IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 11, 2020, accepted December 27, 2020, date of publication December 31, 2020,
date of current version January 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3048492

A New Performance Degradation Evaluation Method Integrating PCA, PSR and KELM

MINGYANG LV[1], (Member, IEEE), CHUNGUANG ZHANG[1], (Member, IEEE), AIBIN GUO[2], (Member, IEEE), AND FANG LIU[3], (Member, IEEE)

[1]School of Electronics and Information Engineering, Dalian Jiaotong University, Dalian 116028, China
[2]Haifeng General Aviation Technology Company Ltd., Beijing 100070, China
[3]Institute of Information, Zhejiang University of Finance and Economics, Hangzhou 310018, China

Corresponding author: Chunguang Zhang (zcg413@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51605068 and Grant 51475065, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY19F020011, in part by the China Education Ministry Humanities and Social Science Research Youth Fund Project under Grant 18YJCZH192, and in part by the Liaoning BaiQianWan Talents Program.

ABSTRACT In order to better characterize the performance degradation trend of rolling bearings, a new performance degradation evaluation method based on principal component analysis (PCA), phase space reconstruction (PSR) and kernel extreme learning machine (KELM), namely PAPRKM is proposed to evaluate the performance degradation of rolling bearings in this paper. In the PAPRKM method, the time-domain and frequency-domain features of the vibration signal are extracted to construct the high-dimension feature matrix. Then the PCA is used to reduce the dimension of the feature matrix in order to represent the running state and the declining trend of rolling bearings, so as to eliminate the redundancy and information conflict among these features. Nextly, the PSR is adopted to obtain more relevant information from the time series. By determining the delay time and embedding dimension, the time series are reconstructed to obtain a new performance degradation index, which is regarded as the input data to input into KELM, and the degradation trend prediction model is established to realize the performance degradation trend prediction. Finally, the actual vibration signals of rolling bearings are applied to prove the effectiveness of the PAPRKM. The obtained experimental results show that the PAPRKM method can effectively predict the performance degradation trend of rolling bearings. The predicted results are more accurate than the other compared methods.

INDEX TERMS Rolling bearing, performance degradation trend prediction, feature extraction, principal component analysis, phase space reconstruction, kernel extreme learning machine.

I. INTRODUCTION

As one of the most widely used parts in mechanical system, the operation condition of rolling bearings directly affects the whole performance of mechanical system. Under the comprehensive application of internal and external factors, the actual service life of rolling bearings is far lower than the designed life. According to statistics, the proportion of accidents caused by rolling bearing failure in mechanical system has increased year by year [1]. Therefore, it is necessary to monitor the state of the rolling bearings and predict the remaining useful life (RUL). In this way, the faults of rolling bearings can be found in advance, which effectively improve the reliability and safety of mechanical equipment operation. In addition, the RUL prediction of rolling bearings is the basis of the reasonable maintenance plan of the unit.

With the continuous development and practice of theoretical research on the useful life of rolling bearings, the research topic of performance degradation assessment and remaining useful life prediction of rolling bearings has a very mature theoretical basis and application value. A series of prediction models are proposed and established in the past few decades. Root mean square (RMS) is used as the feature vector to evaluate the damage of rolling element [2]. The time-domain and frequency-domain feature indexes of bearing vibration signals are extracted as the input of state prediction model to predict the life trend of rolling bearings [3]. The kurtosis is used as the degradation performance index to establish a rolling bearing residual life prediction model based on SVM [4]. Based on the normalized state deviation, the method of state evaluation and residual life prediction of

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

rolling bearings is constructed [5]. The RMS and kurtosis coefficients are combined to construct a time-domain index TALAF as the rolling bearing feature and extract the feature vector of rolling bearing vibration signal [6]. The reliability of rolling bearing is predicted by using wavelet packet analysis and state space model (SSM) [7]–[14]. Firstly, the wavelet packet analysis is used to extract the feature of the vibration signal, and then the correlation band energy is combined with the on-line monitoring data. The average filter is used as the input of the estimation model to improve the model parameters. Finally, the SSM prediction method of degradation index is established to predict the probability density distribution of degradation index and obtain the reliability. In addition, some researchers proposed a lot of algorithms, which can be combined with different prediction models [15]–[33].

To sum up, the most of the existing feature extraction methods are based on vibration signals, the features of time-domain and frequency-domain are extracted, and a single time-domain index is selected as the feature vector. But these methods are lack of joint extraction of time-domain and frequency-domain features and analysis of each feature index. When there are fewer features, the relevant features may be missed. When there are more features, the feature redundancy may occur, which will result in certain interference. In view of these problems, the rationality of feature extraction of time-domain and frequency-domain, and the influence of redundancy and correlation between multi-feature indexes on the accuracy of bearing performance degradation evaluation are considered in this paper. The PCA, PSR and KELM are integrated to propose a new method, named PAPRKM for evaluating performance degradation trend of rolling bearings. The PCA is used to reduce the dimension of the selected features. The phase space of the features after dimension reduction and fusion is reconstructed to form the performance degradation index with accurately describing the running state and declining trend of rolling bearings. At the same time, the KELM is a new learning algorithm with single hidden layer feed-forward neural network [34], which has the advantages of strong generalization learning ability and fast training speed, and has been successfully applied to fault diagnosis, power load prediction and wind power prediction [35]–[37]. Therefore, it is used to establish the performance degradation prediction model of rolling bearings. The PAPRKM is applied in the life-cycle vibration data of rolling bearings to evaluate the prediction accuracy.

The main contributions of this paper are summarized as follows:

- A new performance degradation evaluation method of rolling bearings based on creatively combining principal component analysis, phase space reconstruction and kernel extreme learning machine is proposed to realize the performance degradation evaluation.
- The motivation to combine PCA, PSR with KELM is to make full use of the dimension reduction ability of PCA, the nonlinear characteristics of PSR and learning speed and generalization ability of KELM to realize the performance degradation evaluation of rolling bearings.
- The performance of the PAPRKM method is has been extensively investigated by the vibration signals of rolling bearings.
- The PAPRKM method can effectively and accurately evaluate the performance degradation process.

II. BASIC METHODS

A. TIME-DOMAIN AND FREQUENCY-DOMAIN SIGNAL ANALYSIS

Based on feature analysis of time-domain signal, the dimensionless index and dimensional index of time-domain signal are analyzed. The dimensional index is usually affected by the load and speed of rolling bearings, while the dimensionless index is not affected, which is determined by probability density function [38]. So, two indexes are regarded as feature indexes in time-domain. Some important time-domain indexes are mean value, root mean square value, square root amplitude, kurtosis, and so on. In order to effectively represent the running state of rolling bearings, several time-domain features are combined. Therefore, 10-dimensional indexes including mean value, root mean square value, square root amplitude, absolute average value, skewness, kurtosis, variance, maximum value, minimum value and peak-to-peak value are used. 6-dimensionless indexes including waveform index, peak index, pulse index, clearance factor, skewness index and kurtosis value are used.

The feature analysis of frequency-domain signal is the Fourier transform in essence, the vibration signal is converted into some relatively simple harmonic components, such as phase information and frequency structure, so as to make analysis and judgment easily. Common frequency-domain analysis methods include Fourier transform, frequency response function, power spectrum analysis, energy spectrum analysis and wavelet analysis and so on [39], [40]. Fourier transform uses Fast Fourier Transform (FFT) to transform time-domain signal into frequency-domain signal. It is difficult to identify the features of some signals through time-domain signals, but the contained feature information can be easily reflected from the spectrum after they are transformed into frequency-domain signals [15]. The FFT transformation of signal $x(t)$, $t = 0, 1, \ldots, N\text{-}1$ is described.

$$X(k) = \sum_{t=0}^{N-1} x(t) W_N^{kt}, \quad k = 0, 1, \ldots, N-1, W_N = e^{-\mathrm{j}\frac{2\pi}{N}} \tag{1}$$

B. PRINCIPAL COMPONENT ANALYSIS

PCA is a standard method applied to dimensionality reduction and feature extraction [41]. It is a data analysis method based on second-order statistics. Mathematically, PCA relies on the eigen-decomposition or singular value decomposition of the covariance matrix. The PCA algorithm processes are shown in the following.

1) $Z = [z_1, z_2, \ldots, z_n]^T$ is the signal feature matrix and z_i represents the element. The mean $M(z)$ of the samples can be calculated by

$$M(z) = \frac{1}{n}\sum_{i=1}^{n} z_i \tag{2}$$

2) According to the mean of the samples, the covariance matrix C_x can be constructed as

$$C_x = \frac{1}{n}\sum_{i=1}^{n} (z_i - M)(z_i - M)^T \tag{3}$$

3) The feature value λ and feature matrix ξ of C_x are written as

$$C_x\xi = \lambda\xi \tag{4}$$

4) The number of principal components is determined by the cumulative contribution rate η, which represents the proportion of the first k variances in the total variance. At last, we can get the feature matrix U, as shown in (6).

$$\eta = \frac{\sum_{t=1}^{k} \lambda_t}{\sum_{t=1}^{m} \lambda_t} \tag{5}$$

$$U = [u_1, u_2, \ldots, u_k] \tag{6}$$

In (6), u_k represents the k_{th} principal component of the feature matrix. When the cumulative contribution rate of the k principal components is greater than 90%, U contains most of the information of the original data.

