

Article

Data-Driven Method for Predicting Remaining Useful Life of Bearing Based on Bayesian Theory

Tianhong Gao ¹, Yuxiong Li ¹ , Xianzhen Huang ^{1,2,*} and Changli Wang ¹

¹ School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819, China; gaotianhong960315@163.com (T.G.); 13354260226@163.com (Y.L.); a15698841933@163.com (C.W.)

² Key Laboratory of Vibration and Control of Aero-Propulsion System Ministry of Education of China, Northeastern University, Shenyang 110819, China

* Correspondence: xzhuang@mail.neu.edu.cn

Abstract: Bearings are some of the most critical industrial parts and are widely used in various types of mechanical equipment. Bearing health status can have a significant impact on the overall equipment performance, and bearing failures often cause serious economic losses and even casualties. Thus, estimating the remaining useful life (RUL) of bearings in real time is of utmost importance. This paper proposes a data-driven RUL prediction method for bearings based on Bayesian theory. First, time-domain features are extracted from the bearing vibration signal and data are fused to build a health indicator (HI) and a state model of bearing degradation. Then, according to Bayesian theory, a Bayesian model of state parameters and bearing life is established. The parameters of the Bayesian model are updated and bearing RUL is predicted by the Metropolis–Hastings algorithm. The method was validated by the XJTU-SY bearing open datasets and the prediction results are compared with the existing methods. Accuracy of the proposed method was demonstrated.

Keywords: data-driven method; Bayesian model; Metropolis–Hastings algorithm; remaining useful life prediction



Citation: Gao, T.; Li, Y.; Huang, X.; Wang, C. Data-Driven Method for Predicting Remaining Useful Life of Bearing Based on Bayesian Theory. *Sensors* **2021**, *21*, 182. <https://doi.org/10.3390/s21010182>

Received: 3 December 2020

Accepted: 24 December 2020

Published: 29 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 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/>).

1. Introduction

Bearings are some of the most basic yet critical components used in the manufacturing industry and the overall performance and reliability of mechanical equipment are closely related to bearing performance [1]. Although bearings are the most commonly used components in mechanical equipment, bearings are also the most susceptible to failure [2]. Moreover, abnormal operating states can seriously affect production activities and may even lead to catastrophic consequences; therefore, predicting bearing remaining useful life (RUL) is of both theoretical and practical value. Since bearing RUL prediction is an important element of equipment prognostics and health management, extensive research has been carried out. Condition monitoring and RUL prediction of bearings in operation can help guide timely and reasonable maintenance, prolong service life of equipment, improve reliability of mechanical systems, and avoid catastrophic accidents caused by bearing damage [3,4].

Current methods for RUL prediction can be divided into two categories: model-based and data-driven methods. Model-based methods typically establish a degradation model according to the physical structure of the bearing, which is then used to predict the RUL of the bearing [5,6]. Jiang [7] proposed a prediction method for RUL based on the convex optimization-life parameter degradation mechanism model. Sun [8] established a Hertzian contact dynamic theory model of a bearing ball and raceway and showed that optimal damping can improve bearing life. Model-based methods require an accurate degradation model; however, complex structure of components and operation mechanisms, as well as environmental uncertainties in engineering practice, make it difficult to establish an accurate model [9]. Data-driven methods mainly rely on machine learning and deep

learning algorithms to predict bearing RUL in the absence of a physical model of system degradation. This type of method can be universally applied for cases where the physical system cannot be accurately modeled. Numerous effective data-driven methods have been developed for RUL prediction. Wu [10] introduced long short-term memory (LSTM) networks to realize high-precision RUL prediction for complicated industrial objects. Ren [11] proposed a bearing RUL prediction method based on a deep neural network (DNN) and deep autoencoder. Xia [12] presented an innovative two-stage automated approach using a DNN to accurately determine the RUL of bearings. Li [13] constructed a modified health index based hierarchical gated recurrent unit network to improve the accuracy of bearing RUL prediction. Both supervised and unsupervised learning have also been applied in bearing fault diagnosis with good results [14–17].

Data-driven methods can overcome difficulties associated with model construction and can achieve more accurate prediction results. However, uncertainties still exist in practical applications, such as uncertainties in material properties, measurement errors, processing technologies, and operating conditions, which are often ignored. Owing to the high costs of product testing, and limitations of existing data transmission and storage technologies, sufficient data on the typical life cycle of bearings are usually unavailable. Therefore, in practice, data are insufficient to support data-driven prediction methods. Realizing a more accurate bearing RUL real-time prediction with limited bearing vibration signal data and considering the random factors is of great difficulty in related research. Bayesian theory is an effective method for data analysis with uncertain factors, which are regarded as random parameters. Expert knowledge, theoretical analyses, and historical data are used to obtain probability distributions of certain parameters (i.e., prior distributions). Then, updating methods are used to transform real-time data into more accurate distribution information (i.e., posterior distributions). Thus, the quantitative method of uncertainty based on the Bayesian theory has great research value in the field of RUL prediction. Mosallam [18] proposed a Bayesian approach for predicting the RUL of key components in systems with variable operating conditions. Cheng [19] presented a prediction method based on functional principal component analysis and the Bayesian method for Li-ion batteries RUL evaluation. Liu [20] proposed a dynamic data-driven layered Bayesian degradation model to tackle structural damage growth prediction. Tang [21] introduced a Bayesian Monte Carlo method to predict the aging trajectory of Li-ion batteries with significantly reduced experimental tests. Li [22] proposed a sequential Bayesian, which updated the Wiener process model improved the accuracy of RUL prediction. Martha [23] introduced a Bayesian hierarchical model to estimate the RUL of civil aerospace gas turbine engines.

This paper proposes a new data-driven bearing RUL prediction method based on Bayesian theory. A flowchart of the prediction process is illustrated in Figure 1. First, time-domain features are extracted from training bearing vibration signals and are screened. Standardization and dimensional reduction are carried out to build an appropriate health indicator (HI). Then, a state model of the bearing degradation process is established based on the processed data and a Bayesian model of state parameters and bearing life is constructed. Finally, parameters of the Bayesian model are updated according to real-time bearing data using the Metropolis–Hastings (M-H) algorithm to realize real-time prediction of bearing RUL.

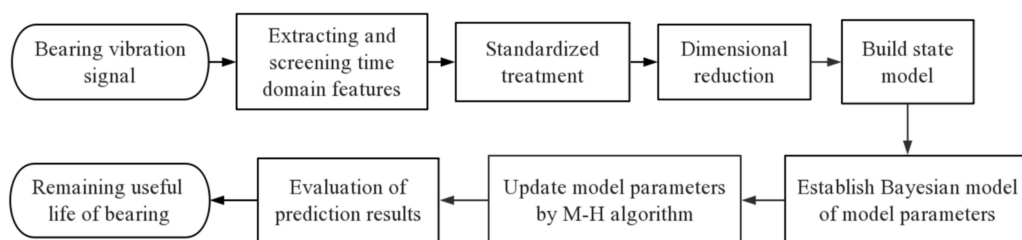


Figure 1. Flowchart of proposed bearing remaining useful life (RUL) prediction method.

