

Research Article

Remaining Useful Life Prediction for Rotating Machinery Based on Optimal Degradation Indicator

Aisong Qin, Qinghua Zhang, Qin Hu, Guoxi Sun, Jun He, and Shuiquan Lin

Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, Maoming 525000, China

Correspondence should be addressed to Qin Hu; huqinbinzhou@163.com

Received 6 September 2016; Revised 1 March 2017; Accepted 8 March 2017; Published 22 March 2017

Academic Editor: Mickaël Lallart

Copyright © 2017 Aisong Qin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Remaining useful life (RUL) prediction can provide early warnings of failure and has become a key component in the prognostics and health management of systems. Among the existing methods for RUL prediction, the Wiener-process-based method has attracted great attention owing to its favorable properties and flexibility in degradation modeling. However, shortcomings exist in methods of this type; for example, the degradation indicator and the first predicting time (FPT) are selected subjectively, which reduces the prediction accuracy. Toward this end, this paper proposes a new approach for predicting the RUL of rotating machinery based on an optimal degradation indicator. First, a genetic programming algorithm is proposed to construct an optimal degradation indicator using the concept of FPT. Then, a Wiener model based on the obtained optimal degradation indicator is proposed, in which the sensitivities of the dimensionless parameters are utilized to determine the FPT. Finally, the expectation of the predicted RUL is calculated based on the proposed model, and the estimated mean degradation path is explicitly derived. To demonstrate the validity of this model, several experiments on RUL prediction are conducted on rotating machinery. The experimental results indicate that the method can effectively improve the accuracy of RUL prediction.

1. Introduction

Rotating machinery are the most widely used mechanical equipment in industry and generally consist of motor, shaft, gear box, bearing, and load [1, 2]. Any fault in rotating machinery can cause the breakdown of the entire machine, which can lead to catastrophic consequences. According to statistics, one-third to one-half of the expenditure is wasted through ineffective maintenance [3]. If the remaining useful life (RUL) can be predicted in advance, catastrophes can be avoided, and predictive maintenance can be implemented to maximize machine uptime and minimize maintenance costs. Therefore, there is a pressing need to continuously develop and improve the current RUL prediction technologies [4].

At present, the main method used for RUL estimation is a physics-based failure model and a data-driven method. The data-driven method has become the mainstream in the field of RUL prediction, and it relies only on the available past observed data and statistical models [5, 6]. The data-driven method can be beneficial when mechanical principles

are not straightforward or when mechanical systems are very complex. Considering that rotation machinery has complex degradation processes due to nonlinearity, stochasticity, and nonstationarity, the data-driven method is suitable for estimating the RUL of rotating machinery. The main existing data-driven methods are based on regression-based models, Gamma processes, Wiener processes, or Markovian-based models. As the Wiener degradation model is a frequently used model, its prognostics have been introduced in recent research. Guan et al. modeled the constant-stress accelerated degradation test as a Wiener process in [7], and simulations were conducted to demonstrate the effectiveness of degradation modeling. Huang et al. proposed a nonlinear heterogeneous Wiener process model to characterize degradation trajectories in [8]. Wang et al. presented an adaptive method of RUL estimation based on a generalized Wiener degradation process in [9]. The nonlinearity and temporal uncertainty were jointly taken into account in the proposed degradation model.

However, there is little research concerning the construction of an optimal degradation indicator and the selection of first predicting time (FPT) based on the Wiener process degradation model. It is well known that the extraction and construction of a degradation indicator plays an important role in RUL prediction. The purpose of a degradation indicator construction for prognostics is to identify a set of degradation indicators that can clearly present the degradation process. However, in most of the current literature, the selection of a degradation indicator usually depends on the authors' subjective judgment and experience. The degradation indicator's ability to capture the trend of the progression process has not been discussed. Usually the degradation curve of an excellent degradation indicator has good monotonicity, which is explained in [10] where an actual gyro's drift data are illustrated. In reality, however, the performance variables at any time take on stochastic properties because of the differences in the products; this is often not the case for most of the individual product units where the degradation curve is monotonous. That is, the performance degradation curve is nonmonotonic in practice. In [11], an actual measurement series for a bearing failure history was collected from a pump in the field. Although the measurement showed a generally increasing trend, there were large fluctuations at multiple places. In [12], the practical degradation path of a ball bearing was shown; the degradation curve was quite steady until ball-bearing failure occurred.

To our knowledge, the studies on an optimal degradation indicator model are very limited. A few exceptions are the works in [13–15]. Gomes et al. presented a method for combining the measured parameters into a single indicator for monitoring the condition of systems in [13]. Liu et al. proposed a health-index extraction and optimization framework, requiring only the operating parameters, for battery degradation modeling and RUL estimation in [14]. To derive optimal prognostic features, Liao employed the Paris model, combined with a genetic programming (GP) method, to predict the RUL of bearings in [15]. In this literature, the FPT was not considered while using GP to generate the optimal prognostic features. The time to start predicting was defined as the FPT, in [16], as shown in Figure 1. If the FPT is not adopted, a large amount of data will be required to build an optimal degradation indicator, which will affect the convergence and complexity of GP and make the optimal degradation indicator nonideal. Through determining an appropriate FPT, the data from the FPT to the failure time can be utilized to construct an optimal degradation indicator, which will greatly reduce the amount of data and be beneficial in constructing an optimal degradation indicator. In addition, Liao assumed in [15] that the fault growth of the bearing followed the Paris law. However, the fault growth of the bearing came into being as the result of the accumulation of many small cracks and not the principal crack extension, which implied that the Paris law did not apply to bearings. It is common knowledge that establishing physical models for complicated rotating machinery is difficult. Thus, only

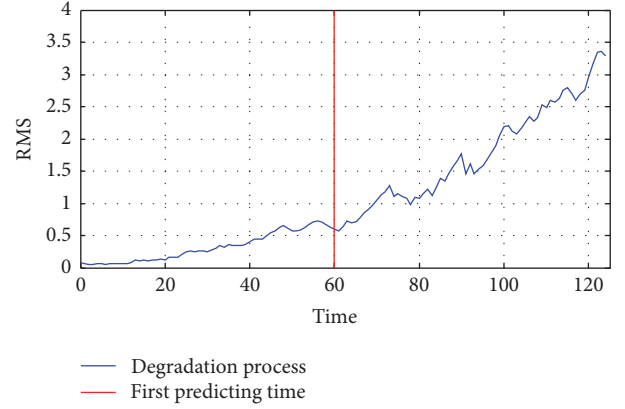


FIGURE 1: Degradation process and first predicting time.

a data-driven approach can be adopted to estimate the RUL. However, little reported literature can be found on constructing an optimal degradation indicator based on the FPT and addressing the Wiener degradation model for rotating machinery.

Based on the above discussions, an integrated RUL prediction method was proposed based on GP and the concepts of FPT and Wiener process degradation modeling, which aimed to solve the problem that the degradation progression might be nonmonotonic.

This research work has the following theoretical and practical contributions:

- (i) Our method can easily construct an optimal degradation indicator using GP owing to the involvement of the FPT, which can greatly reduce the data in the terminator set during GP.
- (ii) The optimal degradation indicator extracted through GP has a better monotonic trend than the conventional degradation indicator. Moreover, the optimal degradation indicator contains more degradation information than a conventional degradation indicator, which characterizes the degradation process well.
- (iii) Based on the optimal degradation indicator, a Wiener-process-based degradation model is established to estimate the RUL. To demonstrate the effectiveness of our method, we apply it to a real test bed. The results show that our developed method can make accurate RUL predictions.