C. PHASE SPACE RECONSTRUCTION

The PSR is a key step in the analysis and processing of time series chaos [42]. It is used to mine the nonlinear characteristics of vibration signal. According to the Takens embedding theorem, a multi-dimensional time series matrix can be reconstructed from one-dimensional time series, and more useful information can be extracted from multi-dimensional time series. The analysis and prediction of time series are carried out in this reconstructed phase space.

The PSR of the measured one-dimensional time series $\boldsymbol{x}(n)$ is carried out to obtain the multi-dimensional time series.

$$\begin{cases} \boldsymbol{X}_1 = [\boldsymbol{x}(1), \boldsymbol{x}(1+\tau), \ldots, \boldsymbol{x}(1+(m-1)\tau)] \\ \boldsymbol{X}_2 = [\boldsymbol{x}(2), \boldsymbol{x}(2+\tau), \ldots, \boldsymbol{x}(2+(m-1)\tau)] \\ \ldots \\ \boldsymbol{X}_N = [\boldsymbol{x}(N), \boldsymbol{x}(N+\tau), \ldots, \boldsymbol{x}(N+(m-1)\tau)] \end{cases} \tag{7}$$

where m is the embedding dimension and τ is the delay time, the number of phase points in the reconstructed phase space is given as follow.

$$M = N - (m-1)\tau \tag{8}$$

The core idea of phase space reconstruction is to select the appropriate embedding dimension m and delay time τ. τ and m are two important parameters for PSR. The rationality selection is directly related to the quality and effect of PSR. In this paper, C-C method is used to select the m and τ to reconstruct the phase space of one-dimensional time series of rolling bearings.

D. EXTREME LEARNING MACHINE

The ELM is proposed based on a single hidden layer feedforward neural network (SLFN), which has extremely generalization and fast learning ability [43]. The advantage of ELM is that it can generate the threshold of hidden layer neurons and the connection weights between input layer and hidden layer randomly. It does not need to adjust the weights of SLFN, so it involves less computational complexity and consumes less time [44].

The network structure of ELM is shown in FIGURE 1.

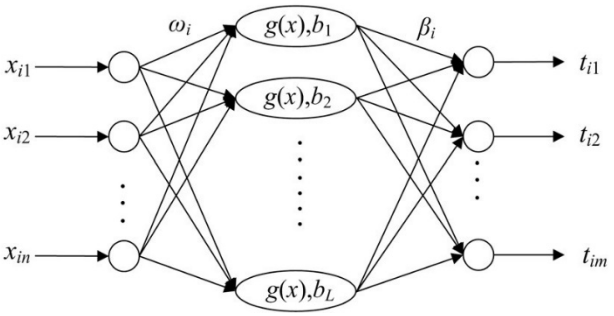


FIGURE 1. Network structure of ELM.

For N different samples (x_j, t_j), $j = 1, 2, \ldots, N$, where $x_j = [x_{j1}, x_{j2}, \ldots, x_{jn}]^T$ is the j_{th} sample, each sample contains n-dimensional features. And $t_j = [t_{j1}, t_{j2}, \ldots, t_{jm}]^T$ is the encoded class label, each sample belongs to m different classes. For ELM with L hidden neurons, the mathematical model can be expressed as:

$$\sum_{i=1}^{L} \beta_i g(w_i \cdot x_j + b_i) = t_j, \quad j = 1, 2, \ldots, N \tag{9}$$

where $g(x)$ excitation function, w_i, b_i, and β_i are the input weights, hidden layer bias and output weights of the i_{th} hidden neuron node respectively. Equation (9) can be expressed in matrix form:

$$H\beta = T \tag{10}$$

where β is the matrix of output layer weights, T is the corresponding coding class label, and H is the hidden layer output matrix:

$$H = \begin{bmatrix} h(x_1) \\ \ldots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) \ldots g(w_L \cdot x_j + b_L) \\ \ldots \\ g(w_1 \cdot x_N + b_1) \ldots g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \tag{11}$$

Since Equation (10) is linear, β can be obtained.

$$\beta = H^{\dagger}T = H^{T}\left(HH^{T}\right)^{-1}T \tag{12}$$

where $H^{\dagger}$ is the Moore-Penrose generalized inverse matrix of the hidden layer output matrix.

III. A PERFORMANCE DEGRADATION EVALUATION METHOD

A. PERFORMANCE DEGRADATION EVALUATION METHOD

The traditional ELM uses explicit nonlinear feature mapping, which often needs more hidden neurons to solve the more complex classification, regression and other nonlinear problems, which result in a very complex network structure. Compared with ELM prediction model, the advantage of KELM only needs to select the kernel function in advance, it does not need to define the mapping function explicitly and set the number of hidden neurons. KELM extends ELM from explicit activation to implicit mapping functions. In many research areas, it has better generalization performance than the traditional ELM algorithm. And KELM replaces random mapping by kernel mapping, which can effectively improve the generalization and stability caused by random assignment of hidden neurons. Therefore, KELM is selected in this paper. The kernel function matrix in the KELM is defined as follow [45]–[56].

$$\Omega_{ELM} = HH^{T} : \Omega_{ELM} = h\left(x_i\right)h\left(x_j\right) = K\left(x_i, x_j\right) \tag{13}$$

where $h\left(x\right)$ is the output function of the hidden layer node; $K\left(x_i, x_j\right)$ is the kernel function, and the types of kernel functions include linear kernel function, polynomial kernel function and RBF kernel function. This paper selects RBF kernel function:

$$K\left(x_i, x_j\right) = \exp\left(-\left|x_i - x_j\right| / \left(2\sigma^2\right)\right) \tag{14}$$

Since the kernel function is calculated in the form of inner product, it is not necessary to set the number of hidden layer neurons, nor to initialize the input weight matrix and offset randomly. Once the training sample data and kernel parameters are determined, the output weight matrix is uniquely determined, which greatly improves the stability and generalization ability of the algorithm.

The output function of the generalized single hidden layer feed-forward neural network model based on KELM is described as follow.

$$\begin{aligned} f\left(x\right) &= h\left(x\right)H^{T}\left(\frac{I}{C} + HH^{T}\right)^{-1}T \\ &= \begin{bmatrix} K\left(x, x_1\right) \\ \cdots \\ K\left(x, x_N\right) \end{bmatrix}^{T}\left(\frac{I}{C} + \Omega_{ELM}\right)^{-1}T \end{aligned} \tag{15}$$

where I is a diagonal matrix; C is a penalty factor. And the weight matrix β is obtained, which makes the ELM more stable and has better generalization performance.

B. MODEL OF PAPRKM

In this paper, in order to make full use of the dimension reduction ability of PCA, the nonlinear characteristics of PSR and learning speed and generalization ability of KELM, a new performance degradation evaluation method of rolling bearings based on creatively combining PCA, PSR and KELM is proposed to realize the performance degradation trend prediction of rolling bearings. The performance degradation evaluation model is show in FIGURE 2.

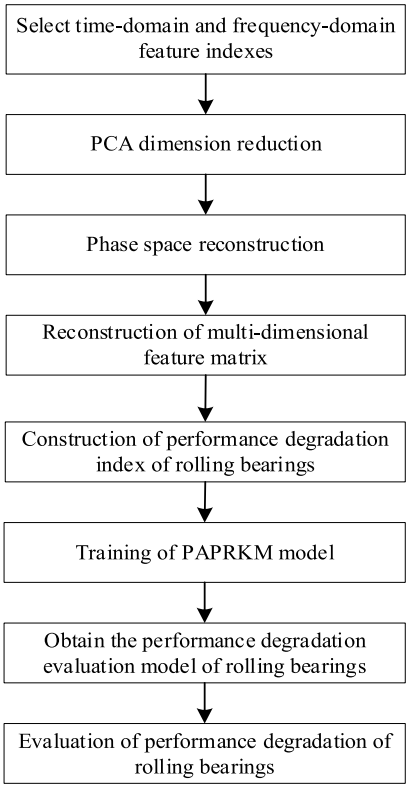


FIGURE 2. Performance degradation prediction process of rolling bearings.

The specific steps of PAPRKM in the performance degradation evaluation of rolling bearings are described as follows.

Step 1: Input the whole life cycle vibration data of rolling bearings and initialize the parameters of PAPRKM model.

Step 2: The time-domain and frequency-domain features of the input rolling bearing life cycle vibration data are extracted to extract the effective features that can reflect the bearing performance changes, and finally 29 time-domain and frequency-domain features are extracted.