2. Data-Driven State Model

2.1. Feature Index Selection

Time-domain indicators are commonly used in equipment fault detection and fault trend prediction, and can fully reflect the overall health degradation process of the system [24]. Common time-domain feature indexes include the mean value, root mean square value, peak value, and absolute mean amplitude. Considering the complex working conditions of bearings, a single feature index offers limited information for characterizing the bearing degradation process. Therefore, 16 time-domain feature indexes of bearing vibration signal data are extracted for the bearing RUL prediction. For details, see Table 1.

Table 1. Selection of feature indexes.

Sequence Number	Feature Index	Expression	Sequence Number	Feature Index	Expression
1	Mean value	$\bar{X} = \frac{1}{N} \sum_{t=1}^N x_i$	2	Standard deviation	$X_\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2}$
3	Variation coefficient	$X_{cv} = X_\sigma / \bar{X}$	4	Peak value	$X_{\max} = \max\{ x_i \}$
5	Square root amplitude	$X_r = \left[\frac{1}{N} \sum_{i=1}^N \sqrt{ x_i } \right]^2$	6	Absolute mean amplitude	$\bar{X}_p = \frac{1}{N} \sum_{i=1}^N x_i $
7	Root mean square	$X_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$	8	Peak-to-peak	$X_{p-p} = \max(x_i) - \min(x_i)$
9	Skewness	$X_{ske} = \frac{\sum_{i=1}^N (x_i - \bar{X})^3}{(N-1)X_\sigma^3}$	10	Kurtosis	$X_{kur} = \frac{\sum_{i=1}^N (x_i - \bar{X})^4}{(N-1)X_\sigma^4}$
11	Skewness factor	$I_{ske} = \frac{X_{ske}}{X_{rms}^3}$	12	Kurtosis factor	$I_{kur} = \frac{X_{kur}}{X_{rms}^4}$
13	Crest factor	$I_p = \frac{X_{\max}}{X_{rms}}$	14	Impulse factor	$I_i = \frac{X_{\max}}{\bar{X}_p}$
15	Waveform factor	$I_w = \frac{X_{rms}}{X_p}$	16	Margin factor	$I_m = \frac{X_{\max}}{X_r}$

2.2. Data Fusion

Due to differences in the dimension and magnitude of each feature index, the multidimensional feature index data must be normalized. The Z-score standardization method is applied [25]:

$$X_i^* : x_i^*(t) = \frac{x_i(t) - \mu_i}{\sigma_i} \quad t = 1, 2, \dots, n \quad (1)$$

where $x_i(t), t = 1, 2, \dots, n$ is the time series of the i th dimensional feature index composed of n data, $X_i^* = (x_i^*(t), t = 1, 2, \dots, n)$ is the normalized i -dimensional time series data, and μ_i and σ_i correspond to the mean and standard deviation of the i th dimensional time series data, respectively.

Although the multidimensional eigenvalue index contains sufficient information about the bearing degradation process, invalid information can be introduced resulting in increased computational complexity and reduced prediction efficiency. To overcome this, the multidimensional feature index data can be fused to construct a single bearing HI associated with the degradation process. Principal component analysis (PCA), which is one of the most widely used methods in data fusion, is applied [26,27]. The basic principle is to replace a large number of related variables with a small set of unrelated variables while retaining as much information as possible about the initial variables. Derived variables are called principal components, and are linear combinations of the initial variables. The basic steps of the PCA can be summarized as follows.

n groups of evaluation samples are set and each sample is evaluated by m indicators. The sample data can be expressed in the following form:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & x_{nm} \end{bmatrix}.$$

1. Calculate the correlation coefficient matrix of standardized data X^* . Then use the correlation coefficient to determine similarity among index variables.
2. Calculate the eigenvalues and eigenvectors of the correlation coefficient matrix. Eigenvalue λ_i is the variance of the i th principal component Y_i . The eigenvector corresponding to each eigenvalue is a linear coefficient of variation, and the principal component Y_i can be defined as

$$Y_i = \sum_{j=1}^m \alpha_{ij} X_j^* \quad i = 1, 2, \dots, n \quad (2)$$

where Y_i is the i th principal component data (i.e., output data), X_j^* is the j th dimensional original time series data (i.e., input data) after standardization, and α_{ij} is the linear transformation coefficient corresponding to the i th principal component and j th dimensional original time series data.

3. Calculate the variance contribution rate and cumulative contribution rate. The variance contribution rate reflects the role of index variables in the evaluation; the larger the value, the more effective the principal components are at retaining information. Generally, a principal component with an 85% cumulative contribution rate will meet calculation requirements.

The variance contribution rate of the p th principal component Y_p can be expressed as

$$W_p = \lambda_i / \sum_{i=1}^m \lambda_i.$$

The cumulative contribution rate of principal components $Y_1 Y_2 \dots Y_p$ can be defined as

$$Z_p = \sum_{i=1}^p \lambda_i / \sum_{i=1}^m \lambda_i.$$

2.3. Establishment of HI and State Model

In this paper, the state model is established in the negative time scale based on two considerations: First, many studies assume the degree of degradation of the system is consistent at the initial time; however, due to complex working conditions and errors in manufacturing and assembly, the initial degradation of different systems will vary quite considerably. Second, many studies have only used system status monitoring data, whereas system life information is ignored. To improve the information utilization rate and facilitate prediction of the remaining life, bearing life in the negative time scale can be taken as one of the Bayesian model parameters. The negative time scale transformation formula is

$$t_i^* = t_i - T_i \quad 0 < i < n \quad (3)$$

where n is the number of training bearings, t_i^* is time in the negative time scale, t_i is time in the positive time scale, and T_i is the system life. Thus, the state model in the negative timescale is

$$H_i = F(t_i^*, \theta_i) \quad t^* \in [1 - T_i, 0] \quad (4)$$

where $t^* \in [1 - T_i, 0]$ is the negative time scale, with 0 representing bearing failure; H_i is the HI of the bearing; $F(\cdot)$ is the state degradation model, which can be determined

according to the trend of the maximum principal component over time; θ_i and are the state model parameters. By transforming the model into the negative time scale using Equation (3), the state model in the positive time scale can be obtained as

$$H_i = F(t_i - T_i, \theta_i) \quad t \in [0, T_i] \quad (5)$$

3. Remaining Useful Life Prediction Model Based on Bayesian Theory

3.1. Bayesian Model

The bearing degradation process will be affected by uncertain factors, such as material properties, manufacturing and assembly processes, complex working conditions, and so on. In this paper, uncertainty in the prediction problem can be effectively dealt with by adopting the Bayesian model and probability method. There is a certain deviation between the bearing HI $H(t)$ and the measured value of HI $Y(t)$ (maximum principal component data), referred to as system noise. Noise usually obeys the standard normal distribution [1]. Therefore, the relationship between the measured value of $Y(t)$ and $H(t)$ is

$$Y(t) = H(t) + \varepsilon \quad (6)$$

where $\varepsilon \sim N(0, \sigma^2)$ is noise with a standard normal distribution. According to the properties of a normal distribution, the measured value of $Y(t)$ satisfies the following normal distribution:

