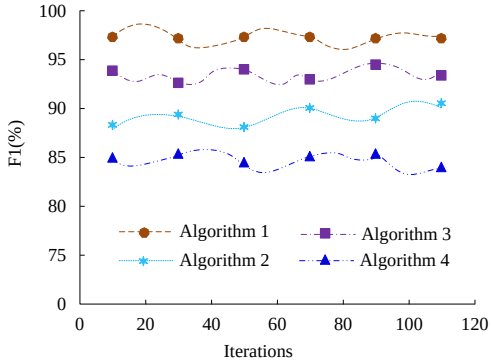


(b) The running time of each method is compared

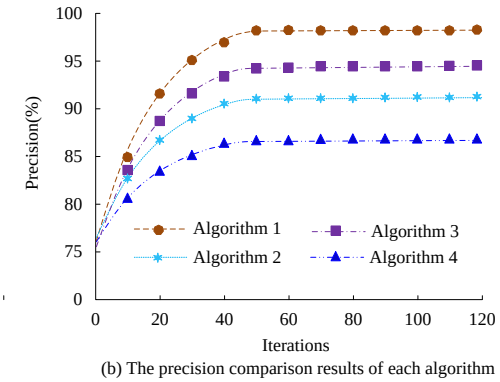
Fig. 11. The fitness curve and the run time comparison results of the algorithm

In Figure 11(a), the fitness curve of Algorithm 1 is the first to converge and starts converging around 118 generations of iterations. The fitness curve of Algorithm 2 starts to converge around 139 generations of iterations. The fitness curve for Algorithm 3 begins to converge at around 196 generations of iterations. The fitness profile of Algorithm 4 starts converging around 203 generations of iterations. In Figure 11(b), the average running times of Algorithm 1, 2, 3, and 4 are 0.86 s, 2.13 s, 3.96 s, and 5.69 s, respectively. The aforementioned findings demonstrate that, in terms of running time and adaption dimension, Algorithm 1 outperforms the comparative algorithms. Figure 12 displays the comparison findings of each algorithm's F1 value and accuracy rate.

In Figure 12(a), the average F1 values of Algorithm 1, 2, 3 and 4 are 97.5%, 88.6%, 94.6% and 85.3%, respectively. Among them, Algorithm 1 proposed in the study has the highest average F1 value. In Figure 12(b), the average accuracy of Algorithm 1 is 97.2%, which is higher than that of Algorithm 2 (91.3%), Algorithm 3 (93.9%) and Algorithm 4 (86.1%). The aforementioned findings demonstrate that, in terms of accuracy and F1 value dimensions, the suggested Algorithm 1 performs better than the comparative algorithms. When the aforementioned findings are combined, they demonstrate that the study's suggested algorithm performs more effectively than the comparative algorithms in terms of running time dimensions, precision rate, F1 value, accuracy rate, loss value, and fitness curve.



(a) The F1 values of each algorithm are compared



(b) The precision comparison results of each algorithm

Fig. 12. F1 value contrast results and the precision of these algorithms

3.2. Performance analysis of the rolling bearing FD model for CWPs

After verifying the superior performance of the PSO-TWSVM algorithm, the study also analyzes the performance of the algorithm-based FD of RB model

(Model 1) for CWP. The study analyzes it experimentally in comparison with GA-SVM based FD of RB model for CWP (Model 2) and FPSO-CNN based FD of RB model for CWP (Model 3). The models are put on the dataset to diagnose four kinds of faults and normal operating conditions. The correct diagnosis rate of each model is shown in Figure 13.