Step 3: The dimension of 29 extracted time-domain and frequency-domain features are reduced by PCA, After PCA dimension reduction and fusion, the first principal component summarizes the degradation trend of rolling bearings during its life cycle. It includes the whole vibration signal process of rolling bearings from stable (normal) state in the early stage to fault occurrence time in the later stage. Therefore, the first principal component after dimension reduction is taken as the

feature index to reflect the performance degradation of rolling bearings.

Step 4: Since the first principal component obtained by PCA dimensionality reduction is a one-dimensional time series, which contains insufficient information, the phase space reconstruction technique is used to increase the data dimension and reconstruct the multi-dimensional feature matrix to increase the useful information of the data.

Step 5: The multi-dimensional feature matrix after phase space reconstruction is used as the performance degradation index of rolling bearings, and input into KELM model to train the model, so as to obtain the performance degradation evaluation model of rolling bearings, namely PAPRKM model.

Step 6: The PAPRKM model is used to evaluate the performance degradation of rolling bearings to verify the theoretical effect.

C. PARAMETERS OPTIMIZATION OF KELM

After the kernel function is selected, the KELM needs to select the penalty parameter c and the parameter σ^2 in the kernel function. The parameter c determines the training error and generalization ability, and the parameter σ^2 reflects the distribution or range features of the training sample data. They are used to determine the width of the local area. The larger parameter σ^2 means lower variance. In this paper, the grid search method is used to optimize the two parameters in order to guarantee the stability and accuracy of KELM model.

D. EVALUATION INDEX

In order to quantitatively analyze the prediction performance of the model, the mean absolute percentage error (MAPE) is used as the evaluation index to analyze the effectiveness of the prediction results. The specific expression is described as follow.

$$X_{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\% \tag{16}$$

where N is the number of prediction samples, y_i is the actual value and $\hat{y}_i$ is the predicted value.

IV. DEGRADED FEATURE EXTRACTION METHOD

A. DATA SOURCE AND DESCRIPTION

The bearing dataset used in this experiment comes from IEEE PHM Challenge 2012. This dataset is collected from the PRONOSTIA test platform, as shown in FIGURE 3. Accelerated degradation experiments can be conducted on this platform to generate the run-to-failure vibration signal.

The PRONOSTIA test platform contains rotating part, load part and data collection part. The motor power of the rotating part is 250W. The power is transferred to the bearing by the axis of rotation. The load part provides a load of 4000N to make the bearing degrade quickly. Vibration signals are collected by means of acceleration sensor and temperature sensor, which are placed in horizontal and vertical directions.

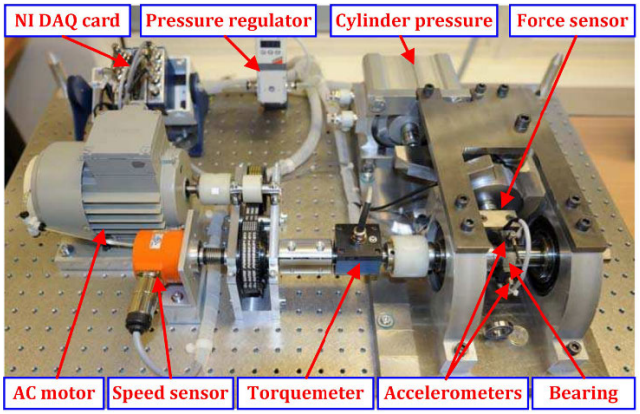


FIGURE 3. The experimental platform of rolling bearings.

The sampling frequency of the acceleration sensor and temperature sensor are 25.6 kHz and 10 Hz respectively.

The data provided by IEEE PHM Challenge 2012 include three working conditions. Under the first condition, the motor's rotation speed is 1800rpm and the load is 4000N. The second condition is a motor speed of 1650rpm and load of 4200N. The third condition is motor speed 1500rpm and load 5000N.

In this experiment, a single bearing is tested for performance degradation, and bearing data bearing1_1 is selected. The experimental environment of the computer is Intel Core i5 2450, 4GB ram, windows 10 and MATLAB 2018b.

B. FEATURE SELECTION

In the time-domain and frequency-domain, the factors reflected by different features are quite different, and the sensitivity to fault degree is also different. In the performance degradation prediction of rolling bearings, the sensitivity of different features to fault degree must be considered. Therefore, it is necessary to remove some undesirable features from the obtained time-domain and frequency-domain features, and several features with better effect are selected.

For bearing data bearing1_1, there are 1428 data files in total, and each data file has 2560 points. For all 1428 data files, the obtained feature map is shown in FIGURE 4 ∼FIGURE 6.

FIGURE 4 shows the dimensional time-domain indexes of rolling bearings, including mean value, minimum value, maximum value, root mean square value, variance, skewness, absolute average amplitude, square root amplitude, kurtosis and peak-to-peak value. In the stage change period of rolling bearings, that is, the sample values from 1100 to 1400, these dimensional time-domain indicators can better reflect the performance change state of rolling bearings in the process of operation, and the stability of each feature index is good, and the experimental effect is good, which should be retained.

FIGURE 5 shows the dimensionless time-domain indexes of rolling bearings, including pulse index, kurtosis index, waveform index, margin index, peak value index and

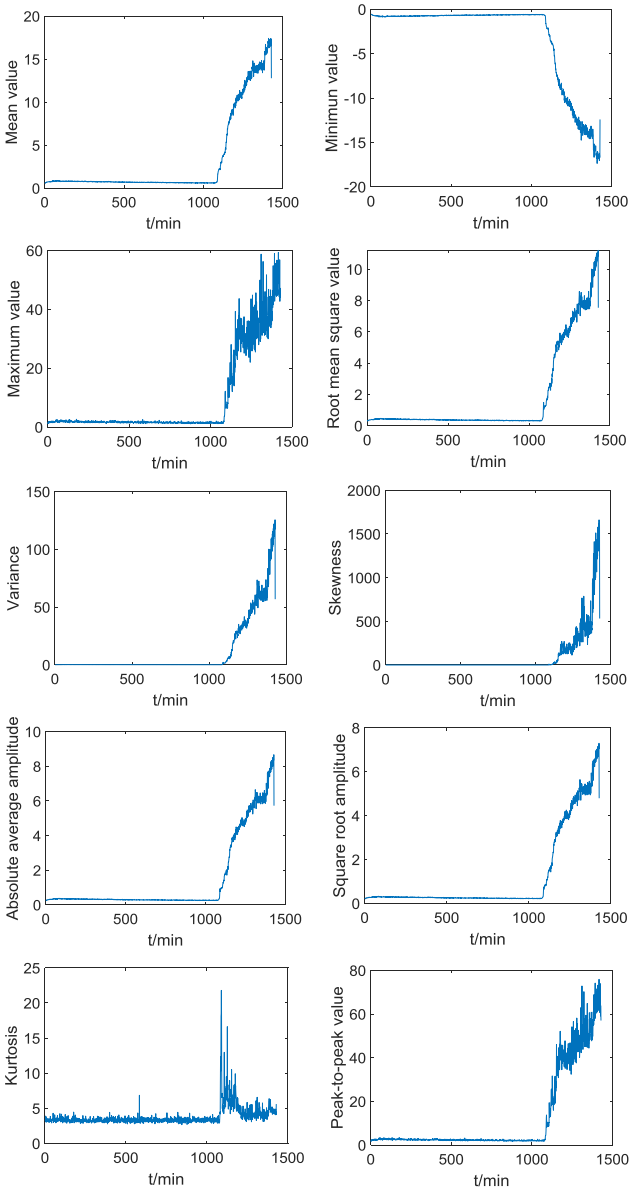


FIGURE 4. Dimensional time-domain index of life cycle.

skewness index. In the stage change period of rolling bearing performance, that is, the sample values from 1100 to 1400, pulse index and skewness index can better reflect the performance change state of rolling bearings in the process of operation, the stability of feature indexes are good, and the experimental effect is good, which should be retained. Although the waveform index, margin index and peak value index show a certain trend in the process of rolling bearing performance degradation, the overall information is noisy, so these feature indexes are eliminated. In addition, the kurtosis index cannot accurately reflect the change process of the rolling bearing in the stage change period of the rolling bearing performance, that is, the sample values from 1100 to 1400, and cannot effectively reflect the performance change state of rolling bearings, the deviation of experimental effect is large, which should be eliminated.

