Rolling Bearing Life Prediction Technology Based on Feature Screening and LSTM Model

Yujun Zhao
School of Intelligent Manufacturing, Nanyang Institute of Technology, Nanyang, 473000, China

***Abstract*—As one of the important components of industrial equipment, the health condition of rolling bearings will directly affect the operational effectiveness of the equipment. Therefore, to ensure equipment safety and reduce maintenance costs, an intelligent rolling bearing life prediction technology is proposed. Firstly, it extracts the fault information of rolling bearings and introduces Fisher score for feature selection. Simultaneously, a variational modal analysis method on the grounds of improved particle swarm optimization is introduced to achieve denoising of rolling bearing signals. Finally, an improved bidirectional long short-term model is introduced to construct a prediction model and achieve the life prediction of rolling bearings. In the performance analysis of the denoising model, the optimal modal component K value of the denoising model was obtained through experimental analysis as 3, and the optimal penalty factor number was 1000. In the time-domain signal analysis of the two models, the proposed model possesses a more excellent decomposition effect on the original signal compared to the comparative model, and the signal denoising ability is improved by 26.35%. In the prediction of rolling bearing life, the proposed model can accurately predict the early and late life of rolling bearings. For example, when the collection time is 100, the actual remaining life is 0.712, and the proposed model is 0.721, which is better than other models. In the comparison of average absolute error, the proposed model is 0.035, which outperforms other models. This indicates that the proposed rolling bearing life prediction model has excellent predictive performance. The research provides essential technical references for the maintenance of industrial machinery and equipment, as well as equipment life monitoring.**

***Keywords*—Features; rolling bearings; prediction; fisher score; bidirectional long short term model**

I. Introduction

As an important component of industrial equipment, rolling bearings (RB) are widely used in various mechanical equipment, such as automobiles, airplanes, trains, motors, etc. Its function is to provide support and rotation between the shaft and bearing seat, so the operating status of RBs will directly affect the efficiency and service life of the equipment [1]. However, the lifespan of RBs is influenced by various factors, like load, speed, temperature, lubrication status, etc. This can lead to premature wear, failure, and even damage of RBs, causing serious consequences to the equipment, increasing maintenance costs and production losses [2]. At present, commonly used RB life prediction techniques mainly include vibration analysis, acoustic diagnosis, temperature detection, etc. Vibration analysis is one of the most commonly used methods for detecting rolling bearing faults [3]. By measuring the vibration signal generated by rolling bearings to determine whether there is a fault, in actual diagnosis, the vibration signal is affected by environmental noise interference. In addition, the cost of collecting and analyzing vibration signals is high, and both need to be addressed [4]. Acoustic diagnosis is a method of diagnosing faults by analyzing the sound signals generated by rolling bearings. In practical applications, this method is susceptible to environmental noise and interference from complex working environments [5]. In order to solve the problem of insufficient diagnosis of traditional rolling bearings, a denoising technique was proposed and a model was established. Meanwhile, an intelligent RB life prediction technology was proposed. The technology proposed by the research institute has good resistance to noise and environmental interference, and is more stable than traditional technologies. This technology utilizes machine learning and data analysis methods for predicting the lifespan of RBs on the grounds of a large amount of real-time operating data. The innovation of the research lies in the introduction of a fusion improved feature selection and denoising technology, which significantly improves the training effect of the model on RBs. Secondly, an optimized prediction model is proposed for enhancing the accuracy and stability of the model by training the selected feature data. The research has two significance. Firstly, intelligent RB life prediction technology can detect faults and problems of RBs in advance, reduce equipment maintenance costs and production losses; On the other hand, this technology can provide reference for the design and manufacturing of RBs, optimize the structure and materials of bearings, and improve the service life and performance of bearings.

The research content consists of six sections. Section I is the introduction. Related work is given in Section II. Algorithm model is discussed in Section IV. Discussion is given in Section V. Section VI summarizes the entire text, and elaborates on the improvement direction.

II. Related Work

Life prediction can help people have a clearer understanding of the operating status of mechanical equipment and improve its working efficiency. In recent years, artificial intelligence has experienced rapid development, and deep learning based life prediction of RBs has become a research hotspot. Cheng et al. found that nearly half of motor failures are caused by the degradation of rolling element bearings. So, a new data-driven framework was proposed for bearing life prediction. Firstly, the original vibration of the training bearing is processed using the Hilbert Huang transform. Then, it uses convolutional models to identify feature data and

achieve degradation estimation of the test bearings. Finally, the effectiveness of the method was verified through specific experiments, and it is superior to similar methods and has good application effects [6]. Liu H et al. proposed a new end-to-end residual service life prediction method on the grounds of feature attention. This study directly applies feature attention mechanism to input data, dynamically assigning greater attention weights to more important features, thereby improving predictive performance. Next, the study employs bidirectional gated recursive units for feature association learning. The experiment showcases that the proposed method possesses more advantages compared to similar technologies [7]. Liu Y Q et al. found that smooth wear is often overlooked, but it is one of the main causes of bearing failure, especially in locomotives. So, a locomotive track model with traction power transmission was adopted to study motor wear from the perspective of contact stress and relative slip at the raceway interface. The outcomes demonstrate that the wear of the inner ring of the motor bearing is relatively uniform. In addition, the increase in surface wear of the motor will increase internal power and worsen the vibration of the traction motor and its adjacent components in the locomotive [8].

In recent years, artificial intelligence has experienced rapid development, and deep learning based life prediction of RBs has become a research hotspot. Qin Y et al. proposed a new gated dual attention unit model for forecasting the remaining service life of RBs. Firstly, a series of root mean square values are calculated as health indicator vectors on the grounds of the full life vibration data of RBs. Then, it uses health indicators to predict lifespan. Through experimental analysis, it is shown that the proposed technology can accurately predict the health status of bearings, providing important suggestions and references for the maintenance and use of industrial bearings [9]. Althubaiti A et al. proposed an intelligent bearing prediction model. This model uses convolutional models to extract motor fault feature data, and uses a degradation index for label training. By training labels through the model, the bearing life prediction is achieved. Through relevant experimental analysis, it has been shown that the proposed technology can accurately predict bearing life and performs better than similar technologies [10]. Hu R et al. proposed a data-driven bearing life diagnosis technology, which uses a guided deep subdomain adaptive network for data feature collection, and then introduces a transfer learning model for parameter feature training. Finally, it selects four mainstream technologies for comparison. The experiment shows that the proposed technology possesses excellent application effects in practical scenarios and is superior to other life prediction models [11]. Wang B et al. found that the accuracy and generalization ability of bearing prediction technology are affected due to the lack of clear learning mechanisms. Therefore, they proposed a new multi-scale convolutional attention network framework. Firstly, it constructs a self attention module for fusing multi-sensor data, and then adopts a multi-scale learning strategy to learn representations at different time scales. Comparing the proposed technology with relevant technologies, the proposed technology has excellent application effects [12].

