


## Article

# Residual Life Prediction of Rolling Bearings Based on a CEEMDAN Algorithm Fused with CNN–Attention-Based Bidirectional LSTM Modeling

Xinggang Zhang <sup>1</sup>, Jianzhong Yang <sup>2,3,\*</sup>  and Ximing Yang <sup>1</sup><sup>1</sup> College of Naval Architecture and Ocean Engineering, Beibu Gulf University, Qinzhou 535011, China<sup>2</sup> College of Electronic and Information Engineering, Beibu Gulf University, Qinzhou 535011, China<sup>3</sup> College of Computer Science and Engineering, Guangxi Normal University, Guilin 535011, China

\* Correspondence: gxyjz@bbgu.edu.cn

**Abstract:** This paper presents a methodology for predicting the remaining usability of rolling bearings. The method combines a fully adaptive ensemble empirical modal decomposition of noise (CEEMDAN), convolutional neural network (CNN), and attention bidirectional long short-term memory network (ABiLSTM). Firstly, a finite number of intrinsic mode functions (IMFs) are obtained from breaking down the initial vibration signals using CEEMDAN. The IMFs are further screened by combining the correlation criterion and the craggy criterion. Then, time-frequency domain features, which are extracted from the screened IMFs, are reconstructed into a feature set. The SPT is recognized through some features, like the root mean square (RMS), variance, and kurtosis. Secondly, the deterioration character of rolling bearings was extracted using CNN and used to train the ABiLSTM network. Based on the output of the ABiLSTM network, it forecasts how long rolling bearings will last during use. Finally, the XJTU-SY rolling bearing dataset validated the validity of the suggested rolling bearing remaining life prediction method. We compare our algorithm with other algorithms, such as GRU, LSTM, and CNN–BiLSTM, in which the accuracy of MAE, MSE, RMSE, MAPE, and R2\_score is significantly improved. Thus, the results of the validation experiments demonstrate that our proposed algorithm has excellent prediction accuracy.



**Citation:** Zhang, X.; Yang, J.; Yang, X. Residual Life Prediction of Rolling Bearings Based on a CEEMDAN Algorithm Fused with CNN–Attention-Based Bidirectional LSTM Modeling. *Processes* **2024**, *12*, 8. <https://doi.org/10.3390/pr12010008>

Academic Editor: Iqbal M. Mujtaba

Received: 25 October 2023

Revised: 11 December 2023

Accepted: 15 December 2023

Published: 19 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** rolling bearings; CEEMDAN; remaining service life prediction; bidirectional long short-term memory networks; convolutional neural networks; attention mechanism

## 1. Introduction

Several parts cooperate in contemporary industrial production systems to accomplish predetermined objectives [1]. Rolling bearings, as basic components in industrial complex equipment, have a significant impact on the mechanical system's overall performance. Powerful predictive health-monitoring tools are needed to ensure the health of rolling bearings during operation [2]. Predictive operating condition monitoring tools indicate impending failures, providing ample lead time for maintenance programs. Nevertheless, it is worth noting that rolling bearings frequently encounter challenging and fluctuating working conditions, rendering them vulnerable to potential damage within engineering applications. The service life of rolling bearings might exhibit significant variation while being employed in conjunction with product components functioning under identical conditions. The ability to make precise predictions and accurate assessments of the remaining service life of rolling bearings is crucial for ensuring the smooth and effective functioning of mechanical equipment. Additionally, it enables the immediate detection and elimination of unexpected failure events [3]. Therefore, the primary objective of this study is to establish a system that can effectively forecast the remaining operational lifespan of bearings. This will be achieved by leveraging real-time sensor data and conducting a thorough evaluation of the trend in performance degradation, which holds significance.

Within the realm of contemporary research, there exist two primary methodologies for the prediction of the rolling bearing remaining useful life (RUL), like the methods based on failure physics (failure mechanism/mechanics) modeling and the data-driven life prediction methods. Since complex engineering equipment often consists of multiple subsystems and components, the corresponding failure mechanisms are difficult to obtain; therefore, the failure physics models are usually difficult to construct. However, the advancement of big data technologies and artificial neural networks has been characterized by rapid growth in recent years; moreover, artificial neural networks based on big data-driven artificial neural networks are being given an increasingly more important role in rolling bearing life prediction.

The approach of prediction based on models mostly relies on the in-depth study of the internal materials and structure of rolling bearings, as well as the operating conditions. The physical and mechanism characteristics of the interior materials, structure, and operating conditions of rolling bearings are further analyzed to build the prediction models [3]. There are three dominant approaches to model-based prediction, namely, mechanistic modeling, empirical modeling, and equivalent physical modeling [4]. Ruiz et al. proposed a discrete multi-physics model for simulating the fluid–structure interaction [5], and the acquired findings exhibited a higher degree of accuracy; nonetheless, the process and selection of the model parameters were quite intricate. Nevertheless, it is important to acknowledge that these three model-based forecasting approaches also possess notable limitations. On the one hand, they are strongly influenced by external environmental conditions, such as the temperature of the experiment, the fatigue of the material, and the change in the working conditions, which can affect the prediction results. On the other hand, this method cannot guarantee the effective extraction of rolling bearing features in complex environments, which typically results in the occurrence of modeling failures and an inability to accurately generate forecasts.

The utilization of a data-driven technique is, presently, the prevailing method employed by a multitude of individuals for the purpose of forecasting. The methods can be categorized into four main groups: optimization algorithm methods, sample entropy methods, machine learning methods, and performance characterization methods. The prevailing and widely adopted approaches primarily rely on machine learning techniques. These approaches can avoid the need to study the intricate internal configuration of rolling bearings. The health parameters for remaining useful life (RUL) prediction are derived from the original dataset using appropriate algorithms. There exists a multitude of distinct algorithms that are tailored for the implementation of this particular technique, such as the machine learning algorithms commonly used in research, including neural networks [6], support vector machines [7], Gaussian process regression, etc. In recent years, scholars have proposed a large number of methods based on vibration signal feature extraction and time series prediction, which have been successfully applied regarding the estimation of the remaining useful life (RUL) of bearings, which does not need to take into account mechanical structure, operating conditions, and failure mechanisms, and hence significantly enhancing the precision of residual life estimation. In situations characterized by a scarcity of available data. Kong et al. [8] proposed a rolling bearing remaining service life prediction based on the utilization of an adaptive time series feature window in conjunction with a multi-step forward method. Zhuang et al. [9] adopted a multi-source confrontational online regression in the presence of online unknown conditions for residual bearing life prediction. Yao et al. [10] presented a robust and lightweight neural network for the diagnosis of bearing faults. Yan et al. [11] brought forward a bearing fault diagnosis algorithm with the comprehensive lifecycle of a bearing, encompassing its progression from a state of normal functioning to eventual failure. The lifecycle was divided into several degradation stages, and the hidden Markov model combined with the support vector machine improved using particle swarm optimization was used to predict the degradation stages, respectively. Deep learning, through its multi-level internal structure and the advantages of re-learning training methods, can overcome the shortcomings of traditional manual feature extraction.

In addition, an obvious feature of RUL prediction is temporal correlation. Nevertheless, the conventional methods outlined above are incapable of leveraging the temporal correlation properties of time series data to facilitate more accurate and precise predictions of RUL. In order to tackle this issue, Han et al. [12] employed a methodology that involved the utilization of stacked auto-encoders and recurrent neural networks to make predictions regarding the remaining lifespan of bearings. Liu et al. [13] proposed an RUL prediction method for lithium-ion battery capacity based on the GRU method. Guo et al. [14] proposed the utilization of long short-term memory (LSTM) neural networks for the prediction of time series related to mechanical breakdown. The LSTM model is a type of recurrent neural network (RNN) that has demonstrated successful applications across diverse domains.

Zhu et al. [15] proposed a BiLSTM prediction model and integrated the backpropagation part in the original prediction model to improve prediction accuracy. However, due to the lack of further preprocessing and decomposition processing of the raw data, the overall prediction was ineffective. Regarding the specific instance of mechanical multi-structured data acquisition with multiple noises overlaid, multiple noises are often included in the acquired data, affecting the subsequent analysis and processing of the signals [16]. Guo et al. [17] proposed EMD combined with the LSTM algorithm. This algorithm first decomposes the signal into several modal components for prediction and then adds up the individual component predictions and reconstructs them to produce the prediction result. The algorithm has some improvement in RUL but suffers from modal aliasing. To solve this problem, Torres et al. [18] introduced an enhanced technique known as the entire empirical modal decomposition (EEMD) with adaptive noise ensemble empirical modal decomposition. This solution efficiently addresses the issue of moving white noise from higher to lower frequencies. The health factors utilized in this study are the capacity and equal-voltage-drop discharge duration, as proposed by Chen et al. [19]. To enhance the accuracy of the remaining useful life (RUL) forecast, the health factors were preprocessed using integrated empirical mode decomposition (EEMD). This preprocessing technique effectively mitigates the impact of capacity recovery and the noise present in the raw data. Subsequently, the preprocessed health factors were combined with support vector regression (SVR) to accomplish the RUL prediction. Lin et al. [20] used the CEEMDAN algorithm, and a refined algorithm for EEMD was developed to effectively address the issues of pattern aliasing and white noise elimination during the processing of the raw data. Combined with the LSTM model, the redundant rolling bearing vibration data are processed using the CEEMDAN. Subsequently, the processed data are fed into the long short-term memory (LSTM) model for the purpose of training. It was discovered to be significantly better than using the LSTM model alone. However, the related CEEMDAN algorithm has been less utilized for rolling bearing life prediction.

