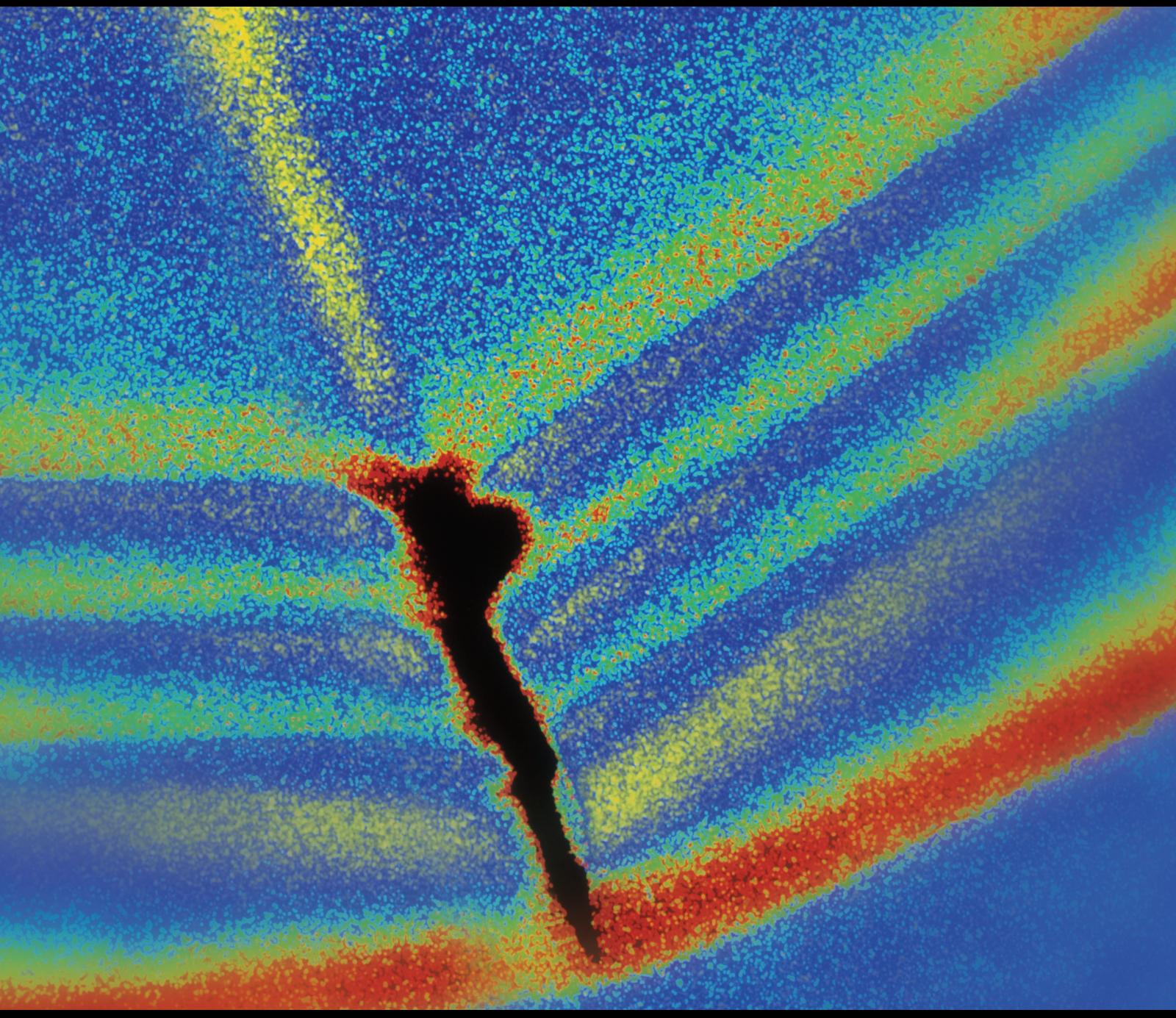
Shock and Vibration

Prognostics and Performance Evaluation for Mechanical Systems

Lead Guest Editor: Qing Li
Guest Editors: Yang Gao, Shulin Liu, Yonggang Xu, and Yanfeng Peng



Prognostics and Performance Evaluation for Mechanical Systems

Shock and Vibration

Prognostics and Performance Evaluation for Mechanical Systems

Lead Guest Editor: Qing Li
Guest Editors: Yang Gao, Shulin Liu, Yonggang Xu, and Yanfeng Peng

Contents

Internal Leakage Diagnosis of a Hydraulic Cylinder Based on Optimization DBN Using the CEEMDAN Technique
Peng Zhang and Xinyuan Chen
Research Article (10 pages), Article ID 8856835, Volume 2021 (2021)

Remaining Useful Life Prediction of High-Frequency Swing Self-Lubricating Liner
Xiuhong Hao, Shuqiang Wang, Mengfan Chen, and Deng Pan
Research Article (12 pages), Article ID 8843374, Volume 2021 (2021)

Tool Wear Assessment Approach Based on the Neighborhood Rough Set Model and Nearest Neighbor Model
Ren Sheng and Xiaoran Zhu
Research Article (15 pages), Article ID 8876187, Volume 2020 (2020)

Application of Feature Fusion Using Coaxial Vibration Signal for Diagnosis of Rolling Element Bearings
Jing Jiao, Jianhai Yue, Di Pei, and Zhunqing Hu
Research Article (14 pages), Article ID 8831723, Volume 2020 (2020)

Hindawi
Shock and Vibration
Volume 2021, Article ID 8856835, 10 pages
https://doi.org/10.1155/2021/8856835

Research Article

Internal Leakage Diagnosis of a Hydraulic Cylinder Based on Optimization DBN Using the CEEMDAN Technique

Peng Zhang [1,2] **and Xinyuan Chen** [1,2]

[1] *Key Laboratory of Metallurgical Equipment and Control Technology (Wuhan University of Science and Technology), Ministry of Education, Wuhan, China*
[2] *Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China*

Correspondence should be addressed to Xinyuan Chen; chenxinyuan@wust.edu.cn

Received 19 September 2020; Revised 30 January 2021; Accepted 6 March 2021; Published 13 March 2021

Academic Editor: Li Qing

Internal leakage diagnosis in a hydraulic cylinder is a key technique for the maintenance of hydraulic systems. However, it is difficult to diagnose the internal leakage under different low loads. To solve this problem, a novel fault diagnosis method based on the optimization deep belief network (DBN) combined with the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) technique is proposed to treat the collected AE signals. The raw AE signals are decomposed into a set of intrinsic mode functions (IMFs) by using CEEMDAN. Subsequently, according to the decreasing order of the Pearson correlation coefficient values, the first five IMFs are selected for signal reconstruction to suppress the abnormal interference from noise. The reconstructed signals are regarded as the input of the optimization DBN, and the particle swarm optimization simulated annealing (PSOSA) algorithm is adopted to identify the four internal leakage levels. The experimental results show that the proposed method exhibits a higher classification accuracy than other methods under different low loads. This result validates the effectiveness and superiority of the proposed approach to realize internal leakage diagnoses under different low loads.

1. Introduction

Hydraulic systems are widely applied in industry as actuators in hydraulic systems [1–3]. Internal leakage is a major failure mode caused by the wear of the piston seal that closes the gap between the movable piston and cylinder wall [4]. The detection of internal leakage is critical in practical applications because such leakage can affect the dynamic performance of the system. However, internal leakage cannot be detected until the actuator seal is completely damaged, and the actuator fails to respond to a control signal. Moreover, the detection of internal leakage is challenging due to the associated concealment. In addition, online nonintrusive diagnosis of internal leakage cannot be easily realized because of the dynamic working conditions of hydraulic cylinders. Hence, this paper proposes a method to realize automatic internal leakage diagnoses under different loads.

The existing methods for the fault diagnoses of internal leakage in hydraulic cylinders can be categorized into two types, namely, model-based and data-driven methods. Among model-based methods, a linear model-based approach with an adaptive threshold [5] and extended Kalman filter- (EKF-) based fault detection scheme [6] have been developed to detect different types of leakage in hydraulic actuators. However, due to the uncertainty and nonlinearity in hydraulic systems, it is difficult to establish a precise model. Consequently, researchers have adopted data-driven methods to realize fault diagnoses. Data-driven methods usually include two main steps: fault feature extraction and selection using signal processing techniques, and fault classification using classifiers. Many features based on pressure signals have been found to be sensitive to internal leakage, such as the root mean square (RMS) values of level two detailed coefficients obtained using the wavelet

transform [7–9], five features extracted through wavelet packet analysis [10], and the instantaneous amplitude of the first IMF obtained through the EMD technique [11]. However, to implement the abovementioned schemes, a baseline value must be predetermined, and this value cannot be automatically adjusted with the change in the working conditions. Hence, Li et al. proposed a featured temporal segmentation method and an AdaBoost-BP detector to automatically evaluate the internal leakage faults in hydraulic actuators [4]. Nevertheless, such traditional data-driven methods exhibit several notable deficiencies: (1) the features input to the classifiers are designed by diagnosticians depending on prior knowledge regarding the signal processing techniques and diagnostic expertise. (2) The adopted classifiers have shallow architectures, which limit the capacity of such classifiers to learn the complex nonlinear relationships in fault diagnosis scenarios [12]. These limitations can be overcome by using deep learning (DL) algorithms. The key property of deep learning models is the multiple layer structure. Using the multiple hidden layers stacked hierarchically, a deep learning model can realize the highly complicated transformation and abstraction of raw signals [13, 14]. Guo et al. adopted a convolutional neural network (CNN) that employed raw strain signals to output the internal leakage of the hydraulic cylinder [15]. However, the strain caused by microflows is extremely small under low pressure, that is, for pressure values lower than 7 MPa, and the abovementioned method cannot effectively diagnose the internal leakage in such cases. In other words, the existing methods cannot effectively diagnose internal leakage under different low loads.

In contrast to the internal leakage detection technologies based on the dynamic pressure or strain, the acoustic emission (AE) technology, as a nondestructive testing (NDT) technique with an extremely high sensitivity, can be used to detect weak internal leakage signals without hindering operations [16–20]. Furthermore, the AE technology exhibits several advantages over other nondestructive testing (NDT) techniques such as the vibration technology [21]. First, the defects are easier to detect because of the enhanced signal to noise ratio. Second, because this approach can detect mechanical abnormalities at high frequencies, it is highly sensitive compared to vibration techniques. In addition, the AE technology can be applied to all types of machinery, regardless of speed [22]. Nevertheless, owing to the use of a highly sensitive transducer, the measured raw signal, which often includes noise and irrelevant and redundant signals, must be subjected to noise removal to ensure reliable fault diagnosis. CEEMDAN-based denoising methods can self-adaptively decompose a signal into several monocomponents according to the local characteristic time scales of the signal, which can ensure a high decomposition performance in nonlinear and nonstationary processes [23]. Moreover, under different working loads of the hydraulic cylinder, the difference in the raw leakage signals of identical patterns increases, which hinders the fault diagnosis process. In this case, Pearson correlation coefficients can be used to account for the main characteristics that differentiate the fault modes while eliminating the difference caused by the different working conditions [24, 25]. Therefore, in this study, a reconstruction denoising strategy combining CEEMDAN with Pearson correlation coefficients is established to obtain the denoised signal. First, the sampled AE signals are decomposed into different IMFs by using CEEMDAN. Subsequently, the Pearson correlation coefficients between the raw signals and each IMF are calculated. The reconstructed data based on the IMFs with large Pearson correlation coefficients are obtained for further analysis.

The deep belief network (DBN) is a promising tool to realize fault characteristic mining and intelligent diagnosis in the case of massive data [19, 26–28]. In contrast to the CNN, the DBN avoids the process of transforming one-dimensional time series data. Moreover, the DBN does not rely on manual feature extraction and selection, and thus, it can automatically learn effective information from the raw signals. To reduce the influence of manual parameter setting on fault diagnosis accuracy, the hybrid particle swarm optimization simulated annealing (PSOSA) algorithm [29, 30] is adopted to identify the optimal structure of the trained DBN [31]. Moreover, experiments are conducted using a gas-hydraulic pressurized cylinder, which is easy to disassemble and can operate under low load conditions.

Overall, this paper proposes a method involving the optimization DBN in combination with the CEEMDAN technique. First, the CEEMDAN combined with the Pearson correlation coefficients is employed to eliminate noise from the raw AE signals for four different internal leakage levels. Subsequently, the denoised signals are directly fed to the optimization DBN by the PSOSA to identify the levels of internal leakage.

2. Theoretical Framework

2.1. CEEMDAN Theory. The process flow of the CEEMDAN algorithm can be summarized as follows [32, 33]:

Step 1: Add a white noise series to the investigated signal, that is, $x^{(i)} = x + \beta w^{(i)}$, where $i = 1, 2, 3, \ldots, I$. Decompose each $x^{(i)}$ by using the EMD to obtain the corresponding $d_1^{(i)}$. Next, calculate the first CEEMDAN mode $\overline{d_1} = (1/I)\sum_{i=1}^{I} d_1^{(i)}$ and first residue $r_1 = x - \overline{d_1}$.

Step 2: Decompose each $r_1 + \varepsilon_1 E_1(w^{(i)})$ by using the EMD, where $i = 1, 2, 3, \ldots, I$, E_k represents the k-th mode component of the given signal obtained through the EMD calculation. Define the second CEEMDAN mode as $\overline{d_2} = (1/I)\sum_{i=1}^{I} E_1(r_1 + \varepsilon_1 E_1(w^{(i)}))$. For $k = 2, \ldots, K$, calculate the kth residue, $r_k = r_{k-1} - \overline{d_k}$;

Step 3: Decompose each $r_k + \varepsilon_k E_k(w^{(i)})$ by using the EMD, where $i = 1, \ldots, I$. Calculate the $k+1$ CEEMDAN mode as $\overline{d_{k+1}} = (1/I)\sum_{i=1}^{I} E_1(r_k + \varepsilon_k(w^{(i)}))$.

2.2. DBN Principle. The DBN consists of one visible layer and three hidden layers, as shown in Figure 1. The learning process of DBN includes two stages: the unsupervised pretraining stage and the supervised fine-tuning stage. In the supervised fine-tuning stage, the labelled target data are input to the model, and the parameters are updated [34, 35].

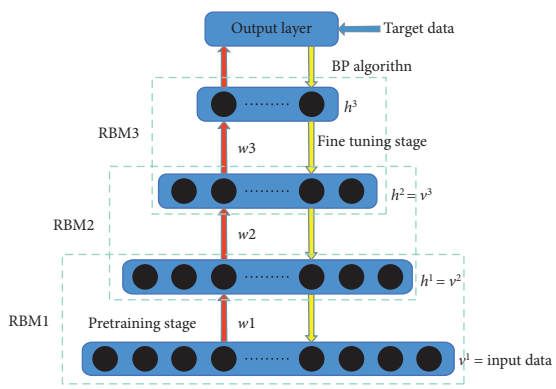


Figure 1: Architecture of the DBN.

Suppose that there exist I visible units and J hidden units in an RBM. The states of the visible layer and hidden layer are represented as v and h, respectively. The RBM energy function can be defined as follows:

$$E(v,h|\theta) = -\sum_{i=1}^{I} a_i v_i - \sum_{j=1}^{J} b_j h_j - \sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij} v_i h_j, \quad (1)$$

where $\theta = \{w_{ij}, b_j, a_i\}$; θ represents the model parameters; v_i is the visible unit i; h_j is the hidden unit j; w_{ij} is the weight connecting the visible unit i and hidden unit j; a_i is the offset of the visible unit i; and b_j is the offset of the hidden unit j.

The contrastive divergence (CD) algorithm is adopted to update the model parameters as follows:

$$\begin{aligned} \Delta w &= \varepsilon_w\left(\langle vh^T\rangle_{\text{data}} - \langle vh^T\rangle_{\text{recon}}\right), \\ \Delta a &= \varepsilon_a\left(\langle v\rangle_{\text{data}} - \langle v\rangle_{\text{recon}}\right), \\ \Delta b &= \varepsilon_a\left(\langle v\rangle_{\text{data}} - \langle v\rangle_{\text{recon}}\right), \\ w &\leftarrow w + m\Delta w, \\ a &\leftarrow a + m\Delta a, \\ b &\leftarrow b + m\Delta b, \end{aligned} \quad (2)$$

where $\varepsilon_w, \varepsilon_a, \varepsilon_b$ denote the learning rates of the weights between 0 and 1; a and b represent the biases of the hidden layer and visible layer, respectively; $\langle\cdot\rangle_{\text{data}}$ and $\langle\cdot\rangle_{\text{recon}}$ refer to the expectation derived using the training data distribution and reconstructed data, respectively.

2.3. Hybrid PSO-SA Optimization Method. The following steps are involved in the hybrid PSOSA algorithm [36–38]:

Step 1: Randomly initialize the parameters, including the position x_i^0, velocity v_i^0, and iterations. Calculate the fitness value of each particle $F(x_i)$ by using the fitness function defined in

$$F(x_i) = \frac{n_{\text{error}}}{N}, \quad (3)$$

where n_{error} refers to the number of misclassified data, and N refers to the total number of data. Next, update the best position of a particle $x_{ip\text{best}}^k$ and the best particle position of the swarm $x_{g\text{best}}^k$ according to the history. Initialize temperature $T = (-F(x_{g\text{best}}^k)/\ln(0.2))$.

Step 2: Update the positions and velocities of all the particles by using the following function:

$$\begin{aligned} v_i^{k+1} &= wv_i^k + c_1 r_1\left(x_{ip\text{best}}^k - x_i^k\right) + c_2 r_2\left(x_{g\text{best}}^k - x_i^k\right), \\ x_i^{k+1} &= x_i^k + v_i^{k+1}, \\ w &= w_{\max} - \frac{(w_{\max} - w_{\min})k}{T}, \\ c_1 &= c_{\max} - \frac{(c_{\max} - c_{\min})k}{T}, \\ c_2 &= c_{\max} + \frac{(c_{\max} - c_{\min})k}{T}, \end{aligned} \quad (4)$$

where w is the inertia coefficient, $w \in [0,1]$. c_1 and c_2 are cognitive parameters. r_1 and r_2 are random values in the interval [0, 1].

Step 3: Evaluate whether the new solution is accepted by using the metropolis acceptance criteria as follows:

$$\Delta f = F\left(x_i^{k+1}\right) - F\left(x_i^k\right); \quad \text{if } \Delta f < 0;\ x_i^k = x_i^{k+1}. \quad (5)$$

Otherwise, calculate the acceptance probability $p = \exp(-\Delta f/T)$.

$$\text{if } p > \text{rand}(0,1),\ x_i^k = x_i^{k+1}; \quad \text{else } x_i^k = x_i^k. \quad (6)$$

Step 4: Update the best position of a particle $x_{ip\text{best}}^k$ and the best particle position of the swarm $x_{g\text{best}}^k$ considering the history.

Step 5: If $F(x_{g\text{best}}^k) \le 0.01$, terminate the cycle, and output the best solution and objective. Otherwise, gradually decrease the temperature $T = T * 0.95$, and return to step 3.

3. Experimental Setup

3.1. Test Rig. As shown in Figure 2(a), the test rig consists of a gas-hydraulic pressurized cylinder 1, Jack 2, and pneumatic device 3.

The operating principle of the gas-hydraulic pressurized cylinder is shown in Figure 3. The red, green, blue, purple, and black regions correspond to the high-pressure air, low-pressure air, pressure-free hydraulic oil, preloading hydraulic oil, and high-pressure hydraulic oil, respectively. AE sensor 3 with an electromagnet is fixed on the mounting plane. The piston of rod 2 uses a y-sealing ring. Notches with different sizes are set on the sealing rings for a load of 1 MPa and 2 MPa to simulate four internal leakage levels corresponding to no internal leakage, small internal leakage, medium internal leakage, and large internal leakage. The seal rings with different size notches under 1 MPa are illustrated in Figure 2(b). The leaking oil flows out through the P1 port, and the internal leakage rate of the flow is determined using

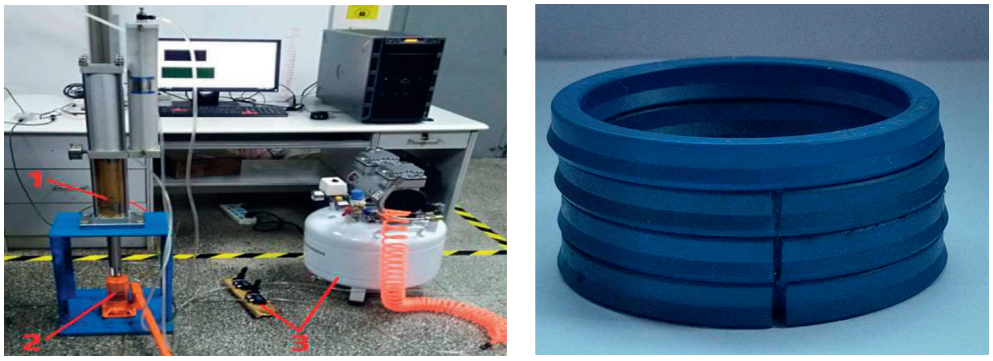


Figure 2: (a) Image of the test rig. (b) Image of the seal rings with notches of different sizes.

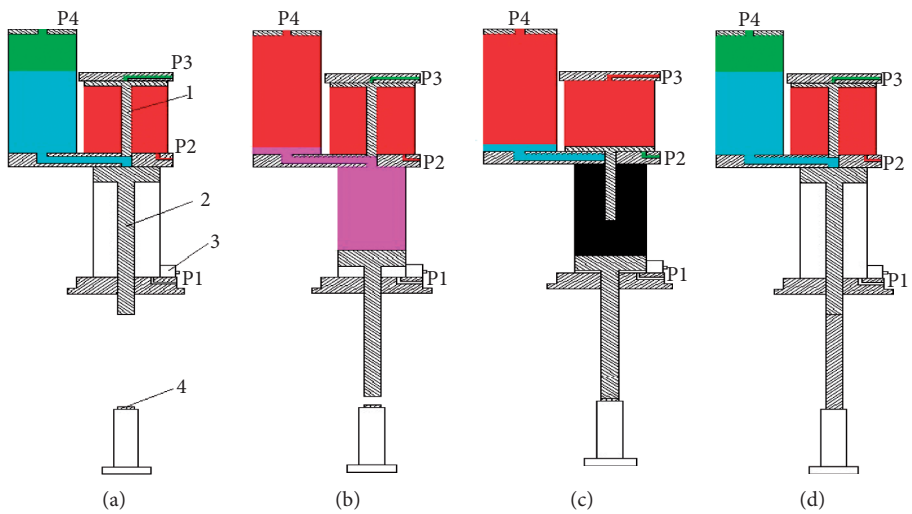


Figure 3: Schematic of the gas-hydraulic pressurized cylinder.

a measuring cup and stopwatch. The operating principle of the gas-hydraulic pressurized cylinder can be defined as follows:

(1) As shown in Figure 3(a), the P2 port connects the high-pressure air, the P3 and P4 ports connect the low-pressure air, the P1 port is the drain port, rods 1 and 2 retract, and the hydraulic oil is located in the oil tank.

(2) As shown in Figure 3(b), the P4 port connects the high-pressure air and preloading hydraulic oil. The hydraulic oil enters into nonrod chamber of the cylinder, and rod 2 extends.

(3) As shown in Figure 3(c), the P3 port connects the high-pressure air, and P2 port connects the low-pressure air. Rod 1 extends, thereby pressurizing the hydraulic oil. The leaking hydraulic oil flows out through the P1 port.

(4) As shown in Figure 3(d), the P2 port connects the high-pressure air, P3 port connects the low-pressure air, rod 1 retracts, and pressure relief occurs. Rod 2 is compressed by Jack 4, and the hydraulic oil flows back to the oil tank.

3.2. AE Sensor and Data Acquisition Device. As shown in Figure 4(a), AE sensor 1 is the PXR04 device manufactured by Changsha Pengxiang Electronic Technology Co. LTD. This device has a frequency range and resonant frequency of 30–140 kHz and 40 kHz, respectively. Amplifier 2 shown in Figure 4(b) has a fixed gain of 100 dB. Data acquisition card 3 pertains to NI-6340 manufactured by National Instruments; this device has a maximum acquisition frequency and output voltage amplitude of 400 kHz and 10 V, respectively.

4. Internal Leakage Diagnosis Method Based on the Optimization DBN

The proposed diagnosis method based on the optimization DBN can be described as follows:

Step 1: Define four internal leakage levels of the hydraulic cylinder under two loads (1 and 2 MPa). Collect raw AE signals and decompose the signals into a set of IMFs through the CEEMDAN.

Step 2: Calculate the Pearson correlation coefficient between each IMF and the raw AE signals. Next, select the first five sensitive IMFs to obtain the reconstructed

Figure 4: (a) Image of the AE sensor. (b) Image of the data acquisition device.

signals by considering the decreasing correlation coefficient values.

Step 3: Divide the reconstructed signals into training and testing samples. Train the optimization DBN by using the PSOSA algorithm and training samples. Use this network to classify the internal leakage levels by considering the testing samples. These steps are implemented to apply the proposed method to diagnose four internal leakage levels, and the results are recorded.

4.1. Signal Reconstruction and Data Description. According to the Shannon sampling theorem, the sampling frequency of the data acquisition card is set as 400 kHz. The AE signals in different internal leakage levels under a load of 1 and 2 MPa are shown in Figures 5 and 6, respectively.

The original signals are preprocessed using CEEMDAN. The Pearson correlation coefficient represents the correlation between each IMF and the original signal. The Pearson correlation coefficient between $\mathrm{IMF}_k(t)$ and the original signal data sequence $S(t)$ is calculated as follows [24]:

$$R_k = \frac{\sum_{t=1}^{T}\left(\mathrm{IMF}_k(t) - \overline{\mathrm{IMF}_k(t)}\right)\left(S(t) - \overline{S(t)}\right)}{\sqrt{\sum_{t=1}^{T}\left(\mathrm{IMF}_k(t) - \overline{\mathrm{IMF}_k(t)}\right)^2\left(S(t) - \overline{S(t)}\right)^2}}, \quad (7)$$

where T is the number of points in the original signal data sequence; $\overline{S(t)}$ and $\overline{\mathrm{IMF}_k(t)}$ denote the mean values of $S(t)$ and $\mathrm{IMF}_k(t)$, respectively.

To eliminate the noise and enhance the signal to noise ratio, the first five IMFs are used for signal reconstruction. The reconstruction process can be expressed as

$$\widehat{S}(t) = \mathrm{IMF}_a + \mathrm{IMF}_b + \mathrm{IMF}_c + \mathrm{IMF}_d + \mathrm{IMF}_e, \quad (8)$$

where $\widehat{S}(t)$ denotes the reconstructed signal; IMF_a, IMF_b, IMF_c, IMF_d, and IMF_e are the first five IMFs.

In this case study, datasets A and B are established to verify the performance of the proposed method. Datasets A and B include the data for the four conditions under loads of 1 and 2 MPa, respectively. Both datasets contain 100 training samples (each sample has 2000 data points) and 50 testing samples for each condition. The sample distributions for the four conditions in datasets A and B are presented in Table 1.

4.2. Comparison Approach. In this case study, PSOSA is applied to develop the optimization DBN based on the training samples. The parameter settings for the hybrid PSOSA are presented in Table 2. Table 3 lists the values of the parameters for the optimization DBN derived using the PSOSA. The optimal architecture of the DBN is 2000-860-284-78-4. The input and output layers consist of 2000 and 4 neurons, respectively, which are set based on the dimensions of the input data and fault patterns. The learning rate and momentum obtained using the PSOSA are 0.48 and 0.23, respectively.

To illustrate the effectiveness of the proposed method in actual applications of the fault diagnosis for the internal leakage, comparative experiments are conducted considering the following six methods:

(1) DBN method: The original AE signals are used directly as the input of the optimization DBN.

(2) EEMD + DBN: The reconstructed AE signals based on the EEMD method are used as the input of the optimization DBN.

(3) CEEMDAN + DBN: The reconstructed AE signals based on the CEEMDAN method are used as the input of the optimization DBN.

(4) CEEMDAN + BP: The reconstructed AE signals based on the CEEMDAN are used as the input of the BP network. To enable a fair comparison, the BP network and optimization DBN used the same data, architecture, learning rate, momentum, batch size, and epochs.

(5) T + BP: the original AE signals are preprocessed to extract the time domain features including the mean value, root mean square, peak, skewness, kurtosis value, crest factor, clearness factor, shape factor, and impulse factor. Next, the 9 selected features are used as the input of the BP network with 1 hidden layer. The architecture of the BPNN is 9-78-4.

(6) $T + F$ + BP: The original AE signals are preprocessed to extract the time and frequency domain features. The 9 time domain features are the same as those for method 5. The 4 frequency domain features include the mean frequency, frequency centre, root mean frequency, and standard deviation frequency. Next, the 13 selected features are used as the input of the

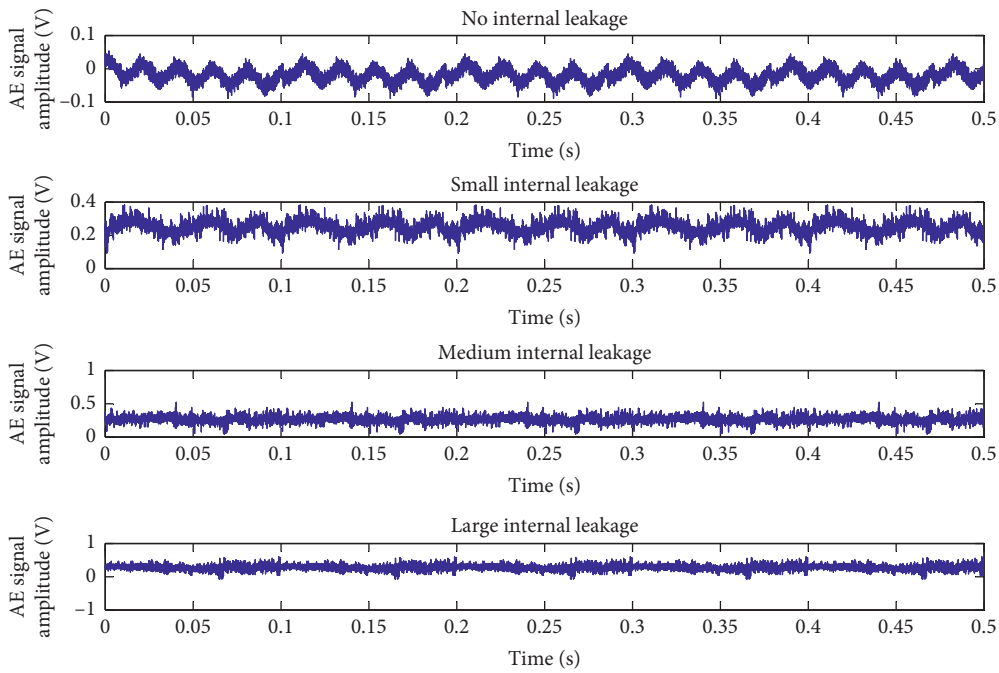


Figure 5: AE signals under a load of 1 MPa in four different internal leakage states.