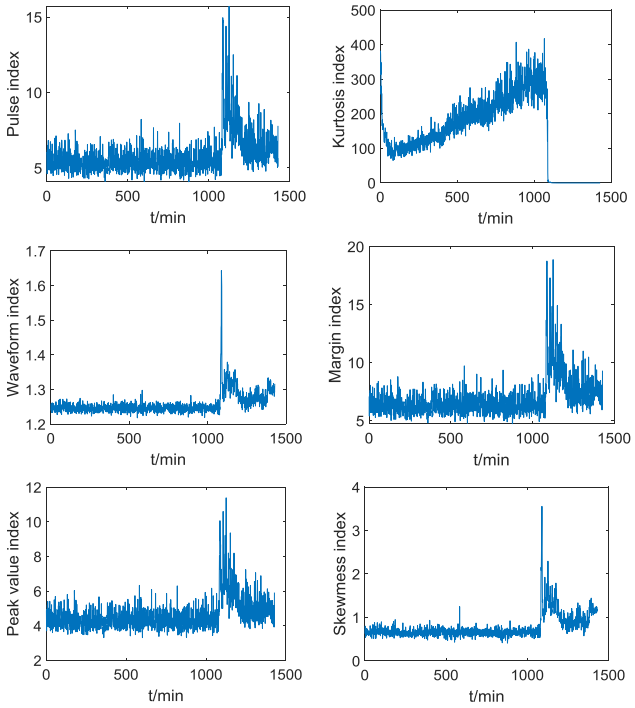


FIGURE 5. Dimensionless time-domain index of life cycle.

It can be seen from FIGURE 4 and FIGURE 5 that different feature indexes reflect different operation states of rolling bearings. Since the performance of rolling bearings shows obvious degradation trend in the whole life cycle, it is necessary to screen these feature indexes in order to find out the feature indexes which can better reflect the degradation trend of rolling bearings. Finally, 12 time-domain feature indexes are selected through comparison and selection, they are mean value, minimum value, maximum value, root mean square value, variance, skewness, absolute average amplitude, square root amplitude, kurtosis, peak-to-peak value, waveform index and skewness index respectively.

In the FIGURE 6, the frequency-domain feature (P1) reflects the magnitude of vibration energy in frequency-domain, the frequency-domain features (P2∼P4, P6, and P10∼P13) reflect the degree of dispersion or concentration of frequency spectrum, and the frequency-domain features (P5 and P7∼P9) reflect the change of main frequency band position. All frequency-domain features show a rising or falling trend, while the frequency-domain features (P7∼P9) show the degradation trend in the event of bearing failure, but the overall information is noisy and the stability of the feature index is poor, so they should be removed. Finally, the frequency-domain features (P1∼P6 and P10∼P13) are selected through comparison and analysis in here.

C. ESTABLISHMENT OF DEGRADATION PERFORMANCE INDEX

Based on the PCA, dimension reduction fusion is carried out for the time-domain and frequency-domain features of

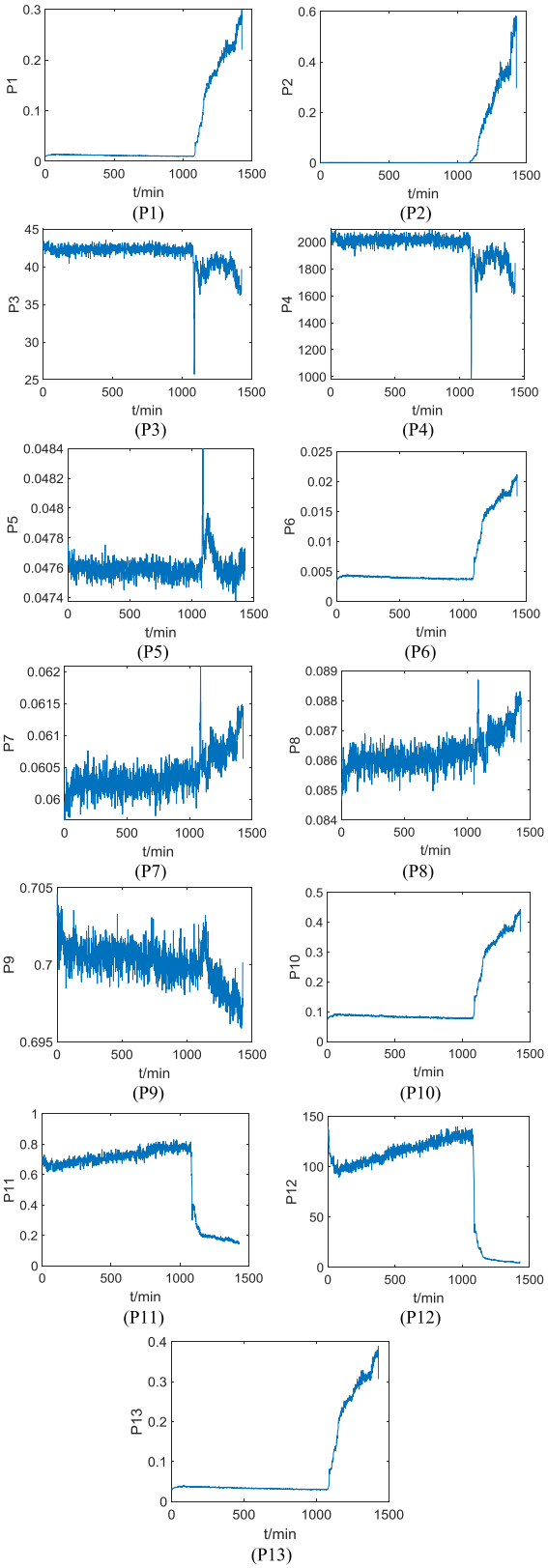


FIGURE 6. Frequency features of rolling bearings.

rolling bearings are selected in the first section. The first principal component is selected to represent the multi-feature information after fusion and as the degradation performance

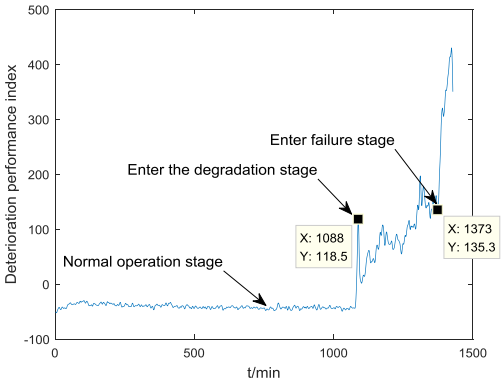


FIGURE 7. Performance degradation index of rolling bearings.

index of rolling bearings. The first principal component after fusion is shown in FIGURE 7.

In FIGURE 7, the curve of performance index from 0 to1088 minutes is relatively smooth, and it can be seen that rolling bearings is in normal period. There is an obvious inflection point of fluctuation at 1088 minute. The main component value suddenly increases, and the amplitude fluctuates up and down. At this time, the rolling bearing operation enters the degradation stage. After 1373 minutes, the main component value suddenly increased to suddenly decreased, which indicates that the rolling bearing has been damaged. At this time, the rolling bearing operation enters the failure stage. It can be seen that the degradation performance index has a process of rising first and then falling after entering the degradation stage. This reason is when the surface defect is just formed, continuous rolling will cause small peeling or crack, but then it will be smoothed by continuous rolling contact. When the damage extends to a wider area, the vibration level will rise again. This phenomenon is called healing.

D. DETERMINATION OF PSR PARAMETERS

For the ideal infinite and noise-free one-dimensional time series, the embedding dimension m and the delay time τ can take any value. But for the vibration signal time series of rolling bearings, they are all finite length sequences with noise, so the embedding dimension m and the delay time τ should be properly selected. According to Part II, the C-C method is used to determine the delay time τ and the embedding dimension m in this paper.

The correlation integral is defined as follow.

$$C(m, N, r, \tau) = \frac{2}{M(M-1)} \sum_{1 \le i < j < M} \theta\left(r - d_{ij}\right), r > 0 \tag{17}$$

where, m is the embedding dimension, n is the number of time series, r is the search radius taken in the calculation, τ is the delay time, $d_{ij} = \left\| x_i - x_j \right\|_\infty$, θ is the Heaviside function.

The test statistics are defined as follow.

$$S_1(m, N, r, \tau) = C(m, N, r, \tau) - C^m(1, N, r, \tau) \tag{18}$$

In the formula (18), the block average method is used to calculate.

$$S_2(m,N,r,\tau)=\frac{1}{\tau}\sum_{s=1}^{\tau}[C_s(m,N/\tau,r,\tau) -C_s^m(1,N/\tau,r,\tau)] \quad (19)$$

The difference is defined as follow.

$$\Delta S\,(m,N,r,\tau)=\max\{S\,(m,N,r,\tau)\} -\min\{S\,(m,N,r,\tau)\} \quad (20)$$

Take $N=1428$, $m=2,3,4,5$, $r=k\sigma/2$, $k=1,2,3,4$, σ is used as the standard deviation of time series $x\,(n)$, $\tau=1,2,\ldots,200$. The mean value is $\overline{S_2}\,(\tau)$ and $\overline{S_1}\,(\tau)$, the mean value of difference is $\Delta\overline{S_2}\,(\tau)$, The $\Delta\overline{S_1}\,(\tau)$ of $S_2\,(m,N,r,\tau)$ and $S_1\,(m,N,r,\tau)$ are calculated respectively. The first local minimum point of $\Delta\overline{S_1}\,(\tau)$ is selected as the delay time τ.