In summary, the above research analyzed the commonly used life prediction techniques and explored their application effects. Meanwhile, the current popular intelligent bearing life prediction technology is analyzed, and the above technologies have excellent application effects in practical scenarios. However, the above research techniques have not fully utilized the fault characteristics of bearings, so an intelligent bearing life prediction technology is proposed to meet the requirements of industrial manufacturing development.

III. Construction of a Life Prediction Model for RBs

This section mainly extracts and screens the time-domain features of RBs, and proposes an improved signal denoising technique to address the problem of characteristic signal noise points. Finally, it constructs an improved RB life prediction technology and establishes relevant models.

A. Extraction and Screening of RB Features

RBs are one of the important components in industrial manufacturing, with a large number of applications in automotive engines, motors, and industrial machinery. RBs use rolling elements to roll between the inner and outer rings, reducing friction and sliding on the contact surface, thereby achieving smooth rotational motion [13]. The relevant diagram of the RB is indicated in Fig. 1.

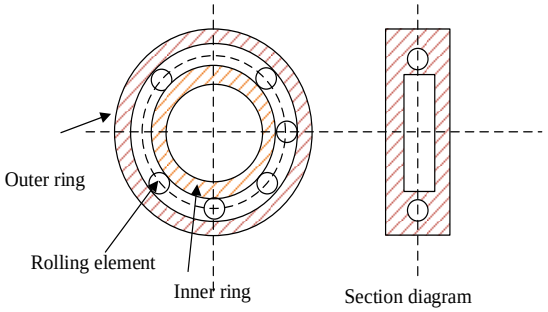

Fig. 1. Schematic diagram of RBs.

The service life of RBs will continuously decrease during long-term operation, and effective prediction of the service life of RBs is the key to ensuring the stable use of mechanical equipment. Therefore, a method for predicting the lifespan of RBs is proposed, which first extracts the time-domain feature information of RB features [14]. Time domain features are a series of indicators used to describe the characteristics of a signal in the time domain, mainly including dimensional features and non-dimensional features. These features include the mean of the bearing signal, the fluctuation amplitude of the signal, the vibration amplitude of the signal, and other characteristics, which can effectively reflect the operating status of the bearing [15]. It assumes the vibration signal $x(t)$ of a RB, where t represents time. The goal of the study is to extract some key time-domain features from these signals to reflect the state of the bearings. Firstly, the study can calculate the mean μ of the signal, which represents the concentrated trend of the signal, as shown in Eq. (1).

$$\mu = \frac{1}{T}\int_0^T x(t)dt \tag{1}$$

Among them, T is the duration of the signal. Next, the study can calculate the variance σ^2 of the signal, which represents the degree of signal dispersion, as shown in Eq. (2).

$$\sigma^2 = \frac{1}{T}\int_0^T (x(t)-\mu)^2 dt \tag{2}$$

The larger the variance, the greater the degree of dispersion of the signal. In addition to mean and variance, this study can also calculate other time-domain features such as Peak Value, Peak-to-Peak Value, Skewness, and Kurtosis. The peak represents the maximum or minimum value of the signal, as shown in Eq. (3).

$$PeakValue = \max(x(t)) - \min(x(t)) \tag{3}$$

The Peak-to-Peak Value represents the amplitude range of signal vibration, as shown in Eq. (4).

$$PeakToPeakValue = \max(x(t)) - \min(x(t)) \tag{4}$$

Skewness measures the asymmetry of signal distribution, as shown in Eq. (5).

$$Skewness = \frac{1}{T}\int_0^T \left(\frac{x(t)-\mu}{\sigma}\right)^3 dt \tag{5}$$

By calculating these time-domain features, research can extract information related to the bearing state from the signal. However, bearings work in harsh environments for a long time, and the environment is complex and variable, resulting in non-stationary characteristics in the signals monitored by sensors, which affects the effectiveness of feature extraction [16]. There are many commonly used feature selection methods, including variance selection and chi square test, to further screen the extracted features. This study adopted the Fisher score feature selection method mainly because it can handle non-linear features and has better feature processing effects [17]. The given feature set is $\{f_1, f_2, \ldots, f_n\}$, with c class feature samples $x_j \in R^m$, where $j = 1, 2, \ldots, N$. If the class distance between the i features of the training sample is defined as $S_b(f_i)$, and the class distance between the k-class sample and the i-th feature f_i is defined as $S_\omega^{(k)}(f_i)$, then $S_b(f_i)$ is expressed as Eq. (6).

$$S_b(f_i) = \sum_{k=1}^{c} n_k (m_i^{(k)}) - m_i) \tag{6}$$

In Eq. (6), n_k represents the number of k-class samples, $m_i^{(k)}$ represents the mean of k-class samples under the i-th feature, and m_i is the mean under the i-th feature. The expression of $S_\omega^{(k)}(f_i)$ is shown in Eq. (7).

$$S_\omega^{(k)}(f_i) = \sum_{j=1}^{n_k} (x_{ij}^{(k)} - m_i^{(k)})^2 \tag{7}$$

In Eq. (7), $x_{ij}^{(k)}$ serves as the value of the j-th sample in i features f_i of the k-th class of samples. When the value of class distance $S_b(f_i)$ is larger, the better the distance $S_\omega^{(k)}(f_i)$ between c classes, and the i-th feature value of the training sample can be obtained, as shown in Eq. (8).

$$F(f_i) = \frac{S_b(f_i)}{\sum_{k=1}^{c} S_\omega^{(k)}(f_i)} = \frac{\sum_{k=1}^{c} n_k (m_i^{(k)} - m_i)}{\sum_{k=1}^{c}\sum_{j=1}^{n_k} (x_{ij}^{(k)} - m_i^{(k)})^2} \tag{8}$$

By using the above methods, the identification of imbalanced feature categories can be achieved. On the grounds of the differences in degraded features of the same class, redundant features can be removed to filter out the main features.