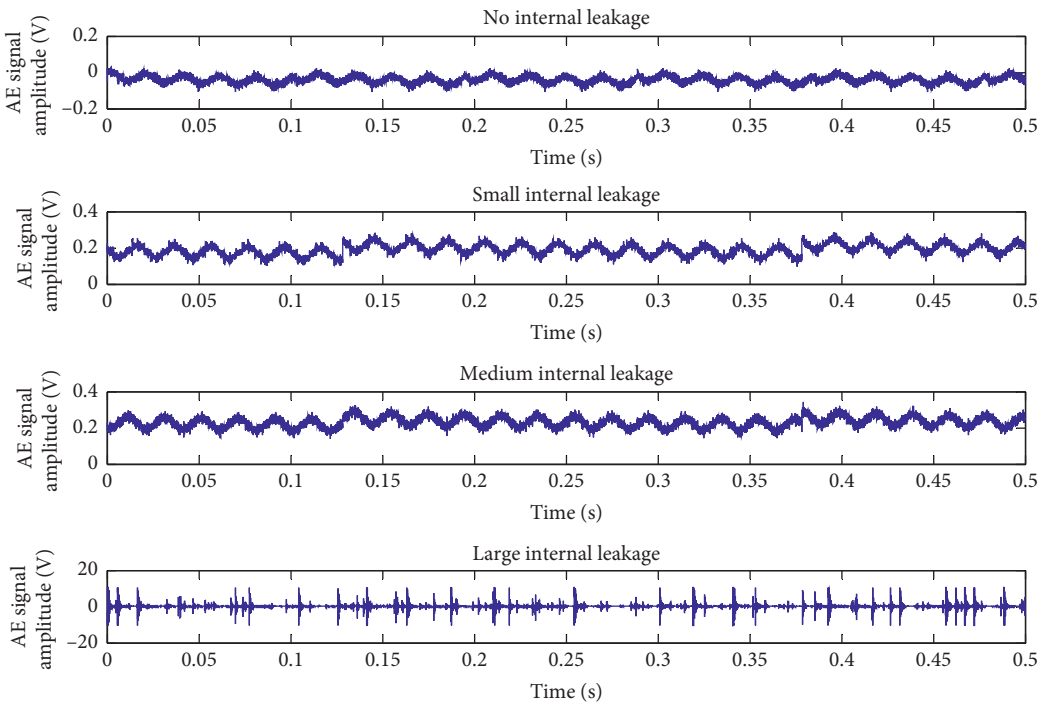


Figure 6: AE signals under a load of 2 MPa in four different internal leakage states.

Table 1: Sample distributions under four different conditions.

Datasets	Load (MPa)	Number of training samples/testing samples	Fault type	Label
A	1	100/50	No internal leakage	1
		100/50	Small internal leakage	2
		100/50	Medium internal leakage	3
		100/50	Large internal leakage	4
B	2	100/50	No internal leakage	1
		100/50	Small internal leakage	2
		100/50	Medium internal leakage	3
		100/50	Large internal leakage	4

Table 2: Parameter settings for the hybrid PSO-SA.

Parameters	Value
Population size of PSO	20
Number of iterations of PSO	50
Initial temperature T	100
Annealing factor	0.95
Max/min position of a particle	500/10
Max/min speed of a particle	1/−1
Max/min weight of PSO	1.5/0.5
Max/min acceleration factor of PSO	1.5/0.5

Table 3: Optimization parameters of the DBN.

Parameters	Value
Number of nodes in the input layer	2000
Number of nodes in the output layer	4
Number of RBMs	3
Number of iterations for each RBM	200
Number of nodes in the first hidden layer (obtained using the PSO-SA)	860
Number of nodes in the second hidden layer (obtained using the PSO-SA)	284
Number of nodes in the third hidden layer (obtained using the PSO-SA)	78
Learning rate of the deep belief network (obtained using the PSO-SA)	0.48
Momentum of the deep belief network (obtained using the PSO-SA)	0.23

BP network with 1 hidden layer. The architecture of the BPNN is 13-78-4.

The six methods are run 20 times, and the average classification accuracy is determined, as listed in Table 4.

The results show that the proposed method exhibits a considerably higher classification accuracy than that of the traditional methods based on the BP network. The BP network has the same architecture as the DBN, and the same reconstructed AE signals without any feature extraction are used to train the DBN and BP network in methods 3 and 4. The comparison of the diagnosis accuracy of methods 3 and 4 indicates that the accuracy of the DBN is considerably higher than that of the BP network when managing massive data. This phenomenon occurs because deep learning methods can learn the valuable information from the input data through multiple feature transformations. According to the comparison of the diagnosis accuracy of methods 5 and 6, different features yield different results, which indicates that the realization of the accurate classification of BP network requires manual intervention to select the proper features to implement the fault diagnosis.

In addition, the DBN combined with the CEEMDAN to realize the denoising in method 3 exhibits a higher accuracy than that of the DBN based on the original AE signals (method 1). This finding demonstrates that the proposed method can effectively denoise signals and enhance the robustness of the fault diagnosis results. Furthermore, the CEEMD + DBN outperforms the EEMD + DBN, and in contrast to the EEMD, the CEEMDAN can decrease the reconstruction error and avoid the problem of different numbers of modes for different realizations of the signal plus noise. Furthermore, the proposed method can achieve a higher classification accuracy than that of the other methods in datasets A and B. The results demonstrate that the proposed method can effectively diagnose the fault types even under load fluctuations.

5. Results and Discussion

The Pearson correlation coefficients are used to select the sensitive IMFs for further processing. However, no prior knowledge is available to set the number of sensitive IMFs for further reconstruction. Therefore, an investigation is conducted to determine the number of sensitive IMFs. According to the decreasing correlation coefficient values, the number of sensitive IMFs is increased sequentially from 1 to 12. The procedures are repeated 20 times, and the average accuracy for the testing samples of datasets A and B is separately determined.

The average classification accuracy varies with the different numbers of selected IMFs, as shown in Figure 7. Initially, the accuracy increases with the number of selected IMFs. However, the accuracy decreases as the number of selected IMFs increases to 12. The result indicates that, in the initial stage, as the number increases from 1 to 5, the selected IMFs include considerable fault-related information. As the number increases from 5 to 12, the selected IMFs contain an excessive amount of fault-unrelated information or a high degree of overlap. As the number of selected IMFs increases, the accuracy exhibits the opposite trend. Thus, the first five sensitive IMFs with a large correlation coefficient are considered for subsequent reconstruction in this paper.

To observe the clustering result of the proposed method, the principal component analysis (PCA) is performed on the input and output data of the third hidden layer. A certain confusion occurs among the four conditions in the input data, as shown in Figures 8(a) and 9(a). After the 3 hidden layers of the DBN, as shown in Figures 8(b) and 9(b), most

Table 4: Accuracy comparison for the six methods.

Methods	Accuracy for dataset A (%) Training/testing samples	Accuracy for dataset B (%) Training/testing samples
DBN	72.8/71.9	73.4/72.4
EEMD + DBN	86.8/84.7	86.9/85.1
CEEMDAN + DBN (the proposed method)	93.0/92.1	93.6/92.5
CEEMDAN + BP	55.6/54.2	58.2/58.0
T + BP	40.1/39.6	40.6/39.8
$T + F$ + BP	61.8/61.5	63.3/62.4

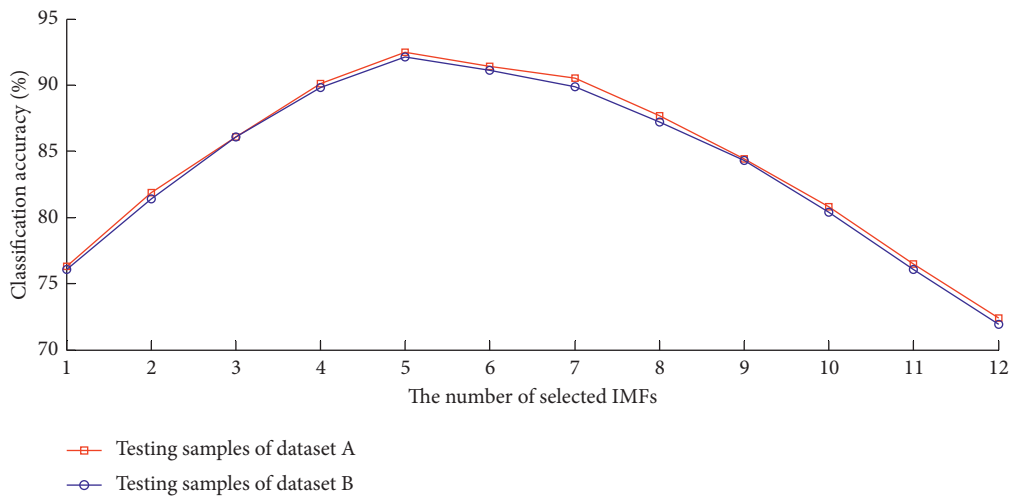


Figure 7: Average classification accuracy under different numbers of selected IMFs.

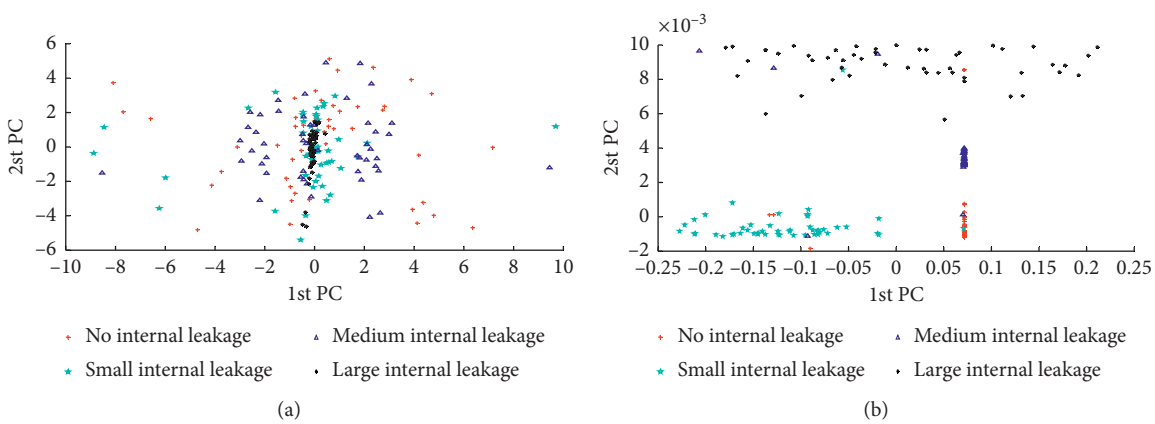


Figure 8: 2D-PCA projection of the testing samples of dataset A. (a) Input data. (b) Third hidden layer.

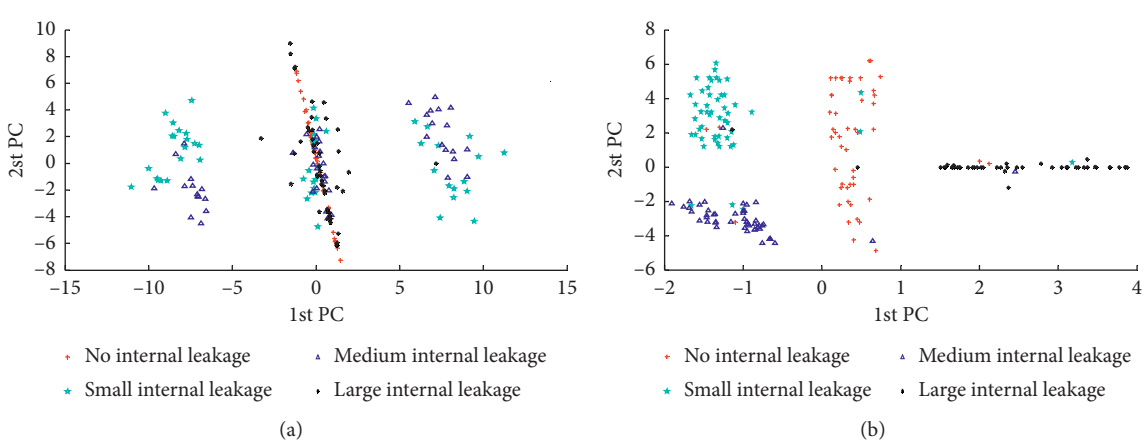


Figure 9: 2D-PCA projection of the testing samples of dataset B. (a) Input data. (b) Third hidden layer.

samples are correctly assigned to the corresponding cluster. The results indicate the effectiveness of the proposed method in mining the fault characteristics and classifying internal leakage conditions.

6. Conclusions

This paper proposes a fusion method that combines the optimization DBN with the CEEMDAN to realize the internal leakage diagnosis of hydraulic cylinders under different low loads. The main contributions of the paper are as follows: (1) first, the AE signals are examined to be used for the hydraulic cylinder internal leakage diagnosis. Based on the AE fault signals, the CEEMDAN-based denoising method and optimal number of representative IMFs for signal reconstruction are investigated. (2) The optimization DBN is obtained using the hybrid PSOSA algorithm instead of subjective human experience. The performance of the proposed method under identical loads is superior to those of the existing state of the art methods. (3) The proposed method can effectively diagnose faults under different low loads, which cannot be easily realized through the existing methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China, grant number (51975425).

References

[1] P. Ranjan, G. Wrat, M. Bhola, S. K. Mishra, and J. Das, "A novel approach for the energy recovery and position control of a hybrid hydraulic excavator," *ISA Transactions*, vol. 99, pp. 387–402, 2020.

[2] J. Koivumäki, W.-H. Zhu, and J. Mattila, "Energy-efficient and high-precision control of hydraulic robots," *Control Engineering Practice*, vol. 85, pp. 176–193, 2019.

[3] J. Zhao, G. Shen, C. Yang, W. Zhu, and J. Yao, "A robust force feed-forward observer for an electro-hydraulic control loading system in flight simulators," *ISA Transactions*, vol. 89, pp. 198–217, 2019.

[4] L. Li, Y. Huang, J. Tao, C. Liu, and K. Li, "Featured temporal segmentation method and adaBoost-BP detector for internal leakage evaluation of a hydraulic cylinder," *Measurement*, vol. 130, pp. 279–289, 2018.

[5] Z. Shi, F. Gu, B. Lennox, and A. D. Ball, "The development of an adaptive threshold for model-based fault detection of a nonlinear electro-hydraulic system," *Control Engineering Practice*, vol. 13, no. 11, pp. 1357–1367, 2005.

[6] L. An and N. Sepehri, "Hydraulic actuator leakage fault detection using extended Kalman filter," *International Journal of Fluid Power*, vol. 6, no. 1, pp. 41–51, 2005.

[7] A. Y. Goharrizi and N. Sepehri, "A wavelet-based approach to internal seal damage diagnosis in hydraulic actuators," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 5, pp. 1755–1763, 2009.

[8] A. Y. Goharrizi, N. Sepehri, and Y. Wu, "A wavelet-based approach for diagnosis of internal leakage in hydraulic actuators using on-line measurements," *International Journal of Fluid Power*, vol. 11, no. 1, pp. 61–69, 2010.

[9] A. Y. Goharrizi and N. Sepehri, "A wavelet-based approach for external leakage detection and isolation from internal leakage in valve-controlled hydraulic actuators," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4374–4384, 2010.

[10] X. Zhao, S. Zhang, C. Zhou, Z. Hu, R. Li, and J. Jiang, "Experimental study of hydraulic cylinder leakage and fault feature extraction based on wavelet packet analysis," *Computers & Fluids*, vol. 106, pp. 33–40, 2015.

[11] A. Y. Goharrizi and N. Sepehri, "Internal leakage detection in hydraulic actuators using empirical mode decomposition and Hilbert spectrum," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 2, pp. 368–378, 2011.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[13] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 303–315, 2016.

[14] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019.

[15] Y. Guo, Y. Zeng, L. Fu, and X. Chen, "Modeling and experimental study for online measurement of hydraulic cylinder micro leakage based on convolutional neural network," *Sensors*, vol. 19, no. 9, 2019.

[16] W. Kaewwaewnoi, A. Prateepasen, and P. Kaewtrakulpong, "Investigation of the relationship between internal fluid leakage through a valve and the acoustic emission generated from the leakage," *Measurement*, vol. 43, no. 2, pp. 274–282, 2010.

[17] H. Y. Sim, R. Ramli, S. Ahmad, and F. S. Ming, "Detection and estimation of valve leakage losses in reciprocating compressor using acoustic emission technique," *Measurement*, vol. 152, pp. 1–15, 2020.

[18] B. Ahn, J. Kim, and B. Choi, "Artificial intelligence-based machine learning considering flow and temperature of the pipeline for leak early detection using acoustic emission," *Engineering Fracture Mechanics*, vol. 210, pp. 381–392, 2019.

[19] X. Lang, Z. Hu, P. Li, Y. Li, J. Cao, and H. Ren, "Pipeline leak aperture recognition based on wavelet packet analysis and a deep belief network with ICR," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6934825, 8 pages, 2018.

[20] G.-Y. Ye, K.-J. Xu, and W.-K. Wu, "Standard deviation based acoustic emission signal analysis for detecting valve internal leakage," *Sensors and Actuators A: Physical*, vol. 283, pp. 340–347, 2018.

[21] G. Thompson and G. Zolkiewski, "An experimental investigation into the detection of internal leakage of gases through valves by vibration analysis," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 211, no. 3, pp. 195–207, 1997.

[22] S. A. Mirhadizadeh, E. P. Moncholi, and D. Mba, "Influence of operational variables in a hydrodynamic bearing on the generation of acoustic emission," *Tribology International*, vol. 43, no. 9, pp. 1760–1767, 2010.

[23] L. Wang and Y. Shao, "Fault feature extraction of rotating machinery using a reweighted complete ensemble empirical mode decomposition with adaptive noise and demodulation analysis," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106545, 2020.

[24] B.-S. Peng, H. Xia, Y.-K. Liu, B. Yang, D. Guo, and S.-M. Zhu, "Research on intelligent fault diagnosis method for nuclear power plant based on correlation analysis and deep belief network," *Progress in Nuclear Energy*, vol. 108, pp. 419–427, 2018.

[25] Y. Zhang, J. Ji, and B. Ma, "Fault diagnosis of reciprocating compressor using a novel ensemble empirical mode decomposition-convolutional deep belief network," *Measurement*, vol. 156, Article ID 107619, 2020.

[26] Y. Zou, Y. Zhang, and H. Mao, "Fault diagnosis on the bearing of traction motor in high-speed trains based on deep learning," *Alexandria Engineering Journal*, vol. 90, Article ID 102085, 2020.

[27] G. Zhao, X. Liu, B. Zhang, Y. Liu, G. Niu, and C. Hu, "A novel approach for analog circuit fault diagnosis based on deep belief network," *Measurement*, vol. 121, pp. 170–178, 2018.

[28] J. Gai, J. Shen, Y. Hu, and H. Wang, "An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing," *Measurement*, vol. 162, Article ID 107901, 2020.

[29] S. Du, J. Lv, and L. Xi, "Degradation process prediction for rotational machinery based on hybrid intelligent model," *Robotics and Computer-Integrated Manufacturing*, vol. 28, no. 2, pp. 190–207, 2012.

[30] Y. Wang, C. Zhang, and M. Li, "Application of a hybrid algorithm-PSOSA in well test parameter estimation," *Petroleum*, vol. 4, no. 4, pp. 430–436, 2018.

[31] H. Shao, H. Jiang, X. Zhang, and M. Niu, "Rolling bearing fault diagnosis using an optimization deep belief network," *Measurement Science and Technology*, vol. 26, no. 11, Article ID 115002, 2015.

[32] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: a suitable tool for biomedical signal processing," *Biomedical Signal Processing and Control*, vol. 14, no. 14, pp. 19–29, 2014.

[33] C. He, P. Niu, R. Yang, C. Wang, Z. Li, and H. Li, "Incipient rolling element bearing weak fault feature extraction based on adaptive second-order stochastic resonance incorporated by mode decomposition," *Measurement*, vol. 145, pp. 687–701, 2019.

[34] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," *Neural Networks: Tricks of the Trade*, pp. 599–619, Springer, Berlin, Germany, 2012.

[35] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," *Reliability Engineering & System Safety*, vol. 115, pp. 124–135, 2013.

[36] L.-l. Li, L. Wang, and L.-h. Liu, "An effective hybrid PSOSA strategy for optimization and its application to parameter estimation," *Applied Mathematics and Computation*, vol. 179, no. 1, pp. 135–146, 2006.

[37] Y. Qi, C. Li, P. Jiang, C. Jia, Y. Liu, and Q. Zhang, "Research on demodulation of FBGs sensor network based on PSO-SA algorithm," *Optik*, vol. 164, pp. 647–653, 2018.

[38] K. Thangavel, J. Bagyamani, and R. Rathipriya, "Novel hybrid PSO-SA model for biclustering of expression data," *Procedia Engineering*, vol. 30, pp. 1048–1055, 2012.

Hindawi
Shock and Vibration
Volume 2021, Article ID 8843374, 12 pages
https://doi.org/10.1155/2021/8843374

Research Article

Remaining Useful Life Prediction of High-Frequency Swing Self-Lubricating Liner

Xiuhong Hao,[1,2] **Shuqiang Wang**,[1,2] **Mengfan Chen**,[1,2] **and Deng Pan**[1,2]

[1]*School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China*
[2]*Key Laboratory of Self-Lubricating Spherical Plain Bearing Technology of Hebei Province, Yanshan University, Qinhuangdao 066004, China*

Correspondence should be addressed to Deng Pan; pandeng1896@sina.com

Received 8 August 2020; Revised 6 December 2020; Accepted 12 January 2021; Published 30 January 2021

Academic Editor: Li Qing

The remaining useful life (RUL) prediction of self-lubricating spherical plain bearings is essential for replacement decision-making and the reliability of high-end equipment. The high-frequency swing self-lubricating liner (HSLL) is the key component of self-lubricating spherical plain bearings under high-frequency oscillation conditions. In this study, a RUL prediction method was proposed based on the Wiener process and grey system theory. First, the predictive processing of the wear depth was carried out using the grey model GM(1,1) to reduce the randomness and enhance the inherent regularity of the life test data. A degradation process model was established and the RUL was predicted online with the model parameter estimates based on the Bayesian updating strategy. Finally, examples were provided to elaborate the RUL prediction of the HSLL. The results show that the prediction accuracy of the proposed RUL prediction model is higher than that of the simple Wiener process during the entire residual life cycle of the HSLL. Based on the original wear data, the prediction accuracy of the RUL exhibited a strong dependence on prior samples and was relatively low owing to the larger deviation of the wear rate between the test sample and prior samples.

1. Introduction

As key activity connectors that are applied extensively to aerospace, engineering machinery, and water projects, self-lubricating spherical plain bearings (SSPB) offer numerous advantages, such as a compact structure, no need for supplementary lubricant, and long life. Moreover, it directly affects the safety and reliability of high-end equipment [1, 2]. A woven self-lubricating liner is critical for ensuring the service performance index of SSPB. The remaining useful life (RUL) prediction of the self-lubricating liner serves as the core and foundation of the service life assessment of SSPB, as well as the fault prediction and health management of high-end equipment [3–5].

With the rapid development of machine learning theory, RUL predictions based on support vector machines have been developed [6, 7]. The performance degradation parameters of the mechanical parts are fitted and the degradation laws are obtained based on the artificial intelligence algorithm. The RUL is calculated with the failure threshold. However, the probability distribution function (PDF), which can embody the random uncertainty of the RUL, cannot be obtained. That is, it is impossible to achieve the probabilistic prediction of the mechanical parts and the evaluation of the RUL prediction rapidly for batch products [8].

As a superior degradation modeling of the random process, RUL predictions based on the Wiener process have been used extensively in the key components of aerospace, railway, and electronic applications. Deng et al. carried out a turbofan engine degradation simulation and improved the near-failure prediction accuracy using the surrogate Wiener propagation model and a long short-term memory network [9]. Guan et al. presented a combined model with a Wiener process-based wear model and unequal interval weighted grey linear regression. The RUL of the Guangzhou Metro was analyzed with high precision [10]. Pan et al. proposed a RUL prediction method based on the Wiener degradation model by considering temporal uncertainty, measurement

uncertainty, and unit-to-unit heterogeneity. The effectiveness of the Wiener degradation model was verified by the degradation dataset of a light-emitting diode [11]. Kong et al. constructed a random process model by inserting the jump at the change point in the degradation process, which was described by the linear Wiener process. The degradation processes of the bearings were simulated [12].

Because numerous interference factors exist during life testing or equipment running, excessive randomness appears in the degradation data and the accuracy of the RUL prediction is significantly reduced. The grey data can be whitened by the accumulated generating operation (AGO) based on the grey system theory [13]. The randomness of the original data is subsequently significantly reduced, and the inherent regularity is excavated. The grey system theory has been used extensively in areas such as the RUL prediction of mechanical parts and trend predictions of social phenomenon development [14, 15]. Sun et al. predicted the nitrous oxide emissions for 2030 in six countries based on three advanced mathematical grey prediction models [14]. Huang et al. proposed a grey online modeling surface roughness monitoring system to predict the surface roughness in end milling operations [16]. Li et al. constructed a grey model for concrete acidification prediction; they analyzed the effects of the pH value, concrete cover thickness, and surface coating on the service life [17]. Ene and Öztürk established a forecasting system for discarded end-of-life vehicles based on the grey system theory [18]. The prediction accuracy was improved by parameter optimization, Fourier series, and Markov chain correction.

In this study, life tests were carried out on a high-frequency swinging self-lubricating liner (HSLL). A RUL model was proposed based on the grey system theory and Wiener process. To improve the prediction accuracy, the original wear depth of the HSLL was conducted, and the prediction data of the wear depth were obtained. Thereafter, the degradation process model of the wear depth was established and the RUL was predicted at any moment with the model parameter estimates based on the Bayesian updating strategy. Examples were provided to illustrate the effectiveness of the RUL model.

This paper introduced grey system theory into the simple Wiener process. The inherent regularity of the original wear data was clearly reflected through predictive processing based on the GM(1,1) model. Compared to the simple Wiener process, the prediction accuracy of the proposed method for the RUL was improved significantly in the entire residual life cycle. And the uncertainty of the model in this paper is smaller than the simple Wiener process.

2. Life Test of HSLL

The HSLL used in the life test was developed by Yanshan University. Brown polytetrafluoroethylene fibers and Kevlar49 fibers were woven into a satin structure, impregnated, semicured, and cured using a modified phenolic resin to obtain the test material. The HSLL had a maximum swing frequency of 15 Hz, a maximum load of 2.8 kN, and a maximum swing angle of ±10°. According to SAE AS 81819A, the maximum wear depth of the HSLL after 25,000 oscillations should be ≤ 0.127 mm [19].

The test samples were in the form of a cylinder contact, as illustrated in Figure 1. Two half-rings and a test shaft made of no. 45 steel with chrome plating on the surface were placed in a group. The failure threshold of the HSLL was $X_f = 0.254$ mm. The life tester was developed by Yanshan University and is shown in Figure 2. The maximum swing frequency of the tester was 30 Hz, the maximum load was 80 kN, and the maximum swing angle was ±15°. The operating parameters for the life test were the most stringent conditions of the self-lubricating liner; that is, $f = 15$ Hz, $F = 2.8$ kN, and $\theta = \pm 10^\circ$. The wear depth data was measured through the displacement sensor and transmitted to a computer. The online measurement and collection of the data can be realized. The life tester had the function of self-compensation of the wear depth; that is, it could compensate for the measurement error caused by the change in the environmental temperature and deformation of the parts under loading.

Based on the HSLL, all samples were prepared under the same bonding process and were tested under the same conditions using the same tester. The wear depth, which is the most critical characteristic parameter, was measured and evaluated as the degradation parameter. Considering that the chrome face on no. 45 steel has a higher degree of hardness and excellent wear resistance, the wear depth between the friction pairs was mainly located on the liner. Therefore, the measured wear depth was taken as the HSLL wear depth. The life tester had an automatic alarm function for failure and excessive wear. The experimenter periodically judged the running state of the life tester and the increasing tendency of the wear depth to avoid the occurrence of abnormal test results. Life tests were conducted on the six groups of samples and the curves of the wear depth were measured, as illustrated in Figure 3.

3. Original Data Processing Based on Grey System Theory

In HSLL life tests, vibration and other random factors still existed, and random errors of the wear depth were inevitable, although displacement compensation with the environmental temperature and loading were considered. These will have a certain influence on the prediction accuracy of the RUL. Grey systems with several uncertain factors are the main objects of studies in the grey system theory. The correct description and effective monitoring of the system operation behavior and evolution law can be realized when partially known information is generated and developed; for example, by using the AGO. Once the test data of the wear depth are processed based on the grey system theory, the regularity of the wear depth can be enhanced, and the prediction accuracy of the RUL can be improved.

3.1. Grey Model GM(1,1). The grey model GM(1,1) is the core model of the grey system theory, and the corresponding differential equations can be obtained by generating the

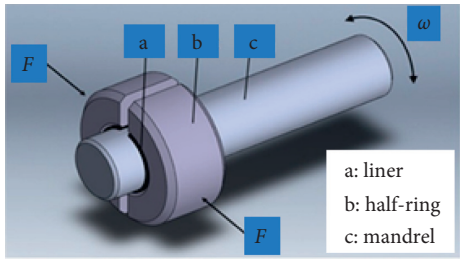


Figure 1: Friction pair for life test.

(a) (b)

Figure 2: Testing machine and self-lubricating fabric liner under test. (a) Testing machine. (b) Self-lubricating fabric liner under test.

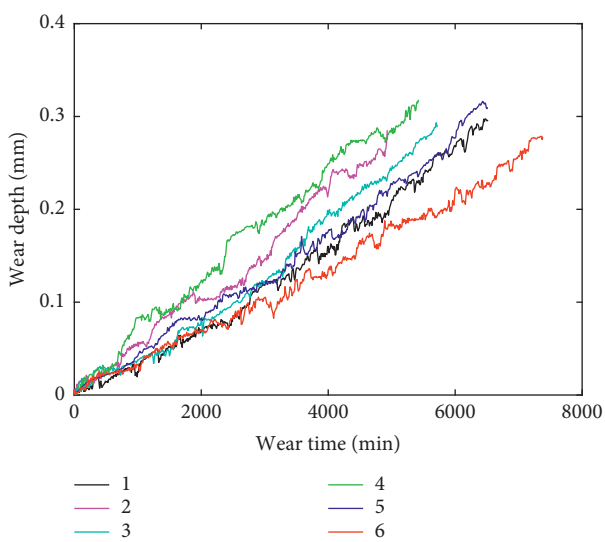


Figure 3: Wear curves of the self-lubricating liner.

discrete grey data into the regular white data [20, 21]. The original series of wear depths for the HSLL can be defined as

$$\mathbf{X}^{(0)} = \left\{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\right\}, \quad (1)$$

where the superscript (0) of $\mathbf{X}^{(0)}$ is the original series of the wear depth.

The first-order AGO can be expressed as $\mathbf{X}^{(1)}$ and is generated from $\mathbf{X}^{(0)}$:

$$\mathbf{X}^{(1)} = \left\{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\right\}, \quad (2)$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, $k = 1, 2, \ldots, n$.