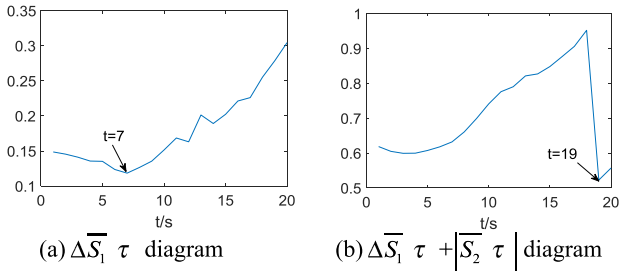
$$\overline{S_2}\,(\tau)=\frac{1}{16}\sum_{m=2}^{5}\sum_{k=1}^{4}S_2\,(m,N,r,\tau) \quad (21)$$

$$\Delta\overline{S_1}\,(\tau)=\frac{1}{4}\sum_{m=2}^{5}\Delta S_1\,(\tau) \quad (22)$$

The optimal embedding window τ_w is selected as the global minimum point of $\Delta\overline{S_1}\,(\tau)+\left|\overline{S_2}\,(\tau)\right|$, so that the embedding dimension m is obtained as follow.

$$m=\tau_w/\tau+1 \quad (23)$$

As shown in FIGURE 8, the first local minimum point of $\Delta\overline{S_1}\,(\tau)$ is 7, and the global minimum point of $\Delta\overline{S_1}\,(\tau)+\left|\overline{S_2}\,(\tau)\right|$ is 19. After calculation, the delay time $\tau=7$ and the embedding dimension $m=4$ are selected as the parameters of phase space reconstruction.



(a) $\Delta\overline{S_1}\ \tau$ diagram (b) $\Delta\overline{S_1}\ \tau+\left|\overline{S_2}\ \tau\right|$ diagram

FIGURE 8. $\Delta\overline{S_1}\,(\tau)$ and $\Delta\overline{S_1}\,(\tau)+\left|\overline{S_2}\,(\tau)\right|$ diagram.

V. EXPERIMENTAL VERIFICATION AND ANALYSIS

A. TEST VERIFICATION ANALYSIS

In order to verify effectiveness of the proposed PAPRKM, the experiments are carried out. The life cycle data of rolling bearings in Part IV is used for the test data. The test sample point is the state after the rolling bearing enters the degradation stage. The test results and evaluation index MAPE values are shown in Table 1, and the prediction effect is shown in FIGURE 9. The abscissa represents the sample number and the ordinate represents the test set output.

TABLE 1. Test results.

Model	MAPE	Time(s)
PAPRKM	0.0563	0.2260

From the Table 1, it can be seen that the value of the evaluation index MAPE is 0.0563, which is very small error. The experimental result shows the stability of PAPRKM model in the performance degradation evaluation of rolling bearings. From the perspective of model running time, the running time of using PAPRKM model to evaluate the performance degradation of rolling bearings is 0.2260s, and the running time is short.

In the FIGURE 9, it can be seen that the degradation performance index formed by multi-feature index after dimension reduction fusion and phase space reconstruction is input into the KELM model, and the prediction output can well reflect the performance degradation trend of rolling bearings. The experimental result shows that the degradation performance index constructed by phase space reconstruction technology can comprehensively reflect the performance degradation process of rolling bearings, which shows the superiority of phase space reconstruction in the construction of degradation performance index.

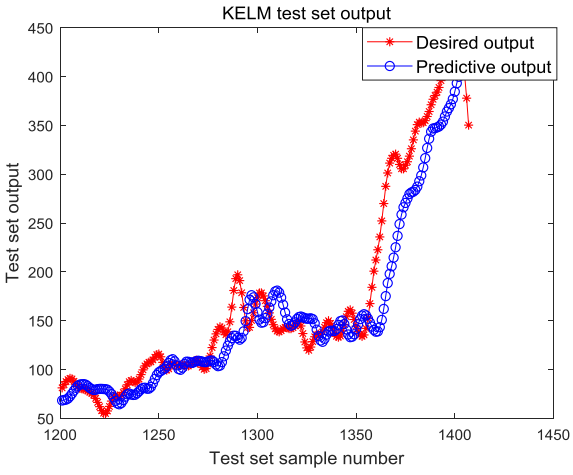


FIGURE 9. The performance degradation trend using the proposed PAPRKM.

B. COMPARATIVE ANALYSIS

In order to verify the superiority of the proposed PAPRKM, several different models are selected to test and compare.

(1) 29 time-domain and frequency-domain feature indexes are directly input into ELM and KELM models.

(2) PSR-ELM is based on single feature index, phase space reconstruction and ELM.

(3) PSR-KELM is based on single feature index, phase space reconstruction and KELM.

(4) PCA-ELM is based on multi-feature dimension reduction fusion and ELM.

(5) PCA-KELM is based on multi-feature dimension reduction fusion and KELM.

(6) PCA-PSR-ELM is based on multi-feature dimension reduction fusion, phase space reconstruction and ELM.

The comparison results of performance index for several different methods and running times of these methods are shown in Table 2. Among them, single feature index selects variance index and frequency-domain feature 1.

TABLE 2. Comparison test results of different evaluation models.

Model	MAPE	Time(s)
ELM	0.2874	0.2021
KELM	0.2470	0.2258
PSR-ELM (Time-domain)	0.3781	0.2024
PSR-ELM (Frequency-domain)	0.1885	0.2006
PSR-KELM (Time-domain)	0.2148	0.2259
PSR-KELM (Frequency-domain)	0.1251	0.2321
PCA-ELM	0.2363	0.2010
PCA-KELM	0.2214	0.2235
PCA-PSR-ELM	0.1632	0.2066
PAPRKM	0.0563	0.2260

As can be seen from the Table 2, from the perspective of each model running time, the running time for each model to evaluate the performance degradation of rolling bearings is between 0.20s and 0.24s, and the difference is not more than 0.04s.

For the performance degradation evaluation model of rolling bearings, the MAPE of KELM is 0.2470 and that of elm is 0.2874, which indicates that the performance degradation evaluation effect of KELM model is better than that of ELM model when 29 time-domain and frequency-domain indexes are directly input into ELM and KELM models. The MAPE of PCA-PSR-ELM is 0.1632, while the MAPE of PCA-ELM is 0.2363, which shows that the performance degradation evaluation effect of PCA-PSR-ELM model is better than that of PCA-ELM model under the same feature index. It is proved the superiority of phase space reconstruction theory in the construction of decay performance index; For PSR-ELM (Time-domain) model, its MAPE value is 0.3781, while that of PSR-ELM (Frequency-domain) model is 0.1885. The MAPE values of the two models are quite different, which indicates that in PSR-ELM model, single domain feature index cannot fully reflect the performance degradation trend of rolling bearings in the whole life cycle. However, the MAPE of PSR-KELM (Time-domain) model is 0.1251, and that of PSR-KELM (Frequency-domain) model is 0.2148, which indicates that PSR-KELM model is better than PSR-ELM model in performance degradation evaluation of rolling bearings under single domain feature index, and PSR-KELM model can better reflect the performance degradation trend of rolling bearings in the whole life cycle. Compared with single domain features, the multi-feature index should be able to reflect the performance change state of rolling bearings more comprehensively. The MAPE of PCA-KELM is 0.2214, which shows that the performance degradation evaluation effect of KELM is better than that of ELM under the same feature index. The MAPE of PAPRKM is 0.0563, which shows the performance degradation evaluation effect of PAPRKM is better than those of the compared models. The PAPRKM has smaller deviation and higher stability. The superiority of KELM and phase space reconstruction theory in the performance degradation evaluation of rolling bearings is verified. It shows that the PAPRKM can better reflect the performance degradation features of rolling bearings, and effectively improve the evaluation accuracy of the performance degradation, and more accurately reflect the performance degradation trend of rolling bearings.

The performance degradation trend effect of rolling bearings is shown in FIGURE 10∼FIGURE 18. The horizontal coordinate represents the sample number, and the vertical coordinates represents the output of test set.

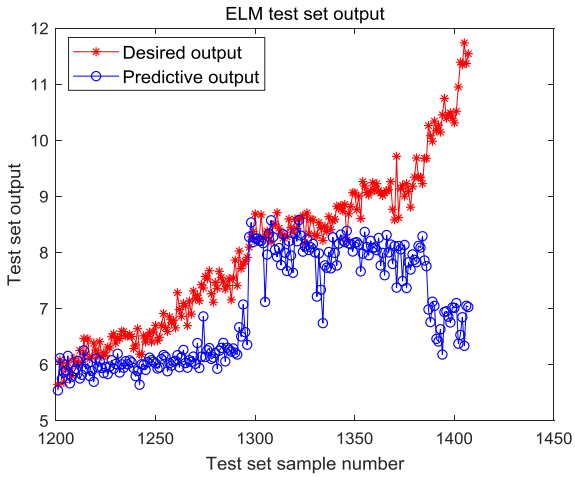


FIGURE 10. The performance degradation trend using ELM.