$$Y(t) \sim N(H(t), \sigma^2). \quad (7)$$

Bayesian theory can be used to deduce the probability of unknown events based on the probability of known events [28]. The basic ideas behind Bayesian theory can be summarized as follows. An unknown parameter is regarded as a random variable. According to existing empirical information, a probability distribution of the variable, i.e., prior distribution, is obtained. The likelihood function is established and the distribution of variables is updated by fusing new information with existing prior distribution information to obtain a new posterior distribution. As the new data are gradually updated, the posterior distribution of the parameter will be closer to the real distribution. The process can be summarized as: the posterior distribution is proportional to the product of the prior distribution and the likelihood function. Considering uncertainty of bearing properties and randomness of service conditions, a Bayesian model of state parameters in the positive time scale can be established. The posterior distribution of state parameter θ and bearing life T can be expressed as

$$f_{PIY}(\theta, T|Y(t)) \propto f_{YIP}(Y(t)|\theta, T) \cdot f_P(\theta, T) \quad (8)$$

where $f_{YIP}(Y(t)|\theta, T)$ is the likelihood function of the measured value of $Y(t)$ under state parameter θ and bearing life T ; $f_P(\theta, T)$ is the prior distribution of state parameter θ and bearing life T , which can be obtained from historical data on the bearing life cycle.

Since the vibration signal of the bearing is independently measured at each time point and the measured value $Y(t)$ of the HI is obtained by fusing multidimensional feature indexes extracted from the vibration signal, data are independent at each time for each value of HI. According to the nature of the independent variable, the total likelihood function is equal to the product of the likelihood functions at each time. From Equation (8), the posterior distribution of state parameters θ and system life T at the predicted time k can be obtained as

$$f_{PIY}(\theta, T|(Y(t)|t = 1 : k)) \propto \left(\prod_{t=1}^k f_{YIP}(Y(t)|\theta, T) \right) \cdot f_P(\theta, T). \quad (9)$$

3.2. Remaining Useful Life Prediction

To predict the RUL of the bearing at time k , parameters are updated according to the Bayesian model established in Equation (5) and measured HI values before time k . Then, the posterior distribution of model parameters (θ, T) at time k is obtained. In general, calculating the posterior distribution of Bayesian model parameters is difficult. To solve this problem, the Markov chain Monte Carlo (MCMC) method can be applied to solve the posterior distribution [29].

The MCMC method is a sampling technique that can be used to extract samples from a probability density function (pdf). Posterior distribution samples are generated through stationary distribution of the Markov Chain and a Monte Carlo simulation is conducted. Here, the M-H algorithm is applied to calculate the posterior distribution [30]. The M-H algorithm constructs the proposal distribution $q(x)$ and generates samples, which are accepted or rejected according to a certain probability. Thus, a sample set conforming to the target distribution $p(x)$ is achieved. The specific steps of the algorithm are as follows:

1. Initialize starting point x^0 .
2. For $N - 1$ iterations, complete the following four steps:
 - a. Draw a sample, x^* , from the proposal distribution; the pdf value is $q(x^*|x^i)$ where i denotes the current iteration and the distribution mean is x^i with a selected standard deviation.
 - b. Sample u from a uniform distribution with a lower limit of zero and an upper limit of 1, $U(0,1)$.
 - c. Compute the acceptance ratio, $A = \min(1, (p(x^*)q(x^i|x^*)/p(x^i)q(x^*|x^i)))$, where $q(x^i|x^*)$ is the pdf value of the proposal distribution at x^i for the selected standard deviation, $p(x^*)$ is the pdf value of the target distribution at x^* , and $p(x^i)$ is the pdf value of the target distribution at x^i .
 - d. If $u < A$, set the new value of x , i.e., $x^{i+1} = x^*$. Otherwise, x remains unchanged, $x^{i+1} = x^i$.

In theory, any proposal distribution chain will gradually converge to the target distribution. Therefore, the M-H algorithm has good sampling effects for any target distribution. The selection of the proposed distribution affects the acceptance probability of the sample and convergence rate of the chain. In this paper, the proposal distribution is chosen as a uniform distribution based on empirical considerations and expert information. In addition, if the number of iterations is large enough, the initial value of the chain has no effect on the final sampling result [31]. Samples of the initial iteration are usually discarded as the training process and after the Markov chain is stable, the distribution can be taken as the sampling result.

4. Application of Proposed Method

4.1. Bearing Data

Datasets containing the complete run-to-failure data of 15 rolling element bearings (XJTU-SY) under accelerated degradation experiments were provided by the Institute of Design Science and Basic Component at Xi'an Jiaotong University (XJTU) and the Changxing Sumyoung Technology Co., Ltd., Zhejiang, China, (SY) [32]. For accelerated degradation experiments, a total of three different operating conditions were set and five bearings were tested under each operating condition. The sampling period was 1 min and the sampling frequency was set to 25.6 kHz. A total of 32,768 data points were recorded in 1.28 s during each sampling process. In our analysis, the bearing dataset of working condition 1 was selected for RUL prediction. Bearing 1_1 was taken as the test bearing and bearings 1_2, 1_3, 1_4, and 1_5 were selected as training bearings 1–4, respectively. Specific information extracted from the bearing datasets are shown in Table 2.

Table 2. Bearing datasets.

Operating Condition	Bearing Dataset	Number of Files	Bearing Lifetime	Fault Element
Condition1 (35 Hz/12 kN)	Bearing 1_1 (Test bearing)	158	2 h 38 min	Outer race
	Bearing 1_2 (Training bearing 1)	161	2 h 41 min	Outer race
	Bearing 1_3 (Training bearing 2)	123	2 h 3 min	Outer race
	Bearing 1_4 (Training bearing 3)	122	2 h 2 min	Cage
	Bearing 1_5 (Training bearing 4)	52	52 min	Inner and outer race

An HI and state model were established using signal data of the training bearings and prior information for the Bayesian model of state parameters and bearing life were obtained. Bearing 1_1 data were used as the test data to carry out the real-time RUL prediction. Finally, the prediction results were compared with the real RUL to verify the method.

4.2. Data Processing

As listed in Table 1, 16 time-domain feature indexes of the four training datasets were extracted and plotted in the negative time scale to observe the degradation trends over time. To ensure the feature indexes accurately reflected the degradation process of the bearings, feature indexes 2, 4, 5, 6, 7, 8, and 12 with obvious degradation trends were selected for further analysis, as shown in Figure 2.