The background value of GM(1,1) is obtained from $\mathbf{X}^{(1)}$:

$$\mathbf{Z}^{(1)} = \left\{z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)\right\}, \quad (3)$$

where $z^{(1)}(k) = \mu(x^{(1)}(k) + x^{(1)}(k-1))$, $k = 2, 3, \ldots, n$, and μ is generally 0.5.

Thereafter, the grey differential equation of GM(1,1) can be obtained as

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad (4)$$

where a and b are the developing coefficient and grey controlled variable, respectively.

The grey derivative of $\mathbf{X}^{(1)}$ is defined in backward difference quotient [22]:

$$\frac{\mathrm{d}x^{(1)}(k)}{\mathrm{d}k} = \frac{\Delta x^{(1)}(k)}{\Delta k} = x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1), \quad k = 2, 3, \ldots, n. \tag{5}$$

For the GM(1,1) presented in equation (4), if the time $k = 2, 3, \ldots, n$, of $x^{(0)}(k)$ is regarded as a continuous variable t, $\mathbf{X}^{(1)}$ can be regarded as a function of t, and it can be expressed as $x^{(1)} = x^{(1)}(t)$. Meanwhile, let $x^{(0)}(k)$ correspond to $\mathrm{d}x^{(1)}/\mathrm{d}t$, and let the background value $z^{(1)}(k)$ correspond to $x^{(1)}(t)$. Then, the whitening differential equation of GM(1,1) can be obtained as follows:

$$\frac{\mathrm{d}x^{(1)}(t)}{\mathrm{d}t} + ax^{(1)}(t) = b. \tag{6}$$

3.2. Original Data Processing. First, inspection is necessary for the original wear data to ensure the feasibility of the modeling method. Taking sample 1 as an example, the class ratio of the original series $\mathbf{X}^{(0)}$ can be calculated as follows:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, \quad k = 2, 3, \ldots, n. \tag{7}$$

If all class ratios $\lambda(k)$ are in the capacitable coverage range $\Theta = (e^{-2/(n+1)}, e^{2/(n+1)})$, the original series $\mathbf{X}^{(0)}$ can be solved by establishing GM(1,1). Otherwise, it is necessary to transform the original series $\mathbf{X}^{(0)}$ to make it fall into the capacitable coverage range, that is, by taking the appropriate constant c for the translation transformation. The transformation can be defined as

$$y^{(0)}(k) = x^{(0)}(k) + c, \quad k = 1, 2, \ldots, n. \tag{8}$$

The minimum and maximum values can be calculated, which are 0.9841 and 1.0173, respectively, by substituting the original series $\mathbf{X}^{(0)}$ of sample 1 into equation (7). Because the capacitable coverage range is $\Theta = (0.9744, 1.0260)$, the translation transformation need not be carried out.

Moreover, a and b in GM(1,1) can be estimated by the least-squares method and the grey differential equation of the original series $\mathbf{X}^{(0)}$:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left[\mathbf{D}^{\mathrm{T}}\mathbf{D}\right]^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{Y}, \tag{9}$$

where

$$\mathbf{D} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}. \tag{10}$$

By substituting the original series $\mathbf{X}^{(0)}$ of sample 1 into equations (9) and (10), a and b can be calculated as 5.17×10^{-4} and 98.31, respectively. The first-order AGO can be obtained by substituting a and b into equation (6):

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}, \quad k = 0, 2, \ldots, n-1. \tag{11}$$

Then, the predicted value of the original degradation series can be obtained as follows:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), \quad k = 0, 2, \ldots, n-1. \tag{12}$$

Finally, the average relative residual can be determined as follows:

$$\bar{\varepsilon} = \sum_{k=2}^{n} \frac{1}{n-1}\left|\frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}\right|. \tag{13}$$

If the average relative residual $\bar{\varepsilon}$ is less than 0.2, the fitting precision of the grey model GM(1,1) to the original series $\mathbf{X}^{(0)}$ meets the general requirements. If $\bar{\varepsilon}$ is less than 0.1, the GM(1,1) model exhibits excellent fitting precision on the original series $\mathbf{X}^{(0)}$.

The predicted values and corresponding average relative residuals can be obtained by substituting the wear data of samples 1, 2, and 3 into the above equations. In this case, $\bar{\varepsilon}$ of samples 1, 2, and 3 are 0.0577, 0.0377, and 0.1402, respectively. Therefore, the GM(1,1) models exhibit high fitting precision, and the prediction curves of samples 1, 2, and 3 are illustrated in Figure 4. The figure indicates that the influence of the external disturbance factors was significantly weakened and the inherent laws of the degenerate variables were extracted effectively.

4. RUL Prediction Based on Wiener Process

4.1. Degradation Process Model of HSLL

4.1.1. Model of Wear Process. The degradation process model of the wear depth for the HSLL could be constructed using the Wiener process. The wear depth at time t is defined as $X(t)$ and is expressed as

$$X(t) = X(0) + \alpha t + \sigma_B B(t), \tag{14}$$

where $X(0)$ is the wear depth at the initial time and $X(0)$ is generally equal to 0. However, for the HSLL, $X(0)$ is not equal to 0 and data conversion must be conducted by $X(t)$-$X(0)$. Moreover, α is the drift coefficient and follows a normal distribution, that is, $\alpha \sim N(\mu_\alpha, \sigma_\alpha^2)$ [23, 24]; σ_B is the diffusion coefficient, which can describe the inconsistency and instability of the HSLL life test as well as the effects of certain random factors such as vibration; $B(t)$ is the standard Brownian motion and follows a normal distribution, that is, $B(t) \sim N(0, t)$; α is the random parameter; and σ_B is the common parameter.

During the online RUL prediction of the single sample, the parametric estimated value must be accurately obtained,

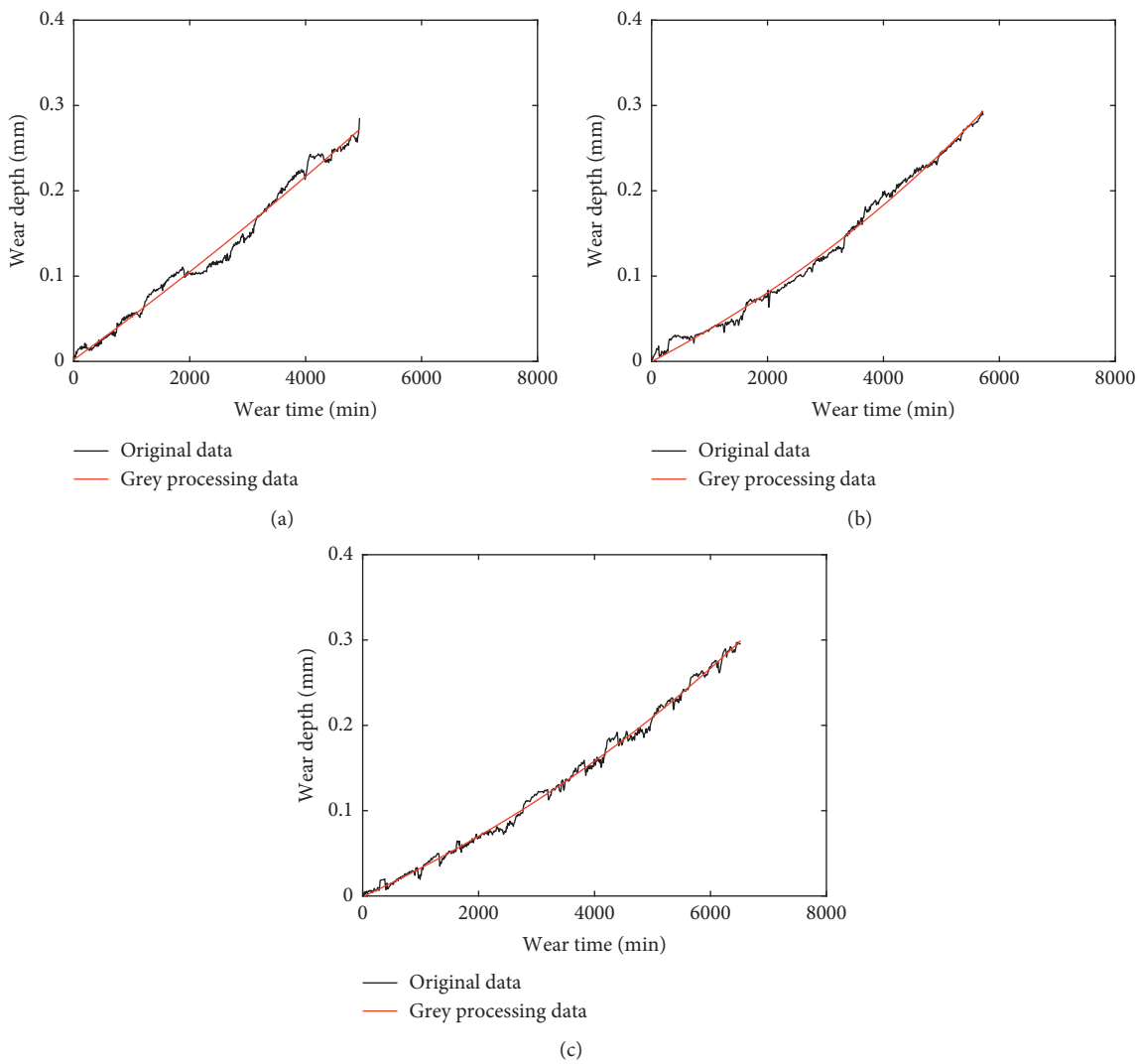


Figure 4: Prediction results of samples 1, 2, and 3.

and it is difficult to meet the requirements depending only on the wear depth of the selected sample. The distributed hyperparameter of α and the estimated value of σ_B should be obtained from the wear data of historical prior samples. Then, the random parameter α of the selected sample can be realized by online updating using the Bayesian method.

4.1.2. Parameter Estimation of Wear Process Model

(1) Prior Parameter Estimation. The parameter estimation of the random parameter α and common parameter σ_B in the wear process model were executed with samples 1, 2, and 3 considered as historical prior samples. The sampling time and sampling number were the same for all samples. The sampling numbers were defined as m. The wear depth of the historical prior samples was fitted by the GM(1, 1) model, and the corresponding predicted values are given as follows:

$$\mathbf{X} = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,m} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,m} \\ X_{3,1} & X_{3,2} & \cdots & X_{3,m} \end{bmatrix}, \tag{15}$$

where $X_{i,j}$ is the prediction value of sample i at time $t_{i,j}$, $1 \le j \le m$.

Following the grey data processing, the degradation process model of the wear depth was established:

$$X_{i,j} = \alpha_0 t + \sigma_B B\left(t_{i.j}\right), \tag{16}$$

where α_0 is the prior value of the random parameter α and follows a normal distribution; that is, $\pi_0(\alpha_0) \sim N(\mu_{\alpha,0}, \sigma^2_{\alpha,0})$. The hyperparameters $\mu_{\alpha,0}$ and $\sigma^2_{\alpha,0}$ could be calculated from historical prior samples.

The wear depth of sample i, $X_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,m})^{\mathrm{T}}$, follows a multivariate Gaussian distribution. The mean and covariance can be obtained as follows:

$$\begin{cases} \boldsymbol{\mu} = \mu_{\alpha,0}\mathbf{t}, \\ \sum = \sigma_{\alpha,0}^2 \mathbf{t}\mathbf{t}^{\mathrm{T}} + \boldsymbol{\Omega}, \end{cases} \tag{17}$$

where $\mathbf{t} = (t_1, t_2, \ldots, t_m)^{\mathrm{T}}$, $\boldsymbol{\Omega} = \sigma_B^2 \mathbf{Q}$, and

$$\mathbf{Q} = \begin{bmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & \cdots & t_2 \\ \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & \cdots & t_m \end{bmatrix}. \tag{18}$$

The log-likelihood function can be generated based on the wear data of the historical prior samples:

$$\begin{aligned} L(\theta|\mathbf{X}_{1:3}) = &-\frac{3m}{2}\ln(2\pi) - \frac{3}{2}\ln|\sum| \\ &-\frac{1}{2}\sum_{i=1}^{3}\left(X_i - \mu_{\alpha,0}\right)^{\mathrm{T}}\sum^{-1}\left(X_i - \mu_{\alpha,0}\right), \end{aligned} \tag{19}$$

where θ is the estimated parameter, $\theta = (\mu_{\alpha,0}, \sigma_{\alpha,0}, \sigma_B)$, and $\mathbf{X}_{1:3}$ denotes the wear data of the historical prior samples.

To simplify the terms in the log-likelihood, we use the results

$$\begin{aligned} |\sum| &= |\boldsymbol{\Omega}|\left(1 + \sigma_{\alpha,0}^2 \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}\right), \\ \sum^{-1} &= \boldsymbol{\Omega}^{-1} - \frac{\sigma_{\alpha,0}^2}{1 + \sigma_{\alpha,0}^2 \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}\boldsymbol{\Omega}^{-1}\mathbf{t}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}. \end{aligned} \tag{20}$$

The first partial derivatives of $L(\theta|\mathbf{X}_{1:3})$ for $\mu_{\alpha,0}$ and $\sigma_{\alpha,0}$ can be expressed as

$$\frac{\partial}{\partial \mu_{\alpha,0}} L(\theta|\mathbf{X}_{1:3}) = \frac{\sum_{i=1}^{3} \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}_i - 3\mu_{\alpha,0}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}{1 + \sigma_{\alpha,0}^2 \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}, \tag{21}$$

$$\frac{\partial}{\partial \sigma_{\alpha,0}} L(\theta|\mathbf{X}_{1:3}) = -\frac{3\sigma_{\alpha,0}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}{1 + \sigma_{\alpha,0}^2 \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}} + \frac{\sigma_{\alpha,0}\sum_{i=1}^{3}\left(\mathbf{X}_i - \mu_{\alpha,0}\mathbf{t}\right)^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\left(\mathbf{X}_i - \mu_{\alpha,0}\mathbf{t}\right)}{\left(1 + \sigma_{\alpha,0}^2 \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}\right)^2}. \tag{22}$$

The expressions of $\mu_{\alpha,0}$ and $\sigma_{\alpha,0}$ can be obtained using equations (21) and (22) equal to 0

$$\widehat{\mu}_{\alpha,0} = \frac{\sum_{i=1}^{3} \mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}_i}{3\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}, \tag{23}$$

$$\widehat{\sigma}_{\alpha,0} = \left[\frac{\sum_{i=1}^{3}\left(\mathbf{X}_i - \mu_{\alpha,0}\mathbf{t}\right)^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\left(\mathbf{X}_i - \mu_{\alpha,0}\mathbf{t}\right)}{3\left(\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}\right)^2} - \frac{1}{\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}\right]^{1/2}. \tag{24}$$

By substituting equations (23) and (24) into equation (19), the contour likelihood function of σ_B can be given by [25]

$$\begin{aligned} L(\sigma_B|\mathbf{X}_{1:3}) = &-\frac{3m}{2}\ln(2\pi) - \frac{3}{2} - \frac{3}{2}\ln|\boldsymbol{\Omega}| \\ &-\frac{1}{2}\left[\sum_{i=1}^{3}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}_i - \frac{\sum_{i=1}^{3}\left(\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}_i\right)^2}{\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}\right] \\ &-\frac{3}{2}\ln\left[\frac{\sum_{i=1}^{3}\left(\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}_i\right)^2}{3\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}} - \frac{\left(\sum_{i=1}^{3}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}\right)^2}{3^2\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{t}}\right]. \end{aligned} \tag{25}$$

The maximum of equation (25) can be calculated by the optimization toolbox of MATLAB, and the maximum likelihood estimate of the common parameter σ_B can be determined as follows:

$$\widehat{\sigma}_B = 5.7095 \times 10^{-5}. \tag{26}$$

By substituting equation (26) into equations (23) and (24), the estimated values of $\mu_{\alpha,0}$ and $\sigma_{\alpha,0}$ can be obtained as follows:

$$\begin{cases} \widehat{\mu}_{\alpha,0} = 4.9750 \times 10^{-5}, \\ \widehat{\sigma}_{\alpha,0} = 4.1708 \times 10^{-6}. \end{cases} \tag{27}$$

(2) Real-Time Parameter Updating. The wear data of samples 4, 5, and 6 were selected as the test sets and the random parameter α was updated in real time. Following the grey data processing, the wear process model can be expressed as

$$X_k = \alpha_k t_k + \sigma_B B(t_k), \tag{28}$$

where X_k is the wear depth at time t_k of the single sample online and α_k is the posterior value of the random parameter α and follows a normal distribution; that is, $\pi_k(\alpha_k) \sim N(\mu_{\alpha,k}, \sigma_{\alpha,k}^2)$.

It was assumed that the wear depth before t_k is $\mathbf{X}_{1:k} = (X_1, X_2, \ldots, X_k)^{\mathrm{T}}$ of the single sample online, and the time interval is Δt. Based on the Bayesian method, the posterior distribution of the random parameter α can be determined by [26, 27]

$$\pi(\alpha_k|\mathbf{X}_{1:k}) \propto L(\mathbf{X}_{1:k}|\alpha_k)\pi_0(\alpha_0), \tag{29}$$

where the prior information follows a normal distribution, that is, $\pi_0(\alpha_0) \sim N(\mu_{\alpha,0}, \sigma^2_{\alpha,0})$, and $L(\mathbf{X}_{1:k}|\alpha_k)$ is a likelihood function with given random parameters.

Based on equation (14) and the basic nature of the Brownian motion, $L(\mathbf{X}_{1:k}|\alpha_k)$ can be obtained as follows:

$$L(\mathbf{X}_{1:k}|\alpha_k) = \frac{\exp\left[-\sum_{q=1}^{k}\left(X_q - X_{q-1} - \alpha_k\Delta t\right)^2/2\sigma_B^2\Delta t\right]}{\prod_{q=1}^{k}\sqrt{2\pi\sigma_B^2\Delta t}.} \tag{30}$$

Because both $L(\mathbf{X}_{1:k}|\alpha_k)$ and $\pi_0(\alpha_0)$ follow a normal distribution, $\pi(\alpha_k|\mathbf{X}_{1:k})$ also follows a normal distribution. Based on equations (29) and (30), the mean and variance of $\pi(\alpha_k|\mathbf{X}_{1:k})$ can be obtained as follows, respectively:

$$\begin{aligned}\mu_{\alpha,k} &= \frac{\mu_{\alpha,0}/\sigma^2_{\alpha,0} + \sum_{q=1}^{k}\left(\left(X_q - X_{q-1}\right)/\sigma_B^2\right)}{\sum_{q=1}^{k}\left(\Delta t^2/\sigma_B^2\Delta t\right) + \left(1/\sigma^2_{\alpha,0}\right)},\\ \sigma_{\alpha,k} &= \sqrt{\frac{1}{\sum_{q=1}^{k}\left(\Delta t^2/\sigma_B^2\Delta t\right) + \left(1/\sigma^2_{\alpha,0}\right).}}\end{aligned} \tag{31}$$

4.2. RUL Prediction. The RUL distribution could be obtained using the concept of the first arrival times during the wear process. When the random process $\{X(t), t \geq 0\}$ reached the failure threshold, the time was defined as the sample life:

$$T = \inf\left\{t\colon X(t) \geq X_f | X(0) < X_f\right\}, \tag{32}$$

where X_f is the failure threshold, $X(0)$ is the wear depth of the initial moment, and $X(0) = 0$.

The residual useful life L_k in the current runtime t_k can be defined as

$$L_k = \{l_k\colon T - t_k | T > t_k\}, \tag{33}$$

where l_k is the online RUL value.

Because the RUL of the Wiener process with linear drift follows an inverse Gaussian distribution [28], the PDF of L_k can be obtained as follows:

$$f_{L_k}(l_k) = \frac{X_f - X(k)}{\sqrt{2\pi l_k^3\left(\sigma^2_{\alpha,k}l_k + \sigma_B^2\right)}}\exp\left\{-\frac{\left(X_f - X(k) - \mu_{\alpha,k}l_k\right)^2}{2l_k\left(\sigma^2_{\alpha,k}l_k + \sigma_B^2\right)}\right\}. \tag{34}$$

Based on the observed data, the PDF of the RUL for sample 4 could be calculated and is illustrated in Figure 5.

Figure 5 indicates that the PDF curves become narrower with wear data information increases, which indicates the uncertainty of RUL prediction is increasingly small, and the results of RUL prediction will be increasingly accurate.

The RUL of a single sample can be written as

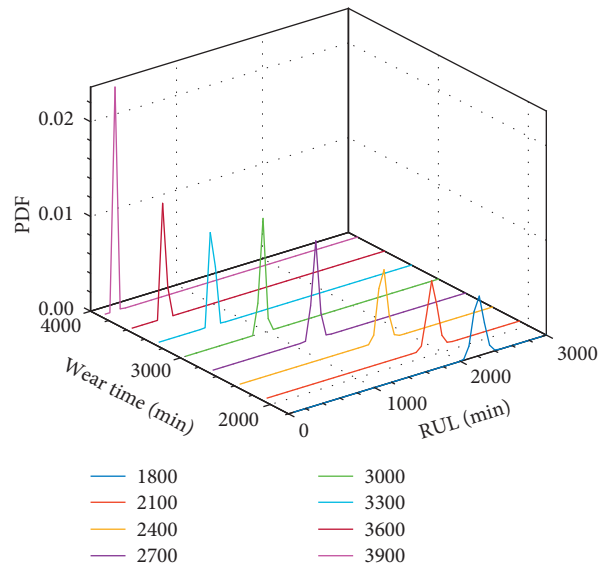


Figure 5: PDF of RUL.

$$\begin{aligned}E(L_k) &= E\left(E\left(L_k|\left(\mu_{\alpha,k}, \sigma_{\alpha,k}\right)\right)\right)\\ &= E\left(\frac{X_f - X(k)}{\left(\mu_{\alpha,k}, \sigma_{\alpha,k}\right)}\right)\\ &= \frac{X_f - X(k)}{\sigma^2_{\alpha,k}}\exp\left(\frac{-\mu^2_{\alpha,k}}{2\sigma^2_{\alpha,k}}\right)\int_0^{\mu_{\alpha,k}}\exp\left(\frac{x^2}{2\sigma^2_{\alpha,k}}\right)\mathrm{d}x\\ &= \frac{\sqrt{2}\left(X_f - X(k)\right)}{\sigma_{\alpha,k}}D\left(\frac{\mu_{\alpha,k}}{\sqrt{2}\,\sigma_{\alpha,k}}\right),\end{aligned} \tag{35}$$

where $D(z)$ is the Dawson function with respect to z:

$$\begin{aligned}D(z) &= \exp\left(-z^2\right)\int_0^z \exp\left(x^2\right)\mathrm{d}x,\\ z &= \frac{\mu_{\alpha,k}}{\sqrt{2}\sigma_{\alpha,k}}.\end{aligned} \tag{36}$$

When z is sufficiently large, according to the approximate property of the integral, $D(z) \approx 1/(2z)$. The RUL can be obtained as

$$E(L_k) \approx \frac{X_f - X(k)}{\mu_{\alpha,k}}. \tag{37}$$

600 min of the wear data was selected as the initial prediction point. The random parameters were updated every 300 min. The RUL of samples 4, 5, and 6 could be predicted with the origin wear depth and grey processing data, respectively. The curves of the actual RUL and the prediction RUL were depicted in Figure 6.

Figure 6(a) indicates that the predicted values of the three samples were close at the initial prediction point. The actual and predicted values of sample 5 exhibited better consistency. For samples 4 and 6, which had lower/higher

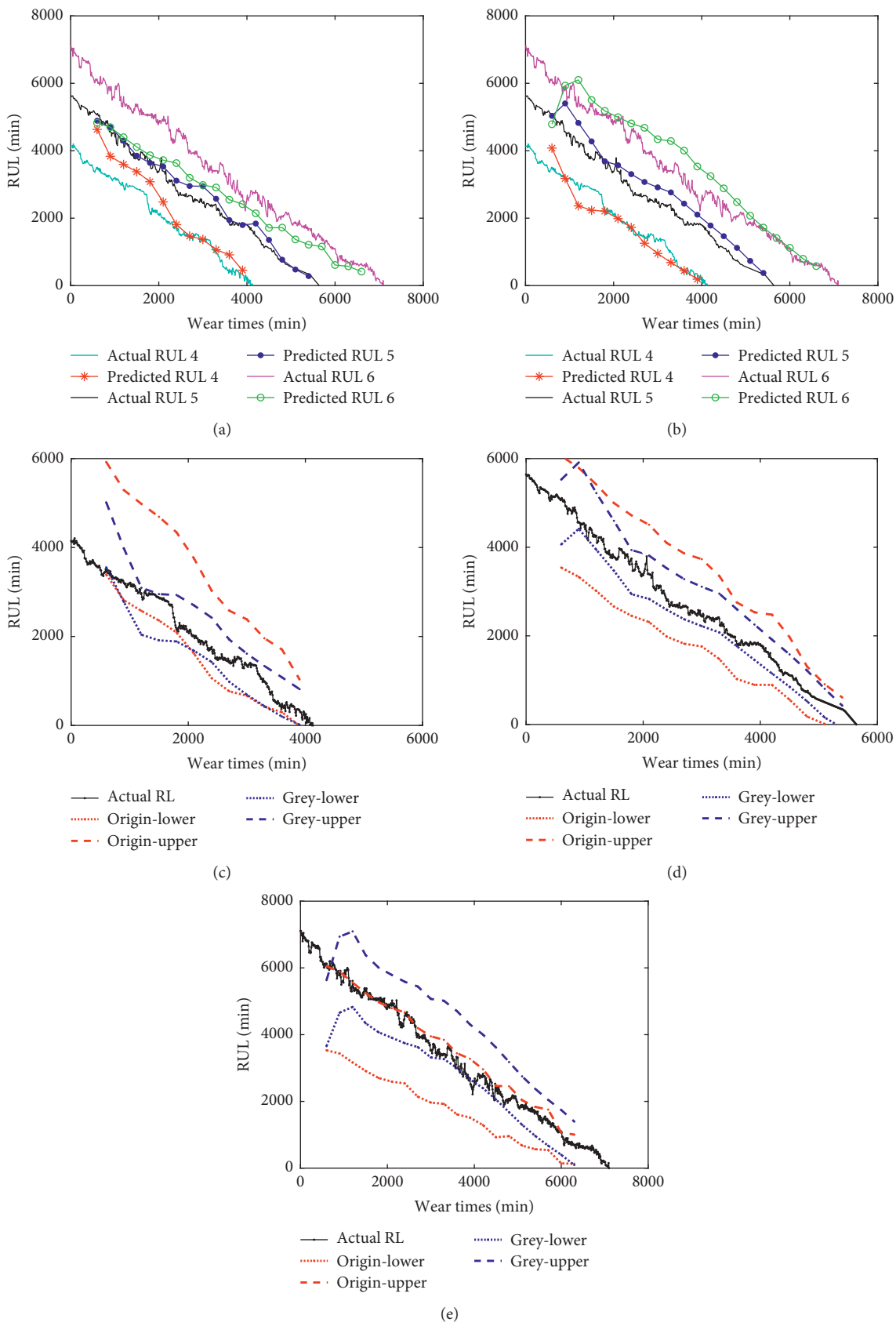


Figure 6: RUL and confidence interval for HSLL based on original wear depth and grey processing data. (a) Origin wear depth. (b) Grey processing data. (c) Confidence interval of samples 4. (d) Confidence interval of samples 5. (e) Confidence interval of samples 6.

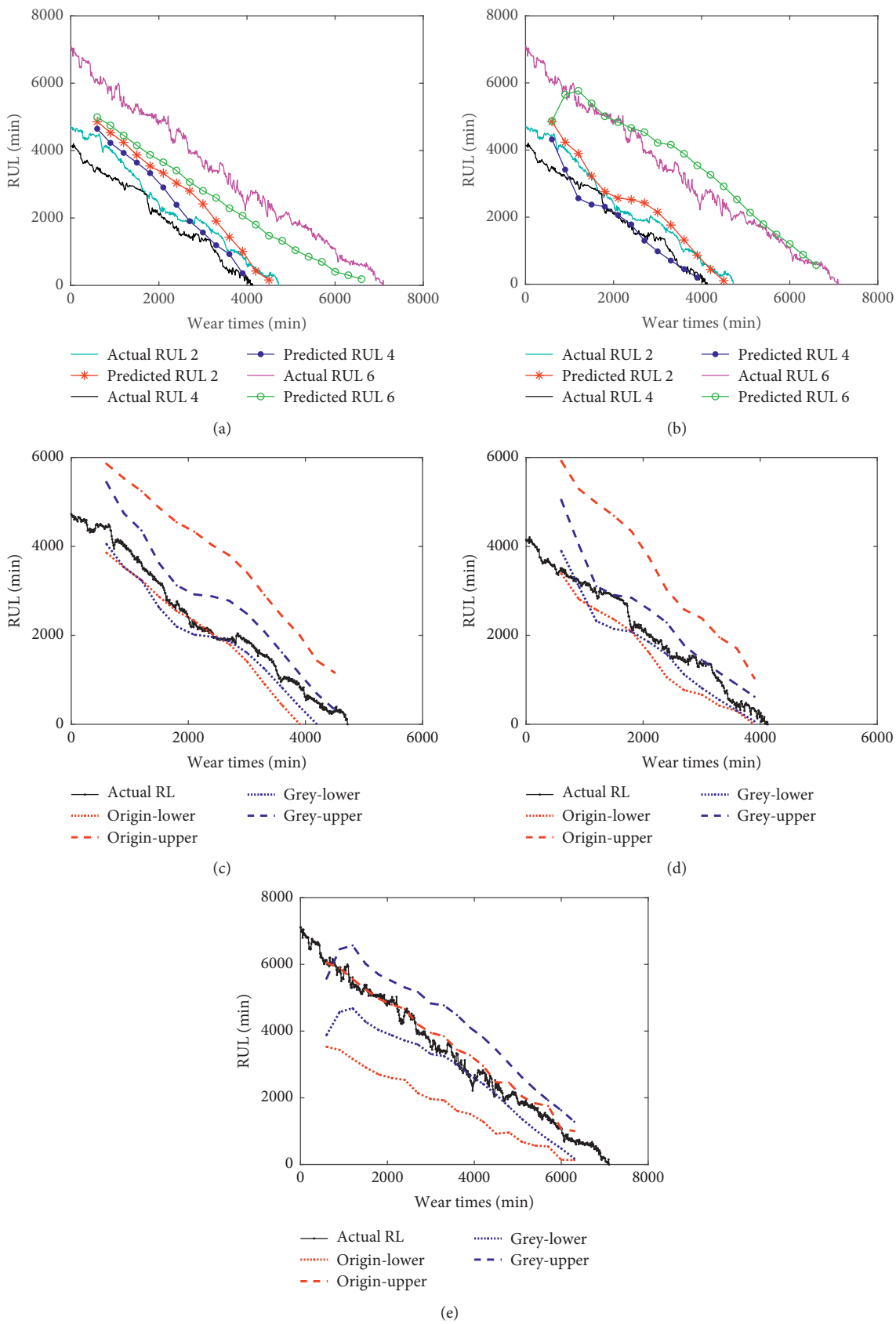


Figure 7: RUL and confidence interval for HSLL based on original wear depth and grey processing data. (a) Origin wear depth. (b) Grey processing data. (c) Confidence interval of samples 2. (d) Confidence interval of samples 4. (e) Confidence interval of samples 6.

