

Research Article

An Integrated Cumulative Transformation and Feature Fusion Approach for Bearing Degradation Prognostics

Lixiang Duan¹, Fei Zhao,² Jinjiang Wang,¹ Ning Wang,¹ and Jiwang Zhang¹

¹School of Mechanical and Transportation Engineering, China University of Petroleum, Beijing 102249, China

²Sichuan Special Equipment Inspection and Research Institute, Chengdu 610051, China

Correspondence should be addressed to Lixiang Duan; duanlx@cup.edu.cn

Received 4 August 2017; Revised 21 December 2017; Accepted 15 January 2018; Published 18 February 2018

Academic Editor: Rafał Burdzik

Copyright © 2018 Lixiang Duan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aimed at degradation prognostics of a rolling bearing, this paper proposed a novel cumulative transformation algorithm for data processing and a feature fusion technique for bearing degradation assessment. First, a cumulative transformation is presented to map the original features extracted from a vibration signal to their respective cumulative forms. The technique not only makes the extracted features show a monotonic trend but also reduces the fluctuation; such properties are more propitious to reflect the bearing degradation trend. Then, a new degradation index system is constructed, which fuses multidimensional cumulative features by kernel principal component analysis (KPCA). Finally, an extreme learning machine model based on phase space reconstruction is proposed to predict the degradation trend. The model performance is experimentally validated with a whole-life experiment of a rolling bearing. The results prove that the proposed method reflects the bearing degradation process clearly and achieves a good balance between model accuracy and complexity.

1. Introduction

In order to improve production efficiency, quality, and flexibility, modern manufacturing highly depends on the trouble-free operation of components in manufacturing machines [1, 2]. Therefore, the timely monitoring and estimation of the running status of important machine components are necessary [3]. Prognostics, as an important process of monitoring the condition of components or a system [4], is regarded as a predictive maintenance strategy that can anticipate the occurrence of component defects to ensure the availability, reliability, and security of critical components (such as bearings, gears, or cutting tools) [5, 6]. A bearing is at the heart of rotating machinery and is viewed as a frequent contributor to the breakdown of such machinery [7, 8]; hence, accurate prognostics of bearing degradation trends is of substantial practical relevance.

In general, prognostics can be roughly classified into model-based, data-driven, and hybrid approaches [9]. Among these, data-driven approaches are easier to deploy and have been widely used [10]. They usually follow a road map of data acquisition, feature extraction, and prognostic

modeling. Of the three steps above, the key step is feature extraction, because it has a direct influence on predictor performance. Features that can properly reflect the degradation progression could yield accurate and simple prognostics [11–13].

Various features extracted from time domain, frequency domain, and time-frequency domain have been investigated for bearing degradation prognostics. Time domain methods extract statistical features, such as root mean square (RMS), kurtosis, crest factor, and peak-to-peak value. The energy of bearing defect frequency is extracted as features by frequency domain methods, such as ball-pass frequency of outer ring f_{BPFO} , ball-pass frequency of inner ring f_{BPFI} , and ball-spin frequency f_{BSF} . Time-frequency domain methods extract features, such as short time Fourier transform (STFT) [14], wavelet transform (WT) [15, 16], and empirical mode decomposition (EMD) [17].

These features are extracted to serve as degradation prognostics; however, the original features often show nonideal evolution, fail to reflect the degradation trend, and even cause problems for the prognostics tasks.

In the current research, these questions are studied from different aspects. Some methods are proposed from the smoothing aspect to reduce the influence of fluctuation and obtain a smoothed trend feature. For example, Sutrisno et al. [18] used a moving average filter to reduce the fluctuation of original features and the smoothed features were put into the least squares support vector regression (LS-SVR) for bearing remaining life estimation. Following a similar idea, Loutas et al. [19] employed a moving average filter to suppress the high noise and volatility of condition-based maintenance features. Ben Ali et al. [20] proposed a feature extraction method based on a combination of a sliding average and Weibull distribution to avoid feature fluctuation for accurate bearing remaining useful life prediction. Although smoothing methods are easy to implement, they could lose some useful information. More generally, new indexes have been constructed using some further processing methods on the original features. For instance, Qiu et al. [21] developed a robust degradation indicator based on a self-organizing map neural network to evaluate the bearing degradation performance. Sassi et al. [22] developed two new scalar indicators by combining six conventional features to evaluate the severity of bearing degradation. Zhang et al. [23] used a continuous hidden Markov model combined with various features to construct an effective degradation indicator of rolling bearings for residual life prediction. Xi et al. [24] established a virtual health index based on a transformation formula to quantify the health degradation of an engineering system and further performed a health state prediction process. Liao [25] discovered prognostic features to represent the fault progression by using genetic programming, in which monotonicity is used as fitness function. Sun et al. [26] proposed a nonprobabilistic metric to assess the operational reliability of aeroengines by utilizing the cosine function to map kernel principal angles extracted from condition information into a similarity index.

The above feature extraction methods process only a single signal fragment and ignore the relationship of the whole signal. The extracted features are easily affected by other factors such as noise and show some local fluctuations, so they do not clearly describe the state of the machinery degradation. Bearing degradation is a cumulative process, and thus we can consider the entire degradation process from a cumulative perspective. By continuously accumulating data, each extracted feature is able to take advantage of all the previous data, thereby achieving the purpose of mining the overall data, to reduce the influence of fluctuations and extract trend feature with more reliable trend characteristics.

Considering these problems, this study proposes a novel prognostics method based on a cumulative transformation technique and feature fusion approach for bearing degradation trend prognostics. To consider the overall relationship of the data and obtain features with a good monotonic trend characteristic, a novel cumulative transformation algorithm is proposed. However, each cumulative feature only contains partial information of bearing status. To more comprehensively reflect the degradation process, an effective and popular feature fusion method KPCA is used to fuse multi-dimensional cumulative features and develop a degradation

index. Furthermore, the degradation index is reconstructed by phase space reconstruction (PSR) to better grasp its nature and input into an extreme learning machine (ELM) for bearing degradation trend prognostic. The effectiveness of the proposed method is validated with a whole-life experiment of a rolling bearing.

The main merits of this study are as follows: (1) a novel cumulative transformation algorithm is proposed to extract features with better trend characteristics; (2) an integrated cumulative transformation and feature fusion method is investigated for bearing degradation prognostics; (3) a prediction model based on PSR and ELM is built to perform the prognostic task. The remainder of this article is organized as follows: The theoretical background of cumulative transformation is introduced in Section 2. The proposed method is then described in Section 3. In Section 4, the bearing experimental test demonstrated the effectiveness of the method. Finally, the conclusions of this study are drawn in Section 5.

2. Theoretical Framework

2.1. The Theory of Cumulative Damage. Generally, degradation of many units or systems, such as parts or machinery, is a damage accumulation process under different running environments, which grows as time passes. To explore the relationship between the degradation process and cumulative damage and find an effective way to reflect the degradation process, some methods have been investigated from the perspective of accumulation. These methods can be roughly divided into two categories: model-based and data-driven.

Model-based methods suppose that the degradation process can be expressed by a series of mathematical equations regarding cumulative system damage and time, and a model is set up by investigating the failure mechanism and the degradation paths. For instance, in [27, 28], different cumulative damage models are constructed to describe the system degradation process by considering different degradation paths and distributions, and a preventive maintenance policy is obtained to minimize the maintenance cost. In [29], a nonparametric model called cumulative incidence functions and logical analysis are integrated to solve the problem of multi-failure-mode in prognosis. The curves of cumulative incidence functions are calculated to reflect the running time on the degradation status of the monitored bearing by using the lifetime data of different failure modes. Model-based cumulative damage methods can describe the degradation process well using an accurate model, although an accurate model is hard to build in most cases, and the applicability of the model is limited for different systems.

Compared with model-based cumulative methods, data-driven methods are easier to achieve by transforming condition monitoring data into an appropriate cumulative form to infer the system status and estimate the useful remaining life. In [30], cumulative energy functions and mathematical morphology gradient in both the time and frequency domains from partial discharge waveforms are calculated as feature parameters to detect defects and assess the insulation conditions of high-voltage equipment. In [31], a smoothed accumulation method for bearing remaining

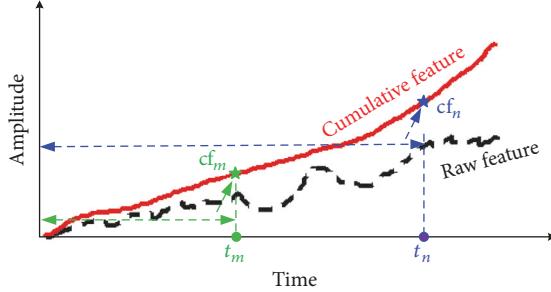


FIGURE 1: Diagram of cumulative transformation.

useful life estimation is proposed, in which the acceleration values instead of directly measured vibration values were used to estimate the bearing degradation. In [4], a cumulative method is applied to describe the degradation process from a data-driven perspective, in which the condition monitoring lifetime data are used to obtain cumulative descriptors to reflect the bearing degradation process.

In summary, a system degradation process is usually regarded as a continuous damage accumulation process over time. Therefore, we can reveal the essence of the degradation process from the perspective of accumulation by building a cumulative damage model or extracting a cumulative feature.

