IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 2 January 2024, accepted 14 January 2024, date of publication 19 January 2024, date of current version 8 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3355978

RESEARCH ARTICLE

Remaining Life Prediction of Bearings Based on Improved IF-SCINet

JING ZHANG, CHAO ZHANG, SHUAI XU, GUIYI LIU, HONGBO FEI, AND LE WU

School of Mechanical Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China
Inner Mongolia Key Laboratory of Intelligent Diagnosis and Control of Mechatronic System, Baotou 014010, China

Corresponding author: Chao Zhang (zhanghero@imust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51965052 and Grant 52365014, and in part by the Central Government's Guidance in Local Science and Technology Development under Grant 2022ZY0221.

ABSTRACT In the field of health management, predicting the remaining useful life (RUL) of a device becomes critical. However, the RUL prediction process is often affected by a various of confounding factors, resulting in reduced prediction accuracy. To improve the accuracy of RUL prediction, this study first extracts the root mean square, skewness, and Kurtosis from the bearing characteristics, and adopts the multidimensional scale change features to construct a health indicator that fully reflects the bearing degradation trend. Then, a combination of the Isolation forest algorithm and the 3σ criterion was used to adaptively determine the first prediction time (FPT) of the bearings. Subsequently, the time series model SCINet is introduced for the first time into the field of bearing life prediction and used to predict the RUL of bearings. Finally, a series of multi-step prediction experiments are conducted on two publicly available datasets, PHM2012 and XJTU-SY, and compared with LSTM, GRU, Informer, and TCN models. The results show that the improved IF-SCINET has a stronger prediction capability compared to the traditional model, which significantly improves the accuracy and stability of bearing RUL prediction.

INDEX TERMS First prediction time, isolation forest, remaining life prediction, rolling bearing, sample convolutional interaction.

I. INTRODUCTION

Rolling bearings play an important and crucial role in rotating machinery. With the increasing complexity of the mechanical structure, the incidence of equipment failure has increased significantly. Once a failure occurs, it not only leads to equipment downtime and maintenance, but can also lead to more serious safety issues [1], [2]. According to relevant statistics, about 40% to 50% of equipment failures are caused by anomalies in the operation of rolling bearings [3]. Consequently, the practice of prognostics and health management (PHM) for bearings has become indispensable [4], [5]. One essential aspect of PHM is the prediction of remaining bearing life, which is crucial for ensuring equipment reliability and preventing potentially severe safety incidents [6], [7], [8], [9].

Presently, existing methodologies for Remaining Useful Life (RUL) prediction can be generally classified into two principal categories: model-based prediction and data-driven prediction [10]. Model-based approaches conventionally rely on mathematical or physical models to elucidate the process of bearing degradation [11], [12], [13], [14]. Nevertheless, in light of the escalating intricacy of contemporary equipment, attaining a precise physical model to elucidate the operational conditions of bearings is becoming progressively arduous. In contradistinction, a data-driven approach can directly harness the information garnered from sensors to prognosticate the trajectory of bearing degradation, obviating the imperative for an exhaustive understanding of the intricacies of bearing failure mechanisms [15], [16]. This makes the data-driven approach particularly suitable for complex systems, where fault conditions can be analyzed by detecting changes in measurement data. Data-driven methodologies can be further categorized into statistical data-based [17], machine learning, and deep learning paradigms [18], [19].

The associate editor coordinating the review of this manuscript and approving it for publication was Jiajie Fan.

Although statistical approaches exhibit commendable predictive performance within the confines of linearity assumptions, they tend to falter in capturing non-linear features inherent in sequential data [20]. Consequently, in recent times, data-driven predictive techniques have progressively transitioned into the domains of machine learning and deep learning. Within the machine learning domain, research predominantly revolves around methodologies such as Support Vector Machines (SVM) [21], Markov Decision Process (MDP) model [22], and Random Forests (RF) [23]. In contrast, deep learning, characterized by the utilization of intricate neural network architectures to bolster mapping capabilities, has exhibited notable advantages in terms of predictive accuracy compared to conventional machine learning methodologies. In the sphere of sequence modeling, deep neural networks encompass Long Short-Term Memory Networks (LSTMs) [24], [25], [26], Gated Recurrent Unit Networks (GRUs), Temporal Convolutional Networks (TCNs) [27], as well as Transformer-based architectures, including its variant, Informer. Recent years have witnessed the emergence of a multitude of deep learning methodologies for Remaining Useful Life (RUL) prediction, as enumerated below.

In a previous study, Yang et al. [28] constructed a generalized regression neural network (GRNN) model based on health metrics for predicting the RUL of rolling bearings. The method showed good performance in terms of prediction accuracy and reliability. Wang et al. [29] proposed a CNN-LSTM-PSO tool residual life prediction method based on multi-channel feature fusion to address the problems of weak tool wear state features, difficult extraction, and low prediction precision and accuracy. Cao et al. [30] used a parallel GRU, a comprehensive strategy of a two-stage attention mechanism and nonparametric uncertainty quantification methods to obtain accurate and reliable prediction results. Wang et al. [31] combined improved TCN and migration learning to construct a bearing life prediction model that better exploits the inherent degradation tendency of the bearings, thus effectively improving the prediction accuracy of the remaining life. In addition, Liu et al. [32] proposed a novel autonomous optimization Transformer, which combines a cyclic mechanism and a positional embedding algorithm to achieve accurate prediction of long sequences.

Nonetheless, these sequence modeling networks are not without inherent limitations. Recurrent Neural Networks (RNNs), for instance, exhibit vulnerability to issues such as error accumulation, gradient vanishing, and gradient explosion. While Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures offer partial mitigation of these concerns, they do not provide complete resolutions. Transformer and Informer models, on the other hand, typically require substantial computational resources, constraining their applicability in resource-constrained settings. Additionally, Temporal Convolutional Networks (TCNs) require raw sequences with valid historical lengths, potentially demanding increased memory resources during evaluation. During the 36th Conference on Neural Information Processing Systems (NeurIPS 2022, Liu et al. [33] introduced an inventive neural network framework known as the Sample Convolution and Interaction Network (SCINet). Its central concept revolves around the comprehensive exploitation of a unique temporal property intrinsic to time series data, specifically, the robust preservation of temporal relationships even after undergoing downsampling into two distinct subsequences. When considered from the perspective of computational cost, SCINet significantly outperforms the Transformer model in terms of efficiency. Furthermore, SCINet effectively addresses two notable limitations associated with Temporal Convolutional Networks (TCNs), namely, the absence of a requirement for output sequences of equal length as input and the prevention of inadvertent leakage of future information. SCINet surpasses traditional RNNs, Transformer models, and TCNs in the realm of time series prediction. However, it is noteworthy that, to date, SCINet has not been employed in the domain of bearing life prediction. Consequently, this study pioneers the incorporation of SCINet into the prediction of RUL for bearings.

The inherently microscopic scale of the bearing seal structure renders direct, non-destructive observation of its degradation exceedingly challenging [34]. Consequently, the imperative arises for the establishment of a Health Indicator (HI) capable of characterizing the state and extent of degradation in bearings. This degradation indicator serves as a means to encapsulate the operational condition and performance of the bearing during its service life, and it lays the essential groundwork for the accurate prediction of Rolling Bearings' Remaining Useful Life (RUL). An ideal HI not only elucidates the trend of machinery deterioration but also significantly enhances the precision of RUL prediction [35]. Within the domain of machine RUL prediction, the Root Mean Square (RMS) stands as the most widely adopted and utilized HI [36]. Liu et al. [37] employed the Root Mean Square (RMS) as the degradation index for rolling bearings. They implemented a statistical process control (SPC) methodology to segment the degradation process into distinct stages, enabling adaptive model switching, and utilized the expectation maximization (EM) algorithm for real-time parameter updates. This approach facilitated the prediction of Remaining Useful Life (RUL) distribution across various degradation stages. Yang et al. [38] on the other hand, adopted RMS as the Health Indicator (HI) and introduced two novel evaluation metrics, the global gain indicator and local gain indicator, designed to discern diverse states of rolling bearings. However, relying solely on RMS as the exclusive degradation indicator for bearings may prove insufficient for capturing localized degradation patterns, potentially undermining the accuracy of overall RUL predictions. Alternatively, the extraction of multiple features from the raw vibration signals, followed by their consolidation into a unified feature, can heighten the sensitivity of health state information. This approach not only

enhances the precision of bearing predictions but also averts the potential for inaccuracies in the results [39]. Ni et al. [40] devised a novel Health Index (HI) for tracking degradation. This innovative HI incorporated information from spectral correlation, Wasserstein distance, and linear rectification, enabling it to represent all periodic power spectral probability distributions over time. In recent years, some methods for the adaptive extraction of health metrics have been gradually developed. Chen et al. [41] proposed a deep convolutional autoencoder based on quadratic function for constructing health metrics of bearings. Qin et al. [42] proposed a new degradation trend-constrained variational self-encoder to construct health indicators with obvious degradation trends. Zhou et al. [43] innovatively constructed a health indicator characterizing the degradation process by distributing the contact ratio metric. In the present investigation, a novel and all-encompassing health metric is introduced to delineate the bearing's degradation trend. This metric is grounded in the utilization of the RMS to capture global characteristics, with kurtosis and skewness employed to encapsulate local features. Subsequently, a distance matrix is assembled through the application of standardized Euclidean distances, and the amalgamation of these health metrics is accomplished via a multidimensional scaling transformation.