Table 1: The RMSE, MAE, and SMAPE of the predicted values.

Indicator	4		5		6	
	Original wear data	Grey processing data	Original wear data	Grey processing data	Original wear data	Grey processing data
RMSE	0.6293	0.2516	0.2054	0.2331	0.9360	0.4874
MAE	0.5498	0.1581	0.1846	0.1886	0.8192	0.3597
SMAPE	34.60%	9.30%	9.60%	8.64%	26.59%	13.33%

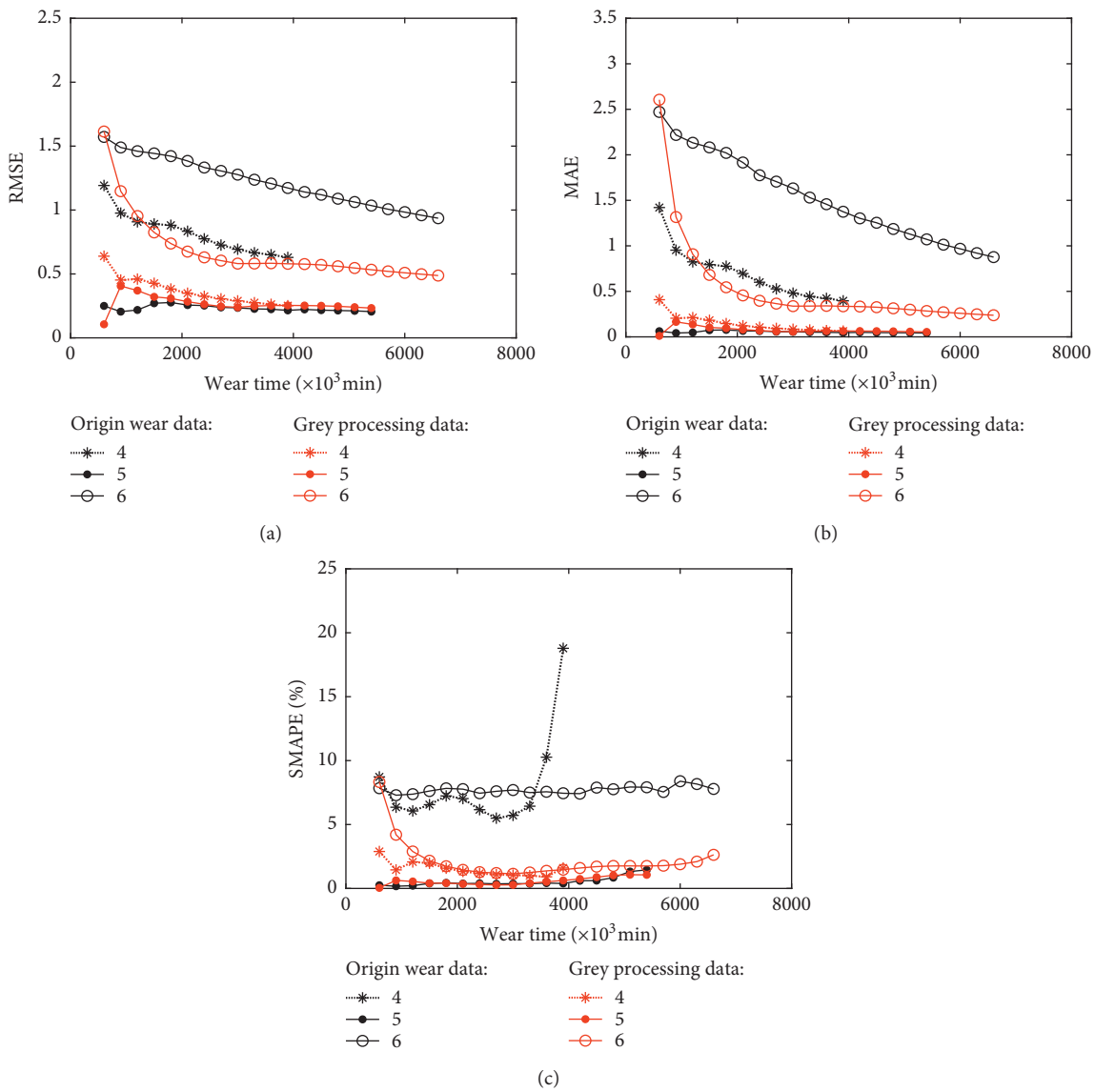


Figure 8: RMSE MAE and SMAPE of predicted RUL values. (a) RMSE. (b) MAE. (c) SMAPE.

wear rates, the prediction deviation was relatively large. However, the predicted life values gradually became closer to the actual life values and were almost equal to the actual life values at the end of the wear process. Following the wear depth being processed by the GM(1,1) model, as shown in Figure 6(b), the initial prediction point of predicted values were also basically close, but the prediction accuracy of the RUL was improved significantly in the entire residual life cycle.

Simultaneously, Figures 6(c)–6(e) display the 95% pointwise confidence intervals for the RUL based on these prediction points. Basically, all confidence intervals can cover the actual values of the RUL. The actual RUL line of sample 5 was positioned centrally over the confidence interval, for the origin wear depth, and the confidence intervals of samples 4 and 6 were relatively biased. And compared to the simple Wiener process, the 95% confidence interval of the proposed method in the paper is narrower, which can

help reduce uncertainty in residual life prediction results and can better adapt to the individual needs of residual life prediction. These further indicate that the errors of the proposed method are lower.

Meanwhile, the RULs of samples 2, 4, and 6 when samples 1, 3, and 5 were selected as the prior samples are shown in Figure 7.

Figure 7 had the RUL prediction curves and the 95% pointwise confidence intervals similar to Figure 6. When the differences of the wear rates for samples 2, 4, and 6 were larger, the prediction deviations were relatively large at the initial phase. But the predicted life values gradually became closer to the actual life values with the wear process. Similarly, the overall prediction accuracy of the proposed method is higher than the original wear data.

Synthesizing Figures 5–7, the PDF of RUL became narrower with the continuous accumulation of data. That is, the uncertainty of RUL prediction becomes smaller and smaller. And the relative error between the prediction of the RUL and the actual RUL is smaller. These happened because the results of model parameter estimation were gradually closer to the reality of sample degradation with the test data increasing.

To analyze the prediction precision of the proposed RUL method, multiple evaluation indicators, namely, the root mean square error (RMSE), mean absolute error (MAE), and symmetric mean absolute percentage error (SMAPE), were calculated and can be given as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(h\left(t_i\right) - y_i\right)^2}, \tag{38}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|h\left(t_i\right) - y_i\right|, \tag{39}$$

$$\mathrm{SMAPE} = \frac{1}{m}\sum_{i=1}^{m}\frac{\left|h\left(t_i\right) - y_i\right|}{\left(\left|h\left(t_i\right)\right| + \left|y_i\right|\right)/2} \times 100\%, \tag{40}$$

where m is a positive integer, $h(t_i)$ is the predicted value at point i, and y_i is the actual RUL at point i.

The evaluation indicators of samples 4, 5, and 6 can be calculated using equations (38)–(40), as shown in Table 1.

Table 1 shows that the assessment results were similar for RMSE and MAE. For sample 5, the prediction accuracy based on the original wear data was slightly higher than that based on the grey processing data. But for samples 4 and 6, the prediction accuracy of the proposed method was much higher. Based on the SMAPE, the predicted trends of the proposed method were all more accurate.

The broken line graph of RMSE, MAE, and SMAPE shown in Figure 8 reflects the changes in the RUL prediction accuracy of the predicted values at point i. All three indicators decreased gradually, and the prediction accuracy all gradually increased. Because of the great difference in the wear rates among samples, the values of the three indicators based on the original wear data varied greatly. However, all three indicators based on the grey processing data remained at relatively low levels, and the effectiveness of the proposed method in this paper was demonstrated.

5. Conclusions

The predictive processing of the wear depth was executed based on the GM(1,1) model. The randomness of the original wear data was reduced and the inherent regularity was clearly reflected. The RUL prediction for the HSLL was accomplished using the Wiener process model and data whitening operation on the original wear depth. Based on the original wear data and simple Wiener process, the prediction accuracy of the RUL exhibited a strong dependence on the prior samples, and the prediction accuracy was relatively low owing to the deviation of the wear rate between the test sample and prior samples. Compared to the simple Wiener process, the prediction accuracy of the proposed method for the RUL was improved significantly in the entire residual life cycle. For 95% confidence interval, the width of an interval for the improved method is narrower than the simple method, which indicates the smaller uncertainty of the model proposed.

Although the RUL prediction of self-lubricating spherical plain bearings based on the Wiener process and grey system theory has exhibited a relatively high prediction accuracy, it is important that the origin data is nonnegative and performs well on the smooth performance. And depending on the gradual degradation data rather than the saltation data, the application of the proposed method may be limited; hence, further study is required.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This project was funded by the National Natural Science Foundation of China (Grant no. 51605418). The authors would like to thank Editage (http://www.editage.cn) for English language editing.

References

[1] B. Suresha, K. Shiva Kumar, S. Seetharamu, and P. Sampath Kumaran, "Friction and dry sliding wear behavior of carbon and glass fabric reinforced vinyl ester composites," *Tribology International*, vol. 43, no. 3, pp. 602–609, 2010.

[2] J. Aguirrebeitia, M. Abasolo, J. Vallejo, I. Coria, and I. Heras, "Methodology for the assessment of equivalent load for self-lubricating radial spherical plain bearings under combined load," *Tribology International*, vol. 105, pp. 69–76, 2017.

[3] M. Qiu, Y. Miao, Y. Li, and J. Lu, "Film-forming mechanisms for self-lubricating radial spherical plain bearings with hybrid

PTFE/aramid fabric liners modified by ultrasonic," *Tribology International*, vol. 87, pp. 132–138, 2015.

[4] D. Xiang, W. Shu, and K. Li, "Friction and wear behavior of a new 40Cr steel-PTFE fabric composite under heavy loads," *Materials Science and Engineering: A*, vol. 484, pp. 365–368, 2008.

[5] F. Camci and R. B. Chinnam, "Health-state estimation and prognostics in machining processes," *IEEE Transactions on Automation Science and Engineering*, vol. 7, no. 3, pp. 581–597, 2010.

[6] C. Duan, M. Shao, S. Li et al., "The research progress in polyimide self-lubricating composites under extreme conditions," *Scientia Sinica Chimica*, vol. 48, no. 12, pp. 1561–1567, 2018.

[7] L. Liao and F. Kottig, "Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction," *IEEE Transactions on Reliability*, vol. 63, no. 99, pp. 191–207, 2014.

[8] T. H. Loutas, D. Roulias, and G. Georgoulas, "Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression," *IEEE Transactions on Reliability*, vol. 62, no. 4, pp. 821–832, 2013.

[9] Y. J. Deng, B. A. Di, and M. Pechenizkiy, "Controlling the accuracy and uncertainty trade-off in RUL prediction with a surrogate Wiener propagation model," *Reliability Engineering & System Safety*, vol. 196, pp. 1–10, 2020.

[10] Q. L. Guan, X. K. Wei, and L. M. Jia, "RUL prediction of railway PCCS based on Wiener process model with unequal interval wear data," *Applied Sciences-Basel*, vol. 10, no. 5, pp. 1–22, 2020.

[11] D. Pan, S. Lu, Y. Liu, W. Yang, and J.-B. Liu, "Degradation data analysis using a Wiener degradation model with three-source uncertainties," *IEEE Access*, vol. 7, pp. 37896–37907, 2019.

[12] D. Kong, N. Balakrishnan, and L. Cui, "Two-phase degradation process model with abrupt jump at change point governed by Wiener process," *IEEE Transactions on Reliability*, vol. 66, no. 4, pp. 1345–1360, 2017.

[13] E. Scarlat and C. Delcea, "An overview on the hybrid intelligent systems from the grey systems theory and knowledge perspective," *Journal of Grey System*, vol. 28, no. 2, pp. 13–26, 2016.

[14] H. Sun, J. Jiang, M. Mohsin, J. Zhang, and Y. A. Solangi, "Forecasting nitrous oxide emissions based on grey system models," *Environmental Geochemistry and Health*, vol. 42, no. 3, pp. 915–931, 2020.

[15] T.-Y. Pai, S.-H. Lin, P.-Y. Yang, D.-H. Chang, and J.-L. Kuo, "Predicting hourly ozone concentration time series in Dali area of Taichung city based on seven types of GM (1,1) model," *Time Series Analysis, Modeling and Applications*, vol. 47, pp. 369–383, 2013.

[16] P. B. Huang, H.-J. Zhang, and Y.-C. Lin, "Development of a grey online modeling surface roughness monitoring system in end milling operations," *Journal of Intelligent Manufacturing*, vol. 30, no. 4, pp. 1923–1936, 2019.

[17] B. Li, L. Cai, and W. Zhu, "Predicting service life of concrete structure exposed to sulfuric acid environment by grey system theory," *International Journal of Civil Engineering*, vol. 16, no. 9, pp. 1017–1027, 2018.

[18] S. Ene and N. Öztürk, "Grey modelling based forecasting system for return flow of end-of-life vehicles," *Technological Forecasting and Social Change*, vol. 115, pp. 155–166, 2017.

[19] L. L. Jing, H. M. Zhu, and Z. Z. Sun, "Standard analysis of self lubricating plain bearing," *Aeronautic Standardization & Quality*, vol. 10, pp. 34–38, 2010.

[20] J. L. Deng, *Grey Forecasting and Decision*, Huazhong University of Science and Technology Press, Wuhan, China, 1986.

[21] J. L. Deng, *The Primary Method of Grey System Theory*, Huazhong University of Science and Technology Press, Wuhan, China, 2002.

[22] J. L. Deng, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.

[23] Z.-S. Ye, Y. Wang, K.-L. Tsui, and M. Pecht, "Degradation data analysis using wiener processes with measurement errors," *IEEE Transactions on Reliability*, vol. 62, no. 4, pp. 772–780, 2013.

[24] G. Jin, D. E. Matthews, and Z. Zhou, "A Bayesian framework for on-line degradation assessment and residual life prediction of secondary batteries inspacecraft," *Reliability Engineering & System Safety*, vol. 113, pp. 7–20, 2013.

[25] C. Y. Peng and S. T. Tseng, "Mis-specification analysis of linear degradation models," *IEEE Transactions on Reliability*, vol. 58, no. 3, pp. 444–455, 2009.

[26] C. H. Hu, H. D. Fan, and Z. Q. Wang, *Equipment Remaining Life Prediction and Optimal Maintenance Decision*, National Defense Industry Press, Beijing, China, 2018.

[27] N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, "Residual-life distributions from component degradation signals: a Bayesian approach," *IIE Transactions*, vol. 37, no. 6, pp. 543–557, 2005.

[28] X.-S. Si, W. Wang, C.-H. Hu, D.-H. Zhou, and M. G. Pecht, "Remaining useful life estimation based on a nonlinear diffusion degradation process," *IEEE Transactions on Reliability*, vol. 61, no. 1, pp. 50–67, 2012.

Hindawi
Shock and Vibration
Volume 2020, Article ID 8876187, 15 pages
https://doi.org/10.1155/2020/8876187

Research Article

Tool Wear Assessment Approach Based on the Neighborhood Rough Set Model and Nearest Neighbor Model

Ren Sheng and Xiaoran Zhu

School of Mechanical Engineering, Yellow River Conservancy Technical Institute, Kaifeng, Henan, China

Correspondence should be addressed to Xiaoran Zhu; xiaoran.zhu@foxmail.com

Received 15 July 2020; Revised 11 November 2020; Accepted 5 December 2020; Published 17 December 2020

Academic Editor: Li Qing

In order to assess the degree of wear of tool for milling process quantitatively, a new assessment approach is proposed. Firstly, making full use of the neighbor information, two sensitive features are selected by using the neighborhood rough set model, and then, boundary curves are established by using the nearest neighbor model with noncounter data in two dimension spaces. Secondly, the intersection area or expanding area is used to describe the difference between two boundary models because the intersection area or expanding area can consider the effect of distance and angle simultaneously in two dimension spaces. Thirdly, after determining a baseline state, a new quantitative assessment indicator (QAI) can be calculated based on the intersection area or expanding area. The QAI can directly measure the difference between the model of baseline state and the model of unknown state and indirectly measure the degree of wear of tool. Finally, the effectiveness of the assessment approach is proven by using the Milling Dataset which was provided by the NASA Ames Research Center.

1. Introduction

In order to guarantee the quality and productivity of the machining process, tool condition monitoring has received more and more attention. Zhou and Xue [1] summarized the monitoring methods and monitoring algorithms in the milling process, and Prashant Waydande and Chinchanikar [2] reviewed the monitoring system of tool wear. Generally speaking, the condition monitoring methods and systems of tool could be divided into two parts: one is physics-driven models, and the other is data-driven models. Physics-driven models explore a physical understanding of the tool in order to estimate the running state and wear degree. For example, Yen et al. [3] used Taylor's formula to estimate the tool wear. Zhu and Zhang [4] developed a generic explicit model with adjustable coefficients for flank wear of tool and then established a relationship between milling force and flank wear of tool. Therefore, it could use the tool wear model to predict the life. For example, Nouri et al. [5] proposed a method to monitor end milling tool wear in real-time by tracking the force model coefficient during the cutting process to predict the life to tool. However, the major challenges of physics-driven models included the following [6]: sufficient knowledge of the failure mechanism of tool wear was lacked, and the values of some parameters in the formulations were difficultly obtained. In order to avoid the comprehensive or complicated physical models, data-driven models were chosen and developed in recent years. Data-driven models tried to model or simulate the tool wear process by using the actual running data, including force, vibration, acoustic emission, spindle motor and feed currents, and so on. For example, Shankar et al. [7] used sound pressure and machining force to train a neural network to predict flank wear of tool during the milling process. Kalvoda and Hwang [8] used the Hilbert–Huang transform to process the acceleration vibration signal, and the results showed that the changing of frequency of the marginal spectra could reflect the geometric changing of the cutter tool. After carefully analyzing the acoustic emission signal by using statistical analysis, spectrum analysis, and wavelet packet, Leng et al. [9] proposed a method to detect the tool wear for drilling

the composite material. Sun et al. [10] studied the relationship between inverter input current and tool wear condition based on current virtual values which were calculated from the line current on the inverter input side. Proteau et al. [11] proposed a specific cutting energy, which used selected multifeatures to train a recurrent neural network, to predict the tool wear. In general, the data-driven models used machine learning algorithms or artificial intelligence algorithms to train a model, and then, the model could automatically recognize the complex patterns. But, the modeling process needed a large amount of historical data [12], especially failure data or degraded data, whereas the failure data or degraded data were very difficult to get. If we can only use one-class data to train a model, the modeling process would become much easier. Therefore, the concept of anomaly detection could help us.

Anomaly detection refers to the process of finding patterns in data that do not conform to expected behavior [13]. The nonconforming patterns include anomalies, outliers, discordant observations, exceptions, aberrations, surprises, and contaminants. The anomaly detection usually used normal data to train models because the nonconforming patterns were unknown and unexpected. So, if we consider the state of tool wear as nonconforming patterns, we could use anomaly detection methods to detect the state of tool wear. The models of anomaly detection included the density model, reconstruction model, and boundary model [14]. For example, the Gaussian model was a density model. If the distribution of unknown data was different with the distribution of the model, the unknown data could be considered as outlier. But, the density model needed typical data from the true data distribution and needed a large amount of training data for high dimension data. The self-organizing map (SOM) was a reconstruction model. It tried to reconstruct a new model in the new space after mapping the original data into a high-dimension space. But, it needed to assume the clustering characteristics of the training data. For the boundary model, it tried to find an optimized boundary around the target data, and the typical models were the support vector data description (SVDD) [15] and the nearest neighbor model. Because the nearest neighbor model tried to directly embrace the entire target data based on the distance information, the model was easily understood and used. Therefore, we choose the nearest neighbor model to describe the running data of tool. Moreover, the redundant and irrelevant features could affect the accuracy of the nearest neighbor model. Therefore, the sensitive features should be selected firstly. The rough set theory [16] could be used to select sensitive features based on the concept of lower approximation, upper approximation, and boundary region. Unfortunately, the rough set only processed discrete features, and continuous features should be discretized. But, the discretization could cause the loss of information. Then, Hu et al. [17] proposed the neighborhood rough set model. The neighborhood rough set can directly process continuous features based on the concept of neighborhood relationship. Therefore, we choose the neighborhood rough set model to select the sensitive features.

In order to monitor the wear condition of tool in process, a new assessment method based on the neighborhood rough set model and the nearest neighbor model is proposed. After selecting sensitive features by using the neighborhood rough set model, the wear condition of tool could be described by using the boundary curves which are trained from the nearest neighbor model. Finally, the quantitative assessment indicator is designed based on the areas which are created by the boundary curves in the two-dimension spaces.

2. Neighborhood Rough Set Model and Nearest Neighbor Model

One key definition of the neighborhood rough set model is the dependency degree. For example, the dependency degree of D to B is defined as the ratio of consistent objects:

$$r_B(D) = \frac{\left|\mathrm{POS}_B(D)\right|}{|U|}, \tag{1}$$

where U is the sample space. $\mathrm{POS}_B(D)$ calls the positive region. The positive region is the sample set which can be classified into one of the classes without uncertainty. $r_B(D)$ reflects the capability of B to approximate D.

Another key definition is the significance of a feature:

$$\mathrm{SIG}(a, B, D) = r_B(D) - r_{B-a}(D), \tag{2}$$

and if $\mathrm{SIG}(a, B, D) = 0$, feature a is superfluous in B; otherwise, a is indispensable.

Based on the concept of the neighborhood rough set model, the dependency would decrease if any sensitive feature is missed, and then, we could use the significance of a feature to evaluate each feature. At the beginning, the pool of selected features is empty. Then, we could traverse all features, finding the sensitive features which could make the SIG get maximum values and $\mathrm{SIG} > 0$. Finally, the sensitive features were added to the pool of selected features. The detailed description of feature selection algorithm could be found in [17].

The nearest neighbor model is a one-class classifier. It tries to estimate an hypersphere, which is centered around the test object $\mathbf{z}$, in d dimensions. The volume of the hypersphere is grown until it captures k objects from the training set. The local density could be estimated as follows:

$$p_{\mathrm{NN}}(z) = \frac{(k/N)}{V_k\left(\left\|z - \mathrm{NN}_k^{\mathrm{train}}(z)\right\|\right)}, \tag{3}$$

where $\mathrm{NN}_k^{\mathrm{train}}(\mathbf{z})$ is the k nearest neighbor of $\mathbf{z}$ in the training set and V_k is the volume of the hypersphere containing this object.

For the one-class classifier, a test object $\mathbf{z}$ is accepted when its local density is larger or equal to the local density of the first nearest neighbor in the training set $\mathrm{NN}^{\mathrm{train}}(\mathbf{z}) = \mathrm{NN}_1^{\mathrm{train}}(\mathbf{z})$.

3. Assessment Approach

It is very important how to quantitatively assess the degree of tool wear. Therefore, this section would focus on the quantitative indicator firstly. Pan et al. [18] proposed a quantitative assessment indicator based on the SVDD. After training a hypersphere of baseline state by using SVDD with the data of baseline state, the formulation of the center of the hypersphere of baseline state could be obtained. Then, the distance of the data of unknown state from the center of the hypersphere of baseline state could be calculated, and the value of distance could be considered as the quantitative assessment indicator. Although the SVDD and the nearest neighbor model all belong to the boundary model, there is no formulation of the center of the nearest neighbor model. So, it is difficult to calculate the distance. Besides, the distance cannot distinguish two different running states which have the same distance.

After getting the boundary curves of the baseline state and the unknown state by using the nearest neighbor model, the spatial position of two boundary curves could help us to determine the running condition. As shown in Figure 1(a), the intersection area could associate with the degree of similarity for two boundary models. The bigger the intersection area, the greater the similarity for two boundary models. As shown in Figure 1(b), the expanding area could associate with the degree of nonsimilarity for two boundary models. The bigger the expanding area, the greater the nonsimilarity for two boundary models. Therefore, the value of area could be considered as a quantitative indicator to assess the running condition. Moreover, the intersection area or expanding area can consider the effect of distance and angle simultaneously. The formulation of QAI is

$$\mathrm{QAI}_i = \begin{cases} \mathrm{Area}_{\mathrm{baseline}} + \mathrm{Area}_{i\mathrm{th\ state}} - \mathrm{Area}_{\mathrm{intersection}} \text{ for intersection situation,} \\ \mathrm{Area}_{\mathrm{baseline}} + \mathrm{Area}_{i\mathrm{th\ state}} + \mathrm{Area}_{\mathrm{expanding}} \text{ for expanding situation.} \end{cases} \tag{4}$$

The area is estimated by using the Monte Carlo method, taking intersection area as an example, as shown in Figure 2. Firstly, a square is created, and then, the area of square could be easily estimated, as $\mathrm{Area}_{\mathrm{square}}$. Secondly, uniformly scattering points over the square, the number of points is N_{total}. Thirdly, counting the number of points inside the square, the number is N_{inside}. Finally, the ratio of the intersection area and square area is equal to the ratio of the inside count and the total sample count, so the intersection area is $\mathrm{Area}_{\mathrm{intersection}} = (N_{\mathrm{inside}}/N_{\mathrm{total}})\mathrm{Area}_{\mathrm{square}}$.

The assessment process is shown in Figure 3. Firstly, the Milling Dataset is analyzed by using the neighborhood rough set model, and then, two sensitive features are selected. Secondly, we suppose that the degree of tool wear is very small at beginning stage, especially for first-time use. Therefore, the first test is selected as baseline state. Then, the boundary curve of baseline state can be obtained by using the nearest neighbor model. Thirdly, for unknown state, using same two sensitive features to train the nearest neighbor model, the boundary curve of unknown state can be obtained. Finally, the QAI which corresponds to the unknown state can be calculated, and the value of QAI can quantitatively describe the degree of tool wear.

4. Experimental Dataset, Results, and Discussion

4.1. Milling Dataset. In order to evaluate whether the assessment method is effective, the Milling Dataset [19], which is developed and provided by the NASA AMES and UC Berkeley, is used. The dataset contains 16 cases in which the milling parameters were varied for each case. In this paper, Case 1 and Case 5 are selected to research, and the experimental conditions of Case 1 and Case 5 are shown in Table 1.

Three different types of sensors, which include the acoustic emission sensor (WD925, PHYSICAL ACOUSTIC GROUP), vibration sensor(7201-50, ENDEVCO), and current sensor(OMRON K3TB-A1015 and CTA 213), were acquired at several positions. An acoustic emission sensor and a vibration sensor are each mounted to the table and the spindle of the machining center. The signals from all sensors are amplified and filtered and then fed through two RMS before they enter the computer for data acquisition. The signal from a spindle motor current sensor is fed into the computer without further processing. A 70 mm face mill with 6 inserts was chosen as the tool. The type of inserts is KC710 which is coated with multiple layers of titanium carbide, titanium carbonitride, and titanium nitride in sequence. Generally, the flank wear VB is used for evaluating tool wear. In experiments, VB is measured as the distance from the cutting edge to the end of the abrasive wear on the flank face of the tool. After the insert was taken out of the tool, the wear was measured by using a microscope. But, VB was measured between runs at irregular intervals and was not always measured. The description of the dataset is shown in Table 2.

Taking first dataset of Case 1 as an example, the signals of AC, DC, vibration table, vibration spindle, AE table, and AE spindle are shown in Figure 4. It can be found that there are three phases in the signal, which include the starting phase, operating phase, and stopping phase. Because we only want to focus on the tool wear during the operation phase, the data of the starting phase and stopping phase should be removed firstly.

Now, there are six features. However, the proposed new assessment indictor QAI only can be calculated in two-

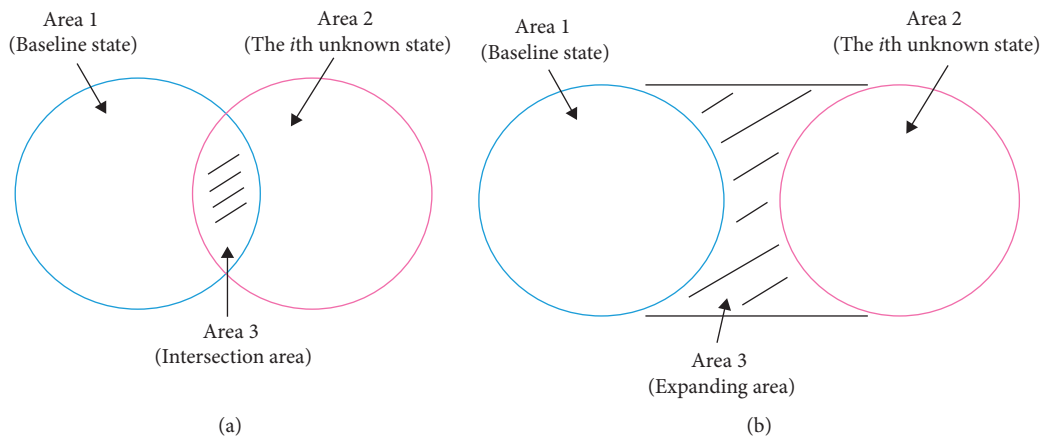


Figure 1: Schematic diagram of the assessment indicator. (a) Intersection situation. (b) Expanding situation.

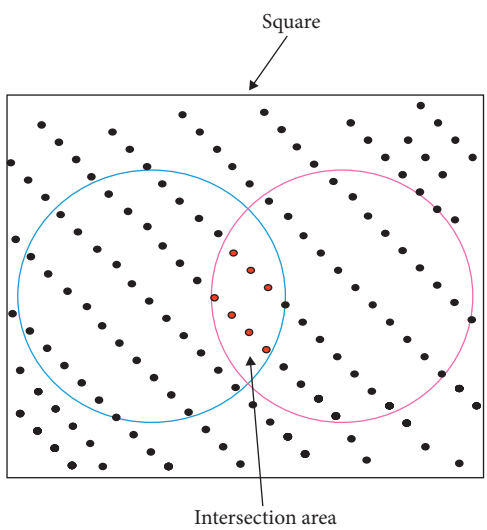


Figure 2: Schematic diagram of estimating the intersection area by using the Monte Carlo Method.

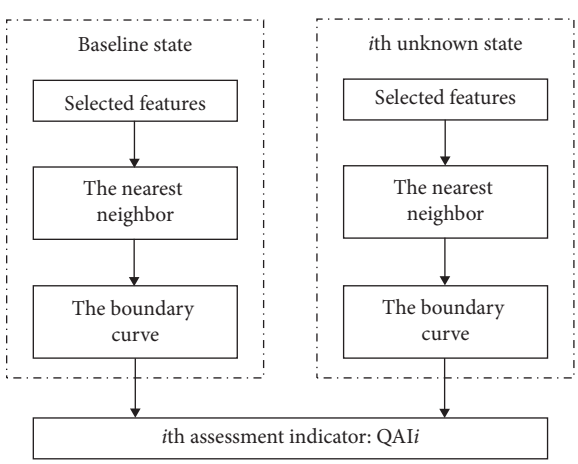


Figure 3: Flow chart of the assessment process.

dimension space. Therefore, two sensitive features should be selected secondly. AE spindle and vibration spindle as the sensitive features are selected by using the neighborhood rough set model. Taking the first dataset and seventh dataset of Case 1 as an example, multiboundary models are trained by using the nearest neighbor model with different pair of features, respectively, as shown in Figures 5 and 6. Comparing Figure 5(a) with Figure 6(a), it can be found that the value range of AE spindle and vibration spindle increases with time, implying that AE spindle and vibration spindle could more accurately reflect the state of tool wear. Comparing Figure 5(b) with Figure 6(b), it can be found that the value range of AE table and vibration table barely increases with time. Besides, the dispersion degree of the boundary model which uses AE table and vibration table is greater than that of the boundary model which uses AE spindle and vibration spindle. Moreover, it can be found that the distortion phenomenon for DC is as shown in Figure 6(c).