In the past, RUL prediction has been contingent upon a single CNN or LSTM. However, those methods cannot capture the abrupt changes which leads to poor prediction accuracy. Thus, some researchers have introduced the attention mechanism for improving their algorithm. For instance, Zhang et al. [21] presented a novel end-to-end RUL prediction and a proposed framework for the prediction of the estimation of the remaining lifespan of a rolling bearing using a convolutional recursive attention network. They adopted manual feature extraction that could only extract simple features; however, some complicated or deep features are easily ignored or hardly extracted. Mohd Saufi et al. [22] integrated an LSTM neural network with the Laplace methodology for accurately forecasting the remaining service life of mechanical components. Yao et al. [23] used the attention mechanism to improve 1D-CNN with simple recurrent units for the RUL of roller bearings. This model has a significant improvement in reducing prediction time while satisfying certain prediction accuracy criteria. To a certain degree, these techniques have enhanced the precision and accuracy when predicting the remaining lifespan of rolling bearings. Nevertheless, in more intricate work settings, and in today's rapidly evolving world, the precision and accuracy of these methods often necessitate enhancement to fulfill the requirements of practical engineering.

The paper's contributions can be outlined as follows:

- The rolling bearing RUL prediction model introduces the starting prediction point (SPT) identification structure component in the preprocessing stage. This helps reduce the amount of data used to train models for the initial phase of rolling bearing health life data. This results in shorter prediction times, giving projects a competitive advantage and increasing productivity.
- The preprocessed data are combined with the CEEMDAN algorithm to decompose the rolling bearing data into reasonable modal components and residuals and reconstruct them to obtain a better dataset.
- Accuracy: Overcoming the noise while combining traditional feature extraction with the deep learning feature extraction method so that the degraded features are extracted more fully. Meanwhile, in the prediction part, instead of simply increasing the number of network layers, an attentional mechanism structure is added to capture the mutation points and suppress the unnecessary parts, thus improving the accuracy of the prediction.

In this paper, a method for predicting the remaining life of rolling bearings is presented by combining techniques such as CEEMDAN, CNN, BiLSTM, and an attention mechanism. In this paper, preprocessing is performed before introducing the data into the neural network, such as normalization, sliding average filtering, and a more accurate selection of SPT points, which suppresses the noise of the original data to a certain extent. The preprocessed data are then combined with the CEEMDAN algorithm to decompose the rolling bearing data into reasonable modal components and residuals and reconstruct them to obtain a better dataset. Then, the advantage of the CNN model in deep feature extraction is utilized to enhance the predictability and effectiveness of traditional manual feature extraction. Finally, BiLSTM, with an attention mechanism, is utilized to predict the RUL of rolling bearings. The accuracy and precision improvement values of the RUL predictions are more pronounced in this paper compared to the state-of-the-art and classical methods used in the related literature within acceptable prediction timescales. This helps reduce calculation and forecasting times while improving the accuracy of RUL forecasts. Accordingly, this will give the project a competitive advantage and increase productivity.

The subsequent sections of this paper are organized in the following way. Section 2 describes the advantages of CEEMDAN, CNN, BiLSTM, and attention mechanisms, as well as the construction of the model and the corresponding evaluation metrics. Section 3 describes the acquisition of raw data and data preprocessing, as well as the detailed implementation process and results of the model. We also conduct a corresponding analysis to confirm the performance effect of the method by comparing it with classical and state-of-the-art methods. Finally, we summarize the paper in Section 4.

## 2. Model Construction

The utilization of CNN is prevalent owing to its robust capacity for feature extraction. BiLSTM is adept at processing long-term dependencies and exhibits superior performance on data with pronounced nonlinear relationships. Consequently, it is well-suited for handling extensive time series data. This section offers a concise overview of CEEMDAN, CNN, BiLSTM, attention mechanisms, and the model construction employed, as well as the performance metrics used in this paper.

### 2.1. CEEMDAN

The CEEMDAN algorithm, which stands for Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, was introduced by Torres et al. [18] in 2011 as an innovative method for signal decomposition. It can better solve the phenomenon of modal aliasing that exists in EEMD. The exact decomposition process is described below.

Step 1: The to-be-decomposed signal  $x(t)$  is added when the sequence  $x_i(t)$  is constructed using multiplying Gaussian white noise with a mean value of 0 by a factor of



$K$ , resulting in a total of  $K$  experiments. Here,  $i$  represents the index of each experiment, ranging from 1 to  $K$ . The calculation process is shown below.

$$x_i(t) = x(t) + \varepsilon \delta_i(t) \quad (1)$$

where  $\varepsilon$  is the Gaussian white noise weighting factor; the variable  $\delta_i(t)$  represents the Gaussian white noise that is generated during the processing of the  $i$ th component.

Step 2: The first modal component (IMF) and its mean value refer to the average value of the first IMF determined using the CEEMDAN technique. The specific calculation process is shown below.

$$IMF_1(t) = \frac{1}{K} \sum_{i=1}^K IMF_1^i(t) \quad (2)$$

$$r_i(t) = x(t) - IMF_1(t) \quad (3)$$

where  $IMF_1(t)$  means the first modal component derived from the CEEMDAN decomposition;  $r_i(t)$  is the residual signal that remains after the first decomposition.

Step 3: The EMD decomposition is continued after increasing particular noise to the stage  $j$  residual signal obtained after decomposition.

$$IMF_j(t) = \frac{1}{K} \sum_{i=1}^K E_1(r_{j-1}(t) + \varepsilon_{j-1} E_{j-1}(\delta_i(t))) \quad (4)$$

$$r_j(t) = r_{j-1}(t) - IMF_j(t) \quad (5)$$

where  $IMF_j(t)$  is the  $j$ th modal component derived from the CEEMDAN decomposition;  $E_{j-1}(\cdot)$  is the IMF component of the EMD decomposition of the sequence at index  $j - 1$ st;  $\varepsilon_{j-1}$  is a CEEMDAN weighting coefficient for increasing noise to the residual signal at stage  $j - 1$ ;  $r_j(t)$  represents phase  $j$  margin signals.

Step 4: The process of iteration comes to a halt. The CEEMDAN algorithm decomposition concludes when the EMD stopping condition is met and the residual signal  $n(t)$  of the  $n$ th decomposition exhibits monotonic behavior.

In addition, it is possible to select the IMF components obtained from the decomposition, especially for certain signals with localized tilts. The CEEMDAN algorithm removes localized tilts with the following selection rules. The specific calculation process is shown below.

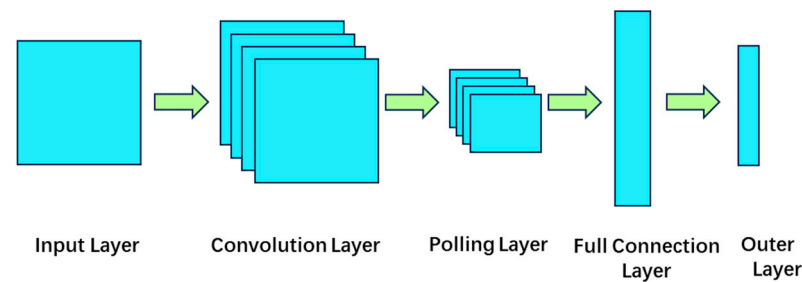
$$F = \frac{1}{n} \sqrt{\sum_{k=1}^n [I_k(i) - I(i)]^2} \quad (6)$$

In this context,  $i$  represents the variable denoting the number of data points, whereas  $n$  represents the variable representing the total number of data points;  $I_k(i)$  is the  $k$ th order IMF component;  $I(i)$  is the original signal.

The minimum value of  $F$  corresponds to the component signal that should be selected.

## 2.2. CNN

The convolutional neural network (CNN), equipped with a feedback mechanism [24], has high performance in the domain of large-scale image processing. Figure 1 shows a classical CNN mode. A CNN typically comprises three main components, namely, a convolutional layer, a pooling layer, and a fully connected layer [25].



**Figure 1.** Structure of the CNN model.

The convolutional layer of a CNN employs convolution as a means to decrease the dimensionality of the input image and extract relevant features. Convolutional operations are linear operations; however, neural networks necessitate the accommodation of nonlinear functions and the utilization of activation functions. In CNN, the SoftMax function is commonly employed in the context of classification tasks, while the sigmoid function is frequently utilized for regression purposes within CNN [26]. The primary purpose of the pooling layer [27] is to decrease the dimensionality of the picture following the convolution process. This reduction in dimensionality serves several academic objectives, including minimizing the model size, improving computational efficiency, mitigating the risk of overfitting, and enhancing the reliability of feature extraction. The fully connected layer serves as a classifier within the convolutional neural network, facilitating dimensionality reduction to preserve relevant information. One of the features of CNN models is the ability to derive the distinctive characteristics inherent in the input data, and it is necessary to extract local features [28].

### 2.3. BiLSTM

The main method of RUL prediction involves learning spatiotemporal information characterizing equipment performance degradation from mechanical equipment sensor data. RNN is capable of effectively using the temporal correlation inherent in sequential data for processing time domain sequences. However, when the length of a single time series datum is large or short, it is known as RNN, and will encounter challenges such as the issue of gradient vanishing and gradient explosion [29]. LSTM is a model of the network structure of memory cells implemented by Hochreiter [30] to address the investigation of gradient explosion and gradient vanishing as a subject of scholarly concern within the domain of deep learning. This issue specifically relates to the difficulties faced by recurrent neural networks while handling time series data of substantial duration.