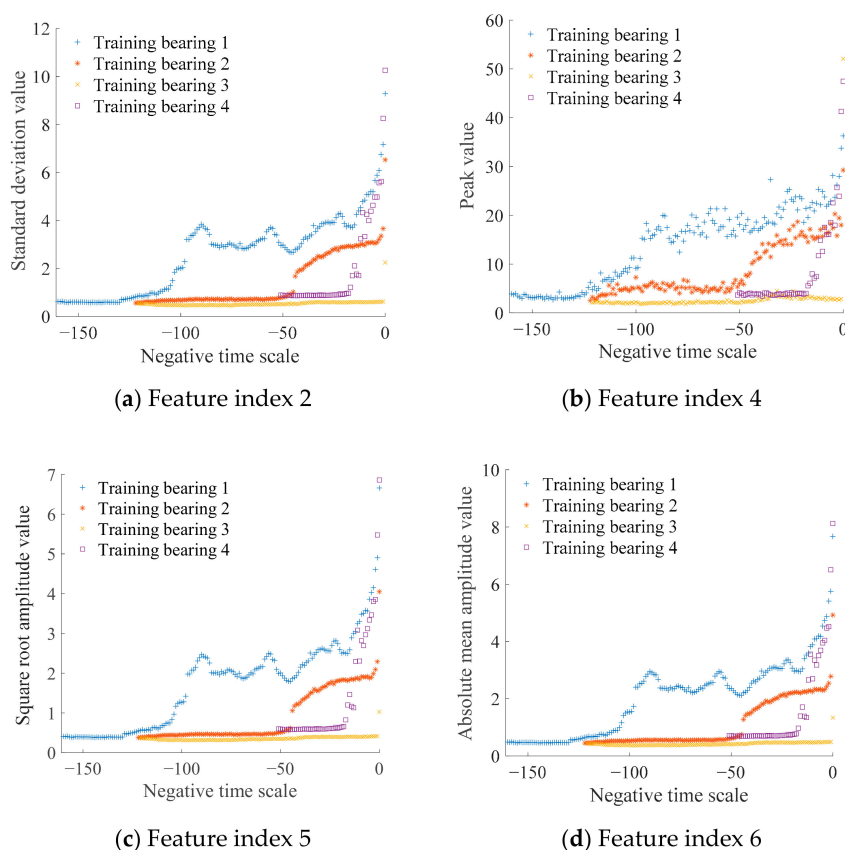


Figure 2. Cont.

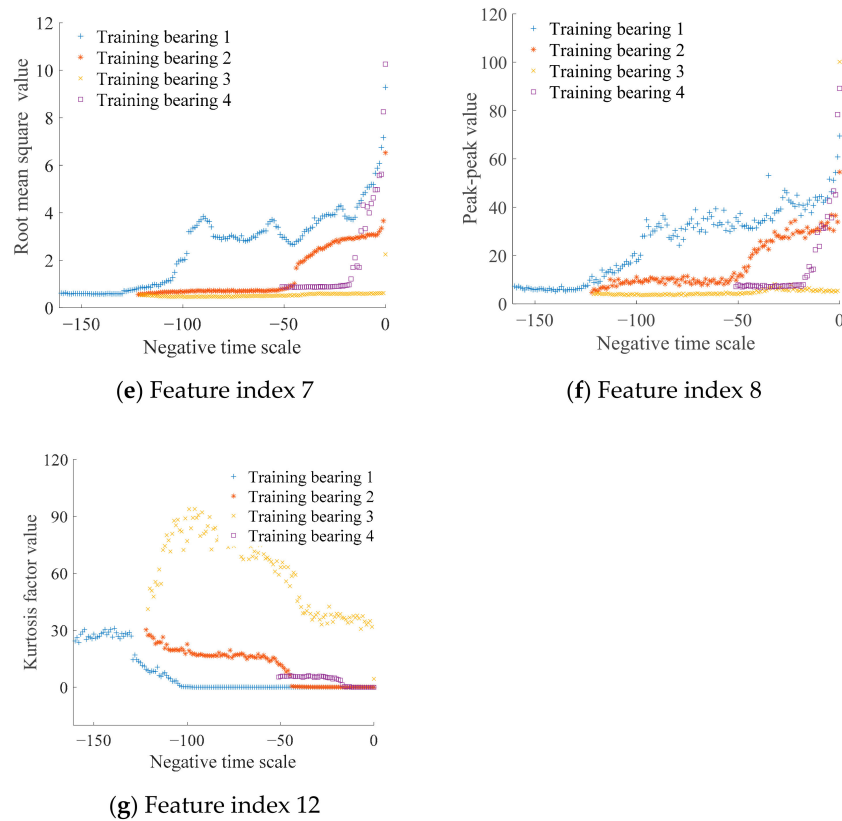


Figure 2. Data trends of selected feature index.

The seven selected feature indexes were standardized using Equation (1), and then the PCA method was applied to effectively achieve dimension reduction. Figure 3 shows the contribution rate of each principal component after data fusion. Principal component 1 retains sufficient information from the original data as the variance contribution rate of principal component 1 is up to 89.1969%. Therefore, principal component 1 was selected as the measurement value of the HI. Table 3 shows the corresponding linear transformation coefficients of 7 selected feature indexes in principal component 1. The principal component 1 can be calculated by introducing the linear transformation coefficient into Equation (2).

$$Y_1 = \sum_{j=1}^7 \alpha_{1j} X_j^* \quad (10)$$

where α_{1j} represents the transformation coefficient of principal component 1 corresponding to the j selected feature indexes. Figure 4 shows the trends of principal component 1 of the four training bearings. Principal component 1 clearly changes over time, which can adequately reflect degradation of the bearing and further verifies the rationality of selecting principal component 1 to build HI and state model.

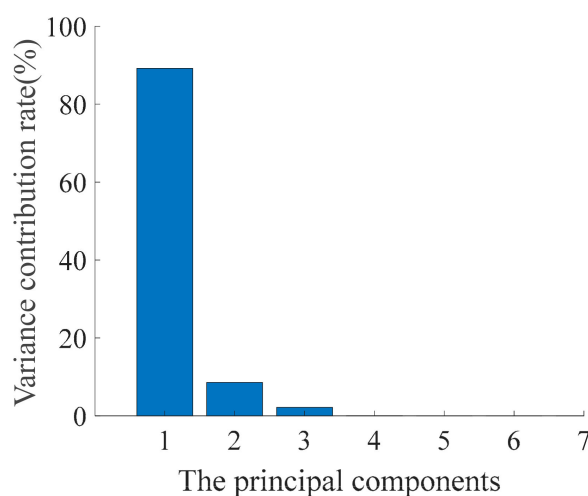


Figure 3. Variance contribution rates of each principal component.

Table 3. Linear transformation coefficients of principal component 1.

Selected Feature Index	X_σ	X_{\max}	X_r	\bar{X}_p	X_{rms}	X_{p-p}	X_{kur}
Coefficient value α_{1j}	0.3960	0.3940	0.3947	0.3959	0.3959	0.3859	−0.2692

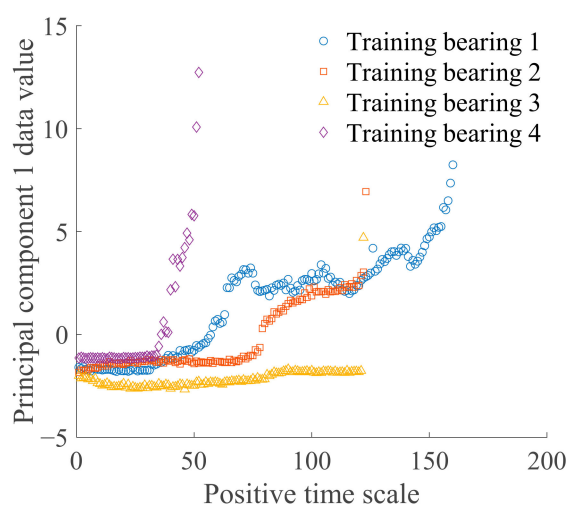


Figure 4. Principal component 1 data.