In the realm of traditional data-driven Remaining Useful Life (RUL) prediction research, due attention to the First Prediction Time (FPT) has often been disregarded. Nevertheless, FPT plays a pivotal role in the precision of RUL prediction. An incorrect determination of FPT can lead to the omission of critical information regarding early failures or the inclusion of early failure data within the RUL estimation. Kong et al. [44] employed the 3σ criterion technique to delineate distinct health stages. Wang et al. [45] performed health stage identification by measuring the gradient of degradation within a sliding window to characterize the degradation trend and to identify jump points. Wasim Ahmad et al. [36] utilized the rate of growth of a health indicator to ascertain the onset of health state degradation in a bearing. Nonetheless, it should be noted that these methods may be susceptible to substantial impacts when subjected to noise disturbances. In the present study, we propose an approach that combines the Isolation Forest algorithm with the 3σ method to determine the initial degradation point of the bearing.

A rolling bearing RUL prediction method based on IF-SCINet is proposed to address the above problems. The main contributions of this study include the following three aspects:

(1) A health indicator that more comprehensively reflects the bearing degradation trend is proposed for RUL prediction. The indicator is based on the seuclidean distance to construct the distance matrix, and the root mean square, skewness, and kurtosis are fused by a multidimensional scale transformation to realize a comprehensive representation of the bearing degradation trend.

(2) The anomaly detection algorithm Isolation Forest is applied to detect the initial degradation point of the bearing, and combined with the 3σ method criterion to determine the range of anomaly proportion, and finally combined with the grid search algorithm to determine the FPT.

(3) In this study, SCINet was used for the first time for RUL prediction of bearings and validated with PHM2012 and XJTU-SY datasets, and good experimental results were obtained.

The rest of the paper is organized as follows: Section II describes the basic theoretical background of the proposed method; Section III presents the details of the proposed prediction model; Section IV validates the effectiveness of the proposed method using the PHM2012 and XJTU-SY bearing datasets; and Section V concludes the paper.

II. BASIC THEORETICAL BACKGROUND

A. DESCRIPTION OF SELECTED FEATURES

In the bearing vibration index, the root mean square is a very important index, which reflects the strength and stability of the bearing vibration signal. For the time series data $x_1, x_2, \cdots, x_N$, there is the following formula:

$$x_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{1}$$

Kurtosis reacts to the shock characteristics of bearing vibration signals. Kurtosis is particularly sensitive to shock signals and is especially suitable for surface damage type faults, especially for early fault characterization.

$$\alpha = \frac{1}{N}\sum_{i=1}^{N} x_i^4 \tag{2}$$

Skewness reflects the asymmetry of the vibration signal and is used to measure the asymmetry of the probability distribution of a random variable. Normally the vibration signal is symmetric about the x-axis, when the skewness should be close to 0. If the friction or collision in a certain direction of the equipment is large, it will cause the vibration asymmetry and make the skewness larger.

$$\beta = \frac{1}{N}\sum_{i=1}^{N} x_i^3 \tag{3}$$

In the above equation: x_i denotes each data point in the dataset and N denotes the total number of data points.

B. MDS ALGORITHM

The Multidimensional Scaling (MDS) algorithm serves as a powerful tool for data reduction and visualization. It effectively converts high-dimensional data into a lower-dimensional space while preserving the intrinsic distance relationships among data points. At its essence, MDS leverages a distance matrix to encapsulate the similarities and correlations among data points. The dimensionality reduction process of the MDS algorithm consists of:

1) CONSTRUCT A DISTANCE MATRIX D

Calculate the distance between data points in the original space. By calculating the distance between each pair of data points, a distance matrix can be constructed. The constructed distance matrix is $D = (d(x, y))_{M\times M}$.

2) COMPUTE THE INNER PRODUCT MATRIX B

The inner product matrix B is computed by centering the distance matrix. The purpose of centering the distance matrix is to make the data points symmetric with respect to the origin. The centered distance matrix A is calculated by:

$$A = E - \frac{1}{N}jj^T \tag{4}$$

where E is a unit matrix of size $M \times M$ and j is an M-dimensional all-1 vector.

The distance matrix D is converted to the inner product matrix B,which is calculated as follows:

$$B = -\frac{1}{2}AD^2A \tag{5}$$

3) COMPUTE THE LOW-DIMENSIONAL MATRIX C

After obtaining the inner product matrix B, compute its eigenvalues and eigenvectors. The eigenvalues indicate the main direction of change of the data, and the eigenvectors indicate the magnitude in the corresponding direction. The matrix B can be decomposed as:

$$B = SVS^T \tag{6}$$

where V is the diagonal matrix corresponding to the eigenvalues of matrix B and S is the corresponding eigenvector.

The largest k eigenvalues and their corresponding eigenvectors are selected as the base of the k-dimensional space after dimensionality reduction. Calculate the coordinates after dimensionality reduction: project the original data onto the selected k-dimensional base to get the coordinates after dimensionality reduction. The specific calculation formula is:

$$Z = SV^{\frac{1}{2}} \tag{7}$$

The low-dimensional matrix C can be obtained by extracting the first x column vectors of the matrix Z, of size $M\times x$.

C. ISOLATION FOREST ALGORITHM

The Isolation Forest algorithm, introduced by Liu et al. [46] at the 8th IEEE International Conference on Data Mining in 2008, constitutes a machine learning approach for the detection of anomalies. The fundamental premise of this algorithm is rooted in the principle that data points exhibiting higher degrees of clustering should be partitioned more frequently, whereas data points with greater isolation merit less frequent partitioning. Isolation Forest leverages the count of segmentations to discern whether a given data point is tightly grouped (representative of normalcy) or relatively isolation (indicative of abnormality). The core assumption underlying this algorithm posits that anomalous data points tend to exhibit a sparser distribution within the feature space when compared to their normal counterparts. The primary objective of the Isolation Forest algorithm is the rapid isolation of anomalous data points in the upper strata of the constructed tree structure, while normal data points necessitate a greater number of splits to achieve the same isolation. This objective is particularly pertinent to the detection of initial deterioration points in bearing systems.

The algorithm initiates by creating an isolation tree, essentially a binary search tree. This isolation tree's construction process unfolds as follows: first, a dimension is randomly chosen for partitioning, and the space is divided using a randomly determined threshold situated between the minimum and maximum values of that dimension. Subsequently, points smaller than the threshold are assigned to the left subtree, while those larger are assigned to the right subtree. This partitioning operation is iteratively repeated within each subspace until a predetermined maximum tree height is reached, or until all leaf nodes contain only one data point. Each data point's depth within the tree is recorded as a score. Given the stochastic nature of generating a single isolation tree, multiple isolation trees are needed to be generated, and their average depths are calculated. To create diverse isolation trees, each one can be constructed by randomly sampling a subset N_i from the training set N. Then, by traversing each isolation tree within the isolation forest, the path length of a sample x in each isolation tree is computed, leading to the calculation of the corresponding anomaly score. Finally, to classify a sample as anomalous or not, a comparison is made between the anomaly scores of the samples. The formula for calculating the anomaly score of sample x is:

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{8}$$

In the formula: n denotes the number of leaf nodes of the isolation tree, $h(x)$ denotes the total path from the root node to the leaf node where sample x is located in the isolation tree, $E(h(x))$ denotes the expectation of $h(x)$ in the set of isolation trees, $c(n)$ denotes the average path length of the isolation tree.

D. SCINet MODEL

SCINet, called Sample Convolutional Interaction Network, is a neural network model for time series forecasting. The model was first proposed in 2021 and has demonstrated its superior performance on several datasets, as well as its lower time cost relative to other models. The overall structure of SCINet is shown in Fig. 1.

Fig. 1(a) shows the basic modules that make up the SCINet network: the SCI-Block, which downsamples the input features and divides them into two subsequences, F_{odd} and F_{even}, which retain most of the information of the original sequence. The features are then extracted from the two subsequences using different convolutional kernels respectively. In order to compensate for the feature loss caused by downsampling, the two subsequences are used for interactive learning by element product, and finally two updated sub-features F_{odd} and F_{even}

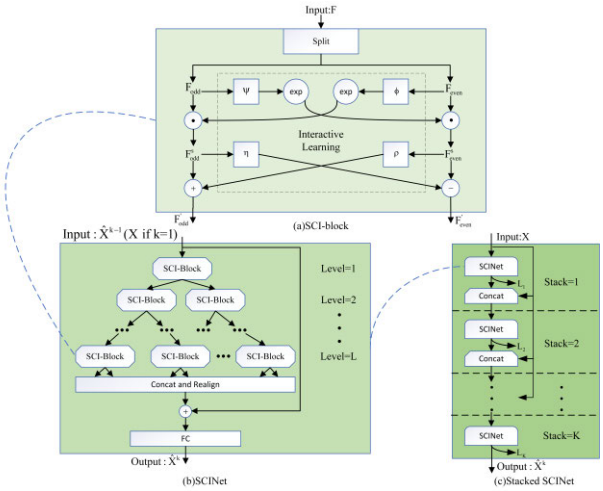

FIGURE 1. SCINet network structure.

are obtained, which can be expressed as:

$$F^s_{odd} = F_{odd} \odot \exp(\phi(F_{even})), F^s_{even} = F_{even} \odot \exp(\varphi(F_{odd})) \tag{9}$$

$$F'_{odd} = F^s_{odd} \pm \rho(D^s_{even}), F'_{even} = F^S_{even} \mp \eta(F^s_{odd}) \tag{10}$$

where ϕ, φ, ρ, and η denote four different one-dimensional convolutional modules, and $\odot$ denotes element-by-element convolution.