The remaining parts of this paper are organized as follows: In Section 2, the related works are surveyed and presented. In Section 3, we describe the RUL prediction framework. Section 4 briefly reviews the GP algorithm and the time-domain parameters, and Section 5 describes the Wiener process degradation modeling for RUL prediction. Introduction to the experimental equipment and the construction of an optimal degradation indicator are given in Section 6. The RUL prediction and the comparison of the proposed method with other prediction methods are presented in Section 7, and Section 8 concludes this paper.

2. Related Works

Researchers had conducted considerable efforts on RUL estimation and developed a variety of prediction methods. Owing to their favorable mathematical properties and physical interpretations, Wiener processes have been widely adopted to characterize degrading systems [10, 17–19]. Si et al. presented a more general nonlinear Wiener-process-based model for RUL estimation in [20]. Unlike the usual monotonic assumption, they formulated an analytical approximation of the RUL distribution. Wang and Carr, in [21], considered a Wiener process with random effects, but without measurement errors, for the lifetime modeling of bridge beams. Tseng and Peng proposed an integrated Wiener process to model the cumulative degradation path of a product's quality characteristics in [22]. Balka et al. reviewed some methods of cure rate models based on the first passage times, using Wiener processes, in [23]. Mishra and Vanli, in [24], used principal component regression and Wiener process degradation modeling for predicting the RUL of a structure from the Lamb wave sensor data. The principal component regression was used for extracting damage-sensitive features, and the Wiener process was developed to model the random growth. Tang et al. addressed the effects of model misspecification of the linear Wiener process for RUL estimation in [25]. Tang et al., in [26], proposed real-time RUL prediction based on nonlinear Wiener-based degradation processes with measurement errors. The simulation results showed that considering measurement errors could significantly improve the accuracy of real-time RUL prediction. Wang et al. proposed an additive Wiener-process-based prognostic model for hybrid deteriorating systems, in [27], by employing an additive Wiener process model that consisted of a linear degradation part and a nonlinear part. The above survey of the related works and discussions poses an interesting challenge for prognostic studies through Wiener-process-based degradation modeling.

In terms of feature extraction, Kotani et al. used GP to search for the terms of polynomial expressions, which were the features extracted to improve the performance of pattern recognition in [28]. Muni et al., in [29], presented an online feature selection algorithm using GP. The proposed methodology simultaneously selected a good subset of features and constructed a classifier. Guo et al., in [30], proposed an approach-based GP for feature extraction from raw vibration data. The created features were used as the inputs to a neural classifier for the identification of six bearing conditions. Bechhoefer et al. fused condition indicators, in [31], to obtain a degradation indicator, achieving fault detection and threshold setting for gears; state-space models were used to estimate the RUL. Zhou et al. proposed an indirect degradation indicator construction method for estimating the rotating machinery's health conditions, in [32], and used a Monte Carlo approach to predict the RUL. Sun et al., in [33], presented the application of a state-space model for the prognostics of an engineering system subject to degradation. A health index was inferred from a set of sensor signals to characterize the hidden health state of the system. Zeng et al. proposed a new equipment degradation state recognition

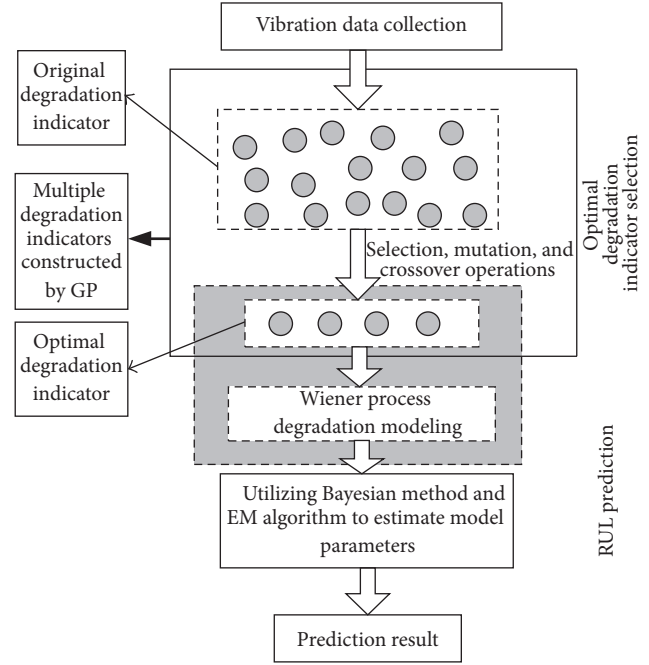


FIGURE 2: Remaining useful life prediction framework.

and fault prognostics method based on the kernel principal component analysis and hidden semi-Markov model in [34]. The kernel principal component analysis method was used to construct an optimal set of prognostic parameters while preserving the most important characteristics of the input data.

From the above survey of related works, we can observe that there is a continuing trend to develop an integrated method for RUL prediction, which can fully utilize the advantages of multiple methods. Owing, in part, to the low prediction accuracy of the rotating machinery RUL results, this naturally leads to our primary objective of this paper, which is to improve the RUL prediction accuracy for rotating machinery.

3. Remaining Useful Life Prediction Framework

To estimate the RUL of the rotating machinery, an integrated RUL prediction framework is proposed based on the GP and Wiener process degradation modeling, as shown in Figure 2. The proposed framework can be implemented as per the following steps.

Step 1 (vibration data collection). For monitoring rotating machinery, a number of sensors of various types (velocity transducer, acceleration transducer, and displacement transducer) are mounted on the bearings of the rotating machinery to measure the initial vibration signals. Run-to-failure vibration data are collected, providing an overall indication of the mechanical system's health. By calculating the probability density functions of the vibration signals, the most commonly used vibration parameters can be obtained.

Step 2 (optimal degradation indicator extraction). Based on the acquired condition monitoring data, the dimension parameters (e.g., root mean square, mean value, and peak value) and dimensionless parameters (e.g., impulse index, margin index, and kurtosis index) in the time domain are chosen as the original degradation indicators. Then, multiple degradation indicators are constructed through GP with the introduction of FPT. According to this approach, the initial features are combined and optimized, and optimal degradation indicators are formed after the run-to-failure data processing under the FPT. Then, the fitness function is adopted to measure the performances of the generated features. The optimal degradation indicator can well characterize the system health condition.

Step 3 (building Wiener-process-based degradation model). To achieve degradation modeling and RUL estimation for rotating machinery, the optimal degradation indicator is modeled as a Wiener process. The Wiener-process-based degradation is utilized to describe the equipment degradation process.

Step 4 (utilizing Bayesian method and EM algorithm to estimate model parameters). Based on the monitored degradation data, a parameter estimation approach for a degradation model obtained through the collaboration between Bayesian updating and the expectation maximization (EM) algorithm is presented. The Bayesian method is used to update the drift coefficient and the EM algorithm is utilized to update all other parameters. The obtained estimation in each iteration is unique and optimal.

Step 5 (prediction result). Finally, a practical case study is provided to show that the presented approach models the degradation process, estimates the model parameters, and generates a prediction result.

4. Brief Review of Genetic Programming Algorithm and Time-Domain Parameters

4.1. Genetic Programming Algorithm. GP has been proposed as a machine learning method in different fields, and it has the advantage of selecting and constructing features. The basic idea is described below. First, GP randomly creates an initial population (generation 0), which consists of a number of individuals in a tree structure. Then, a fitness function is assigned to calculate the fitness value of each individual. According to the principle of selecting the superior and eliminating the inferior, the proximate optimum solution or the optimal solution for one generation can be found by selecting genetic operators (selected operator, crossover operator, mutation operator, etc.) to optimize the population adaptively.

