Research Article

Data-Driven Bearing Fault Diagnosis of Microgrid Network Power Device Based on a Stacked Denoising Autoencoder in Deep Learning and Clustering by Fast Search without Data Labels

Fan Xu ^{1,2}, Xin Shu,^{3,4} Xin Li,⁵ and Xiaodi Zhang⁶

¹School of Automation, China University of Geosciences, Wuhan 730072, China

²School of Data Science, City University of Hong Kong, Tat Chee Avenue, Kowloon 990777, Hong Kong, China

³China Railway Major Bridge Engineering Group Co., Ltd., Wuhan 730072, China

⁴State Key Laboratory for Health and Safety of Bridge Structures, Wuhan 730072, China

⁵School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 730072, China

⁶State Grid Beijing Electric Power Company, Beijing 100080, China

Correspondence should be addressed to Fan Xu; fanxu8@cityu.edu.hk

Received 18 May 2020; Revised 6 September 2020; Accepted 5 October 2020; Published 2 November 2020

Academic Editor: Atila Bueno

Copyright © 2020 Fan Xu 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.

The traditional health indicator (HI) construction method of electric equipment devices in microgrid networks, such as bearings that require different time-frequency domain indicators, needs several models to combine. Therefore, it is necessary to manually select appropriate and sensitive models, such as time-frequency domain indicators and multimodel fusion, to build HIs in multiple steps, which is more complicated because sensitivity characteristics and suitable models are more representatives of bearing degradation trends. In this paper, we use the stacked denoising autoencoder (SDAE) model in deep learning to construct HI directly from the microgrid power equipment of raw signals in bearings. With this model, the HI can be constructed without multiple model combinations or the need for manual experience in selecting the sensitive indicators. The SDAE can extract the representative degradation information adaptively from the original data through several nonlinear hidden layers automatically and approximate complicated nonlinear functions with a small reconstruction error. After the SDAE extracts the preliminary HI, a model is needed to divide the wear state of the HI constructed by the SDAE. A cluster model is commonly used for this, and unlike most clustering methods such as k-means, k-medoids, and fuzzy c-means (FCM), in which the clustering center point must be preset, cluster by fast search (CFS) can automatically find available cluster center points automatically according to the distance and local density between each point and its clustering center point. Thus, the selected cluster center points are used to divide the wear state of the bearing. The root mean square (RMS), kurtosis, Shannon entropy (SHE), approximate entropy (AE), permutation entropy (PE), and principal component analysis (PCA) are also used to construct the HI. Finally, the results show that the performance of the method (SDAE-CFS) presented is superior to other combination HI models, such as EEMD-SVD-FCM/k-means/k-medoids, stacked autoencoder-CFS (SAE-CFS), RMS, kurtosis, SHE, AE, PE, and PCA.

1. Introduction

The microgrid power equipment (bearings) is a very commonly used mechanical device in the industrial field, but it wears down easily. Its status and reliable operation are of great significance in ensuring power system safety and reducing equipment operating costs. As bearing running time increases, performance will gradually degrade. The quality of

bearings also affects the operation of adjacent entire power system [1]. An indicator is commonly used to assess the health status of the bearing, which can provide a sound foundation for bearing performance degradation assessment (PDA) [2]. The signals are often used as an indicator to monitor the status because the quality of the bearing vibration signal can indicate bearing health in the PDA [3, 4]. Many models, including various statistical parameters and

mechanical signal-processing methods, are often used to extract useful degradation features for constructing bearing health indicators. Root mean square (RMS) and kurtosis are the most commonly used time-domain statistical parameters and can be considered for monitoring the health status of bearings by using a vibration signal. Williams et al. used RMS and kurtosis and demonstrated that they could effectively reflect and extract the bearing fault features [5]. Tse and Wang developed a method based on RMS to construct a health indicator for bearing PDA models, after the original vibration signal has been filtered from a ranged frequency band [6]. In [6], RMS is used to track the degradation status at the point when the vibration energy has changed. Shen et al. used multiple time indicators, including the RMS, to extract the useful degradation characteristics of bearings. They used a multivariable support vector machine to predict the remaining useful life of microgrid power equipment (bearings) [7]. Lei et al. also considered RMS for extracting the degradation trend to evaluate the degradation status of bearings. They demonstrated that the RMS could provide useful status information in the degradation stage, but not under normal conditions [8]. In [9], RMS and kurtosis were used to monitor bearings with a low filter, to filter out the useless frequency band and retain the useful band, according to the bearing working frequency. Lei et al. used multidimensional time-frequency, including RMS and kurtosis, for bearing fault feature extraction and fault diagnosis [10]. Other statistical indicators such as entropy models, including Shannon entropy (SHE), approximate entropy (AE), and permutation entropy (PE), are useful ways to assess the gradation trend of a mechanical device. AE can represent the regularity of multidimensional time series and contains more time-related information by using a coarse-graining operation for a time series [11, 12]. Yan et al. presented a health indicator for bearing PDA based on AE. As the working condition of the bearing deteriorates, the degree of wear increases and the number of frequency components contained in the vibration signal will increase, eventually causing its regularity to decrease and its corresponding AE value to increase. A detailed analysis of the impact of parameters on the AE model is described in this study [13]. The computational efficiency of PE is higher than that of AE as it only uses the order of the values and it is robust under a nonlinear distortion of the signal [14]. Yan et al. used PE as a health indicator to track the degradation of bearings, as it can describe the complexity of a vibration signal measured in a physical system by using phase-space reconstruction and takes into account the nonlinear behavior of the vibration signal [15]. Many traditional mechanical signal-processing models, such as wavelet transfer, EMD [16], and ensemble empirical mode decomposition (EEMD) [17], are often used to construct health indicators for the bearing PDA model. Qiu et al. used wavelet transfer to filter the noise, which could mask the bearing vibration signal, and then SHE was used to optimize the Morlet wavelet shape to extract the weak fault feature [18]. Lou et al. used the wavelet and fuzzy inference to extract useful fault features for bearing fault diagnosis [19]. Compared with wavelet transfer, EMD can decompose the original vibration signal into intrinsic mode

functions (IMFs) adaptively without wavelet bias function selection and decomposition level choices. Wang et al. [20] used the EMD to decompose the vibration signal into IMFs and then applied singular value decomposition (SVD) to calculate the singular values of the IMFs. After these two steps, the Mahalanobis distance was used to construct a health indicator for bearing PDA. Rai et al. [21] considered the EEMD for decomposing the bearing vibration signal into IMFs and then SVD was used to calculate the singular values (SVs). Finally, k-medoids were used to find the available cluster center points, which can reflect the bearing degradation status by using a health indicator, known as the confidence value (CV), to build the bearing PDA model.

These have achieved significant success in health indicator construction and are used for bearing PDA. However, there are some problems with these PDA models.

- (1) Time-frequency domain indicators are commonly used methods and must select sensitive indicators to show the difference between different faults and to improve the accuracy of fault classification. Wei et al., for example, considered the self-weight to evaluate and judge the quality of the time-frequency domain indicators for fault diagnosis [22]. The optimized indicators can then be used to diagnose faults and to improve the identification accuracy. Tse et al. selected commonly used sensitivity time-frequency indicators to extract the oil sand pump degradation trend with principal component analysis (PCA) [23]. Thus, manual experience can be applied in filtering the sensitivity indexes selected, ensuring they can achieve and improve the performance of fault diagnosis and the PDA of a device in a complex mechanical system.
- (2) The operating environment of complex mechanical systems is variable, and the commonly used time-frequency domain indicators will exhibit different advantages and disadvantages depending on the operating conditions. Therefore, relying on manual experience to select sensitive features applicable to complex and varied mechanical equipment operating environments is difficult.
- (3) In addition, commonly used models should be combined to extract useful degradation information from the raw vibration, such as EEMD combined with SVD and a clustering model to construct the health indicator for evaluating the PDA of bearings. Thus, these combined models to some extent lack versatility.