Table 1: Experimental conditions of Case 1 and Case 5 of the Milling Dataset.

Case	Depth of cut (mm)	Feed (mm/rev)	Material
1	1.5	0.5	Cast iron
5	1.5	0.5	Stainless steel J45

Taking Case 1 and Case 5 as an example, we assess the tool wear in process by using the method which is proposed in Section 3. The parameters of the nearest neighbor model are shown in Table 3.

4.2. The Assessment Result of Case 1. We choose AE spindle and vibration spindle as sensitive features to train the boundary model, and then, two different boundary curves are compared to find the intersection area or expanding area. Partial results are shown in Figures 7–12. Based on the intersection area or expanding area, we could calculate the assessment indicator QAI, as shown in Figure 13. The VB which was measured in experiment is shown in Figure 14.

4.3. The Assessment Result of Case 5. We compare two different boundary curves to find the intersection area or expanding area, as shown in Figures 15–19. Based on the intersection area or expanding area, we could calculate the

Table 2: Description of the dataset.

Name	Description
VB	Flank wear, measured after runs. Measures were not taken after each test
AC	AC spindle motor current
DC	DC spindle motor current
Vibration table	RMS of vibration at table
Vibration spindle	RMS of vibration at spindle
AE table	RMS of acoustic emission at table
AE spindle	RMS of acoustic emission at spindle

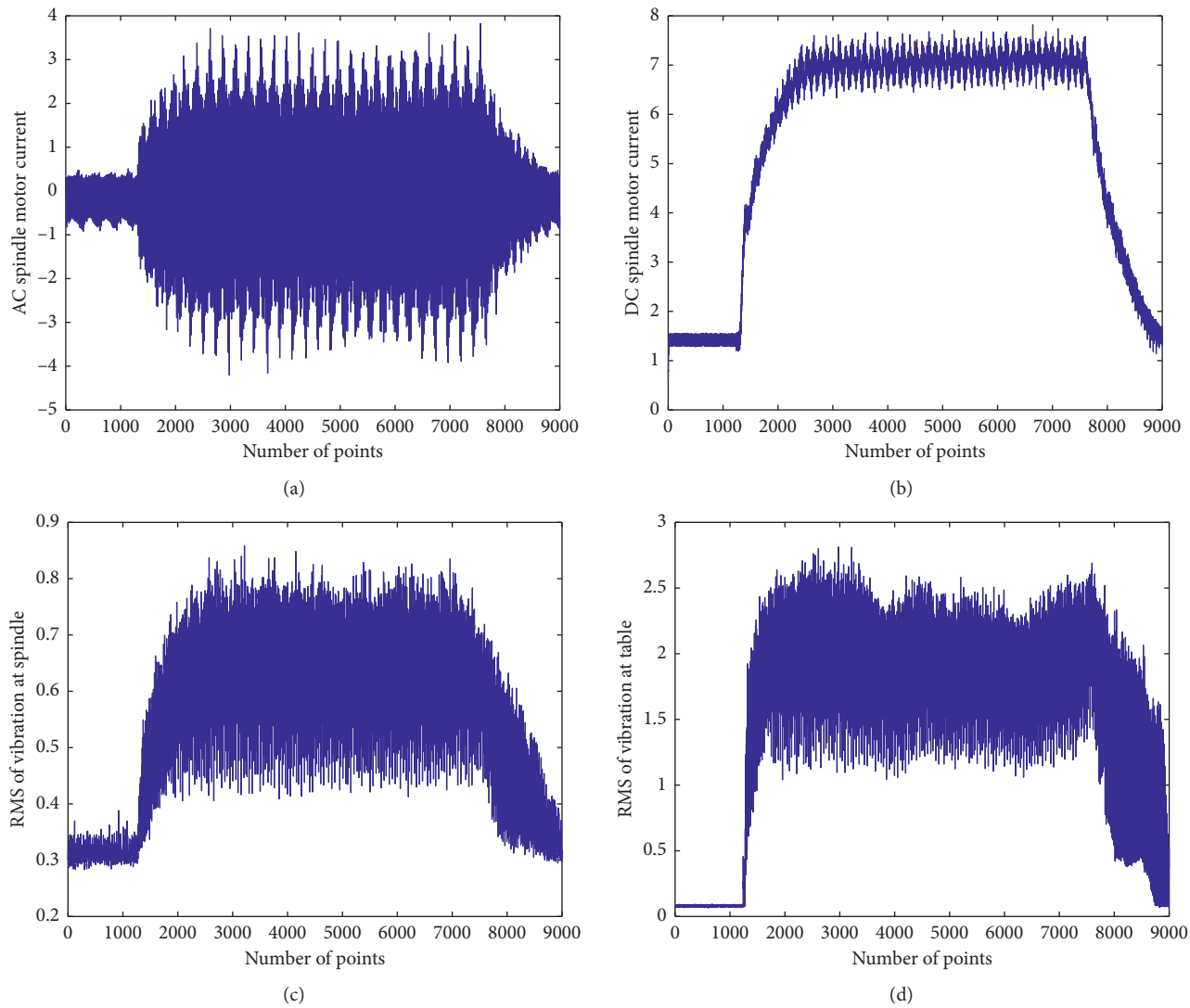


Figure 4: Continued.

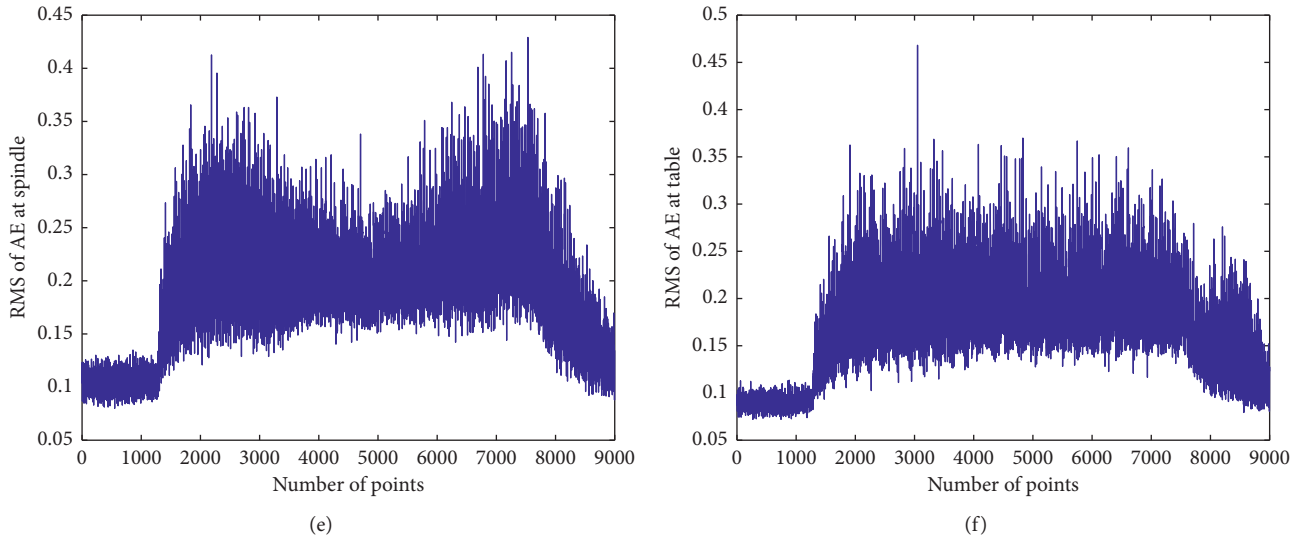


Figure 4: First dataset of Case 1 of the Milling Dataset. (a) AC spindle motor current. (b) DC spindle motor current. (c) RMS of vibration at spindle. (d) RMS of vibration at table. (e) RMS of AE at spindle. (f) RMS of AE at table.

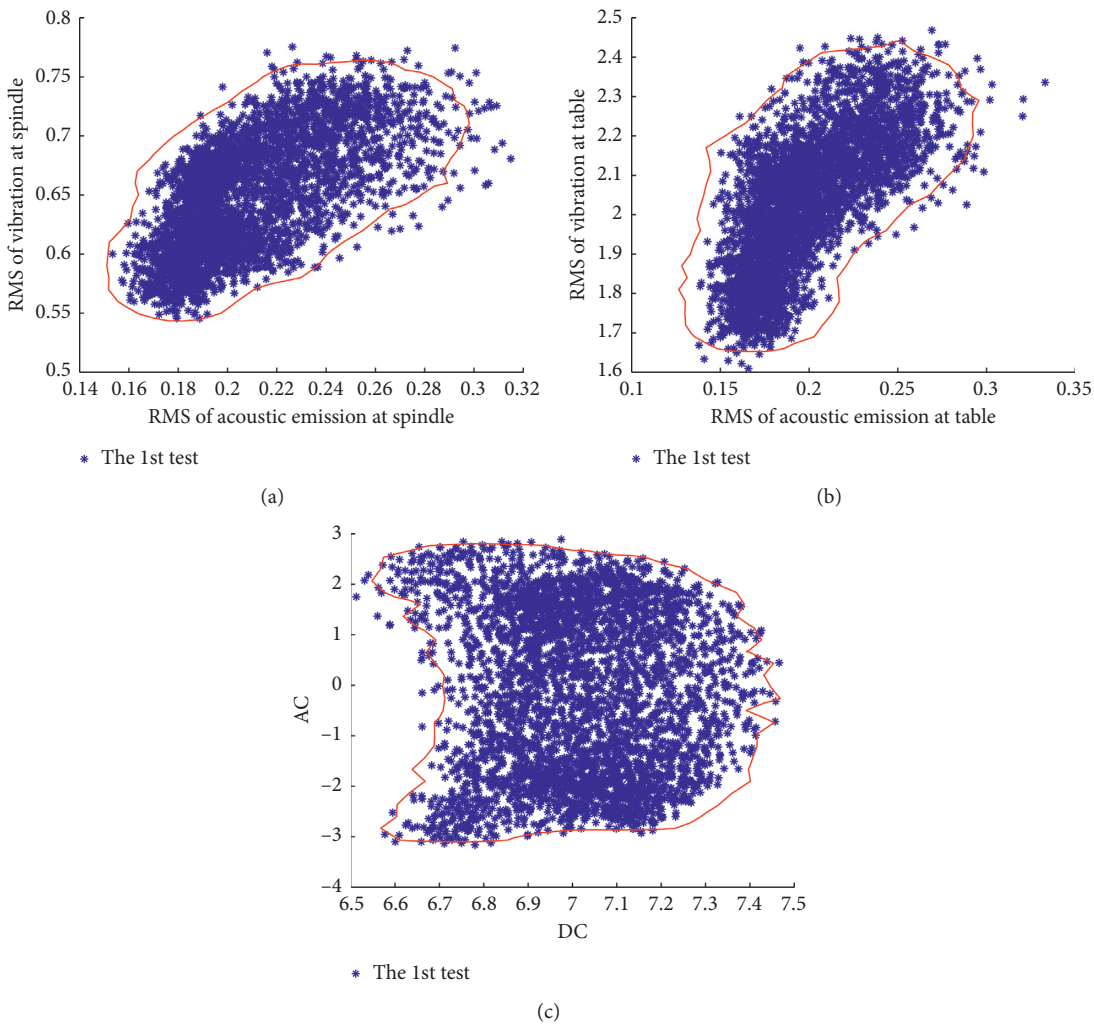


Figure 5: Different boundary curves of the first dataset of Case 1. (a) Features of the boundary curve are AE spindle and vibration spindle. (b) Features of the boundary curve are AE table and vibration table. (c) Features of the boundary curve are DC and AC.

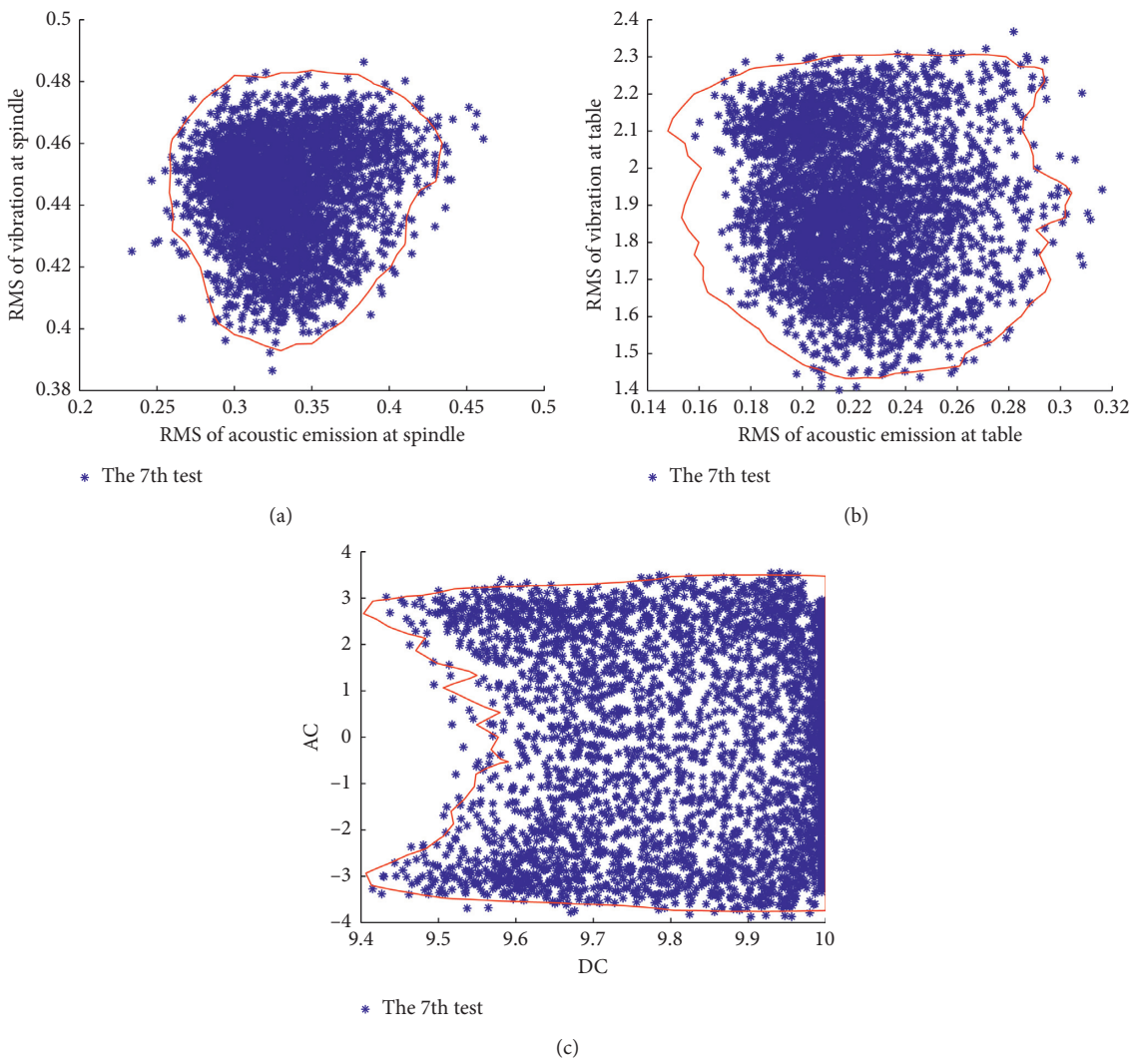


Figure 6: Different boundary curves of the seventh dataset of Case 1. (a) Features of the boundary curve are AE spindle and vibration spindle. (b) Features of the boundary curve are AE table and vibration table. (c) Features of the boundary curve are DC and AC.

Table 3: The parameters of the nearest neighbor model.

Features	The number of training samples	k nearest neighbor	Distance
AE spindle vibration spindle	3500	15	Euclidean distance

assessment indicator QAI, as shown in Figure 20. The VB which was measured in experiment is shown in Figure 21.

5. Discussion

In order to evaluate the degree of tool wear quantitatively, two new assessment indicators QAI are calculated for Case 1 and Case 2, as shown in Figures 13 and 20. Comparing with the corresponding VB which are shown in Figures 14 and 21, it can be found that the trend of QAI is basically consistent with VB.

For Case 1, the wear condition of tool could be divided into three parts. At the beginning stage which is from the 1st test to 6th test, the abrasion wear occurred because the insert should remove large chunks, and the wear of tool is very serious. The second stage is from the 7th test to 15th test; although the wear of tool is still increasing, the degree of wear slows down because the tool is in stable running condition. The third stage is from the 16th test to 17th test, and the tool is in accelerated wear condition.

No matter Case 1 or Case 5, it can be found that the test models gradually go away from the baseline mode with the increase in time. The intersection areas gradually decrease, and then, the expanding areas gradually increase. Moreover, except the changing about distance, the angle between the baseline model and test model is changing too. The angle informant is implicit in the

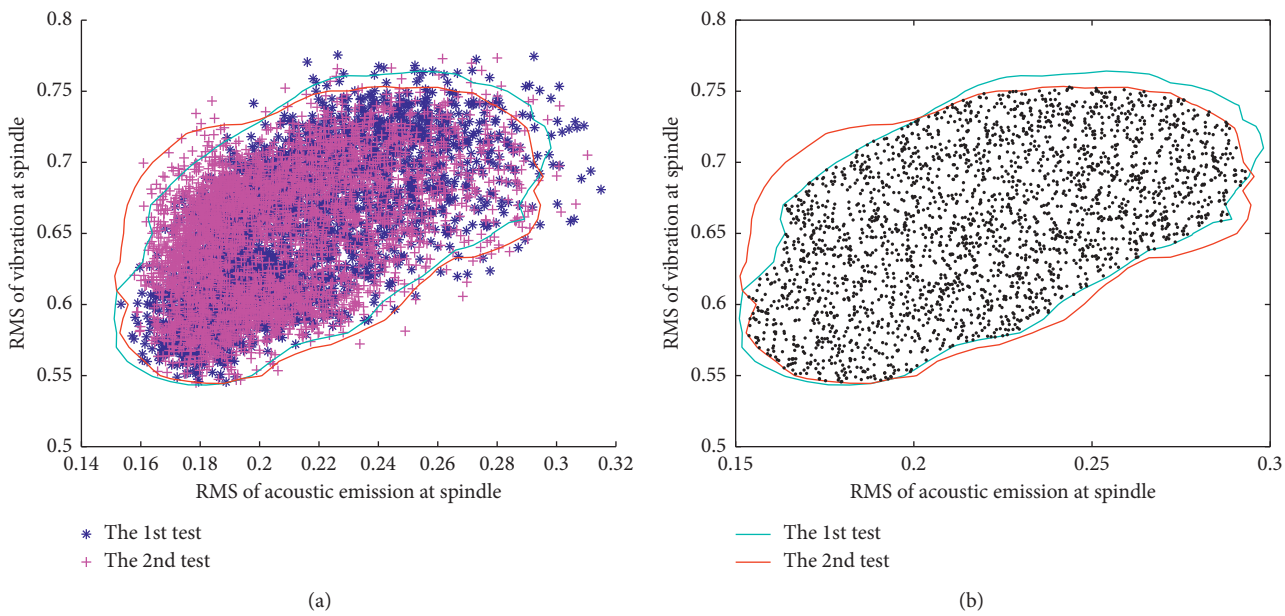


Figure 7: The 1st test vs. 2nd test for Case 1. (a) Two boundary models. (b) The intersection area.

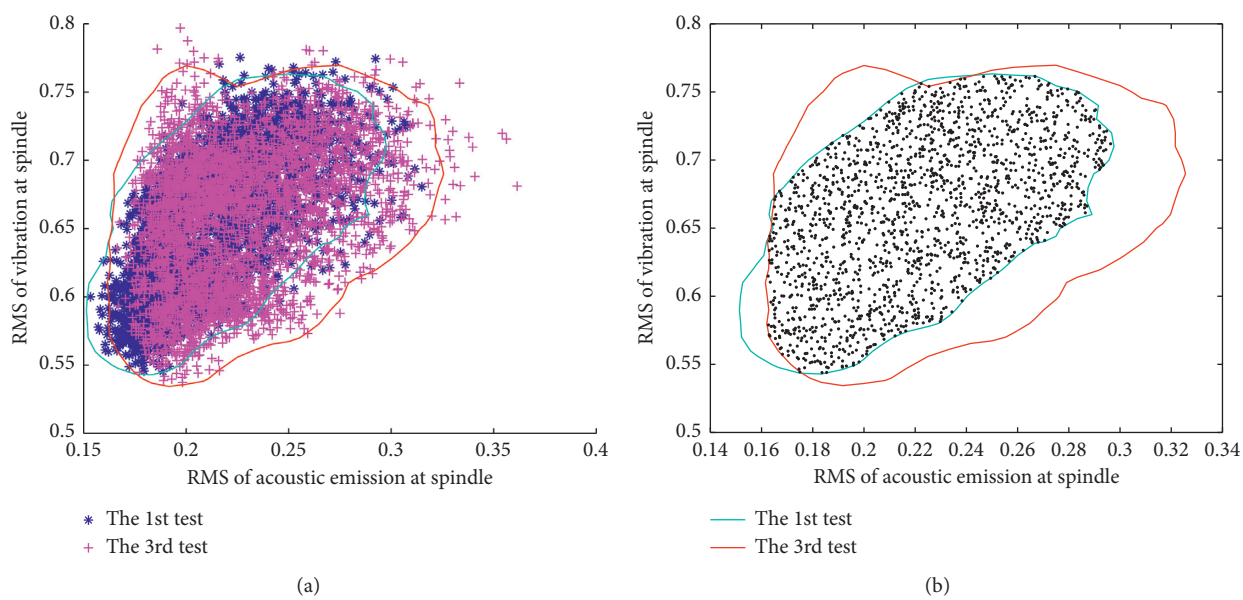


Figure 8: The 1st test vs. 3rd test for Case 1. (a) Two boundary models. (b) The intersection area.

intersection area or expanding area. For example, the comparing boundary curve of the 6th test with the boundary curve of the 5th test for Case 5 in Figures 18 and 19, it could be found that the boundary curve of the 6th test only expands in the X-direction because only the value of RMS of AE spindle changed significantly, whereas the value of RMS of vibration spindle almost remained constant. This phenomenon could be considered that the angles between two boundary curves are different. Mapping the angles into the expanding areas, the two expanding areas are different. Therefore, the intersection area or expanding area can reflect the effect of distance and angle simultaneously.

In order to evaluate the effect of different features, two new QAIs are calculated, as shown in Figures 22 and 23. One is using AE table and vibration table as features, and the other is using AC and DC as features. Comparing with Figure 20, it could be found that the trend of two QAIs is not good; especially as shown in Figure 22, the predicted degree of tool wear is not true.

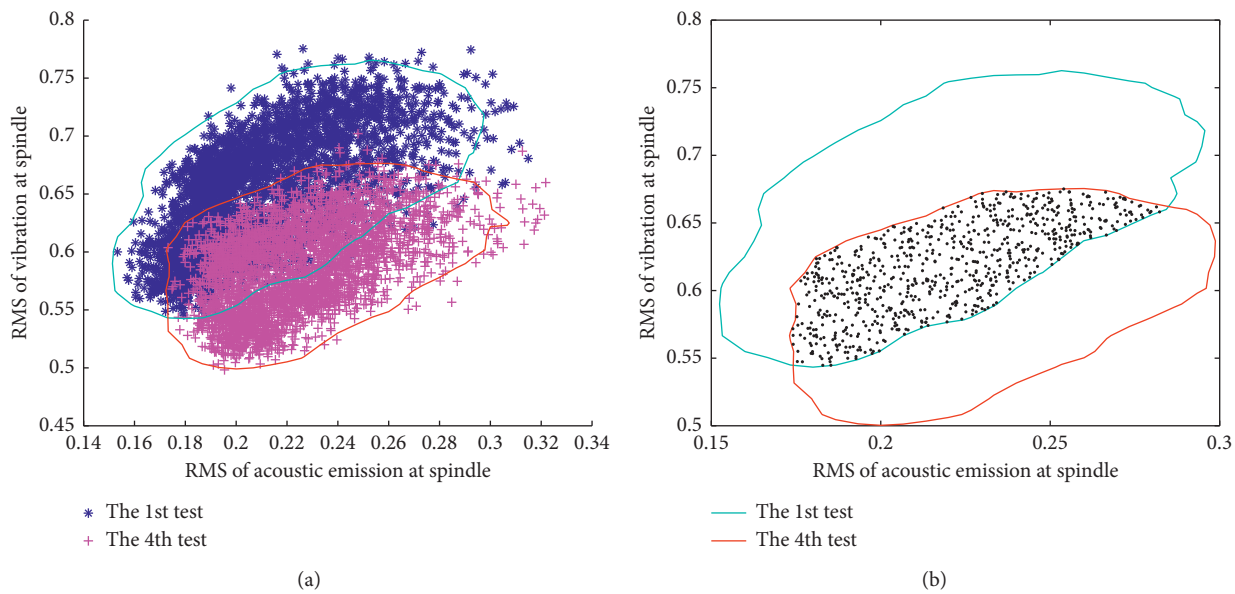


Figure 9: The 1st test vs. 4th test for Case 1. (a) Two boundary models. (b) The intersection area.

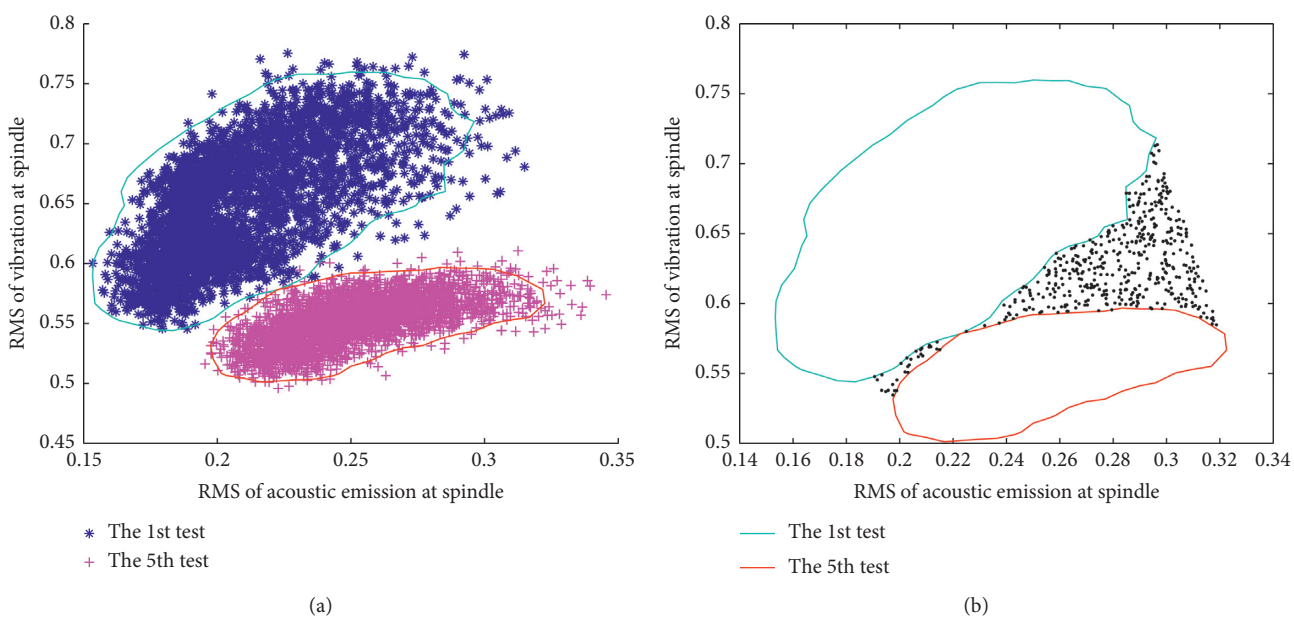


Figure 10: The 1st test vs. 5th test for Case 1. (a) Two boundary models. (b) The expanding area.

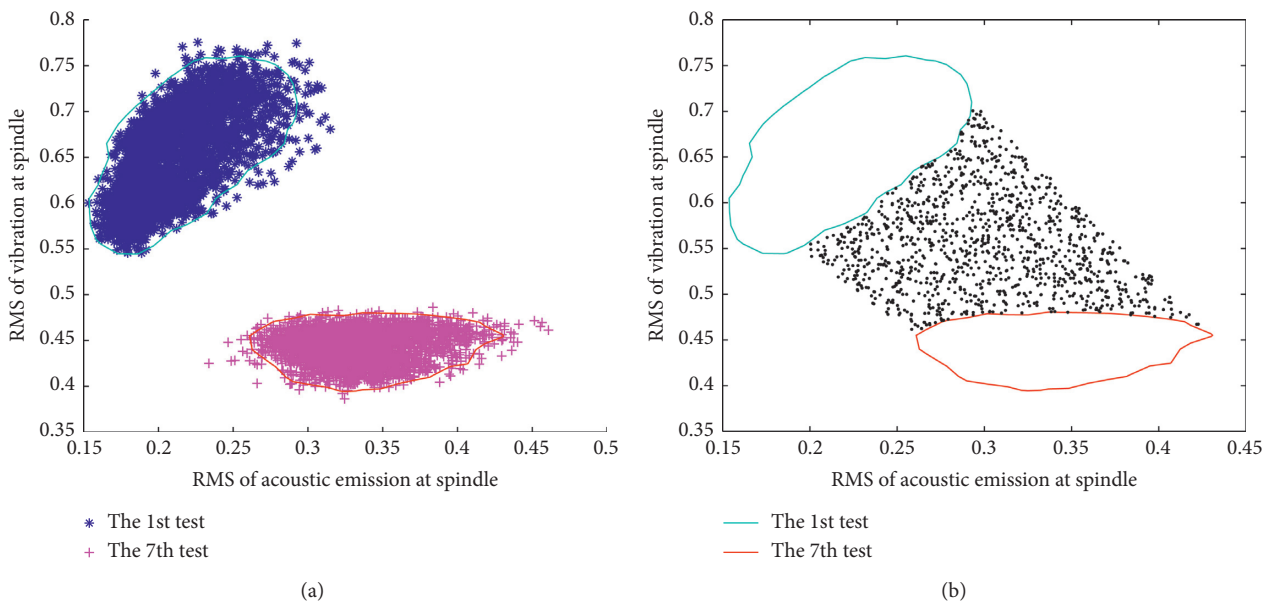


Figure 11: The 1st test vs. 7th test for Case 1. (a) Two boundary models. (b) The expanding area.