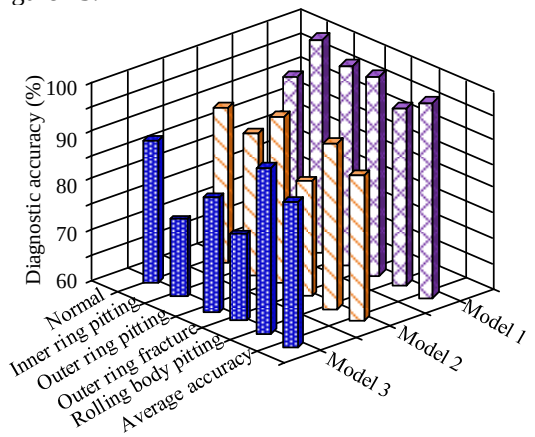


Fig. 13. Diagnostic accuracy rate of each model

In Figure 13, the average diagnostic accuracy of Model 1, 2, and 3 for the same fault on the dataset are 94.7%, 86.1%, and 79.8%, respectively. Among them, Model 1 has the highest average diagnostic accuracy. The findings suggest that the model suggested by the research is feasible and outperforms the comparison models when it comes to the diagnosis dimension for the same issue on the dataset. To further validate the superior performance of the proposed Model 1, the study conducts visualization classification effect comparison experiments with Model 2 and 3, and the MCFRNN-based FD model for the RB of CWP. The visual classification effect of each model is shown in Figure 14.

In Figure 14(b), the classification effect of Model 2 is the worst, and the four types of faults are confused together to differentiate inconspicuously, and the visualization effect is the worst. In Figure 14(c), the classification effect of Model 3 is better than that of Model 2, but there is a situation that three fault types are mixed together. In Figure 14(d), the classification effect of Model 3 is better than that of Model 2 and 3, but a small amount of dispersion and a slightly worse aggregation of the same type. In Figure 14(a), Model 1 has the best classification effect, with significant aggregation of each fault type and obvious differentiation. The above results show that from the dimension of visualization classification effect, the proposed model of the study has good classification effect and its performance is better than the comparison model. When the aforementioned findings are combined, the research-proposed model performs better than the comparison model in terms of fault accuracy and visualization classification impact.

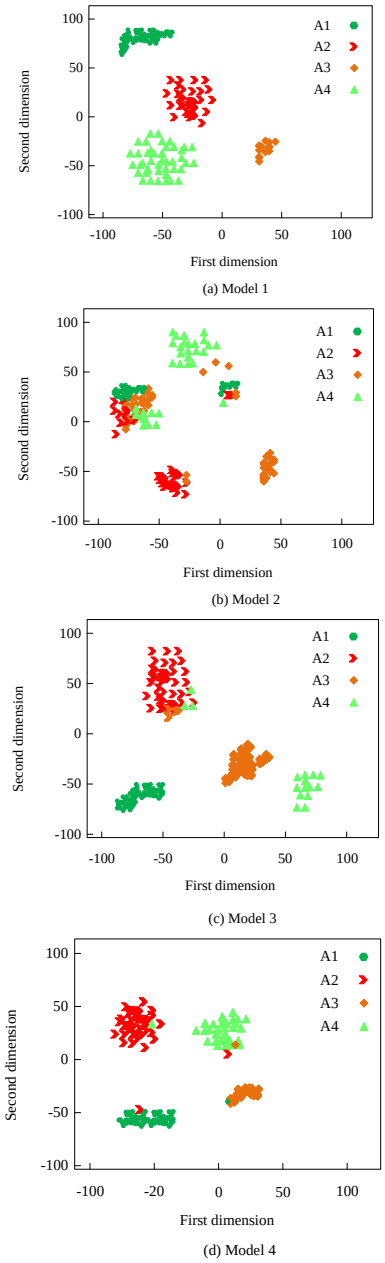


Fig. 14. Each model visualization classification effect

4. DISCUSSION

In this study, the performance of PSO-TWSVM algorithm was analyzed in comparative experiments and the performance of FD of RB model for CWP based on PSO-TWSVM algorithm was analyzed in comparative experiments. The outcomes indicated that the PSO-TWSVM algorithm had significant advantages in the dimensions of accuracy rate, loss value and running time. In the accuracy comparison experiments, the average accuracy of PSO-TWSVM, GA-SVM, PSO-CNN, and MCFRNN algorithms were 97.8%, 95.1%, 94.2%, and 89.6%, respectively. Among them, the PSO-TWSVM algorithm proposed in the study achieved the highest accuracy rate, indicating that the optimization feature of the PSO algorithm in this algorithm optimized the performance of the algorithm and improved its classification effect. The results were