Therefore, these fusion time-frequency indicators and combined models are complicated and are reliant on manual experience.

To overcome these drawbacks, in this study, the stacked denoising autoencoder (SDAE) [24, 25] from deep learning is used to extract the initial degradation level for bearing PDA directly from the raw vibration signal, without selecting various indicators and model combinations. The deep architectures of the SDAE enable it to extract the

representative information adaptively from the original data through several nonlinear hidden layers and approximate complicated nonlinear functions, with a small reconstruction error and without manual experience and data labels [26]. Thus, there is no need for manual experience or prior knowledge of the SDAE. The bearing degradation trend can be extracted by using encoder and decoder processes and reconstructing the input through several hidden layers. The SDAE is an unsupervised model, which is robust when there is noise in the original vibration signal. It is an improved model based on a stacked autoencoder (SAE), which is a basic deep learning model that is widely used in different domains, such as fault feature extraction and fault diagnosis.

Feng et al. [26] used the SAE to extract the bearing fault information from the frequency-domain signal directly after fast Fourier transform (FFT) without time-frequency indicator selection. Lv et al. used a weighted time series fault diagnosis method based on the SAE. The proposed model in this study not only captures the high-order correlations among monitoring variables but also uses the time correlations among samples [29]. To further explore more representative fault characteristics using the SAE network, Qi et al. combined EEMD and the autoregressive (AR) model to preprocess the collected nonstationary vibration signals and obtain AR parameters based on intrinsic extraction. The decomposed mode function components are selected as an input feature of the SAE network [27, 28].

The vibration data gathered from these various engineering devices contain noise, and the empty and destroyed data collected will make the analysis difficult. The SDAE is robust because it destroys the original data into zero according to the denoising probability and reconstructs the input data by using encoding and decoding. This denoising operation improves the SAE so that it can learn a more robust representation. Therefore, the SDAE is more robust and stable than the SAE. Xu et al. used the SDAE to train the bearing vibration signal after a fast Fourier transform (FFT) and extracted useful fault features under various conditions. The SDAE was found to reduce the feature dimension to 2 directly without PCA. The clustering model was then used for bearing fault diagnosis [29]. The authors also demonstrated that the SDAE is more robust than the SAE. Additionally, the SDAE has been used successfully in other domains, including multimodal video classification, physiological signal processing, and 3-D object identification [30–34]. However, few studies have focused on the PDA of bearings when using the SDAE, and most instead consider r bearing fault diagnosis. The SDAE has had many successful applications, so it is used in this study to extract the initial degradation trend for bearings directly from the original vibration signal.

After the SDAE has extracted the degradation trend to construct a health status without a data label, the clustering model is a common method of building a health indicator for bearing PDA and for determining the degradation status by calculating the distance between each sample point and its cluster center point. Pan et al. proposed a model based on wavelet transfer and fuzzy c -means (FCM) for bearing PDA [31] and developed a method based on EEMD, SVD, and

k -medoids clustering models to construct a health indicator for bearing PDA. They demonstrated that the proposed model was better than other models such as EEMD-SVD- k -means [21], RMS, and kurtosis. However, clustering models such as FCM, k -means, and k -medoids require cluster number selection before calculation. Typically, the three degradation statuses of Normal, Slight, and Severe are suitable for bearing degradation division. These clustering models have successfully been used in the PDA of bearings.

However, these clustering models need to preset the number of clusters. Several states may sometimes exceed these predefined three statuses. Manually selecting the number of degraded trend states by using and selecting the number of cluster center points will lead to erroneous judgments. For example, a bearing may have only two degradation stages such as Normal and Severe at one time, and not the three statuses. This preset three-degradation status method cannot adaptively meet the requirements of dynamic changes in different situations of data acquisition under complex project operating conditions.

To solve this problem, in this paper, we use a clustering by fast search (CFS) model for bearing PDA. CFS can find the available cluster center number automatically according to the local density and the distance between any two samples [34]. CFS has been successfully used in other domains. To solve most current clustering techniques, only static data can be processed into clusters. Zhang et al. proposed a CFS model based on the peak of k -medoids to integrate the current model into the previous model to achieve final clustering and applied it to industrial dynamic acquisition data analysis. Effectively analyzing these data can help improve industrial services and ensure the system has no possibility of symptomatic failure [35]. Xu et al. used the EEMD with base-scale entropy to extract the useful fault information of bearings under different conditions. The base-scale entropy-based feature vector is then used as the input of CFS for fault diagnosis. CFS has been successfully applied in various areas, but few studies report that CFS has been used for the PDA of bearings [36]. Thus, in this study, CFS is used to find the available cluster center point and then construct a health indicator, as in [21], to evaluate the degradation status.

As mentioned above, an unsupervised method based on the SDAE and CFS is proposed to construct a health indicator for bearing PDA without data labels or prior knowledge.

The contribution of this paper is as follows:

- (1) The SDAE extracts the bearing degradation from the original vibration signal directly without manual intervention, as is often used to select sensitive time-frequency indicator and combine several available models to construct HI. The SDAE and CFS are used in this paper with bearing PDA because another research has investigated bearing PDA using the SDAE and AP.
- (2) To demonstrate that the model proposed is better than other combined models, including EEMD-SVD-FCM/ k -means/ k -medoids, SAE-CFS, PCA,

RMS, kurtosis, SHE, AE, and PE, a detailed comparative analysis is provided.

The rest of this paper is organized as follows. The basic theories of the SDAE and CFS are given in Section 2. Section 3 describes the procedures of the method proposed, the experiment comparison analysis is given in Section 4, and Section 5 concludes the paper.

2. Basic Theories of the SDAE and CFS

2.1. Basic Theory of the SDAE. In this section, the basic unit in the SDAE, DAE, which is based on AE, is used. Then, the basic structure of the SDAE is then described, which is stacked from DAE.

2.1.1. Autoencoder (AE). The main idea of AE [25] is to build constant functions between input X and output Z and to achieve dimensionality reduction and preserve data feature information. It can be divided into two parts: encoder and decoder.

(1) *Encoder.* Figure 1 illustrates how encoding is the process of implementing an input dataset X mapped into a low-dimensional space by an activation function. The encoder performs a mapping conversion from the input vector $X = \{x_1, x_2, \dots, x_n\}$ to the output representation $Y = \{y_1, y_2, \dots, y_n\}$ by using an active function. n is the total number of samples. The calculation expression is as follows:

$$Y = f_\delta(X) = s(Wx + b), \quad (1)$$

where f_δ and s are the sigmoid activation functions, where $f_\delta(X) = 1/1 + e^{-X}$, and $f_\delta(X)$ is the abbreviation of $s(Wx + b)$ for input X . W is the weight vector in the neural network between the former and latter layers, and b denotes the bias item.

(2) *Decoder.* Decoding is the procedure of mapping Y from a high-dimensional space into a high low-dimensional space $Z = \{z_1, z_2, \dots, z_n\}$ and reconstructing the input sample X into Z . The calculation is as follows:

$$Z = g_\delta(Y) = s(Wy + b), \quad (2)$$

where g_δ and s are the sigmoid activation functions, where $g_\delta(Y) = 1/1 + e^{-Y}$. $g_\delta(Y)$ is the abbreviation of $s(Wy + b)$. Therefore, the reconstructed error (lost function) between Z and X is defined as

$$J(X, Z) = \frac{1}{n} \|Z - X\|^2. \quad (3)$$

The backpropagation algorithm is used to adjust the weight vector W and bias item b and train the autoencoder network to reduce the reconstructed error. Hence, the restrictive error $J(X, Z)$ is converged and minimized until it meets the termination condition, i.e., it exceeds the maximum iteration.

