

A Novel Product Remaining Useful Life Prediction Approach Considering Fault Effects

Jingdong Lin, Zheng Lin, Guobo Liao, and Hongpeng Yin

Abstract—In this paper, a novel remaining useful life prediction approach considering fault effects is proposed. The Wiener process is used to construct the degradation process of single performance characteristic with the fault effects. The first passage time based remaining useful life distribution is calculated by assuming fault occurrence moment is a random variable and follows a certain distribution. Expectation maximization algorithm is employed to estimate model parameters, where the fault occurrence moment is considered as a missing data. Finally, a Copula function is used to describe the dependence between the multiple performance characteristics and derive joint remaining useful life (RUL) distribution of product with the fault effects. The effectiveness of the proposed approach is verified by the experiments of turbofan engines.

Index Terms—Degradation process, fault effects, fault occurrence moment (FOM), performance characteristic (PC), remaining useful life (RUL).

I. INTRODUCTION

REMAINING useful life (RUL) prediction is a key technology of ensuring the product's reliability and safety, it is also an effective path to reduce the maintenance costs for practical product [1], [2]. Therefore, the researches on the RUL prediction methods have drawn much attention from both academia and engineering circles in recent years. In general, RUL prediction means to find the probability distribution function (PDF) of the RUL or the expectation value of the RUL [3], [4]. Most of products have their own performance characteristics (PCs) that increase or decrease over time, and PCs are collectively referred to as degradation data. As degradation data is directly related to the health condition of the products, the prediction methods based on degradation data become more and more popular [5]–[8].

In past few decades, there are many methods of RUL prediction based on degradation data that have been researched, these are generally summarized into three types:

Manuscript received May 27, 2020; revised June 28, 2020 and July 26, 2020; accepted August 28, 2020. This work was supported by General Program of National Natural Science Foundation of China (61773080), China Central Universities Foundation (2019CDYGD001), Scientific Reserve Talent Programs of Chongqing University (cqu2018CDHB1B04), and Graduate Research and Innovation Foundation of Chongqing (CYB20065). Recommended by Associate Editor Xiaou Li. (Corresponding author: Hongpeng Yin.)

Citation: J. D. Lin, Z. Lin, G. B. Liao, and H. P. Yin, "A novel product remaining useful life prediction approach considering fault effects," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 11, pp. 1762–1773, Nov. 2021.

The authors are with the School of Automation, Chongqing University, Chongqing 400044, China (e-mail: linzhending@163.com; 18380581987@163.com; 727187457@qq.com; yinhongpeng@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JAS.2021.1004168

physical, data-driven, and hybrid methods [9]. Among these categories, the methods of data-driven become the mainstream direction in research of RUL prediction, because the methods only need to establish a statistical model to fit the observations without any assumption of physical parameters or additional expertise. Machine learning has been extensively studied in the past decade, it does not need to understand the degradation mechanism of the product, but uses a large amount of data to train the prediction model. In order to improve the accuracy of remaining life prediction, Gebraeel and Lawley [10] were the first to use the trained neural network to calculate the remaining life distribution function of the device based on the monitored data. Ren *et al.* [11] used the self-encoding multi-dimensional feature extraction to characterize the health degradation trend of the battery, and then trained the life prediction model based on the deep neural network, and successfully applied the method to the prediction of the remaining life of the lithium battery. Wu *et al.* [12] used the least squares support vector machine (SVM) method to fit the performance degradation process of equipment for the problem of only a small number of samples in reliability analysis. Tran *et al.* [13] used the minimum mean variance to establish a degenerated index of equipment, and then used the established index as training data to complete the remaining life prediction test of related mechanical equipment by combining with SVM. However, because of random and dynamic characteristics of product's degradation process, the stochastic-process based prediction approaches have attracted great attention in the data-driven methods. According to Si *et al.* [14], the data-driven based RUL prediction methods are classified into two main types: indirect data and direct data based methods, and the scope of application of stochastic process in the prediction of RUL is illustrated in detail. The stochastic process mainly includes Markov chains, Wiener processes, and Gamma processes. Among them, for its important physical interpretation and nice mathematical properties, Wiener process can describe non-monotonic degradation process of many typical products. It has been widely applied to establish the degradation process and predict the RUL of a variety of industrial components. Tseng *et al.* [15] used a Wiener process to determine the lifetime for the light intensity of light emitting diode (LED) lamps of contact image scanners. The Wiener process was even extended to the adaptive Wiener process model by Zhai and Ye [16]. More detail about Wiener process applied in RUL prediction can refer to [17].

