

Research Article

Life Prediction of the Gear Transmission System with Multicharacteristics

Junliang Li,¹ Bin Ren ,¹ Zhanpu Xue ,² Bowen Yin,¹ and Hao Zhang²

¹School of Mechanical Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050018, China

²School of Mechanical Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China

Correspondence should be addressed to Bin Ren; renbin@stdu.edu.cn

Received 20 July 2022; Accepted 20 September 2022; Published 25 October 2022

Academic Editor: Yu-Ling He

Copyright © 2022 Junliang Li 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.

This study obtains and predicts multifault data in the key transmission and connection systems with gears. Model building is based on the multikernel extreme learning machine with the method of maximum correlation kurtosis deconvolution and variational mode decomposition. To this end, the realization form of the life prediction is first studied by enhancing the low-frequency signal. Then, the larger correlation coefficient is selected as the sensitive feature parameter aiming at mapping to a feature space by the randomly initialized hidden layer in the learning machine, and the weight value of output layer is obtained using the least square method. A case study on the fault diagnosis of gear transmission system is conducted in the end to illustrate the proposed approach.

1. Introduction

The condition of the key transmission connection systems, such as gears, has an important effect on the safe performance of large engineering equipment. In the process of long-term variable load service and under the influence of various uncertain excitations, the interaction between the functional components of the entire system elongates the failure degradation trend cycle. Moreover, the degradation data volume is large, and many interference frequency components exist. As a result, finding failure timely and making an early warning is difficult. Therefore, many researchers have focused on extracting the more characteristic information of the fault as possible from the vibration signal to realize accurate identification of the fault. They have also focused on ensuring the safety and stability of the equipment operation process.

An ELM is a simple, easy-to-use, widely used, and effective single hidden layer feedforward neural network [1–3]. In recent years, several scholars have used the ELM method to diagnose faults and predict the life of vulnerable parts, such as gears. Wei Chao et al. proposed the EEMD which is singular value decomposition method to extract

fault feature and identify the gear fault of the EEMD–ELM method based on the study of the gearbox fault diagnosis and identification models [4]. The comparison shows that an ELM has a faster running speed and higher classification accuracy than an SVM. Yang Lu et al. proposed an optimizing algorithm of the ELM parameters based on the simulated annealing particle swarm algorithm to diagnose the faults of the wind turbine gearbox and solve the poor network structure stability and classification accuracy of ELMs [5]. The results show that the method has better stability and reliability. Qin et al. proposed a gear fault diagnosis method based on the KELM [6]. The experimental results show that the KELM rolling gear fault diagnosis classification model has higher accuracy and stability than the SVM and ELM fault classification models.

Zhou et al. used the EEMD method to decompose the vibration signal for obtaining a fault feature matrix composed of IMF, and a new ELM algorithm combining an integrated ELM and an evolutionary learning machine was proposed [7]. The artificial bee colony algorithm was used to optimize the input weights and the hidden layer bias. Rodriguez et al. used the stationary wavelet singular entropy to obtain high-quality fault features and input the obtained

features into the KELM classifier for solving the difficulty of the current fault diagnosis method to accurately diagnose the fault that occurs in the variable-speed rotation phenomenon [8]. For the vibration characteristics and common faults of the gears, Zhou et al. proposed a cascaded feature reduction method based on global supervision Laplacian score and kernel principal component analysis and a multiple fault recognition method of a binary tree KELM based on the particle swarm optimization [9]. However, the vibration signal of the gear in this method is only used for feature extraction, which reduces the accuracy of fault diagnosis. Considering the problem of nonstationary and nonlinearity in the gearbox vibration signal, Wang proposed an extracting method of feature fault based on the complementary set empirical mode decomposition and the multiscale permutation entropy [10]. Wang and Wang proposed a method combining the variational modal decomposition and ELM methods to predict the life of the gear, but the accuracy is insufficiently high [11]. Gu et al. used the discrete particle swarm optimization to optimize the multiscale wavelet kernel function of the KELM [12]. This way improves the convergence speed and classification accuracy of the algorithm and has significant effects on the life prediction of the wind gearbox bearing. Liu and Huang proposed a personalized diagnosis method of the gear faults based on finite element simulation and ELM [13]. Liu et al. proposed a failure early warning method of the wind turbine gearbox based on security projection, nuclear ELM, and information entropy [14]. Chen et al. combined the integrated convolutional neural network and ELM in fault diagnosis to accurately detect different faults and predict their life [15].

In summary, the ELM algorithm has been widely used in the fault diagnosis and the life prediction of the key transmission connection systems, such as gears, and has significant effects. However, the following problems still exist: (1) the prediction results of the ELM regression model are greatly affected by the input parameters; the error is larger, and the quality is poor; (2) the accuracy of the failure trend prediction is unstable, the learning ability is poor, and the extrapolation ability is weak. On the basis of the abovementioned analysis, a new algorithm combining the MCKD-VMD method with the multikernel ELM is proposed and used in analyzing the gear transmission system [16, 17]. The effectiveness of the algorithm is verified by numerical simulation and gear broken tooth fault experiment.

2. Obtaining Methods of Multiple Features

2.1. Basic Principles of the MCKD Method. The maximum correlation kurtosis deconvolution is a new convolution technique proposed on the basis of the minimum entropy deconvolution to enhance the periodic impact component of the signal. It fully utilizes the periodic characteristics of impact faults, which can effectively suppress the influence of noise and other interference components; its essence is to find an FIR filter $f(L)$ (L is the length of the filter) for maximizing the correlation kurtosis of the original impact sequence, and this way, it restores its characteristics to enhance the signal [18]. Correlation kurtosis is defined as follows:

$$K_{C,M}(T) = \frac{\sum_{k=1}^N \left(\prod_{m=0}^M y_{k-mT} \right)^2}{\left(\sum_{k=1}^N y_k^2 \right)^{M+1}}, \quad (1)$$

where y_k is the original signal the fault signal collected by the sensor; $y_k = \sum_{l=1}^L f_l x_{k-l+1}$ when $k \neq 1, 2, \dots, N$, $x_k = 0$, and $y_k = 0$; N is data length; and $\mu_y = 0$.

To maximize the correlation kurtosis by selecting an optimal filter $f(l)$, the objective function is given as

$$\begin{aligned} \text{MCKD}_M(T) &= \max_f [K_{C,M}(T)] \\ &= \max_f \frac{\sum_{k=1}^N \left(\prod_{m=0}^M y_{k-mT} \right)^2}{\left(\sum_{k=1}^N y_k^2 \right)^{M+1}}, \end{aligned} \quad (2)$$

where T is the period of the impact signal; M is the number of displacements; L is the length of the filter; f is the filter vector; and $f = [f_1, f_2, \dots, f_L]^T$.

To solve the objective function, formula (2) can be differentiated as

$$\begin{aligned} \frac{d}{df_{kl}} K_{C,M}(T) &= \frac{d}{df_{kl}} \frac{\sum_{k=1}^N \left(\prod_{m=0}^M y_{k-mT} \right)^2}{\left(\sum_{k=1}^N y_k^2 \right)^{M+1}} \\ &= 0 \quad (k=1, 2, \dots, L). \end{aligned} \quad (3)$$