The tree structure is a common representation of a GP individual. Each individual is a math expression using the tree structure. As shown in Figure 3, the individual is represented as $\sqrt{z} \times (y + 2) + \ln x$.

The nodes of the tree are classified into two types. A node located inside the tree is the operator and the nodes at the

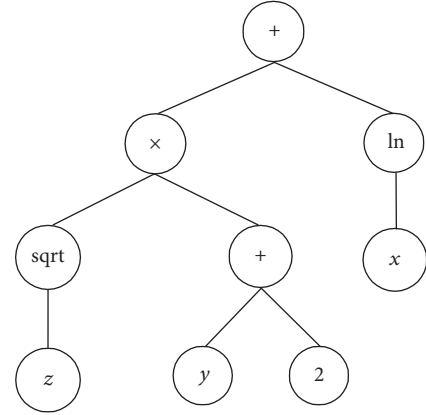


FIGURE 3: Form of genetic programming individual.

leaves of the tree are the terminators, making mathematical expressions easy to evolve and evaluate. For example, the operators in Figure 3 are the basic operators such as “+,” “x,” “ln,” and “sqrt.” The terminators contain variables “x,” “y,” and “z,” and the constant “2.” We can use these operators and terminators to combine more complicated expressions.

In practical application, a few details must be determined before running the GP algorithm:

- (1) Selection of the operator set: this paper uses “+,” “−,” “*,” “/,” “sqrt,” “exp,” “log,” and “fabs” as the operator set
- (2) Selection of the terminals set: five conventional dimensionless parameters are used as the terminals set in this paper
- (3) Design of the fitness function: the fitness function is a measure for evaluating the fitness levels of the individuals
- (4) Setting the population size, evolutionary generation, probability settings of genetic operators, and termination criterion

4.2. Time-Domain Parameters. Most of the existing prognostic techniques use condition-monitoring indexes to represent health. Using the vibration signals in the time domain to extract the degradation indicators is the most common method. In this paper, the dimension parameters and dimensionless parameters are utilized to characterize the dynamics and nonlinearity of the degradation progression. The dimension parameters are the square root of amplitude, root mean square, mean value, and peak value, which are defined as follows [35]:

$$x_d = \left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l} = \begin{cases} X_r, & l = \frac{1}{2}, \\ \bar{X}, & l = 1, \\ X_{rms}, & l = 2, \\ X_p, & l \rightarrow \infty, \end{cases} \quad (1)$$

where x denotes the vibration amplitude, $p(x)$ denotes the probability density function of vibration amplitude, X_r is the square root of amplitude, \bar{X} is the mean value, X_{rms} is the root mean square, and X_p is the peak value.

$$\begin{aligned} X_r &= \left[\int_{-\infty}^{+\infty} \sqrt{|x|} p(x) dx \right]^2 = \left(\frac{1}{N} \sum_{i=1}^N \sqrt{|x_i|} \right)^2, \\ \bar{X} &= \int_{-\infty}^{+\infty} |x| p(x) dx = \frac{1}{N} \sum_{i=1}^N |x_i|, \\ X_{\text{rms}} &= \sqrt{\int_{-\infty}^{+\infty} x^2 p(x) dx} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}, \\ X_p &= \lim_{l \rightarrow \infty} \left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l} = x_{\max}. \end{aligned} \quad (2)$$

The dimensionless parameter is a better diagnostic characteristic. The dimensionless parameters are defined as follows:

$$\zeta_x = \frac{\left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l}}{\left[\int_{-\infty}^{+\infty} |x|^m p(x) dx \right]^{1/m}}. \quad (3)$$

Specifically,

- (i) if $l = 2, m = 1$ have waveform index S_f ;
- (ii) if $l \rightarrow \infty, m = 1$ have impulse index I_f ;
- (iii) if $l \rightarrow \infty, m = 1/2$ have margin index CL_f ;
- (iv) if $l \rightarrow \infty, m = 2$ have peak index C_f ;
- (v) kurtosis index $K_v = \beta / X_{\text{rms}}^4$. In the formula, β is a dimension parameter named kurtosis, which is defined as $\beta = \int_{-\infty}^{+\infty} x^4 p(x) dx$. X_{rms} is a dimension parameter named root mean square.

In practice, the above-mentioned dimensionless parameters are frequently used. The dimension parameters vary in different work conditions and may be influenced by disturbances (e.g., speed, load, and sensitivity of the instrument) that cause data deviation. However, the dimensionless parameters are sensitive to faults, instead of work conditions. Moreover, they can classify some types of faults correctly and efficiently. In consideration of the advantages of dimension and dimensionless parameters, an approach to apply GP combining dimension parameters with dimensionless parameters is proposed to construct the optimal degradation indicator, which can characterize the degradation process more comprehensively.

5. Wiener Process Degradation Modeling with Random Effects

As mentioned above, many variants of the Wiener model have been reported in literature. In the model considered in this work, the parameters are estimated using Bayesian updating

and the EM algorithm. The Wiener process is typically used for modeling degradation processes where the degradation increases linearly in time with random noise. The rate of degradation is characterized by the drift coefficient.

5.1. Degradation Model. In general, the Wiener process degradation modeling with random effect can be represented as

$$X(t) = \beta t + \sigma B(t), \quad (4)$$

where β is the drift coefficient, σ is the diffusion coefficient, and $B(t)$ is the standard Brownian motion representing the stochastic dynamics of the degradation process. β is considered as a random variable following a normal distribution of $N(\mu_\beta, \sigma_\beta^2)$, which represents individual differences between the pieces of equipment. Thus, the estimated parameters in the model have μ_β , σ_β^2 , and σ^2 . For convenience, we denote $\Theta = (\mu_\beta, \sigma_\beta^2, \sigma^2)$ as a parameter vector to represent the model parameters.

Lifetime is usually defined as $T = \inf\{t \mid X(t) \geq \eta\}$, which means that the random degradation process $\{X(t), t \geq 0\}$ first reaches a prespecified failure threshold η . The probability density function and the expectation of lifetime T can be directly obtained as

$$\begin{aligned} f_T(t) &= \frac{\eta}{\sqrt{2\pi t^3 (\sigma_\beta^2 t + \sigma^2)}} \\ &\quad \times \exp\left(-\frac{(\eta - \mu_\beta t)^2}{2t (\sigma_\beta^2 t + \sigma^2)}\right), \\ E(T) &= E(E(T \mid \beta)) = E\left(\frac{\eta}{\beta}\right) \\ &= \frac{\eta}{\sigma_\beta^2} \exp\left(-\frac{\mu_\beta^2}{2\sigma_\beta^2}\right) \int_0^{\mu_\beta} \exp\left(-\frac{x^2}{2\sigma_\beta^2}\right) dx \\ &= \frac{\sqrt{2}\eta}{\sigma_\beta} D\left(\frac{\mu_\beta}{\sqrt{2}\sigma_\beta}\right), \end{aligned} \quad (5)$$

where $D(z) = \exp(-z^2) \int_0^z \exp(x^2) dx$ is the Dawson integral for a real z , which is known to exist.

5.2. Model Parameter Estimation Based on EM Algorithm [36]. To estimate the parameters, we assume that there are n tested items and that the degradation data $X_i = (x_{i1}, \dots, x_{im})'$ of the i th item are available, where x_{i1}, \dots, x_{im} denotes the degradation observations at time t_1, \dots, t_m . For simplicity, let $X = (X_1, \dots, X_n)'$ and $\Omega = (\beta_1, \dots, \beta_n)$, where X denotes the degradation dataset and β_i denotes the drift coefficient of the i th item.