2.2. Cumulative Transformation Algorithm. For most mechanical components, such as bearings, gears, and rotors, the degradation is a constant damage accumulation process without self-recovery when these parts exceed the service life, so the appropriate degradation features should exhibit good trend characteristics, such as monotonicity, trendability, and robustness. Monotonicity characterizes the potential monotonic increasing or decreasing trend of the feature, and it is an essential characteristic of a degradation feature. Trendability is related to the characteristic form and its correlation with time, which reflects how the characteristic sequence changes with time and has some universality. Robustness is used to measure the fluctuation of the feature sequence. However, the features extracted from the raw vibration data often do not show good trend characteristics. Thus, a new concise and effective strategy is proposed to get monotonic and trendable features.

A new cumulative transformation algorithm is introduced by transforming an extracted feature into its corresponding cumulative form. The sketch map of cumulative transformation is shown in Figure 1. Mainly, the cumulative function is defined as the accumulation of a given time series, to which a pointwise running sum of squares of difference and scaling operations are simultaneously applied to achieve the transformation task as follows:

$$cf_n = \sqrt{\sum_{i=1}^n (f(i) - f_{\text{nor}})^2}, \quad (1)$$

where f_{nor} represents the normal value, which is defined as the corresponding average of a segment stationary trend of feature $f(i)$ in normal conditions, and cf_n represents the

cumulative total of feature up to n observations. It should be noted that the cumulative sum of a feature is sensitive to noise. To get a cumulative feature with better monotonicity and trendability, the smoothing should be applied a priori, before accumulation.

To quantitatively evaluate the suitability of the extracted features, the trend characteristics, such as monotonicity, trendability, and robustness, are further investigated. Monotonicity is given by the absolute difference between the number of positive and negative derivatives of each feature. It can be seen that the value of monotonicity is in the range from 0 to 1. The higher the monotonicity is, the greater the fitness of the feature is presented. The range of trendability is from 0 to 1, and the larger the trend index is, the higher the linear correlation degree of the feature sequence with time is. The range of robustness is also from 0 to 1; the more the feature fluctuates, the smaller the robustness is, and the greater the uncertainty is when the trend prognostics is carried out.

The criteria of monotonicity, trendability, and robustness are, respectively, defined as follows: [32]:

Mon

$$= \left| \frac{\text{number of } d/df > 0}{n-1} - \frac{\text{number of } d/df < 0}{n-1} \right|,$$

Tre (F, T)

$$= \frac{|n \sum_i f_i t_i - \sum_i f_i \sum_i t_i|}{\sqrt{[n \sum_i f_i^2 - (\sum_i f_i)^2] [n \sum_i t_i^2 - (\sum_i t_i)^2]}}, \quad (2)$$

$$\text{Rob } (F) = \frac{1}{n} \sum_i \exp \left(-\frac{f_i - \tilde{f}_i}{f_i} \right),$$

where f is the original feature curve, n is the feature number of observations, d/df is the differentiation operator which is the differential of the original feature, t is time index, and \tilde{f} is the feature curve after smooth processing of f .

A single evaluation metric can only measure the suitability of the degradation feature from a certain aspect. In order to evaluate the feature more comprehensively, a linear weighted comprehensive indicator is proposed to realize the effective fusion of evaluation metrics. The specific form of weighted fusion is as follows:

$$\begin{aligned} W_{F \in \Omega} &= \omega_1 \text{Tre}(F, T) + \omega_2 \text{Mon}(F) + \omega_3 \text{Rob}(F) \\ \text{s.t. } \omega_i &> 0 \\ \sum_i \omega_i &= 1, \end{aligned} \quad (3)$$

where W is a comprehensive evaluation index, F is the feature sequence, and ω is the weight. ω is usually determined by empirical knowledge. It can be seen from the derivation that W is positively related to the three evaluation indexes, so the greater W is, the more effective the description of the degradation process is, and the more favorable to the prediction the degradation trend is.

3. Proposed Method

3.1. Framework of Proposed Method. The goal of the present study is to develop a degradation index with a better trend and achieve simpler and more accurate prognostics. The complete procedure of bearing degradation trend prognostics is depicted in Figure 2. First, some typical features in the time and time-frequency domains are extracted based on raw vibration data. Second, cumulative transformation is proposed to gain the cumulative features with better trend characteristics from the original features. Following that, the feature fusion KPCA [33] is used to get a degradation trend index to reflect degradation process. Next, the mutual information method is adopted to choose the time delay and the process of CAO is employed to determine the embedded dimension. Meanwhile, the degradation index is reconstructed by PSR [34]. Finally, the reconstructed degradation index is input into ELM for bearing degradation trend prognostic.

To quantitatively assess the proposed method, three criteria are investigated, including mean absolute percent error (MAPE) and root mean square error (RMSE). The mathematical expressions are as follows:

$$\begin{aligned} e_{\text{MAPE}} &= \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \\ e_{\text{RMSE}} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \end{aligned} \quad (4)$$

where n is the number of data and y_i and \hat{y}_i represent the true and predicted values, respectively.

3.2. Cumulative Feature Extraction and Fusion. The failure of machinery represents the procedure of an abnormal phenomenon from the incipient failure to deterioration. Many types of signals are used to reflect the abnormal phenomena. Vibration signals acquired via sensors are the most widely used for condition monitoring, although they usually contain redundant dimensions, and the vibration signal generated by a certain bearing is often overwhelmed by noise or other components' vibrations; hence, it is seldom used directly.

In order to solve this problem, the useful features are extracted from the time domain, frequency domain, and time-frequency domain for fault diagnosis and prognostics. In this study, 12 time domain features and four frequency domain features are extracted, as shown in Table 1. In Table 1, x_i is the original time domain signal set (see Section 4.1), n is the sample points of x_i , μ and σ represent the mean value of x_i and standard deviation, respectively, $s(k)$ is the frequency spectrum of x_i , K is the number of spectral lines of $s(k)$, and X_{fm} represents the mean value of $s(k)$.

However, time domain and frequency domain analyses cannot simultaneously deal with the signals in the time and frequency domains; furthermore, some useful information could be discarded. To solve this problem, more information is obtained by time-frequency domain analysis. The most

TABLE 1: Features and corresponding formulas.

Feature	Formula
Mean absolute value (MAV)	$X_{\text{am}} = \frac{1}{n} \sum_n x_i $
Maximum	$X_{\text{max}} = \max(x)$
Variance	$X_{\text{var}} = \frac{1}{n-1} \sum_n (x_i - \mu)^2$
Skewness	$X_{\text{skew}} = E \left[\left(\frac{x_i - \mu}{\sigma} \right)^3 \right]$
Shape factor	$X_{\text{sf}} = \frac{X_{\text{rms}}}{X_{\text{abs}}}$
Clearance factor	$X_{\text{clf}} = \frac{X_{\text{max}}}{X_{\text{smr}}}$
Average frequency	$X_{\text{fm}} = \frac{\sum_{k=1}^K s(k)}{K}$
Frequency skew	$X_{\text{fs}} = \frac{\sum_{k=1}^K (s(k) - X_{\text{fm}})^3}{K X_{\text{fa}}^3}$
RMS	$X_{\text{rms}} = \sqrt{\frac{1}{n} \sum_n x_i^2}$
Peak-to-peak	$X_{\text{p-p}} = \max(x) - \min(x)$
Square mean root (SMR)	$X_{\text{smr}} = \left(\frac{1}{n} \sum_n \sqrt{ x_i } \right)^2$
Kurtosis	$X_{\text{kurt}} = \frac{1}{n} \sum_n \left(\frac{x_i - \mu}{\sigma} \right)^4$
Crest factor	$X_{\text{cf}} = \frac{X_{\text{max}}}{X_{\text{rms}}}$
Impulse factor	$X_{\text{if}} = \frac{X_{\text{max}}}{X_{\text{abs}}}$
Standard deviation frequency (SDF)	$X_{\text{fa}} = \sqrt{\frac{\sum_{k=1}^K (s(k) - X_{\text{fm}})^2}{K-1}}$
Frequency kurtosis	$X_{\text{fk}} = \frac{\sum_{k=1}^K (s(k) - X_{\text{fm}})^4}{K X_{\text{fa}}^4}$

commonly used time-frequency analysis methods are STFT, WT, and EMD. Among them, the difficulty of STFT is how to select the proper window function, and the main weakness of EMD is its high sensitivity to noise and mode mixing. In addition, WT is reported to have better applicability in dealing with vibration data from rotating machinery such as bearings [35]. In this study, a specific wavelet "db4" from Daubechies family wavelets in three levels is applied to extract the eight features of wavelet packet energy $P_{d,n}$.