FIGURE 1: Structure of AE.

2.1.2. Denoising Autoencoder (DAE). The data collected in actual engineering often contain noise, and hence the characteristics obtained by the autoencoder will cause errors owing to the presence of noise. The denoising autoencoder (DAE) [37] solves this problem by destroying the noise-containing data into zero according to the denoising probability P and reconstructing the destroyed input X_1 into output Z by using the encoder and decoder in AE. The basic structure of DAE is shown in Figure 2.

The black round node in Figure 2 is the damaged data point in X_1 , and P denotes the denoising probability. Some parts of the input data X are therefore set as zero and then X is changed to dataset X_1 . f and g denote the sigmoid activation functions in formulas (1) and (2). The following calculation steps are the same as for AE when an encoder and a decoder are used to reconstruct the output Z into the original input data X .

2.1.3. Stacked Denoising Autoencoder (SDAE). The SDAE contains three layers: (1) an input layer; (2) several hidden layers; and (3) an output layer. It uses several hidden layers, which are stacked from several DAE units, to extract the useful information. Therefore, the output Z from the previous DAE hidden layer is regarded as the input of the next DAE hidden layer. The connection weight matrix W and bias vector b are then iteratively updated during the pretraining period. After the training has been completed, the entire network is fine-tuned, and after the above steps, the SDAE is formed. The basic structure of the SDAE is shown in Figure 3. Here, N is the number of DAE hidden layers in the SDAE.

2.2. Basic Theory of CFS. The CFS clustering algorithm is mainly based on the characterization of the cluster center.

- (a) The cluster center point itself has a high density and is surrounded by data points whose densities are not more than those of its own neighbors.
- (b) The distance between the cluster center point and other data points in another cluster is better.

The detailed computational procedures of CFS clustering are given in [32], and its calculation process is as follows:

FIGURE 2: Basic structure of DAE.

FIGURE 3: Basic structure of the SDAE.

- (1) For a given dataset $X = \{x_1, x_2, \dots, x_n\}$
The distance between any two points x_i and x_j can be calculated by

$$d_{ij} = \text{dist}(x_i, x_j). \quad (4)$$

- (2) The local density ρ_i for a data point x_i is calculated by

$$\rho_i = \sum_{j=1}^N e^{-(d_{ij}/d_c)^2}, \quad (5)$$

where ρ_i denotes the total number of distances that are less than the cutoff distance d_c between the data point x_i and other points and N is the total number of samples.

- (3) Calculate the distance δ_i , assuming that the $\{q_i\}_{i=1}^N$ is the descending order of $\{\rho_i\}_{i=1}^N$, where $\{\rho_i\}_{i=1}^N$ meets the condition: $\rho_{q1} \geq \rho_{q2} \geq \dots \geq \rho_{qN}$, and hence δ_i can be defined as

$$\delta_{q_i} = \begin{cases} \min_{j < i} \{d_{q_i q_j}\}, & i \geq 2, \\ \max_{j \geq 2} \{\delta_{q_j}\}, & i = 1, \end{cases} \quad (6)$$

where δ_i indicates the greatest distance between any one point and point x_i when x_i has the greatest local density. Otherwise, δ_i is the smallest distance between any one point and point x_i when x_i has the smallest local density.

- (4) Compute γ according to the following equation:

$$\gamma_i = \rho_i * \delta_i. \quad (7)$$

- (5) Use γ to determine the potential clustering center point when γ is in descending order. The authors suggest that the greater the γ_i , the greater the possibility of point x being a cluster center point. The stepped data points with the greater γ values are sequentially selected as the cluster center points, that is, the γ value shows a significant jump when the cluster center point transitions to the noncluster center point. These points with skipping are thus selected as cluster center points, according to this characteristic. Finally, the number of distances between each data point and the cluster center is less than the number of cutoff distances required to achieve classification.

3. Procedure of the Presented Model

The procedure of the method comprises four steps: (1) data preprocessing; (2) preliminary degradation trend generation; (3) degradation trend dimension reduction; and (4) bearing degradation assessment and confrontation analysis. The details of these steps are as follows.

- (1) Data preprocessing: to extract the useful bearing degradation trend and preprocess the data more easily, the absolute amplitude values of all original vibration signals are regarded as the input of the SAE and SDAE for training, after standardizing into $[0, 1]$.
- (2) Preliminary degradation trend generation: for the SAE and SDAE, there are nine hidden layers from which the useful initial degradation trend of the bearing can be extracted. The dimension of the extracted bearing degradation trend by each hidden layer, except the last, is not 2. To demonstrate that the SDAE is more robust and stable than the SAE, PCA is used to reduce the dimension of the extracted degradation for the first eight hidden layers. After EEMD is decomposed using the original vibration signal, SVD is used to calculate the SVs to identify the degradation trend. In addition, to show that the model presented is superior to others, RMS, kurtosis, SHE, AE, PE, and PCA are also used for extraction of the degradation feature.
- (3) Degradation trend dimension reduction: for data visualization, the number of neural nodes at the last hidden layer in the SAE and SDAE is set directly to 2. For EEMD, the two IMF components are selected according to the two largest correlation coefficients, which are calculated from each IMF and the original vibration signal. The greater the correlation coefficient value, the greater the amount of useful vibration information it contains. Then, SVD is used to compute the SVs for dimension reduction. For PCA, the first two principal components (PCs) are selected as the extracted features.
- (4) Bearing degradation severity assessment and confrontation analysis:
- (a) Bearing degradation severity assessment: the two-dimensional degradation features extracted using the SAE and SDAE are selected as the input of CFS to find the available cluster center point. The health indicator, or confidence value (CV), is then used to build a PDA model. The details of the CV calculation are as follows:

$$CV = \exp\left(\frac{-DI}{c}\right), \quad (8)$$

where DI is the Euclidean distance, which is often used to compute the distance between each point A (x_1, y_1) and its cluster center point B (x_2, y_2). DI is then calculated by

$$DI = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (9)$$

The main purpose of DI is to transfer all of the CVs to $[0, 1]$ by using one cluster center point. c denotes the scale factor. The CV is close to 1 when the "Normal" clustering center point is used, which indicates that

the sample belongs to “Normal” [21]. For the ease of the comparison analysis, all CVs are normalized to $[0, 1]$, and then the CV degradation trend curve is smoothed by using a smoothing function through the four-time window.

- (b) Confrontation analysis: the method proposed is demonstrated to be superior to other models such as the SAE-CFS, EEMD-SVD-FCM/k-means/k-medoids in [21], PCA, RMS, kurtosis, SHE, AE, and PE, through the detailed analysis given in the following section. The SVs in SVD obtained from EEMD and SVD are regarded as the input of FCM, k-means, and k-medoids to find the available cluster center points. Then, the CVs are calculated according to equation (8).

Through the above steps, all CVs are obtained from the proposed method, SAE-CFS, EEMD-SVD-FCM, k-means, and k-medoids. The steps of the method are shown in Figure 4.

4. PDA Building and Comparison Analysis

In this section, the experimental data and the data collected platform are first introduced, and in the following step, the SAE and SDAE are used to extract the degradation; hence, the last step is the smoothness comparison. The extracted features are then used to find the clustering center points with CFS. Finally, a comparison analysis is given.

4.1. Original Vibration Signals. The experimental data acquisition platform is shown in Figure 5. The operating conditions of the bearing depend on the instantaneous measurement of the radial force exerted on the bearing, the rotation speed of the shaft that manipulates the bearing, and the magnitude of the torque exerted on the bearing. The bearing degradation feature is based on two sensors, vibration and temperature. The vibration sensor consists of two micro accelerometers that are perpendicular to each other. The first is in the vertical position, and the other is in the horizontal position. In addition, the vibration sensor is fixed on the outer ring of the bearing. The data sampling frequency is 25.6kHz. The temperature sensor is not described in detail here. The vibration data in the horizontal position are used in the experiment.