For a given X_i and $\Theta^{(k)}$, the sampling of the i th item is distributed by the following expression:

$$f(X_i | \beta_i, \Theta) = \frac{1}{\prod_{j=1}^m \sqrt{2\pi\sigma^2\Delta t_j}} \times \exp\left(-\sum_{j=1}^m \frac{(\Delta x_{ij} - \beta_i \Delta t_j)^2}{2\sigma^2\Delta t_j}\right) \quad (6)$$

with $\Delta x_{ij} = x_{ij} - x_{i,j-1}$ and $\Delta t_j = t_j - t_{j-1}$, where Δx_{ij} denotes the degradation incremental of the i th item from time t_{j-1} to t_j and Δt_j denotes the time interval.

The log-likelihood function can be written as

$$\begin{aligned} \ell(\Theta | X, \Omega) &= \sum_{i=1}^n [\ln p(X_i | \beta_i, \Theta) + \ln p(\beta_i | \Theta)] \\ &= -\frac{1}{2} \sum_{i=1}^n \left[(m+1) \ln 2\pi + \sum_{j=1}^m \ln \Delta t_j + m \ln \sigma^2 \right. \\ &\quad \left. + \sum_{j=1}^m \frac{(\Delta x_{ij} - \beta_i \Delta t_j)^2}{\sigma^2 \Delta t_j} + \ln \sigma_\beta^2 + \frac{(\beta_i - \mu_\beta)^2}{\sigma_\beta^2} \right]. \end{aligned} \quad (7)$$

Let us assume that $\Theta^{(k)} = \{\mu_\beta^{(k)}, \sigma_\beta^{2(k)}, \sigma^{2(k)}\}$ is the estimate in the i th step based on X . With X_i and $\Theta^{(k)}$ known, the posterior distribution of β_i will still be normal; that is, $\beta_i \sim N(\mu_i^{(k)}, \sigma_i^{2(k)})$.

In the Bayesian framework, the posterior distribution of β_i can be updated via the Bayesian rule as follows:

$$\begin{aligned} p(\beta_i | X_i, \Theta^k) &\propto p(X_i | \beta_i, \Theta^k) p(\beta_i | \Theta^k) \\ &\propto \exp\left[-\sum_{j=1}^m \frac{(\Delta x_{ij} - \beta_i \Delta t_j)^2}{2\sigma^{2(k)}\Delta t_j}\right] \exp\left[-\frac{(\beta_i - \mu_\beta^{(k)})^2}{2\sigma_\beta^{2(k)}}\right] \\ &\propto \exp\left\{-\frac{1}{2} \left[\sum_{j=1}^m \frac{\beta_i^2 \Delta t_j}{\sigma^{2(k)}} - \sum_{j=1}^m \frac{2\beta_i \Delta x_{ij}}{\sigma^{2(k)}} + \frac{\beta_i^2 - 2\beta_i \mu_\beta^{(k)} + (\mu_\beta^{(k)})^2}{\sigma_\beta^{2(k)}} \right]\right\} \\ &\propto \exp\left\{-\frac{1}{2} \left[\left(\frac{t_m}{\sigma^{2(k)}} - \frac{1}{\sigma_\beta^{2(k)}} \right) \beta_i^2 - 2 \left(\frac{x_{im}}{\sigma^{2(k)}} + \frac{\mu_\beta^{(k)}}{\sigma_\beta^{2(k)}} \right) \beta_i \right]\right\} \\ &\propto \exp\left\{-\frac{\left[\beta_i - (x_{im}\sigma_\beta^{2(k)} + \mu_\beta^{(k)}\sigma^{2(k)}) / (t_m\sigma_\beta^{2(k)} + \sigma^{2(k)}) \right]^2}{2\sigma^{2(k)}\sigma_\beta^{2(k)} / (t_m\sigma_\beta^{2(k)} + \sigma^{2(k)})}\right\}. \end{aligned} \quad (8)$$

Owing to the property of the normal distribution of $\beta_i | X_i, \Theta^k$, we obtain

$$p(\beta_i | X_i, \Theta^k) = \frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}} \exp\left[-\frac{(\beta_i - \mu_i^{(k)})^2}{2\sigma_i^{2(k)}}\right] \quad (9)$$

with

$$\begin{aligned} \mu_i^{(k)} &= \frac{x_{im}\sigma_\beta^{2(k)} + \mu_\beta^{(k)}\sigma^{2(k)}}{t_m\sigma_\beta^{2(k)} + \sigma^{2(k)}}, \\ \sigma_i^{2(k)} &= \frac{\sigma_\beta^{2(k)}\sigma^{2(k)}}{t_m\sigma_\beta^{2(k)} + \sigma^{2(k)}}, \end{aligned} \quad (10)$$

where we can learn that the posterior estimation of β_i can be easily updated once a new observation is available. Now, let us focus on calculating the maximum-likelihood estimation $\hat{\Theta} = (\hat{\mu}_\beta, \hat{\sigma}_\beta^2, \hat{\sigma}^2)$ using the EM algorithm. Θ can be estimated through two steps: the E-step and the M-step.

In the E-step, the expectation $E[\ell(\Theta | X, \Theta^{(k)})]$ can be computed as follows:

$$\begin{aligned} E[\ell(\Theta | X, \Theta^{(k)})] &= -\frac{1}{2} \sum_{i=1}^n \left[(m+1) \ln 2\pi + \sum_{j=1}^m \ln \Delta t_j + m \right. \\ &\quad \cdot \ln \sigma^2 \\ &\quad \left. + \sum_{j=1}^m \frac{(\Delta x_{ij})^2 - 2\mu_i^{(k)}\Delta x_{ij}\Delta t_j + (\Delta t_j)^2((\mu_i^{(k)})^2 + \sigma_i^{2(k)})}{\sigma^2 \Delta t_j} \right. \\ &\quad \left. + \ln \sigma_\beta^2 + \frac{(\mu_i^{(k)})^2 + \sigma_i^{2(k)} - 2\mu_i^{(k)}\mu_\beta + \mu_\beta^2}{\sigma_\beta^2} \right]. \end{aligned} \quad (11)$$

Then, in the M-step, letting $\partial E[\ell(\Theta | X, \Theta^{(k)})]/\partial \theta = 0$, we obtain $\Theta^{(k+1)}$ as follows:

$$\begin{aligned} \mu_\beta^{(k+1)} &= \frac{1}{n} \sum_{i=1}^n \mu_i^{(k)}, \\ \sigma_\beta^{2(k+1)} &= \frac{1}{n} \sum_{i=1}^n \left[(\mu_i^{(k)})^2 + \sigma_i^{2(k)} - 2\mu_i^{(k)}\mu_\beta^{(k+1)} + (\mu_\beta^{(k+1)})^2 \right], \\ \sigma^{2(k+1)} &= \frac{1}{nm} \\ &\quad \cdot \sum_{i=1}^n \sum_{j=1}^m \frac{(\Delta x_{ij})^2 - 2\mu_i^{(k)}\Delta x_{ij}\Delta t_j + (\Delta t_j)^2((\mu_i^{(k)})^2 + \sigma_i^{2(k)})}{\Delta t_j}. \end{aligned} \quad (12)$$

The above steps are iterated multiple times to produce a sequence $\{\Theta^{(0)}, \Theta^{(1)}, \Theta^{(2)}, \dots\}$ of increasingly good approximations $\hat{\Theta} = (\hat{\mu}_\beta, \hat{\sigma}_\beta^2, \hat{\sigma}^2)$. For each iteration, the analytical solution for updating the model parameters is derived. The iterations are usually terminated when the EM algorithm converges.