Fig. 1(b) shows a SCINet that uses a binary tree to combine the SCI-Blocks, and for k>1, downsampling and processing through different SCI-Block blocks is performed to learn valid features with different temporal resolutions. Previous information will be accumulated gradually, by which short-term and long-term temporal dependencies in the time series can be captured. The information will be combined and added to the original time series through residual linkage at the Concat&Realign layer, and finally the prediction output will be made through the fully connected layer.

Fig. 1(c) shows the SCINet stacked up to form a more powerful network Stacked SCINet, such that the network can be used to predict more complex time series, simplifying the learning of intermediate time features by applying intermediate supervision.

For training K ($K \geq 1$) Stacked SCINets, the loss of the Kth prediction result can be expressed as:

$$L_k = \frac{1}{\tau} \sum_{i=0}^{\tau} ||\hat{x}^k_i - x_i|| \tag{11}$$

The loss for the Kth SCINet depends on whether it is a single-step prediction or a multi-step prediction.

The single-step prediction of L_k can be expressed as:

$$L_k = \frac{1}{\tau - 1} \sum_{i=0}^{\tau-1} ||\hat{x}^k_i - x_i|| + \lambda ||\hat{x}^k_\tau - x_\tau|| \tag{12}$$

where, $\lambda \in (0, 1)$ denotes a balancing hyperparameter, and adjusting λ improves the prediction performance of the network.

The multistep prediction of L_k can be expressed as:

$$L_k = \frac{1}{\tau - 1} \sum_{i=0}^{\tau-1} ||\hat{x}^k_i - x_i|| \tag{13}$$

The full loss of Stacked SCINet can be expressed as:

$$L = \sum_{k=1}^{K} L_k \tag{14}$$

III. DETAILED INFORMATION ON THE PROPOSED LIFE PREDICTION MODEL

A. CONSTRUCTION OF RSK

In order to better reflect the degradation trend of the bearings, the three features described in Section II were selected in this study to construct the indices characterizing the degradation state of the bearings. In the process of bearing degradation, the root mean square can well represent the overall trend of bearing degradation. However, it is not obvious enough in the representation of the local degradation characteristics, so the kurtosis is introduced to describe the initial wear phenomenon of the bearing. When the failure is gradually aggravated, the skewness is again used to capture the localized features at the later stage, so as to more accurately represent the whole degradation process of rolling bearings. These three features characterizing the degradation trend are fused using the MDS algorithm. During the fusion process, it is necessary to consider how to construct a suitable distance matrix.

Through the comparison of the effects of Cosine distance, Euclidean distance, Mahalanobis distance, and Seuclidean distance (as shown in Fig. 2), it can be observed that both the Euclidean distance graph and the Seuclidean distance graph exhibit enhanced monotonicity. It is worth noting that the Seuclidean distance graph demonstrates fewer random disturbances. Therefore, the final decision is to use Seuclidean distance to construct the distance matrix. The formula for Seuclidean distance in n-dimensional space is as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (\frac{x_i - y_i}{s_i})^2} \tag{15}$$

where x_i ($1 \leq i \leq M$) and y_i ($1 \leq i \leq M$) denote the values of the two data points in the ith dimension and is the standard deviation in the ith dimension.

The seuclidean is a variant of the eucldean that is used to resolve scale differences between dimensions. More specifically, seuclidean distance is equivalent to dividing the values on each dimension by the standard deviation on that dimension before calculating the Euclidean distance. This helps to eliminate the effects of different scales of values in different dimensions and ensures that the contribution of each dimension to the distance calculation is relatively balanced.

At this stage, the construction of the health indicator, Root Mean Square-Skewness-Kurtosis (RSK), is complete.

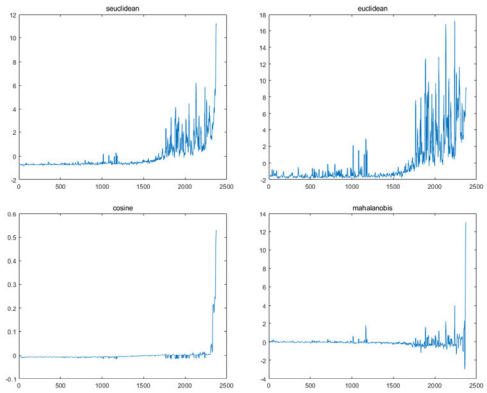

FIGURE 2. Comparison of four distances.

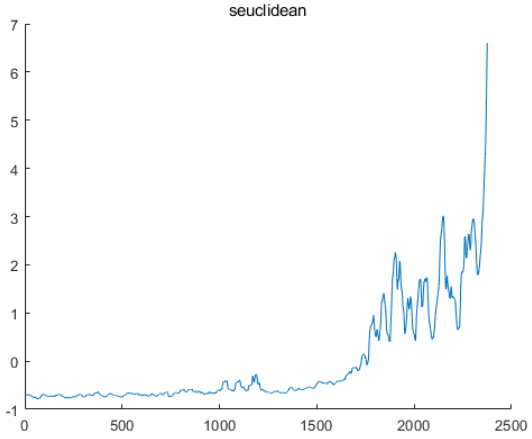

FIGURE 3. Distance map after filtering.

B. RSK HEALTH INDICATOR PREPROCESSING

To efficiently mitigate the impact of random noise and high-frequency components, this study employs a moving average filter with a 5-step delay window to preprocess HI contaminated by noise. The moving-average filter is a widely adopted signal processing method achieved by computing the mean value of data points within a fixed window. More precisely, for each data point, it is aggregated with all other data points within the window, and the result is divided by the window's size to derive the average value, which subsequently replaces the original data points. Fig. 3 provides a visual representation of the filtered distance plot.

C. IMPROVED IF-SCINET

After constructing the bearing health indicator RSK and normalizing it, it was then necessary to determine a suitable initial degradation point in order to determine the time of the first prediction of the model. The determination of the initial degradation point is in a way equivalent to bearing anomaly detection.

1) IMPROVED IF DETERMINATION OF FPT

Existing anomaly detection methods often rely on the modeling of normal data samples to establish their distribution within the feature space. Consequently, any data falling outside this established region is classified as an anomaly. However, a significant drawback of these techniques lies in their excessive emphasis on modeling normal data, with relatively less attention given to abnormal data. This imbalance can result in high false positive rates or the failure to detect anomalies when they occur. The task of identifying the initial degradation point in bearings fundamentally aligns with anomaly detection. Therefore, this study employs the Isolation Forest algorithm for the purpose of detecting the initial degradation point in bearings.

In traditional isolation forest algorithms, it is necessary to manually specify a hyperparameter, which is the proportion of anomalous data in the dataset. This brings a lot of inconvenience. In order to find the most appropriate proportion of anomalies, this study introduces the 3σ criterion into the method, i.e., the probability that the distribution of values is in the range of $(\mu - \sigma, \mu + \sigma)$ is 0.68, and the proportion of anomalies is set as $(0, 0.68]$. Then, the grid search algorithm is used to select the optimal value of the anomaly proportion. So far, the improved IF realizes the adaptive determination of the initial bearing degradation point. Fig. 4 demonstrates the comparison of the initial degradation points of the bearings determined by the isolation forest algorithm and the 3σ method. As can be seen in Fig. 4: For Bearing1_3, the initial degradation point identified is not very different from the degradation status shown in the corresponding amplitude diagram. For Bearing2_6, although there are no obvious signs of degradation in the corresponding amplitude diagram, there is a more obvious increase in the magnitude of change in the corresponding health indicators after the initial degradation point is determined. For bearing 3_3, the initial degradation point identified is slightly advanced from the amplitude plot, but within reasonable limits. This may be due to the random nature of isolation forests when constructing isolation trees. The improved isolation forest algorithm is able to adaptively detect the initial degradation points of the bearings, while the traditional method is less effective. This may be due to the introduction of a delay factor in the moving average filtering, which causes the initial degradation point determined by the traditional method to appear at the stage of bearing failure, which further illustrates the effectiveness of the proposed method.