Assuming that $E_{d,n}$ is the energy of the n th frequency band in the d th layer, which is defined as [36]

$$E_{d,n} = \sum_{k=1}^M \|W_{d,n}^k\|^2, \quad (5)$$

where $W_{d,n}^k$ is the wavelet coefficient of the k th discrete point of the decomposition signal $x_{d,n}(i)$ and M is the number of $x_{d,n}(i)$, the contribution of the n th band in the d th layer wavelet coefficient to the energy of the signal is defined as

$$P_{d,n} = \frac{E_{d,n}}{\sum_{n=1}^{2^d} E_{d,n}}. \quad (6)$$

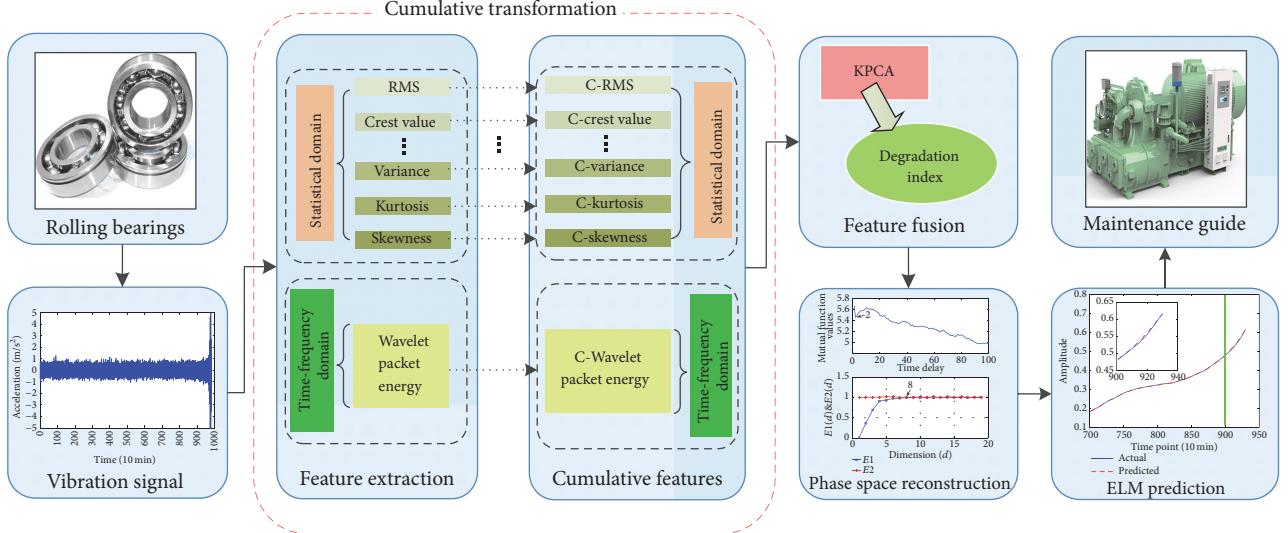


FIGURE 2: Flowchart for bearing degradation trend prognostic.

After the original features are extracted, each feature is mapped to its respective cumulative form by cumulative transformation algorithm, and a total of 24 cumulative features are obtained. Each cumulative feature contains partial information of bearing status and reflects the bearing degradation process from different aspects. To comprehensively describe the degradation process, KPCA is used to fuse all cumulative features and obtain a degradation index.

3.3. PSR and ELM. In the data-driven prognostic methods, a number of methods are used to build prediction models, for example, ARIMA, artificial neural network (ANN), and support vector machine (SVM). Among these methods, ELM is a new, simple, and effective single hidden layer feedforward neural network (SLFN) learning algorithm, which was first proposed by Huang et al. for both classification and regression purposes [37]. The structure parameter of ELM is randomly chosen, and iteration is not required. Compared with other gradient-based machine learning algorithms, ELM has the advantages of high computational efficiency, easy implementation, and good generalization performance. Owing to these advantages, ELM is employed to build a model to perform the prediction tasks in this study.

In addition, the time sequences prognostic supposes that the future values are determined by certain past values. The bearing degradation index is a one-dimensional time sequence, and the challenge is how to train the ELM model. PSR is able to extend a one-dimensional time sequence to a high-dimensional phase space that has an equivalent space with the original dynamic system in topology, and it can effectively grasp the nature of time series. Therefore, it is employed to deal with the model's import problem. The time lag and embedded dimension of the one-dimensional degradation index are determined by mutual information method and CAO's method, respectively. The degradation index reconstructed by PSR is then used as an import to the ELM model.

4. Experiment and Results

4.1. Experimental Setup. To assess the effectiveness of the proposed method, the vibration signals originating from Center for Intelligent Maintenance Systems (IMS), University of Cincinnati [38], are used. The designed experiment platform is shown in Figure 3 and the experiments are carried out under constant load and speed conditions. Four Rexnord ZA-2115 double-row bearings are erected on the axis of the test rig. The rotation speed is maintained at 2000 rpm. The axis and bearings bear a radial load of 6000 lbs through a spring mechanism, and all bearings are forced lubricated. The acceleration sensors are installed on the bearing housing. The new bearings are tested under this condition for full life cycle test. All failures occurred after exceeding designed lifetime of the bearing which is more than 100 million revolutions. It is found that the data in 10000 minutes before bearing failure can well reflect the whole failure process of the bearing, so the data in this period are analyzed. Each data set (x_i) of bearing 1 used in this study is 1-second vibration signal snapshots recorded at 10-minute intervals and consists of 20480 points with the sampling rate of 20 kHz.

4.2. Data Processing. The 12 time domain features, four frequency domain features, and eight time-frequency domain features are extracted from the vibration signal as described in Section 3.2. The original features are shown in Figures 4–7. From Figures 4–7, we can see some of the extracted original features (e.g., square mean root, RMS, skewness, standard deviation frequency, and WPE2), which can reflect the bearing degradation process to some extent. By contrast, some other features cannot reflect the bearing degradation trend appropriately. For example, frequency kurtosis and WPE1 are almost constant; the crest factor, clearance factor, and so on contain a lot of noise information and show no trendability. Moreover, the evolution of the original feature curves always

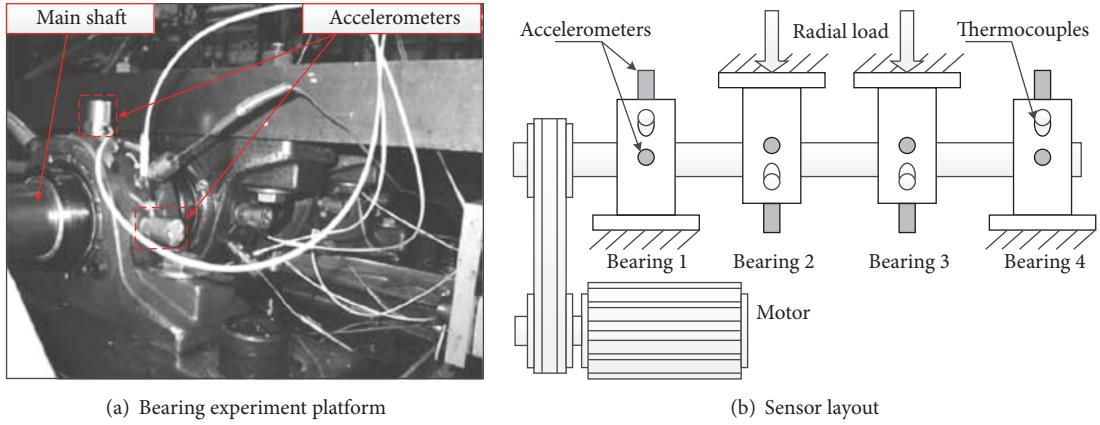


FIGURE 3: Illustration of the bearing experiment platform.

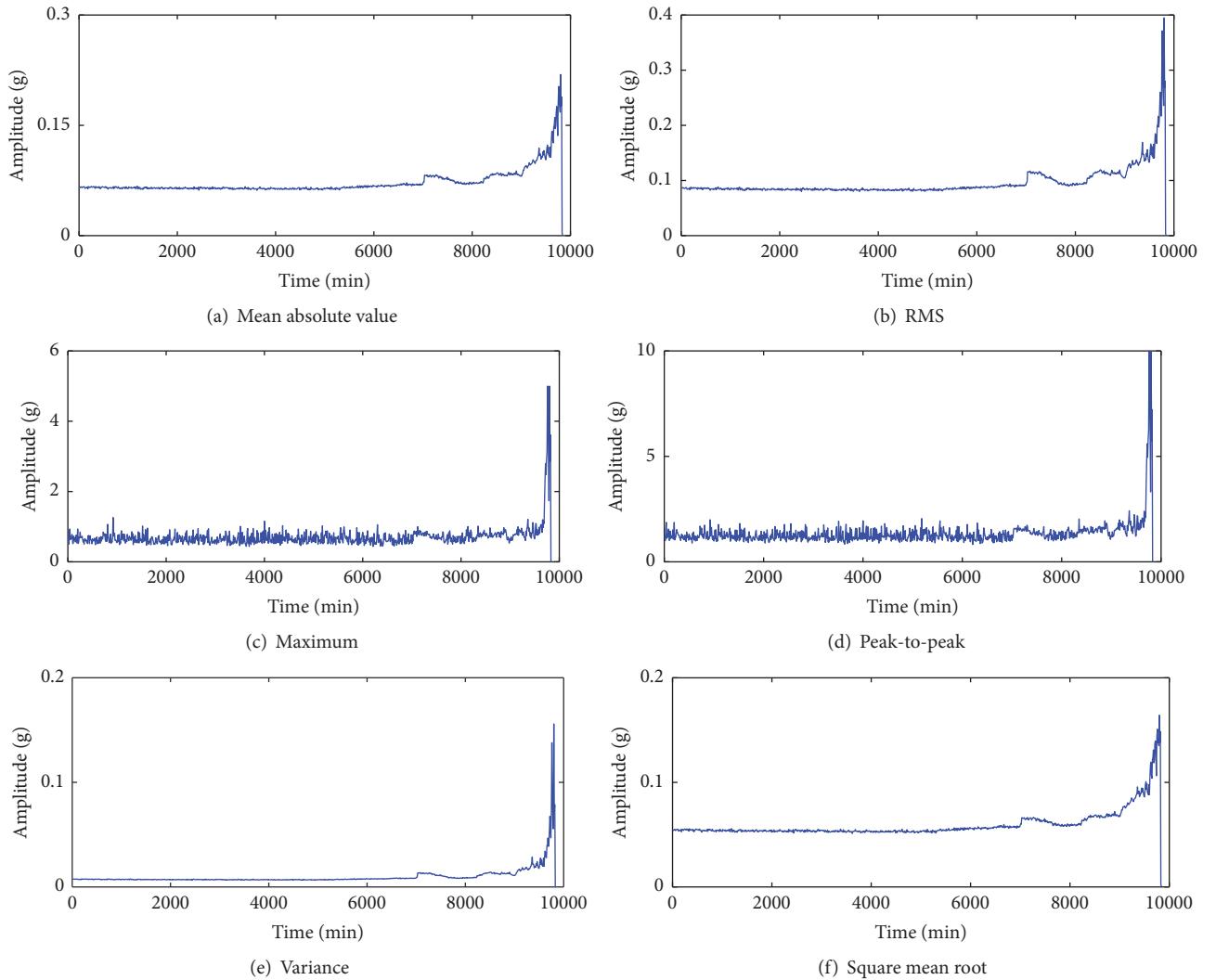


FIGURE 4: Six dimensional time domain features.