BiLSTM is a neural network in which the states of both past and future hidden layers can be recursive with feedback. Therefore, it is good at uncovering hidden connections between time series data [31]. Figure 1 shows the outcomes of the model unfolding along the time axis for the BiLSTM network at moments  $1, 2, \dots, t$ , where  $x$  is the model input,  $\hat{A}_t$  is the status of the concealed layer, and  $h_t$  is the output. Because BiLSTM has concealed layers in both forward and inversion directions, the model can handle both onward and toward time flow directions. As shown in the figure, there is no interaction between the forward-propagating hidden layer and the backward-propagating hidden layer, which can be detached. They can be treated as two separate and opposite data flow networks.

Given the assumption that  $A$  represents the hidden layer state of the forward LSTM network at a specific time, the corresponding mathematical expression is depicted in Equation (7). The architecture under consideration can be seen as a single-layer LSTM network that computes the process of computing the state  $A_t$  at moment  $t$  from the state  $A_{t-1}$  at moment  $t-1$ , where  $x_t$  is the input at instant  $t$ .

$$A_t = LSTM(x_t, A_{t-1}) \quad (7)$$

where:

$A_t$ —hidden layer state of the forward LSTM network at time  $t$ ;  
 LSTM—LSTM unit;  
 $x_t$ —input at moment  $t$ ;  
 $A_{t-1}$ —hidden layer state of state-positive LSTM network at moment  $t - 1$ .

Similarly,  $A_t^*$  is the hidden layer state of the inverse LSTM network at moment  $t$ . Then, the computational equation is displayed in Equation (8).

$$A_t^* = LSTM(x_t, A_{t-1}^*) \quad (8)$$

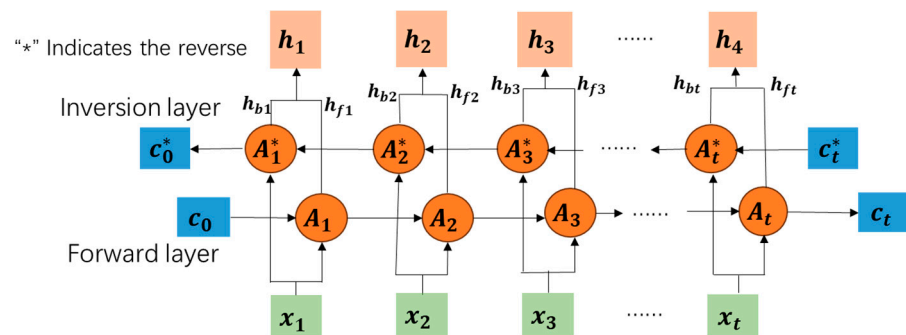
where:

$A_t^*$ —hidden layer state of the inverse LSTM network at time  $t$ ;  
 LSTM—LSTM unit;  
 $x_t$ —input at moment  $t$ ;  
 $A_{t-1}^*$ —hidden layer state of the inverse LSTM network at time  $t$ .

The BiLSTM network output is a combination of two parts of the hidden layer state  $A_t$  and  $A_t^*$ , thus constituting the aggregate latent state of the neural network  $\overset{\leftrightarrow}{A}_t$ .

By forming a forward propagation chain structure with some memory units, information can be coordinated and transmitted [32]. From Equation (4), the memory unit can precisely control the information flow. By integrating all moments of internal state information and input information, the stability of gradient descent during model training is guaranteed.

To enhance the extent of feature extraction from the initial time series and enhance the precision of the model's output, a BiLSTM network is constructed by combining two separate LSTMs with distinct orientations. The diagram illustrating the particular arrangement is depicted in Figure 2.



**Figure 2.** BiLSTM network structure diagram.

In Figure 2,  $x_t$  is input to the anterior layer, and the output  $h_f$  of the forward hidden layer is computed forward from the moment 0 to the moment  $t$ . Then, the input is fed to the reverse layer, and the output  $h_b$  of the backward implicit layer is computed from moment  $t$  to moment 0 in reverse. Finally, the outputs of the forward and reverse layers are passed via the fully linked layer in order to acquire the ultimate output  $h$ .

$$h = f(h_f, h_b) \quad (9)$$

where  $f()$  is mapping functions for the fully connected layer.

#### 2.4. Attention Mechanism

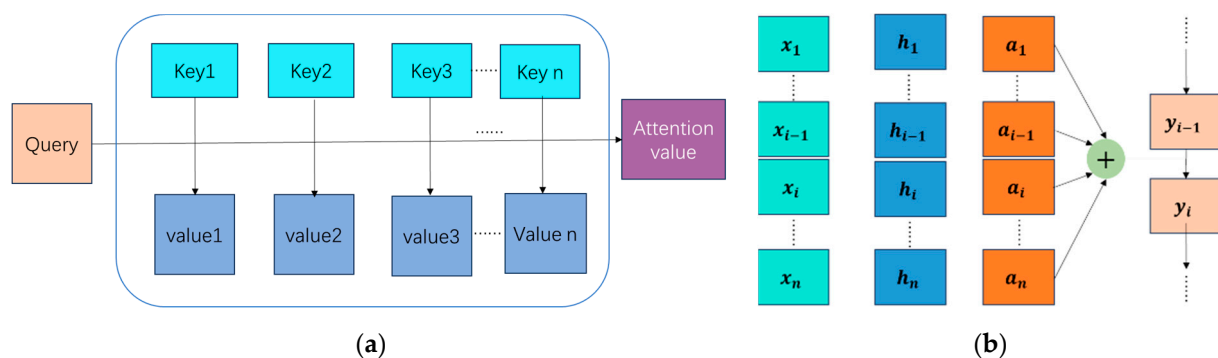
The attention mechanism (AM) has recently become widely used in various fields of deep learning. AM in deep learning is an approach that mimics the human visual and cognitive system by allowing neural networks to focus on relevant parts of the input data as it is processed. By adding the attention mechanism, neural networks can automatically learn

and selectively focus on important information in the input, improving the performance and generalization of the model.

The central idea of the attention mechanism [33] lies in the assignment of weights, namely, assigning a high weight to the important information to reasonably change the external attention to information, ignore irrelevant information, and amplify the desired information. In essence, the attention mechanism attempts to retrieve the values of items in the source through weighted summation; meanwhile, the query database and coefficient value key are utilized to generate the relevant weight coefficient:

$$A(\text{Query}, \text{Source}) = \sum_{i=1}^{L_x} \text{Similarity}(\text{Query}, \text{Key}_i) * \text{Value}_i \quad (10)$$

where  $L_x$  is the variable “length” and denotes the magnitude of the data source. Figure 3 illustrates the conceptual framework and corresponding structure of the attention mechanism.



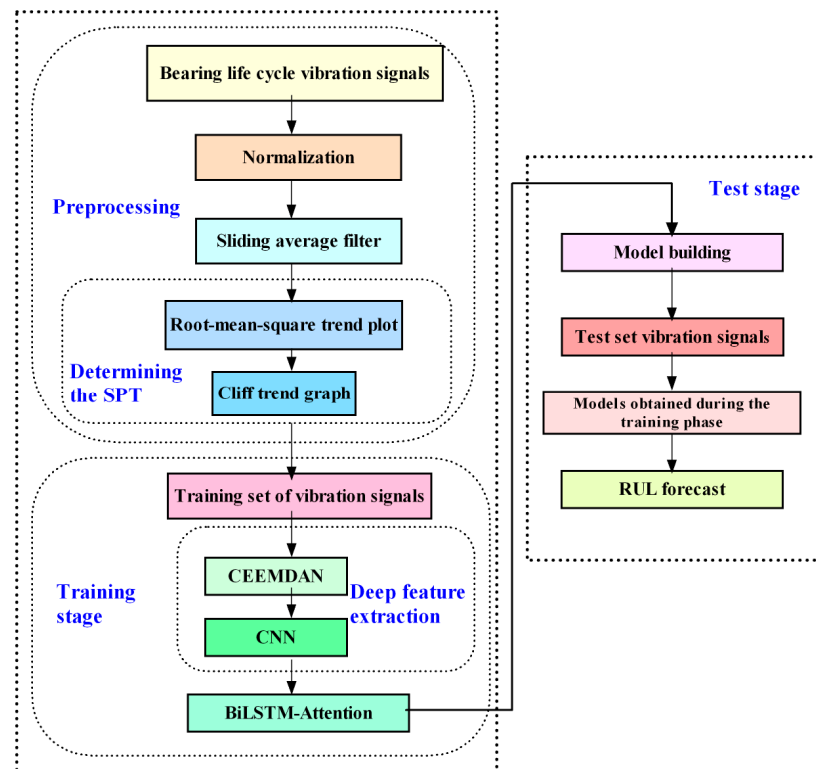
**Figure 3.** (a) The fundamental concept underlying the attention mechanism. (b) Attention mechanism.

### 2.5. Model Construction

Here, the advantage of the CNN multilevel perceptual machine structure was fully utilized, and the preservation of the properties of the original data was enhanced in order to address the challenge of feature extraction in the context of rolling bearing vibration data. The proposed method addressed the limitations of prior approaches in bearing the signal feature extraction that heavily relied on human expertise. BiLSTM in the net could make full use of forward and backward features to better predict future trends, which improved the accuracy of the prognostication of the remaining service life of rolling bearings. In this paper, a lifetime prediction model based on CEEMDAN–CNN–BiLSTM–attention was built, and the flowchart is shown in Figure 4. The relevant steps are as follows.