B. Construction of Denoising Model on the Grounds of Improved VMD

Predicting the service life of RBs is beneficial for the stable operation of equipment and improving its service life. In the prediction of rolling shaft bearings, it is necessary to evaluate the bearing life by extracting bearing vibration signals and characterizing degradation features. However, the original signal features contain a lot of noise, so a Variational Mode Decomposition (VMD) model is used for denoising [18]. VMD is an adaptive signal processing method with advantages such as high accuracy and strong noise resistance. Among them, the Intrinsic Mode Function (IMF) is represented as an unstable amplitude modulated frequency modulation signal, as shown in Eq. (9).

$$\begin{cases} u_k(t) = A_k(t)\cos(\phi_k(t)) \\ \omega_k(k) = \phi'(t) = \dfrac{d\phi_k(t)}{d(t)} \end{cases} \tag{9}$$

In Eq. (9), $u_k(t)$ represents the harmonic signal, $A_k(t)$ represents the amplitude of $u_k(t)$, and $\omega_k(k)$ represents the frequency of $u_k(t)$. There are two parts in the construction of VMD model, including variational model and variational model solving. When constructing a variational model, it is necessary to convert $u_k(t)$ into an explanatory variable, as shown in Eq. (10).

$$u_k(t) = (\delta(t) + \frac{j}{\pi t} \times u_k(t)) \tag{10}$$

In Eq. (10), $\delta(t)$ is the analytical component, and the modal component bandwidth is obtained by the time gradient L2 norm of the component, as showcased in Eq. (11).

$$\| \partial_t \left[\delta(t) + \frac{j}{\pi t} \times u_k(t) \right] e^{-j\omega_k t} \|_2^2 \tag{11}$$

In Eq. (11), $e^{-j\omega_k t}$ is the exponential signal and ∂_t is the derivative symbol. The penalty factor α and the guaranteed constraint multiplication operator $\lambda(t)$ are introduced to solve the variational model, as shown in Eq. (12).

$$L(u_k,\omega_k,\lambda)=\alpha\sum_k\left\|\partial_t\left[\delta(t)+\frac{j}{\pi t}\right]e^{-j\omega_k t}\right\|_2^2+\left\|f(t)-\sum_k u_k(t)\right\|_2^2+\langle\lambda(t),\ f(t)-\sum_k u_k(t)\rangle \quad (12)$$

In Eq. (12), k represents the k-th modal component, where, $f(t)$ is the signal target. When VMD denoises signals, the number of decompositions K and the penalty coefficient have a significant impact on the signal decomposition effect. Both large or small values can easily filter important fault information. Therefore, to better denoise bearing signals in VMD, an improved Self-adaptation Particle Swarm Optimization (SPSO) algorithm is introduced for optimizing the problem [19]. The process of constructing the SPSO model is shown in Fig. 2.

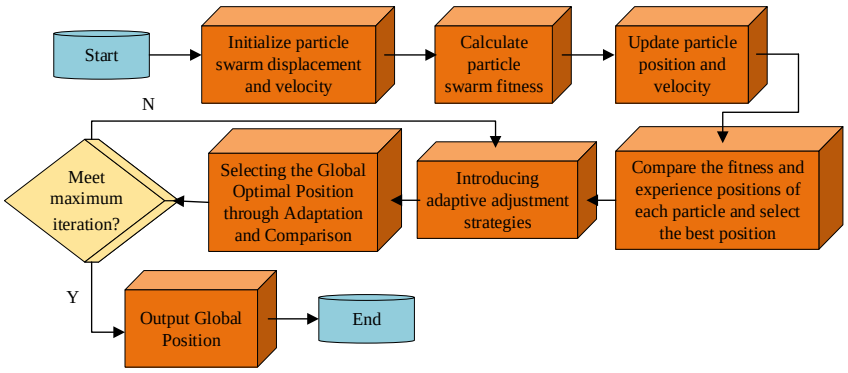

Fig. 2. Schematic diagram of SPSO model process.

The SPSO model adaptively updates the weights and particle velocities of the original particle swarm model, improving the overall training accuracy of the model. To optimize the main information loss problem of VMD signal denoising, the weighted average of bearing signal differences is used as the function optimization objective. The signal difference method is used to calculate the difference between signal data. The smaller the signal difference, the higher the signal similarity, thus screening useful information. The expression of signal difference is showcased in Eq. (13).

$$SDA=\frac{1}{N}\sum_{i=1}^{n}(y_{IMFS}-y_s) \quad (13)$$

In Eq. (13), y_{IMFS} is the sum of the K signal components, y_s represents the original signal, and the final signal average is obtained through weighted processing, as shown in Eq. (14).

$$WSDA=SDA+\frac{K\times SDA}{\beta} \quad (14)$$

In Eq. (14), β represents the weighted penalty parameter. The entire bearing signal denoising process is shown in Fig. 3.

C. Construction of a RB Life Prediction Model on the Grounds of Improved LSTM

To effectively predict the service life of RBs, a Bi-directional Long Short-Term Memory (BiLSTM) model is introduced on the grounds of the extracted RB feature data to construct a prediction model. Compared to traditional one-way LSTM models, it can capture more feature details and has better feature expression ability. Therefore, the extracted feature parameters of RBs are used as inputs for BiLSTM, and the life prediction of RBs is achieved through training. The structural principle of the BiLSTM model is shown in Fig. 4.