For more information on the experimental platform, refer to the literature [38].

The experimental dataset is an accelerated degradation test of the bearing under various operating conditions to obtain the measured data in the bearing life cycle for fault detection and prediction of the bearing’s remaining life [38]. The three load conditions are 4000 N, 4200 N, and 5000 N. The corresponding speeds are approximately 1800 rpm, 1650 rpm, and 1500 rpm. The experimental device samples the data every 0.1 seconds. The data length of each sample is 2560. The details of the experimental data for bearing 1 are given in Table 1. The original vibration time-wave for bearings 11–15 is shown in Figure 6. Bearings 11–14 have 2 or 3 degradation statuses. For bearings 11 and 13, the

amplitude of the vibration signal is gradually increased. The marked red rectangle denotes the Severe status in Figure 6. Bearing 12 clearly shows a jump and some noise under the Normal condition. Hence, it has only two degradation statuses (Normal and Severe). Compared with bearing 12, bearing 14 has two obvious jumping points; hence, bearing 14 contains three statuses. The vibration signal in the blue rectangle denotes the Slight condition, and the Severe condition is shown in the red rectangle. It is difficult to identify the status at first glance without extensive manual experience. The degradation status of bearing 15 is even more problematic as it cannot be seen clearly when extracted manually because there is massive noise. We also take one sample to show the frequency result. The FFT results are shown in Figure 7. The signal is more prominent at the frequency of the range $[350, 450]$ except bearing 15, which are not the approximate integer times the working frequency (25.6 Hz). Therefore, this result indicates that the frequency-domain signal contains no more useful degradation trend information. Moreover, there is no useful information from frequency domain for bearing 15 because the massive noise result makes the degradation trend in the frequency domain not good. Therefore, we use the absolute amplitude from original signal to extract the degradation of bearing and reduce the manual experience.

In the following section, the SDAE and SAE are used to extract the preliminary degradation trend and CFS is then used to find some available clustering center points, which are in turn used to determine the degradation trends and construct the health indicator CV for assessing the bearing’s PDA.

4.2. Degradation Trend Extracted by the SAE and SDAE.

In this part, the SAE and SDAE are used to extract the preliminary degradation trend through several hidden layers. Before preliminary degradation extraction, the absolute values of the vibration amplitudes for all bearings are considered as the input of the SAE and SDAE to reduce the data dimension for convenient data visualization. The hidden layers in the SAE and SDAE have a triangular structure, that is, the number of hidden layer nodes is half of the former adjacent hidden layer. In [29], the authors use a triangular hidden layer structure to extract the potential fault feature and confirm the validity of the proposed model. Note that the triangular structure often results in the number of latter hidden layer nodes being half of the number of former adjacent hidden layer nodes.

Therefore, the number of the first nine hidden layers’ neural nodes is set at $[1280, 640, 320, 160, 80, 40, 20, 10, \text{and } 2]$ in the SAE and SDAE. The maximum iteration number for each layer is 50. With regard to the learning rates in the SAE and SDAE, the lower the learning rate is, the more slowly the update speed changes for the cost function. A small value will result in a local minimum [37, 39]; hence, we use 0.1 in this study. In SDAE, if the value of the denoising probability P is too great, more information will be lost from the original data. The authors suggest that the parameter P is typically fixed below 0.5 [37, 39]. Therefore, in this study, a low value of 0.1 is used for denoising probability P .

FIGURE 4: Flowchart of the method presented.

FIGURE 5: The experimental data acquisition platform.

To demonstrate that the robustness and feature extraction performance of the SDAE are superior to those of the SAE, the degradation trends of all bearings from the first

eight hidden layers are used for comparison. For easy visualization of the data, PCA is used to reduce the dimensions of the extracted degradation vectors to two for the first eight

TABLE 1: Detailed information on the experimental data.

Working conditions	Bearings				
	Bearing 11	Bearing 12	Bearing 13	Bearing 14	Bearing 15
Load (N)	4200	4200	4200	4200	4200
Speed (rpm)	1800	1800	1800	1800	1800
Length	2560	2560	2560	2560	2560
Sample numbers	2803	873	1802	1139	2302

FIGURE 6: Time-wave figure of the original vibration for bearings 11–15. (a) Bearing 11. (b) Bearing 12. (c) Bearing 13. (d) Bearing 14. (e) Bearing 15.

FIGURE 7: Continued.

FIGURE 7: The frequency-domain signal through FFT for bearings 11–bearing 15. (a) Bearing 11. (b) Bearing 12. (c) Bearing 13. (d) Bearing 14. (e) Bearing 15.

hidden layers. The results of the first two principal components (PC1-PC2) are shown in Figures 8 and 9 (due to limited space, only bearing 11 is given as an example in this study).

Figures 8 and 9 show that the degradation trend can be extracted successfully from the original vibration signal after the original vibration data have been trained by the SAE and SDAE. All curves show monotonic growth or reduction at each hidden layer, and compared with the SAE, they are more stable and less noisy in the SDAE at each layer, and there is little or no noise before the first 1500 points. In Figure 9, they look like a straight line without fluctuation when the SDAE is used, but there is a small amount of curve fluctuation at the seventh hidden layer in the SAE. Take the 391st data point as an example. In Figure 8, as the number of hidden layers increases, the 391st data contains obvious noise, even though the 391st data point in the trend is extracted from the last hidden layer. But Figure 9 shows that with the increase of number of hidden layers, the noise contained in the 391st data point gradually weakens, for example, starting from the sixth hidden layer, the noise of the 391st data point is obviously weakened. SDAE sets part of the input data to 0 and reconstructs it through the denoising rate; this can reduce the noise, and hence the denoising effect of SDAE is better than SAE.

The SAE and SDAE are also used to extract the degradation trend for different bearings (11–15), and the corresponding results through the ninth hidden layer are shown in Figures 10 and 11, respectively. Similar to bearing 11, as the number of hidden layers increases, the noise of the HI curve extracted through the final hidden layer by SDAE is not obvious, while the noise of the HI curve extracted by SAE is still obvious, such as bearing 14 in Figures 10 and 11.

In Figures 10 and 11, all curves show a monotonous increase and decrease, except for bearing 15. Compared with the SAE, there is an obvious monotonous increase or decrease curve for bearings 11–14 when the SDAE is used in Figure 11. Starting from around the 1000th point, the curve shows a stable status for bearing 15 in Figure 11. Before the 1000th data point, there is an evident rising and falling trend when SDAE is used. There is no conspicuous trend at the ninth hidden layer when the SAE is used in Figure 8, as it is submerged in massive noise. The SDAE destroys and reconstructs the input to improve its robustness, confirming

that the SDAE has a denoising ability. The trend for bearings 12 and 14 looks like a straight line and is not rising and falling because the vibration amplitude is very smooth at each stage. This is in accordance with Figure 6. In Figure 10, particularly for bearing 14, there is some noise in the SAE under the Normal condition, while there is a stable line in Figure 11 when the SDAE is used. The line pattern for bearing 13 is similar. These results demonstrate that the robustness and stability of the SDAE are superior to those of the SAE. In addition, the SAE and SDAE can extract the preliminary degradation characteristics well without extensive manual experience and tagged data labels.

However, at first glance, bearings 12 and 14 have only two states in the manual process. However, there are two degradation statuses: Normal and Severe, and perhaps three: Normal, Slight, and Severe, for bearing 14. Identifying how many degradation statuses there are for bearing 15 is difficult with the naked eye, and determining the number of degradation statuses a bearing should have using manual experience and the naked eye can in general be erroneous. Therefore, CFS is used to find the cluster center point under different degradation statuses and can be an option an engineer can use to determine the degradation status of a bearing.