similar to the results of the improved PSO-TWSVM algorithm proposed by Nan et al (17). This result indicated that the PSO-TWSVM algorithm had high accuracy in recognition and classification in the real application environment, thus improving the classification ability of the model. The average loss values of PSO-TWSVM, GA-SVM, PSO-CNN, and MCFRNN algorithms in the loss value comparison experiments were 1.39, 1.62, 1.69, and 1.52, respectively. Among them, the research-proposed PSO-TWSVM exhibited the lowest average loss value. This result indicated that the proposed PSO-TWSVM algorithm of the study improved the fitting effect of the algorithm and improved the performance of the algorithm. In the running time comparison experiments, the average running time of PSO-TWSVM, GA-SVM, PSO-CNN, and MCFRNN algorithms were 0.86 s, 2.13 s, 3.96 s, and 5.69 s, respectively. Among them, the average running time of PSO-TWSVM proposed in the study was the shortest. This result indicated that the introduction of PSO algorithm improved the computational efficiency of TWSVM algorithm. Meanwhile, in the F1 value and accuracy rate comparison experiments, the average F1 value and average accuracy rate of the research-proposed TWSVM algorithm were 97.5% and 97.2%, respectively. The results were better than the comparison algorithms, which further verified the superior performance of the algorithm. Wang and Chen obtained similar conclusions in the proposed PSO-TWSVM algorithm (18). Secondly, in the comparative analysis of the performance of the FD of RB model for CWP, it was found that the PSO-TWSVM-based FD model has good performance in terms of visualization classification effect and diagnosis accuracy. In the comparative analysis of visualization classification effect, its classification effect was the best. This result indicated that the introduction of the improved VMD enhanced the fault features and improved the classification effect of the model. Cheng's team obtained similar conclusions in a related study in 2022 (19). In the comparative FD analysis, the average diagnosis correctness of the proposed model of the study was 94.7%, which was significantly better than the comparison model. This conclusion was consistent with the related study done by Sun et al. in 2023 (20). This result indicated that the introduction of KPCA improved the model FD performance.

5. CONCLUSIONS

Aiming at the current FD method of RB of CWP in NPP, which has problems such as inaccurate diagnosis, the study presented a method that combined PSO and TWSVM and in an attempt to solve the problem of difficulty in extracting the features of RB. Penalty function, WHNR and 1.5 Dimensional Energy Spectrum were introduced to improve the VMD for decomposing the VSs. Meanwhile, in an attempt to better extract the VS features, MPE was combined with the improved VMD vibration decomposition method. Moreover, to increase the calculation speed of the algorithm and monitor the RB operation status, EMD was combined with the KPCA algorithm, so as to construct the condition monitoring model based on KPCA-improved VMD-EMD. Finally, the PSO-TWSVM algorithm was applied to the condition monitoring model to construct a data dimensionality reduction and TWSVM FD model of RB of CWP in NPP. The average loss values of the proposed PSO-TWSVM, GA-SVM, PSO-CNN and MCFRNN algorithms are 1.39, 1.62, 1.69 and 1.52, respectively. The average running time was 0.86s, 2.13s, 3.96s and 5.69s, respectively. The average accuracy was 97.8%, 95.1%, 94.2% and 89.6%, respectively. The average F1 values were 97.5%, 88.6%, 94.6% and 85.3%, respectively. The above results show that the performance of the proposed PSO-TWSVM algorithm is significantly better than the comparison algorithm in terms of running time, accuracy and loss value. Then, through the comparative analysis of FD model experiments, it was found that the average diagnostic accuracy of the FD model, GA-SVM, PSO-CNN and MCFRNN models proposed in this study were 94.7%, 86.1% and 79.8%, respectively, among which the diagnostic accuracy of the proposed model was the highest. The aforementioned findings show that the PSO-TWSV FD model put forth in the study is quite successful at fostering the growth of FD of RB of CWP and enhancing its accuracy. The limitation of this study is that the operational environment of the CWP in a NPP is more complex in the real world. Complex working environment such as media characteristics, temperature and pressure changes, environmental factors and operating conditions and other aspects. This complex real-world environment represents a promising avenue for future research.