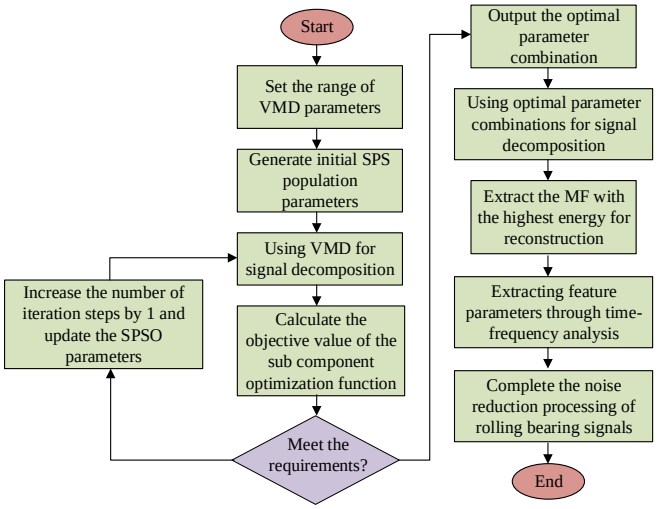

Fig. 3. Bearing signal denoising process on the grounds of SPSMO-VMD

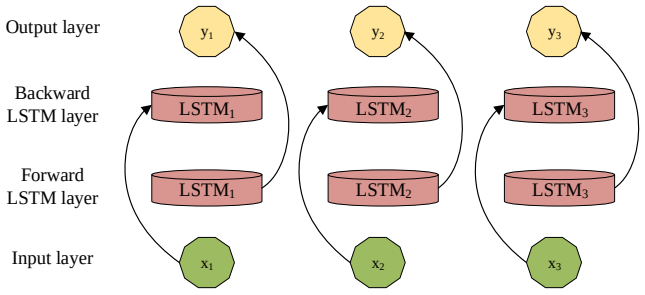

Fig. 4. Schematic diagram of BiLSTM structure.

The BiLSTM model draws on the characteristics of bidirectional recurrent networks, and can effectively learn time series features through the LSTM model. In model construction, it mainly consists of a forward LSTM and a backward LSTM. The outputs of the two LSTM models are weighted, and the model training prediction results are obtained by learning past information from the data. The output expression of the forward LSTM model is shown in Eq. (15).

$$\overrightarrow{H_t} = \overrightarrow{LSTM}(h_{t-1}, x_t, c_{t-1})\, t_r \in [1, T_r] \tag{15}$$

In Eq. (15), h_{t-1} represents the output at the previous time point, x_t represents the direct input, c_{t-1} represents the cell state at the previous time point, and $t_r \in [1, T_r]$ represents the time interval. The backward LSTM output is shown in Eq. (16).

$$\overleftarrow{H_t} = \overleftarrow{LTM}(h_{t+1}, x_t, c_{t+1}) t \in [1, T] \tag{16}$$

According to Eq. (16) and Eq. (15), the updated expression of the BiLSTM model can be obtained, as shown in Eq. (17).

$$H_t = \left[\overrightarrow{h_t}, \overleftarrow{h_t}\right] \tag{17}$$

In Eq. (17), $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ represent forward and backward LSTM outputs, respectively. In actual model training, BiLSTM is also prone to parameterization problems due to structural feature factors. Therefore, the study introduces the Adam algorithm to optimize the BiLSTM prediction model. The Adam algorithm has advantages such as high computational efficiency and easy implementation, and can obtain the optimal solution of the model in a short period of time. Parameter optimization is shown in Eq. (18).

$$\delta^* = arm \min_{\delta} S(f(x_r, \delta)) \tag{18}$$

In Eq. (18), δ^* represents the optimized BiLSTM model parameters, δ represents the current model parameters, x_r represents the input parameters, $S(\cdot)$ represents the objective function, and $f(\cdot)$ represents the network output. Then it initializes the first-order moment g_t and the second-order moment h_t at time t, and obtains the result as shown in Eq. (19).

$$\begin{cases} g_t = \mu * g_{t-1} + (1-\mu) * m_t \\ h_t = v * n_{t-1} + (1-\mu) * m_t^2 \end{cases} \tag{19}$$

In Eq. (19), μ is a value of 0.9, v defaults to a value of 0.999, and m_t represents the gradient at time t. The expression of parameter δ is shown in Eq. (20).

$$\delta = \delta - \eta = \frac{g_t}{\sqrt{\hat{h}_t} + \varepsilon} \tag{20}$$

In Eq. (20), η represents the learning rate, ε ranges from 10 to 8, and $\hat{h}_t$ is expressed as shown in Eq. (21).

$$\hat{h}_t = \frac{h_t}{(1-v)} \tag{21}$$

Through the above research, the initial parameter optimization of the BiLSTM model can be achieved. Meanwhile, it is necessary to consider the problem of too many network parameters in the research. Model training with too many parameters can easily fall into overfitting problems, which can affect the actual prediction performance of the model. In this regard, the Dropout mechanism is introduced for solving the model overfitting. By performing the above operations, the problem of overfitting in the model can be solved, thereby constructing a RB prediction model on the grounds of improved BiLSTM [20]. The life prediction of the entire RB is showcased in Fig. 5.

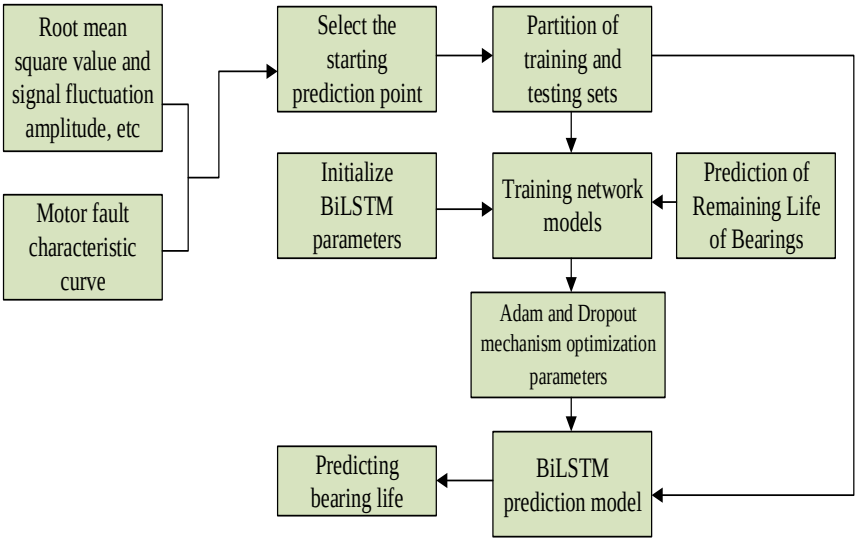

Fig. 5. Residual life prediction process for RBs.