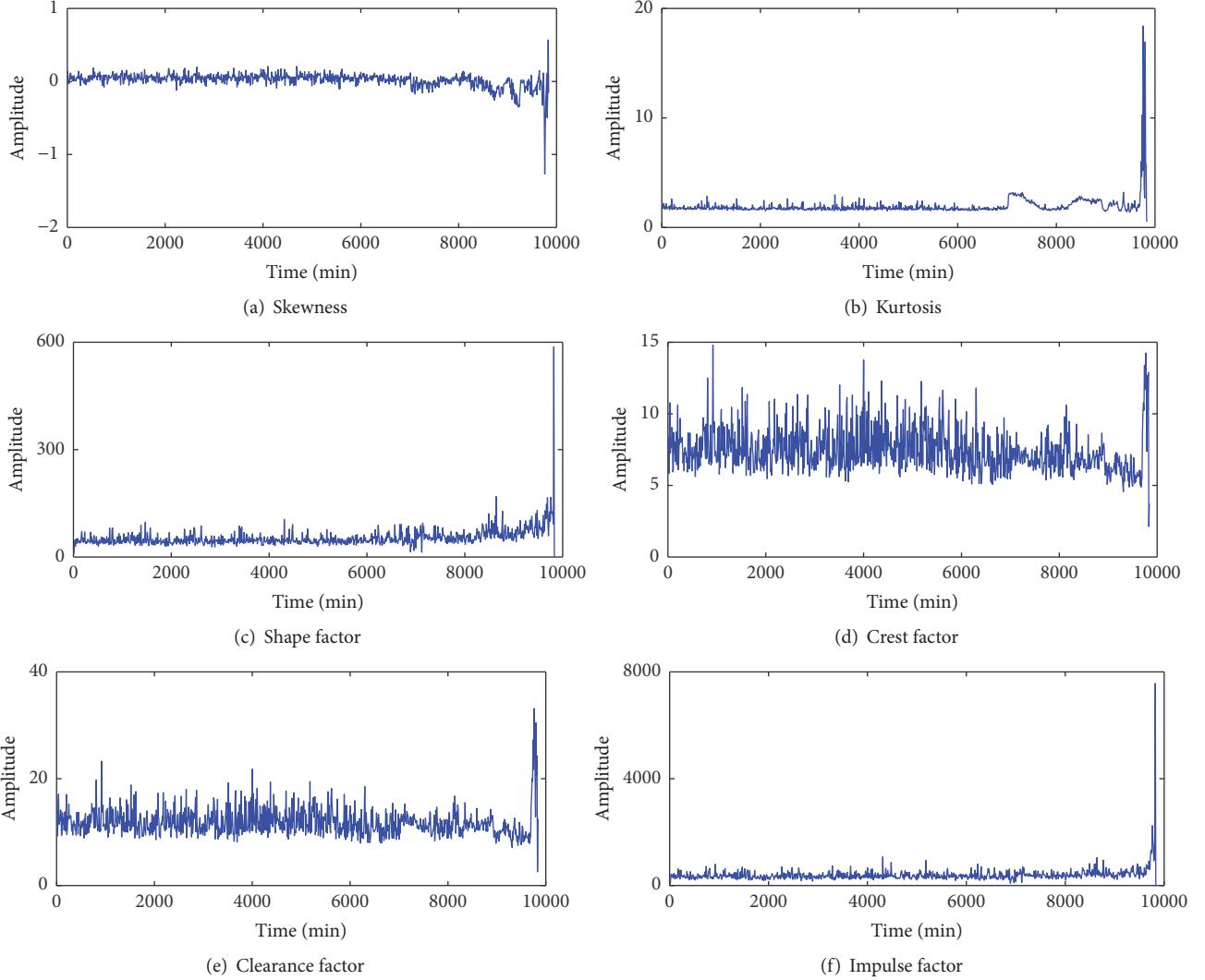


FIGURE 5: Six dimensionless time domain features.

shows some fluctuation and low trend characteristics owing to the background noise or some significantly stronger signals (e.g., gears and bars), which fail to track degradation trend effectively and even cause problems for the prognostic task.

In order to quantitatively describe the trend characteristics of the original features, a comprehensive evaluation indicator is proposed for feature evaluation. Table 2 shows the evaluation results of the eight optimal features, which indicate that the evaluation results and waveforms can show good agreement. It is found that even the comprehensive index of the eight optimal primitive features is relatively small.

To obtain the features that have better trend characteristics and can better reflect the bearing degradation process, the original features are transformed to build respective cumulative features. The eight optimal features and their corresponding cumulative features are shown in Figures 8 and 9. Compared with the original features, all the cumulative features show a smooth, monotonically increasing trend and can more clearly reflect the bearing degradation process.

The evaluation results of cumulative features are shown in the Table 3. Compared with Table 2, it can be seen that the monotonicity, trendability, and robustness of the eight features are obviously improved by the cumulative transformation. The monotonicity of each cumulative feature is increased to 1. Although there are some differences in the trendability, they have been greatly improved. The difference in robustness is small, and it remains at a high level.

In order to verify that the cumulative transformation has strong versatility, 24 original features of a rolling bearing are transformed to obtain the corresponding cumulative features, and the original and cumulative features are normalized and displayed by waterfalls. The feature waveforms are shown in Figures 10 and 11.

It can be seen that the trend between original features is different, and the characteristics of each feature also show more violent fluctuations, whereas cumulative features show a similar monotonically increasing trend, and the waveforms of cumulative features are very smooth.

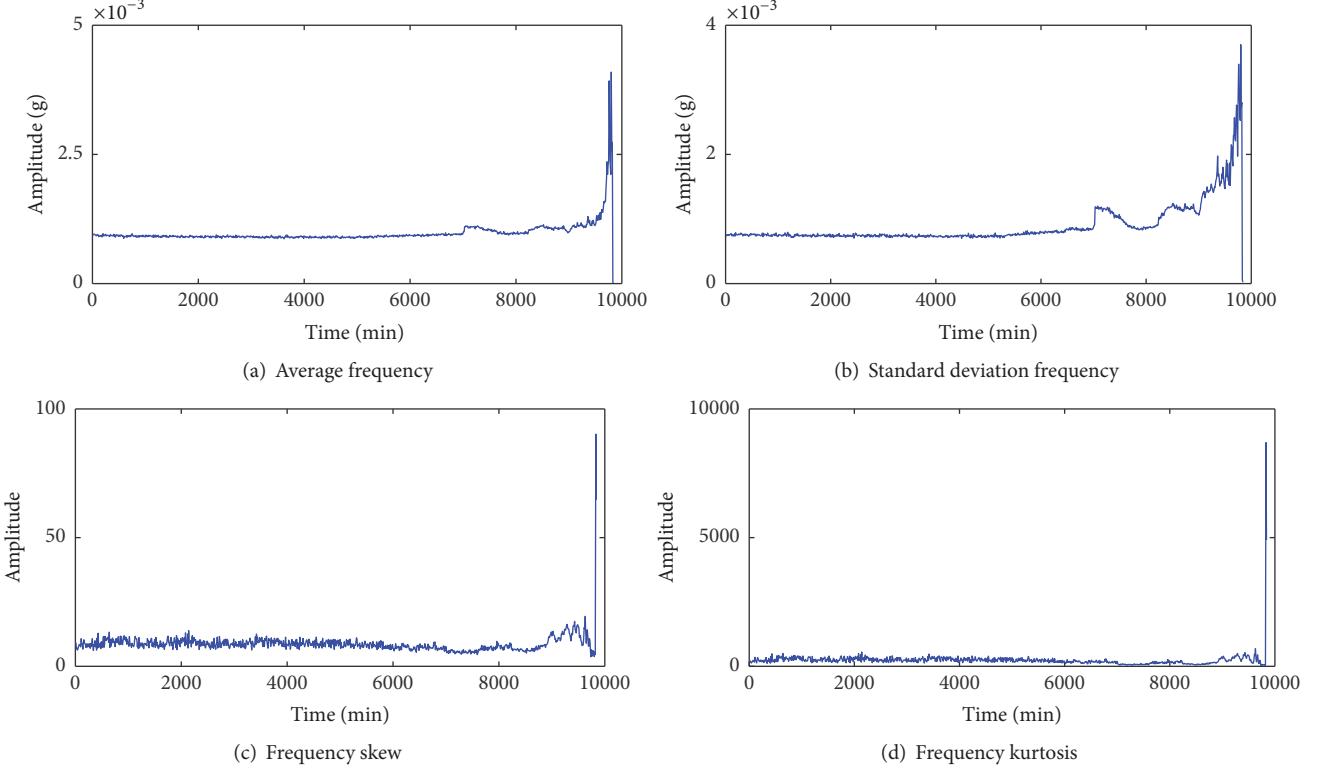


FIGURE 6: Four frequency domain features.