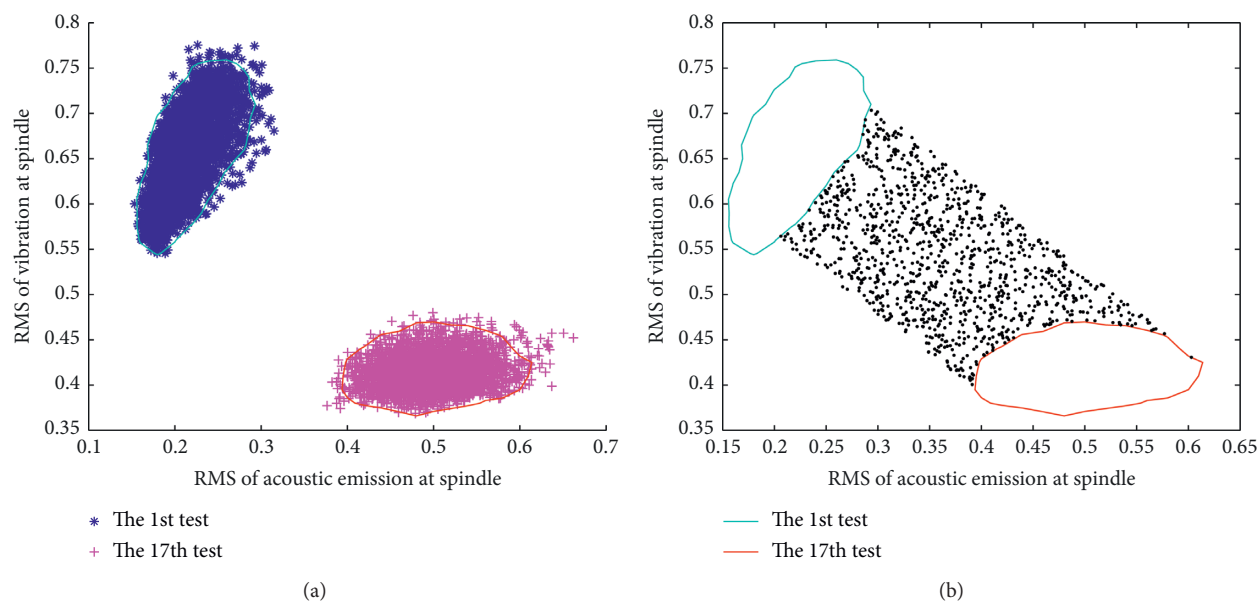


Figure 12: The 1st test vs. 17th test for Case 1. (a) Two boundary models. (b) The expanding area.

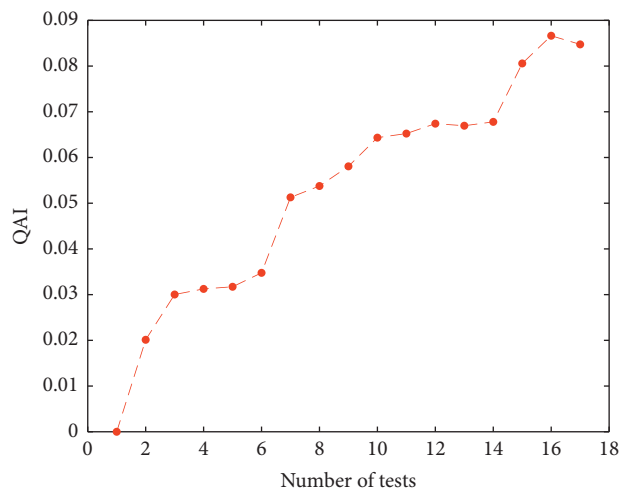


Figure 13: The QAI changing with the number of test for Case 1 (features are AE spindle and vibration spindle).

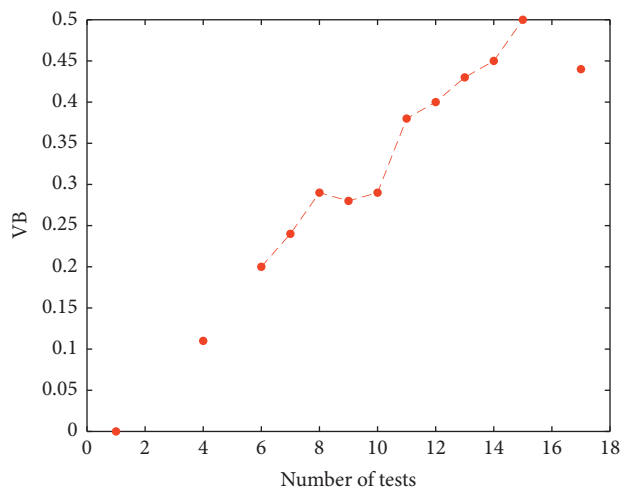


Figure 14: The VB changing with the number of test (missing values of 2nd, 3rd, 5th, and 16th test) for Case 1.

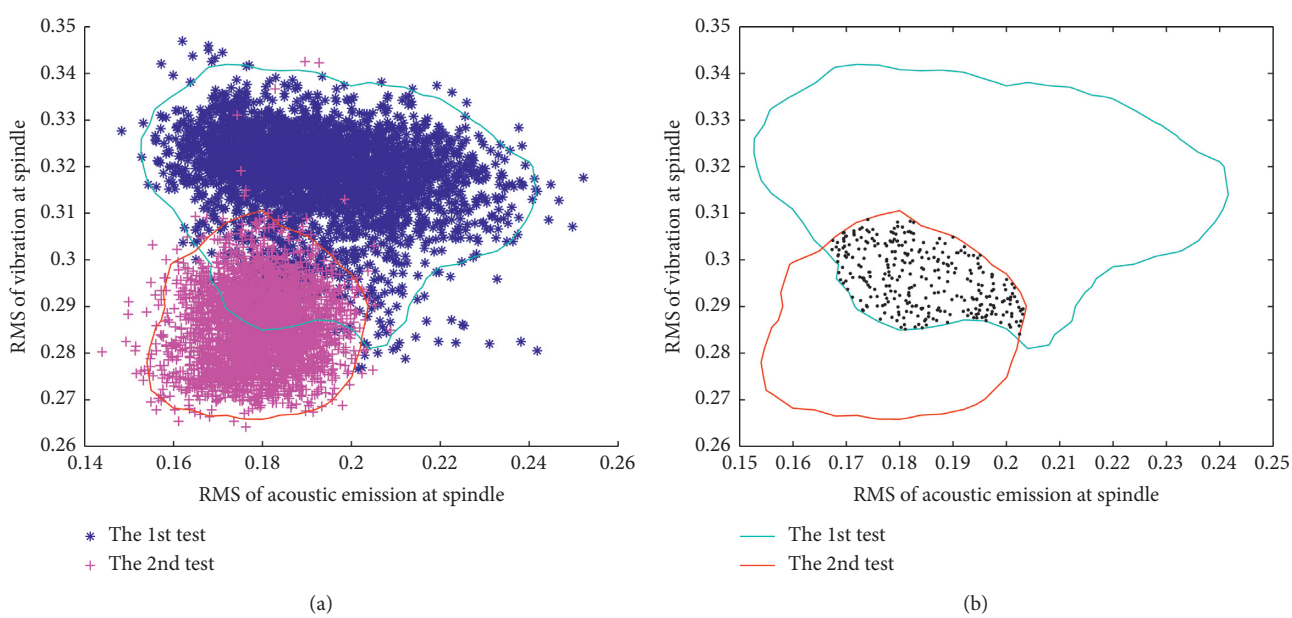


Figure 15: The 1st test vs. 2nd test for Case 5. (a) Two boundary models. (b) The intersection area.

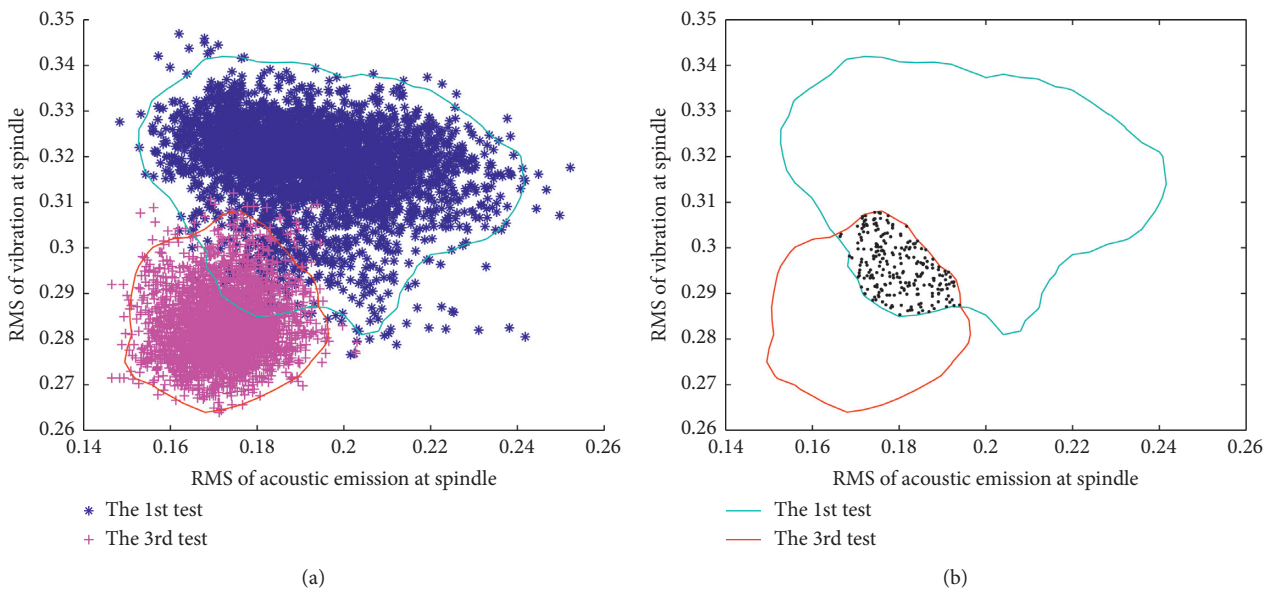


Figure 16: The 1st test vs. 3rd test for Case 5. (a) Two boundary models. (b) The intersection area.

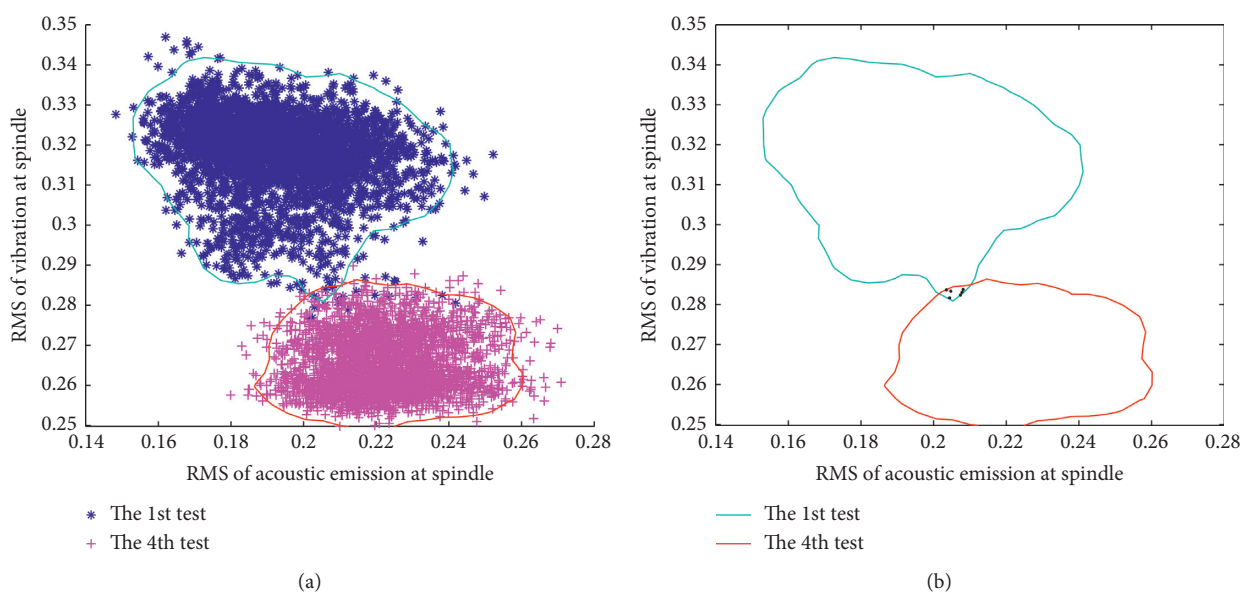


Figure 17: The 1st test vs. 4th test for case 5. (a) Two boundary models. (b) The intersection area.

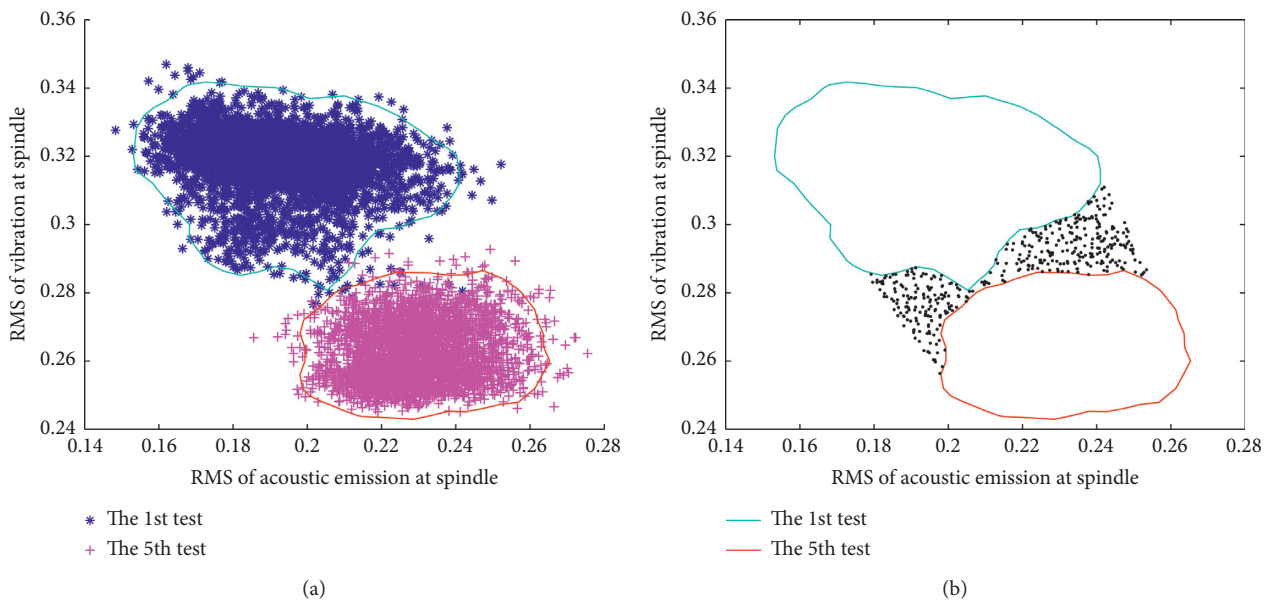


Figure 18: The 1st test vs. 5th test for Case 5. (a) Two boundary models. (b) The expanding area.

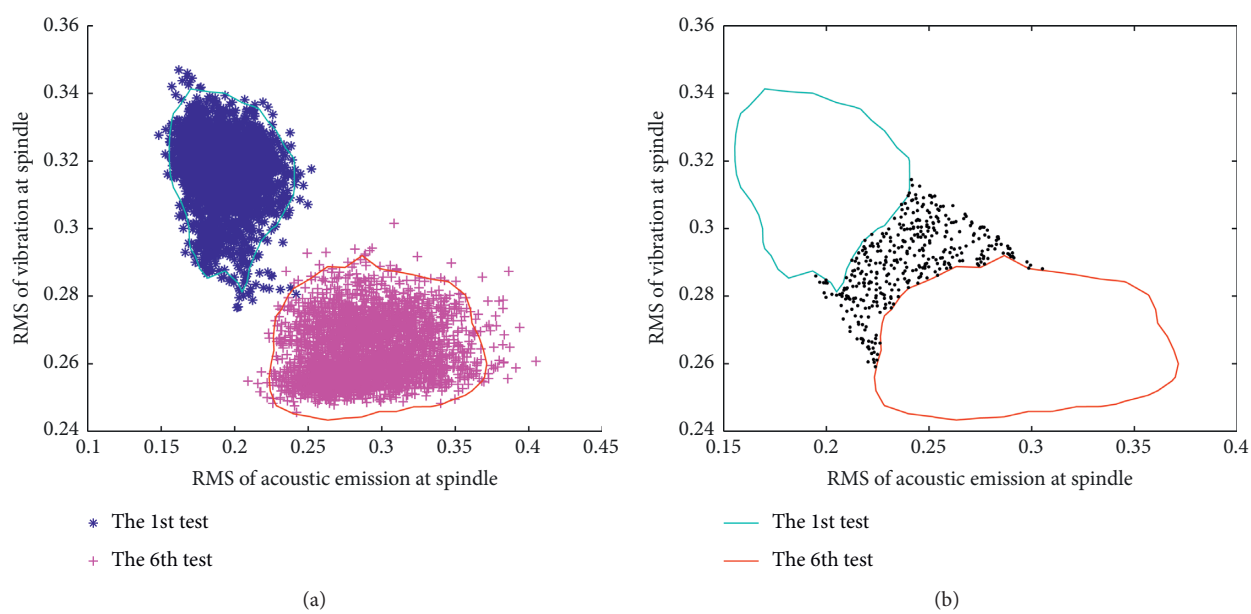


Figure 19: The 1st test vs. 6th test for Case 5. (a) Two boundary models. (b) The expanding area.

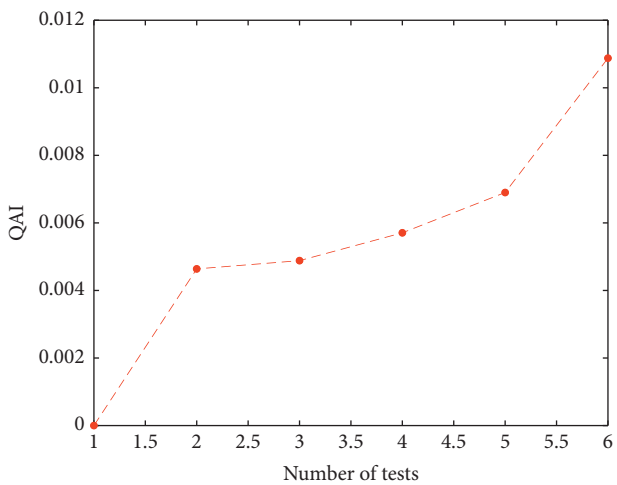


Figure 20: The QAI changing with the number of test for Case 5 (features are AE spindle and vibration spindle).

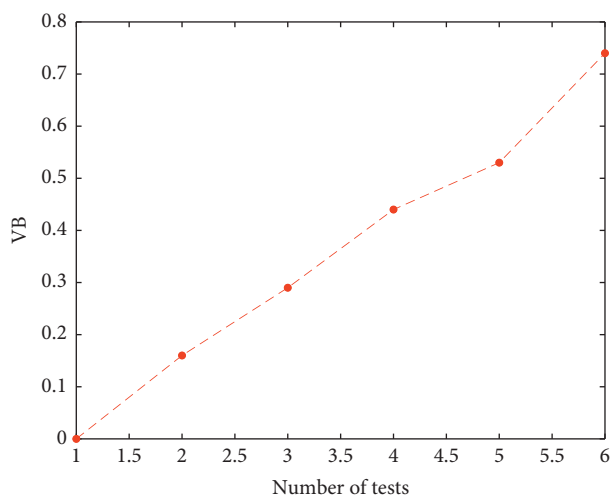


Figure 21: The VB changing with the number of test for Case 5.

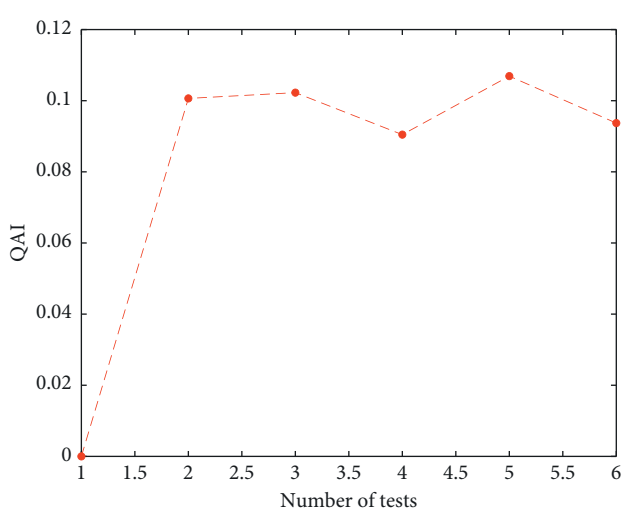


Figure 22: The QAI changing with the number of test for Case 5 (features are AE table and vibration table).

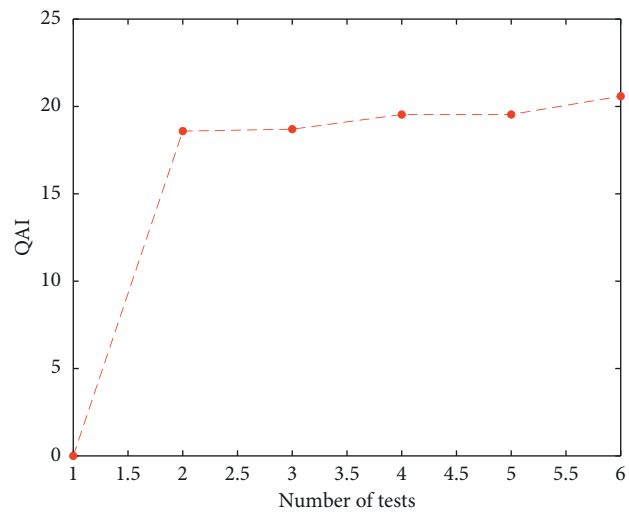


Figure 23: The QAI changing with the number of test for Case 5 (features are AC and DC).

6. Conclusions

After selecting two sensitive features by using the neighborhood rough set model, the boundary curves can be obtained based on the nearest neighbor model. And then, a new assessment process and a quantitative assessment indicator QAI are proposed. The main conclusions includethe following: (1) The boundary curve, which is trained based on selected features, can clearly describe the running state in two-dimension spaces. (2) The intersection area or expanding area can comprehensively consider the effect of distance and angle between the baseline curve and unknown curve. (3) The QAI can quantitatively describe the degree of wear of tool. (4) Unlike with the traditional indicator VB which only is measured offline when we disassemble the inserts from the tool, it is easy to monitor the tool wear in process using the QAI.

Data Availability

The previously reported Milling Dataset was used to support this study and is available at the NASA Ames Prognostics Data Repository. The prior studies (and datasets) are cited at relevant places within the text as references: A. Agogino and K. Goebel (2007), BEST lab, UC Berkeley. "Milling Dataset," NASA Ames Prognostics Data Repository (http://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/), NASA Ames Research Center, Moffett Field, CA.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors thank UC Berkeley and the NASA Ames Prognostic Data Repository for providing the Milling Dataset.

References

[1] Y. Zhou and W. Xue, "Review of tool condition monitoring methods in milling processes," *The International Journal of Advanced Manufacturing Technology*, vol. 96, no. 5–8, pp. 2509–2523, 2018.

[2] N. A. Prashant Waydande and S. Chinchanikar, "A review on tool wear monitoring system," *Journal of Mechanical Engineering and Automation*, vol. 6, no. 5A, pp. 49–53, 2016.

[3] Y.-C. Yen, J. Söhner, B. Lilly, and T. Altan, "Estimation of tool wear in orthogonal cutting using the finite element analysis," *Journal of Materials Processing Technology*, vol. 146, no. 1, pp. 82–91, 2004.

[4] K. Zhu and Y. Zhang, "A generic tool wear model and its application to force modeling and wear monitoring in high speed milling," *Mechanical Systems and Signal Processing*, vol. 115, pp. 147–161, 2019.

[5] M. Nouri, B. K. Fussell, B. L. Ziniti, and E. Linder, "Real-time tool wear monitoring in milling using a cutting condition independent method," *International Journal of Machine Tools and Manufacture*, vol. 89, p. 1, 2015.

[6] Ö. Eker, F. Camci, and I. K. Jennions, "Major challenges in prognostics: study on benchmarking prognostics datasets," in *Proceedings of the 1st European Conference of the Prognostics and Health Management Society*, pp. 148–145, Nantes, France, July 2012.

[7] S. Shankar, T. Mohanraj, and R. Rajasekar, "Prediction of cutting tool wear during milling process using artificial intelligence techniques," *International Journal of Computer Integrated Manufacturing*, vol. 32, no. 2, pp. 174–182, 2018.

[8] T. Kalvoda and Y.-R. Hwang, "A cutter tool monitoring in machining process using Hilbert-Huang transform," *International Journal of Machine Tools and Manufacture*, vol. 50, no. 5, pp. 495–501, 2010.

[9] S. Leng, Z. Wang, T. Min, Z. Dai, and G. Chen, "Detection of tool wear in drilling CFRP/TC4 stacks by acoustic emission," *Journal of Vibration Engineering & Technologies*, vol. 8, no. 3, pp. 463–470, 2019.

[10] W. Sun, M. Huang, Y. He, and K. Li, "Design of tool-state monitoring system based on current method," *The Journal of Engineering*, vol. 2019, no. 23, pp. 9026–9030, 2019.

[11] A. Proteau, A. Tahan, and M. Thomas, "Specific cutting energy: a physical measurement for representing tool wear," *The International Journal of Advanced Manufacturing Technology*, vol. 103, no. 1–4, pp. 101–110, 2019.

[12] X. Jin, Z. Que, Y. Sun, Y. Guo, and W. Qiao, "A data-driven approach for bearing fault prognostics," *IEEE Transactions on Industry Applications*, vol. 55, no. 4, pp. 3394–3401, 2019.

[13] R. A. Ariyaluran Habeeb, F. Nasaruddin, A. Gani, I. A. Targio Hashem, E. Ahmed, and M. Imran, "Real-time big data processing for anomaly detection: a survey," *International Journal of Information Management*, vol. 45, pp. 289–307, 2019.

[14] D. Tax, *One-Class Classification; Concept-Learning in the Absence Of Counter-examples*, Technische Universiteit Delft, Delft, Netherlands, 2001.

[15] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[16] Z. Pawlak, "Rough set theory and its applications to data analysis," *Cybernetics and Systems*, vol. 29, no. 7, pp. 661–688, 1998.

[17] Q. Hu, D. Yu, and Z. Xie, "Neighborhood classifiers," *Expert Systems with Applications*, vol. 34, no. 2, pp. 866–876, 2008.

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

[19] A. Agogino and K. Goebel, *BEST Lab, UC Berkeley, "Milling Data Set," NASA Ames Prognostics Data Repository*, NASA Ames Research Center, Moffett Field, CA, USA, 2007.

Hindawi
Shock and Vibration
Volume 2020, Article ID 8831723, 14 pages
https://doi.org/10.1155/2020/8831723

Research Article

Application of Feature Fusion Using Coaxial Vibration Signal for Diagnosis of Rolling Element Bearings

Jing Jiao, Jianhai Yue, Di Pei, and Zhunqing Hu

School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Jing Jiao; jingjiao@bjtu.edu.cn

Received 22 July 2020; Revised 5 September 2020; Accepted 16 September 2020; Published 1 October 2020

Academic Editor: Baoqing Li

The research of rolling element bearings (REBs) fault diagnosis based on single sensor vibration signal analysis is very common. However, the information provided by an individual sensor is very limited, and the robustness of the system is poor. In this paper, a novel fault diagnosis method based on coaxial vibration signal feature fusion (CVSFF) is proposed to fully analyze the multisensor information of the system and build a more reliable diagnostic system. An ensemble empirical mode decomposition (EEMD) method is used to decompose the original vibration signal into a number of intrinsic mode functions (IMFs). Then the autocorrelation analysis is introduced to reduce the random noise remaining in IMFs. After that, the Rényi entropy is calculated as the feature of bearings. Finally, the features of coaxial vibration signal are fused by a multiple-kernel learning support vector machine (MKL-SVM) to classify bearing conditions. In order to verify the effectiveness of the CVSFF method in REB diagnosis, eight data sets from the Case Western Reserve University Bearing Data Center are selected. The fault classification results demonstrate that the proposed approach is a valuable tool for bearing faults detection, and the fused feature from coaxial sensors improves fault classification accuracy for REBs.

1. Introduction

The state of bearing is very important to efficient operation of mechanical equipment. Fault bearing will cause periodic impacts on vibration, which will lead to problems of other parts of mechanical system. Therefore, it is of great significance to find out bearing fault in time and replace bearing to avoid the breakdown of machine.

With the development of information technology, judging bearing state by signal analysis has become an important trend of condition-based monitoring (CBM) [1]. In recent years, many methods have been applied to fault detection of REBs, such as vibration signal, oil, temperature, and acoustic emission analysis. Among all these methods, vibration signal analysis is the most widely used and effective method owing to the great information that vibration signal contains [2, 3]. If bearing is operating with local spall, it will cause vibration impulse [4]. Through analysis of the collected vibration signals, the fault impact characteristics can be obtained, so as to realize bearing fault diagnosis. In the research of bearing fault diagnosis based on vibration signals, many papers focus on the signal collected from a single sensor. However, for the complex mechanical system, it is uncertain to use the information of an individual sensor for fault diagnosis, which will lead to unreliable diagnosis results in many cases.

Information fusion is a technology that merges data to obtain more consistent, informative, and accurate information than the original raw data that are mostly uncertain [5]. Some scholars have made achievements in bearing fault diagnosis based on multisensor information fusion technology. These research studies are mainly divided into multisource data fusion relating to the data from the same kind of sensors and different kinds of sensors. Owing to the advantages of vibration signals, many scholars adopt the method of multiple vibration sensor signal fusion in bearing fault diagnosis. These signal fusion studies can be classified as data-level fusion, feature-level fusion, and decision-level fusion.

Yan et al. [6] proposed a concept of space-time fragments, and vibration signals captured by multiple sensors are fused at data level. For feature-level fusion, Jiang et al. [7] developed a feature-level information fusion methodology of fault diagnosis in rotating machinery. Tao et al. [8] presented a novel bearing fault diagnosis method using multivibration signals and deep belief network (DBN). Banerjee and Das [9] suggested a hybrid model for an on board fault-tolerant control system by vibration data fusion. Wang et al. [10] fused vibration data features as the health indicator of bearing status and got a good representation of bearing defect conditions. Zhou et al. [11] addressed a feature fusion approach based on NCA and coupled hidden Markov model. Hong et al. [12] introduced a preprocessing model of bearing using wavelet packet-empirical mode decomposition (WP-EMD) for feature extraction of vertical and horizontal vibrations. For decision-level fusion, Hui et al. [13] proposed an automated bearing fault diagnosis model that employs SVM and the Dempster–Shafer evidence theory in classification.

In the aspect of different kinds of signals fusion, some scholars have carried out bearing fault diagnosis based on the fusion of vibration signals and other types of signals. Safizadeh and Latifi [14] presented a new method of bearing fault diagnosis using the fusion of two primary sensors: an accelerometer and a load cell. Lu et al. [15] designed a sound-aided vibration signal adaptive stochastic resonance (SAVASR) method for bearing fault detection. All those works indicate that multisensor information fusion methods have higher classification accuracy than single sensor data analysis in bearing fault diagnosis.