2) IF-SCINET MODEL

After adaptively determining the FPT of RSK using the improved Isolation Forest algorithm, it needs to be imported into the SCINet model for prediction. In this paper, we did not improve the structure of the SCINet model, but instead used only a grid search to determine the optimal combination of the number of level tiers for the models SCI-Block and Stack SCINet, i.e., the model reaches an optimal solution when level=2 and Stack=1. Fig. 5 illustrates the improved IF-SCINet model. First, a basic Isolation Forest model is defined. Next, a 3σ method was used to determine the proportion of anomalies in the data, followed by a grid search to determine the smallest negative average anomaly score. Finally, the Isolation Forest model was retrained using the

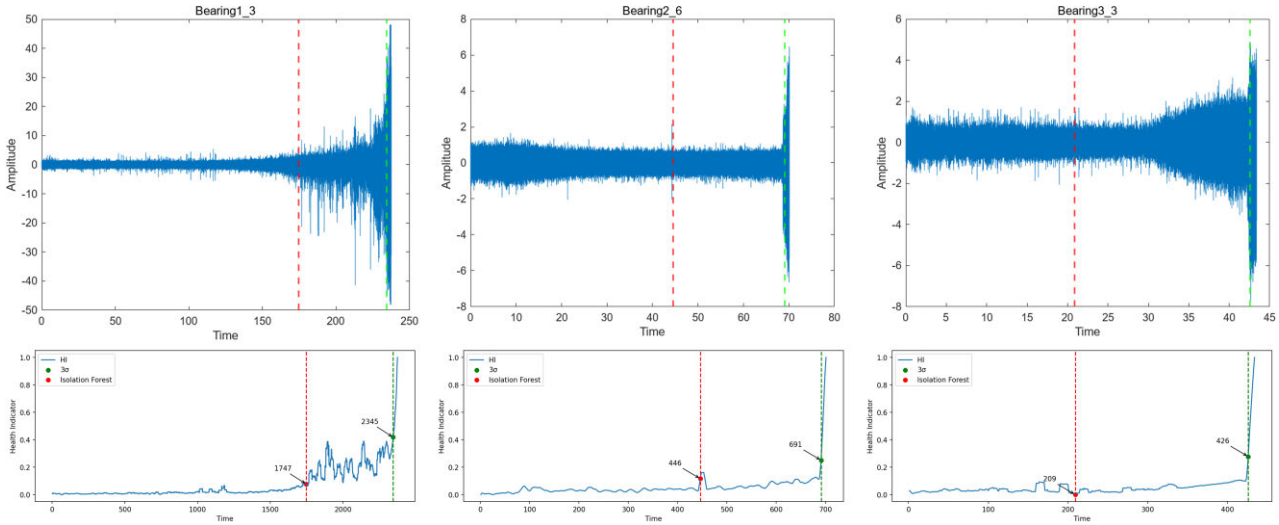

FIGURE 4. Comparison chart for determining the first prediction point.

optimal proportion of anomalies to determine the bearing FPT. After that, the training and prediction phases of the SCINet model are initiated. The initial step is to downsample the training data followed by interactive learning to compensate for the loss due to downsampling to capture local and global features of the time series data. Subsequently, short-term and long-term dependencies in the time series are captured by arranging multiple SCI-Blocks in a tree structure. Next, a new sequence re-formed by an inverse parity splitting operation is added to the original sequence by residual joining. After joining via fully connected layers, intermediate supervision is added to simplify the learning of intermediate temporal features. Finally, the trained model is used to predict the output of the test data. In contrast to SCINet, Stack SCINet introduces intermediate supervision. Intermediate supervision helps the model understand the dynamics of time series data more effectively and improves the ability to extract important features in the time series. Specifically, intermediate supervision enables the model to perform more frequent parameter updates when learning time series data representations by introducing additional supervisory signals or targets during the training process. This helps alleviate the gradient vanishing problem, accelerates the convergence process of the model, and improves the model's ability to capture dynamic changes in the data.

Fig. 6 shows the overall prediction flow of the proposed method. The overall process consists of feature extraction of vibration signals and construction of health indicators, followed by determination of the first prediction time of the bearing using the improved IF, and finally life prediction using the SCINet network.

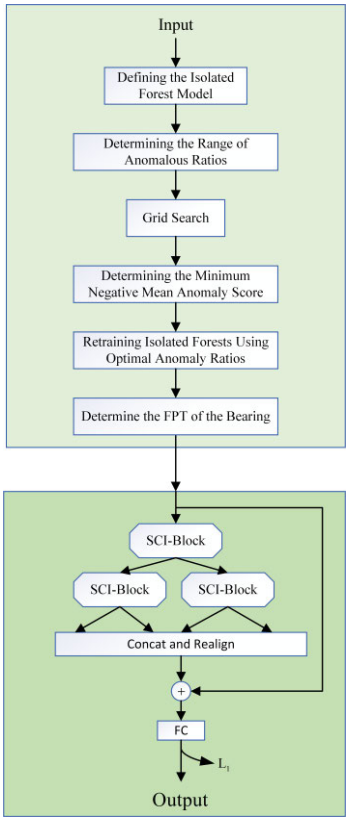

FIGURE 5. Improved IF-SCINet modeling.

IV. CASE STUDIES

A. EVALUATION INDICATORS

The evaluation metrics used in this study are commonly used in the field of PHM research are Mean Absolute Error (MAE), Normalized Root Mean Square Error (NRMSE) and Mean Absolute Percentage Error (MAPE). The calculation formula is as follows:

$$MAE(X, h) = \frac{1}{\tau}\sum_{i=1}^{m}\left|\hat{x}_i - x_i\right| \tag{16}$$

$$NRMSE = \frac{\sqrt{\frac{1}{\tau}\sum_{i=0}^{\tau}(\hat{x}_i - x_i)^2}}{X} \tag{17}$$

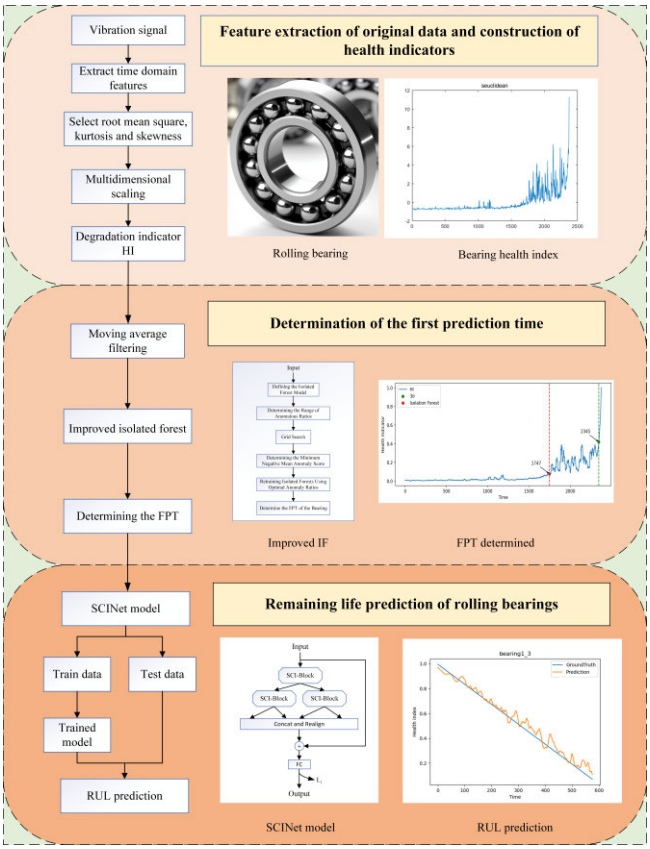

FIGURE 6. Prediction flow of the proposed method.

$$MAPE = \sqrt{\frac{1}{\tau}\sum_{I=0}^{\tau} |(\hat{x}_i - x_i)/x_i|} \tag{18}$$

where τ denotes the number of test samples, $\hat{x}_i$ denotes the predicted value of the ith sample, x_i denotes the true value of the ith sample, and X denotes the difference between the maximum and minimum values of the true value of the sample.

MAE focuses on the absolute value of the prediction error and MAPE focuses on the percentage of error. NRMSE is more useful in comparing model performance across different datasets because it normalizes the range of data [47].

B. MODEL HYPERPARAMETERS AND MODEL COMPARISON SETTINGS

In order to make an assessment of the network performance, typical time series prediction models LSTM, TCN, GRU, and the Informer model were used to compare with the SCINet and IF-SCINet models. Although single-step prediction can obtain higher prediction accuracy, it cannot show the long-term degradation trend of the bearing. Therefore, multi-step prediction is used in this study, and the prediction step sizes are set to 12, 24, 36, and 48, respectively. On the XJTU-SY dataset, only 12 and 24 steps were selected for prediction due to the amount of bearing data. A grid search is performed for all tunable hyperparameters in the network model to determine the optimal parameters. As shown in Table 1.

TABLE 1. Hyperparameter settings for the IF-SCINet model.

Parameter	Value (PHM2012)/(XJTU-SY)
Epoch	100
Batch size	(32)/(16)
Dropout	0.3
Learning rate	0.0001
Level	2
Stack	1
Patience	5
Kernel	(7)/(5)
Hidden size	1

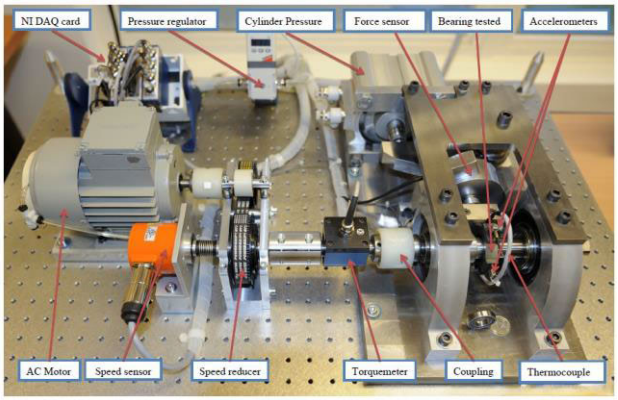

FIGURE 7. Pronostia experimental platform.

C. DATASET 1

1) INTRODUCTION TO THE DATASET1

To validate the effectiveness of the bearing life prediction method proposed in this study, the first set of bearing degradation data used the PHM2012 Challenge dataset [48]. This dataset was obtained by performing accelerated degradation experiments on bearings on the PRONOSTIA experimental platform. The structure of the experimental setup is shown in Fig 7. The setup consists of an asynchronous motor, a bearing under test, and a variety of sensors. The data provided by the PRONOSTIA platform corresponds to some extent to normally degraded bearings, and each degraded bearing contains virtually all types of defects, including balls, rings, and cages.