6. Experimental Demonstrations

In this section, we provide a practical case study to illustrate the application of our model and compare the performance of our model with that of other models.



FIGURE 4: Test bed.



FIGURE 5: Rolling element bearing.

6.1. Introduction to the Experimental Equipment and Data Acquisition. The bearing is a key device in rotating machinery, and its operating state has a direct influence on the rotating machinery condition. The rotating machinery operation at very high speeds can lead to bearing wear. As the wear accumulates, the bearing will become deformed and such deformation may lead to incipient faults in the rotating machinery. The increasing faults result in the failure of the bearing and of the rotating machinery. Past data show that almost 80% of the failures in rotating machinery result from the wear of rolling element bearings, which were extensively investigated in the literature.

A picture of a test rig is provided in Figure 4, and it shows the major components of the test rig, such as the industrial motor, shaft, and test bearing. The bearings are instrumented with accelerometers in both the axial and radial directions on the bearing housing. An illustration of a bearing is provided in Figure 5.

We use the vibration accelerometer to gather the vibration signals. The sampling frequency is 1 kHz and each sample contains 4096 data points. The failure threshold is determined by the vibration level ISO 2372 and ISO 10816. The rotation speed is maintained at 1800 RPM. A radial load of 2.5 MPa pressure is applied onto the bearing, and twelve failure data of bearings are collected. Among them, eight bearings resulted in outer race failures, which are

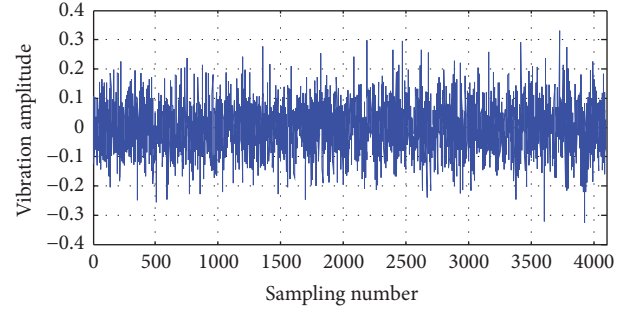


FIGURE 6: Vibration wave under the normal condition.

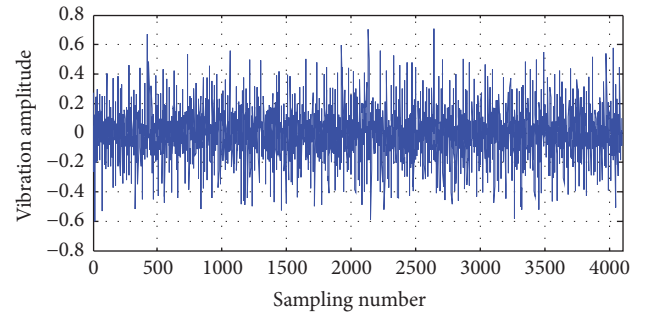


FIGURE 7: Vibration wave in the failure stage.

TABLE 1: Failure data of bearings.

Bearing number	1	2	3	4	5	6	7	8
Failure time (h)	900	918	954	876	922	894	988	1002

summarized in Table 1. The failure times are recorded as the times at which the observed values cross the threshold. The total life expectancy of the eight bearings is 932 h.

The monitoring data of the first sample are illustrated in Figures 6 and 7 and show the time-domain vibration waves under the normal condition and in the failure stage, respectively. For our monitored rotating machinery with a terminated life of 900 h, 450 monitoring data were collected at regular condition monitoring intervals of 2 h.

By calculating the vibration signal, four dimension parameters and five dimensionless parameters were measured. For convenience, the square root of amplitude, mean value, root mean square, peak value, waveform index, impulse index, margin index, peak index, and kurtosis index are represented by S , M , R , P , W , I , T , L , and K , respectively. By visually inspecting the nine features, no obvious trend was found in an individual feature. Among these features, the square root of amplitude showed the best increasing trend, as shown in Figure 8. It showed a smooth trend at the beginning, which, however, fluctuated toward the end of life.

The degradation parameter possessing the worst monotonicity was the mean value, as shown in Figure 9. A sudden increase in the mean value appeared only just before the bearing failed.

6.2. Optimal Degradation Indicator Extraction. The FPT selection results are shown in Figures 10 and 11. It is observed

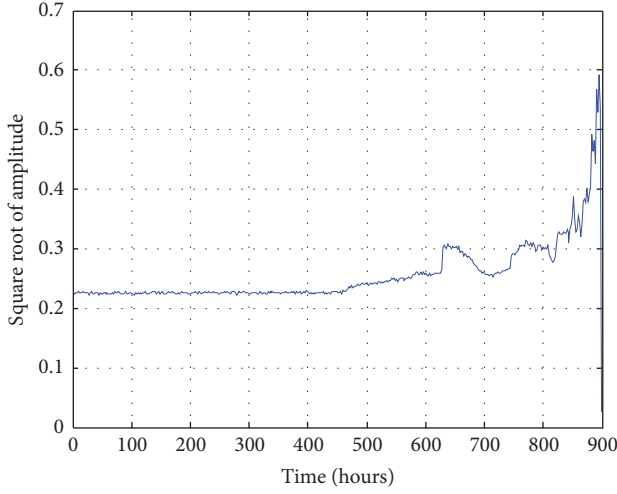


FIGURE 8: Square root of amplitude.

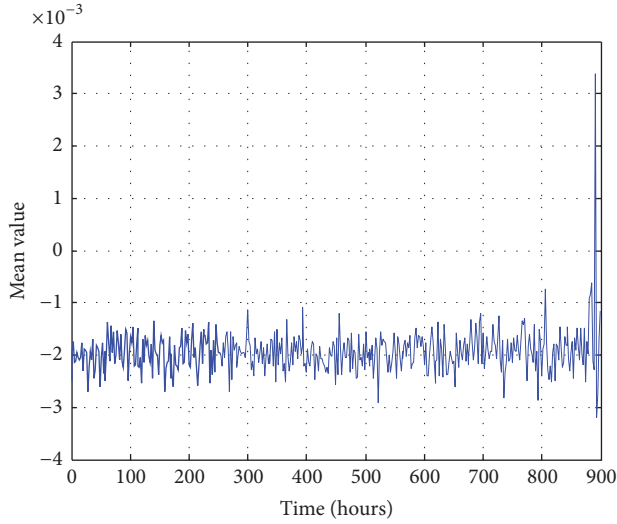
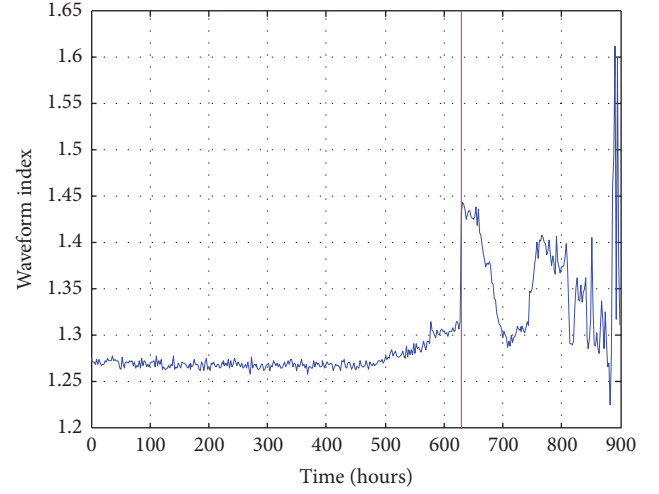
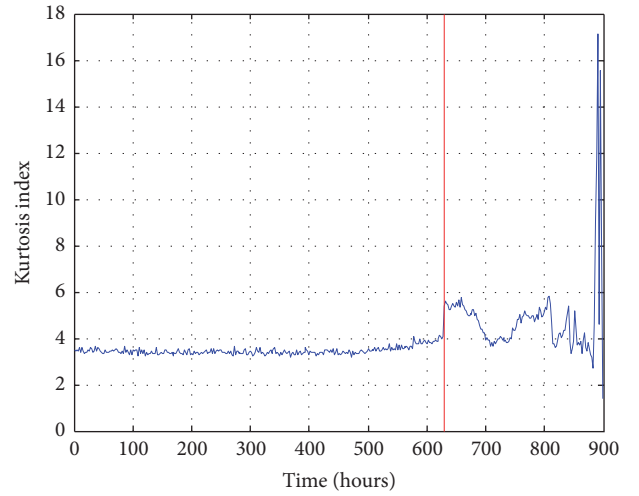


FIGURE 9: Mean value.