4.3. Constructing a Health Indicator CV and Judging Degradation by Using CFS. Before CFS calculation, some parameters need to be preset, such as the cutoff d_c . In [32], the authors advise that the average number of a neighbor data point for other points should not exceed d_c . In general, the average neighbor data point accounts for about 1–2% of the total number of data points. Hence, the average number of a neighbor data point is often set at about 1–2% of the total sample size. If the local density ρ_i of point x_i is too great, it will result in low discrimination; if d_c is too small, the same cluster class will be split into multiple parts [32]. The results of local density ρ_i and the distance δ_i when the SAE/SDAE-CFS is used for bearings 11–15 are shown in Figure 12. The corresponding two-dimensional clustering results when the SAE/SDAE-CFS is used are also shown in Figures 12 and 13.

- (1) As shown in Figures 14(a) and 14(c) and Figures 15(a) and 15(c), three points (A, B, and C) are filtered by using the cutoff distance d_c and the local density ρ_i . The greater the value of the cutoff distance

FIGURE 8: Extracted degradation through every hidden layer for bearing 11 by using the SAE.

d_c and the local density p_i , the more likely it is to become the cluster center point. Hence, these three points with obvious jumping are selected as the clustering center points, and there are only two points with obvious jumping.

(2) Compared with the SAE, the selected clustering center points with jumping for bearings 13 and 15 in Figures 16(c) and 16(e) are more obvious when the SDAE is used. For example, in Figure 14(c), bearing 13 contains three clustering center points, but the

FIGURE 9: Extracted degradation through every hidden layer for bearing 11 by using the SDAE.

third clustering center candidate point (C) is close to the fourth (D). The third clustering center candidate point (C) is separated and filtered far away from the fourth in Figure 16(c). It is easier to choose and

determine the clustering center points by using the SDAE. This is also shown in Figure 16(e), where four selected clustering center points are very clearly filtered far away from other points that are close to

FIGURE 10: The extracted two-dimensional degradation features for bearings 11–15 through the final hidden layer by using the SAE.

FIGURE 11: Extracted two-dimensional degradation features for bearings 11–15 through the final hidden layer by using the SDAE.

the horizontal axis. However, in Figure 14(c), many data points are scattered because there is massive noise in Figure 8 when the SAE is used. Thus, the feature extraction of the SDAE is superior to that of the SAE.

(3) Figures 12 and 13 show that compared with the SAE-CFS, the SDAE-CFS performs better at clustering. In Figures 13(b) and 13(d), all samples are separated well by using the SDAE, as few points are scattered around its clustering center point in Figure 12(d),

FIGURE 12: Two-dimensional clustering result by using the SAE and CFS for bearings 11–15. A: Normal1; B: Normal2; C: Slight; D: Severe. (a) CFS-bearing 11. (b) CFS-bearing 12. (c) CFS-bearing 13. (d) CFS-bearing 14. (e) CFS-bearing 15.

FIGURE 13: Continued.

FIGURE 13: Two-dimensional clustering result by using the SDAE and CFS for bearings 11–15. A: Normal1; B: Normal2; C: Slight; D: Severe. (a) CFS-bearing 11. (b) CFS-bearing 12. (c) CFS-bearing 13. (d) CFS-bearing 14. (e) CFS-bearing 15.

FIGURE 14: Result of local density ρ_i and the distance δ_i by using the SAE-CFS for bearings 11–15. (a) SAE-CFS-bearing 11. (b) SAE-CFS-bearing 12. (c) SAE-CFS-bearing 13. (d) SAE-CFS-bearing 14. (e) SAE-CFS-bearing 15.

but there is only one point identifiable under the Normal condition in Figure 13(d) using the naked eye. This is consistent with the situation in Figure 10 because when the SAE is used, the extracted degradation curve contains some noise under the Normal condition throughout the entire life.

- (4) It should be noted that bearing 15 clearly contains four clustering center points when the SDAE is used in Figures 12(e) and 13(e). Thus, bearing 15 has four degradation statuses. Identifying the degradation trend of bearing 15 from Figure 5 is difficult using the naked eye or manual experience, but CFS can enable

us to determine the number of degradation statuses for bearing 15 without prior knowledge.

- (5) In Figures 12(e) and 13(e), four symbols, A to D, denote the different degradation statuses for bearing 15: A—Normal1, B—Normal2, C—Slight, and D—Severe. We use RMS as a reference to determine the degradation trend and demonstrate that CFS can find the available and suitable cluster center point number that is the same as RMS. The RMS obtained from bearing 15 is shown in Figure 17, and it is evident that starting from the 66th point, the RMS curve increased sharply until the 141st point. After

FIGURE 15: CVs for bearings 11–15 by using the SAE-CFS under various conditions such as “Normal,” “Slight,” and “Severe.”

FIGURE 16: Continued.

FIGURE 16: Result of local density ρ_i and the distance δ_i by using the SDAE-CFS for bearings 11–15. (a) SDAE-CFS-bearing 11. (b) SDAE-CFS-bearing 12. (c) SDAE-CFS-bearing 13. (d) SDAE-CFS-bearing 14. (e) SDAE-CFS-bearing 15.

FIGURE 17: RMS for bearing 15 under various conditions such as “Normal,” “Slight,” and “Severe.”

the 141st point, there is another increase until the 359th point. The curve decreased sharply from the 359th point to the 1104th point, while it became stable from the 1104th point to the endpoint. Hence, there are four degradation statuses for bearing 15: A—Normal1 (from 1 to 66), B—Normal2 (from 66 to 141), C—Slight, and D—Severe. Finally, the CVs calculated by using different clustering center points are shown in Figures 15 and 18. These results demonstrate that CFS can give us an option to determine the number of degradation statuses for bearing 15 without prior knowledge.

- (6) In Figure 18(e), the degradation trend is similar to that of RMS, and several turning points, such as the 66th point, the 359th point, and the 1104th point, can reflect the degradation trend of bearing 15 well by using our presented model. Each subfigure in Figure 18(e) can be regarded as a reference to determine the degradation trend together with other curves. The first three subfigures in Figure 18 show that starting from the 359th point, these three curves clearly become stable, so it is easier to judge the Severe status for bearing 15 after the 1104th

point. Furthermore, from the 1104th point, all CV curves are more stable than the RMS curve. Before the 359th point, most CVs are close to 1 and clustering center points (A and B) are used under the Normal condition. After the 359th point, there is an obvious increase when clustering center point C is used, so some points under the Slight condition are close to 1. These results indicate that our proposed model can reflect the degradation trend well and more precisely than RMS without prior knowledge.

- (7) In Figure 18, starting from the 1532nd point until the 2750th point, bearing 11 changes status from Normal to Slight. After the 2750th point until the endpoint, the status for bearing 11 changes from Normal to Severe. These state turning points for bearings 13 and 14 are [1275, 1720] and [1088, 1093], respectively. For bearing 12, they are [827, 830].

The above results show that the SDAE is more robust than the SAE and that CFS has a good ability to choose the clustering center point well under different conditions without prior knowledge. In the following sections, other

FIGURE 18: CVs for bearings 11–15 resulting from using the SDAE-CFS under various conditions such as “Normal,” “Slight,” and “Severe.”

models presented in [21], such as the time-frequency indicators RMS and kurtosis and PCA, are considered and compared with our proposed model.

4.4. The Presented Method Compared with RMS, Kurtosis, and PCA. In this section, time-frequency indicators, including RMS and kurtosis, and PCA are compared with the SDAE-CFS.