IV. Algorithm Model Simulation Testing

This section mainly conducts simulation performance analysis on the proposed denoising model and RB life prediction model. This includes parameter optimization testing of VMD models, performance comparison analysis of denoising models, and life prediction analysis of RBs.

A. Simulation Analysis of Denoising Model Performance

To verify the proposed residual life prediction technology for RBs, experimental testing will be conducted on the WINDOWS 10 64 bit platform. The processor is INTEL i7, with a running memory of 64GB, and simulation experiment analysis is completed on the Matlab platform. The initialization parameters of the experimental model are showcased in Table I.

In the performance analysis of denoising models, the SPSO algorithm is used to optimize the parameter combination of the VMD model. Meanwhile, Empirical Mode Decomposition (EMD) is introduced as the testing benchmark. In the experiment, the original signal was decomposed at 1KHz. The original time-domain waveform signal is shown in Fig. 6.

TABLE I. Model Initial Parameters

Parameter indicator type	Numerical value
Sampling rate	1HKz
Number of input nodes	3
Number of output layer nodes	1
Number of hidden layer nodes	256
Prediction step size	40
Learning rate	0.01

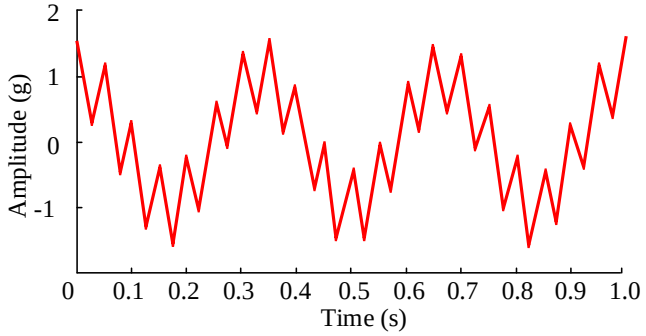

Fig. 6. Time domain waveform of the original signal.

In actual signal filtering and decomposition, the number of modal components K will have a direct impact on the decomposition effect of the VMD model. Setting the K value too small can lead to incomplete signal decomposition, while setting the K value too large can cause partial information loss in the RB. Therefore, the decomposition effect of VMD under different K values will be analyzed, as shown in Fig. 7.

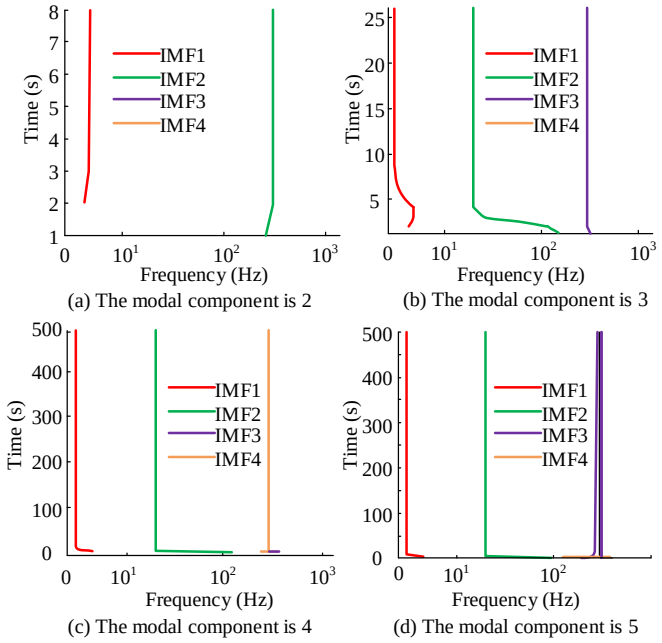

Fig. 7. Center frequency plots of each component under different modes K.

Fig. 7 (a) shows the center frequency maps of each component when K value is 2. It is evident that the cosine signal with a center frequency of 3Hz and the cosine signal with a center frequency of 300Hz can be seen, but the decomposition is not complete, and mode mixing phenomenon occurs. Fig. 7 (b) shows the center frequency plots of each component at K=3. It can be clearly seen that the center frequencies of the IMF can be effectively separated, and there is no overlap of the four IMF center frequencies. The overall decomposition effect is excellent. Fig. 7 (c) shows the center frequency plots of each component at K=4. According to the data results, IMF1 represents a 3Hz cosine signal, while IMF3 represents a 20Hz cosine signal. However, during the actual decomposition process, excess IMF2 appeared, and the center frequencies between IMF2 and IMF3 overlapped. Fig. 7 (d) shows the center frequency maps of each component at K=5, which is consistent with Fig. 7 (c). During the decomposition process, there were excess IMF3 signals and IMF4. In addition, the value of the penalty factor α will also affect the decomposition effect of VMD, as shown in Fig. 8.

Fig. 8 (a) to 8 (d) were tested with penalty factors of 125, 250, 500, and 1000, respectively. When the penalty factor is 125, there is aliasing between IMF2 and IMF3, resulting in VMD being unable to effectively decompose the original signal. When the penalty factor is 250, both IMF2 and IMF3 exhibit aliasing phenomenon. When dividing the penalty factors into 500 and 1000, as the quantity of penalty factors grows, the problem of overlapping center frequencies reduces, but it will increase the difficulty of VMD operation. When the penalty coefficient is 500, the center frequency overlap problem has been significantly improved. Considering the model decomposition effect and computational time, the penalty factor will be set to 1000 in subsequent experiments, and the modal component K value will be set to 3. It selects EMD for signal decomposition comparison, as shown in Fig. 9.

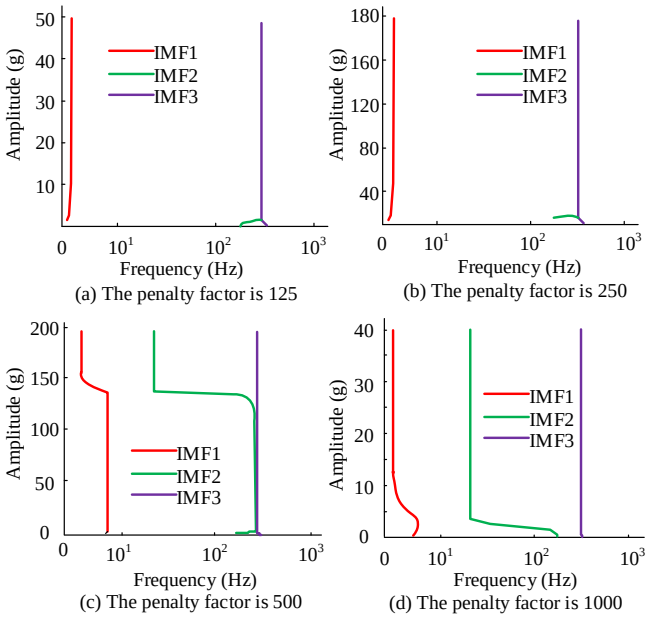

Fig. 8. VMD decomposition center frequency results under different penalty factors.