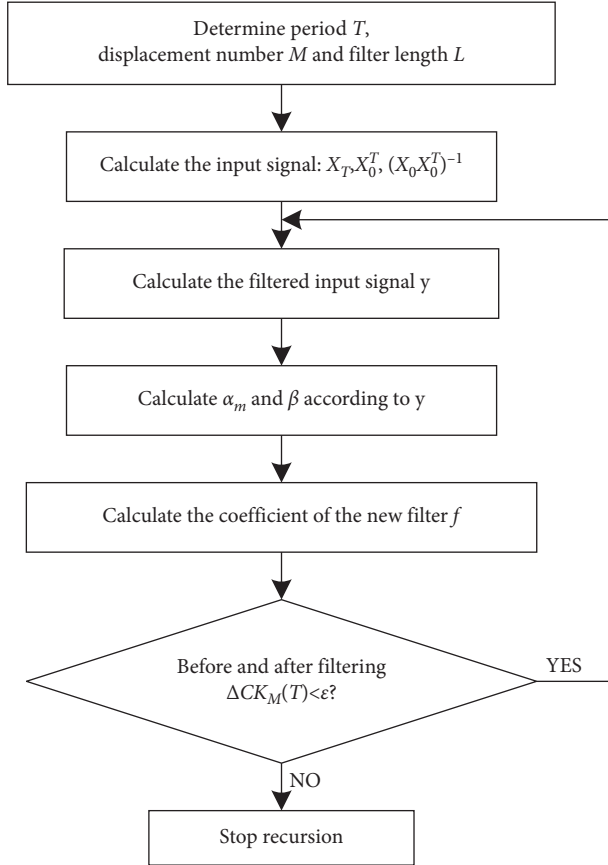
Then, the matrix representation method of filter $f(l)$ coefficients is given as

$$f = \frac{\|y\|^2}{(M+1)\|\beta\|^2} (X_0 X_0^T)^{-1} A, \quad (4)$$

where $f = [f_1, f_2, \dots, f_L]^T$ and $A = \sum_{m=0}^M X_{mT} \alpha_m$, then

$$\begin{aligned} X_0 &= \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ 0 & x_1 & \cdots & x_{N-1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & x_{N-L+1} \end{bmatrix}_{L \times N}, \\ \beta &= \begin{bmatrix} y_1 & y_{1-T} & \cdots & y_{1-MT} \\ y_2 & y_{2-T} & \cdots & y_{2-MT} \\ \cdots & \cdots & \cdots & \cdots \\ y_N & y_{N-T} & \cdots & y_{N-MT} \end{bmatrix}_{N \times 1}, \\ X_{mT} &= \begin{bmatrix} x_{1-mT} & x_{2-mT} & \cdots & x_{N-mT} \\ 0 & x_{1-mT} & \cdots & x_{N-mT-1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & x_{N-mT-L+1} \end{bmatrix}_{L \times N}, \\ \alpha_m &= \begin{bmatrix} y_{1-mT}^{-1} (y_1 y_{1-T} \cdots y_{1-MT})^2 \\ y_{2-mT}^{-1} (y_2 y_{2-T} \cdots y_{2-MT})^2 \\ \cdots \\ y_{N-mT}^{-1} (y_N y_{N-T} \cdots y_{N-MT})^2 \end{bmatrix}_{N \times 1}. \end{aligned} \quad (5)$$

The filter parameters $f(l)$ are solved by the iterative method. The specific process is shown in Figure 1.


 FIGURE 1: Iterative process diagram of the filter $f(l)$ parameter.

2.2. *Basic Principles of VMD.* The VMD is a signal decomposition estimation method based on constructing variational problems and solving variational problems. It has significant effects in many aspects compared with the empirical modal decomposition. In the method, the frequency center and bandwidth of each component are determined by iteratively searching for the optimal solution of the variational model in the process of obtaining the decomposed components. This way adaptively realizes the frequency domain division of the signal and the effective separation of the components.

Specifically, the signal $x(t)$ is first decomposed into KUs components by the VMD algorithm, and the variation is constructed by seeking the bandwidth of the modal component Us with the help of methods, such as Hilbert transformation. Then, the Lagrangian multiplication operator $\lambda(t)$ and the quadratic penalty factor α are introduced. The alternating direction multiplier method is used to continuously update each mode and its center frequency and gradually demodulates each mode to the corresponding base frequency band [19]. The corresponding center frequency of each mode is extracted. The expression of the solution process of modal $\hat{y}_k^{n+1}(\omega)$ is

$$\hat{y}_k^{n+1}(\omega) = \frac{\hat{Y}(\omega) - \sum_{i \neq k} \hat{Y}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2}, \quad (6)$$

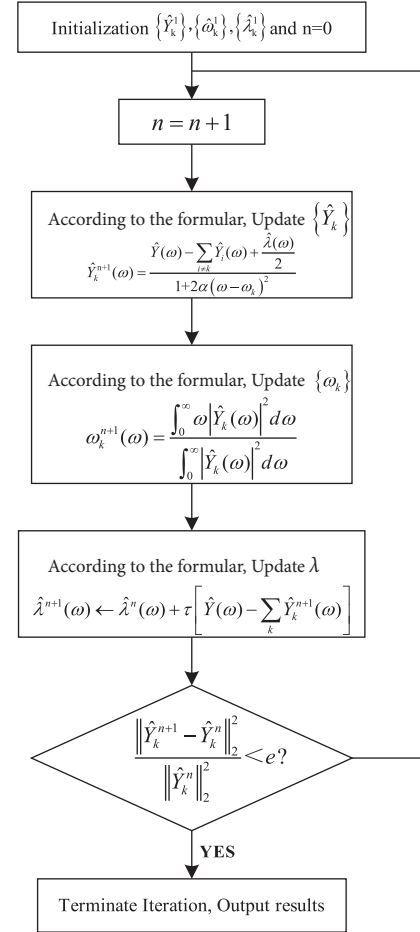


FIGURE 2: VMD decomposition algorithm flowchart.

where $\hat{Y}(\omega)$ is the Fourier transforms of $y(t)$; $\hat{Y}_k(\omega)$ is the Fourier transforms of $y_i(t)$; $\hat{\lambda}(\omega)$ is the Fourier transform of $\lambda(t)$.

The time domain signal of each modal component can be obtained by performing the Fourier transform on the filtered signal. The center of gravity of the updated current modal power spectrum is

$$\omega_k^{n+1}(\omega) = \frac{\int_0^\infty \omega |\hat{y}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{y}_k(\omega)|^2 d\omega}, \quad (7)$$

where $y_k(\omega)$ is the modal function in the $n+1$ th cycle; $\omega_k^{n+1}(\omega)$ is the center of gravity of the power spectrum of the updated modal function; and $\lambda^{n+1}(\omega)$ is the multiplication operator in the $n+1$ th cycle.

The realization process of the algorithm is shown in Figure 2.

The fault signal data of the key transmission connection systems, such as gears, have large volume and difficult extraction and are nonlinear and nonstationary when the variational modal decomposition method is used to directly decompose the fault signal. This method cannot accurately identify the fault frequency and completely extract the hidden information. The maximum correlation kurtosis deconvolution is a new convolution technique that enhances

the periodic impact components of the signal, which can make the low-frequency submerged components of the signal obvious. Therefore, the method of combining MCKD and VMD is used in this study to extract the multiple characteristics of faults. First, the collected vibration signals are fused with the correlation functions, and the signals with higher correlation are merged together to effectively remove the interference signals. Second, the fusion signal is enhanced by the MCKD method to make the low-frequency signal more obvious. Third, the signal enhanced by MCKD is decomposed by VMD to obtain several modal components, and the components with larger correlation coefficients are reconstructed. Finally, the power spectrum analysis is performed to identify the gear failure frequency characteristics.

3. Failure Prediction Method Based on Multikernel ELM

3.1. Prediction Model. We suppose N samples (x_i, y_i) , where $x_i = [x_{i1}, x_{i2}, \dots, x_{ij}]^T \in R^j$ and $y_i = [y_{i1}, y_{i2}, \dots, y_{ij}]^T \in R^k$. Then, the network output of an ELM with activation function $g(x)$ and hidden layer nodes L is

$$f_L(x) = \sum_{i=1}^L \beta_i g(a_i \cdot x_i + b_i), x_i \in R^j, \beta_i \in R^k, \quad (8)$$

where $\alpha_l = [a_{l1}, a_{l2}, \dots, a_{lj}]^T$ is the input weight from the input layer to the l -th hidden layer node; b_l is the deviation of the l -th hidden layer node; $\beta_l = [\beta_{l1}, \beta_{l2}, \dots, \beta_{lk}]^T$ is the output weight connecting the l -th hidden layer node; $a_l \cdot x_i$ is the inner product of the vector sum; and the excitation function $g(x)$ includes "sigmoid," "RBF," or "sine."