The dataset contains data for 3 different operating conditions, with 7, 7 and 3 sub-datasets corresponding to each load. As shown in Table 2.

2) RUL PREDICTION RESULTS

The comparative results of the model evaluation indicators are shown in Fig. 8 through 10.

First of all, from Fig. 8 to 10 as a whole, it can be seen that the prediction performance of the SCINet and IF-SCINet models is significantly better than that of the other four typical time series models in most cases, and although the prediction performance is not the optimal prediction

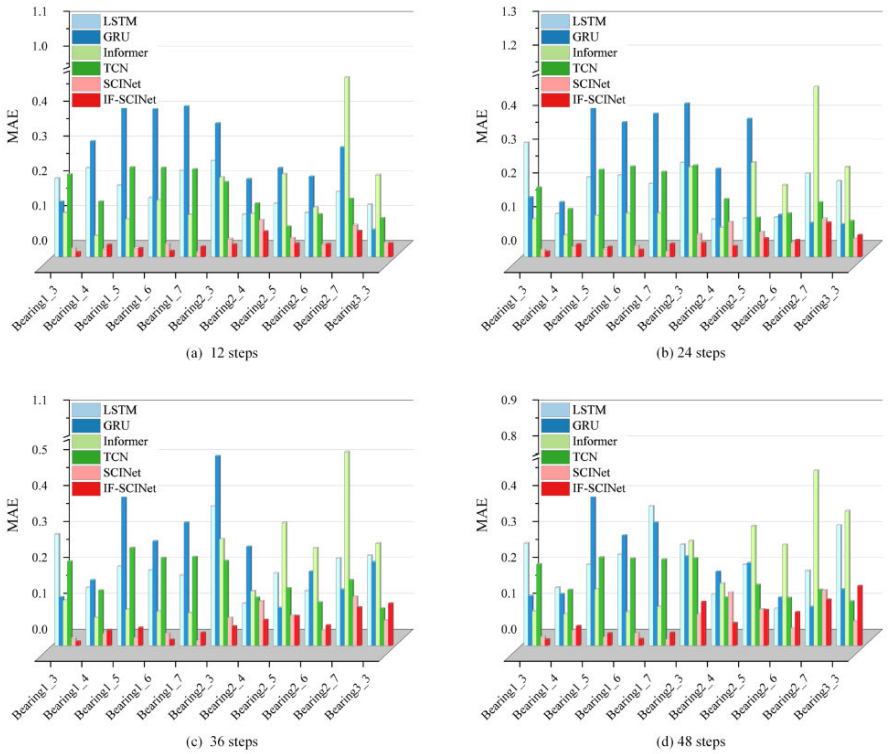

FIGURE 8. Multi-step prediction comparison chart (MAE).

TABLE 2. Display of bearing parameters.

Operating conditions	Rotating Speed(rpm)/ Load(N)	Bearing datasets
Condition1	1800/4000	Training(Bearing1_1—1_2) Test(Bearing1_3—1_7)
Condition2	1650/4200	Training(Bearing2_1—2_2) Test(Bearing2_3—2_7)
Condition3	1500/5000	Training(Bearing3_1—3_2) Test(Bearing3_3)

performance in some cases, the deviation from the optimal prediction is not very large. As the number of forecasting steps increases, the SCINet and IF-SCINet networks still perform well in most cases, which further illustrates the superiority of SCINet and IF-SCINet networks in long time series forecasting. Secondly, it can be seen from Fig. 8 that the MAE values of both the SCINet model and the IF-SCINet model increase as the number of prediction steps increases. However, they are still able to maintain lower MAE values, i.e., smaller prediction errors, on most bearings compared to the other four conventional models. The overall decreasing MAE values of the GRU network model as the number of prediction steps increases may be attributed to the fact that the gating mechanism designed in its network model enables it to capture and memorize the long-term dependencies in the sequences more efficiently. As can be seen in Fig. 9, from 12-step prediction to 48-step prediction, the SCINet model and the IF-SCINet model maintain low NRMSE values in most cases, i.e., the model fits well. For Bearing2_4, SCINet showed a poor fit, but the improved IF-SCINet showed a superior model fit. This further illustrates the effectiveness of the method. From Fig. 10, it can be seen that the SCINet model and the IF-SCINet model exhibit lower MAPE values, i.e., show better prediction results, compared with the four traditional prediction models. And the prediction effect of IF-SCINet model is better than SCINet model in most cases.

Fig. 11 shows the remaining life prediction plots of Bearing1_3, Bearing2_6 and Bearing3_3 after determining the first prediction time, and on the whole, it can be seen that the proposed method is basically able to accurately track the predicted real RUL of the bearings. for Bearing1_3, the overall prediction effect is relatively stable, and basically able to track the real life value. For Bearing2_6 and Bearing3_3, although there is a slight disturbance at the beginning of the prediction, it basically fits the real life curve well at the late stage of prediction.

D. DATASET 2

1) INTRODUCTION TO THE DATASET2

The second set of bearing degradation data utilizes the XJTUSY dataset [49]. This dataset was collected from accelerated life tests of 15 rolling bearings under three operating conditions. The test platform is shown in Fig. 12. The test rig provides full-life data for a total of 15 bearings under three operating conditions with a sampling frequency of 25.6 KHz.

In this paper, we have selected the bearings in the second condition for generalizability experiments, as shown in Table 3.

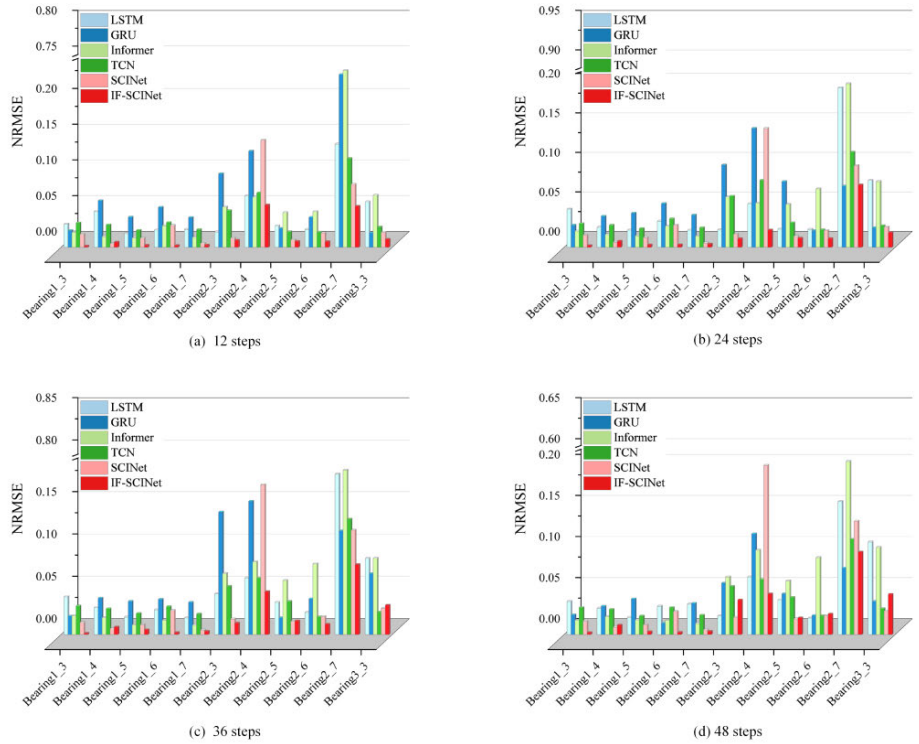

FIGURE 9. Multi-step prediction comparison chart (NRMSE).

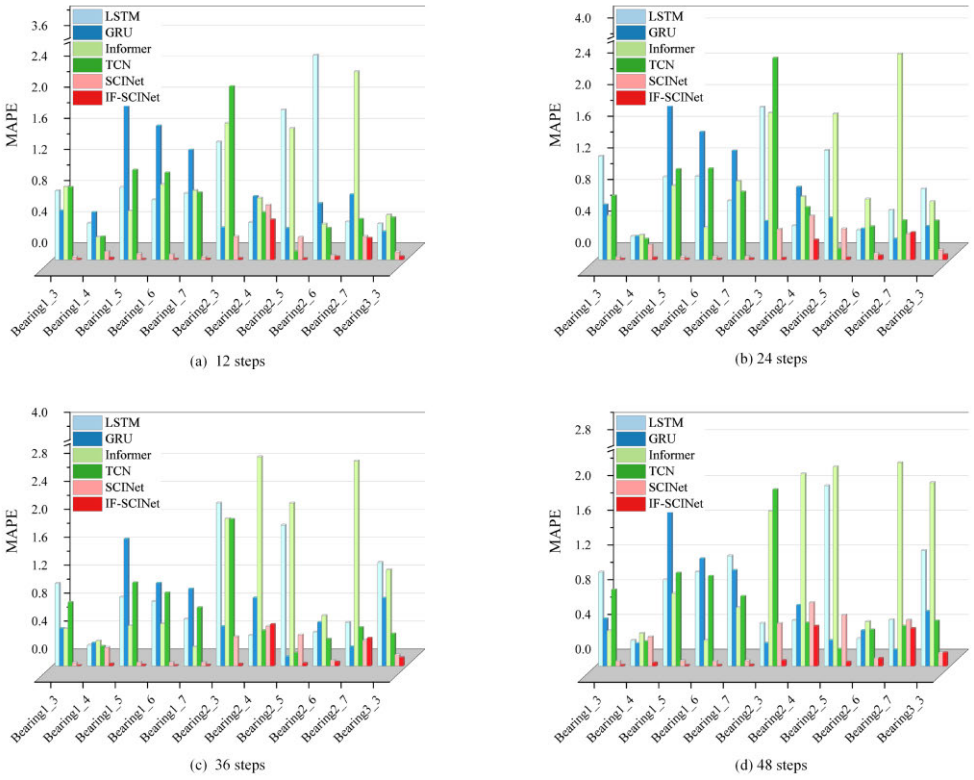

FIGURE 10. Multi-step prediction comparison chart (MAPE).