4.4.1. RMS and Kurtosis. The results of RMS and kurtosis calculated from bearings 11–15 are shown in Figures 19 and 20.

- (1) In Figure 20, the degradation trend is obviously obscured by noise, which will easily result in the degradation status being misjudged, particularly for bearings 11 and 12. Unlike kurtosis and RMS, the SDAE-CFS, shown in Figure 18, can reflect the degradation well.
- (2) Some RMS curves have small fluctuations and few noises. Thus, all of the stable and smooth curves of RMS are inferior to those of the SDAE-CFS. In

Figure 18, most of the CV curves under different conditions look like straight lines at first glance, especially under the Normal condition; these CV curves can be used to identify the degradation status more easily than those of RMS when the degradation status has changed. For example, the status change from Normal to Severe in bearing 12 is very clear, starting from the 827th point; there is an obvious jump in Figure 18, while the RMS curves in Figure 19 show a gradual change, not a jump.

- (3) For bearing 15, the degradation trend must be judged manually, while CFS can find the available number of the degradation status automatically. Therefore, these results demonstrate that the proposed model is better than RMS and kurtosis. In addition, CFS can provide the available clustering center point to assess the number of the degradation statuses.

4.4.2. PCA. In this section, PCA is used to extract the bearings' degradation trend. The first two components (PC1-PC2) are used to show the results of the

FIGURE 19: CVs for bearings 11–15 by using RMS under various conditions such as “Normal,” “Slight,” and “Severe.” (a) RMS-bearing 11. (b) RMS-bearing 12. (c) RMS-bearing 13. (d) RMS-bearing 14. (e) RMS-bearing 15.

FIGURE 20: CVs for bearings 11–15 by using kurtosis under various conditions such as “Normal,” “Slight,” and “Severe.” (a) Kurtosis-bearing 11. (b) Kurtosis-bearing 12. (c) Kurtosis-bearing 13. (d) Kurtosis-bearing 14. (e) Kurtosis-bearing 15.

degradation. In Figure 21, there is no obvious degradation for bearing 15 because much noise masks the trend of the degradation. The SDAE used the DAE to destroy the data to zero and then reconstruct it. Therefore, the SDAE is more robust and stable than PCA and CFS in finding suitable clustering center points to determine the degradation status. In addition, most of the CV curves obtained from PCA are not as stable and smooth as those of the SDAE-CFS.

4.5. The Presented Method Compared with EEMD-SVD-k-Means/k-Medoids/FCM. In this section, the proposed model is compared with other models in [21], such as EEMD-SVD-k-medoids/k-means/FCM. Some parameters should be set before calculating EEMD and k-medoids/k-means/FCM.

- (1) EEMD: two parameters must be selected before the EEMD calculation; m is the ensemble number and

FIGURE 21: CVs for bearings 11–15 by using PCA under various conditions such as “Normal,” “Slight,” and “Severe.”

the amplitude of the added white noise $n_i(t)$ [17]. The added white noise is calculated from the standard deviation (SD) of the original vibration signal. In [17], the authors advise that the white noise should be set at 20% of the standard deviation from the original data [17]. For parameter m , a few hundred numbers will result in greater accuracy. Hence, the parameter $m=100$ is selected in this study.

- (2) k-Medoids/k-means/FCM: in FCM, the iteration convergence termination tolerance $\varepsilon = 1e-6$. Euclidean distance is used to calculate the similarity between any two samples in FCM/k-means/k-medoids. The parameter c is the number of clustering center points. For bearings 11, 13, and 15, $c=3$, for bearing 12, $c=2$, and for bearing 15, $c=4$.

First, EEMD is used to decompose the original signals to IMFs. As space is limited, here we only use bearings 11 and 12 as examples. IMFs obtained from EEMD are shown in Figure 22. The amplitude of the first two IMFS is greater than that of the others because all IMFs are decomposed in order of frequency from high to low. In addition, the correlation coefficient is used to calculate the degree of relevance between each IMF and the original signal. The values of the corresponding correlation coefficients are shown in Figure 23. The two highest values are for IMF1

and IMF2. This indicates that these first two IMFs contain useful information about the original signal. Therefore, IMF1 and IMF2 are used to calculate the SVs (SV1 and SV2) through SVD. The results for SV1 and SV2 are shown in Figure 24. In this figure, bearing 11 has 3 statuses while bearing 12 has 2. The black rectangle denotes the Severe status. Therefore, these two extracted feature vectors [SV1, SV2] are regarded as the input of k-medoids/k-means/FCM for finding the available clustering center points. The two-dimensional clustering figure of the bearings when EEMD-SVD-k-medoids/k-means/FCM is used for bearings 11–15 is shown in Figure 25. The corresponding CVs for bearings 11–15 obtained from EEMD-SVD-k-medoids/k-means/FCM under various conditions are shown in Figure 26.

- (1) Figures 12 and 13 show that the SAE/SDAE-CFS performs better at clustering when compared with EEMD-SVD-k-medoids/k-means/FCM. In Figure 13(b), all samples are separated well when CFS is used, for example, bearing 12. At first glance, the points simply look like a point overlapping with its clustering center point under the Normal condition. However, some points are scattered around their clustering center points in Figures 25(b), 25(g), and 25(l). The same holds in Figures 25(d), 25(i), and 25(n) for bearing 14.

FIGURE 22: Results of IMFs when using EEMD for bearings (a) 11 and (b) 12.

- (2) For bearing 15, the number of clustering center points is set at 4 according to the CFS clustering result referred to above. Thus, CFS can provide us with an available option to determine the number of degradation statuses for bearing 15 without prior knowledge, but k-medoids/k-means/FCM cannot do this.
- (3) In Figures 26(b), 26(g), and 26(l), there is some noise under the Severe status for bearing 12. This noise may be mistakenly assessed when judging the state of degradation. In Figure 18(b), there are only two straight lines to divide the trend statuses.
- (4) In Figure 18(a), the CV line shows an obvious increase between the Normal and Slight statuses for bearing 11, and it is easy to identify these statuses.

However, in Figure 26, they are similar under the Normal and Slight conditions when EEMD-SVD-k-medoids/k-means/FCM is used.

- (5) All CV lines in Figure 18 are more stable than those in Figure 26, particularly under the Normal status. In addition, there is some noise in Figure 26 when different models are used.

As mentioned, the proposed method (SDAE-CFS) exceeds the good results of EEMD-SVD-k-medoids/k-means/FCM, RMS, kurtosis, and PCA.

4.6. The Presented Method Compared with SHE, AE, and PE. In this section, typical health indicators such as SHE, AE, and PE are used to assess the bearing degradation trend.

FIGURE 23: Correlation coefficients between each IMF and the original vibration signal. (a) Bearing 11. (b) Bearing 12.

FIGURE 24: SV features extracted for bearings 11 and 12 from the first two IMFs by using SVD.

FIGURE 25: Continued.

FIGURE 25: Two-dimensional clustering figure of bearings using EEMD-SVD-k-medoids/k-means/FCM. (a) Bearing 11. (b) Bearing 12. (c) Bearing 13. (d) Bearing 14. (e) Bearing 15. A: Normal1; B: Normal2; C: Slight; D: Severe.

Some parameters should be preconfigured before AE and PE calculation.

AE: the two parameters that should be set before calculation are embedded dimension and tolerance. Increasing the embedded dimension will cause the AE to include more useful information in the calculation, but it will also increase the computational cost. The authors suggest that the embedded dimension is often fixed at 2 [13]. Tolerance is often set at $(0.1 \sim 0.25) * SD$, where SD is the standard deviation from the original data [13].