We suppose that the ELM has L hidden layer nodes (given in advance); if this feedforward neural network can approximate these N samples with zero error, then we have a_l, b_l , and β_l :

$$f_L(x) = \sum_{i=1}^L \beta_i g(a_i \cdot x_i + b_i) = y_i, i = 1, 2, \dots, N. \quad (9)$$

They are simplified as

$$H\beta = Y, \quad (10)$$

where H is the hidden layer output matrix, which is given as

$$H = \begin{bmatrix} g(a_1 \cdot x_1 + b_1) & \cdots & g(a_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(a_1 \cdot x_N + b_1) & \cdots & g(a_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \quad (11)$$

where Y is the desired output matrix, and β is the output weight matrix.

The input weights and deviations in the ELM algorithm can be given randomly, H is the deterministic matrix, and the connection weights of the hidden and output layers can be obtained by the least square solution as

$$\beta = H^* Y, \quad (12)$$

where H^* is the Moore–Penrose generalized inverse matrix of H (hidden layer output) matrix.

The KELM mainly uses kernel mapping to replace the random mapping of ELM and uses the kernel function to map the input samples to high-order space operations. This way greatly improves the stability of the model. Thus, the classification and regression capabilities are better than those of the kernelless ELM algorithm. The connection weight of the input and hidden layers and the threshold of the hidden layer can be set randomly and do not need to be iteratively adjusted. They are determined by solving the equation set at one time, which is faster when the learning accuracy is ensured. The expression of the KELM is

$$f(x) = \begin{bmatrix} K(x, x_1) \\ \cdots \\ K(x, x_N) \end{bmatrix} (1/c + K(x_i, x_j))^{-1} T, \quad (13)$$

where K is the kernel function; C is the penalty function; I is the diagonal matrix; and T is the expected output vector.

The kernel function of KELM is the core of the KELM algorithm, which represents the kernel mapping ability of the KELM algorithm [20]. The selection of the kernel function parameters in the model is determined by the modeling data and is difficult to select accurately. The commonly used forms are shown as follows:

Linear kernel function (linear) is given as

$$K(x_i, x_j) = x_i \cdot x_j. \quad (14)$$

Polynomial kernel function (polynomial) is

$$K_{\text{poly}}(x_i, x_j) = (x_i \cdot x_j + b)^d, b \geq 0. \quad (15)$$

Radial basis kernel function (RBF) is

$$K_{\text{RBF}}(x_i, x_j) = \exp(-\delta \|x_i - x_j\|^2), \delta > 0. \quad (16)$$

S-type kernel function (sigmoid) is

$$K(x_i, x_j) = \tanh(ax_i \cdot x_j + b), a > 0, v > 0. \quad (17)$$

Among the abovementioned four kernel functions, the linear kernel function has a weaker learning ability and simpler algorithm. The polynomial kernel function is a nonstandard kernel function, which is suitable for orthogonal normalized data and can effectively solve the "dimension disaster" problem in the high-dimensional feature space operations. However, the parameters are too many. The sigmoid kernel function is derived from a neural network, and the calculation amount is more complicated when used as an activation function. Nevertheless, the function is smooth and easy to obtain. Considering that the performance degradation process of the multifeature failure is more complicated, the combination of polynomial and sigmoid is selected and applied to the life trend prediction:

$$K_{\text{multi-kernel}} = \mu K_{\text{polynomial}}(x, x_i) + (1 - \mu) K_{\text{Sigmoid}}(x, x_i), \quad (18)$$

where $K_{\text{polynomial}}$ is the polynomial kernel function; K_{Sigmoid} is the sigmoid type kernel function; and μ is the proportion of the polynomial and sigmoid kernel function ($0 \leq \mu \leq 1$).

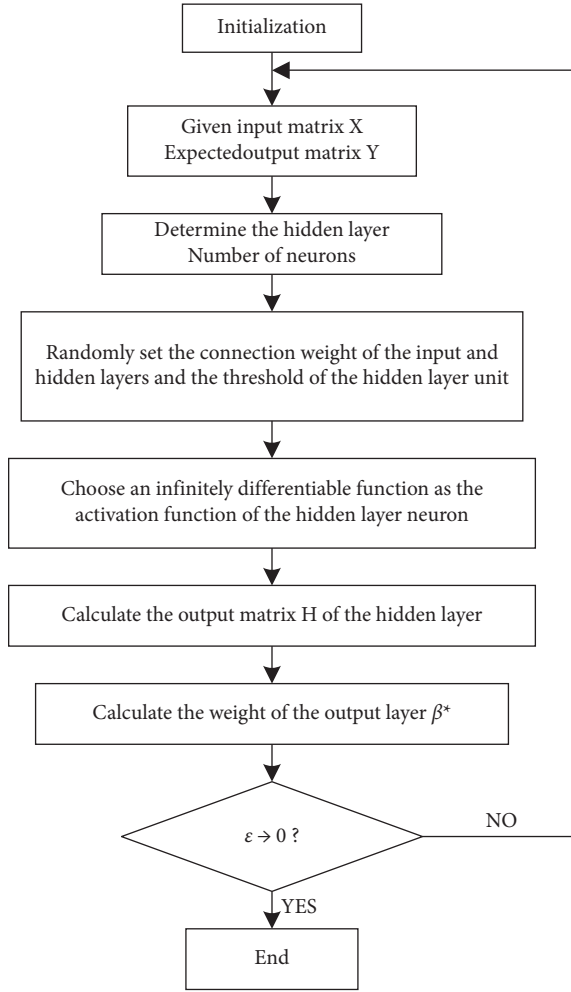


FIGURE 3: Process of the (P-S) KELM algorithm.

The multikernel ELM model (referred to as (P-S) KELM) is

$$f(x) = \begin{bmatrix} K_{\text{multi-kernel}}(x, x_1) \\ \dots \\ K_{\text{multi-kernel}}(x, x_N) \end{bmatrix} \left(\frac{1}{c} + K_{\text{multi-kernel}}(x_i, x_j) \right)^{-1} T. \quad (19)$$

The abovementioned formula shows that the number of hidden layer nodes, initial weights, offsets, and the calculation of the output matrix H are no longer considered in the solving process of the (P-S) KELM model. However, the output function value can be solved in the form of a kernel function.

3.2. Predicting Algorithm. First, the parameters, such as samples, input weights, and hidden layer node deviations, are initialized. The specific algorithm flowchart is shown as Figure 3:

3.3. Evaluating Index. To accurately reflect the prediction accuracy of the (P-S) KELM, three evaluating indicators are proposed: the mean square error, the mean absolute error, and the R squared.

(1) Mean square error (MSE) is given as

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2, \quad (20)$$

where y_i is the true value in the test set, and \hat{y}_i is the predicted value in the test set. The mean square error is equivalent to the loss function of linear regression, which can be applied to the test set, and the final result is equivalent to the loss value.

(2) Mean absolute error (MAE) is given as

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|. \quad (21)$$

The average absolute error is the direct calculation of the residual error, which represents the average value of the absolute error between the predicted value and the true value. It is a linear fraction, and all individual differences have the same weight on the average.

(3) Accuracy (R squared) can be given as

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2 \text{residual}}{\sum_i (\hat{y}_i - y_i)^2 \text{total}}, \quad (22)$$

where $\sum_i (\hat{y}_i - y_i)^2 \text{residual}$ is the sum of squares of all the errors predicted by the trained model, and $\sum_i (\hat{y}_i - y_i)^2 \text{total}$ is the square of the sum.

Accuracy is the standard measured in the classification algorithm. The value is between (0, 1); if the value is closer to 1, then the accuracy becomes higher; otherwise, it becomes lower.

4. Simulation Signal Analysis

To verify the feasibility of this method, a fault simulation signal model is established for a gear as

$$\begin{aligned} S_1(t) &= \sin(90\pi t) (1 + 0.2 \cos(90\pi t)), \\ S_2(t) &= 0.8 \sin(100\pi t + \sin(5\pi t)), \\ S(t) &= S_1(t) + S_2(t) + 0.5 \sin(1000\pi t) + w(t). \end{aligned} \quad (23)$$

The simulated signals are shown in Figure 4.

The fault pulse is more obvious when the fault feature is extracted from the fault simulation signal. The fault frequency can be observed in its simple spectrum diagram. The low-frequency signal can be observed through the spectrum diagram before and after the MCKD signal is increased, as shown in Figure 5(b). The fault can be identified more clearly by strengthening, and the fault features are most obvious at 49.99 and 499.99 Hz. The time domain waveform and frequency spectrum diagrams are obtained by further VMD decomposition of the simulated signal. They are shown in Figure 6.