— Degradation process
— FPT

FIGURE 10: Waveform index.



— Degradation process
— FPT

FIGURE 11: Kurtosis index.

that the waveform index and kurtosis index are sensitive to the incipient faults in the bearing. The waveform index and kurtosis index were divided into two stages by the FPT. Before FPT, the values of the waveform index and kurtosis index were stable; however, they increased suddenly after the FPT. Therefore, $FPT = 630$ h indicates the initial time of the degradation process. The degradation data after FPT are used to construct new degradation parameters.

The specific steps to achieve an optimal degradation indicator for application in RUL prediction are as follows.

Step 1. Experimental dataset acquisition is as follows. Run-to-failure data of bearings are obtained.

Step 2. Basic parameters of GP are determined. Table 2 lists the basic parameters of GP.

Step 3. M initial individuals are generated randomly. The initial individual is expressed in $N = f(S, M, R, P, W, I, M, L, K)$ represented new degradation parameter, which means

that the new degradation parameter is a function that depends on the square root of amplitude, mean value, root mean square, peak value, waveform index, impulse index, margin index, peak index, and kurtosis index.

Step 4. The data of the degradation index after $FPT = 630$ h are computed as the inputs to N . The monotonicity is used for evaluating the fitness levels of the individuals. In this paper, the fitness function [15] is defined by

$$\text{fitness} = \left| \frac{\# \text{of } d/dF > 0}{n-1} - \frac{\# \text{of } d/dF < 0}{n-1} \right|, \quad (13)$$

where n is the number of observations in a period. F represents a feature and d/dF is the derivative. The larger

TABLE 2: Basic parameters of GP.

Parameters	Settings
Object	Building the new degradation indicator
Set of terminals	S (square root of amplitude), M (mean value), R (root mean square), P (peak value), W (waveform index), I (impulse index), T (margin index), L (peak index), K (kurtosis index)
Set of functions	$+$, $-$, $*$, $/$, \exp , \log , $\sqrt{}$
Parameters	Population size $M = 1000$ Evolutionary generation $G = 50$
Probability settings of genetic operator	Crossover probability 0.85 Mutation probability 0.15
Method of selection	Tournament selection method, the size of tournament is five
Termination criterion	Required maximum of evolutionary generation
Max depth of tree	Six

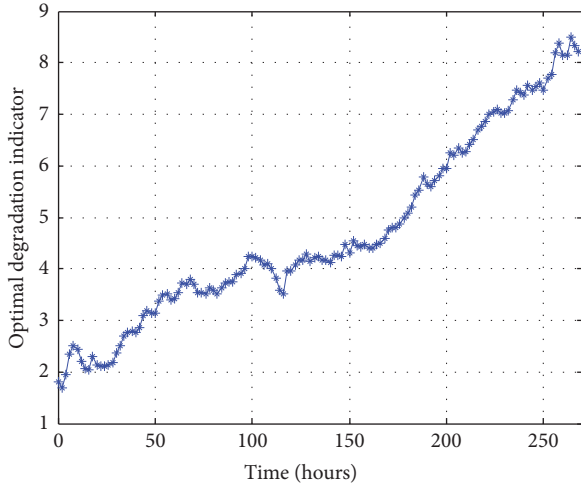


FIGURE 12: Optimal degradation indicator.

the fitness value, the better the monotonicity of the new degradation parameter.

Step 5. According to the genetic parameters, new individuals are generated using the following series of actions: reproduction, crossover, and mutation.

Step 6. Steps 3, 4, and 5 are executed repeatedly until the algorithm running termination criterion is met.

Analyzing the operation results of GP, the best individual is

$$N = \ln(T - e^S) + \frac{(W/P + R * \sqrt{T})}{S}, \quad (14)$$

where S , R , P , W , and T represent the square root of amplitude, root mean square, peak value, waveform index, and margin index, respectively.

Figure 12 shows the plot of the optimal degradation indicator obtained from the FPT-to-failure test. A clear increasing trend is exhibited by the optimal degradation parameter, from