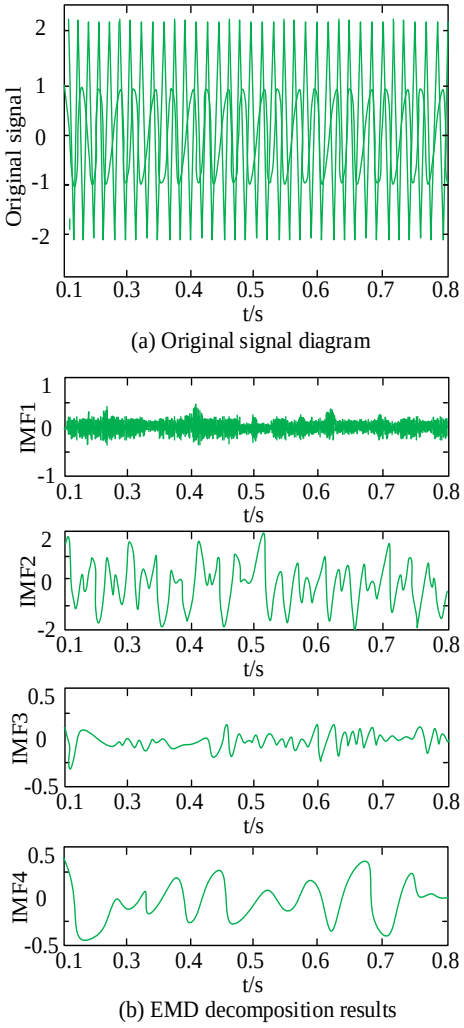

Fig. 9. Time domain diagram of EMD decomposition simulation signal.

Fig. 9 shows the time-domain simulation results of EMD decomposition. Among them, Fig. 9 (a) is the original signal graph, and Fig. 9 (b) is the EMD decomposition result. According to the data results, setting the number of modal layers to 4 resulted in four types of mixed signals. In IMF1 signal analysis, obvious fluctuations can be seen from the image. As the sampling time increases, there are still significant fluctuations in the waveform and mode aliasing phenomenon exists. In IMF2, IMF3, and IMF4, as the sampling time increases, there are still significant fluctuations in the three components of the signal, which cannot maintain the original waveform of the signal. In addition, incomplete decomposition and mode aliasing occur during the decomposition process. It uses VMD to analyze the original signal, as shown in Fig. 10.

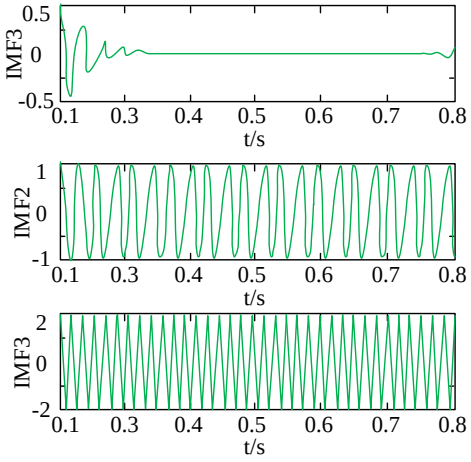

Fig. 10. Time domain diagram of VMD decomposition simulation signal.

Fig. 10 shows the time-domain results of VMD decomposition simulation signals. Due to setting the modal component K to 3, three modulus decomposition results were obtained, corresponding to IMF1, IMF2, and IMF3, respectively. Compared with the original image 9 (a), it can be found that VMD can accurately decompose the original signal, and there is no aliasing problem during the decomposition process. It can maintain the original signal waveform, and the decomposed component waveform is basically consistent with the original time-domain waveform. Therefore, it can be concluded that VMD has excellent decomposition performance in the original time-domain signal decomposition, and performs better in signal decomposition compared to the EMD model, with a 26.35% improvement in signal denoising ability.

B. Simulation Performance Analysis of RB Life Prediction

In the prediction of the service life of RBs, a total of 1356 samples will be selected from the data collected and feature screened during the experimental process. The samples include early degradation data of RBs, mid-term degradation data, and severe degradation data, accounting for 40%, 30%, and 30%, respectively. In the sample data, radial force will have a negative impact on the vibration of RBs, so the lateral vibration signal is selected for reflecting the degradation phenomenon of RBs. In the life cycle monitoring of RBs, the characteristic changes in the early stage are relatively stable,

and there will be obvious fluctuations in the later stage, which will have a significant impact on the life prediction of different models. Therefore, the Particle Swarm Optimization (PSO) optimized VMD combined Long Short Term Memory (LSTM) model, PSO optimized VMD combined BiLSTM model, PSO optimized VMD combined Recurrent Neural Network (RNN) model, and the proposed SPSO-VMD BiLSTM model were selected for comparison. The predicted life of the roller bearing is shown in Fig. 11.

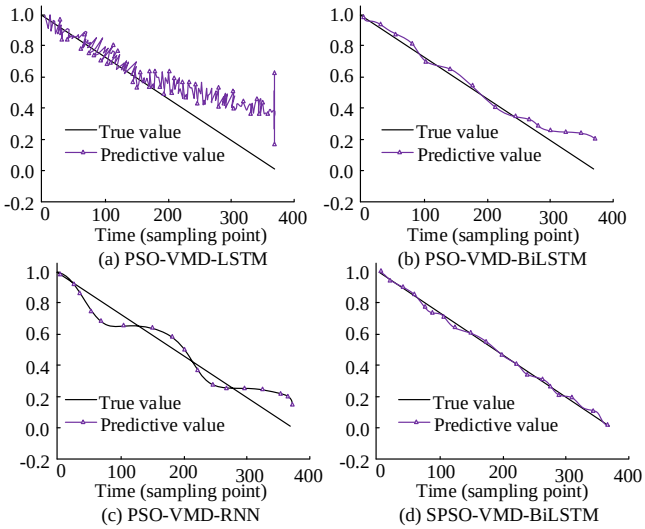

Fig. 11. Comparison of life prediction for RBs using multiple models.