PE: in references [15, 40], the authors demonstrate that the embedded dimension should be in the range of 3~7. If the embedded dimension is more than 8, the corresponding calculation efficiency is poor because the reconstruction of phase space will homogenize the vibration signals. If the time delay is more than 5 and the embedded dimension is less than 4, the calculation cannot accurately detect small changes in the vibration signal. In addition, the experiment result demonstrates and the authors suggest that fixing the embedded dimension at 6 and the time delay at 3 could provide a suitable PE

FIGURE 26: Continued.

FIGURE 26: CVs for bearings 11–15 obtained from EEMD-SVD-k-medoids/k-Mean/FCM under various conditions such as “Normal,” “Slight,” and “Severe.”

calculation [16]. Therefore, we also use an embedded dimension fixed at 6 and a time delay fixed at 3. The results of the degradation curve when SHE, AE, and PE are used are shown in Figures 25–27.

- (1) In Figures 27 and 28, there is no clear degradation for bearing 15, so it is difficult to identify the degradation status. Although the SHE values for bearing 15 are close to a gentle straight line, there are some jump points

where it is easier to misjudge different degradation statuses, and there is a lot of noise in Figure 28. Compared with SHE and PE, when AE is used, the curve line in Figure 29 has a blurred degradation trend for bearing 15, and AE cannot provide a suggestion for the number of degradation statuses with which to identify the degradation trend, while CFS can. In Figure 18, the four different CV curves can be used to

FIGURE 27: CVs for bearings 11–15 by using SHE. (a) Bearing 11. (b) Bearing 12. (c) Bearing 13. (d) Bearing 14. (e) Bearing 15.

FIGURE 28: CVs for bearings 11–15 by using PE. (a) Bearing 11. (b) Bearing 12. (c) Bearing 13. (d) Bearing 14. (e) Bearing 15.

identify the differences and can be combined to divide the state by using different clustering center points.

- (2) In Figure 27, not all curves show a monotonous increasing and decreasing trend, such as the SHE curves for bearings 11, 13, and 15. The PE values for bearing 11 are similar to those of SHE, while all CV curves in Figure 18 are monotonous increases and decreases. The noise in Figures 28 and 29 is obvious, for example, in Figure 29(b).
- (3) For PE, starting from the 200th point, the curves for bearings 12 and 14 are close to stable in Figure 28,

but the degradation trend for bearing 12 is from Normal to Severe around the 820th point, not around the 200th point. In addition, bearing 14 has three degradation statuses, but after the 200th point, the curve is stable until the end, as in Figure 28. The status of bearing 15 is similar to that of bearings 12 and 14.

To further demonstrate that the denoising effect of SDAE is good, we also use monotonicity (*Mon*) index. *Mon* uses the difference of any two adjacent HI points to calculate and assess the monotonicity of extracted HI [41]. If the difference

FIGURE 29: CVs for bearings 11–15 by using AE. (a) Bearing 11. (b) Bearing 12. (c) Bearing 13. (d) Bearing 14. (e) Bearing 15.

TABLE 2: *Mon* result of different models by using bearing 12 dataset.

Model	<i>Mon</i>
SDAE-CFS	0.9310
SAE-CFS	0.0288
EEMD-SVD-k-medoids	0.0161
EEMD-SVD-k-means	0.0092
EEMD-SVD-FCM	0.0230
RMS	0.0253
Kurtosis	0.0207
AE	0.0196
PE	0.0081
SHE	0.0334

value by using two adjacent HI data points is greater than 0, then the HI curve rises monotonically and vice versa. If the curve rises and falls monotonically within short spans, then the HI curve has significant noise and oscillations. The calculation of *Mon* is as follows:

$$Mon = \left| \frac{DFA}{t-1} - \frac{DFB}{t-1} \right| = \left| \frac{DFA - DFB}{t-1} \right|, \quad (10)$$

DFA = number of $dF > 0$,

DFB = number of $dF < 0$,

$$dF = \frac{HI_t - HI_{t-1}}{\Delta t}, \quad 2 \leq t \leq N,$$

where dF is the difference between any two adjacent HI points. The closer *Mon* is to 1, the better the performance is [41–53]. We take bearing 12 as an example; the *Mon* result of different models is shown in Table 2.

Table 2 shows that the *Mon* of SDAE-CFS is higher than that of other models. The SDAE sets the input data of each

hidden layer according to the denoising rate and then reconstructs the input data. Therefore, the SDAE can denoise the extracted HI curve well.

5. Conclusions

The original vibrations over the entire life of the bearings under different conditions were used as the input for the SAE and SDAE to extract the degradation trend and directly reduce the dimension of the extracted feature to two without PCA. The results demonstrate that the SDAE was more robust and had better feature extraction performance than the SAE. CFS was then implemented to find the available clustering center point, which was used to assess the health status through the CV index without data labeling or prior knowledge. To verify the performance of the proposed method, it was compared with other combination models, such as EEMD-SVD-k-medoids/k-means/FCM, RMS, kurtosis, SHE, AE, PE, and PCA. The experimental results confirmed that the SDAE-CFS was more robust and stable than the other models. Finally, we also use the *Mon* index to evaluate the denoising effect of different models. The larger the value of *Mon*, the smaller the noise of the extracted HI curve, and vice versa. The model proposed in the article can significantly increase the value of *Mon* by a quantitative level, as shown in Table 2. SDAE-CFS can improve the value of *Mon* from two decimal places to one decimal point.

Data Availability

Previously reported bearing data were used to support this study and are available at <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>.

Conflicts of Interest

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence their work. There is no professional or other personal interest of any nature or kind in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, the manuscript.

Acknowledgments

This research was partly supported by the Green Intelligent Inland Ship Innovation Programme and the National Natural Science Foundation of China (grant no. 51909199).