TABLE 2: Evaluation results of the first eight best bearing features.

Feature	Monotonicity	Trendability	Robustness	Comprehensive indicator
WPE2	0.0081	0.7834	0.9813	0.4353
WPE5	0.0163	0.7262	0.9837	0.4228
WPE8	0.0203	0.6945	0.9811	0.4147
SDF	0.0163	0.6211	0.9739	0.3893
SMR	0.0224	0.5842	0.9826	0.3830
MAV	0.0183	0.5807	0.9820	0.3798
RMS	0.0163	0.5527	0.9790	0.3698
Skewness	0.0142	0.4845	0.9564	0.3437

Different cumulative features can reflect the bearing degradation process from different aspects; the entire bearing degradation process cannot be described comprehensively by only one cumulative feature. Therefore, these cumulative features are fused to construct the fusion degradation index by the KPCA method, and the first principal component is selected as the degradation index. It can be seen from Figure 12 that the degradation index is monotonically increasing, and the different life states clearly characterize the entire life of the bearing. Accordingly, the index constructed by KPCA can reflect the degradation more effectively and comprehensively.

As can be seen from Figure 12, the bearing is in normal operation before the first 7000 minutes, after which the condition of the bearing suddenly changed and the curve increased sharply, which shows there are certain faults occurring in the bearing.

In the degradation trend prognostics, an improper input may lead to a bad prognostics result, and thus the key problem is how to validly import the one-dimensional degradation index into the model. PSR can effectively extend the one-dimensional time signal to its corresponding high-dimensional equivalent, which can be used to deal with the model's import problem. In the present study, the time delay τ is set as 2 through the mutual information method. Then, the embedded dimension d is set to 8 through the CAO method (as shown in Figure 13.). Based on the selected optimal delay time and embedded dimension, the ELM model is used to predict the bearing degradation trend.

4.3. Results and Discussion. A nonlinear function, that is, a sigmoid function, is selected as the activation function in ELM for bearing degradation trend prediction, and the number of hidden nodes is set as 10. As mentioned above, the

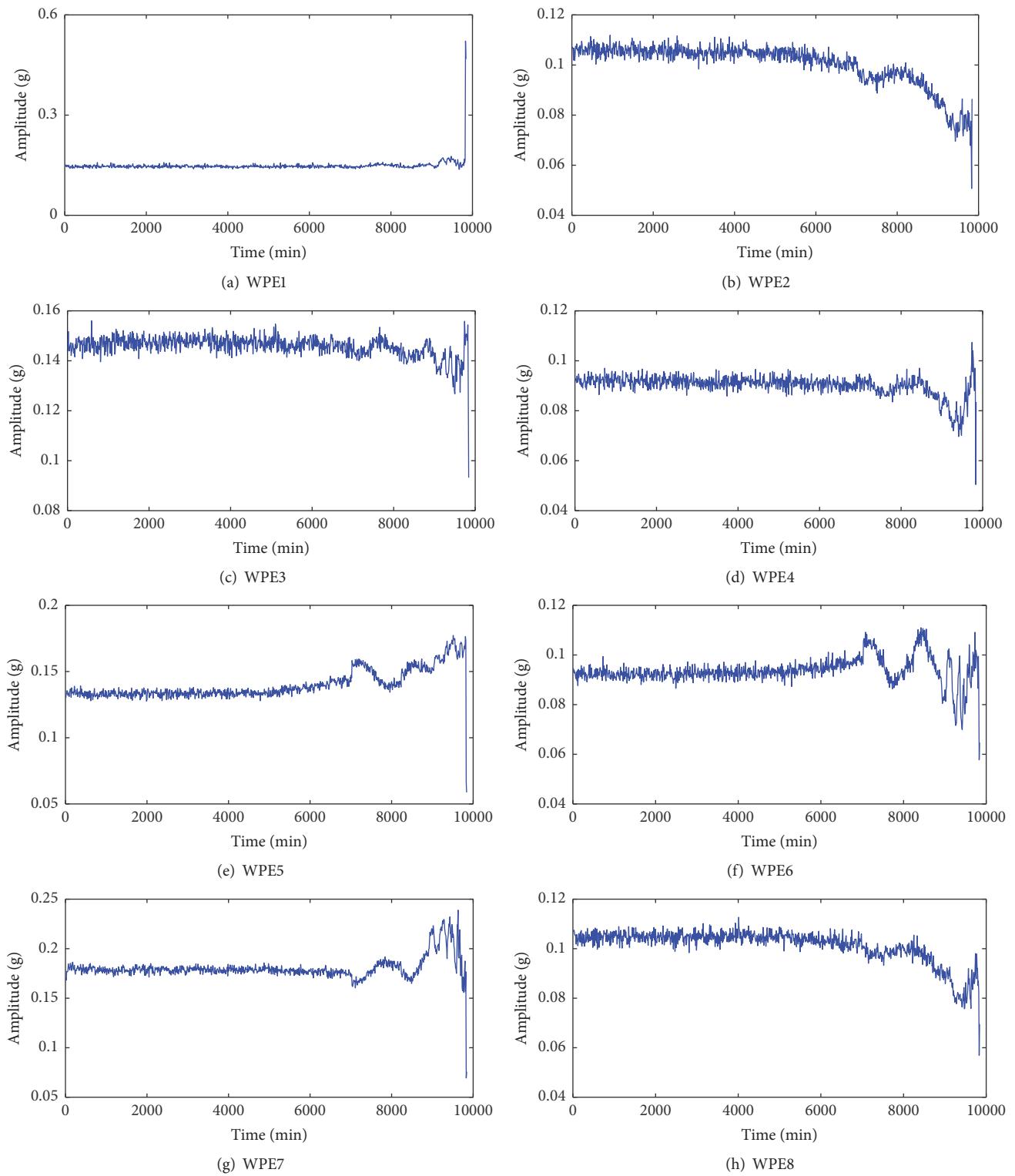


FIGURE 7: Eight wavelet packet energy features.

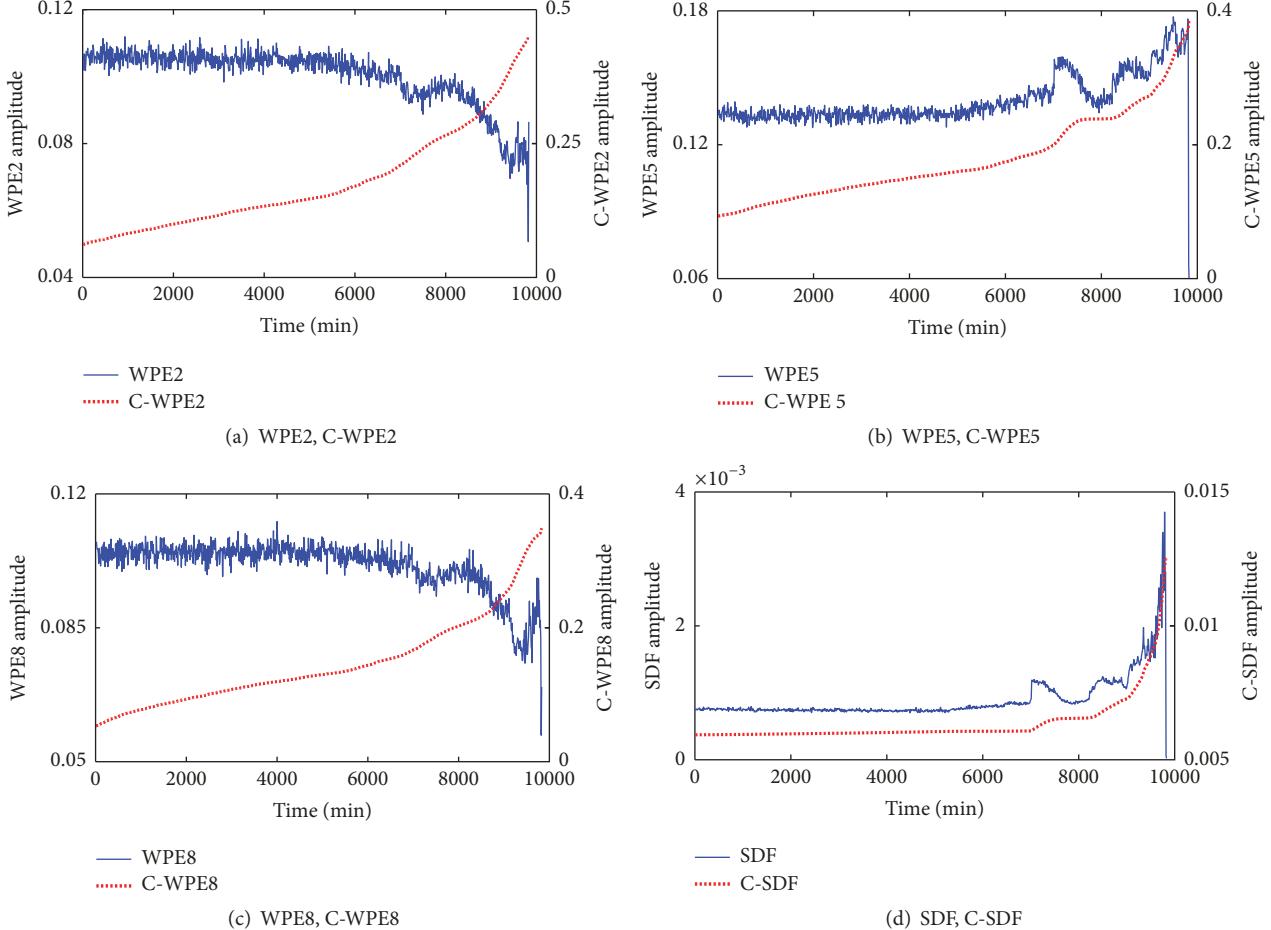


FIGURE 8: First four preferred features and corresponding cumulative features.