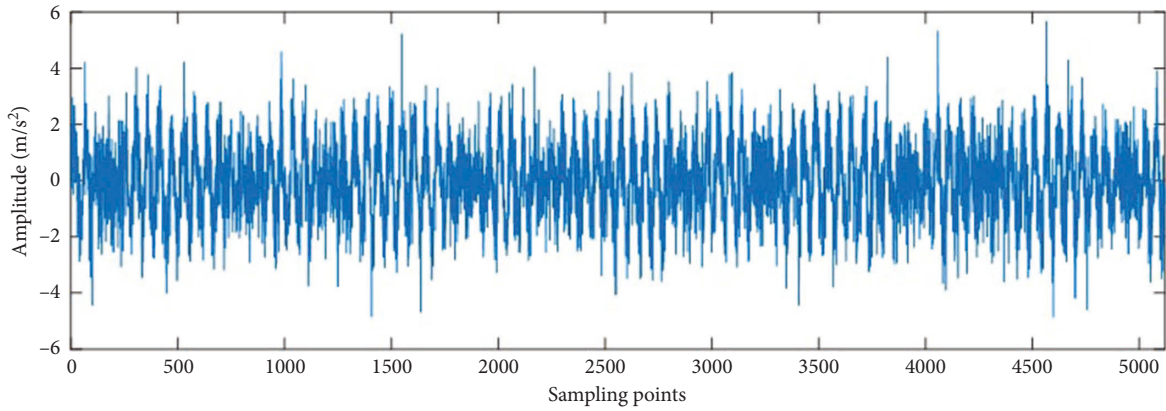
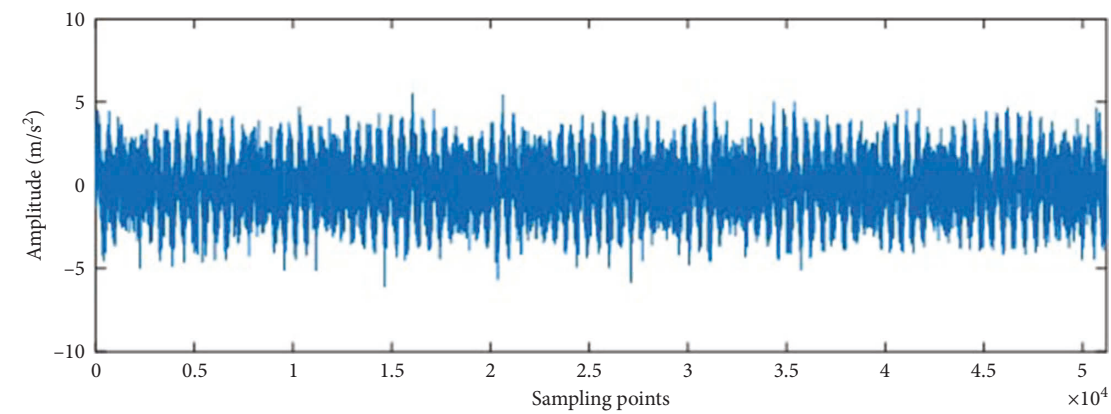
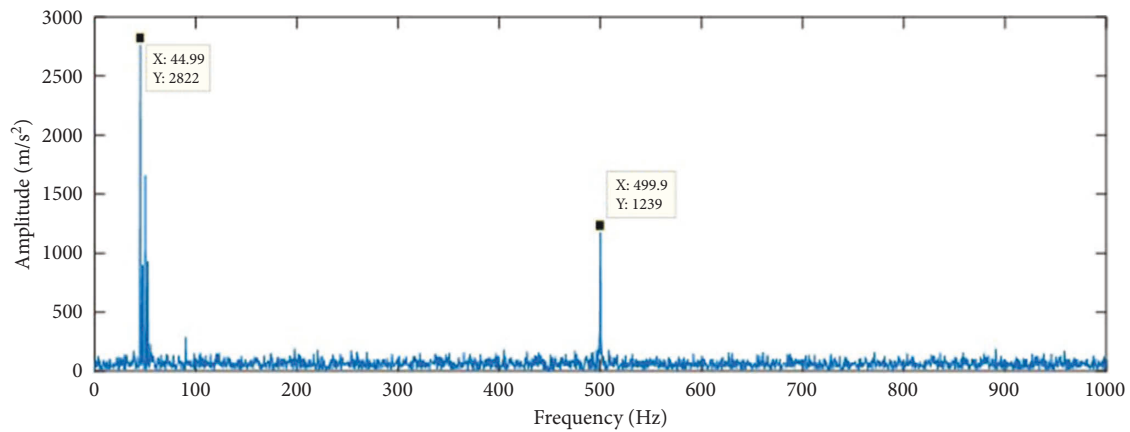


FIGURE 4: Time domain waveform of the simulated signal.



(a)

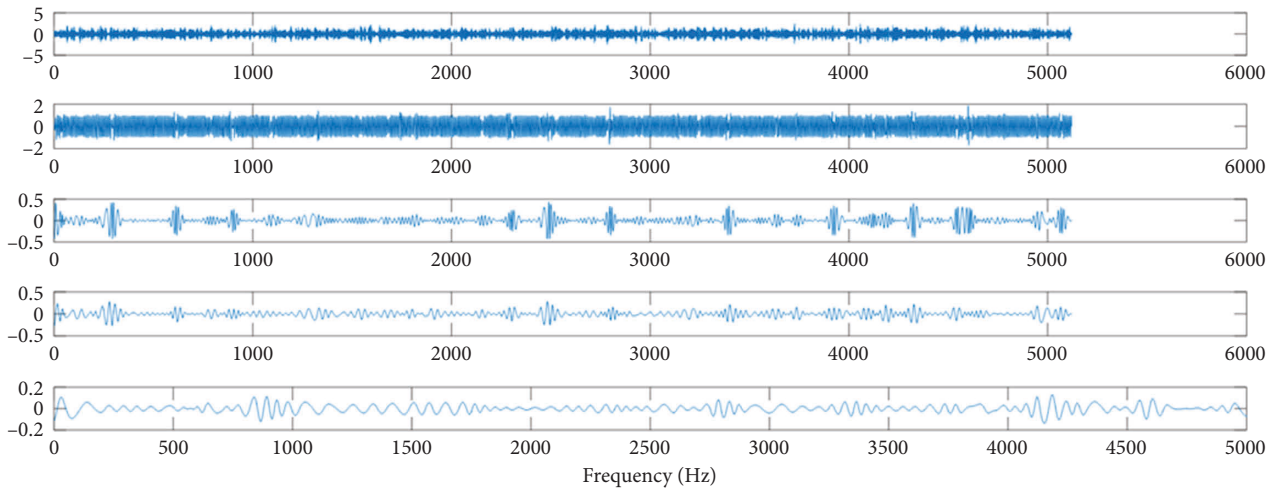


(b)

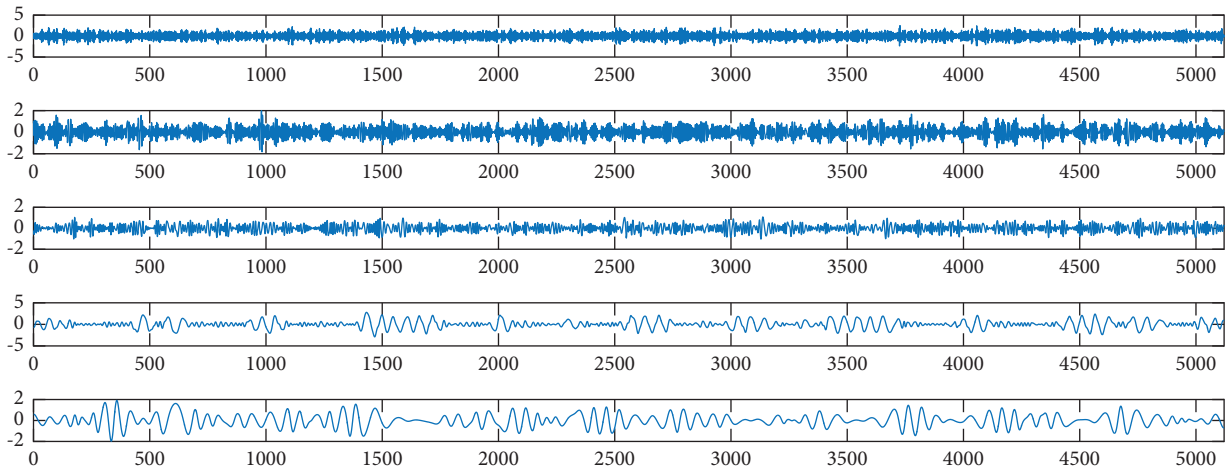
FIGURE 5: Time domain waveform and frequency spectrum of the simulated signal after MCKD. (a) Time domain waveform after MCKD. (b) Frequency spectrogram after MCKD.