2) DETERMINATION OF THE FPT

Fig. 13 shows the determination of bearing 2_4 FPT, from which it can be seen that although there is no obvious degradation trend on the amplitude graph, there is a clear sign of increase on the constructed health indicator after the determination of FPT, which also further illustrates that the

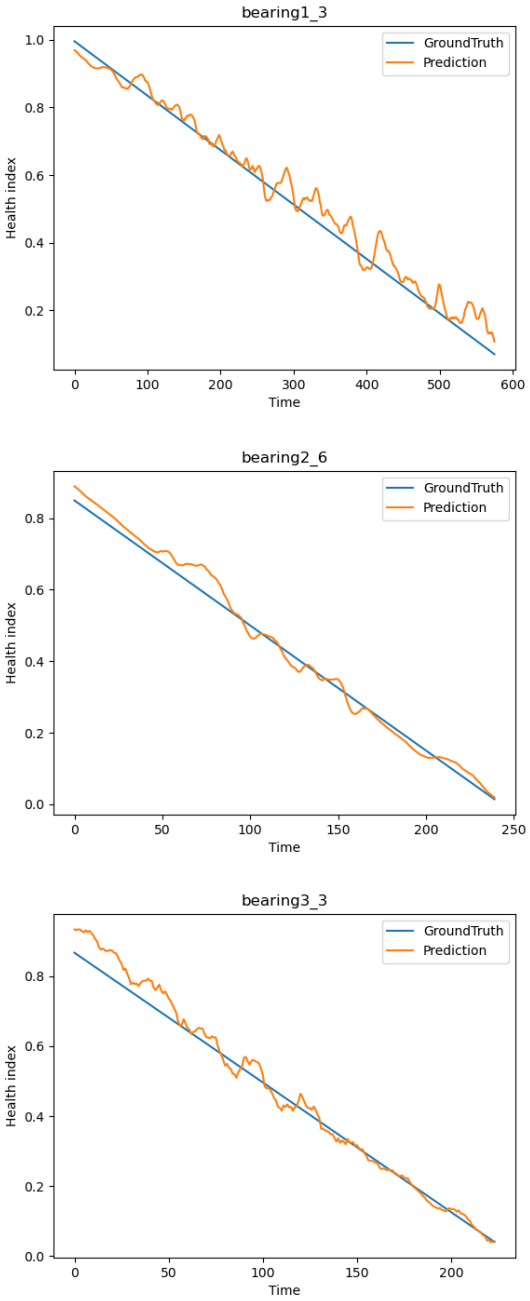

FIGURE 11. Partial bearing forecast charts.

TABLE 3. Display of bearing parameters.

Operating conditions	Rotating Speed(r/min)/ Load(KN)	Bearing datasets
Condition II	2250/11	Training(Bearing2_1—2_2) Test(Bearing2_3—2_5)

proposed way of constructing the health indicator is more capable of reflecting the degradation characteristics of the bearing.

3) RESULTS OF MODEL EVALUATION

Fig. 14 exhibits the evaluation effectiveness graphs of the six models on the XJTU-SY dataset. From the overall view of the

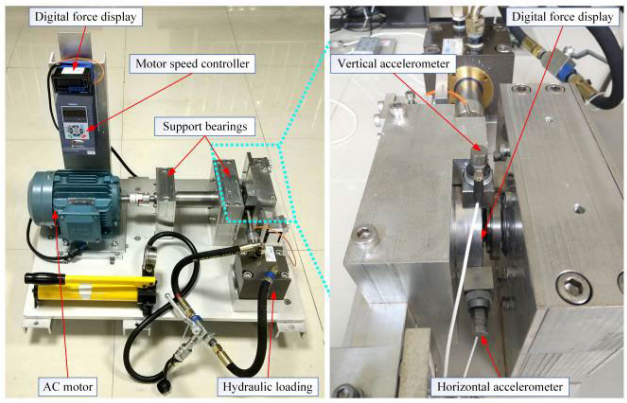

FIGURE 12. Bearing test rig of XJTU-SY.

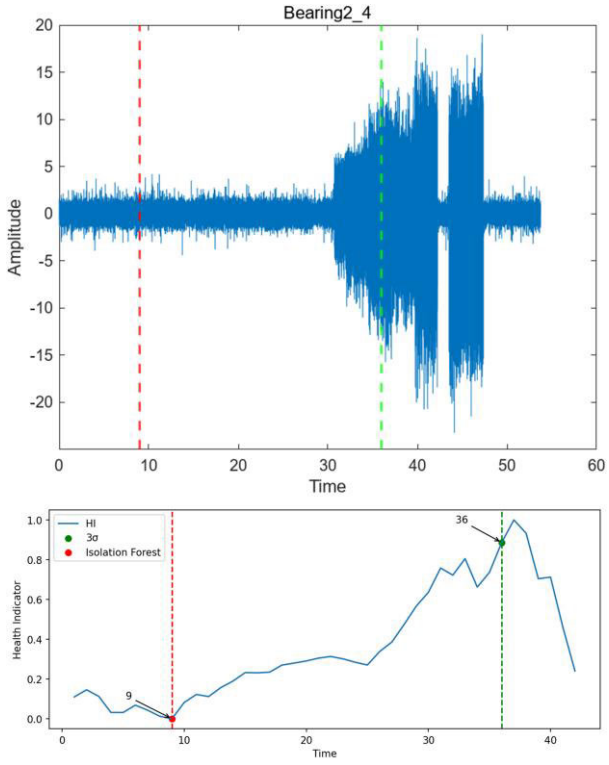

FIGURE 13. FPT determination for Bearing2_4.

three evaluation index comparison graphs, IF-SCINet shows superior performance in most cases. From the MAE comparison plot, it can be seen that the evaluation performance of the Informer model varies greatly on different bearings, while the IF-SCINet model performs more stably, which also indicates the model's higher robustness. From the NRMSE plot, it can be seen that the TCN model and the IF-SCINet show not much difference in the evaluation results. the GRU network even shows more superior results on some bearings. From the MAPE plot, it can be seen that for Bearings 2_3, the original SCINet model exhibits moderate evaluation results compared to the other models, but the IF-SCINet model exhibits optimal results, which further validates the effectiveness of the proposed method. In response to the comparison of the NRMSE

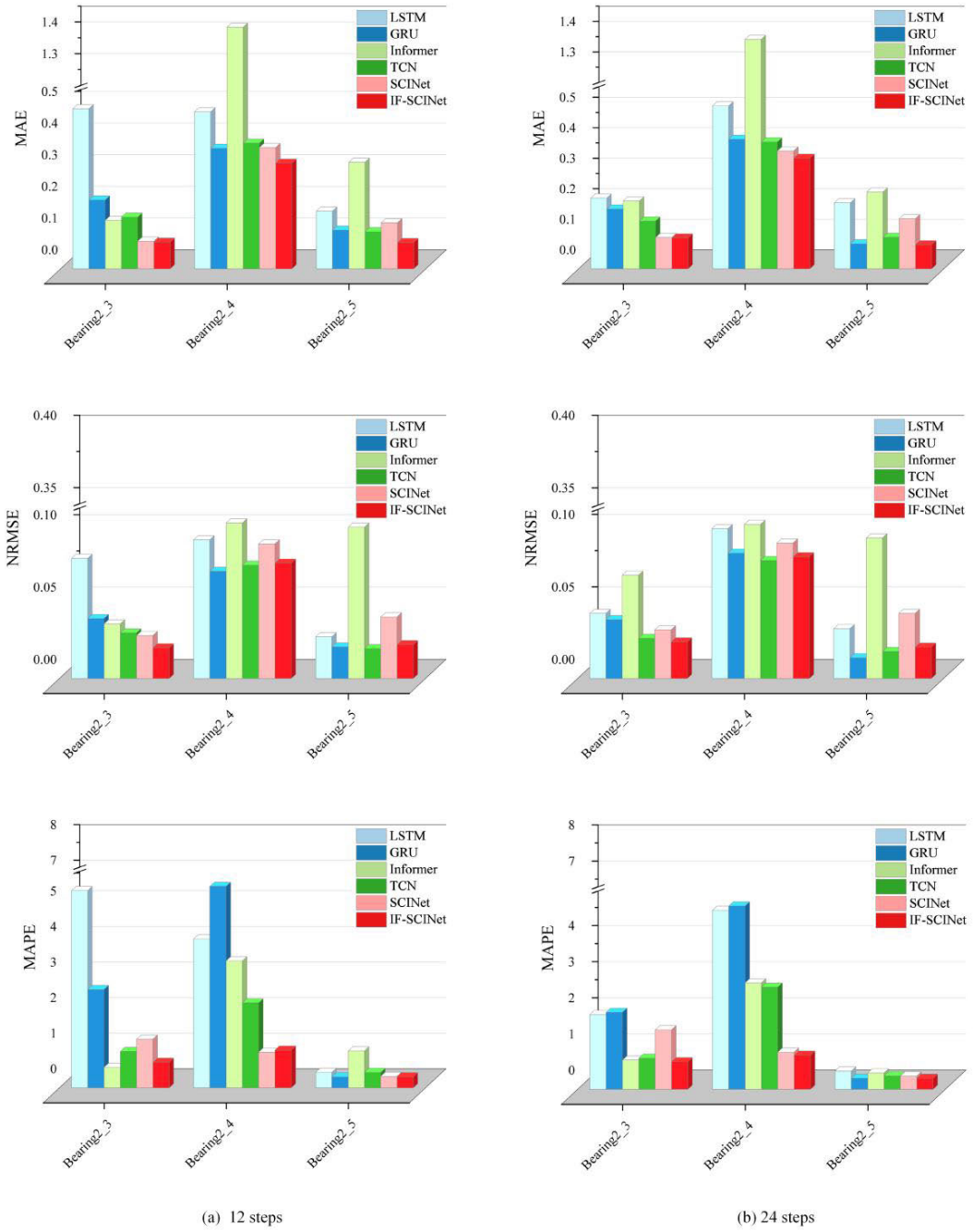

FIGURE 14. Comparison of multi-step forecasts for the three evaluation indicators.