As can be seen from FIGURE 10 to FIGURE 18 that the performance degradation evaluation results of different models are different.

As can be seen from FIGURE 10, the ELM model has a large deviation in the prediction result, so it cannot effectively predict the trend and state of the performance degradation of rolling bearings. Test set sample number ranged from 1250 points to 1428 points, the predictive output could not accurately reflect the changing process, resulting in a large deviation in the experimental effect.

As can be seen from FIGURE 11, compared with ELM model, KELM model has better prediction results. Test set sample number ranged from 1200 points to 1428 points, although the predictive output can better reflect the performance degradation trend of rolling bearings, there is a large deviation between the predictive output and the desired output.

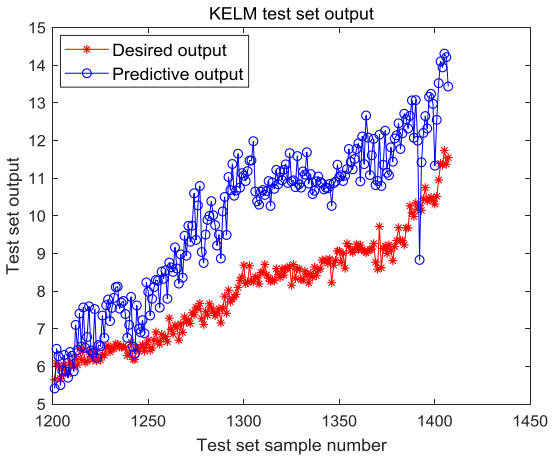


FIGURE 11. The performance degradation trend using KELM.

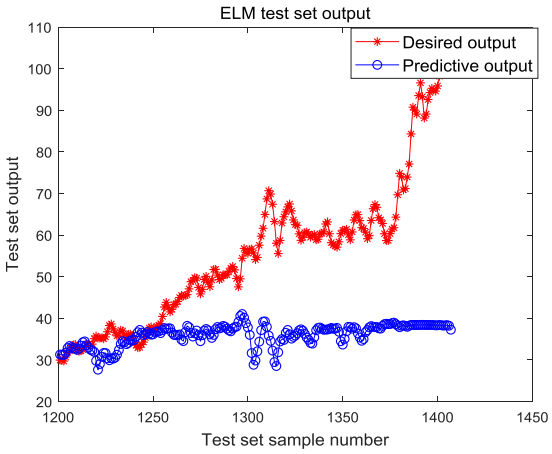


FIGURE 12. The performance degradation trend using PSR-ELM (Time-domain).

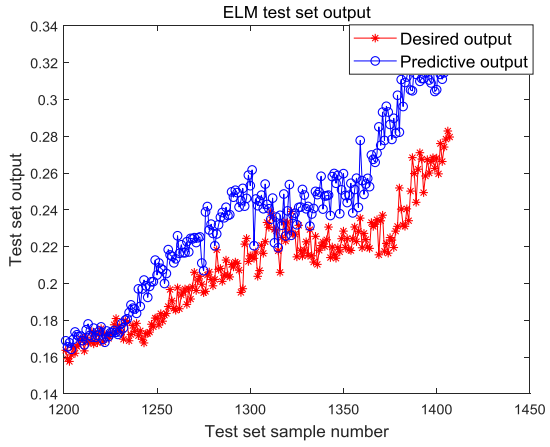


FIGURE 13. The performance degradation trend using PSR-ELM (Frequency-domain).

As can be seen from FIGURE 12, the prediction effect of PSR-ELM (Time-domain) model is not good that is impossible to predict the performance degradation of rolling bearings.

As can be seen from FIGURE 13, in the PSR-ELM (Frequency-domain) model, test set sample number ranged from 1200 points to 1428 points, although the predictive output can better reflect the performance degradation trend of rolling bearings, there is a certain deviation between the predictive output and the desired output.

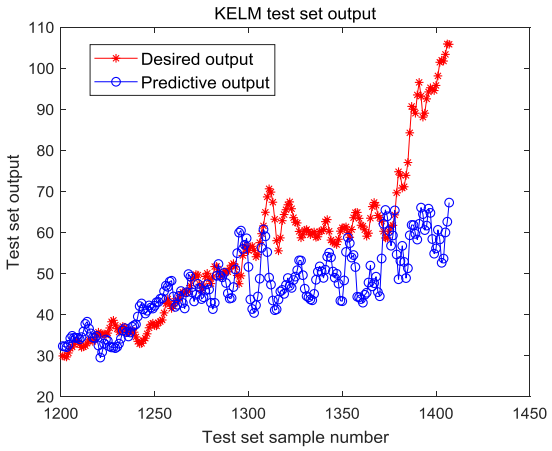


FIGURE 14. The performance degradation trend using PSR-KELM (Time-domain).

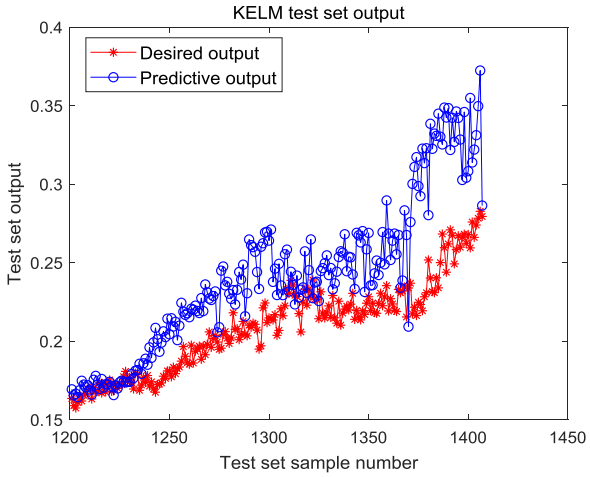


FIGURE 15. The performance degradation trend using PSR-KELM (Frequency-domain).

As can be seen from FIGURE 14, in the PSR-KELM (Time-domain) model, test set sample number ranged from 1200 points to 1300 points, the predictive output can better reflect the performance degradation trend of rolling bearings. Test set sample number ranged from 1301 points to 1428 points, the predictive output cannot reflect the performance changing process of rolling bearings effectively.

As can be seen from FIGURE 15, in the PSR-KELM (Frequency-domain) model, test set sample number ranged from 1200 points to 1428 points, although the predictive output can better follow the desired output trend and generally reflect the performance degradation trend of rolling bearings, there is a large deviation between the predictive output and the desired output.

FIGURE 12∼FIGURE 15 show the performance degradation trend of a single feature index for the ELM model after reconstruction. The experimental results show that the single domain feature index cannot fully reflect the whole life and

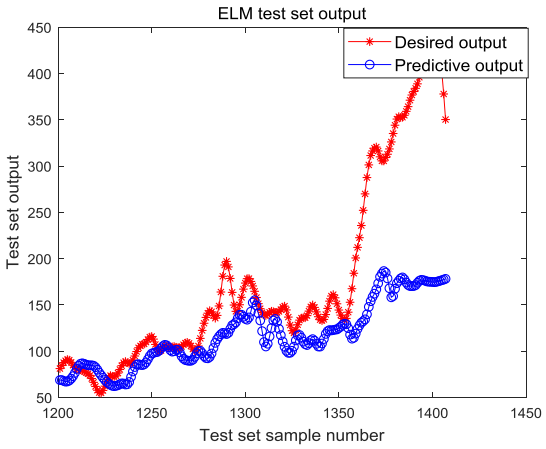


FIGURE 16. The performance degradation trend using PCA-ELM.

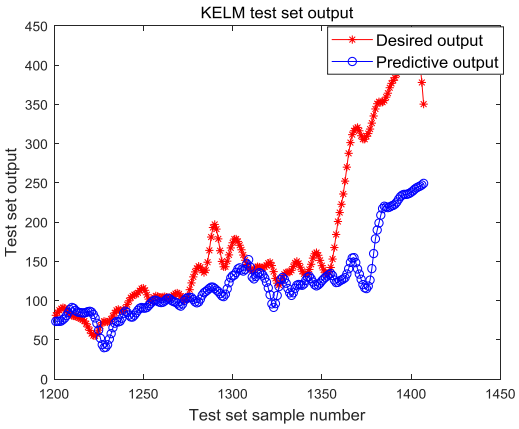


FIGURE 17. The performance degradation trend using PCA-KELM.

fate features of rolling bearings. Therefore, the multi-feature index is used to reflect the degradation features of rolling bearings.

As can be seen from FIGURE 16 to FIGURE 17, although the FIGURE 16 and FIGURE 17 can generally reflect the performance degradation trend of rolling bearings, the predictive output after the sample at 1350 cannot well reflect the performance degradation trend of rolling bearings.

As can be seen from FIGURE 18, the FIGURE 18 is more accurate than FIGURE 16 and FIGURE 17 in reflecting the performance degradation trend of rolling bearings. It shows that the feature index can better describe the performance degradation trend of rolling bearings after the phase space reconstruction, which reflects the accuracy of the phase space reconstruction in establishing the performance degradation index of rolling bearings and the superiority of the phase space reconstruction in the performance degradation evaluation of rolling bearings.