The correlation coefficient values of the five modal components and the simulation signal are calculated, as shown in Figure 7. The figure shows that the correlation between components U_1 , U_3 , and U_4 , and the original signal is relatively higher. This condition indicates that they are more sensitive to the gear failures. Therefore, the three modal components are selected as the frequency domain characteristic values for the trend prediction of the gear failures.

Specifically, the three components U_1 , U_3 , and U_4 are merged into a matrix to form a new frequency domain eigenvalue as the input value of the (P-S) KELM model. The simulation signals after 6 s are intercepted to be used as the true value of the trend prediction of the simulated signal. As observed, the true trend of the simulated signal is the same as the predicted trend. However, individual data have errors, and the error reaches a maximum of 1500 Hz as the



(a)



(b)

FIGURE 6: Time domain waveform and frequency spectrum after VMD decomposition. (a) Time domain waveform. (b) Frequency spectrogram.

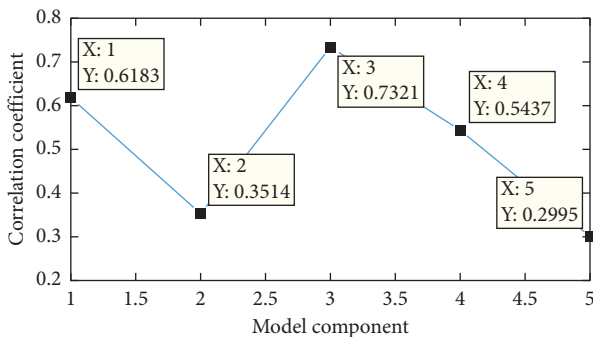


FIGURE 7: Correlation coefficient value between VMD decomposition result of the simulated and original signals.

frequency gradually rises. According to the results, the (P-S) KELM model predicts the trend of gear failure signals and is therefore feasible and accurate.

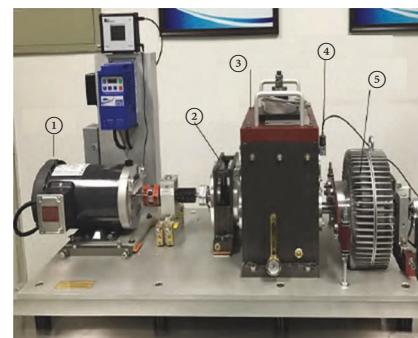


FIGURE 8: Fault diagnosis comprehensive test bench.

5. Experimental Verification

5.1. Acquisition of the Gear Failure Experiment Data. The gear fault diagnosis experiment in the fixed-axis gearbox on the American DDS power transmission fault diagnosis comprehensive test bench (Figure 8) is conducted to verify

TABLE 1: Gear teeth number and transmission ratio of each stage in the gearbox.

Transmission system	Primary transmission (planetary gear)	Secondary transmission (fixed-axis gear)	Thirdly transmission (fixed-axis gear)
Number of teeth	Sun gear28 ring gear100	High-speed shaft gear29 intermediate shaft gear100	Intermediate shaft pinion36 output shaft gear90
Transmission ratio	4.5714	3.4483	2.5

TABLE 2: Relevant frequencies of the transmission system (unit: Hz).

Rotation frequency of the input shaft	Rotation frequency of the high-speed shaft	Rotation frequency of the intermediate shaft	Rotation frequency of the output shaft	Meshing frequency in the first level	Meshing frequency in the third level	Meshing frequency in the second level
40	2.5375	8.7501	1.015	1120	91.35	253.75

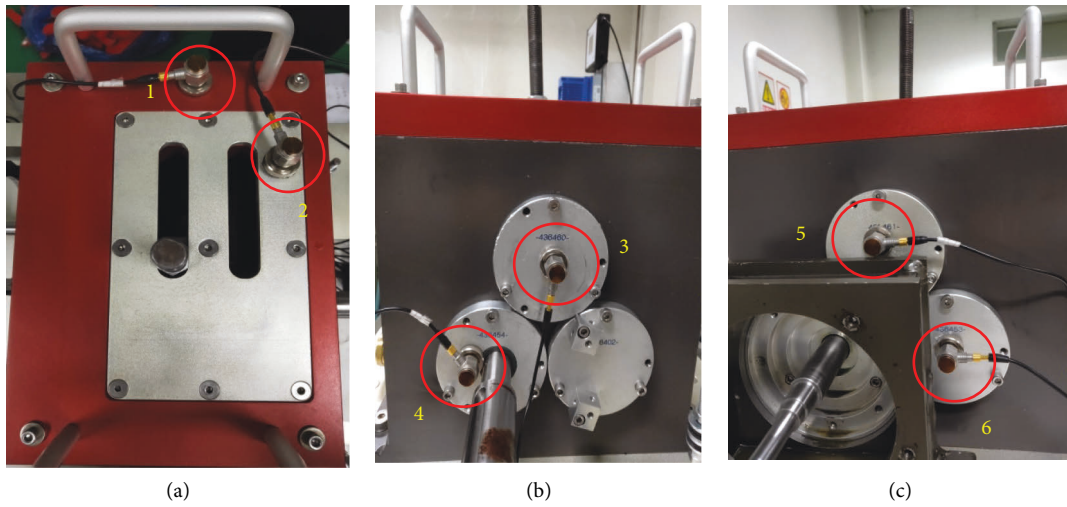


FIGURE 9: Measuring point position of the sensor. (a) Measuring point position on the upper box. (b) Measuring point position on the right box. (c) Measuring point position on the left box.

the effectiveness and practicability of the proposed method in extracting the fault features of the gear. The test bench is the electric motor①, the planetary gearbox②, the fixed-axis gearbox③, the position of the sensor④, and the magnetic powder brake⑤ from left to right. The number of the teeth and the transmission ratios of each transmission gear are shown in Table 1, and the relevant frequencies of the transmission system are shown in Table 2. The sensor adopts an acceleration sensor (SN178383).

The experiment is based on the state of normal, wear, and broken tooth simulation failure experiments of the gear in the gearbox. Six sensors are used for measuring point arrangement. Each sampling time is 20 s. Each sampling obtains six sets of vibration data, which are collected 10 times: $[x_{11}(t), \dots, x_{16}(t)]$, $[x_{21}(t), \dots, x_{26}(t)]$, \dots , $[x_{61}(t), \dots, x_{66}(t)]$. By using the data correlation fusion algorithm, every six sets of data are merged into one set of data. The faulty gear is located at the third-stage meshing position of the planetary gear in the intermediate shaft, the sampling frequency is 25600 Hz, and the state of each gear is intercepted for 30 s. The locations of the sensor are shown in Figure 9.

5.2. Predicting Analysis of the Gear Failure. The input model data should be accurate and of high quality to predict the trend of the faulty gear. The redundant data should be eliminated when analyzing the fault signal. The gear fault signal has the characteristics of weakness, nonlinearity, and nonstationary. Thus, the cross-correlation function fusion calculation of the collected fault signal is first conducted, and the signals with high correlation degree can be retained and merged together while removing the excess noise signal. Then, the signal is enhanced by the MCKD method and decomposed by the VMD method to obtain a series of modal components. Finally, the component with the larger correlation coefficient that is most sensitive to the gear fault characteristics are extracted as the frequency domain characteristic parameters of the gear broken tooth by the calculation of the correlation coefficient between the components and the original signal to predict the fault trend.