4.3. Establishment of HI and State Model

Selection of a reasonable model is the foundation of high-precision prediction. To reflect changes in the value of HI (principal component 1) over time (Figure 4), an exponential model was selected and used to establish the state model, expressed as

$$H_i(t^*) = \exp[a_i \cdot t^* + b_i] + c_i \quad t^* \in [1 - T_i, 0] \quad (11)$$

where $H_i(t^*)$ is the HI and a_i , b_i , and c_i are the state parameters of the i th training bearing. According to Equation (3), the state model can be expressed in the positive time scale as

$$H_i(t) = \exp[a_i \cdot (t - T_i) + b_i] + c_i \quad t \in [1, T_i]. \quad (12)$$

Measured values of health indicators (HIs) and the bearing life of the training bearings were introduced into Equation (12) and the parameter estimation values of each training

bearing state model were obtained using the least squares method. Figure 5 shows state curves of the four training bearings. The degradation process of training bearing 3 is stable until unexpected breakdown occurs at the last sampling point, which shows a lack of universality. Thus, data of training bearing 3 are not considered when building the Bayesian model. Differences can be observed in the degradation process and state parameters of training bearings 1, 2, and 4. Therefore, three prior samples of state parameters a , b , and c and bearing life T provide a suitable reference for the Bayesian model. In addition, the mean and variance of noise were obtained using statistics. Since noise is assumed to follow a normal distribution, the mean value of noise was zero and the three prior samples of noise variance were obtained.

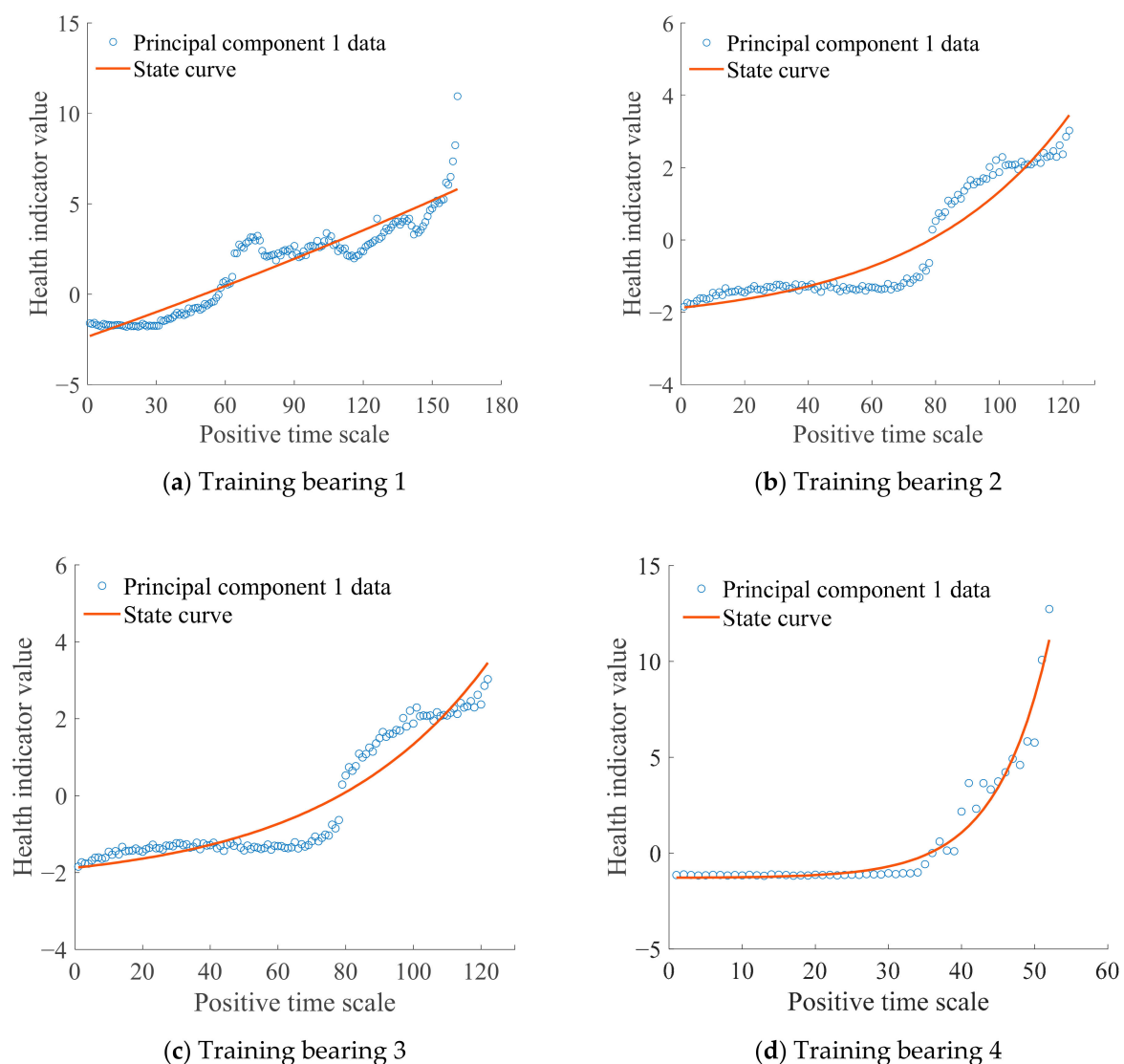


Figure 5. Training bearing state model.

4.4. Bayesian Model and RUL Prediction

The feature indexes screened in Section 4.2 were extracted from the test data and standardized using Equation (1). Based on the linear transformation coefficients presented in Table 3 and Equation (2), PCA dimensional reduction was carried out on the standardized feature index data. One-dimensional data obtained from the transformation were used as the measured values of $Y(t)$. A Bayesian model in the positive time scale was established

using Equation (9) and Equation (12). Since the posterior distribution of parameters a , b , c , and T cannot be solved directly, the M-H algorithm was applied to perform independent sampling. For the test bearing, the prior distribution of parameters a , b , and c was set as a uniform distribution. According to three prior samples obtained from the historical data, the range of uniform distribution was determined to be $[\bar{a} - z, \bar{a} + z]$, where \bar{a} is the mean value of three prior samples and z is a known constant used to describe differences among bearing degeneration. The prior distribution of b and c is determined in the same way. When the M-H algorithm was applied, the prior distribution of parameters a , b , and c were taken as the proposal distribution. The prior distribution of parameter T was set as a normal distribution. The mean value depends on the current value of the Monte Carlo chain and the standard deviation is 0.2. For example, the current value of the Monte Carlo chain is T^i with prior distribution $N(T^i, 0.2)$ in the i th iteration. A candidate sample T^* is randomly selected from the proposed distribution $N(T^i, 0.2)$. If the acceptance ratio A is greater than random variable u , sample T^* is accepted with $T^{i+1} = T^*$; otherwise, $T^{i+1} = T^i$. Following the above method, the prior distribution of parameter T was selected as the normal random walk distribution and the prior distribution was taken as the proposal distribution of parameter T sampling.