1. Collection of bearing vibration data. The provided data generally encompass details pertaining to the overall condition of the bearing in terms of its health, such as wear, cracks, etc.
2. The process of data preprocessing. The raw signals were normalized, and noise was reduced by sliding average filtering [34]. In this paper, the SPT is determined by calculating correlation characteristics such as standard deviation (SD), root mean square (RMS), and cliffside.
3. The vibration data were subjected to decomposition using the CEEMDAN technique, resulting in a set of intrinsic mode function (IMF) scores. The individual components are inputted into a convolutional neural network (CNN) in order to extract features, resulting in a feature vector for each intrinsic mode function (IMF). All IMF feature vectors were connected to form a complete feature vector.
4. Deep feature extraction. To enhance the extraction of deep features from the bearing vibration data, it is crucial to optimize the number of layer nodes in the four-layer CNN structure of the convolutional neural network.

5. Training stage or training phase. The optimized CNN-extracted deep features and related operational features are inputted into the ABiLSTM network, and the prediction algorithm is built using ABiLSTM training according to Equations (2)–(5) and the uniqueness of the storage unit structure of the BiLSTM–attention network. The unsupervised learning training process is, therefore, completed.
6. The study involves conducting tests. The partitioned test set was utilized as an input for the trained predictive model in order to derive the remaining useful life (RUL) values of the bearings.



**Figure 4.** The flowchart of CEEMDAN–CNN–ABiLSTM.

## 2.6. Performance Metrics

To ascertain the efficacy of the proposed predictive model, it is necessary to conduct a validation process, mean squared error (MSE), and root mean squared error (RMSE)), and the mean absolute error (MAE) and adjusted r-square ( $R^2\_score$ ) are chosen as the evaluation metric for the expected values of the remaining useful life (RUL), where  $y_i$  denotes the true observation,  $y_i^2$  denotes the mean value of the accurate observations, and  $\hat{y}_i$  denotes the predicted value. The aforementioned formulas can be represented as follows:

$$MSE = \frac{1}{N} \sum_{i=1}^N ((y_i - \hat{y}_i))^2 \quad (11)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |(y_i - \hat{y}_i)| \quad (12)$$

$$R^2\_score = 1 - \frac{\sum_{i=1}^N ((y_i - \hat{y}_i))^2 / N}{\sum_{i=1}^N ((y_i - y^2))^2 / N} \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N ((y_i - \hat{y}_i))^2} \quad (14)$$



$$MAPE = \frac{1}{n} \sum_{i=1}^n |(y_i - \hat{y}_i)/y_i| * 100\% \quad (15)$$

$$MA = 1 - MAPE \quad (16)$$

### 3. Experimental Process and Result

The operating system used in this experimental environment is Windows 11 with an operating system based on  $\times 64$  processor. The CPU is an Inter(R) Core (TM) i7-12700H 2.3 GHZ; the dedicated GPU is a NVIDIA GEFORCE RTX 3050 processor, and GPU parallel computation is adopted. The code is based on Python 3.10 and the compiler PyCharm 2022.2.1. The version of the virtual environment, Anaconda Navigator 3, is 2022.9.0. The deep learning framework was built using Keras 2.0.2 based on TensorFlow 2.10.0.

#### 3.1. Introduction to the Dataset and Data Preprocessing

##### 3.1.1. Raw Data Acquisition Platform

Experimental validation was conducted using the XJTU-SY dataset (as shown in Figure 5) [35]. It contains life data for 15 bearings under three different operating conditions. It is possible to obtain full lifecycle monitoring data for test bearings [36]. The test bearing is LDK UER204 rolling bearing (Deyuan Bearing Industrial Company Limited in Quanzhou, Fujian, China), and its important specifications are provided in Table 1. The bearing life experiments involved the development of three distinct operating situations, as outlined in Table 2. The many types of bearing failures and normal conditions are illustrated in Figure 6 presented below.

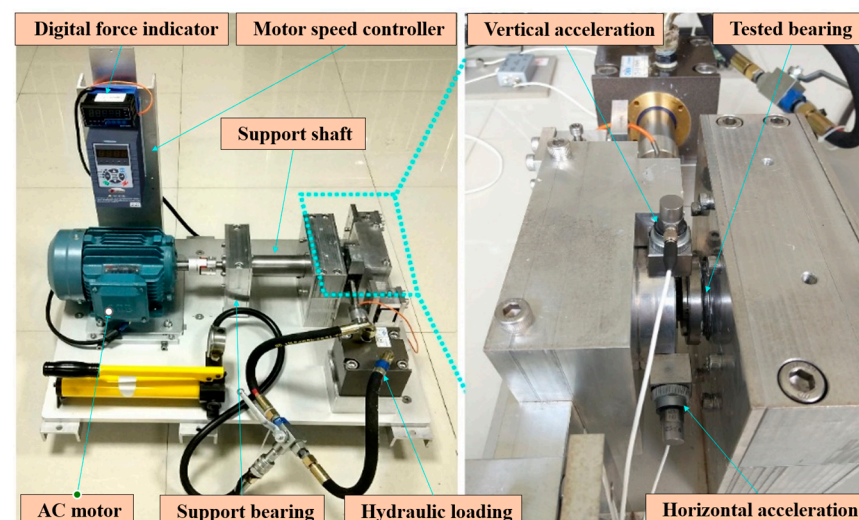


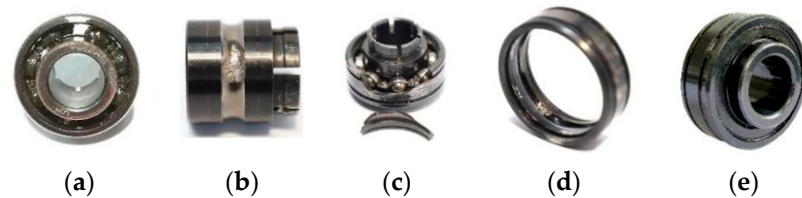
Figure 5. Bearing accelerated life test platform (provided in the XJTU-SY dataset [35]).

Table 1. LDK UER204 bearing parameters.

Parameter	Value
Raceway diameter of inner ring/mm	29.30
Raceway diameter of external ring/mm	39.80
Bearing medium diameter/mm	34.55
Basic dynamic load rating/N	12,820
Diameter of ball/mm	7.92
Number of balls	8
Contact angle/(°)	0
Basic static load rating/KN	6.65

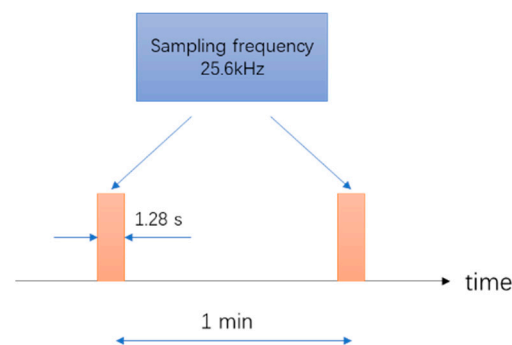
**Table 2.** Accelerated bearing life test conditions.

No.	1	2	3
Speed/(r/min)	2100	2250	2400
Radial force/KN	12	11	10

**Figure 6.** (a) Cage fracture, (b) inner race wear, (c) outer race fracture, (d) outer race wear, (e) normal bearing photograph [37].

### 3.1.2. Data Acquisition

Sensor PCB 352C33 (Manufactured by PCB Piezotronics Incorporated in Buffalo, New York, NY, USA) relevant test parameters are shown in Figure 7 [37]. The acquisition frequency during bearing operation is set to 25.6 kHz; the time between each acquisition is 1 min, and the time during acquisition is 1.28 s.

**Figure 7.** Sensor data sampling process [37].

### 3.1.3. Description of Data Samples

The provided data includes information pertaining to the tested bearing, such as its associated operating environment, the total number of data samples, the basic rated life  $L_{10}$ , the actual life, and the failure location.

The rated life of rolling bearings refers to the existence belonging to the identical category of bearing that can be reached or exceeded under the same operating conditions when the reliability is 90%. It is calculated using the following formula:

$$L_{10} = \frac{10^6}{60n} \left( \frac{C}{P} \right)^\varepsilon \quad (17)$$

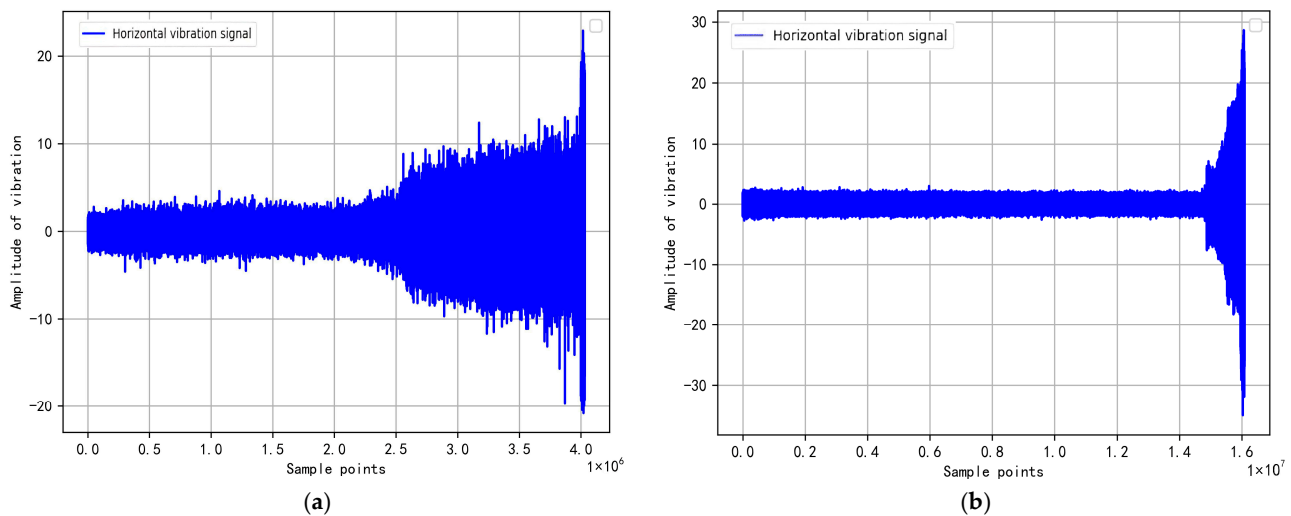
where  $L_{10}$  is the basic rated life,  $C$  is the rated dynamic load,  $n$  is the working speed of the bearing,  $\varepsilon$  is the life index. The relevant test bearings in this paper are ball bearings, and the  $\varepsilon$  reference standard value is 3 [38].