C. ANALYSIS OF FITTING DEGREE

The fitting degree between the predicted value and the actual value obtained by the proposed PAPRKM model is shown in

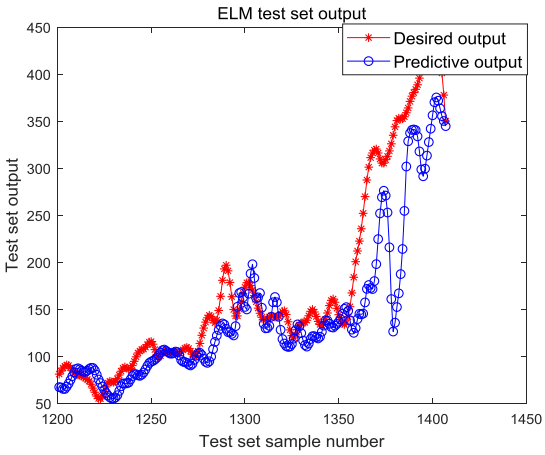


FIGURE 18. The performance degradation trend using PCA-PSR-ELM.

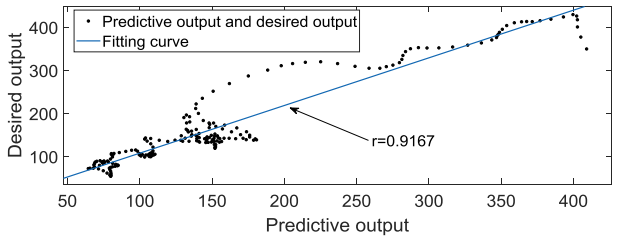


FIGURE 19. Fitting degree between predicted value and actual value.

FIGURE 19. The abscissa represents the actual value and the ordinate represents the predicted value.

As can be seen from the FIGURE 19 that r = 0.9167 is closed to 1, which indicates that the regression line has a better fitting degree for the used data. The experimental results prove that the PAPRKM method can effectively achieve the better expected effect for performance degradation evaluation of the rolling bearings.

VI. CONCLUSION

In this paper, in order to better characterize the performance degradation trend of rolling bearings, a new method of performance degradation evaluation for rolling bearings based on PCA, PSR and KELM, namely PAPRKM is proposed based on the analysis of feature indexes in time-domain and frequency-domain of the whole life cycle data of rolling bearings. Firstly, feature extraction in time-domain and frequency-domain is carried out for the whole life cycle data of rolling bearings. Then, PCA method is used to reduce the dimension of the extracted feature indexes to eliminate redundancy and information conflict. The first principal component is taken as the performance degradation index of rolling bearings. Due to the lack of information, the multi-dimensional feature matrix is constructed by using the phase space reconstruction technology to increase the dimension. More effective information is obtained from time series, and the reconstructed multi-dimensional feature matrix is input into KELM model as the performance degradation index of rolling bearings, and the trained model is used to evaluate the performance degradation of rolling

bearings. Through experimental verification and analysis, the MAPE value of PAPRKM is 0.0563 and the running time is 0.2260s. In addition, the prediction performance of ELM, KELM, PSR-ELM (Time-domain), PSR-ELM (Frequency-domain), PSR-KELM (Time-domain), PSR-KELM (Frequency-domain), PCA-ELM, PCA-KELM and PCA-PSR-ELM are compared. The MAPE values are 0.2874, 0.2470, 0.3781, 0.1885, 0.2148, 0.1251, 0.2363, 0.2214 and 0.1632 respectively. The running times are 0.2021s, 0.2258s, 0.2024s, 0.2006s, 0.2259s, 0.2321s, 0.2010s, 0.2235s and 0.2066s, respectively. The experimental results show that the prediction performance of PAPRKM is the best, which verifies the superiority and effectiveness of PAPRKM in the performance degradation evaluation of rolling bearings. Therefore, the performance degradation evaluation method proposed can effectively evaluate the performance degradation process of rolling bearings.

REFERENCES

[1] S. Haidong, C. Junsheng, J. Hongkai, Y. Yu, and W. Zhantao, "Enhanced deep gated recurrent unit and complex wavelet packet energy moment entropy for early fault prognosis of bearing," *Knowl.-Based Syst.*, vol. 188, Jan. 2020, Art. no. 105022.

[2] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *J. Sound Vibrat.*, vol. 289, nos. 4–5, pp. 1066–1090, Feb. 2006.

[3] S. Du, J. Lv, and L. Xi, "Degradation process prediction for rotational machinery based on hybrid intelligent model," *Robot. Comput.-Integr. Manuf.*, vol. 28, no. 2, pp. 190–207, Apr. 2012.

[4] F. Zhao, J. Chen, and W. Xu, "Condition prediction based on wavelet packet transform and least squares support vector machine methods," *Proc. Inst. Mech. Eng., E, J. Process Mech. Eng.*, vol. 223, no. 2, pp. 71–79, May 2009.

[5] H. Li *et al.*, "Prediction of rolling bearing degradation trend based on binary multiscale entropy," *China Mech. Eng.*, vol. 28, no. 20, pp. 2420–2425, 2017.

[6] L. WF, D. HM, and X. AQ, "New index in time-domain and application of PNN in fault diagnosis of rolling bearing," *Mech. Sci. Technol.*, vol. 35, no. 9, pp. 1382–1386, 2016.

[7] Y. Pan, J. Chen, and X. Li, "Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means," *Mech. Syst. Signal Process.*, vol. 24, no. 2, pp. 559–566, Feb. 2010.

[8] W. Deng, J. Xu, Y. Song, and H. Zhao, "An effective improved co-evolution ant colony optimization algorithm with multi-strategies and its application," *Int. J. Bio-Inspired Comput*, vol. 16, no. 3, pp. 158–170, 2020.

[9] R. Chen, S.-K. Guo, X.-Z. Wang, and T.-L. Zhang, "Fusion of multi-RSMOTE with fuzzy integral to classify bug reports with an imbalanced distribution," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 12, pp. 2406–2420, Dec. 2019.

[10] T. Jin, H. Xia, and H. Chen, "Optimal control problem of the uncertain second-order circuit based on first hitting criteria," *Math. Methods Appl. Sci.*, vol. 44, no. 1, pp. 882–900, Jan. 2021.

[11] W. Deng, J. Xu, Y. Song, and H. Zhao, "Differential evolution algorithm with wavelet basis function and optimal mutation strategy for complex optimization problem," *Appl. Soft Comput.*, to be published, doi: 10.1016/j.asoc.2020.106724.

[12] H. Zhao, J. Zheng, J. Xu, and W. Deng, "Fault diagnosis method based on principal component analysis and broad learning system," *IEEE Access*, vol. 7, pp. 99263–99272, 2019.

[13] Y. Xu, H. Chen, J. Luo, Q. Zhang, S. Jiao, and X. Zhang, "Enhanced moth-flame optimizer with mutation strategy for global optimization," *Inf. Sci.*, vol. 492, pp. 181–203, Aug. 2019.

[14] H. Shao, M. Xia, G. Han, Y. Zhang, and J. Wan, "Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer CNN and thermal images," *IEEE Trans. Ind. Informat.*, early access, Jun. 30, 2020, doi: 10.1109/TII.2020.3005965.

[15] Y. J. Song, D. Q. Wu, A. W. Mohamed, X. B. Zhou, B. Zhang, and W. Deng, "Enhanced success history adaptive DE for parameter optimization of photovoltaic models," *Complexity*, vol. 2021, 2021, Art. no. 6660115.

[16] Y. Xue, B. Xue, and M. Zhang, "Self-adaptive particle swarm optimization for large-scale feature selection in classification," *ACM Trans. Knowl. Discovery Data*, vol. 13, no. 5, pp. 1–27, Oct. 2019.

[17] H. Zhao, D. Li, W. Deng, and X. Yang, "Research on vibration suppression method of alternating current motor based on fractional order control strategy," *Proc. Inst. Mech. Eng., E, J. Process Mech. Eng.*, vol. 231, no. 4, pp. 786–799, Aug. 2017.

[18] T. Jin, X. Yang, H. Xia, and H. Ding, "Reliability index and option pricing formulas of the first hitting time model based on the uncertain fractional-order differential equation with caputo type," *Fractals*, 2020, doi: 10.1142/S0218348X21500122.

[19] W. Deng, J. Xu, X.-Z. Gao, and H. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Trans. Syst., Man, Cybern. Syst.*, early access, Nov. 4, 2020, doi: 10.1109/TSMC.2020.3030792.

[20] Z. He, H. Shao, X. Zhang, J. Cheng, and Y. Yang, "Improved deep transfer auto-encoder for fault diagnosis of gearbox under variable working conditions with small training samples," *IEEE Access*, vol. 7, pp. 115368–115377, 2019.