TABLE 3: Evaluation results of cumulative features of the first eight best original features.

Feature	Monotonicity	Trendability	Robustness	Comprehensive indicator
C-WPE2	1.0000	0.9302	0.9895	0.9770
C-WPE5	1.0000	0.9195	0.9929	0.9744
C-WPE8	1.0000	0.9295	0.9900	0.9769
C-SDF	1.0000	0.6228	0.9961	0.8861
C-SMR	1.0000	0.7268	0.9891	0.9159
C-MAV	1.0000	0.7495	0.9882	0.9225
C-RMS	1.0000	0.7797	0.9864	0.9312
C-Skewness	1.0000	0.9559	0.9868	0.9841

bearing is in a normal state before the 7000 minutes, so the points from 701 to 900 are used to train the ELM model and the following 30 points are used for testing.

The actual value and the predicted result of the original and cumulative fused index are shown in Figure 14. It can be seen that the predicted result of the cumulative fused index is consistent with the actual degradation curve.

The present method is compared with other methods to verify the advantages in the bearing degradation trend prognostics. Kurtosis, RMS, and WPE2 of the original and cumulative form are, respectively, used as the degradation

index to feed input to the ELM model, determining the import parameters of the ELM model with the mutual information method and CAO method. For illustration, Figures 15–17 show different prediction results.

From Figures 15–17, it can be seen that, owing to the fluctuation and bad trend, the degradation progression cannot be reflected clearly by the original features; furthermore, the prediction difficulty has also been increased, which leads to bad performance of the prognostic model. Compared with the original features, the cumulative features achieve better prediction result than the corresponding original

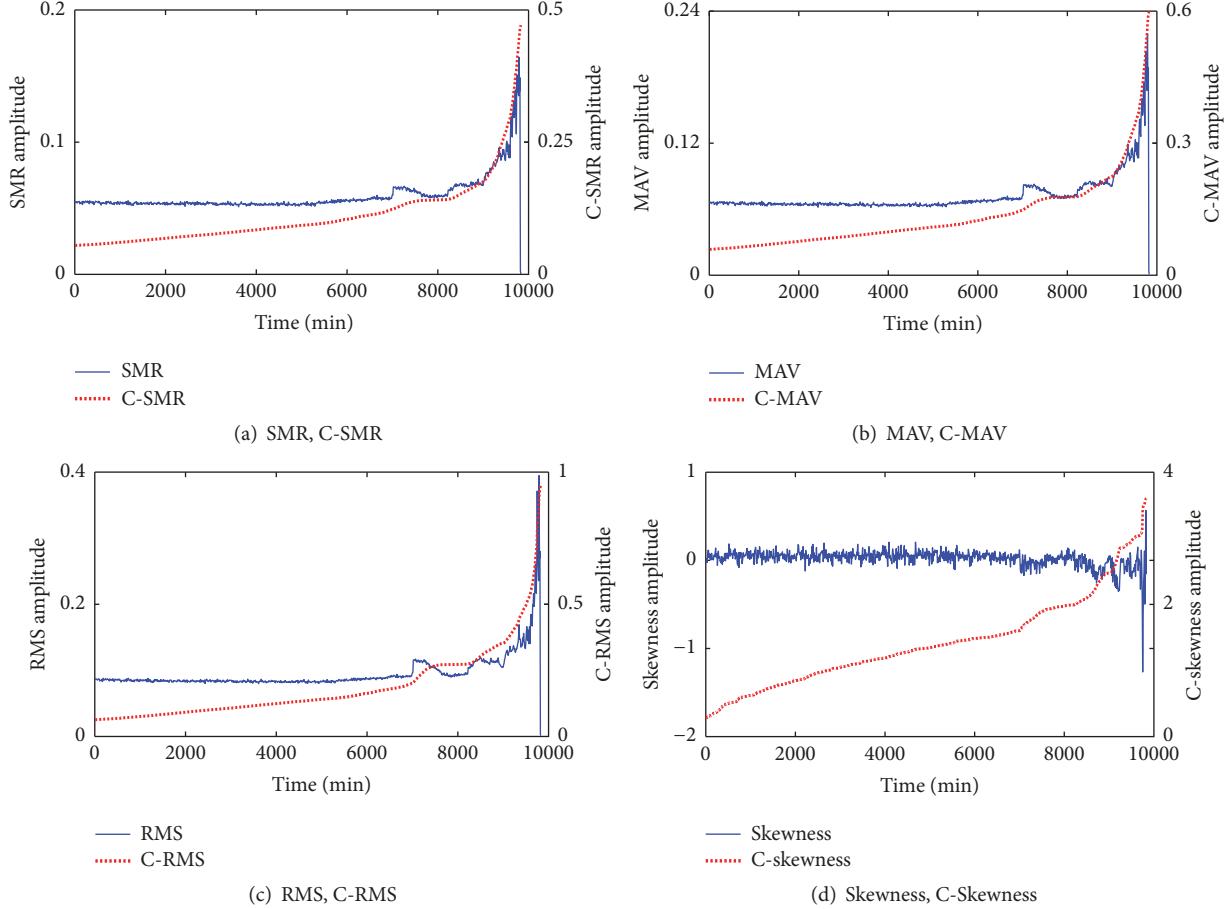
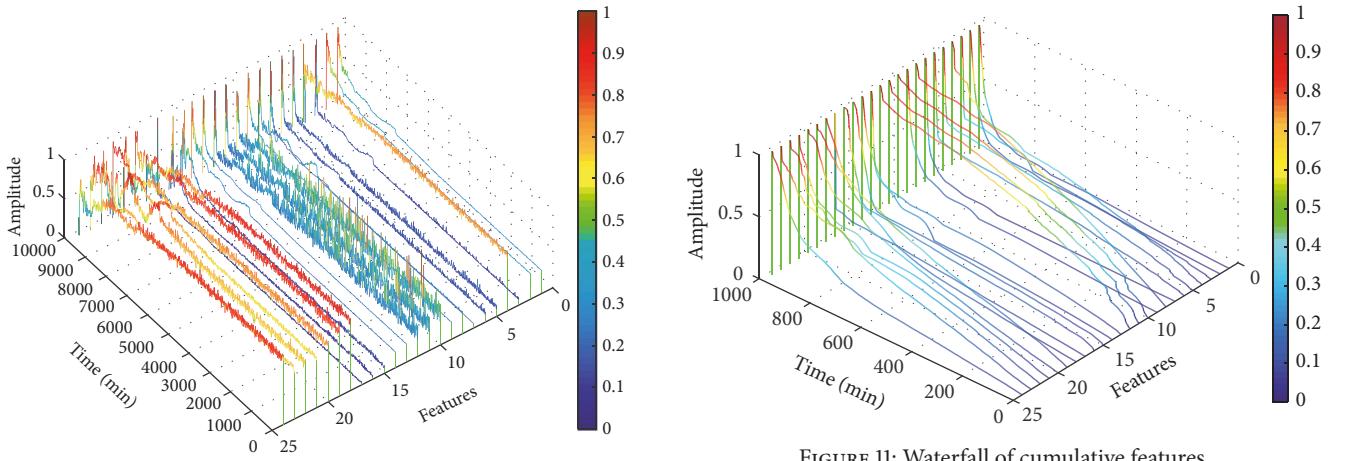


FIGURE 9: Fifth to eighth preferred features and corresponding cumulative features.



features. On the whole, the predicted results have the same trend as the actual cumulative features, and there are only minor differences between the predicted results and actual cumulative features.

To quantitatively evaluate the performance of different methods, the MAPE and RMSE values of different degradation indexes are summarized in Table 4. We note that the cumulative features can achieve higher accuracy than the original features.

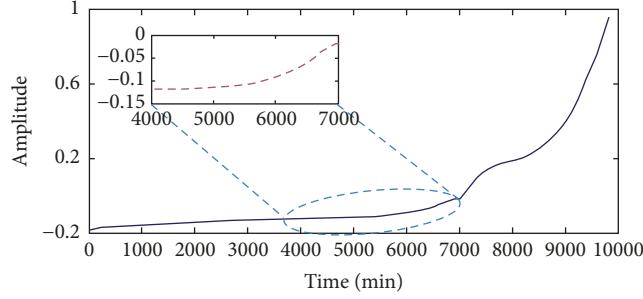
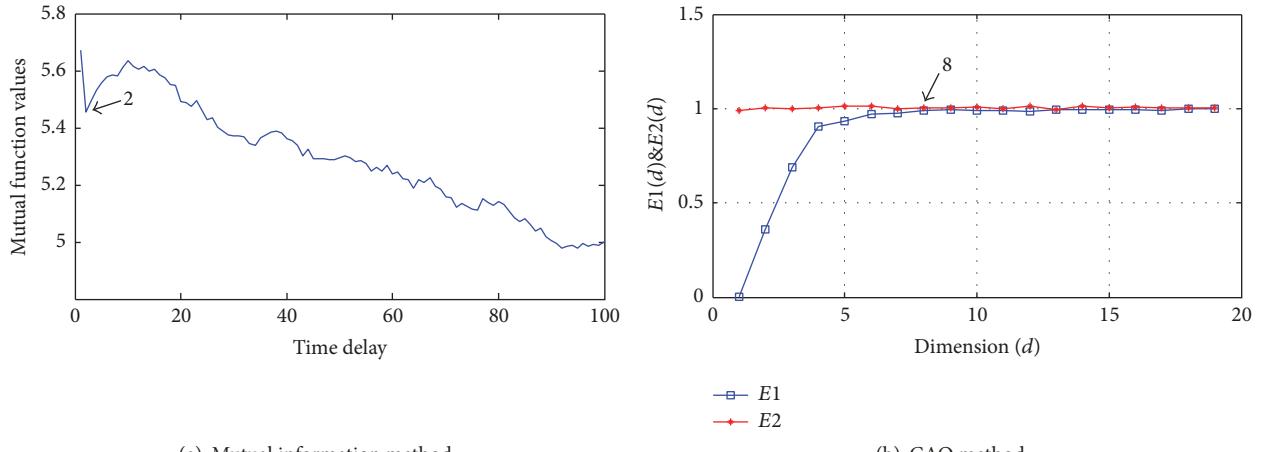


FIGURE 12: Degradation index constructed by KPCA.