Fig. 11(a) to 11(b) show PSO-VMD-LSTM, PSO-VMD-BiLSTM, PSO-VMD-RNN, and SPSO-VMD-BiLSTM, respectively. In the comparison of the four models, it can be found that except for the PSO-VMD-RNN model, the other three models have high prediction accuracy in the early stages. When the sampling time is 100, the actual remaining lifespan is 0.712. PSO-VMD-LSTM predicts a remaining lifespan of 0.735, PSO-VMD-BiLSTM predicts a remaining lifespan of 0.731, and PSO-VMD-RNN predicts a remaining lifespan of 0.698. The proposed model possesses a prediction accuracy of 0.721, and the overall prediction accuracy of the proposed model is the best. In the later stage, the problem of roller bearing failures increases, and only SPSO-VMD-BiLSTM can accurately predict the model life. Compared with PSO-VMD-RNN, PSO-VMD-LSTM, and PSO-VMD-BiLSTM, the prediction accuracy of SPSO-VMD-BiLSTM increased by 35.65%, 18.65%, and 12.35%, respectively. It introduces Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for comparing the predictive performance of different models, as shown in Fig. 12.

Fig. 12(a) and 12(b) show the comparison results between RMSE and MAE, respectively. In RMSE error comparison, the RMSE errors of SPSO-VMD-BiLSTM, PSO-VMD-LSTM, and PSO-VMD-BiLSTM when they tend to converge are 0.031, 0.062, 0.078, and 0.083, respectively. In MAE error comparison, the error values of SPSO-VMD-BiLSTM, PSO-VMD-LSTM, and PSO-VMD-BiLSTM towards convergence are 0.035, 0.051, 0.071, and 0.078. It demonstrates that the proposed model possesses lower errors and more excellent prediction performance in the prediction of roller bearings. Finally, a comprehensive performance comparison is conducted, as shown in Table II.

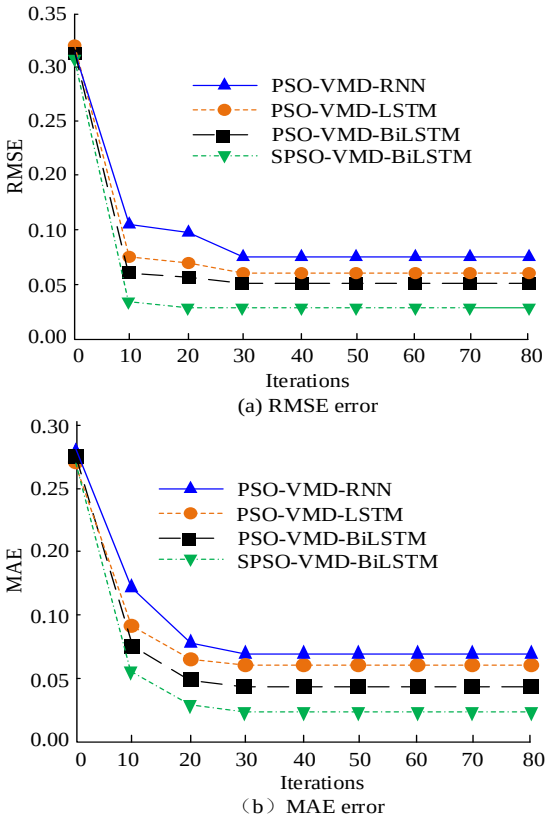

Fig. 12. Comparison of errors among different models.

TABLE II. COMPREHENSIVE COMPARISON RESULTS OF DIFFERENT METHODS

Period	Method Type	Prediction accuracy	Time consuming（s）
Rolling bearings in the early stage	PSO-VMD-RNN	0.856	3.25
	PSO-VMD-LSTM	0.863	3.54
	PSO-VMD-BiLSTM	0.905	2.54
	SPSO-VMD-BiLSTM	0.966	2.42
Rolling bearing mid-term	PSO-VMD-RNN	0.873	3.56
	PSO-VMD-LSTM	0.886	3.45
	PSO-VMD-BiLSTM	0.925	3.05
	SPSO-VMD-BiLSTM	0.965	2.75
Late stage of rolling bearings	PSO-VMD-RNN	0.856	3.65
	PSO-VMD-LSTM	0.878	3.45
	PSO-VMD-BiLSTM	0.895	3.12
	SPSO-VMD-BiLSTM	0.956	2.89

Table II compares the life prediction results of different methods in three periods of rolling bearings. According to the test results, whether in the early, middle, or later stages, the research model still has the highest prediction accuracy, all above 0.950. In addition, the prediction time of different models was compared, and the overall time consumption of the research model was shorter, indicating that the technology proposed in the study has better performance effects.

V. Discussion

Fault diagnosis of rolling bearings is an important step in ensuring the normal operation of industrial equipment. Compared with similar technologies, the proposed research method has excellent diagnostic performance for rolling bearing faults. In addition, this study will also compare the proposed model with the techniques proposed in study [9] and [10]. Compared with the technique proposed in reference [10], this technique proposes a sparse data based rolling bearing state prediction technique, which can reduce the dependence of diagnostic models on sparse data and achieve bearing diagnosis under sparse data. Compared with study [9], due to better noise reduction processing of the data in the early stages and increased analysis of rolling bearing fault factors, the research technology for fault diagnosis is more accurate than reference [9]. In the later life prediction of rolling bearings, the predicted life of the research model is 0.256, while the study [9] is 0.268. The research model is closer to the actual results and has higher accuracy. Meanwhile, a comparison was made between the diagnostic techniques in study [10]. Reference [10] proposed a rolling bearing life prediction technique based on a deep learning framework, which achieves fault diagnosis through the fusion of multi-sensor data. The technology proposed in study [10] is highly dependent on signal data and requires a large amount of data for training during the diagnostic process, which is slower compared to the data processing of research models. In the early prediction of rolling bearings, the prediction accuracy of study [10] reached 0.956, which is close to the research model. However, in the mid to late stage of prediction, the accuracy of the research model can still reach above 0.90. Due to the complexity of data processing, the prediction accuracy in study [10] has significantly decreased.