Nevertheless, there are still some problems in vibration signal source of the above bearing fault diagnosis research studies. The authors of [6, 9, 10] do not specify the installation location of vibration sensors. The experiments of papers [7, 9] have too many sensors, which make the system more complex. The researchers [11, 12, 14, 15] only collect vibration signals of one or multidirections of a single bearing. In the experiment Case 2 of [11], vibration sensors are only installed on the base of the test rig. As a matter of fact, the location of vibration sensor is very important in fault diagnosis of mechanical equipment. Vibration data are composed of multiple vibration source signals and noise signals. It must pass through multiple interfaces to reach an accelerometer, which will definitely cause energy dissipation. Because a shaft and bearing inner race are rigidly connected, the shaft plays a role of vibration transmission between coaxial bearings. The fault signals of coaxial bearings are usually similar in a frequency domain. Therefore, the coaxial vibration signal feature fusion (CVSFF) algorithm proposed in this paper takes vibration sensors of coaxial bearings as data sources. Enough information about REBs can be obtained from limited multisource data. Then, the characteristics of this information are fused to realize bearing fault diagnosis.

In the research of bearing fault diagnosis based on information fusion technology, feature-level fusion is widely used. Among the existing algorithms, a support vector machine (SVM) has been widely used because of its good classification performance. The researchers of papers [7, 9, 13] used SVM as a classifier for fault classification and have achieved good results. However, the classification performance of SVM is greatly affected by a kernel function. The determination of kernel function depends on human experience [16]. To solve this problem, multiple kernel learning (MKL) methods have been proposed. MKL learns the kernel function and classifier parameters simultaneously, which can effectively solve the problem of kernel function selection. Meanwhile, the SVM trained by MKL has more flexibility and higher classification accuracy. Many scholars have applied MKL to SVM model optimization and obtained good results [16, 17].

Based on the abovementioned analysis, a new method of fault diagnosis based on CVSFF is presented in this paper. First, an EEMD method is used to decompose the original vibration signals collected from bearings at both ends of a shaft. Then, autocorrelation is carried out to reduce the random noise in IMF components. After that, an energy ratio of each IMF component is calculated to extract probability mass function (PMF). The Rényi entropy feature matrix of coaxial sensors is obtained based on PMF. Finally, different states of REBs are classified by MKL-SVM.

This paper is organized as follows: the proposed method of CVSFF is presented in the next section. In Section 3, the algorithm is validated and analyzed by experimental data. Meanwhile, MKL-SVM classification based on an individual sensor and SVM, genetic algorithm-optimized SVM (GA-SVM), and particle swarm optimization SVM (PSO-SVM) based on coaxial features are carried out to evaluate the effectiveness of CVSFF. Section 4 presents the conclusion.

2. Methods

2.1. EEMD. There are many time-frequency data analysis methods to decompose nonlinear and nonstationary time series into a set of components, such as empirical mode decomposition (EMD) [18], ensemble empirical mode decomposition (EEMD) [19], variational mode decomposition (VMD) [20], and broadband mode decomposition (BMD) [21, 22]. Among all those methods, EEMD is the most widely used algorithm in CBM of REBs. EEMD can significantly improve the decomposition effect by reducing mode mixing. The two important parameters used in the EEMD algorithm are ratio k for standard deviation of white noise to standard deviation of signal and the total number M of EMD. Huang suggested that k is generally set as 0.2 [19]. Moreover, the amplitude of white noise should be reduced appropriately for the signal mainly composed of high-frequency components and increased for the signal mainly composed of low-frequency components. Besides, Huang found that M follows the statistical law of the following equation.

$$\varepsilon_n = \left(\frac{k}{\sqrt{M}}\right) \text{or} \ln \varepsilon_n + \left(\frac{k}{2}\right) \ln M = 0, \tag{1}$$

where ε_n is the maximum relative error of signal decomposition. In this paper, M is taken as 100, and k is taken as 0.2.

The specific decomposition steps of EEMD are as follows:

(1) Add random white noise $k\sigma_x n(t)$ to $x(t)$ as shown in the following equation, where $n(t)$ is the Gaussian white noise with mean value 0 and standard deviation 1 and σ_x is the standard deviation of signal $x(t)$:

$$x_m(t) = x(t) + k\sigma_x n_m(t). \tag{2}$$

(2) EMD decomposition of signals $x_m(t)$.

(3) Repeat Steps 1 and 2 for M times.

(4) The average value of each IMF component is obtained by decomposing M times EMD, and the global average is obtained:

$$\begin{aligned} \bar{c}_i(t) &= \frac{1}{M}\sum_{m=1}^{M} c_{i,m}\,(i = 1, 2, \ldots, n), \\ \bar{r}_n(t) &= \frac{1}{M}\sum_{m=1}^{M} r_{n,m}, \end{aligned} \tag{3}$$

where $\bar{c}_i(t)\,(i = 1, \ldots, n)$ and $\bar{r}_n(t)$ are the ith IMF and residual component, respectively.

2.2. Autocorrelation. An autocorrelation function describes the relationship of a signal at different times. The autocorrelation function of signal $x(t)$ is defined as

$$R_x(\tau) = \lim_{T\longrightarrow\infty}\int_0^T x(t)x(t+\tau)\mathrm{d}t. \tag{4}$$

According to the properties of autocorrelation function, a periodic signal has the same cycle as an original signal after autocorrelation. Furthermore, the autocorrelation function of random noise attenuates quickly and tends to zero with the increase of time delay τ. If the periodic signal contains random noise, the autocorrelation function can be used to reduce noise. In this paper, autocorrelation function is applied to noise reduction of IMFs, so as to retain the useful periodic signals in IMFs and reduce random white noise.

2.3. Rényi Entropy. Self-information $I(x_1)$ refers to the amount of information contained in an event x_i of a physical system. Different events in a physical system contain diverse amounts of information. So $I(x_1)$ is a random variable that cannot be used as a measure of information about the whole system.

Shannon [23] defines the mathematical expectation of information as information entropy, that is, the average amount of information in a source:

$$S(p) = E[-\log_2(p_i)] = -\sum_{i=1}^{N} p_i \log_2 p_i, \tag{5}$$

where p_i is the probability mass function (PMF).

Shannon entropy is a nonparametric measure, and it is well known that it does not have significant sensitivity when dealing with noisy data. For this reason, Rényi entropy [24] has been chosen as another potential entropy measure, which is defined as follows:

$$R(p) = \frac{1}{1-\alpha}\log_2\sum_{i=1}^{N} p_i^{\alpha-1}. \tag{6}$$

The parameter α in Rényi entropy can be used to make the entropy more or less sensitive to particular segments of the probability distributions. $\alpha \longrightarrow 0$ causes the Rényi entropy to become highly sensitive to changes in the tails of the distribution. And for $\alpha \longrightarrow 1$, it reduces to the Shannon entropy, and hence, the Rényi entropy becomes more sensitive in the regions where the bulk of PMF is located [25].

In order to obtain the frequency distribution change of vibration signal's energy for rolling bearing, p_i is the energy ratio of IMFs to total energy:

$$p_i = \left(\frac{E_i}{E}\right), \tag{7}$$

where E_i is the energy of IMF$_i$, $E_i = \sum_{j=1}^{N}(\bar{c}_{i,j}(t))^2$, $(i = 1, \ldots, n;\ j = 1, \ldots, N)$, n is the number of IMFs, N is the number of points of signal $\bar{c}_{i,j}(t)$, and $E = \sum_{i=1}^{n} E_i$.

The Rényi entropy is a parameter that characterizes the statistical properties of random variables and reflects the randomness of variables. The energy distribution of normal rolling bearing vibration signal in each frequency band is uniform. It means that the energy distribution is uncertain, so entropy is relatively large. When rolling bearing spall occurs, the energy is mainly distributed in the resonance frequency. The uncertainty of energy distribution is relatively reduced, so the entropy decreases. Therefore, the Rényi entropy is a sensitive feature for REBs classification.

2.4. MKL-SVM. SVM is a machine learning method based on statistical learning theory and structural risk minimization principle and is developed by Vapnik and his group [26]. The optimization classification function is as follows:

$$f(x) = \sum_{i=1}^{n} \alpha_i^* K(\mathbf{x}, \mathbf{x}_i) + b^*, \tag{8}$$

where $K(\cdot, \cdot)$ is the kernel function associated with a reproducing kernel Hilbert space (RKHS) H, $\mathbf{x}_i$ is the ith training data, $\mathbf{x}$ is the data to be classified, and α_i^* and b^* are the unknown coefficients.

There are many kinds of kernel functions, such as linear kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T\mathbf{x}_j)$; polynomial kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T\mathbf{x}_j)^d$; Gaussian radial basis kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2)$; and sigmoid kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\beta\mathbf{x}_i^T\mathbf{x}_j + \theta)$.

The kernel function and SVM model parameters directly affect the performance of the SVM classifier. However, it is difficult to map the sample into a high-dimensional feature space by using a single kernel function for complex classification problem. In recent years, many scholars have carried out relevant research on MKL-SVM. In order to select the appropriate kernel function, MKL-SVM learns by combining different kernel functions and thus has more flexibility, better generalization ability, and stronger model interpretation ability.

In MKL framework, $K(\cdot, \cdot)$ is a convex linear combination of a set of basic kernels:

$$K(\mathbf{x}, \mathbf{x}_i) = \sum_m d_m K_m(\mathbf{x}, \mathbf{x}_i), \quad \text{s.t. } d_m \geq 0, \sum_m d_m = 1, \tag{9}$$

where d_m is the weight of $K_m(\cdot, \cdot)$ obtained by sample learning. The decision function of MKL is as follows:

$$f(x) = \sum_m f_m(x) + b, \tag{10}$$

where each function f_m belongs to a different RKHS H_m associated with a kernel K_m. Rakotomamonjy et al. [27] proposed a simple MKL method to learn both the coefficients α_i^* and the weights d_m. They adopted the gradient method to solve the MKL problem. The optimal solution is obtained by calculating the gradient of the objective function about to d_m:

$$\begin{aligned} &\min_{\{f_m\}, b, \xi, d} \left(\frac{1}{2}\right) \sum_m \frac{1}{d_m} \|f_m\|_{H_m}^2 + C \sum_i \xi_i, \\ &\text{s.t. } y_i \sum_m f_m(x_i) + y_i b \geq 1 - \xi_i, \ \forall i, \\ &\xi_i \geq 0, \ \forall i, \\ &\sum_m d_m = 1, d_m \geq 0, \forall m. \end{aligned} \tag{11}$$

Finally, the decision functions of MKL-SVM are obtained as follows:

$$f(x) = \sum_i \alpha_i^* y_i \left[\sum_m d_m K_m(x, x_i) \right] + b^*. \tag{12}$$

2.5. Structure of CVSFF. The method process of CVSFF proposed in this paper is as follows:

Step 1: collecting signals from coaxial vibration sensors

Step 2: decompose vibration data through EEMD and obtain respective IMFs

Step 3: denoise the IMFs by the autocorrelation function

Step 4: extract the Rényi entropy of IMF components

Step 5: randomly select the training set and test set from feature matrix

Step 6: train MKL-SVMs

Step 7: input the test set into MKL-SVM models, and output the classification results

The method flowchart is shown in Figure 1.

3. Experiments Analysis and Discussion

3.1. Data Set. The bearing data diagnosed in this paper were obtained from the Case Western Reserve University (CWRU) Bearing Data Center. These data sets have been considered as a benchmark and analyzed by many researchers. As shown in Figure 2, the test rig consists of a 2 hp motor, a torque transducer, and a dynamometer.

Two coaxial accelerometers are installed at both ends of the motor to measure vibration signal of the corresponding bearing. In this paper, these two coaxial vibration signals in the drive end and fan end are used for fault diagnosis.

Smith and Randall [28] used three established bearing diagnostic techniques to provide a benchmark analysis of these widely used data sets. According to the data analysis conclusion in [28], data of eight states that are difficult to diagnose by the benchmark method were selected for analysis in this paper, as shown in Table 1.

The last column in Table 1 means categorisation of the benchmark method proposed in that paper. The explanation of diagnosis category is as follows:

Y1: data clearly diagnosable and showing classic characteristics for the given bearing fault in both the time and frequency domains

Y2: data clearly diagnosable but showing nonclassic characteristics in either or both of the time and frequency domains

P1: data probably diagnosable; e.g., the envelope spectrum shows discrete component at the expected fault frequencies, but they are not dominant in the spectrum

P2: data potentially diagnosable; e.g., the envelope spectrum shows smeared components that appear to coincide with the expected fault frequencies

N1: data not diagnosable for the specified bearing fault but with other identifiable problems (e.g., looseness)

N2: data not diagnosable and virtually indistinguishable from noise, with the possible exception of shaft harmonics in the envelope spectrum

Randall only used the benchmark method to analyze four sets of data with data set number 169, 170, 171, and 172 for inner ring fault of the drive end with 12 kHz sampling frequency. Data set 171 was selected from the four groups with poor diagnostic results. The reason for choosing data set 276 is the same as that for choosing data set 171. Except these two sets and normal bearing data, data sets in Table 1 are not diagnosable by the benchmark method in [28].

3.2. EEMD Analysis. Vibration signals of two coaxial sensors corresponding to eight states were cut into 1600 segments with ten seconds data length (120,000 points) for each state. Each state of bearing has 200 segments from two coaxial sensors (each sensor has 100 segments; each segment has 1200 points). Every segment of data is decomposed into nine IMFs and a residual component by EEMD. Drive-end raw time waveforms and part of decomposition results of eight different states are given in Figure 3. Only original signal and the first seven orders of IMFs are shown in the figure.

In the time-domain distribution of IMF, there is no fault-related impulse component. In order to view the frequency distribution of each IMF, fast Fourier transform (FFT) was

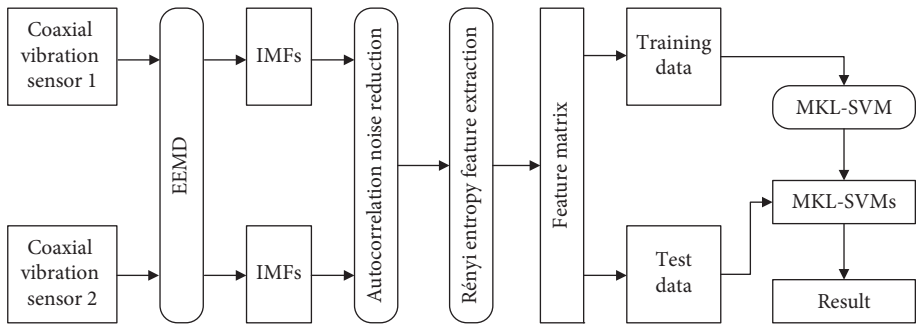


Figure 1: Flowchart of CVSFF.

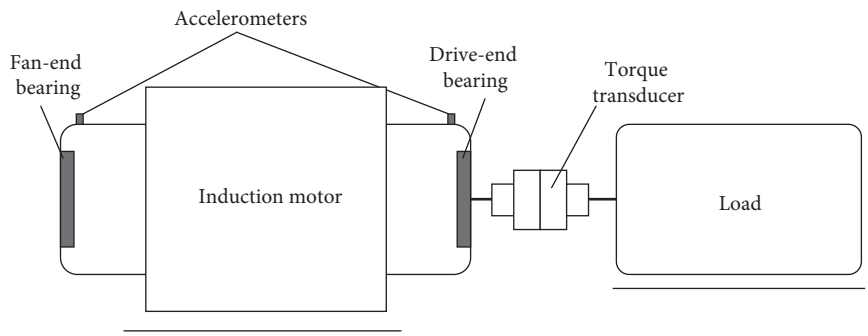


Figure 2: Schematic diagram of the bearing test rig.

Table 1: Data set.

No.	Data set number	Bearing type	Symbol in this paper	Results of [28] DE/FE
1	99	Normal	N	—
2	171	Drive-end inner race fault	I-D	Y2/P1
3	120	Drive-end ball fault	B-D	N1/P2
4	198	Drive-end outer race fault (centred)	O6-D	N2/N2
5	276	Fan-end inner race fault	I-F	N1/Y2
6	292	Fan-end roller fault	B-F	N1/N1
7	301	Fan-end outer race fault (orthogonal)	O3-F	N1/P2
8	307	Fan-end outer race fault (opposite)	O12-F	N1/N1

performed for each component. Take the first segment of data for eight states as examples, IMF frequency-domain distribution of these data is shown in Figure 4.

It can be seen in Figure 4 that the energy of IMFs is mainly concentrated in IMF1. Due to the characteristics of EEMD, the frequency content of IMFs ranges from high to low with the increase of order. Moreover, EEMD separates the frequency bands successfully without mode mixing. It may also be noticed that the frequency distribution of normal bearing coaxial sensor data is very similar by comparing Figures 4(a) and 4(b). In Figures 4(c) to 4(p), which are fault bearing coaxial data, components except IMF1 have similar frequency distribution too. Meanwhile, it can be seen from the red box in Figure 4 that the frequency distribution of IMF components is spread over many frequency bands. The random noise in IMFs submerged part of the fault impulse, thus affecting the subsequent feature extraction effect. Therefore, the method of autocorrelation noise reduction is needed to reduce random noise in IMF components.

3.3. Autocorrelation Noise Reduction. In order to suppress the noise, the autocorrelation function of each IMF is computed in the time domain. Figure 5 is the frequency distribution of the corresponding data in Figure 4 after autocorrelation denoise.

By comparing Figures 5 and 4, it can be seen that the frequency amplitude marked by the red box in Figure 4 significantly decreases after the autocorrelation noise reduction. And the frequency band corresponding to the large amplitude of each component does not change.

The frequency distribution of IMFs is more centralized, which means random noise in the component is effectively suppressed after autocorrelation. However, Res in Figures 5(b), 5(d), and 5(f) show a same low-frequency peak distribution of 10 Hz marked in the red dotted box. We take Res of Figure 5(d) as an example to analyze this phenomenon. The time domain, autocorrelation function, and corresponding frequency-domain distribution of this component are shown in Figure 6.

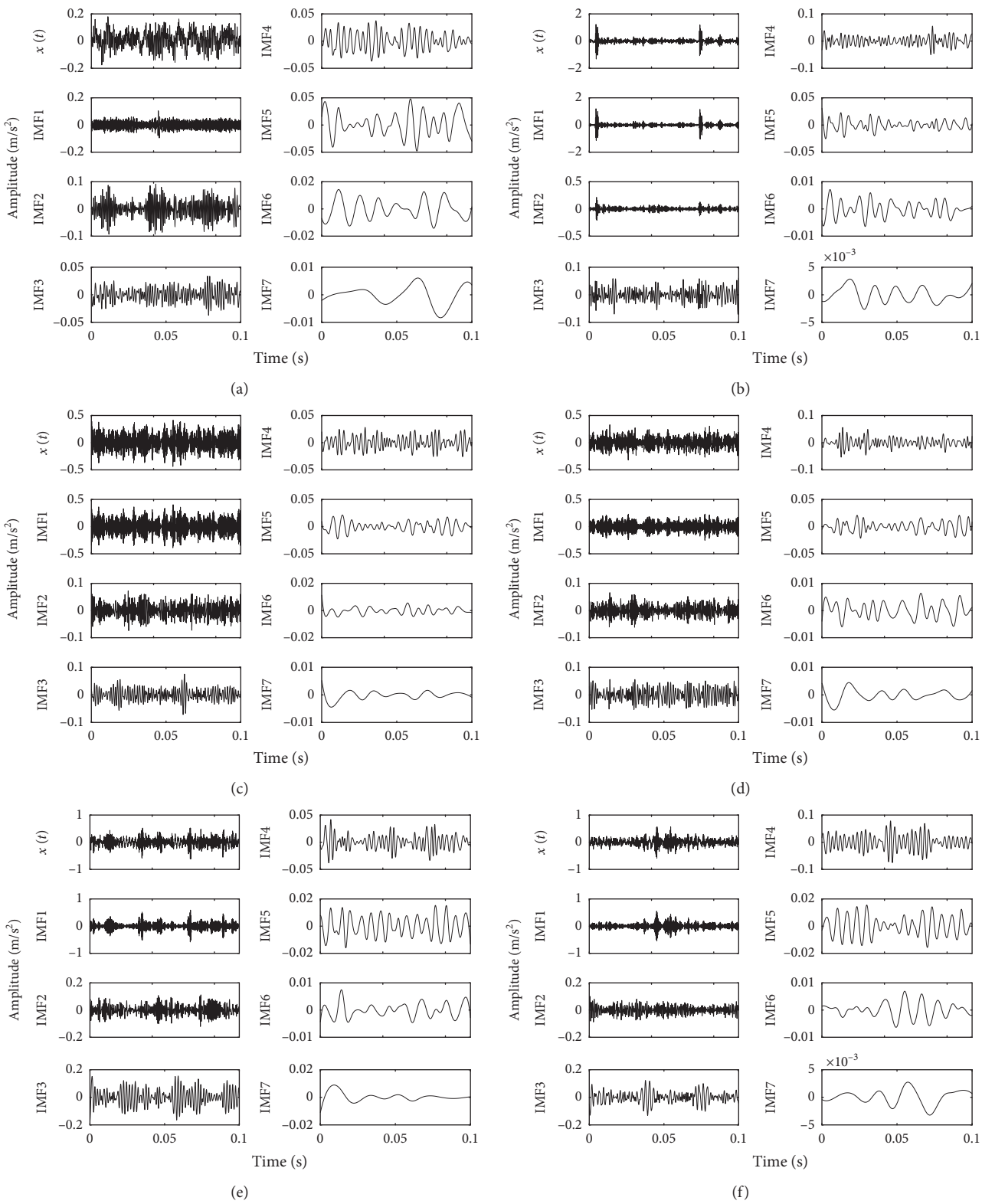
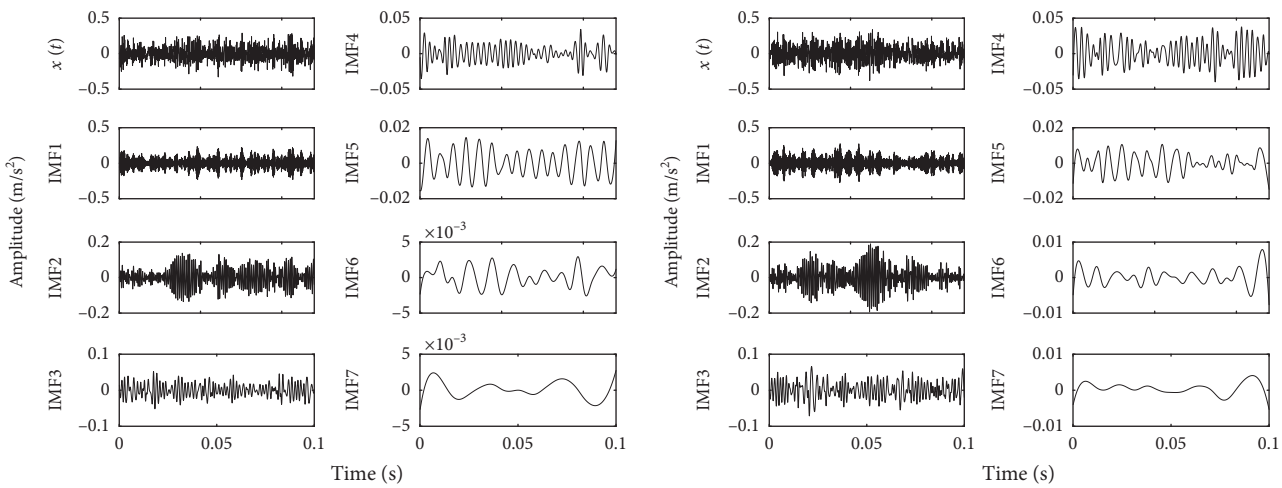
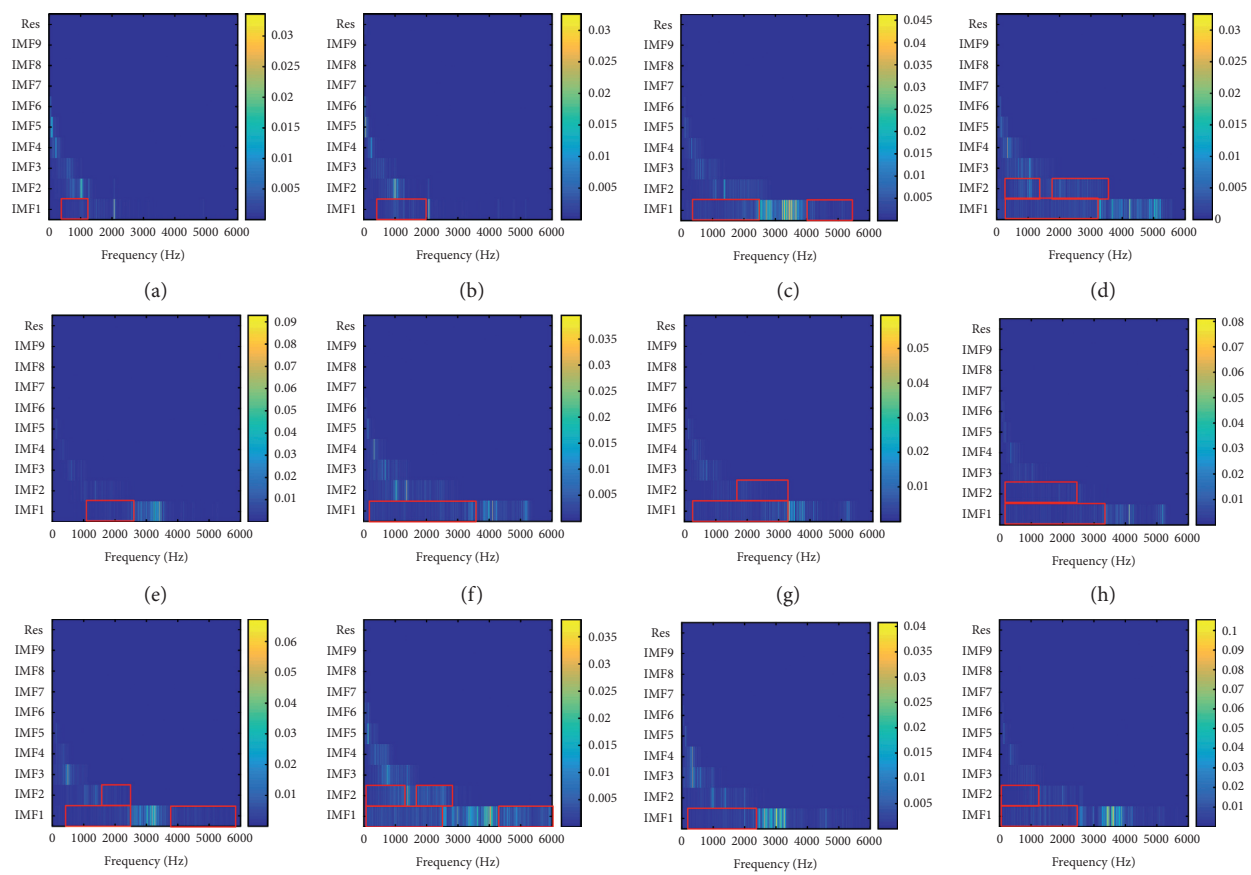


Figure 3: Continued.



(g) (h)

Figure 3: The EEMD results of eight states. (a) N. (b) I-D. (c) B-D. (d) O6-D. (e) I-F. (f) B-F. (g) O3-F. (h) O12-F.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l)

Figure 4: Continued.

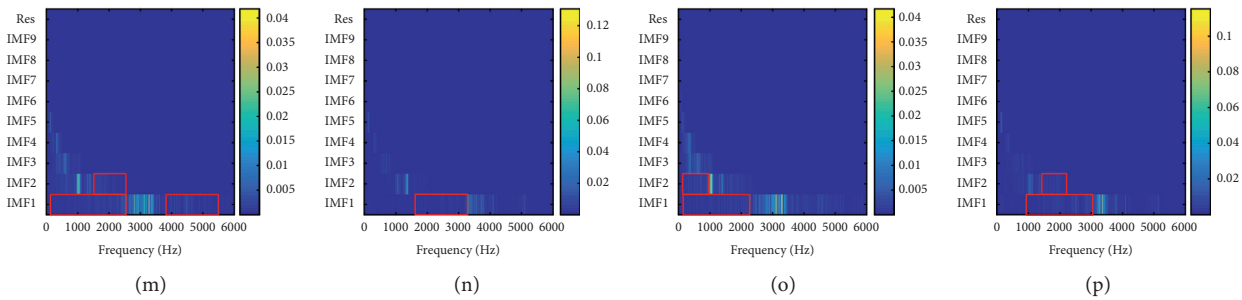


Figure 4: Frequency distribution of each IMF of the original signal. (a) N (drive end). (b) N (fan end). (c) I-D (drive end). (d) B-F (fan end). (e) B-D (drive end). (f) B-D (fan end). (g) O6-D (drive end). (h) O6-D (fan end). (i) I-F (drive end). (j) I-F (fan end). (k) B-F (drive end). (l) B-F (fan end). (m) O3-F (drive end). (n) O3-F (fan end). (o) O12-F (drive end). (p) O12-F (fan end). Res means residual component.