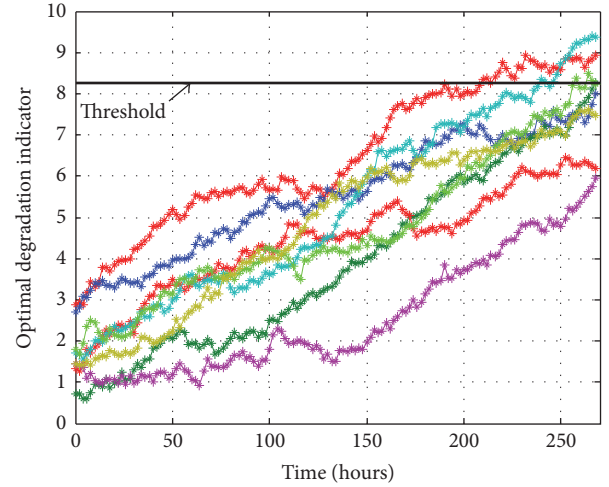


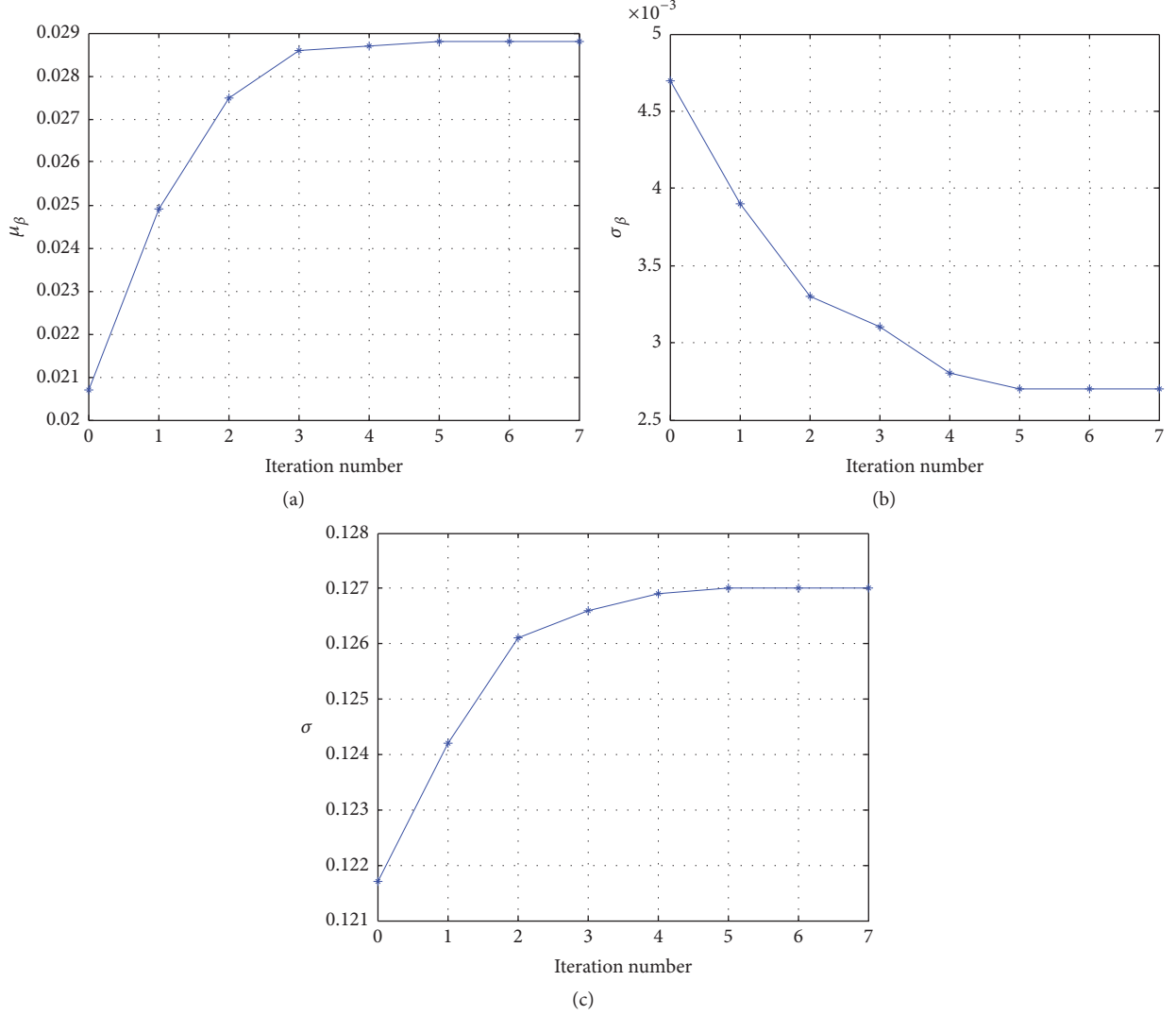
FIGURE 13: Calculated optimal degradation indicator data of eight bearings.

the FPT concept. RUL prediction can be implemented by modeling the optimal degradation indicator.

7. RUL Prediction

For illustration, the optimal degradation parameter is used as the prediction index to collect the monitoring data for the remaining seven bearings. The collected monitoring data of eight bearings are illustrated in Figure 13 and are used to demonstrate the method developed in this work. The values of the optimal degradation indicator after $t = 630$ h are input into the Wiener model for RUL prediction. Now, we model the degradation process as a Wiener process based on the optimal degradation indicator. The failure threshold is 8.25. When the degradation value increases to 8.25, the bearing will be considered to have failed.

Based on the data shown in Figure 13 and the parameter estimation method presented in the above section (parameter estimation is illustrated after each iteration until convergence in Figure 14), we can observe that the parameters are stable

FIGURE 14: Parameter estimation: (a) μ_β , (b) σ_β , and (c) σ .

over four iterations: $\hat{\mu}_\beta = 0.0288$, $\hat{\sigma}_\beta = 0.0027$, and $\hat{\sigma} = 0.1270$.

Once the estimated parameters are obtained, the expectation of RUL can be calculated.

$$E(T) = \frac{\sqrt{2}\eta}{\sigma_\beta} D\left(\frac{\mu_\beta}{\sqrt{2}\sigma_\beta}\right) = 289. \quad (15)$$

The entire life is calculated with $T_{\text{EoL}} = \text{FPT} + \text{RUL}$. We can predict that the entire life of the bearing is 919 h, which agrees well with the actual mean experimental results. This demonstrates that our method can accurately estimate the RUL.

For further verification of the effectiveness of our model for RUL prediction, the Wiener-model-based conventional degradation indicator using the FPT is used to predict the RUL of the eight bearings as well. The model is referred to as Model 1 in this paper. The RMS is used to construct Model 1. We have $\hat{\mu}_\beta = 0.0427$, $\hat{\sigma}_\beta = 0.0210$, and $\hat{\sigma} = 0.0152$

through the parameter estimations, and the prediction result is calculated as 705 h for Model 1, which has a relatively large difference with the experimental result. We can see that the obtained result of our model is better than the one given by the Wiener model based on the conventional prognostic features. This demonstrates that our model can improve the accuracy of RUL estimation. In addition, the estimated mean degradation path based on the established model can be expressed as $E[X(t)] = \hat{\mu}_\beta t$. The estimated mean degradation paths using our model and Model 1 are shown in Figure 15.

We can see from Figure 15 that the predicted results of our model gradually match with the actual sample mean, which illustrates that our model has a good fitting degree. However, Model 1 does not consider the monotonicity of the degradation index, which makes the RUL demonstrate a greater variation than the actual mean experimental result. The experiment demonstrates that our proposed integrated RUL prediction method can work well and efficiently.

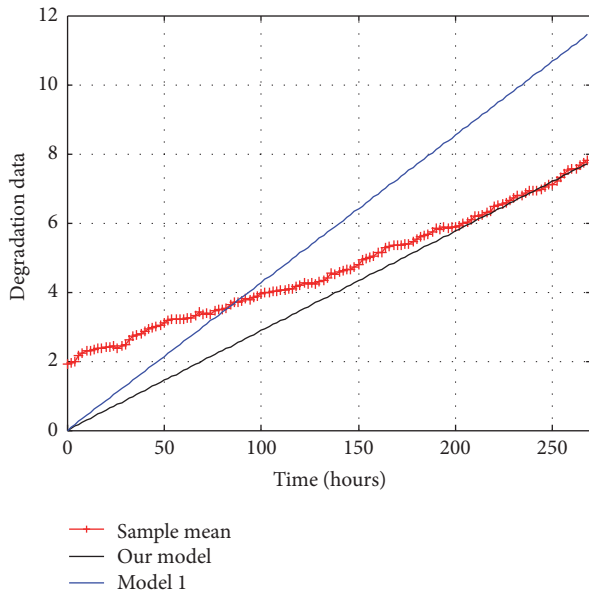


FIGURE 15: Comparison of the estimated mean degradation paths.

8. Conclusions

In order to improve the prediction accuracy of rotating machinery, an integrated RUL prediction method based on GP and Wiener process degradation modeling was proposed in this paper. The GP algorithm was used to find a better degradation indicator using the concept of FPT. By selecting an appropriate FPT value, small amounts of data could be used to construct the optimal degradation indicator, and a better degradation indicator could be obtained. A Wiener model was proposed for RUL prediction based on the obtained optimal degradation indicator. As the input to the degradation model, the optimal degradation indicator fusing multiple vibration features could contain more vibration signatures and provide a more noticeable trend than the conventional degradation indicators. Using the measured data, the parameter estimations for the stochastic parameters in the model were updated recursively, by using the condition monitoring observations based on the Bayesian method and EM. The expectation of the predicted RUL was calculated based on the proposed model and the estimated mean degradation path was explicitly derived. Experimental results indicated that the method could effectively improve the accuracy of RUL prediction.