The time domain waveform and frequency spectrum diagrams of the vibration signal of the gear broken tooth fault are shown in Figure 10. According to the analysis, the fault characteristics cannot be identified and some high-frequency impact components exist due to the influence of

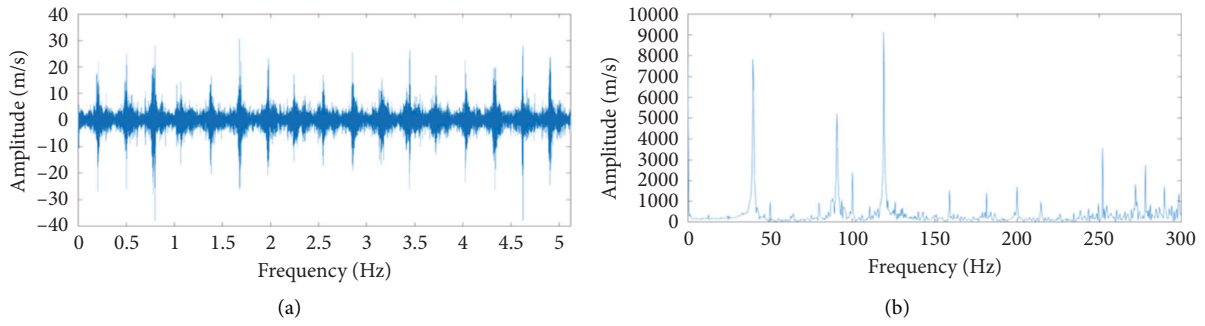


FIGURE 10: Time domain waveform and frequency spectrum of the gear broken tooth fault signal. (a) Time domain waveform. (b) Frequency spectrum.

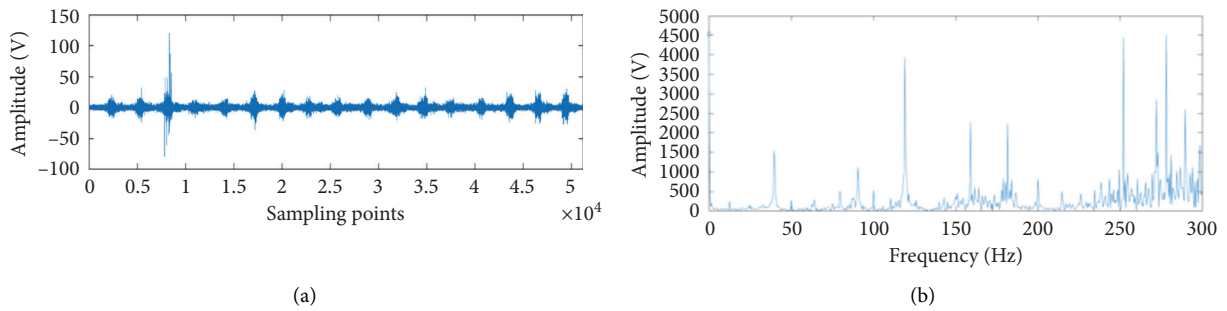


FIGURE 11: The time domain waveform and the frequency spectrum after MCKD processing. (a) Time domain waveform. (b) Frequency spectrum.

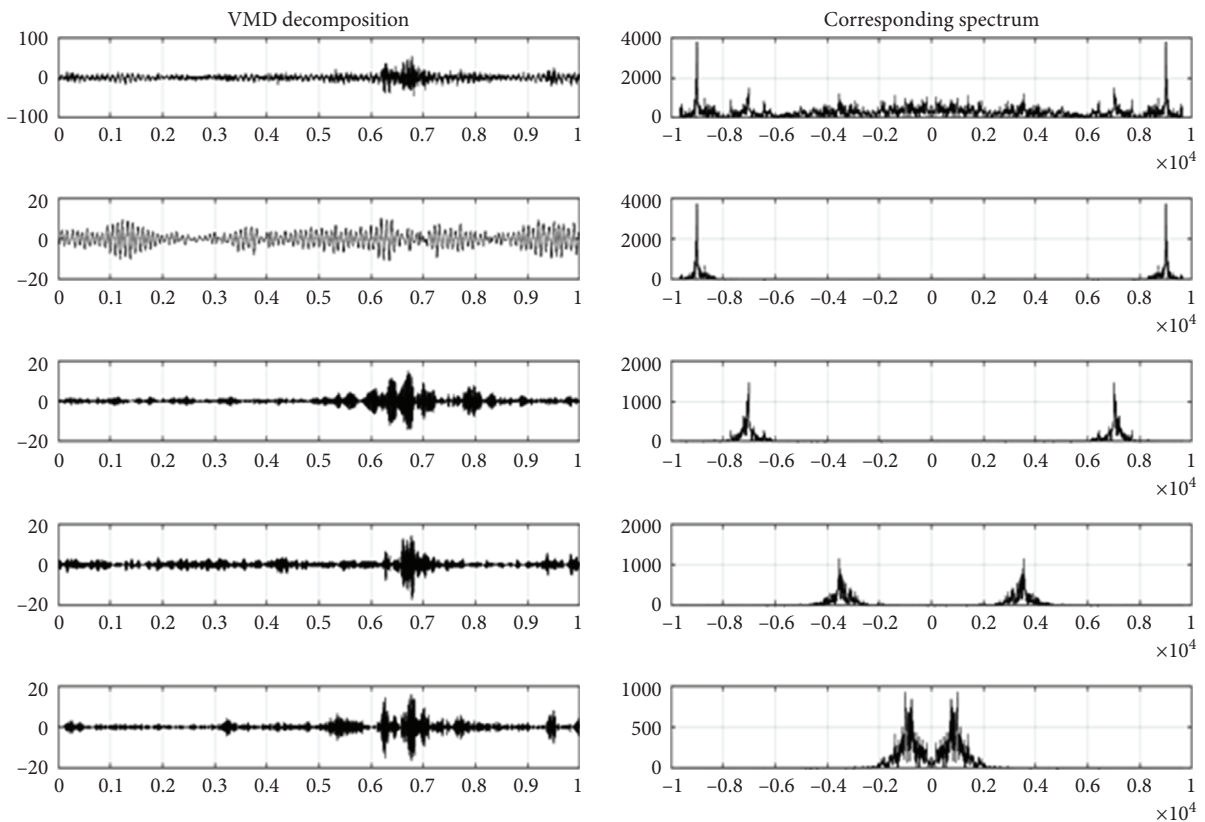


FIGURE 12: Time domain waveform diagram and figure.

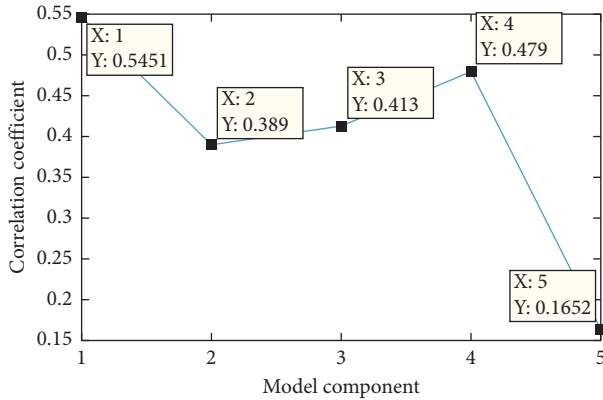


FIGURE 13: Correlation coefficient value between the corresponding spectrum diagram by VMD decomposition result and the original signal.

the noise in the gearbox operation and the accuracy factors of the equipment. The fault signal is enhanced by the MCKD method to obtain the time domain waveform and the spectrogram, which are shown in Figure 11. The figure shows that the time domain waveform of the fault signal has obvious impact components, and the low-frequency components in the spectrogram are enhanced and become obvious.