Source of funding: *This research received no external funding.*

Author contributions: *research concept and design, Z.C.; Collection and/or assembly of data, D.L.; Data analysis and interpretation, H.S.; Writing the article, W.L.; Critical revision of the article, Z.C., D.L.; Final approval of the article, H.S., W.L.*

Declaration of competing interest: *The author declares no conflict of interest.*

REFERENCE

1. Altamimi R, El-Genk MS. Design and analyses of miniature, submersible annular linear induction pump for test loops supporting development of advanced nuclear reactors. Nuclear Science and Engineering. 2024; 198(8): 1620-1644. https://doi.org/10.1080/00295639.2023.2255461.
2. Xu R, Wang H, Peng MJ, Liu, Y.K. An improved regularized particle filter for remaining useful life prediction in nuclear plant electric gate valves. Nuclear

Engineering and Technology. 2022; 54(6): 2107-2119. https://doi.org/10.1016/j.net.2021.12.001.
3. Omri F, Choura O, Hadj Taieb L, Elaoud, S. Prediction of bearing fault effect on the hydraulic performances of a centrifugal water pump. Journal of Vibration Engineering & Technologies. 2022; 10(5): 1905-1915. https://doi.org/10.1007/s42417-022-00490-3.
4. Daiki G, Tsuyoshi I, Takekiyo H, Shota, Y, Keiichi, K, Shigeyuki, T. Failure diagnosis and physical interpretation of journal bearing for slurry liquid using long-term real vibration data. Structural Health Monitoring. 2023;23(2):1201-1216. https://doi.org/10.1177/14759217231184579.
5. Bin L, Rongsheng Z, Qian H, Yongyong, Z., Qiang, F., Xiuli, W. Fault diagnosis of horizontal centrifugal pump orifice ring wear and blade fracture based on complete ensemble empirical mode decomposition with adaptive noise-singular value decomposition algorithm. Journal of Vibration and Control. 2023; 34(2): 1361-6501. https://doi.org/10.1177/10775463231218494.
6. Dutta N, Kaliannan P, Paramasivam S. A comprehensive review on fault detection and analysis in the pumping system. International Journal of Ambient Energy. 2022;43(1):6878-6898. https://doi.org/10.1080/01430750.2022.2056917.
7. Pu H, Wen Z, Sun X, Han L, Na Y, Liu H. Research on the mechanical fault diagnosis method based on sound signal and IEMD-DDCNN. International Journal of Intelligent Computing and Cybernetics. 2023; 16(3): 629-646. https://doi.org/10.1108/IJICC-09-2022-0253.
8. Goyal N, Gupta K. A hierarchical laplacian TWSVM using similarity clustering for leaf classification. Cluster Computing. 2022;25(2):1541-1560. https://doi.org/10.1007/s10586-022-03534-1.
9. Aguiar ALS, Sousa FBC, Melo YVLD. Optical distribution network design using PSO. IEEE Communications Letters. 2023;27(1):239-242. https://doi.org/10.1109/LCOMM.2022.3218530.
10. Liu J, Song Y, Yu X. Risk assessment study of hydrogen energy storage system based on KPCA-TSO-LSSVM. International Journal of Hydrogen Energy. 2024; 79(2): 931-942. https://doi.org/10.1016/j.ijhydene.2024.07.070.
11. Huang H, Yao Z, Wei X, Zhou, Y. Twin support vector machines based on chaotic mapping dung beetle optimization algorithm. Journal of Computational Design and Engineering. 2024;11(3):101-110. https://doi.org/10.1093/jcde/qwae040.
12. Du S, Song R, Qu Q, Zhao Z, Sun H, Chen Y. Surface deformation prediction model of high and steep open-pit slope based on APSO and TWSVM. Elektronika ir Elektrotechnika. 2024;30(1):77-83. https://doi.org/10.5755/j02.eie.36115.