Considering the bearing is in the normal working state and runs smoothly in the early sampling stage, it is necessary to determine the starting point of bearing failure as the initial point of the RUL prediction, which better reflects the degradation trend of the bearing. The HI of the test data increases suddenly when the time series reaches 60 in the positive time scale; therefore, this point was taken as the starting point of bearing failure. Using the measured value of the test HI and the Bayesian model, the MCMC algorithm was applied to update the model parameters. A total of 10,000 iterations were used for each prediction. The first 5000 were used for the training process and the final 5000 update samples were taken as the posterior distribution samples of the model parameters. The posterior distribution of the predicted bearing life T was obtained, and the RUL posterior distribution was then obtained by subtracting the cut-off time of the test data. The expectation of the distribution sample was taken as the final prediction result. Figure 6 shows the distribution of the bearing life posteriori samples at the predicted time of 110.

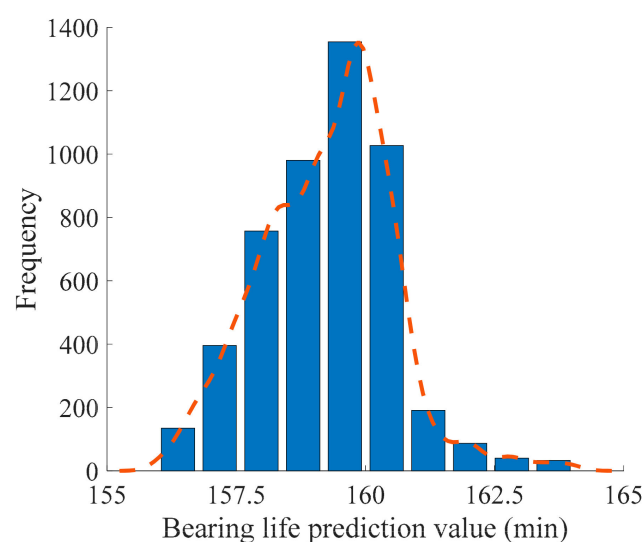


Figure 6. Posteriori distribution diagram of bearing predicted life (predicted time $k = 110$).

After the prediction update, the posterior distribution of bearing RUL is composed of 5000 samples. The probability distribution of the posterior sample is shown in Figure 7. Inspection numbers 1 to 10 correspond to the prediction time series 68, 78, 88, ..., 158. The posteriori samples are represented by the probability density of the normal distribution and the expectation of each posteriori distribution is taken as the final prediction result.

Comparison with the real bearing RUL demonstrates the accuracy of the proposed method. At the same time, the probability density distribution of the prediction results is relatively concentrated, indicating good stability. Moreover, the probability density curve tends to become more concentrated over time, indicating that the impact of uncertainty gradually decreases as the prediction progresses. A certain deviation exists between the predicted result and the actual value at the inspection number 7, which is related to the trend of the predicted bearing vibration signal data. The unstable, fluctuating trend of the original data causes deviation of the predicted result. Figure 8 shows a certain error in the prediction time between 110 and 140, which is caused by divergence of the measured values of the test HI that are used to update the Bayesian model. However, the subsequent prediction deviations gradually decrease, and prediction results are stable around the real RUL value, indicating that the overall prediction result is accurate.

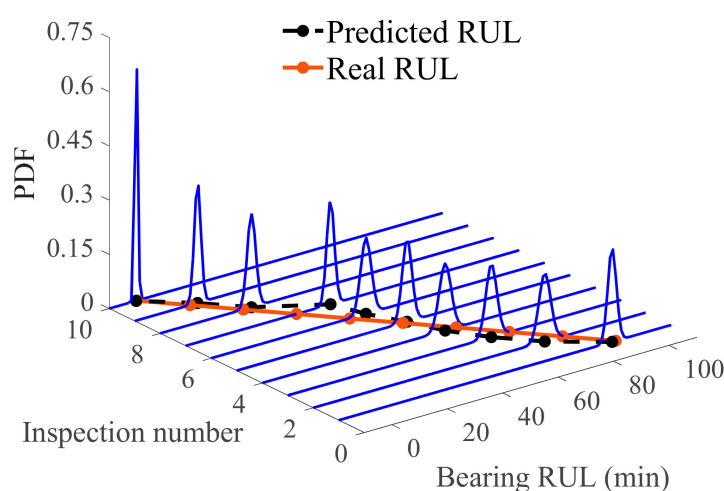


Figure 7. Probability density diagram of bearing RUL.

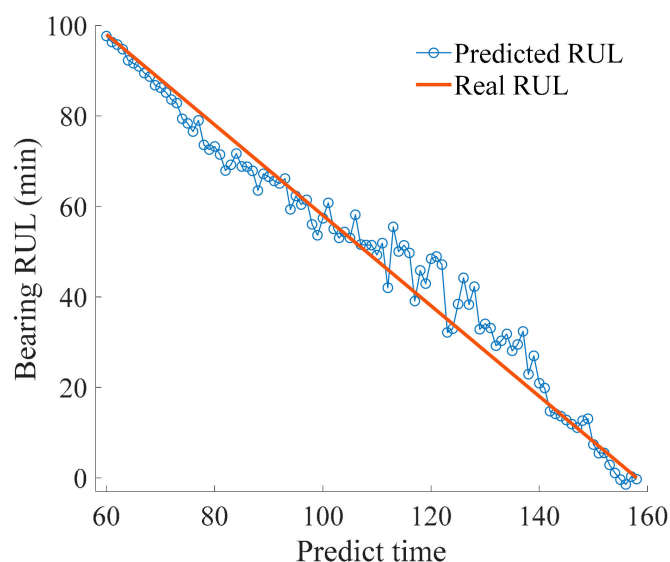


Figure 8. Comparison of predicted RUL and real RUL.

4.5. Evaluation of Prediction Results

In order to quantitatively evaluate the prediction effect, three RUL prediction evaluation methods were adopted in this paper: root mean square error (RMSE), mean relative error (MARE) and error function based on asymmetric exponential [33,34].

RMSE denotes the root mean square of the prediction errors, which can be expressed as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - z_i)^2} \quad (13)$$

where x_i and z_i , respectively, represent the predicted value and the real value of the i th prediction; N is the total number of real-time predictions. A smaller RMSE value means that RUL predicts better effectiveness.

MARE is the mean value of relative error among all time point. The expression of MARE is as follows:

$$MARE = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - z_i}{z_i} \right| \times 100\%. \quad (14)$$

Obviously, the approaches with smaller MARE would be better than others.