The signal enhanced by the MCKD method is decomposed by the VMD method. The VMD decomposition result and the corresponding spectrogram are shown in Figure 12. The correlation coefficient values of five modal components and the original gear broken tooth fault signal are calculated, as shown in Figure 13. The figure shows that the decomposed components U_1 , U_3 , and U_4 have a higher correlation with the original signal. This condition indicates that they are more sensitive to gear failures. The three modal components are selected as the frequency domain characteristic values for the trend prediction of the gear failure. The analysis result is consistent with the abovementioned simulation analysis result. The flowchart of the gear failure trend prediction is shown in Figure 14.

On the basis of the abovementioned analysis, the frequency domain sensitive characteristic values are selected: U_1 , U_3 , and U_4 , which are selected as the standard of gear failure trend prediction. The establishment steps of the (P-S) KELM model are shown in Figure 15. The abovementioned three components (under normal and worn gears) are merged into a matrix to form a new frequency domain eigenvalue as the input value of the (P-S) KELM model, as shown in Figure 16. The predicted result is compared with the fusion frequency domain characteristic value of the measured gear broken state. The prediction effect of the (P-S) KELM model is evaluated by three evaluation indicators: the mean square error (MSE), the mean absolute error (MAE), and the accuracy (R^2). The prediction result is shown in Figure 17.

The actual performance of the gear broken tooth fault is that a large impact signal begins to appear, and the frequency is obviously increased and unstable. The trend comparison charts of the predicting result of the KELM method, the (P-S) KELM method, and the broken tooth fault data are shown

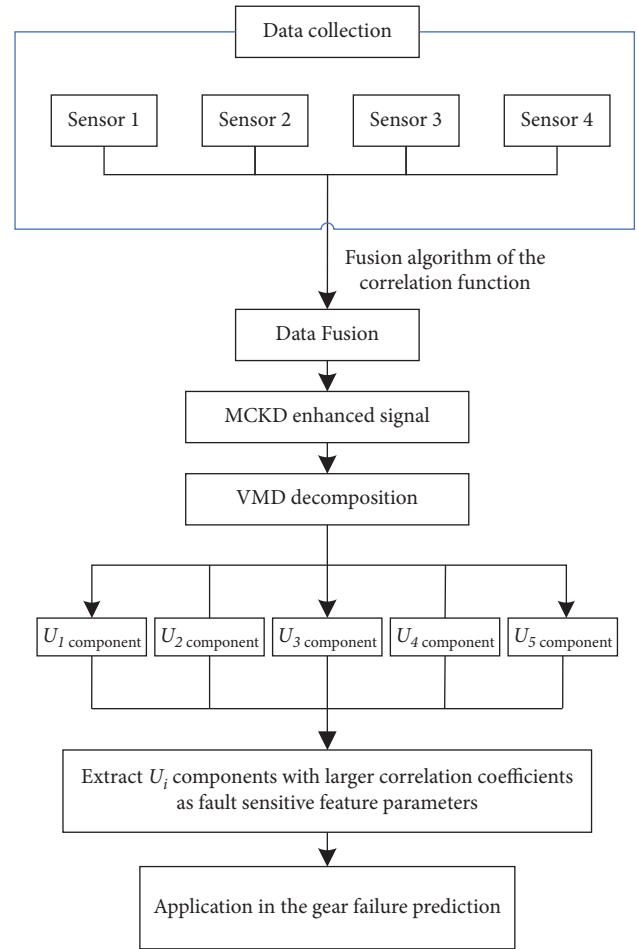


FIGURE 14: Flowchart of the trend forecast.

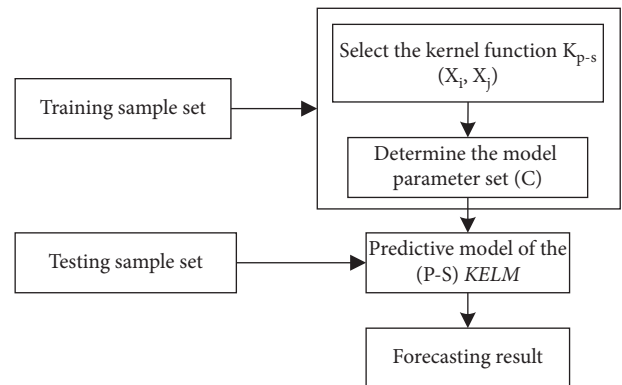


FIGURE 15: Establishment steps of the (P-S) KELM model.

in Figure 17. The analysis shows that the true value curve of the frequency domain characteristic of the gear broken state is consistent with the predicting curve of the (P-S) KELM method. The gear failure becomes more obvious as time increases. The frequency domain characteristic value significantly increases up to 1600 Hz after 12 s. At the same time, the predicting curve of the KELM method is consistent with the trend of the true value. However, a large error in the value exists, and the error becomes obvious with the increase

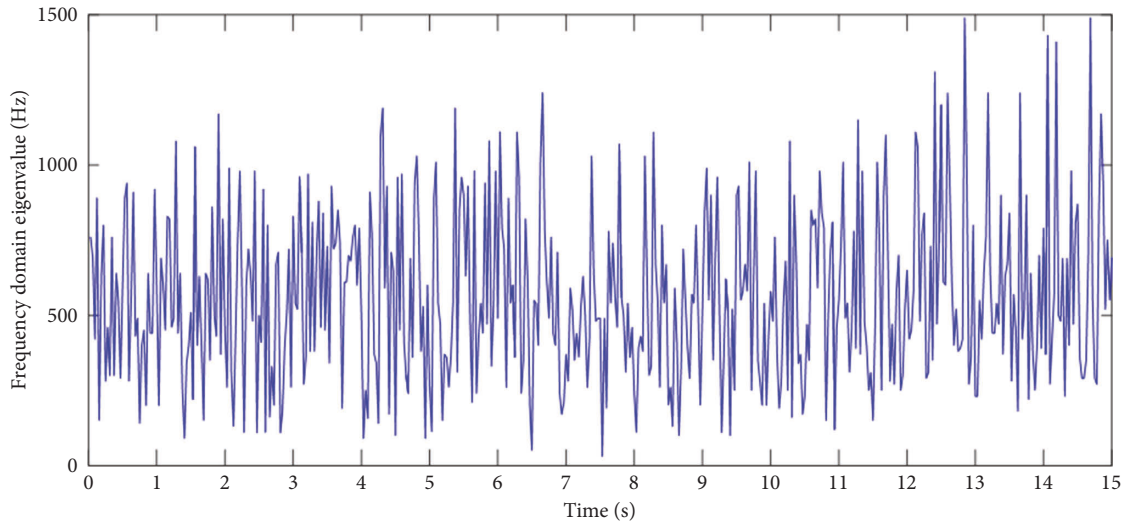


FIGURE 16: Fusions of the gear feature parameter fusion.

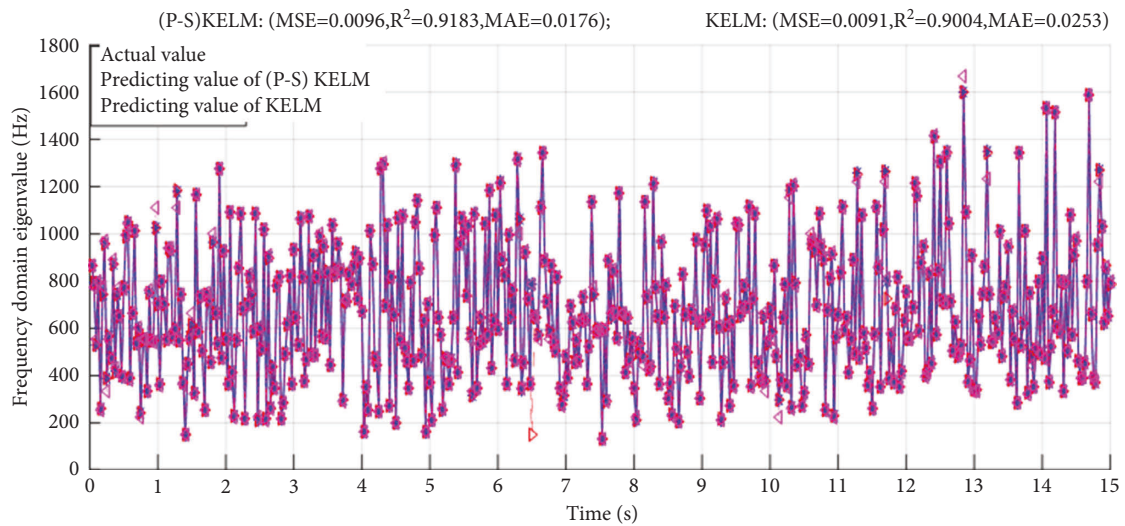


FIGURE 17: Comparisons of the gear trend forecast.