[21] H. Zhao, H. Liu, J. Xu, and W. Deng, "Performance prediction using high-order differential mathematical morphology gradient spectrum entropy and extreme learning machine," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4165–4172, Jul. 2020.

[22] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: A new method for stochastic optimization," *Future Gener. Comput. Syst.*, vol. 111, pp. 300–323, Oct. 2020.

[23] T. Jin, H. Ding, H. Xia, and J. Bao, "Reliability index and asian barrier option pricing formulas of the uncertain fractional first-hitting time model with caputo type," *Chaos, Solitons Fractals*, to be published, doi: 10.1016/j.chaos.2020.110409.

[24] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 7319–7327, Oct. 2020.

[25] Y. Liu, Y. Mu, K. Chen, Y. Li, and J. Guo, "Daily activity feature selection in smart homes based on pearson correlation coefficient," *Neural Process. Lett.*, vol. 51, no. 2, pp. 1771–1787, Apr. 2020.

[26] J. Zheng, Y. Yuan, L. Zou, W. Deng, C. Guo, and H. Zhao, "Study on a novel fault diagnosis method based on VMD and BLM," *Symmetry*, vol. 11, no. 6, p. 747, Jun. 2019.

[27] H. Zhao, J. Zheng, W. Deng, and Y. Song, "Semi-supervised broad learning system based on manifold regularization and broad network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 3, pp. 983–994, Mar. 2020.

[28] Z. Ren, R. Skjetne, A. S. Verma, Z. Jiang, Z. Gao, and K. H. Halse, "Active heave compensation of floating wind turbine installation using a catamaran construction vessel," *Mar. Struct.*, vol. 75, Jan. 2021, Art. no. 102868.

[29] W. Deng, J. Xu, H. Zhao, and Y. Song, "A novel gate resource allocation method using improved PSO-based QEA," *IEEE Trans. Intell. Transp. Syst.*, early access, Oct. 1, 2020, doi: 10.1109/TITS.2020.3025796.

[30] T. Jin, H. Ding, B. Li, H. Xia, and C. Xue, "Valuation of interest rate ceiling and floor based on the uncertain fractional differential equation in caputo sense," *J. Intell. Fuzzy Syst.*, to be published, doi: 10.3233/JIFS-201930.

[31] H. Zhao, S. Zuo, M. Hou, W. Liu, L. Yu, X. Yang, and W. Deng, "A novel adaptive signal processing method based on enhanced empirical wavelet transform technology," *Sensors*, vol. 18, no. 10, p. 3323, Oct. 2018.

[32] W. Deng, W. Li, and X.-H. Yang, "A novel hybrid optimization algorithm of computational intelligence techniques for highway passenger volume prediction," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4198–4205, Apr. 2011.

[33] Y. Song, D. Wu, W. Deng, X.-Z. Gao, T. Li, B. Zhang, and Y. Li, "MPPCEDE: Multi-population parallel co-evolutionary differential evolution for parameter optimization," *Energy Convers. Manage.*, vol. 228, Jan. 2021, Art. no. 113661.

[34] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2004, pp. 985–990.

[35] J. Yu, Y. Xu, G. Yu, and L. Liu, "Fault severity identification of roller bearings using flow graph and non-naive Bayesian inference," *Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci.*, vol. 233, no. 14, pp. 5161–5171, Jul. 2019.

[36] S. Lu, R. Yan, Y. Liu, and Q. Wang, "Tacholess speed estimation in order tracking: A review with application to rotating machine fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 7, pp. 2315–2332, Jul. 2019.

[37] Z. Wang, X. Ren, Z. Ji, W. Huang, and T. Wu, "A novel bio-heuristic computing algorithm to solve the capacitated vehicle routing problem based on Adleman–Lipton model," *Biosystems*, vol. 184, Oct. 2019, Art. no. 103997.
[38] B. Sreejith, A. K. Verma, and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks," in *Proc. IEEE Region 3rd Int. Conf. Ind. Inf. Syst.*, Dec. 2008, pp. 1–6.
[39] Q. Hu, A. Qin, Q. Zhang, J. He, and G. Sun, "Fault diagnosis based on weighted extreme learning machine with wavelet packet decomposition and KPCA," *IEEE Sensors J.*, vol. 18, no. 20, pp. 8472–8483, Oct. 2018.
[40] Y. Lei, Z. He, and Y. Zi, "A new approach to intelligent fault diagnosis of rotating machinery," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1593–1600, Nov. 2008.
[41] L. CY, X. MQ, and S. Guo, "Principal component analysis of acoustic signal applied to rolling bearing fault diagnosis," *Acoustic Technol.*, vol. 27, no. 2, pp. 271–275, 2008.
[42] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.
[43] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.
[44] G.-B. Huang, "An insight into extreme learning machines: Random neurons, random features and kernels," *Cognit. Comput.*, vol. 6, no. 3, pp. 376–390, Sep. 2014.
[45] Z. B. Lu, Z. M. Cai, and K. Y. Jiang, "Parameter selection of phase space reconstruction based on improved C-C method," *J. Syst. Simul.*, vol. 19, no. 11, pp. 2527–2529, 2007.
[46] S. Lu, J. Guo, Q. He, F. Liu, Y. Liu, and J. Zhao, "A novel contactless angular resampling method for motor bearing fault diagnosis under variable speed," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 11, pp. 2538–2550, Nov. 2016.
[47] V. T. Tran, F. AlThobiani, and A. Ball, "An approach to fault diagnosis of reciprocating compressor valves using teager–kaiser energy operator and deep belief networks," *Expert Syst. Appl.*, vol. 41, no. 9, pp. 4113–4122, Jul. 2014.
[48] T. Jin, Y. Sun, and Y. Zhu, "Time integral about solution of an uncertain fractional order differential equation and application to zero-coupon bond model," *Appl. Math. Comput.*, vol. 372, May 2020, Art. no. 124991.
[49] T. Li, Z. Qian, and T. He, "Short-term load forecasting with improved CEEMDAN and GWO-based multiple kernel ELM," *Complexity*, vol. 2020, pp. 1–20, Feb. 2020.
[50] L. Liao and F. Köttig, "A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction," *Appl. Soft Comput.*, vol. 44, pp. 191–199, Jul. 2016.
[51] X. Li, W. Zhang, and Q. Ding, "Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction," *Rel. Eng. Syst. Saf.*, vol. 182, pp. 208–218, Feb. 2019.
[52] Y. Lei, N. Li, and J. Lin, "A new method based on stochastic process models for machine remaining useful life prediction," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 12, pp. 2671–2684, Dec. 2016.
[53] J. Zhu, N. Chen, and W. Peng, "Estimation of bearing remaining useful life based on multiscale convolutional neural network," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3208–3216, Apr. 2019.
[54] Y. Qian, R. Yan, and S. Hu, "Bearing degradation evaluation using recurrence quantification analysis and Kalman filter," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 11, pp. 2599–2610, Nov. 2014.
[55] T. Jin and Y. Zhu, "First hitting time about solution for an uncertain fractional differential equation and application to an uncertain risk index model," *Chaos, Solitons Fractals*, vol. 137, Aug. 2020, Art. no. 109836.
[56] Y. Cheng, H. Zhu, K. Hu, J. Wu, X. Shao, and Y. Wang, "Multisensory data-driven health degradation monitoring of machining tools by generalized multiclass support vector machine," *IEEE Access*, vol. 7, pp. 47102–47113, 2019.

MINGYANG LV (Member, IEEE) was born in 1996. He received the B.S. degree in electrical engineering and automation from the North University of China, Taiyuan, China, in 2018. He is currently pursuing the master's degree with Dalian Jiaotong University, Dalian, China. His research interests include deep learning and fault diagnosis.

CHUNGUANG ZHANG (Member, IEEE) was born in 1978. He received the B.S. degree in electrical technology and the M.S. degree in traffic information control engineering from Dalian Jiaotong University, Dalian, China, in 2001 and 2005, respectively. Since 2014, he has been an Associate Professor with the School of Electronics and Information Engineering, Dalian Jiaotong University. His research interests include signal processing and fault diagnosis.

AIBIN GUO (Member, IEEE) was born in 1977. He received the B.S. degree in mechatronic engineering and the M.S. degree in mechanical engineering and automation from Dalian Jiaotong University, Dalian, China, in 2000 and 2003, respectively, and the Ph.D. degree in mechanical engineering and automation from Beihang University, Beijing, China, in 2008. Since 2008, he has been a Senior Engineer with Haifeng General Aviation Technology Company Ltd., Beijing. He is engaged in aviation emergency rescue and logistics support informatization.

FANG LIU (Member, IEEE) was born in 1976. She received the B.S. degree in computer application technology from Dalian Jiaotong University, Dalian, China, in 2000, and the Ph.D. degree in computer application technology from Zhejiang University, Hangzhou, China, in 2016. Since 2016, she has been a Lecturer with the Zhejiang University of Finance and Economics. Her research interests include artificial intelligence and information processing.