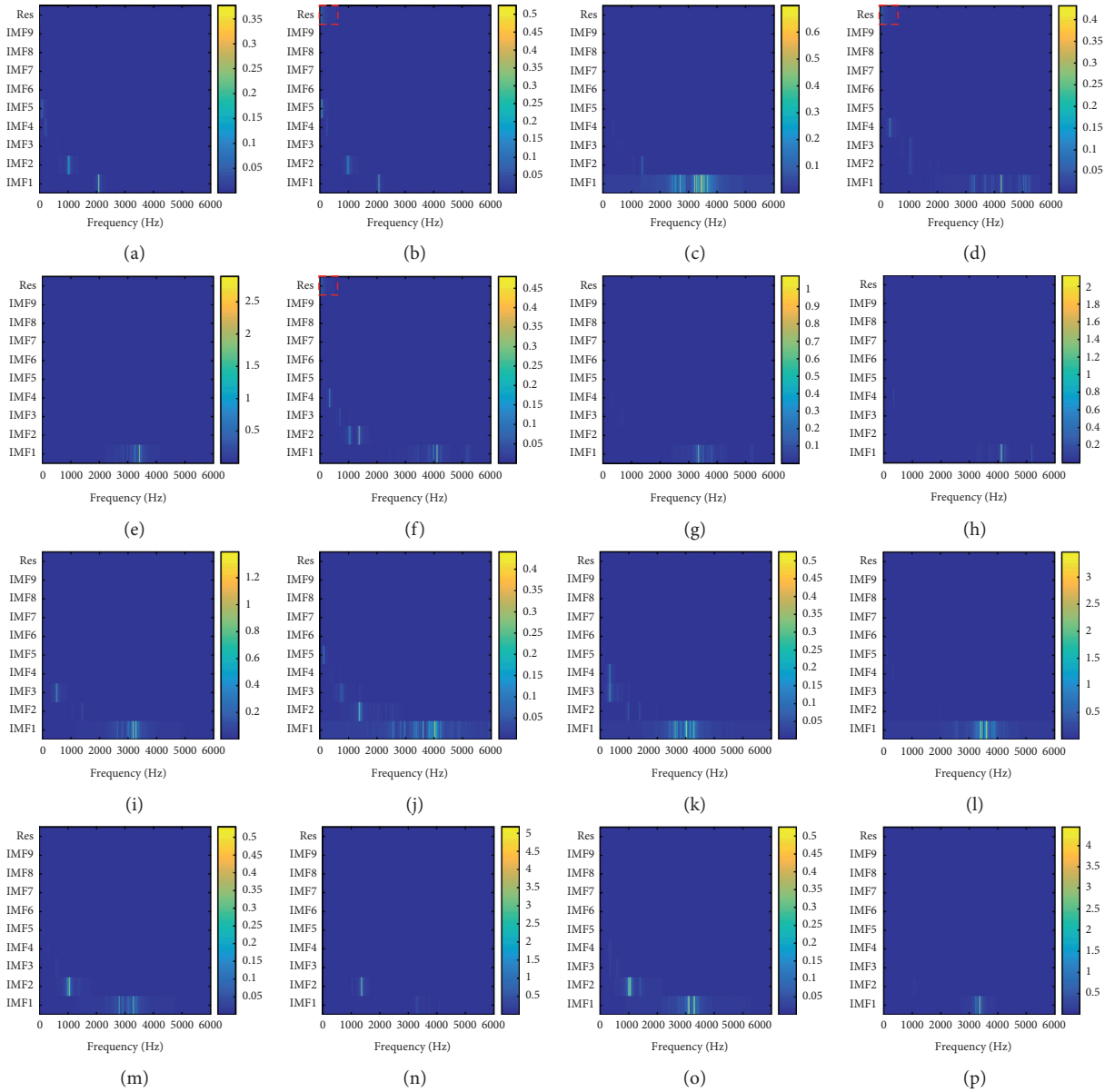


Figure 5: Frequency distribution of each IMFs after autocorrelation denoise. (a) N (drive end). (b) N (fan end). (c) I-D (drive end). (d) B-F (fan end). (e) B-D (drive end). (f) B-D (fan end). (g) O6-D (drive end). (h) O6-D (fan end). (i) I-F (drive end). (j) I-F (fan end). (k) B-F (drive end). (l) B-F (fan end). (m) O3-F (drive end). (n) O3-F (fan end). (o) O12-F (drive end). (p) O12-F (fan end). Res means residual component.

As can be seen from Figure 6, the signal becomes a monotone signal after autocorrelation. So, the frequency peak of 10 Hz in Figures 5(b), 5(d), and 5(f) correspond to the time length 0.1 s of signal. It has no actual physical meaning. Since this paper only uses IMF1-9 components to calculate the entropy value, it has no influence on the feature matrix distribution.

3.4. Entropy Feature Extraction. According to equations (6) and (7), we calculated the Rényi entropy of IMFs. The dimension of entropy multisensor feature matrix of each state is $100 * 2$. The row 100 of feature matrix means 100 segments of each state. The column 2 of feature matrix means 2 coaxial sensors. Table 2 shows the mean values of entropy for eight states which are calculated by taking an average of 100 Rényi entropy values for every state of bearings. The entropy values of IMFs without and with denosie are compared in Table 2.

The variation in Table 2 means the change rate of entropy mean value for the noise-reduced data comparing with original IMFs:

$$\text{variation} = \frac{(R_1 - R_2)}{R_1}, \tag{13}$$

where R_1 is the mean value of entropy for data without denoise and R_2 is mean value of entropy for data with autocorrelation denoise.

As can be seen from Table 2, the entropy of all measuring points decreases after autocorrelation. Arrays with a large reduction (greater than 50%) are highlighted in bold. The maximum reduction is 81.53%. It can also be seen from Table 2 that entropy of data from sensors located in faulty bearing end generally drops greatly. This may be because the vibration signal at the fault bearing end is more susceptible to noise interference.

Figure 7 is the Rényi entropy box plot calculated by vibration signals of two coaxial sensors. Figure 7(a) is the Rényi entropy distribution for data from the drive-end sensor. Figure 7(b) is the Rényi entropy distribution for data from the fan-end sensor. The black part in Figure 7 is the entropy value calculated from raw data, while the red part is the entropy value calculated by IMFs after autocorrelation noise reduction. It is clear that all entropy decreased after noise reduction. Furthermore, the entropy distribution for data without noise reduction from drive-end or fan-end sensor is similar to that after autocorrelation denoise. Although autocorrelation denoise hardly changes the entropy tendency of bearings in different states, it increases the differentiation of eight states. In addition, the entropy values of normal bearing in both sensors are greater than those of all faulty bearings which are consistent with the characteristics of the Rényi entropy.

In order to check the distribution of the outlier in the box diagram, we counted the number of outliers in the box plot of Figure 7. It can be clearly seen from Table 3 that the entropy distribution becomes more centralized and the number of outlier decreases after vibration noise reduction. The overall decrease of outliers was attributed to bearing states of I-F, O3-F, and O12-F. That is, the reduction of the outlier mainly consisted of coaxial data of fan-end fault bearing.

The Rényi entropy of the vibration signal at drive end and fan end constitutes an $800 * 2$ dimensional multisensor feature vector. To compare the distribution of feature matrix without and with noise reduction more distinctly, we draw the feature scatter plot. Their distributions are shown in Figure 8.

As can be observed, features from all conditions except normal bearing have relatively concentrated space distribution in Figure 8(a). In Figure 8(b), although conditions B-F, O3-F, and O12-F still have some similar features, the aliasing of the feature distributions of different states has been significantly improved after noise reduction. The characteristic differentiation degree of bearing in eight different states is increased.

3.5. Result Comparison and Discussion. Through all above signal preprocessing methods carried out in Section 3.2 to 3.4, we obtained the coaxial vibration feature matrix. To classify eight different states of bearing, MKL-SVM is used in this paper. Kernel functions are applied to the features of single sensor and all sensors.

Common kernel functions are divided into two types: local kernel function and global kernel function. The local kernel function has local characteristics, strong learning ability, and weak generalization ability. Linear and Gaussian kernel functions are typical local kernel function. On the contrary, the global kernel function has global characteristics, strong generalization ability, and weak learning ability. Polynomial and sigmoid kernel functions are typical global kernel functions.

In order to make the model have a good classification effect for data with different characteristics, the Gaussian kernel and polynomial kernel functions with different parameters were selected in this paper. Detailed parameters are as follows:

(1) Five Gaussian kernel functions with different bandwidths: bandwidth σ is uniformly sampled at the interval of [0.01, 100] on a logarithmic scale.

(2) Three different levels of polynomial function: the orders d are 1, 2, and 3.

These kernel functions are applied to a single feature and all features, respectively. Therefore, there are 24 kernels in total.

3.5.1. Comparison and Discussion of Coaxial Signals with Individual Signal. In order to verify the sensitivity of coaxial signal features, we also did the same diagnosis based on single sensor. To train MKL-SVMs, 70% of feature matrix is selected as training data. The *one against all* method is adopted to construct multiclassification model. To improve computational efficiency, we jointly optimized eight binary classification problems. A total of eight MKL-SVMs and a combination of kernels are obtained for coaxial signals or

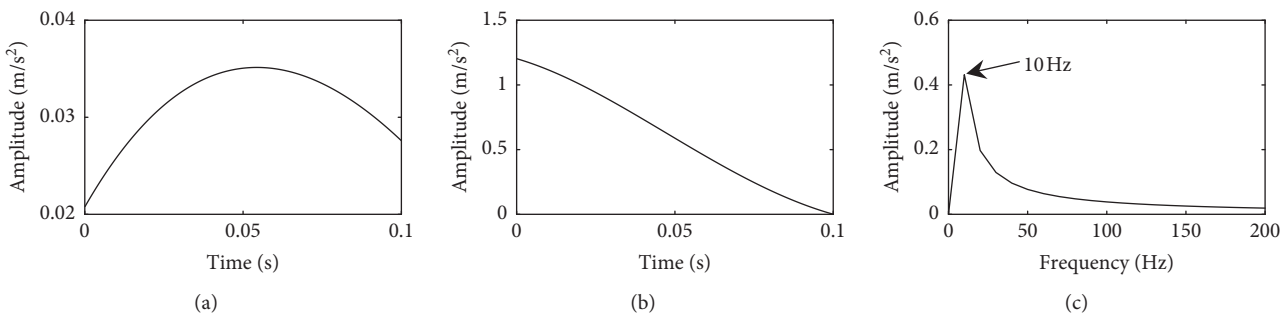


Figure 6: Time domain, autocorrelation function, and corresponding frequency-domain distribution of Res in Figure 5(d).

Table 2: Mean values of the Rényi entropy.

Sensor location	1 N	2 I-D	3 B-D	4 O6-D	5 I-F	6 B-F	7 O3-F	8 O12-F
Without denoise								
Drive end	2.5928	1.2033	1.0874	1.5331	1.6016	1.8180	1.7362	1.6746
Fan end	2.5244	1.9789	2.0922	1.6853	1.9023	1.3981	1.7274	1.4613
Autocorrelation denoise								
Drive end	1.9548	0.4336	0.2008	0.4485	1.0307	1.4850	1.2504	1.0036
Variation (%)	24.61	**63.96**	**81.53**	**70.74**	35.65	18.32	27.98	40.07
Fan end	1.9314	1.4479	1.6025	0.4397	1.2879	0.4887	0.5916	0.5036
Variation (%)	23.49	26.83	23.41	**73.91**	32.30	**65.05**	**65.75**	**65.54**

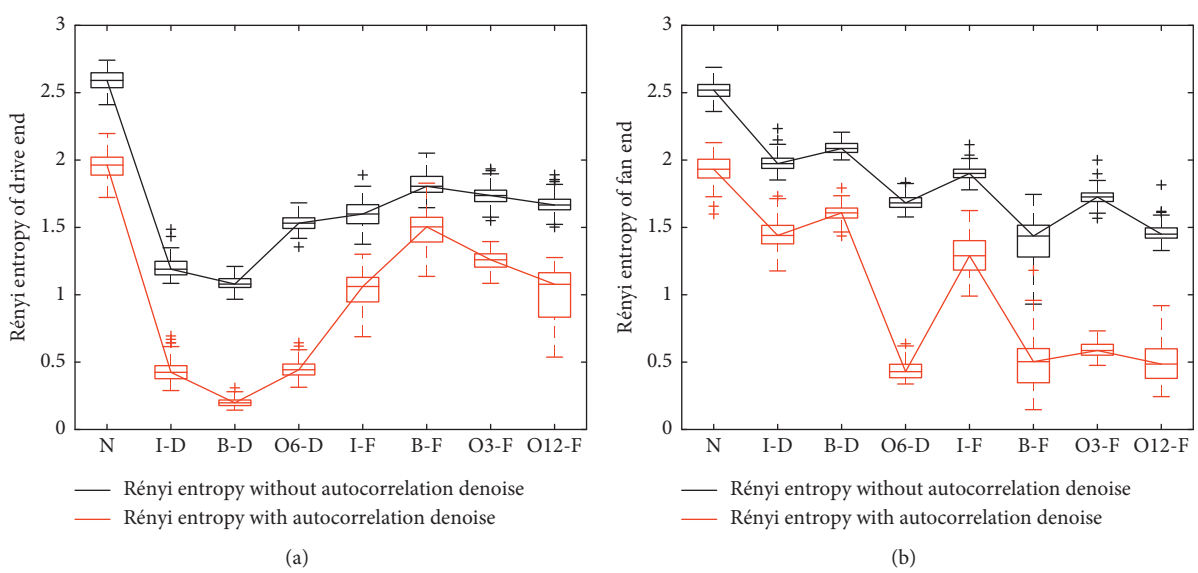


Figure 7: Box plot of the Rényi entropy distribution of (a) drive end and (b) fan end.

Table 3: Box plot outlier quantity of the Rényi entropy.

Signal location	1 N	2 I-D	3 B-D	4 O6-D	5 I-F	6 B-F	7 O3-F	8 O12-F	Total
Without denoise									
Drive end	0	2	0	1	1	0	3	4	11
Fan end	0	2	0	1	2	1	3	2	11
Autocorrelation denoise									
Drive end	0	3	1	2	0	0	0	0	6
Fan end	2	1	2	1	0	1	0	0	7

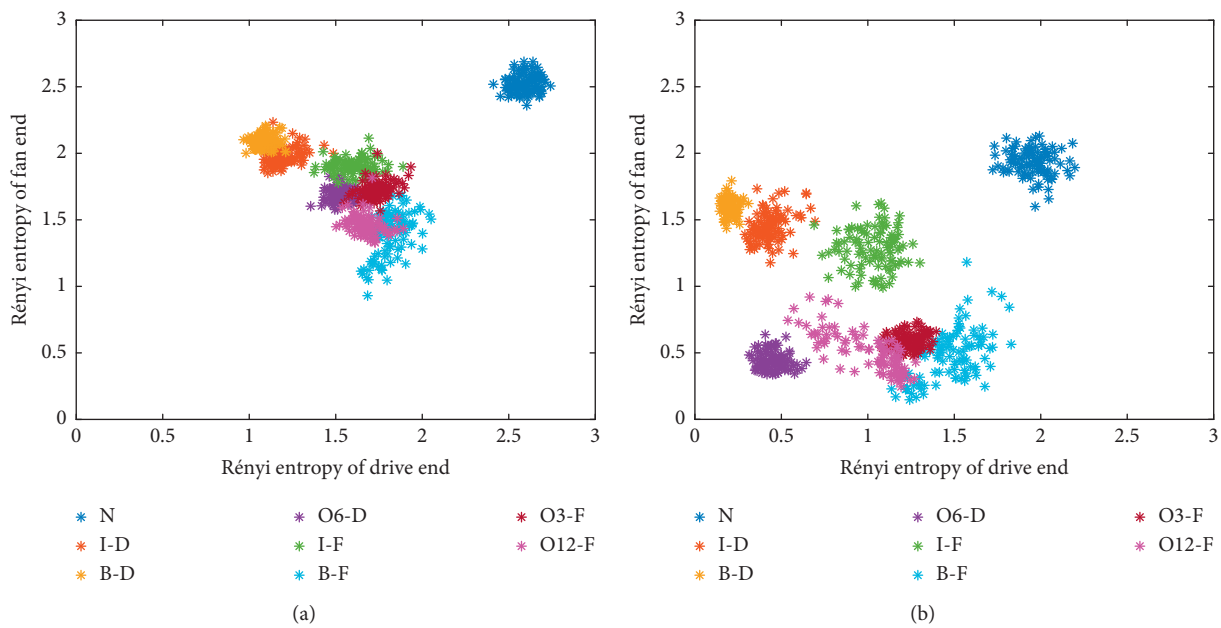


Figure 8: Rényi entropy distributions of (a) raw data and (b) after autocorrelation denosie.

individual signal. The detailed parameters of the models are shown in Table 4 and 5.

As can be observed from Table 4, weight of the kernel function acting on the data of two coaxial sensors accounts for 91.41%, which takes the most part of the MKL model. More specifically, the fusion of two-dimensional coaxial features outperformed feature of a single sensor. It can also be concluded from Table 4 that the weights obtained by MKL realize not only the selection of kernel function but also the selection of data. The kernel function parameters of model based on the data of the drive end in Table 5 are similar to those of the fan end. These selected kernel functions are sorted by weight as follows: Gaussian kernel $\sigma = 0.01$, Gaussian kernel $\sigma = 0.1$, and polynomial kernel $d = 3$. In addition, a first-order polynomial kernel is included based on the drive-end data, but the weight is very small. According to the previous formula, the first-order polynomial kernel is a linear kernel. The classification results of these three models based on different data are shown in Table 6. The classification accuracy based on the CVSFF method is 97.50%. This is much higher than those of single signal-based classification accuracy, which are 66.67% and 60.42%, respectively. All results of three models prove that the fusion of data from two coaxial sensors improved accuracy and robustness in fault detection of bearing.

The huge difference in results is mainly due to the good differentiation of two-dimensional features based on coaxial signals, while the single-dimensional entropy distribution of the individual sensor overlaps, as shown in Figure 7. For the characteristics of drive end, I-F and O12-F are distributed with overlap in the interval [0.6, 1.2]; I-D and O6-D are distributed with overlap in the interval [0.3, 0.5]. For the characteristics of fan end, I-D and I-F are distributed with overlap in the interval [1.1, 1.6]; O6-D, B-F, O3-F, and O12-F are distributed with overlap in the interval [0.3, 0.7].

3.5.2. Comparison and Discussion of MKL-SVM with SVM, GA-SVM, and PSO-SVM. To introduce the effectiveness of MKL-SVM, SVM, genetic algorithm-optimized SVM (GA-SVM), and particle swarm optimization SVM (PSO-SVM) based on single kernel are measured in this section. All parameters of SVM are default. That is to say, the kernel function is the Gaussian kernel; the bandwidth of the Gaussian kernel is 1.

In order to compare the results of CVSFF, we use the same feature matrix that is used in MKL-SVM to train and test the other three SVM models. The classification results are shown in Table 7. The classification accuracies are 97.50%, 95.00%, 95.83%, and 96.67% using the MKL-SVM, original SVM, GA-SVM, and PSO-SVM, respectively. The Gaussian kernel bandwidth obtained by GA and PSO optimization is 8.64 and 5.64. Although the classification accuracy of two optimized algorithms is 0.83% and 1.67% higher than that of SVM, it is still lower than MKL-SVM.

Furthermore, the running time of the models is compared. The experiments are carried out on Intel (R) Core (TM) i5-4210U CPU 2.4 GHz, 4G RAM, Win 7 and MATLAB R2016b. With the exception of the SVM model, the other three models need to optimize the parameters. Thus, the running time of these three models is greatly increased compared with that of SVM. In the first, third, and fourth models, GA-SVM has the shortest running time but the lowest accuracy. The MKL-SVM model adopted in this paper has a higher classification accuracy, although its running time is 86.17 s longer than GA-SVM. Compared with PSO-SVM, this method is not only more efficient but also more accurate in classification.

To analyze the results of these four models in detail, we draw the confusion matrix corresponding to the classification results in Table 7, as shown in Figure 9.

Apparently, the error classification is mainly concentrated in B-F, O3-F, and O12-F. And the predicted states of these

Table 4: Multikernel parameters for coaxial signals model.

No.	1	2	3	4	5
Kernel	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian
Bandwidth	0.01	0.1	1	10	100
Variable	DE + FE	DE + FE	DE + FE	DE + FE	DE + FE
Weight	0.3430	0.3724	0.1903	0	0
No.	6	7	8	9	10
Kernel	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian
Bandwidth	0.01	0.1	1	10	100
Variable	DE	DE	DE	DE	DE
Weight	0	0.0780	0	0	0
No.	11	12	13	14	15
Kernel	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian
Bandwidth	0.01	0.1	1	10	100
Variable	FE	FE	FE	FE	FE
Weight	0.0079	0	0	0	0
No.	16	17	18	19	20
Kernel	Polynomial	Polynomial	Polynomial	Polynomial	Polynomial
Order	1	2	3	1	2
Variable	DE + FE	DE + FE	DE + FE	DE	DE
Weight	0	0	0.0084	0	0
No.	21	22	23	24	
Kernel	Polynomial	Polynomial	Polynomial	Polynomial	
Oder	3	1	2	3	
Variable	DE	FE	FE	FE	
Weight	0	0	0	0	

Table 5: Multikernel parameters for individual signal model.

No.	1	2	3	4	5
Kernel	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian
Bandwidth	0.01	0.1	1	10	100
Weight for individual signal of drive end	0.7169	0.2463	0	0	0
Weight for individual signal of fan end	0.7636	0.1958	0	0	0
No.	6	7	8		
Kernel	Polynomial	Polynomial	Polynomial		
Order	1	2	3		
Weight for individual signal of drive end	0.0053	0	0.0316		
Weight for individual signal of fan end	0	0	0.0405		

Table 6: Results of coaxial signals and individual signal.

No.	Data	Method	Label	Training data	Test data	Accuracy
1	Coaxial signals	MKL-SVM	1–8	560 ∗ 2	240 ∗ 2	97.50% (235/240)
2	Individual signal of drive end	MKL-SVM	1–8	560 ∗ 1	240 ∗ 1	66.67% (160/240)
3	Individual signal of fan end	MKL-SVM	1–8	560 ∗ 1	240 ∗ 1	60.42% (145/240)

Table 7: Results of MKL-SVM and SVM.

No.	Data	Method	Label	Training data	Testing data	Accuracy	Running time (s)
1	Coaxial signals	MKL-SVM	1–8	560 ∗ 2	240 ∗ 2	97.50% (235/240)	170.96
2	Coaxial signals	SVM	1–8	560 ∗ 2	240 ∗ 2	95.00% (228/240)	0.22
3	Coaxial signals	GA-SVM	1–8	560 ∗ 2	240 ∗ 2	95.83% (230/240)	84.79
4	Coaxial signals	PSO-SVM	1–8	560 ∗ 2	240 ∗ 2	96.67% (232/240)	367.18

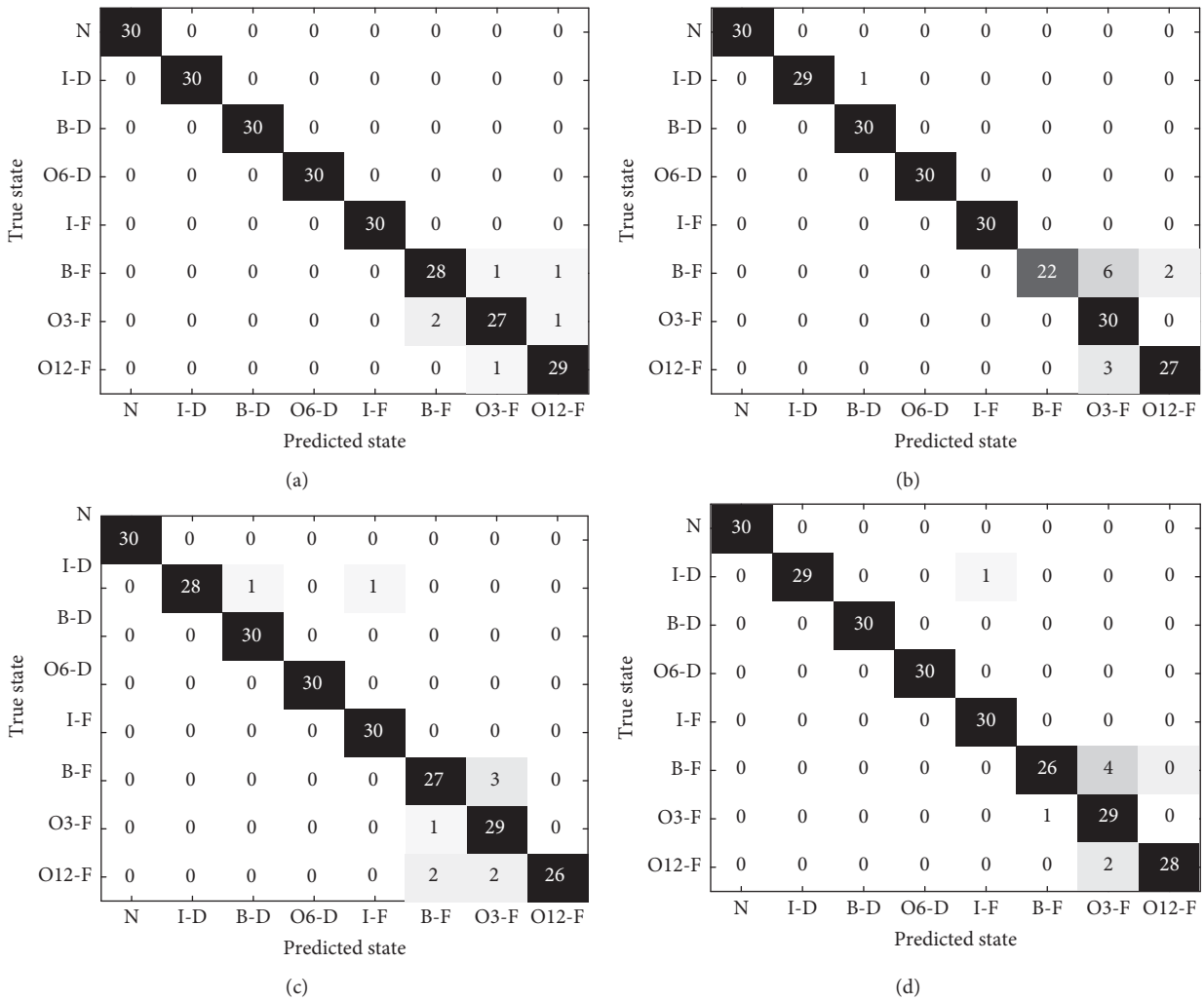


Figure 9: Classification results of (a) MKL-SVM, (b) SVM, (c) GA-SVM, and (d) PSO-SVM.

error classifications are B-F, O3-F, and O12-F. It verifies the characteristics of the feature distribution in Figure 8(b). The lowest classification accuracy of MKL-SVM is 90.00% (O3-F, 27/30) for single state bearing, which is significantly higher than that of SVM (73.33%, B-F, 22/30). In addition, the classification accuracy of MKL-SVM for the first five states is 100%, while SVM, GA-SVM, and PSO-SVM have 1, 2, and 1 misclassification of data in I-D state, respectively.

However, it can also be seen from Figure 9 that there are three groups of misclassification for O3-F based on the method proposed in this paper, while the performance of the other three models is better. The three segments misclassified by MKL-SVM are the 79th, 88th, and 97th segments of data O3-F. GA-SVM misclassified the 71th segment of O3-F. PSO-SVM misclassified the 88th segment of O3-F which is one of the MKL-SVM misclassified groups. The 97th segment is not close to the feature of B-F, but it is still misclassified as B-F. The results that are misclassified by MKL-SVM need further study and improvement. Nevertheless, MKL-SVM outperforms SVM, GA-SVM, and PSO-SVM both in whole and single state classification accuracy.

4. Conclusion

This paper proposed a feature fusion method based on MKL-SVM using coaxial vibration signals to classify REB states. The obtained accuracy of coaxial signals is 97.50% which is much higher than the results of single sensor signal for drive end (66.67%) and fan end (60.42%). Polynomial kernel with global characteristics and Gaussian kernel with local characteristics are selected, which greatly improves the generalization ability of the model. By comparing the CVSFF, SVM, GA-SVM, and PSO-SVM-based feature fusion, 97.50%, 95.00%, 95.83%, and 96.67% accuracies are obtained, respectively. It shows that the method extracted in this paper is more effective for REB fault classification.

Data Availability

The bearing data used to support the findings of this study are obtained from the Case Western Reserve University (CWRU) Bearing Data Center (http://csegroups.case.edu/bearingdatacenter/home).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The research was supported by the China Railway (grant number 2017J004-H).

References

[1] E. P. Carden and P. Fanning, "Vibration based condition monitoring: a review," *Structural Health Monitoring: An International Journal*, vol. 3, no. 4, pp. 355–377, 2004.

[2] B. Li, M.-Y. Chow, Y. Tipsuwan, and J. C. Hung, "Neural-network-based motor rolling bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1060–1069, 2000.

[3] L. Guo, H. Gao, H. Huang, X. He, and S. Li, "Multifeatures fusion and nonlinear dimension reduction for intelligent bearing condition monitoring," *Shock and Vibration*, vol. 2016, Article ID 4632562, 10 pages, 2016.

[4] H. Zhao, S. Zuo, M. Hou et al., "A novel adaptive signal processing method based on enhanced empirical wavelet Transform technology," *Sensors*, vol. 18, no. 10, p. 3323, 2018.

[5] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, "A survey on machine learning for data fusion," *Information Fusion*, vol. 57, pp. 115–129, 2020.

[6] X. Yan, Z. Sun, J. Zhao, Z. Shi, and C.-a. Zhang, "Fault diagnosis of rotating machinery equipped with multiple sensors using space-time fragments," *Journal of Sound and Vibration*, vol. 456, pp. 49–64, 2019.

[7] L.-l. Jiang, H.-k. Yin, X.-j. Li, and S.-w. Tang, "Fault Diagnosis of rotating machinery based on multisensor information fusion using SVM and time-domain features," *Shock and Vibration*, vol. 2014, Article ID 418178, 8 pages, 2014.

[8] J. Tao, Y. Liu, and D. Yang, "Bearing fault diagnosis based on deep belief network and multisensor information fusion," *Shock and Vibration*, vol. 2016, Article ID 9306205, 9 pages, 2016.

[9] T. P. Banerjee and S. Das, "Multi-sensor data fusion using support vector machine for motor fault detection," *Information Sciences*, vol. 217, pp. 96–107, 2012.

[10] J. Wang, Y. Liang, Y. Zheng, R. X. Gao, and F. Zhang, "An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples," *Renewable Energy*, vol. 145, pp. 642–650, 2020.

[11] H. Zhou, J. Chen, G. Dong, H. Wang, and H. Yuan, "Bearing fault recognition method based on neighbourhood component analysis and coupled hidden Markov model," *Mechanical Systems and Signal Processing*, vol. 66-67, pp. 568–581, 2016.

[12] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digital Signal Processing*, vol. 27, pp. 159–166, 2014.

[13] K. H. Hui, M. H. Lim, M. S. Leong, and S. M. Al-Obaidi, "Dempster-Shafer evidence theory for multi-bearing faults diagnosis," *Engineering Applications of Artificial Intelligence*, vol. 57, pp. 160–170, 2017.

[14] M. S. Safizadeh and S. K. Latifi, "Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell," *Information Fusion*, vol. 18, pp. 1–8, 2014.

[15] S. Lu, P. Zheng, Y. Liu, Z. Cao, H. Yang, and Q. Wang, "Sound-aided vibration weak signal enhancement for bearing fault detection by using adaptive stochastic resonance," *Journal of Sound and Vibration*, vol. 449, pp. 18–29, 2019.

[16] V. Mygdalis, A. Tefas, and I. Pitas, "Exploiting multiplex data relationships in support vector machines," *Pattern Recognition*, vol. 85, pp. 70–77, 2019.

[17] M. H. Zangooei and S. Jalili, "PSSP with dynamic weighted kernel fusion based on SVM-PHGS," *Knowledge-Based Systems*, vol. 27, pp. 424–442, 2012.

[18] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[19] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[20] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[21] Y. Peng, Z. Li, K. He et al., "Broadband mode decomposition and its application to the quality evaluation of welding inverter power source signals," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 11, pp. 9734–9746, 2020.

[22] Y. Peng, Z. Wang, K. He et al., "Modulated broadband mode decomposition for the feature extraction of double pulse metal inert gas welding," *IEEE Access*, vol. 8, pp. 134031–134041, 2020.

[23] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[24] A. Rényi, "On measures of entropy and information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, CA, USA, pp. 547–561, 1961.

[25] P. Boškoski and Đ. Juričić, "Fault detection of mechanical drives under variable operating conditions based on wavelet packet Rényi entropy signatures," *Mechanical Systems and Signal Processing*, vol. 31, pp. 369–381, 2012.

[26] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 1995.

[27] A. Rakotomamonjy, F. Bach, S. Canu et al., "Simple MKL," *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.

[28] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study," *Mechanical Systems and Signal Processing*, vol. 64-65, pp. 100–131, 2015.