TABLE 3: Three evaluation index values under the condition of gear broken.

Evaluation index	MSE	MAE	R^2
(P-S) KELM value	0.0096	0.0176	0.9183
KELM value	0.0091	0.0253	0.9004

intime. The error is most obvious when time is 13 s, which is consistent with the true value. The value difference is approximately 1000 Hz.

Table 3 shows that the gear broken tooth failure is consistent with the forecasting trend of the (P-S) KELM method, and the errors are within a very small range. The mean square error (MSE) is 0.0096, the average absolute error (MAE) is 0.0176, and the accuracy (R^2) reaches 91.83%. Regarding the predicted values by the KELM model, the mean square error is 0.0091, the average absolute error reaches 0.0253, and the accuracy is 90.04%.

In summary, the following conclusions can be drawn: the feature parameters based on the MCKD-VMD method are obtained as the sensitive feature parameters of the (P-S) KELM predicting model; the model can greatly improve the predicting accuracy, and its accuracy is much higher than the predicting result by the traditional KELM method. It provides a key technology for early warning and judgment of the gear failure because it avoids economic loss and safety hazards.

6. Conclusion

- (1) The method of acquiring multiple feature parameters based on MCKD-VMD can effectively suppress the interference noise, enhance the impact component of the fault signals, and overcome the modal aliasing and end effects. Thus, the effective sensitive feature parameters are obtained.

- (2) The ELM model based on polynomial and sigmoid kernel functions merge multiple parameters together to solve the multidimensional number problem during high-dimensional feature space operations. The problem that the predicting results of the traditional ELM algorithm models are greatly influenced by the input parameter is solved. As a result, the stability of the learning machine model is improved, and the predicting accuracy of the gear failure trend is ensured.
- (3) The decomposed multicharacteristic parameters are input to predict the life of the gear by the simulation analysis and the gear broken tooth failure experiment. At the same time, three evaluation indicators of mean square error, average absolute error, and accuracy are used to evaluate the predicting results. The results show that the predicting results of this method are consistent with the trend of the actual failure status. This consistency guarantees the validity and accuracy of the input parameters of the prediction model. The accuracy of the failure prediction is greatly improved. Thus, the key technologies for the safety and maintenance of the large-scale engineering equipment are provided.

The proposed algorithm is mainly verified to predict the broken tooth fault trend of the gear transmission system. No experimental verifications for other types of faults are conducted due to factors, such as time and equipment. The wide applicability of the algorithm should be further verified in the future.

Abbreviations

ELM:	Extreme learning machine
EEMD:	Ensemble empirical mode decomposition
SVM:	Support vector machine
KELM:	Kernel extreme learning machine
IMF:	Intrinsic mode function
MCKD:	Maximum correlated kurtosis deconvolution
VMD:	Variational mode decomposition
RBF:	Radial basis kernel function
MSE:	Mean square error
MAE:	Mean absolute error.

Data Availability

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

Conflicts of Interest

The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 11872041, No. 12032017).

References

- [1] G. Huang, G. B. Huang, S. Song, and K. You, "Trends in extreme learning machines: a review," *Neural Networks*, vol. 61, no. 61, pp. 32–48, 2015.
- [2] H. Chen, J. Peng, Y. Zhou, L. Li, and Z. Pan, "Extreme learning machine for ranking: generalization analysis and applications," *Neural Networks*, vol. 53, no. 5, pp. 119–126, 2014.
- [3] Z. Bai, G. B. Huang, D. Wang, H. Wang, and M. B. Westover, "Sparse extreme learning machine for classification," *IEEE Transactions on Cybernetics*, vol. 44, no. 10, pp. 1858–1870, 2014.
- [4] Weichao, *Research on Gear Fault Diagnosis Technology Based on SVD and ELM*, Shenyang: Shenyang Ligong University, Shenyang, China, 2017.
- [5] Yang Lu and C. Wen, "Fault diagnosis of the wind turbine gearbox based on SAPSO-ELM method," *Electro Mechanical Information*, vol. 38, no. 30, pp. 73–75, 2017.
- [6] B. Qin, L. Y. Zhang, and W. J. Sun GD, "The application of permutation entropy and nuclear extreme learning machine in gear fault diagnosis," *China Test*, vol. 43, no. 07, pp. 108–111, 2017.
- [7] J. Z. Zhou, J. Xiao, H. Xiao, W. Zhang, W. Zhu, and C. Li, "Multifault diagnosis for rolling element bearings based on intrinsic mode permutation entropy and Ensemble optimal extreme learning machine," *Advances in Mechanical Engineering*, vol. 6, Article ID 803919, 2014.
- [8] N. Rodriguez, E. Cabrera, G. Cabrera, and C. Lagos, "Stationary wavelet singular entropy and kernel extreme learning for bearing multi-fault diagnosis," *Entropy*, vol. 19, no. 10, 541 pages, 2017.
- [9] S. J. Zhou, *Research on Gear Fault Diagnosis Based on Nuclear Extreme Learning Machine*, Wuhan: Wuhan University of Technology, Wuhan, China, 2018.
- [10] B. Wan, "CuiBZh, research on fault diagnosis of the gearbox based on CEEMD-MPE and ELM method," *Modular Machine Tool and Automatic Manufacturing Technology*, vol. 103, no. 04, pp. 103–106, 2019.
- [11] X. Wang and D. J. Wang, "Bearing life prediction based on the variational modal decomposition and the extreme learning-machine method," *Manufacturing Automation*, vol. 40, no. 11, pp. 36–39, 2019.
- [12] W. X. Gu, X. F. Wang, and J. N. Xu, "Fault diagnosis of Fan Gearbox bearing based on DPSO-MKELM algorithm," *Shandong Electric Power Technology*, vol. 6, no. 49, pp. 63–67, 2019.
- [13] X. Y. Liu, X. J. W. Huang, B. Zhao, and J. Tong, "A Personalized diagnosis method to detect faults in gears using numerical simulation and extreme learning machine," *Knowledge-Based Systems*, vol. 195, 2020.
- [14] S. H. Liu, C. H. L. Liu, and H. Q. Zeng, "Research on fault warning of Wind turbine gearbox based on nuclear extreme learning machine," *China Test*, vol. 45, no. 2, pp. 121–127, 2019.
- [15] Z. Chen, K. Gryllias, and W. H. Li, "Mechanical Fault diagnosis using convolutional neural networks and extreme learning machine," *Mechanical Systems and Signal Processing*, vol. 133, Article ID 106272, 21 pages, 2019.
- [16] C. H. L. Liu, Y. J. Wu, and Z. ChG, "Fault diagnosis of rolling bearing based on variational mode decomposition and fuzzy C-means clustering," *Proceedings of the Chinese Society for Electrical Engineering*, vol. 35, pp. 3558–3565, 2015.
- [17] J. Xia, L. Zhao, Y. Bai, M. Yu, and Z. Wang, "Weak fault feature extraction of rolling bearings based on the MCKD and

- VMD method,” *Vibration and Shock*, vol. 36, no. 20, pp. 78–83, 2017.
- [18] H. M. Zhang, *Research on Weak Fault Diagnosis Method of Rolling Bearing Based on MCKD and Extreme Learning Machine*, pp. 1–57, Kunming University of Science and Technology, Kunming, 2019.
- [19] Z. Ma, Y. Li, Z. Liu, and C. Gu, “Fault feature extraction of rolling bearing based on variational mode decomposition and Teager Energy Operator,” *Journal of Vibration and Shock*, vol. 35, no. 13, pp. 134–139, 2016.
- [20] K. Li, L. Su, J. Wu, H. Wang, and P. Chen, “A rolling bearing fault diagnosis method based on variational mode decomposition and an improved kernel extreme learning machine,” *Applied Sciences*, vol. 7, no. 10, 1004 pages, 2017.