$P$  is the equivalent dynamic load. When the bearing is only subjected to radial load, it can be calculated using Equation (18).

$$P = f_P F_r \quad (18)$$

where  $f_P$  is the load factor and  $F_r$  is the radial load.

When there is no shock or a slight shock,  $F_p$  takes a value in the range of 1.0~1.2. The sample data of the raw vibration signal of bearing 1\_1 under working condition one and bearing 2\_1 under working condition two are shown in Figure 8.



**Figure 8.** Raw vibration signals accelerated experimental full lifecycle diagram of XJTU-SY bearing 1\_1 (a) and bearing 2\_1 (b).

### 3.1.4. Experimental Data Preprocessing

This section describes the normalization of the data and discusses the basis for the selection of SPT points and the associated feature selection.

#### Normalization

Avoiding the effect of differently scaled feature scales on prediction accuracy. In this paper, the vibration signals of the rolling bearing are normalized using the min–max normalization technique [34]. The formula for the min–max normalization process is as follows.

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (19)$$

where the variable  $x$  represents the initial vibration signal of the rolling bearing; the variable  $x_{min}$  represents the minimum value of the initial vibration signal sent by the rolling bearing; the variable  $x_{max}$  represents the highest magnitude observed in the initial vibration signal of the rolling bearing; the variable  $x_{new}$  represents the vibration signal of the rolling bearing that has been normalized.

#### Selection of the Start Prediction Point (SPT)

During the entire lifespan of a bearing, features are usually extracted by applying relevant methods such as signal processing to depict the progression of bearing deterioration from a state of normalcy to one of severe failure [35]. The accelerated full lifecycle vibration time-domain plots in the transverse and longitudinal directions for the working condition I bearing 1\_1 are shown in Figure 9. Figure 10 refers to the standard deviation and variance plots of the transverse and longitudinal vibration signals of bearing 1\_1 under working condition one. The relevant features of the rolling bearings 2\_1 and 3\_1 are shown in Figure 11. The RMS, variance, and kurtosis can better characterize the degradation trend of bearings during their lifetime. Moreover, the selection of the initial point significantly influences the succeeding levels of accuracy and precision in the estimate of the remaining lifespan, which, in the corresponding SPT, is derived from these calculations in this paper. This leads to the basis of the starting point and the benefits of the choice.

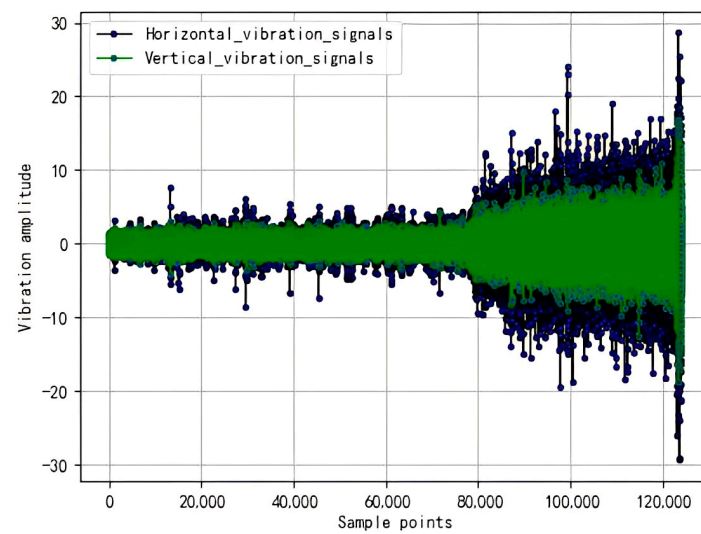


Figure 9. Raw vibration signal obtained from working condition bearing 1\_1 of the XJTUSY.

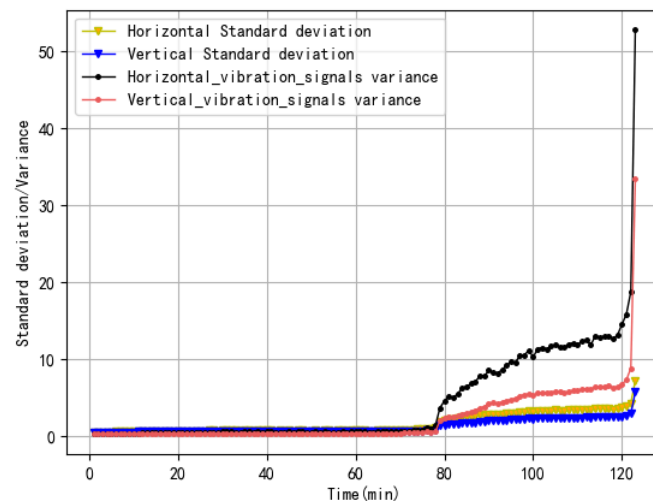


Figure 10. Bearing 1\_1 horizontal vertical standard deviation and variance.

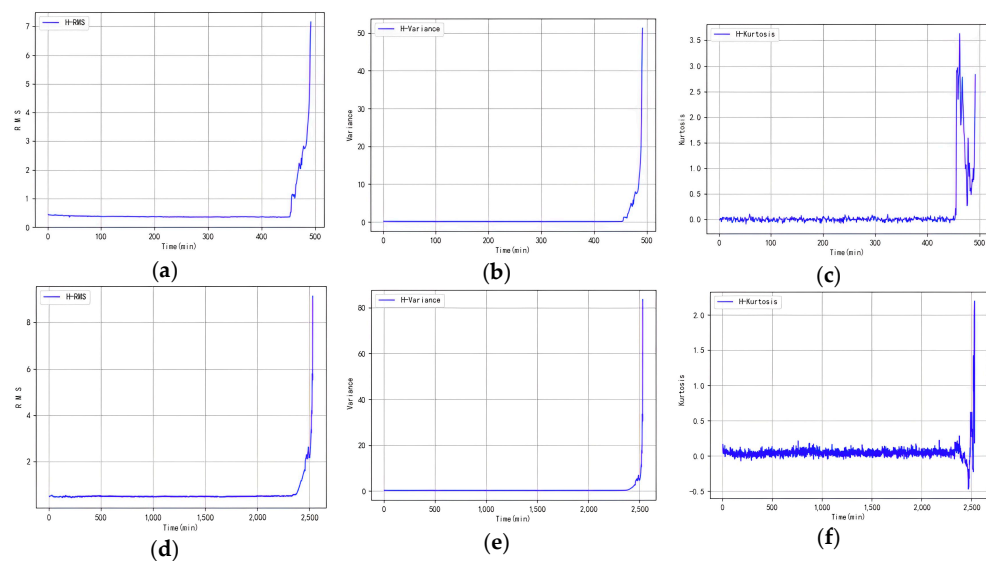


Figure 11. (a–c) The root mean square, variance, and kurtosis of bearing 2\_1. (d–f) The characteristics of root mean square, variance, and kurtosis of bearing 3\_1, respectively.

The parameters of the four parameter comparisons, such as the comparison results from different starting points, are shown in Table 3 below.

**Table 3.** Comparison of prediction from different starting points: Table 2 LDK.

Starting Point Setting	Experimental Time/s	RMSE	MAPE
2484	107.65	0.0242	0.1438
2384	132.41	0.0657	0.2825
2284	196.74	0.0632	0.1632
2184	232.57	0.2322	0.5631
2030	374.32	0.2635	0.4853
1777	486.43	0.3923	0.6855
1523	654.53	0.3504	0.5839
1269	744.82	0.5800	0.8615

### 3.1.5. Reconstructing Datasets

As shown in Table 3, the evaluation of the model involves the consideration of two pertinent parameters, namely, the root mean square error (RMSE) and the mean absolute percentage error (MAPE). The values of the root mean square error (RMSE) and mean absolute percentage error (MAPE) both fall within the interval of  $[0, +\infty)$ . A smaller RMSE value indicates a stronger regression impact; a lower number of MAPE indicates a higher level of accuracy in the prediction model. Moreover,  $y_i$  denotes the true observation,  $y_i^2$  denotes the average of the true observation, and  $\hat{y}_i$  denotes the predicted value. The relevant formula is shown in Equations (14) and (15). Bearing 3\_1 is chosen here as an example under operating condition 3.