Error function based on asymmetric exponential can comprehensively evaluate the accuracy of the prediction method by constructing the exponential error between the predicted value and the true value and synthesizing the prediction accuracy of each prediction time series. According to RUL prediction results, the total evaluation error S can be calculated as follows:

$$S = \sum_{i=1}^N S_i, \quad (15)$$

$$S_i = \begin{cases} \exp(-d_i/13) - 1 & d_i \leq 0 \\ \exp(d_i/10) - 1 & d_i > 0 \end{cases}, \quad (16)$$

$$d_i = x_i - z_i, \quad (17)$$

where S is the total evaluation error of N predictions; d_i is the RUL estimation error of the i th prediction; S_i is the evaluation error of the i th prediction. As the overall error evaluation value S decreases, the prediction accuracy increases.

The support vector machine (SVM) method and Pairs-based particle filter (PF) method were selected to illustrate the accuracy of the proposed method. The SVM is a widely used machine learning algorithm for classification and prediction, and can predict the RUL of bearings under small sample conditions with good prediction accuracy [35]. The Paris-based PF method combines physical model and observation data to identify model parameters, which is a model-based RUL prediction method [36]. Using the same bearing data and characteristic indexes as above, the SVM method and PF method were respectively introduced to predict the RUL of the bearing. The comparison of prediction results and errors is shown in Figures 9 and 10, indicating that the method proposed in this paper has certain stability and accuracy. Then, the prediction errors of several methods are calculated according to Equations (13)–(17). Results of the analysis are presented in Table 4. The proposed method has good prediction accuracy and good stability.

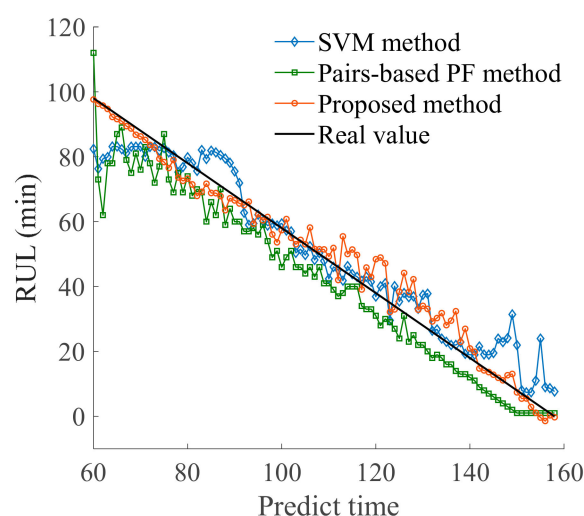


Figure 9. Comparison of prediction result.

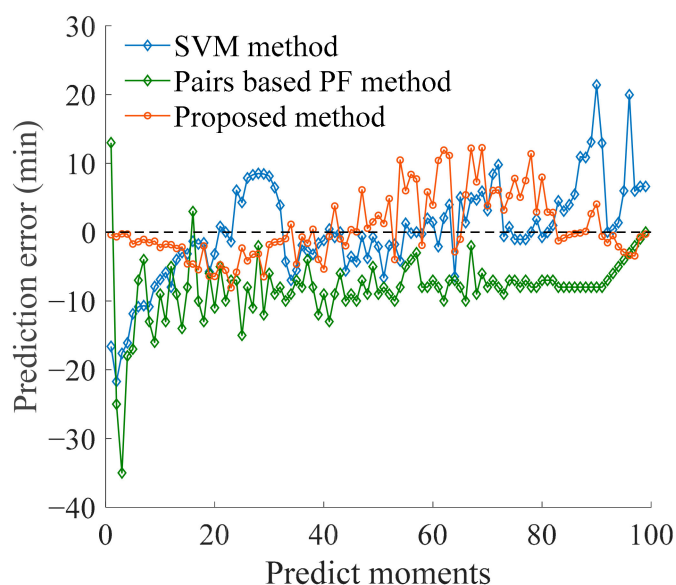


Figure 10. Comparison of prediction error.

Table 4. Comparison of prediction results.

Prediction Method	SVM Method	Pairs Based PF Method	Proposed Method
RMSE	7.1005	8.7427	4.8843
MARE	3.2405	4.7628	2.3079
S value	80.7027	94.8183	46.5061

5. Conclusions

This paper proposed a method of bearing RUL prediction based on Bayesian theory. Feature indexes reflecting the degradation trend were extracted from bearing vibration signals. The corresponding HIs were obtained through PCA and a state model was established. Information was extracted from limited historical data of samples to construct the prior distribution of the model parameters. A Bayesian model of state parameters was established and the MCMC algorithm was applied to update the Bayesian model parameters to obtain the posterior distribution of the RUL and its predicted value. The accuracy and stability of the method were verified using actual bearing data. In addition, the prediction

results were compared with those obtained using an existing prediction method and the advantages of proposed method in terms of prediction accuracy were demonstrated.

Compared with other data-driven life prediction methods, the proposed method can be used to build a degradation model from limited existing data. Furthermore, the Bayesian approach effectively deals with parameter uncertainties in the degradation process. Therefore, prediction error is reduced and prediction accuracy and stability are greatly improved.

Author Contributions: Conceptualization, Y.L. and X.H.; Data curation, T.G.; Formal analysis, T.G.; Funding acquisition, X.H.; Methodology, T.G.; Resources, Y.L.; Software, T.G. and Y.L.; Supervision, C.W.; Validation, T.G.; Visualization, C.W.; Writing—original draft, T.G.; Writing—review & editing, Y.L. and X.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China: 51975110; Liaoning Revitalization Talents Program: XLYC1907171; Fundamental Research Funds for the Central Universities: N2003005.

Data Availability Statement: Data sharing not applicable.