References

- [1] W. Huang, J. Cheng, and Y. Yang, "Rolling bearing fault diagnosis and performance degradation assessment under variable operation conditions based on nuisance attribute projection," *Mechanical Systems and Signal Processing*, vol. 114, no. 1, pp. 165–188, 2019.
- [2] A. Heng, A. C. C. Tan, J. Mathew, N. Montgomery, D. Banjevic, and A. K. S. Jardine, "Intelligent condition-based prediction of machinery reliability," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1600–1614, 2009.
- [3] I. El-Thalji and E. Jantunen, "A summary of fault modelling and predictive health monitoring of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 60, pp. 252–272, 2015.
- [4] A. Rai and S. H. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, vol. 96, pp. 289–306, 2016.
- [5] T. Williams, X. Ribadeneira, S. Billington, and T. Kurfess, "Rolling element bearing diagnostics in run-to-failure lifetime testing," *Mechanical Systems and Signal Processing*, vol. 15, no. 5, pp. 979–993, 2001.
- [6] P. W. Tse and D. Wang, "State space formulation of nonlinear vibration responses collected from a dynamic rotor-bearing system: an extension of bearing diagnostics to bearing prognostics," *Sensors*, vol. 17, no. 2, pp. 1–16, 2017.
- [7] Z. Shen, X. F. Chen, Z. J. He et al., "Remaining life predictions of rolling bearing based on relative features and multivariable support vector machine," *Journal of Mechanical Engineering*, vol. 49, no. 2, pp. 183–189, 2013.
- [8] Y. Lei, N. Li, and J. Lin, "A new method based on stochastic process models for machine remaining useful life prediction," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 12, pp. 2671–2684, 2016.
- [9] B. Y. Kosasih, W. Caesarendra, K. Tieu, A. Widodo, C. A. S. Moodie, and A. K. Tieu, "Degradation trend estimation and prognosis of large low speed slewing bearing lifetime," *Applied Mechanics and Materials*, vol. 493, pp. 343–348, 2014.
- [10] Y. Lei, M. J. Zuo, Z. He, and Y. Zi, "A multidimensional hybrid intelligent method for gear fault diagnosis," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1419–1430, 2010.
- [11] S. M. Pincus, "Approximate Entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.
- [12] S. Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 110–117, 1995.
- [13] R. Yan and R. X. Gao, "Approximate entropy as a diagnostic tool for machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 824–839, 2007.
- [14] T. Ye, Z. L. Wang, and L. Chen, "Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping," *Mechanical Systems and Signal Processing*, vol. 114, no. 1, pp. 658–673, 2019.
- [15] R. Yan, Y. Liu, and R. X. Gao, "Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines," *Mechanical Systems and Signal Processing*, vol. 29, pp. 474–484, 2012.
- [16] 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 of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [17] 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.
- [18] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration*, vol. 289, no. 4–5, pp. 1066–1090, 2006.
- [19] X. S. Lou and K. A. Loparo, "Bearing fault diagnosis based on wavelet transform and fuzzy inference," *Mechanical Systems and Signal Processing*, vol. 18, pp. 1077–1095, 2014.
- [20] Z. Wang, C. Lu, Z. Wang, H. Liu, and H. Fan, "Fault diagnosis and health assessment for bearings using the Mahalanobis-Taguchi system based on EMD-SVD," *Transactions of the Institute of Measurement and Control*, vol. 35, no. 6, pp. 798–807, 2013.
- [21] A. Rai and S. H. Upadhyay, "Bearing performance degradation assessment based on a combination of empirical mode decomposition and K-medoids clustering," *Mechanical Systems and Signal Processing*, vol. 93, pp. 16–29, 2017.
- [22] Z. Wei, Y. Wang, S. He, and J. Bao, "A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection," *Knowledge-Based Systems*, vol. 116, pp. 1–12, 2017.
- [23] P. W. Tse and D. Wang, "Enhancing the abilities in assessing slurry pumps' performance degradation and estimating their remaining useful lives by using captured vibration signals," *Journal of Vibration and Control*, vol. 23, no. 12, pp. 1925–1937, 2017.
- [24] P. Vincent, H. Larochelle, and I. Lajoie, "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, pp. 3371–3408, 2010.
- [25] P. Vincent, H. Larochelle, and Y. Bengio, "Extracting and composing robust features with denoising autoencoders," *International Conference*, pp. 1096–1103, 2008.
- [26] J. Feng, Y. G. Lei, and L. Jing, "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.
- [27] F. Y. Lv, C. L. Wen, and M. Q. Liu, "Weighted time series fault diagnosis based on a stacked sparse autoencoder," *Journal of Chemometrics*, vol. 31, no. 9, pp. 1–16, 2017.
- [28] Y. Qi, C. Shen, D. Wang, J. Shi, X. Jiang, and Z. Zhu, "Stacked sparse autoencoder-based deep network for fault diagnosis of

- rotating machinery," *IEEE Access*, vol. 5, pp. 15066–15079, 2017.
- [29] F. Xu, W. t. P. Tse, and Y. L. Tse, "Roller bearing fault diagnosis using stacked denoising autoencoder in deep learning and Gath-Geva clustering algorithm without principal component analysis and data label," *Applied Soft Computing*, vol. 73, pp. 898–913, 2018.
 - [30] B. Leng, S. Guo, X. Zhang, and Z. Xiong, "3D object retrieval with stacked local convolutional autoencoder," *Signal Processing*, vol. 112, pp. 119–128, 2015.
 - [31] Y. Liu, X. Feng, and Z. Zhou, "Multimodal video classification with stacked contractive autoencoders," *Signal Process.*, vol. 120, pp. 761–766, 2015.
 - [32] J. Li, Z. Struzik, L. Zhang et al., "Feature learning from incomplete EEG with denoising autoencoder," *Neurocomputing*, vol. 165, pp. 23–31, 2014.
 - [33] 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.
 - [34] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.
 - [35] Q. C. Zhang, C. S. Zhu, L. T. Yang, Z. K. Chen, L. Zhao, and P. Li, "An incremental CFS algorithm for clustering large data in industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1193–1201, 2017.
 - [36] F. Xu, Y. J. Fang, R. Zhang, Z. M. Kong, and R. L. Tang, "A fault diagnosis method combined with ensemble empirical mode decomposition, base-scale entropy and clustering by fast search algorithm for roller bearings," *Journal of Vibroengineering*, vol. 18, no. 7, pp. 4472–4490, 2016.
 - [37] C. C. Tan and C. Eswaran, "Using autoencoders for mammogram compression," *Journal of Medical Systems*, vol. 35, no. 1, pp. 49–58, 2011.
 - [38] P. Nectoux, R. Gouriveau, and K. Medjaher, *PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test IEEE International Conference on Prognostics and Health Management*, Denver, CO, USA, 2012.
 - [39] J. Dolz, N. Betrouni, M. Quidet et al., "Stacking denoising auto-encoders in a deep network to segment the brainstem on MRI in brain cancer patients: a clinical study," *Computerized Medical Imaging and Graphics*, vol. 52, pp. 8–18, 2016.
 - [40] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, pp. 174102–174111, 2002.
 - [41] 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.
 - [42] Z. Wu, Q. Li, W. W. Wu, and M. B. Zhao, "Crowdsourcing model for energy efficiency retrofit and mixed-integer equilibrium analysis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4512–4524, 2019.
 - [43] Z. Kong, S. Yang, D. Wang, L. Hanzo, and H. Z. Lajios, "Robust beamforming and jamming for enhancing the physical layer security of full duplex radios," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 12, pp. 3151–3159, 2019.
 - [44] J. Lai and X. Lu, "Nonlinear mean-square power sharing control for ac microgrids under distributed event detection," *IEEE Transactions on Industrial Informatics*, vol. 1, 2020.
 - [45] Z. C. Dong, M. Tian, R. L. Tang, X. Li, and J. G. Lai, "Improving the robustness of spatial networks by link addition: more and dispersed links perform better," *Nonlinear Dynamics*, vol. 100, pp. 2287–2298, 2020.
 - [46] Z. Dong, M. Tian, Y. Lu, J. Lai, R. Tang, and X. Li, "Impact of core-periphery structure on cascading failures in interdependent scale-free networks," *Physics Letters A*, vol. 383, no. 7, pp. 607–616, 2019.
 - [47] F. Xu, Z. L. Huang, F. F. Yang, D. Wang, and K. L. Tsui, "Constructing a health indicator for roller bearings by using a stacked auto-encoder with an exponential function to eliminate concussion," *Applied Soft Computing*, vol. 89, Article ID 106119, 2020.
 - [48] F. Xu, Y. J. Fang, and R. Zhang, "PCA-GG rolling bearing clustering fault diagnosis based on EEMD fuzzy entropy," *Computer Integrated Manufacturing System*, vol. 22, no. 11, pp. 2631–2642, 2016.
 - [49] R. Tang, X. Li, and J. Lai, "A novel optimal energy-management strategy for a maritime hybrid energy system based on large-scale global optimization," *Applied Energy*, vol. 228, pp. 254–264, 2018.
 - [50] J. Lai, X. Lu, X. Yu, and A. Monti, "Stochastic distributed secondary control for ac microgrids via event-triggered communication," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 2746–2759, 2020.
 - [51] R. Tang, Z. Wu, and X. Li, "Optimal operation of photovoltaic/battery/diesel/cold-ironing hybrid energy system for maritime application," *Energy*, vol. 162, pp. 697–714, 2018.
 - [52] J. Lai, X. Lu, X. Yu, and A. Monti, "Cluster-oriented distributed cooperative control for multiple AC microgrids," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 11, pp. 5906–5918, 2019.
 - [53] J. G. Lai, X. Q. Lu, A. Monti, and G. P. Liu, "Stochastic distributed pinning control for co-multi-inverter networks with a virtual leader," *IEEE Transactions on Circuits and Systems-II*, vol. 67, no. 10, pp. 2094–2098, 2019.