The selection of SPT points has a large impact on both the experiment time and the accuracy and precision of rolling bearing RUL prediction, while other relevant parameters remain constant. In Table 3, the SPT value of 2484 has the shortest experimental time; however, its data do not contain all the characteristics of the degradation trend. This situation affects the accuracy and precision of RUL predictions. Within the acceptable time range, and to ensure the accuracy and precision of RUL prediction, the degradation features were concentrated in the late stage of the whole lifecycle, combining factors such as the relevant calculations of RMSE and MAPE. We chose data starting from the 2284th group and data after that as the dataset. Then, the datasets from group 2284 to group 2484 were used as the training set, and the rest of the datasets were used as the test set.

## 3.2. Experiment

### 3.2.1. Network Parameters Setting

#### CNN Module Parameters

The effectiveness of CNN network models for data feature extraction is strongly influenced by factors such as the number of convolutional layers and the number of nodes. The effect of the number of convolutional layers on CNN feature extraction was first determined, and the selection was made about the number of convolutional layers for the purpose of comparison.

The average accuracy (MA) is shown in Equation (16). The number of weights and the training time were chosen as the judgment criteria. The comparison results of the effects of different CNN layers on the prediction results are shown in Table 4. From Table 2, it can be seen that, when the convolutional model had four layers, the MA of the test samples was optimum. With the number of convolutional layers determined, the number of nodes in each convolutional layer affects the prediction accuracy and training and testing time.



**Table 4.** The influence of the number of CNN layers on the prediction results.

CNN Network Layers	MA	Number of Weights	Training Time/s
2	0.828	93,696	298.32
3	0.882	15,923	365.34
4	0.892	22,476	433.24
5	0.887	29,030	743.54
6	0.881	35,584	874.82

### Forecasting Module Parameters

Here, 250 sets of data after the SPT point in the bearing 3\_1 full lifecycle dataset are provided as inputs to the trained BiLSTM–attention network as a test set, and the BiLSTM–attention prediction output is obtained. The relevant parameters of the BiLSTM–attention network framework are shown in Table 5.

**Table 5.** BiLSTM–attention parameters.

Network Parameter	Setting Value
Number of input layers	4
Number of output layers	1
Number of hidden layers	4, 4
Network layer	4
Learning rate	0.01

After relevant experiments, the more desirable dimensional parameters were identified within an acceptable time frame. The window is 150, the epoch value is 350, the filters of Cov1D should be 256, and the units of BiLSTM should be 128.

### 3.2.2. Experimental Validations

The above data have been preprocessed. The IMF component after CEEMDAN decomposition is shown in Figure 12. Owing to the huge volume of data and the long training and testing time, we only extracted ten percent of the data in bearing 1\_1 as the data for this program to conduct relevant experiments. The processed data and CNN feature extraction data are shown in Figure 13. Experiments of significance were conducted utilizing this model, wherein a trade-off between training duration and performance was achieved by setting the epoch value to 350. Rolling bearing 1\_1 is used here as an example. The corresponding experiments are performed on the processed dataset. The observed loss values during the training and validation phases are visually represented in Figure 14.

### 3.3. Experimental Results

Uncertainty in forecast outcomes might arise due to the presence of estimation mistakes and unforeseen disturbances. The presence of uncertainty can result in suboptimal or infeasible solutions to deterministic optimization problems, and, in some cases, it can have significant consequences. It should be noted that (1) we do not have the conditions to collect the original dataset, and we used the XJTU-SY public dataset; therefore, there is no relevant discussion in this paper about the uncertainty of the real data collection in the complex environment; (2) when building the model, the parameters of a good correlation neural network tend to have an important impact on the results. We tried many parameters, and we finally determined excellent network parameters and achieved relatively good prediction results compared with other relevant algorithms, as well as achieving relatively good prediction results. (3) Regarding the uncertainty of the data itself, we carried out preprocessing when importing the data into the model, such as sliding average age filtering, combined with the CEEMDAN algorithm, which has a certain suppression effect on the sudden disturbance noise.

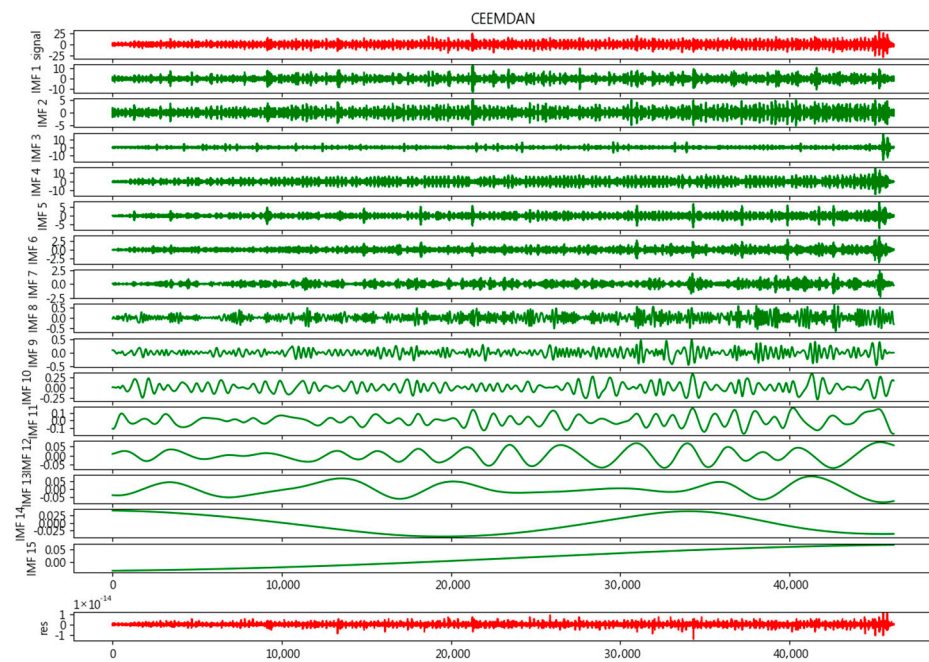


Figure 12. Raw vibration signal CEEMDAN of XJTU-SY bearing 1\_1.

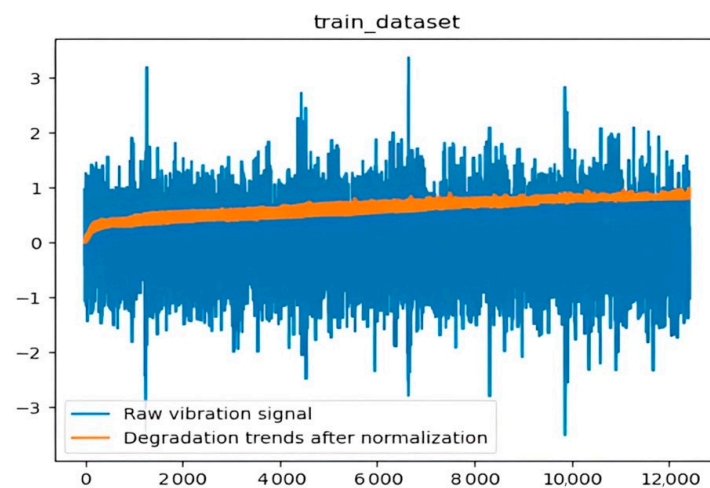


Figure 13. Raw vibration signal train dataset of XJTU-SY bearing 1\_1.

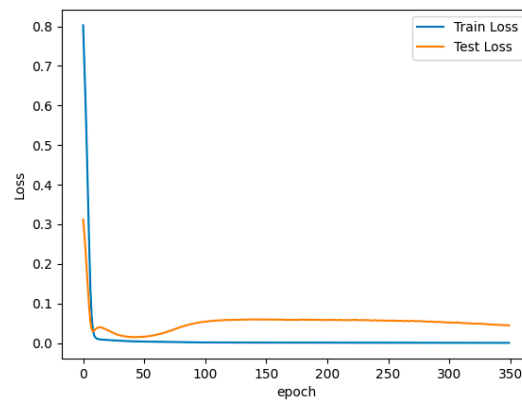
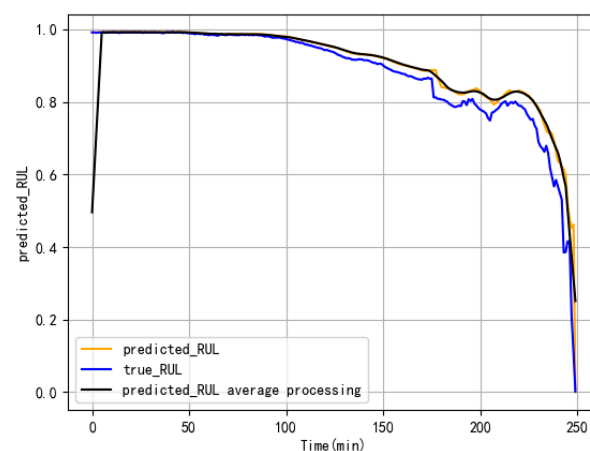


Figure 14. Raw vibration signal train and val loss of XJTU-SY bearing 1\_1.

Here, bearing 3\_1 is selected as the experimental object bearing. Since the degradation trend of the rolling bearing vibration data in the time series has a strong correlation with

the related features such as standard deviation, root mean square, and steepness, these characteristics are selected through relevant calculations. Therefore, these feature values are selected through correlation calculations, which are further used as representations of the relevant features to form the dataset required in this paper. After the normalization of this dataset, the sliding average filter, a more precise selection of SPT points, and raw data were processed into a better training dataset. That is, the sample data were used to derive the CNN.