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

References

- Wang, F.; Su, S. *Fault Diagnosis and Life Prediction of Rolling Bearings*; Science Press: Beijing, China, 2018; pp. 21–25.
- Qiu, H.; Lee, J.; Lin, J.; Yu, G. Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Adv. Eng. Inform.* **2003**, *17*, 127–140. [\[CrossRef\]](#)
- Huang, X.Z.; Li, Y.X.; Zhang, Y.M.; Zhang, X.F. A new direct second-order reliability analysis method. *Appl. Math. Model.* **2018**, *55*, 68–80. [\[CrossRef\]](#)
- Li, Q.; Zuo, M.J.; Liang, S.Y. Prognosis of Bearing Degeneration Using Adaptive Quaternion Least Mean Biquadrate Under Framework of Hypercomplex Data. *IEEE Sens. J.* **2020**, *20*, 2659–2670. [\[CrossRef\]](#)
- Lei, Y.; Li, G.; Jia, N.P.; Lin, F.; Xing, J.S.B. A Nonlinear Degradation Model Based Method for Remaining Useful Life Prediction of Rolling Element Bearings. In Proceedings of the 2015 Prognostics System Health Management Conference, Beijing, China, 21–23 October 2015; Zhao, T., Pecht, M.G., Zhang, S., Eds.; IEEE: New York, NY, USA, 2015.
- Liu, X.L.; Liu, L.S.; Liu, D.T.; Wang, L.L.; Guo, Q.; Peng, X.Y. A Hybrid Method of Remaining Useful Life Prediction for Aircraft Auxiliary Power Unit. *IEEE Sens. J.* **2020**, *20*, 7848–7858. [\[CrossRef\]](#)
- Jiang, Y.Y.; Zeng, W.W.; Shen, J.J.; Chu, J. Prediction of remaining useful life of lithium-ion battery based on convex optimization life parameter degradation mechanism model. *Proc. CSU EPSA* **2019**, *31*, 23–28.
- Liu, O.K.; Li, Q.; Wang, X.; Ding, R. Life prediction method for EMU axle box bearings based on actual measured loadings. *J. Mech. Eng.* **2016**, *52*, 45–54. [\[CrossRef\]](#)
- Zhang, Y.Q.; Zou, J.H.; Ma, J. Rolling bearing residual life prediction based on Grey prediction model with multiple degenerate variables. *J. Detect. Control* **2019**, *41*, 112–120.
- Wu, Y.T.; Yuan, M.; Dong, S.P.; Lin, L.; Liu, Y.Q. Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. *Neurocomputing* **2018**, *275*, 167–179. [\[CrossRef\]](#)
- Ren, L.; Sun, Y.Q.; Cui, J.; Zhang, L. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *J. Manuf. Syst.* **2018**, *48*, 71–77. [\[CrossRef\]](#)
- Xia, M.; Li, T.; Shu, T.X.; Wan, J.F.; de Silva, C.W.; Wang, Z.R. A Two-Stage Approach for the Remaining Useful Life Prediction of Bearings Using Deep Neural Networks. *IEEE Trans. Ind. Inform.* **2019**, *1*, 3703–3711. [\[CrossRef\]](#)
- Li, X.Q.; Jiang, H.K.; Xiong, X.; Shao, H.D. Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network. *Mech. Mach. Theory* **2019**, *133*, 229–249. [\[CrossRef\]](#)
- Jia, F.; Lei, Y.G.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **2016**, *72*, 303–315. [\[CrossRef\]](#)
- Lei, Y.G.; Jia, F.; Lin, J.; Xing, S.B.; Ding, S.X. An Intelligent Fault Diagnosis Method Using Unsupervised Feature Learning Towards Mechanical Big Data. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3137–3147. [\[CrossRef\]](#)
- Thirukovalluru, R.; Dixit, S.; Sevakula, R.K.; Verma, N.K.; Salour, A. Generating Feature Sets for Fault Diagnosis Using Denoising Stacked Auto-Encoder. In Proceedings of the 2016 IEEE International Conference on Prognostics and Health Management, Ottawa, ON, Canada, 20–22 June 2016.
- Ding, X.X.; He, Q.B. Energy-Fluctuated Multiscale Feature Learning with Deep ConvNet for Intelligent Spindle Bearing Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1926–1935. [\[CrossRef\]](#)
- Mosallam, A.; Medjaher, K.; Zerhouni, N. Bayesian Approach for Remaining Useful Life Prediction. In Proceedings of the 2013 Prognostics and Health Management Conference, New Orleans, LA, USA, 14–17 October 2013; Zio, E., Baraldi, P., Pierucci, S., Klemes, J.J., Eds.; Chemical Engineering Transactions: Milano, Italy, 2013; Volume 33, pp. 139–144.

19. Cheng, Y.J.; Lu, C.; Li, T.Y.; Tao, L.F. Residual lifetime prediction for lithium-ion battery based on functional principal component analysis and Bayesian approach. *Energy* **2015**, *90*, 1983–1993. [\[CrossRef\]](#)
20. Liu, Y.H.; Shuai, Q.; Zhou, S.Y.; Tang, J. Prognosis of Structural Damage Growth via Integration of Physical Model Prediction and Bayesian Estimation. *IEEE Trans. Reliab.* **2017**, *66*, 700–711. [\[CrossRef\]](#)
21. Tang, X.P.; Zou, C.F.; Yao, K.; Lu, J.Y.; Xia, Y.X.; Gao, F.R. Aging trajectory prediction for lithium-ion batteries via model migration and Bayesian Monte Carlo method. *Appl. Energy* **2019**, *254*, 113591. [\[CrossRef\]](#)
22. Li, T.M.; Pei, H.; Pang, Z.N.; Si, X.S.; Zheng, J.F. A Sequential Bayesian Updated Wiener Process Model for Remaining Useful Life Prediction. *IEEE Access* **2020**, *8*, 5471–5480. [\[CrossRef\]](#)
23. Zaidan, M.A.; Mills, A.R.; Harrison, R.F.; Fleming, P.J. Gas turbine engine prognostics using Bayesian hierarchical models: A variational approach. *Mech. Syst. Signal Process.* **2016**, *70*, 120–140. [\[CrossRef\]](#)
24. Jin, W.J.; Chen, Y.; Lee, J. Methodology for Ball Screw Component Health Assessment and Failure Analysis. In Proceedings of the ASME 8th International Manufacturing Science and Engineering Conference, Madison, WI, USA, 10–14 June 2013; Volume 2.
25. Foster, P. Exploring multivariate data using directions of high density. *Stat. Comput.* **1998**, *8*, 347–355. [\[CrossRef\]](#)
26. Mosallam, A.; Medjaher, K.; Zerhouni, N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *J. Intell. Manuf.* **2016**, *27*, 1037–1048. [\[CrossRef\]](#)
27. Moghaddass, R.; Zuo, M.J. An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process. *Reliab. Eng. Syst. Saf.* **2014**, *124*, 92–104. [\[CrossRef\]](#)
28. Hamada, M.S.; Wilson, A.; Reese, C.S.; Martz, H. *Bayesian Reliability*; Springer Science & Business Media: Berlin, Germany, 2008.
29. Andrieu, C.; de Freitas, N.; Doucet, A.; Jordan, M.I. An introduction to MCMC for machine learning. *Mach. Learn.* **2003**, *50*, 5–43. [\[CrossRef\]](#)
30. Wang, Z.L. Markov chain Monte Carlo sampling using a reservoir method. *Comput. Stat. Data Anal.* **2019**, *139*, 64–74. [\[CrossRef\]](#)
31. Qi, H.S.; Hu, X.B. Bayesian inference of channelized section spillover via Markov Chain Monte Carlo sampling. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 478–498. [\[CrossRef\]](#)
32. Wang, B.; Lei, Y.G.; Li, N.P.; Li, N.B. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [\[CrossRef\]](#)
33. Son, K.L.; Fouladirad, M.; Barros, A.; Levrat, E.; Lung, B. Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliab. Eng. Syst. Saf.* **2013**, *112*, 165–175. [\[CrossRef\]](#)
34. Saxena, A.; Goebel, K.; Simon, D.; Eklund, N. Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation. In Proceedings of the 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008.
35. Dong, S.J.; Sheng, J.L.; Liu, Z.; Zhong, L.; Wei, H.B. Bearing remain life prediction based on weighted complex SVM models. *J. Vibroeng.* **2016**, *18*, 3636–3653.
36. An, D.; Choi, J.H.; Kim, N.H. Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab. *Reliab. Eng. Syst. Saf.* **2013**, *115*, 161–169. [\[CrossRef\]](#)