plots in Fig. 9 and Fig. 14, it can be seen that, except for the individual bearings, there is not much difference in the NRMSEs on the two datasets, which further illustrates the strong generalization of the model.

V. CONCLUSION

In this study, a bearing RUL prediction method based on IF-SCINet is proposed. Firstly, the root mean square, skewness and kurtosis features are extracted from the raw vibration data, the distance matrix is constructed by using the standardized Euclidean distance, and the health indicators are obtained by multidimensional scale transformation. Then, adaptive determination of FPT is realized by combining isolation forest and 3σ criterion. Finally, the performance of the prediction model is evaluated by MAE, NRMSE and MAPE evaluation metrics, and the results show that the IF-SCINet time series model has better prediction effect.

The following ideas for future research are presented:First of all, the study of adaptive extraction of features to construct health indicators is a popular one, but how to be able to adaptively extract effective features to construct health indicators is still a challenge, and neural network as a black box,

how to make sure that the extracted features are physically meaningful and have good results is a difficult problem. Secondly, most of the current studies only provide a prediction result without a detailed explanation of the uncertainty or error of that result. Therefore, how to reasonably quantify the uncertainty of a prediction result is also a question worth investigating.

REFERENCES

[1] J. Zhuang, M. Jia, Y. Cao, and X. Zhao, "Semi-supervised double attention guided assessment approach for remaining useful life of rotating machinery," *Rel. Eng. Syst. Saf.*, vol. 226, Oct. 2022, Art. no. 108685, doi: 10.1016/j.ress.2022.108685.

[2] Q. Li, C. Yan, G. Chen, H. Wang, H. Li, and L. Wu, "Remaining useful life prediction of rolling bearings based on risk assessment and degradation state coefficient," *ISA Trans.*, vol. 129, pp. 413–428, Oct. 2022, doi: 10.1016/j.isatra.2022.01.031.

[3] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005, doi: 10.1109/TEC.2005.847955.

[4] S. Gawde, S. Patil, S. Kumar, P. Kamat, K. Kotecha, and A. Abraham, "Multi-fault diagnosis of industrial rotating machines using data-driven approach: A review of two decades of research," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106139, doi: 10.1016/j.engappai.2023.106139.

[5] W. Mao, J. Chen, J. Liu, and X. Liang, "Self-supervised deep domain-adversarial regression adaptation for online remaining useful life prediction of rolling bearing under unknown working condition," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1227–1237, Feb. 2023, doi: 10.1109/TII.2022.3172704.

[6] N. Costa and L. Sánchez, "Variational encoding approach for interpretable assessment of remaining useful life estimation," *Rel. Eng. Syst. Saf.*, vol. 222, Jun. 2022, Art. no. 108353, doi: 10.1016/j.ress.2022.108353.

[7] K. Chen, J. Liu, W. Guo, and X. Wang, "A two-stage approach based on Bayesian deep learning for predicting remaining useful life of rolling element bearings," *Comput. Electr. Eng.*, vol. 109, Aug. 2023, Art. no. 108745, doi: 10.1016/j.compeleceng.2023.108745.

[8] Y. Cao, M. Jia, Y. Ding, X. Zhao, P. Ding, and L. Gu, "Complex domain extension network with multi-channels information fusion for remaining useful life prediction of rotating machinery," *Mech. Syst. Signal Process.*, vol. 192, Jun. 2023, Art. no. 110190, doi: 10.1016/j.ymssp.2023.110190.

[9] L. Li, J. Xu, and J. Li, "Estimating remaining useful life of rotating machinery using relevance vector machine and deep learning network," *Eng. Failure Anal.*, vol. 146, Apr. 2023, Art. no. 107125, doi: 10.1016/j.engfailanal.2023.107125.

[10] S. Lu, Z. Gao, Q. Xu, C. Jiang, T. Xie, and A. Zhang, "Remaining useful life prediction via interactive attention-based deep spatio-temporal network fusing multisource information," *IEEE Trans. Ind. Electron.*, Early Access, doi: 10.1109/TIE.2023.3301551.

[11] H. Wang, H. Liao, and X. Ma, "Remaining useful life prediction considering joint dependency of degradation rate and variation on time-varying operating conditions," *IEEE Trans. Rel.*, vol. 70, no. 2, pp. 761–774, Jun. 2021, doi: 10.1109/TR.2020.3002262.

[12] F. Jiang, K. Ding, G. He, H. Lin, Z. Chen, and W. Li, "Dual-attention-based multiscale convolutional neural network with stage division for remaining useful life prediction of rolling bearings," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022, doi: 10.1109/TIM.2022.3210933.

[13] H.-B. Zhang, D.-J. Cheng, K.-L. Zhou, and S.-W. Zhang, "Deep transfer learning-based hierarchical adaptive remaining useful life prediction of bearings considering the correlation of multistage degradation," *Knowl.-Based Syst.*, vol. 266, Apr. 2023, Art. no. 110391, doi: 10.1016/j.knosys.2023.110391.

[14] G. Qiu, Y. Gu, and J. Chen, "Selective health indicator for bearings ensemble remaining useful life prediction with genetic algorithm and Weibull proportional hazards model," *Measurement*, vol. 150, Jan. 2020, Art. no. 107097, doi: 10.1016/j.measurement.2019.107097.

[15] W. Zhao, C. Zhang, J. Wang, S. Wang, D. Lv, and F. Qin, "Research on digital twin driven rolling bearing model-data fusion life prediction method," *IEEE Access*, vol. 11, pp. 48611–48627, 2023, doi: 10.1109/ACCESS.2023.3277040.

[16] S. B. Ramezani, L. Cummins, B. Killen, R. Carley, A. Amirlatifi, S. Rahimi, M. Seale, and L. Bian, "Scalability, explainability and performance of data-driven algorithms in predicting the remaining useful life: A comprehensive review," *IEEE Access*, vol. 11, pp. 41741–41769, 2023, doi: 10.1109/ACCESS.2023.3267960.

[17] N. Li, Y. Lei, X. Liu, T. Yan, and P. Xu, "Machinery health prognostics with multimodel fusion degradation modeling," *IEEE Trans. Ind. Electron.*, vol. 70, no. 11, pp. 11764–11773, Nov. 2023, doi: 10.1109/TIE.2022.3231273.

[18] Y. Wang, M. Wu, R. Jin, X. Li, L. Xie, and Z. Chen, "Local–global correlation fusion-based graph neural network for remaining useful life prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, Early Access doi: 10.1109/TNNLS.2023.3330487.

[19] Y.-I. Park, J. W. Song, and S.-J. Kang, "Pseudo-label-vector-guided parallel attention network for remaining useful life prediction," *IEEE Trans. Ind. Informat.*, vol. 19, no. 4, pp. 5602–5611, Apr. 2023, doi: 10.1109/TII.2022.3202832.

[20] L. Song and H. Wang, "Improved CEEMDAN-based aero-engine gas-path parameter forecasting using SCINet," *J. Mech. Sci. Technol.*, vol. 37, no. 3, pp. 1485–1500, 2023, doi: 10.1007/s12206-023-0234-y.

[21] F. Shen and R. Yan, "A new intermediate-domain SVM-based transfer model for rolling bearing RUL prediction," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 3, pp. 1357–1369, Jun. 2022, doi: 10.1109/TMECH.2021.3094986.

[22] Q. Hu, Y. Zhao, Y. Wang, P. Peng, and L. Ren, "Remaining useful life estimation in prognostics using deep reinforcement learning," *IEEE Access*, vol. 11, pp. 32919–32934, 2023, doi: 10.1109/ACCESS.2023.3263196.

[23] M. G. Alfarizi, B. Tajiani, J. Vatn, and S. Yin, "Optimized random forest model for remaining useful life prediction of experimental bearings," *IEEE Trans. Ind. Informat.*, vol. 19, no. 6, pp. 7771–7779, Jun. 2023, doi: 10.1109/TII.2022.3206339.

[24] T. Zhang, Q. Wang, Y. Shu, W. Xiao, and W. Ma, "Remaining useful life prediction for rolling bearings with a novel entropy-based health indicator and improved particle filter algorithm," *IEEE Access*, vol. 11, pp. 3062–3079, 2023, doi: 10.1109/ACCESS.2023.3234286.