Although we demonstrated the usefulness of our proposed model, there are still many open problems that must be studied. For example, in this case study, eight bearings that were tested resulted in outer race failures. This implied that the health and prediction algorithm was evaluated with respect to only one failure mode of the bearing. In reality, the defect can be either an inner race defect or a roller element defect. Our method has not been fully evaluated with respect to multiple failure modes. In addition, the presented model needs multiple history datasets of similar equipment, which may take a long time to obtain. These problems require considerable further research.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was partially supported by the NSFC under Grants no. 61473094 and no. 61673127, the International and Hong Kong, Macao, & Taiwan Collaborative Innovation Platform, and Major International Cooperation Projects of Colleges in the Guangdong Province (no. 2015KGJHZ026). The work described in this paper was also partially supported by the open foundation of the Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis under Grants no. GDUPTKLAB201604 and no. GDUPTKLAB201603, the Technical Project of Maoming City (no. 201618), and the Guangdong Province Natural Science Fund Project (no. 2016A030313823).

References

- [1] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: state of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724–739, 2009.
- [2] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [3] A. S. Y. Heng, *Intelligent prognostics of machinery health utilising suspended condition monitoring data [Ph.D. thesis]*, Queensland University of Technology, Brisbane, Australia, 2009.
- [4] Z.-Q. Wang, C.-H. Hu, W. Wang, and X.-S. Si, "An additive wiener process-based prognostic model for hybrid deteriorating systems," *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 208–222, 2014.
- [5] J. Z. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1803–1836, 2011.
- [6] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation—a review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, 2011.
- [7] Q. Guan, Y. Tang, and A. Xu, "Objective Bayesian analysis accelerated degradation test based on Wiener process models," *Applied Mathematical Modelling*, vol. 40, no. 4, pp. 2743–2755, 2016.
- [8] Z. Huang, Z. Xu, W. Wang, and Y. Sun, "Remaining useful life prediction for a nonlinear heterogeneous wiener process model with an adaptive drift," *IEEE Transactions on Reliability*, vol. 64, no. 2, pp. 687–700, 2015.
- [9] X. Wang, N. Balakrishnan, and B. Guo, "Residual life estimation based on a generalized Wiener degradation process," *Reliability Engineering and System Safety*, vol. 124, pp. 13–23, 2014.
- [10] X.-S. Si, W. Wang, C.-H. Hu, M.-Y. Chen, and D.-H. Zhou, "A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 219–237, 2013.

- [11] Z. Tian, "An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring," *Journal of Intelligent Manufacturing*, vol. 23, no. 2, pp. 227–237, 2012.
- [12] Z.-X. Zhang, X.-S. Si, and C.-H. Hu, "An age- and state-dependent nonlinear prognostic model for degrading systems," *IEEE Transactions on Reliability*, vol. 64, no. 4, pp. 1214–1228, 2015.
- [13] J. P. P. Gomes, R. K. H. Galvao, T. Yoneyama, and B. P. Leao, "A new degradation indicator based on a statistical anomaly approach," *IEEE Transactions on Reliability*, vol. 65, no. 1, pp. 326–335, 2016.
- [14] D. Liu, J. Zhou, H. Liao, Y. Peng, and X. Peng, "A health indicator extraction and optimization framework for lithium-ion battery degradation modeling and prognostics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 6, pp. 915–928, 2015.
- [15] L. Liao, "Discovering prognostic features using genetic programming in remaining useful life prediction," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2464–2472, 2014.
- [16] N. Li, Y. Lei, J. Lin, and S. X. Ding, "An improved exponential model for predicting remaining useful life of rolling element bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7762–7773, 2015.
- [17] X. Wang, "Wiener processes with random effects for degradation data," *Journal of Multivariate Analysis*, vol. 101, no. 2, pp. 340–351, 2010.
- [18] Z.-S. Ye, Y. Wang, K.-L. Tsui, and M. Pecht, "Degradation data analysis using wiener processes with measurement errors," *IEEE Transactions on Reliability*, vol. 62, no. 4, pp. 772–780, 2013.
- [19] X.-S. Si, W. Wang, M.-Y. Chen, C.-H. Hu, and D.-H. Zhou, "A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution," *European Journal of Operational Research*, vol. 226, no. 1, pp. 53–66, 2013.
- [20] X.-S. Si, W. Wang, C.-H. Hu, D.-H. Zhou, and M. G. Pecht, "Remaining useful life estimation based on a nonlinear diffusion degradation process," *IEEE Transactions on Reliability*, vol. 61, no. 1, pp. 50–67, 2012.
- [21] W. Wang and M. Carr, "An adapted Brownian motion model for plant residual life prediction," in *Proceedings of the Prognostics and System Health Management Conference (PHM '10)*, pp. 1–7, Macao, January 2010.
- [22] S.-T. Tseng and C.-Y. Peng, "Optimal burn-in policy by using an integrated Wiener process," *IIE Transactions*, vol. 36, no. 12, pp. 1161–1170, 2004.
- [23] J. Balka, A. F. Desmond, and P. D. McNicholas, "Review and implementation of cure models based on first hitting times for Wiener processes," *Lifetime Data Analysis*, vol. 15, no. 2, pp. 147–176, 2009.
- [24] S. Mishra and O. A. Vanli, "Remaining useful life estimation with lamb-wave sensors based on wiener process and principal components regression," *Journal of Nondestructive Evaluation*, vol. 35, no. 1, article 11, 2016.
- [25] S. Tang, X. Guo, and Z. Zhou, "Mis-specification analysis of linear Wiener process-based degradation models for the remaining useful life estimation," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 228, no. 5, pp. 478–487, 2014.
- [26] S.-J. Tang, X.-S. Guo, C.-Q. Yu, Z.-J. Zhou, Z.-F. Zhou, and B.-C. Zhang, "Real time remaining useful life prediction based on nonlinear Wiener based degradation processes with measurement errors," *Journal of Central South University*, vol. 21, no. 12, pp. 4509–4517, 2014.
- [27] Z.-Q. Wang, C.-H. Hu, W. Wang, and X.-S. Si, "An additive wiener process-based prognostic model for hybrid deteriorating systems," *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 208–222, 2014.
- [28] M. Kotani, S. Ozawa, M. Nakai, and K. Akazawa, "Emergence of feature extraction function using genetic programming," in *Proceedings of the 3rd International Conference on Knowledge-Based Intelligent Information Engineering Systems (KES '99)*, pp. 149–152, September 1999.
- [29] D. P. Muni, N. R. Pal, and J. Das, "Genetic programming for simultaneous feature selection and classifier design," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 36, no. 1, pp. 106–117, 2006.
- [30] H. Guo, L. B. Jack, and A. K. Nandi, "Feature generation using genetic programming with application to fault classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 1, pp. 89–99, 2005.
- [31] S. Kadry, "Diagnostics and prognostic of engineering systems: methods and techniques," in *Data Driven Prognostics for Rotating Machinery*, E. Bechhoefer, Ed., chapter 6, pp. 1–13, IGI Global, Hershey, Pa, USA, 2012.
- [32] Y. Zhou, Y. Sun, J. Mathew, R. Wolff, and L. Ma, "Latent degradation indicators estimation and prediction: a Monte Carlo approach," *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 222–236, 2011.
- [33] J. Sun, H. Zuo, W. Wang, and M. G. Pecht, "Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 28, no. 3, pp. 585–596, 2012.
- [34] Q. Zeng, J. Qiu, G. Liu, and X. Tan, "Research on equipment degradation state recognition and fault prognostics method based on KPCA-hidden semi-Markov model," *Chinese Journal of Scientific Instrument*, vol. 30, no. 7, pp. 1341–1346, 2009.
- [35] Q. H. Zhang, Q. Hu, G. Sun et al., "Concurrent fault diagnosis for rotating machinery based on vibration sensors," *International Journal of Distributed Sensor Networks*, vol. 9, no. 1, pp. 59–72, 2013.
- [36] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–38, 1977.