(a) Mutual information method

(b) CAO method

FIGURE 13: Delay time and the embedded dimension of phase space reconstruction.

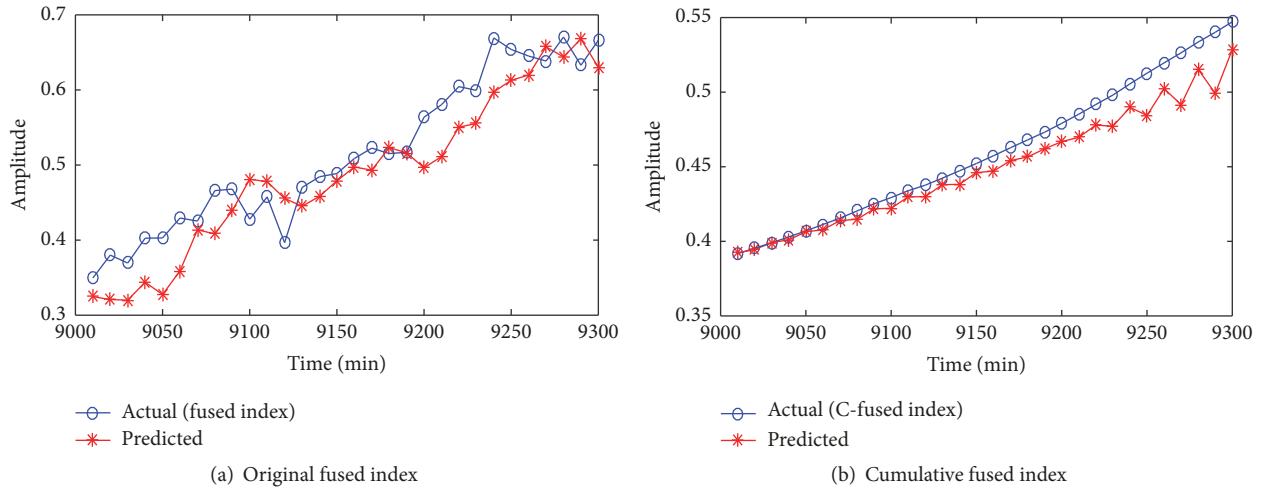


FIGURE 14: Prediction results of original fused index and cumulative fused index.

TABLE 4: Prediction error comparison of different degradation indexes.

Degradation index	MAPE	RMSE	Degradation index	MAPE	RMSE
Kurtosis	0.1643	1.0213	C-Kurtosis	0.0059	0.0625
RMS	0.1425	0.8548	C-RMS	0.0131	0.0051
WPE2	0.0815	0.0056	C-WPE2	0.0052	0.0032
Fused index	0.0799	0.0455	C-Fused Index	0.0226	0.0149

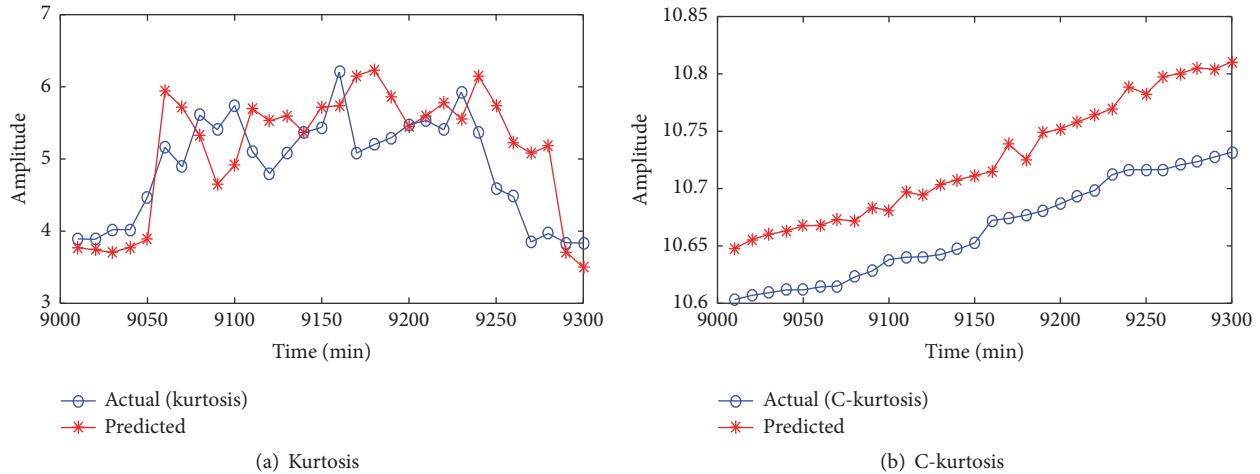


FIGURE 15: Prediction results of kurtosis and C-kurtosis.

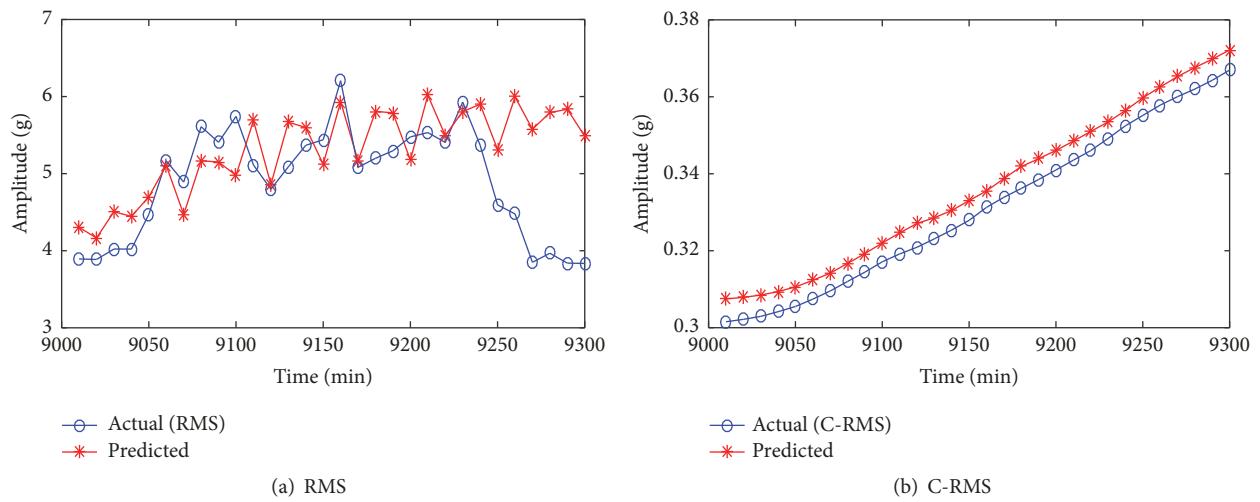


FIGURE 16: Prediction results of RMS and C-RMS.

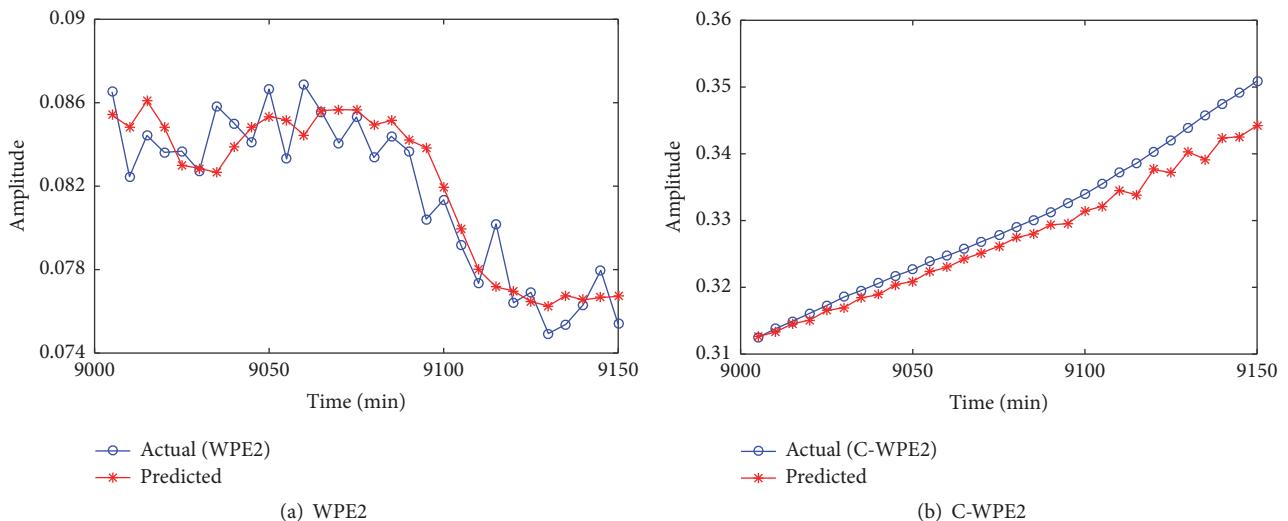


FIGURE 17: Prediction results of WPE2 and C-WPE2.