Through the related experiments, the experimental results are obtained as shown in Figure 15. The blue line represents the real RUL of the rolling bearing, the orange line represents the predicted line, and the black line represents the effect of the predicted RUL after average processing. The MAE, MSE, RMSE, and MAPE of the remaining life prediction based on the CEEMDAN–CNN–ABiLSTM model are 0.84%, 0.09%, 3.02%, and 12.75%, respectively.

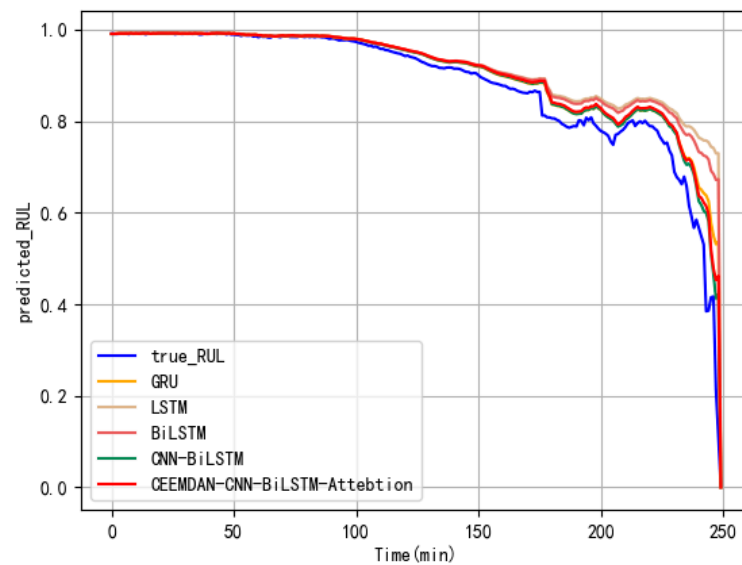


**Figure 15.** CEEMDAN–CNN–ABiLSTM–attention network test results.

From Figure 15, the model presented in this research exhibits a similar trend to the actual data when predicting the remaining useful life (RUL) of the bearing. The correlation between the anticipated and actual data is strong, indicating a high degree of curve fitting. Additionally, the discrepancy or error between the two datasets is minimal. The fluctuation of the predicted values in the late stage of the rolling bearing operation is small, which indicates that the characteristic values selected in this paper can effectively characterize the degradation state of rolling bearings and also indicates that the method of predicting the remaining life of the efficacy of the bearings presented in this research is demonstrated. The method employed in this study accurately predicts the degradation trend of the rolling bearing RUL, demonstrating the feasibility of the approach presented in this work.

### 3.4. Comparison and Analysis with Other Methods

To ascertain the efficacy of deep feature extraction compared to manual feature extraction and the superiority of the BiLSTM network that can fully utilize the degraded historical data with temporal correlation, the GRU neural network, traditional LSTM, BiLSTM, and CNN–BiLSTM were further compared with the model described in this paper. The prediction results of each method are shown in Figure 16. The prediction errors of the six methods are listed in Table 6.



**Figure 16.** Comparison results of four methods.

**Table 6.** Prediction error of six methods (%).

Methodology	MAE	MSE	RMSE	MAPE	$R^2$ _Score
Baseline [23]	11.76	2.16	14.71	-	-
GRU	3.28	0.53	7.16	20.50	70.76
LSTM	3.80	0.74	9.62	20.93	73.25
BiLSTM	3.40	0.57	8.32	19.12	75.15
CNN-BiLSTM	1.44	0.14	4.50	15.63	92.63
CEEMDAN-CNN-BiLSTM-attention	0.84	0.09	3.02	12.75	96.05

Note: “-” means the data are not provided.

### 3.4.1. Predictive Results of Different Models

As can be seen from Figure 16, the GRU neural network, traditional LSTM, and BiLSTM models show large dispersion compared with the real values at the late stage of the bearing run, which indicates that these three methods cannot fully exploit the intrinsic connection of the time series data. What is more, the CNN-BiLSTM network model showed a better fit compared to the first three models but exhibited a certain degree of discretization at the very late stage of the bearing operation. This is because manual feature extraction cannot extract deep features. The model presented in this study, namely, the CEEMDAN-CNN-BiLSTM-attention model, has superior performance in terms of accuracy and fit compared to the other four models. The predicted values of the methodology employed in this research demonstrate consistency with the actual values, which proves the feasibility of this paper’s method.

### 3.4.2. Analysis of the Prediction Results of Different Models

As can be seen from Table 6, under the premise of meeting a certain prediction timeframe, the remaining life prediction precision and accuracy of the associated rolling bearings have been more significantly improved compared with the latest method [23] and the related classical models. The MAE, MSE, RMSE, and MAPE are, respectively, 0.84%, 0.09%, 3.02%, and 12.75% of the remaining life prediction based on the CEEMDAN-CNN-BiLSTM model, and they are smaller than those of several other methods. The  $R^2$ \_score of the method used in this paper is 96.05%, which is the highest among all schemes, indicating a good model fit. Furthermore, it can be concluded from the aforementioned statement that the suggested methodology possesses the capability to accurately and precisely forecast the RUL of rolling bearings.

### 3.4.3. Ablation Analysis

Based on the findings presented in Table 6, it can be inferred that the MAE, MSE, RMSE, and MAPE of the BiLSTM model exhibited reductions of 0.4%, 1.8%, 1.3%, and 1.81%, respectively, compared to the LSTM model. After adding the CNN structure, the CNN–BiLSTM model decreased by 1.9%, 0.43%, 3.82%, and 3.49% in terms of MAE, MSE, RMSE, and MAPE, respectively, compared to the BiLSTM model. After combining the CEEMDAN noise reduction effect and the advantages of the attention algorithm, the model in this paper reduced by 0.6%, 0.05%, 1.48%, and 2.88% in MAE, MSE, RMSE, and MAPE, respectively, compared with CNN–BiLSTM. However, in terms of the  $R^2\_score$ , there is an improvement of 1.9%, 17.48%, and 3.42% in that respective order. This observation suggests that the model presented in this research study exhibits a notable enhancement in the precision and accuracy of predicting the RUL of rolling bearings.

### 3.4.4. Analysis of Running Time

Orthogonal tables are uniformly dispersed and neatly comparable. It selects some representative parameter combinations for experimentation in a comprehensive experiment. It is also one of the methods used to study multi-factorial and multilevel items. Here, the  $L_{16}$  ( $5^4$ ) orthogonal table is used.

Five influential parameters, window, epoch, Cov1D (filters), BiLSTM (units), and SPT, are used here. Some experiments were re-added. Four typical levels were used for each of the five factors, and time results were obtained as shown in Table 7.

**Table 7.** Solution time case in larger dimensions.

Dimensions	Factors					Results
	Window	Epoch	Cov1D (Filters)	BiLSTM (Units)	SPT	Times/s
1	5	350	64	64	1523	37.52
2	5	1000	128	128	2030	71.09
3	5	2000	256	256	2284	235.14
4	5	3000	512	512	2384	312.91
5	20	350	64	256	2384	104.29
6	20	1000	32	512	2284	141.36
7	20	2000	256	64	2030	167.11
8	20	3000	128	128	1523	667.22
9	50	350	128	512	2030	83.90
10	50	1000	256	256	1523	241.37
11	50	2000	32	128	2384	101.23
12	50	3000	64	64	2284	131.06
13	150	350	256	128	2284	34.80
14	150	1000	128	64	2384	51.29
15	150	2000	64	512	1523	744.82
16	150	3000	32	256	2030	351.83

The results show that the value of time is minimum for combination 13 followed by combination 1. Therefore, it is considered here that, to obtain the minimum time, the optimal state when the window is 150, epoch is 350, the filters of Cov1D should be 256, the units of BiLSTM should be 128, and SPT should be 2284. It can be seen that the solution time of the model in more parameters increased with an increase in epoch, Cov1D filters, and units in BiLSTM. However, as the SPT increases, the number of values in the dataset decreases and the time decreases accordingly. At the same time, the window parameter values have little effect on the model runtime.



#### 4. Conclusions

In this research, a rolling bearing residual life prediction model based on the CEEMDAN–CNN–ABiLSTM model was suggested. This model combined the sliding average filtering and incorporated the benefits of CEEMDAN filtering and noise reduction in data processing, CNN extraction deep feature, and the properties of processing time series using BiLSTM, which enhanced the precision of the RUL prediction of rolling bearings. The following specific conclusions were drawn based on the analysis of the experiments related to this paper.

1. The SPT points for bearing life prediction were determined by the trend graphs of characteristics such as RMS and steepness during the entirety of the bearings' lifetime to improve the accuracy of the prediction.
2. The CEEMDAN entered the CNN network after filtering and noise reduction in the data processing process, which could adequately extract deep features characterizing the bearing degradation state from high-dimensional complex data. The previous steps combined the advantage of the autonomous learning capability of BiLSTM–attention networks on time series data to enhance the precision of the forecast and effectively employed degradation history data over time to obtain the degradation state and efficiently implement the nonlinear function mapping.
3. After comparing with other prediction methods, it is found that the bearing degradation prediction approach using CEEMDAN–CNN–BiLSTM–attention demonstrates superior performance compared to the other five methods in terms of its predictive accuracy.