13. Groumpos PP. A critical historic overview of artificial intelligence: issues, challenges, opportunities, and threats. Artificial Intelligence and Applications. 2023; 1(4):197-213. https://doi.org/10.47852/bonviewAIA3202689.
14. Xiao Y, Liu J, Wen K, Liu, B, Zhao, L, Kong, X. A least squares twin support vector machine method with uncertain data. Applied Intelligence. 2023; 53(9): 10668-10684. https://doi.org/10.1007/s10489-022-03897-3.
15. Sangrody R, Taheri S, Cretu AM, Pouresmaeil E. An improved PSO-based MPPT technique using stability and steady state analyses under partial shading conditions. IEEE Transactions on Sustainable Energy. 2023; 15(1): 136-145. https://doi.org/10.1109/TSTE.2023.3274939.
16. Yu Z, Si Z, Li X, Wang, D, Song, H. A novel hybrid particle swarm optimization algorithm for path planning of UAVs. IEEE Internet of Things Journal. 2022; 9(22): 22547-22558. https://doi.org/10.1109/JIOT.2022.3182798.
17. Nan J, Wang J, Wu H, Li, K. Optimized extreme learning machine by an improved Harris hawks optimization algorithm for mine fire flame recognition. Mining, Metallurgy & Exploration. 2023; 40(1): 367-388. https://doi.org/10.1007/s42461-022-00719-5.
18. Wang J, Chen X, Zhang Y. Research on modelling and analysis of factors influencing students' classroom communication ability based on support vector machine. International Journal of Continuing Engineering Education and Life Long Learning. 2022; 32(4):403-417. https://doi.org/10.1504/IJCEELL.2022.124938.
19. Cheng W, Liu X, Xing J, Ding, B., Zhang, R. AFARN: Domain adaptation for intelligent cross-domain bearing fault diagnosis in nuclear circulating water pump. IEEE Transactions on Industrial Informatics. 2022; 19(3): 3229-3239. https://doi.org/10.1109/TII.2022.3177459.
20. Sun C, Song D, Liu H. Pump fault detection based on MFCC-MLCNN. Academic Journal of Science and Technology. 2023; 8(3): 90-97. https://doi.org/10.54097/mcwvm749.

Zimeng CHEN was born in Huanggang, Hubei. P.R. China, in 1989. He received the bachelor's degree from Shandong University, P.R. China. Now, he works in Research Institute of Nuclear Power Operation, His research interests include equipment reliability, prognostic and health management.
E-mail: chenzimeng2018@126.com

Diliang LI was born in Yichun City, Jiangxi Province, China, in 1987. He received the Master degree from Xi'an University of Electronic Science and Technology. Now, He currently employed at Research Institute of Nuclear Power Operation, with a focus on digital nuclear power.
e-mail: 15802786387@163.com

Hua Su was born in Xinyang, Henan, P.R. China, in 1991. She received the Master degree from Beihang University, P.R. China. Now, she works in Research Institute of Nuclear Power Operation. Her research interests include equipment reliability, prognostic and health management.
e-mail: suhua20240703@163.com

Wei Liu obtained his ME in Mechanical and Electronic Engineering from Xi & apos; dian University in 2014. Presently, he is workings as a reliability engineer at Research Institute of Nuclear Power Operation, his area of interest is reliability management of nuclear power plant equipment.
e-mail: 15332380761@163.com