The reason for this is that the cumulative features are calculated based on the whole bearing life data, unlike the traditional features, which are based on a piece of data; thus, the cumulative transformation can partly ignore data fluctuation and extract significant monotonic trend information. Moreover, the KPCA can effectively fuse different features that describe the signal characteristics from different aspects and get more useful information. Therefore, the proposed degradation index can effectively reflect the bearing degradation trend and benefit the prognostic tasks.

5. Conclusion

In this study, a novel rolling bearing degradation prognostics method based on cumulative transformation and KPCA is proposed. More specifically, cumulative transformation and KPCA are integrated to obtain a degradation index to reflect the bearing degradation process. A prediction model based on ELM and PSR is built to perform the prognostics task. The following conclusions can be drawn:

(1) The features of the time domain, frequency domain, and time-frequency domain are extracted, and cumulative transformation is proposed to achieve features with high monotonicity and trendability from the original features, which can benefit the prognostic tasks.

(2) A new degradation index system is constructed, which fuses multidimensional cumulative features by KPCA and reflects the bearing degradation process properly.

(3) A prediction model based on PSR and ELM is proposed to achieve bearing degradation trend prediction. The whole-life experiment of the bearing shows that the proposed method reflects the bearing degradation process clearly and achieves a good balance between model accuracy and complexity.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the financial support provided by National Natural Science Foundation of China (nos. 51674277 and 51504274) and National Key Research and Development Program of China (no. 2017YFC0805803).

References

- [1] M. Ge, Y. Xu, and R. Du, "An intelligent online monitoring and diagnostic system for manufacturing automation," *IEEE Transactions on Automation Science and Engineering*, vol. 5, no. 1, pp. 127–138, 2008.
- [2] X. An, H. Zeng, and C. Li, "Demodulation analysis based on adaptive local iterative filtering for bearing fault diagnosis," *Measurement*, vol. 94, no. 12, pp. 554–560, 2016.
- [3] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 199–221, 2004.
- [4] K. Javed, R. Gouriveau, N. Zerhouni, and P. Nectoux, "A feature extraction procedure based on trigonometric functions and cumulative descriptors to enhance prognostics modeling," in *Proceedings of the 2013 IEEE International Conference on Prognostics and Health Management, PHM 2013*, pp. 1–7, June 2013.
- [5] S. Dong and T. Luo, "Bearing degradation process prediction based on the PCA and optimized LS-SVM model," *Measurement*, vol. 46, no. 9, pp. 3143–3152, 2013.
- [6] X. H. Liang, M. J. Zuo, and M. R. Hoseini, "Vibration signal modeling of a planetary gear set for tooth crack detection," *Engineering Failure Analysis*, vol. 48, pp. 185–200, 2015.
- [7] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—reviews, methodology and applications," *Mechanical Systems and Signal Processing*, vol. 42, no. 1-2, pp. 314–334, 2014.
- [8] C. Shen, D. Wang, F. Kong, and P. W. Tse, "Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier," *Measurement*, vol. 46, no. 4, pp. 1551–1564, 2013.
- [9] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [10] K. Javed, R. Gouriveau, N. Zerhouni, and P. Nectoux, "Enabling health monitoring approach based on vibration data for accurate prognostics," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 647–656, 2015.
- [11] S. Kumar and M. Pecht, "Modeling approaches for prognostics and health management of electronics," *International Journal of Performability Engineering*, vol. 6, no. 5, pp. 222–229, 2010.
- [12] F. Camci, K. Medjaher, N. Zerhouni, and P. Nectoux, "Feature evaluation for effective bearing prognostics," *Quality and Reliability Engineering International*, vol. 29, no. 4, pp. 477–486, 2013.
- [13] B. Zhang, L. Zhang, and J. Xu, "Degradation feature selection for remaining useful life prediction of rolling element bearings," *Quality and Reliability Engineering International*, vol. 32, no. 2, pp. 547–554, 2016.
- [14] J. Yan, C. Guo, and X. Wang, "A dynamic multi-scale Markov model based methodology for remaining life prediction," *Mechanical Systems and Signal Processing*, vol. 25, no. 4, pp. 1364–1376, 2011.
- [15] P. Li, F. Kong, Q. He, and Y. Liu, "Multiscale slope feature extraction for rotating machinery fault diagnosis using wavelet analysis," *Measurement*, vol. 46, no. 1, pp. 497–505, 2013.
- [16] H. Hong and M. Liang, "Fault severity assessment for rolling element bearings using the Lempel-Ziv complexity and continuous wavelet transform," *Journal of Sound and Vibration*, vol. 320, no. 1-2, pp. 452–468, 2009.
- [17] W. Guo, P. W. Tse, and A. Djordjevich, "Faulty bearing signal recovery from large noise using a hybrid method based on spectral kurtosis and ensemble empirical mode decomposition," *Measurement*, vol. 45, no. 5, pp. 1308–1322, 2012.
- [18] E. Sutrisno, H. Oh, A. S. S. Vasan, and M. Pecht, "Estimation of remaining useful life of ball bearings using data driven methodologies," in *Proceedings of the IEEE Conference on Prognostics and Health Management (PHM '12)*, pp. 1–7, Denver, Colo, USA, June 2012.
- [19] 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.

- [20] J. Ben Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network," *Mechanical Systems and Signal Processing*, vol. 56, pp. 150–172, 2015.
- [21] H. Qiu, J. Lee, J. Lin, and G. Yu, "Robust performance degradation assessment methods for enhanced rolling element bearing prognostics," *Advanced Engineering Informatics*, vol. 17, no. 3-4, pp. 127–140, 2003.
- [22] S. Sassi, B. Badri, and M. Thomas, "Tracking surface degradation of ball bearings by means of new time domain scalar indicators," *International Journal of COMADEM*, vol. 11, no. 3, pp. 36–45, 2008.
- [23] S. Zhang, Y. Zhang, and D. Zhu, "Residual Life Prediction for Rolling Element Bearings Based on an Effective Degradation Indicator," *Journal of Failure Analysis and Prevention (JFAP)*, vol. 15, no. 5, pp. 722–729, 2015.
- [24] Z. Xi, R. Jing, P. Wang, and C. Hu, "A copula-based sampling method for data-driven prognostics," *Reliability Engineering & System Safety*, vol. 132, no. 4, pp. 72–82, 2014.
- [25] L. Liao, "Discovering prognostic features using genetic programming in remaining useful life prediction," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2464–2472, 2014.
- [26] C. Sun, Z. He, H. Cao, and Z. Zhang, "A non-probabilistic metric derived from condition information for operational reliability assessment of aero-engines," *IEEE Transactions on Reliability*, vol. 64, no. 1, pp. 1–15, 2014.
- [27] Y. Wang and H. Pham, "Imperfect preventive maintenance policies for two-process cumulative damage model of degradation and random shocks," *International Journal of Systems Assurance Engineering and Management*, vol. 2, no. 1, pp. 66–77, 2011.
- [28] X. Ni, J. Zhao, X. Zhang, X. Lv, and J. Zhao, "System degradation process modeling for two-stage degraded mode," in *Proceedings of the 2014 Prognostics and System Health Management Conference, PHM 2014*, pp. 186–189, August 2014.
- [29] A. Ragab, S. Yacout, M.-S. Ouali, and H. Osman, "Prognostics of multiple failure modes in rotating machinery using a pattern-based classifier and cumulative incidence functions," *Journal of Intelligent Manufacturing*, pp. 1–20, 2016.
- [30] M. X. Zhu, J. N. Zhang, Y. Li, and Y. H. Wei, "Partial discharge signals separation using cumulative energy function and mathematical morphology gradient," *IEEE Transactions on Dielectrics Electrical Insulation*, vol. 23, no. 1, pp. 482–493, 2016.
- [31] S. Porotsky and Z. Bluvband, "Remaining useful life estimation for systems with non-trendability behaviour," in *Proceedings of the 2012 IEEE International Conference on Prognostics and Health Management: Enhancing Safety, Efficiency, Availability, and Effectiveness of Systems Through PHM Technology and Application, PHM 2012*, pp. 1–6, June 2012.
- [32] J. B. Coble, *Merging data sources to predict remaining useful life—an automated method to identify prognostic parameters*, [Ph.D. thesis], University of Tennessee, 2010.
- [33] Q. He, F. Kong, and R. Yan, "Subspace-based gearbox condition monitoring by kernel principal component analysis," *Mechanical Systems and Signal Processing*, vol. 21, no. 4, pp. 1755–1772, 2007.
- [34] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical systems and Turbulence*, D. A. Rand and L. S. Young, Eds., vol. 898 of *Lecture Note in Mathematics*, pp. 366–381, Springer, Berlin, Germany, 1981.
- [35] A. Bellini, F. Filippetti, C. Tassoni, and G.-A. Capolino, "Advances in diagnostic techniques for induction machines," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4109–4126, 2008.
- [36] L. Liao and J. Lee, "A novel method for machine performance degradation assessment based on fixed cycle features test," *Journal of Sound and Vibration*, vol. 326, no. 3-5, pp. 894–908, 2009.
- [37] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [38] <http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository>.