[25] P. V. Kamat, R. Sugandhi, and S. Kumar, "Deep learning-based anomaly-onset aware remaining useful life estimation of bearings," *PeerJ Comput. Sci.*, vol. 7, p. e795, Nov. 2021, doi: 10.7717/peerj-cs.795.

[26] T. Luo, M. Liu, P. Shi, G. Duan, and X. Cao, "A hybrid data preprocessing-based hierarchical attention BiLSTM network for remaining useful life prediction of spacecraft lithium-ion batteries," *IEEE Trans. Neural Netw. Learn. Syst.*, Early Access 2023, doi: 10.1109/TNNLS.2023.3311443.

[27] H. Wang, J. Yang, R. Wang, and L. Shi, "Remaining useful life prediction of bearings based on convolution attention mechanism and temporal convolution network," *IEEE Access*, vol. 11, pp. 24407–24419, 2023, doi: 10.1109/ACCESS.2023.3255891.

[28] C. Yang, J. Ma, X. Wang, X. Li, Z. Li, and T. Luo, "A novel based-performance degradation indicator RUL prediction model and its application in rolling bearing," *ISA Trans.*, vol. 121, pp. 349–364, Feb. 2022, doi: 10.1016/j.isatra.2021.03.045.

[29] S. Wang, Z. Yu, G. Xu, and F. Zhao, "Research on tool remaining life prediction method based on CNN-LSTM-PSO," *IEEE Access*, vol. 11, pp. 80448–80464, 2023, doi: 10.1109/ACCESS.2023.3299849.

[30] L. Cao, H. Zhang, Z. Meng, and X. Wang, "A parallel GRU with dual-stage attention mechanism model integrating uncertainty quantification for probabilistic RUL prediction of wind turbine bearings," *Rel. Eng. Syst. Saf.*, vol. 235, Jul. 2023, Art. no. 109197, doi: 10.1016/j.ress.2023.109197.

[31] Y. Wang, H. Ding, and X. Sun, "Residual life prediction of bearings based on SENet-TCN and transfer learning," *IEEE Access*, vol. 10, pp. 123007–123019, 2022, doi: 10.1109/ACCESS.2022.3223387.

[32] Y. Liu, J. Chen, Y. Chang, S. He, and Z. Zhou, "A novel integration framework for degradation-state prediction via transformer model with autonomous optimizing mechanism," *J. Manuf. Syst.*, vol. 64, pp. 288–302, Jul. 2022, doi: 10.1016/j.jmsy.2022.07.004.

[33] M. Liu, A. Zeng, M. Chen, Z. Xu, Q. Lai, L. Ma, and Q. Xu, "SciNet: Time series modeling and forecasting with sample convolution and interaction," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 5816–5828.

[34] J. Li, Y. Zi, Y. Wang, and Y. Yang, "Health indicator construction method of bearings based on Wasserstein dual-domain adversarial networks under normal data only," *IEEE Trans. Ind. Electron.*, vol. 69, no. 10, pp. 10615–10624, Oct. 2022, doi: 10.1109/TIE.2022.3156148.

[35] D. Chen, Y. Qin, Q. Qian, Y. Wang, and F. Liu, "Transfer life prediction of gears by cross-domain health indicator construction and multi-hierarchical long-term memory augmented network," *Rel. Eng. Syst. Saf.*, vol. 230, Feb. 2023, Art. no. 108916, doi: 10.1016/j.ress.2022.108916.
[36] W. Ahmad, S. A. Khan, and J.-M. Kim, "A hybrid prognostics technique for rolling element bearings using adaptive predictive models," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1577–1584, Feb. 2018, doi: 10.1109/TIE.2017.2733487.
[37] S. Liu and L. Fan, "An adaptive prediction approach for rolling bearing remaining useful life based on multistage model with three-source variability," *Rel. Eng. Syst. Saf.*, vol. 218, Feb. 2022, Art. no. 108182, doi: 10.1016/j.ress.2021.108182.
[38] N. Yang, Z. Wang, W. Cai, and Y. Li, "Data regeneration based on multiple degradation processes for remaining useful life estimation," *Rel. Eng. Syst. Saf.*, vol. 229, Jan. 2023, Art. no. 108867, doi: 10.1016/j.ress.2022.108867.
[39] S. Buchaiah and P. Shakya, "Bearing fault diagnosis and prognosis using data fusion based feature extraction and feature selection," *Measurement*, vol. 188, Jan. 2022, Art. no. 110506, doi: 10.1016/j.measurement.2021.110506.
[40] Q. Ni, J. C. Ji, and K. Feng, "Data-driven prognostic scheme for bearings based on a novel health indicator and gated recurrent unit network," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1301–1311, Feb. 2023, doi: 10.1109/TII.2022.3169465.
[41] D. Chen, Y. Qin, Y. Wang, and J. Zhou, "Health indicator construction by quadratic function-based deep convolutional auto-encoder and its application into bearing RUL prediction," *ISA Trans.*, vol. 114, pp. 44–56, Aug. 2021, doi: 10.1016/j.isatra.2020.12.052.
[42] Y. Qin, J. Zhou, and D. Chen, "Unsupervised health indicator construction by a novel degradation-trend-constrained variational autoencoder and its applications," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 3, pp. 1447–1456, Jun. 2022, doi: 10.1109/TMECH.2021.3098737.
[43] J. Zhou, Y. Qin, J. Luo, and T. Zhu, "Remaining useful life prediction by distribution contact ratio health indicator and consolidated memory GRU," *IEEE Trans. Ind. Informat.*, vol. 19, no. 7, pp. 8472–8483, Jul. 2023, doi: 10.1109/TII.2022.3218665.
[44] X. Kong and J. Yang, "Remaining useful life prediction of rolling bearings based on RMS-MAVE and dynamic exponential regression model," *IEEE Access*, vol. 7, pp. 169705–169714, 2019, doi: 10.1109/ACCESS.2019.2954915.
[45] H. Wang, H. Liao, X. Ma, and R. Bao, "Remaining useful life prediction and optimal maintenance time determination for a single unit using isotonic regression and gamma process model," *Rel. Eng. Syst. Saf.*, vol. 210, Jun. 2021, Art. no. 107504, doi: 10.1016/j.ress.2021.107504.
[46] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Mining*, Pisa, Italy, Dec. 2008, pp. 413–422, doi: 10.1109/ICDM.2008.17.
[47] Y. Qin, J. Yang, J. Zhou, H. Pu, X. Zhang, and Y. Mao, "Dynamic weighted federated remaining useful life prediction approach for rotating machinery," *Mech. Syst. Signal Process.*, vol. 202, Nov. 2023, Art. no. 110688, doi: 10.1016/j.ymssp.2023.110688.
[48] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, and C. Varnier, "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in *Proc. IEEE Int. Conf. Prognostics Health Manag.*, Denver, CO, USA, Jun. 2012, pp. 1–8.
[49] B. Wang, Y. Lei, N. Li, and N. Li, "A hybrid prognostics approach for estimating remaining useful life of rolling element bearings," *IEEE Trans. Rel.*, vol. 69, no. 1, pp. 401–412, Mar. 2020, doi: 10.1109/TR.2018.2882682.

JING ZHANG was born in Xinzhou, Shanxi, China, in 1999. He received the B.S. degree from Jinzhong University, China, in 2022. He is currently pursuing the M.S. degree in mechanical engineering with the Inner Mongolia University of Science and Technology, China. His research interests include life prediction research and the condition monitoring of rotating machinery.

CHAO ZHANG was born in Anguo, Hebei, China, in 1978. He received the B.S. degree in mechanical engineering from the Inner Mongolia University of Science and Technology, China, in 2002, and the M.S. and Ph.D. degrees in mechatronics engineering from Xidian University, China, in 2008 and 2012, respectively.

Since 2018, he has been a Professor with the Mechanical Engineering Department, Inner Mongolia University of Science and Technology. His research interests include vibration signal processing, rotating machinery fault diagnosis and condition monitoring, and digital twinning technology. He was a member of the Fault Diagnosis Committee of the Chinese Society of Vibration Engineering.

SHUAI XU was born in Anshan, Liaoning, China, in 2000. She received the B.S. degree from the University of Science and Technology Liaoning, China, in 2022. She is currently pursuing the M.S. degree in mechanical engineering with the Inner Mongolia University of Science and Technology, China. Her research interests include life prediction research and the condition monitoring of rotating machinery.

GUIYI LIU was born in Luoyang, Henan, China, in 2000. He received the B.S. degree from the College of Arts and Information Engineering, Dalian Polytechnic University, China, in 2022. He is currently pursuing the M.S. degree in mechanical engineering with the Inner Mongolia University of Science and Technology, China. His research interest includes the study of fault diagnosis in variable speed rolling bearings.

HONGBO FEI was born in Jinzhou, Liaoning, China, in 1997. He received the B.S. degree from the University of Science and Technology Liaoning, China, in 2019. He is currently pursuing the M.S. degree in mechanical engineering with the Inner Mongolia University of Science and Technology, China. His research interests include signal processing and fault diagnosis of rotating machinery.

LE WU was born in Fuyang, Anhui, China, in 1995. He received the B.S. degree from Suzhou University, China, in 2020. He is currently pursuing the M.S. degree in mechanical engineering with the Inner Mongolia University of Science and Technology, China. His research interests include signal processing and fault diagnosis of rotating machinery.