In future studies, we may further optimize the model to reduce the time spent by the model in the training process and perform model lightweighting to reduce storage resources. This will help improve the reliability of rolling bearing remaining life prediction and save storage resources and costs. In addition, we can work on improving the information integration capability of the model to further improve the front-end processing and back-end prediction of the data. Thus, a more lightweight, economical, and intelligent idea of the rolling prediction of the remaining life in bearings can be constructed. Additionally, it is imperative to incorporate novel mechanisms into existing models to expand the influence of significant temporal data on the results.

**Author Contributions:** Conceptualization and methodology, J.Y. and X.Z.; data curation X.Y.; writing—original draft preparation, X.Z.; writing—review and editing, J.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data involved in this paper have been presented in the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Rai, A.; Upadhyay, S.H. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribol. Int.* **2016**, *96*, 289–306. [\[CrossRef\]](#)
2. El-Thalji, I.; Jantunen, E. A summary of fault modelling and predictive health monitoring of rolling element bearings. *Mech. Syst. Signal Process.* **2015**, *60–61*, 252–272. [\[CrossRef\]](#)
3. Sadabadi, K.K.; Jin, X.; Rizzoni, G. Prediction of remaining useful life for a composite electrode lithium-ion battery cell using an electrochemical model to estimate the state of health. *J. Power Source* **2021**, *481*, 228861. [\[CrossRef\]](#)
4. Liao, L.; Köttig, F. A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction. *Appl. Soft Comput.* **2016**, *44*, 191–199. [\[CrossRef\]](#)
5. Ruiz-Riancho, I.N.; Alexiadis, A.; Zhang, Z.; Garcia Hernandez, A. A Discrete Multi-Physics Model to Simulate Fluid Structure Interaction and Breakage of Capsules Filled with Liquid under Coaxial Load. *Processes* **2021**, *9*, 354. [\[CrossRef\]](#)
6. Hu, P.; Zhao, C.; Huang, J.; Song, T. Intelligent and Small Samples Gear Fault Detection Based on Wavelet Analysis and Improved CNN. *Processes* **2023**, *11*, 2969. [\[CrossRef\]](#)
7. Zhang, T.; Yu, M.; Li, B.; Liu, Z. Capacity prediction of lithium-ion batteries based on wavelet noise reduction and support vector machine. *Trans. China Electrotech. Soc.* **2020**, *35*, 3126–3136. [\[CrossRef\]](#)

8. Kong, W.; Li, H. Remaining useful life prediction of rolling bearing under limited data based on adaptive time-series feature window and multi-step ahead strategy. *Appl. Soft Comput.* **2022**, *129*, 109630. [\[CrossRef\]](#)
9. Zhuang, J.; Cao, Y.; Jia, M.; Zhao, X.; Peng, Q. Remaining useful life prediction of bearings using multi-source adversarial online regression under online unknown conditions. *Expert Syst. Appl.* **2023**, *227*, 120276. [\[CrossRef\]](#)
10. Yao, D.; Liu, H.; Yang, J.; Li, X. A lightweight neural network with strong robustness for bearing fault diagnosis. *Measurement* **2020**, *159*, 107756. [\[CrossRef\]](#)
11. Yan, M.; Wang, X.; Wang, B.; Chang, M.; Muhammad, I. Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model. *ISA Trans.* **2020**, *98*, 471–482. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Luo, J.; Liu, Z.; Wang, J.; Chen, H.; Zhang, Z.; Qin, B.; Cui, S. Effects of Different Injection Strategies on Combustion and Emission Characteristics of Diesel Engine Fueled with Dual Fuel. *Processes* **2021**, *9*, 1300. [\[CrossRef\]](#)
13. Liu, H.; Li, Y.; Luo, L.; Zhang, C. A Lithium-Ion Battery Capacity and RUL Prediction Fusion Method Based on Decomposition Strategy and GRU. *Batteries* **2023**, *9*, 323. [\[CrossRef\]](#)
14. Guo, J.; Lao, Z.; Hou, M.; Li, C.; Zhang, S. Mechanical fault time series prediction by using EFMSAE-LSTM neural network. *Measurement* **2021**, *173*, 108566. [\[CrossRef\]](#)
15. Zhu, Y.-F.; He, W.-W.; Li, J.-X.; Li, Y.; Li, P. SOC estimation for Li-ion batteries based on Bi-LSTM and Bi-GRU. *Energy Storage Sci. Technol.* **2021**, *10*, 1163–1176. [\[CrossRef\]](#)
16. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. An unconstrained optimization approach to empirical mode decomposition. *Digit. Signal Process.* **2015**, *40*, 164–175. [\[CrossRef\]](#)
17. Guo, R.; Wang, Y.; Zhang, H.; Zhang, G. Remaining useful life prediction for rolling bearings using EMD-RISI-LSTM. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12. [\[CrossRef\]](#)
18. Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 4144–4147. [\[CrossRef\]](#)
19. Chen, L.; Zhang, Y.; Zheng, Y.; Li, X.; Zheng, X. Remaining useful life prediction of lithium-ion battery with optimal input sequence selection and error compensation. *Neurocomputing* **2020**, *414*, 245–254. [\[CrossRef\]](#)
20. Lin, Y.; Yan, Y.; Xu, J.; Liao, Y.; Ma, F. Forecasting stock index price using the CEEMDAN-LSTM model. *N. Am. J. Econ. Financ.* **2021**, *57*, 101421. [\[CrossRef\]](#)
21. Zhang, Q.; Ye, Z.; Shao, S.; Niu, T.; Zhao, Y. Remaining useful life prediction of rolling bearings based on convolutional recurrent attention network. *Assem. Autom.* **2022**, *42*, 372–387. [\[CrossRef\]](#)
22. Mohd Saufi, M.S.R.; Hassan, K.A. Remaining useful life prediction using an integrated Laplacian-LSTM network on machinery components. *Appl. Soft Comput.* **2021**, *112*, 107817. [\[CrossRef\]](#)
23. Yao, D.; Li, B.; Liu, H.; Yang, J.; Jia, L. Remaining useful life prediction of roller bearings based on improved 1D-CNN and simple recurrent unit. *Measurement* **2021**, *175*, 109166. [\[CrossRef\]](#)
24. Mao, W.; He, J.; Tang, J.; Li, Y. Predicting remaining useful life of rolling bearings based on deep feature representation and long short-term memory neural network. *Adv. Mech. Eng.* **2018**, *10*, 1687814018817184. [\[CrossRef\]](#)
25. Xia, T.; Song, Y.; Zheng, Y.; Pan, E.; Xi, L. An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation. *Comput. Ind.* **2020**, *115*, 103182. [\[CrossRef\]](#)
26. Liu, C.; Zhu, L. A two-stage approach for predicting the remaining useful life of tools using bidirectional long short-term memory. *Measurement* **2020**, *164*, 108029. [\[CrossRef\]](#)
27. Pei, H.; Si, X.-S.; Hu, C.-H.; Zheng, J.-F.; Li, T.-M.; Zhang, J.-X.; Pang, Z.-N. An adaptive prognostics method for fusing CDBN and diffusion process: Application to bearing data. *Neurocomputing* **2021**, *421*, 303–315. [\[CrossRef\]](#)
28. Cheng, N.; Chen, D.; Lou, B.; Fu, J.; Wang, H. A biosensing method for the direct serological detection of liver diseases by integrating a SERS-based sensor and a CNN classifier. *Biosens. Bioelectron.* **2021**, *186*, 113246. [\[CrossRef\]](#)
29. Sastrawan, I.K.; Bayupati, I.P.A.; Arsa, D.M.S. Detection of fake news using deep learning CNN–RNN based methods. *ICT Express* **2022**, *8*, 396–408. [\[CrossRef\]](#)
30. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#)
31. Varadharajan, S.K.; Nallasamy, V. P-SCADA-a novel area and energy efficient FPGA architectures for LSTM prediction of heart arrhythmias in BIoT applications. *Expert Syst.* **2022**, *39*, e12687. [\[CrossRef\]](#)
32. Ecker, L.; Schlacher, K. An approximation of the Bayesian state observer with Markov chain Monte Carlo propagation stage. *IFAC-PapersOnLine* **2022**, *55*, 301–306. [\[CrossRef\]](#)
33. Tian, Q.; Wang, H. An Ensemble Learning and RUL Prediction Method Based on Bearings Degradation Indicator Construction. *Appl. Sci.* **2020**, *10*, 346. [\[CrossRef\]](#)
34. Mouellef, M.; Szabo, G.; Vetter, F.L.; Siemers, C.; Strube, J. Artificial Neural Network for Fast and Versatile Model Parameter Adjustment Utilizing PAT Signals of Chromatography Processes for Process Control under Production Conditions. *Processes* **2022**, *10*, 709. [\[CrossRef\]](#)
35. Wang, B.; Lei, Y.; Li, N.; Li, N. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [\[CrossRef\]](#)
36. Lei, Y.; He, Z.; Zi, Y.; Hu, Q. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. *Mech. Syst. Signal Process.* **2007**, *21*, 2280–2294. [\[CrossRef\]](#)

- 
37. Lei, Y.; Han, T.; Wang, B.; Li, N.; Yan, T.; Yang, J. XJTU-SY Rolling Element Bearing Accelerated Life Test Dataset: A Tutorial. *J. Eng. Mech.* **2019**, *55*, 1–6. Available online: <http://www.cjmenet.com.cn/CN/10.3901/JME.2019.16.001> (accessed on 10 September 2023).
  38. *BS ISO 281:2007; Rolling Bearings-Dynamaic Load Ratings and Rating Life*, 2nd ed. BS ISO: London, UK, 2007.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.