Overall, this study proposes a deep learning based rolling bearing life prediction method, which has good application effects. Compared with current research techniques, its stability and accuracy in predictive diagnosis have shown good results. Therefore, this method has broad application prospects in the field of rolling bearing fault diagnosis.

VI. Conclusion

RBs are one of the most critical components in modern industrial fields. Due to the influence of working environment factors, effective life prediction of RBs is crucial for the safe operation of equipment. In this regard, a research proposes an intelligent RB life prediction technology, which first extracts time-domain and another feature information of the RB, and selects the main feature information through Fisher score. Simultaneously improving the SPSO model to optimize VMD and achieve denoising processing of roller bearing information. Finally, on the grounds of BiLSTM, a life prediction model for RBs is constructed for monitoring the life of RBs. In the performance experiment of the denoising model, the optimal modal component K value and penalty factor were obtained through comparison, with values of 3 and 1000, respectively. In time-domain signal decomposition testing, the improved VMD model has a better signal decomposition effect compared to the EMD model, with a 26.35% increase in signal denoising ability. In the prediction of RB life, when the sampling time is 100, the actual remaining life is 0.712. PSO-VMD-LSTM predicts the remaining life as 0.735, PSO-VMD-BiLSTM predicts the remaining life as 0.731, and PSO-VMD-RNN is 0.698. The proposed model possesses a prediction accuracy of 0.721, and the overall prediction accuracy of the proposed model is the best. In addition, the proposed model exhibits the best performance in both RMSE error and MAE error. It demonstrates that the proposed technology possesses excellent predictive performance in predicting the lifespan of RBs. The research model also has limitations. The model mainly focuses on the working conditions of RBs and does not consider other factors that may affect the lifespan of rolling bearings, such as humidity and lubrication conditions. Further analysis is needed in the future to improve the prediction accuracy of the model.

References

[1] Ren X, Wu S, Xing H, Fang X. Fracture mechanics based residual life prediction of railway heavy coupler with measured load spectrum. International Journal of Fracture, 2022, 234(1-2): 313-327.

[2] Chen Z, Wu M, Zhao R, Guretno F. Machine remaining useful life prediction via an attention-based deep learning approach. IEEE Transactions on Industrial Electronics, 2020, 68(3): 2521-2531.

[3] Ma M, Mao Z. Deep-convolution-based LSTM network for remaining useful life prediction. IEEE Transactions on Industrial Informatics, 2020, 17(3): 1658-1667.

[4] Akhenia P, Bhavsar K, Panchal J, Vakharia V. Fault severity classification of ball bearing using SinGAN and deep convolutional neural network. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2022, 236(7): 3864-3877.

[5] Li Y, Song Y, Jia L, Li Q. Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning. IEEE Transactions on Industrial Informatics, 2020, 17(4): 2833-2841.

[6] Cheng C, Ma G, Zhang Y. A deep learning-based remaining useful life prediction approach for bearings. IEEE/ASME transactions on mechatronics, 2020, 25(3): 1243-1254.

[7] Liu H, Liu Z, Jia W, Liu X. Remaining useful life prediction using a novel feature-attention-based end-to-end approach. IEEE Transactions on Industrial Informatics, 2020, 17(2): 1197-1207.

[8] Liu Y Q, Chen Z G, Wang K Y. Surface wear evolution of traction motor bearings in vibration environment of a locomotive during operation. Science China Technological Sciences, 2022, 65(4): 920-931.

[9] Qin Y, Chen D, Xiang S, Zhu C. Gated dual attention unit neural networks for remaining useful life prediction of rolling bearings. IEEE Transactions on Industrial Informatics, 2020, 17(9): 6438-6447.

[10] Althubaiti A, Elasha F, Teixeira J A. Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis–a review. Journal of Vibroengineering, 2022, 24(1): 46-74.

[11] Hu R, Zhang M, Xiang Z, et al. Guided deep subdomain adaptation network for fault diagnosis of different types of rolling bearings. Journal of Intelligent Manufacturing, 2023, 34(5): 2225-2240.

[12] Wang B, Lei Y, Li N, Wang W. Multiscale convolutional attention network for predicting remaining useful life of machinery. IEEE Transactions on Industrial Electronics, 2020, 68(8): 7496-7504.

[13] Seo Y M, Yum S K, Sung I K. Respiratory severity score with regard to birthweight during the early days of life for predicting pulmonary hypertension in preterm infants. Journal of Tropical Pediatrics, 2020, 66(6): 561-568.

[14] Finegan D P, Zhu J, Feng X. The application of data-driven methods and physics-based learning for improving battery safety. Joule, 2021, 5(2): 316-329.

[15] Vollmer I, Jenks M J F, Roelands M C P, Roelands MCP. Beyond mechanical recycling: Giving new life to plastic waste. Angewandte Chemie International Edition, 2020, 59(36): 15402-15423.

[16] Moon B, Lee J, Kim S. Methodology for predicting the durability of aged tire sidewall under actual driving conditions. International Journal of Precision Engineering and Manufacturing, 2022, 23(8): 881-893.

[17] Asyraf M R M, Ishak M R, Sapuan S M. Woods and composites cantilever beam: A comprehensive review of experimental and numerical creep methodologies. Journal of Materials Research and Technology, 2020, 9(3): 6759-6776.

[18] Asgari F, Minooei A, Abdolahi S. A new approach using Machine Learning and Deep Learning for the prediction of cancer tumor. Journal of Simulation and Analysis of Novel Technologies in Mechanical Engineering, 2021, 13(4): 41-51.

[19] Nsugbe E. Toward a Self-Supervised Architecture for Semen Quality Prediction Using Environmental and Lifestyle Factors//Artificial Intelligence and Applications. 2023, 1(1): 35-42.

[20] Cao H, Wu Y, Bao Y, Feng X, Wan S, Qian C. UTrans-Net: A Model for Short-Term Precipitation Prediction//Artificial Intelligence and Applications. 2023, 1(2): 106-113.