In reality, many products usually have complex structure,

and consist of multiple components or a component having multiple PCs. Compared to single PC, multiple PCs can provide more accurate information for RUL prediction. In such a situation, a bivariate or multivariate degradation model is needed to estimate RUL of products, where the existing methods can be roughly classified into multivariate distribution, multivariate stochastic process and Copula function. Multivariate normal distribution as a popular method has been utilized to identify a multivariate joint probability [18]. Wang and Coit [19] conducted a degradation analysis of a system with multiple degradation measurements, and introduce a multivariate s-normal distribution model. Multivariate or bivariate degradation process models based on Gamma processes, Wiener processes are introduced by [20]–[23]. In practice, assuming degradation data follows a multivariate distribution or multivariate stochastic process may not be suitable for all conditions, when degradation mechanisms of multiple PCs are not consistent or independent. As a useful alternative, Copula function is employed to identify a multivariate joint probability [24]. Due to its flexibility, Copula function has been widely used in RUL prediction of a product with multiple PCs. Sari *et al.* [25] investigated the degradation mechanisms of LED lamps, and presented a bivariate degradation model based on Copula function. Copula function is widely applied in degradation modeling, reliability analysis and prognostics, more introductions are in [26]–[29].

In practice, with the wear of the components, overload operation, or the changes of environment factors, many different kinds of faults may occur in the product's degradation process. Most of these faults are not equal to the functional failure, product can continue to work, but the occurrence of the fault will aggravate its degradation degree and change the product's degradation trajectory, ultimately shorten the product's RUL. Meanwhile, there are many researches about that some product's degradation path will be changed at a moment and even more times during the degradation process. More specially, after a product has degraded down to a certain level (which can be called transition time), the initial function form of the condition monitoring signal may be changed, evolve more rapidly or gradually compared to the early stage of usage. It is a quite common phenomenon that degradation trajectory changes during the degradation process for display devices, such as plasma display panels, LED, and organic light emitting diode (OLED) studied by Tang *et al.* [30] and Ng [31]. Si *et al.* [32], research the RUL prediction of systems with operation state switching between the working state and storage state, and the degradation rates of two states are different, the transition time and transition number are derived. Cui *et al.* [33] assumed that the changing location was determinate, and the Wiener process is employed to model the degradation process, the first passage time of the degradation process governed by a discontinuous trend function is studied.

However, the fault is hard to be detected, if it occurs after observations, it means the fault may not occur in observations

but in a future span, and its probability of occurrence increases over time. Therefore, compared to the operation state switches and determinate changing location, the transition time from normal state to fault state (which can be called fault occurrence moment (FOM)) is difficult to be detected directly, so the changing time of the degradation rate with the fault effects is unobservable and random. The effects of a fault in the degradation process cannot be neglected, because the occurrence of FOM will accelerate the failure of the product. Moreover, very few efforts have been made to address the RUL prediction which involves fault effects. Besides, when a fault occurs in the degradation process, the degradation rates of multiple PCs will be changed at the same time, which is shown in Fig. 1. In addition, PCs may be dependent because of the strong coupling between the product's components. Therefore, our final goal is to estimate the RUL joint distribution of multiple dependent PCs considering fault effects.

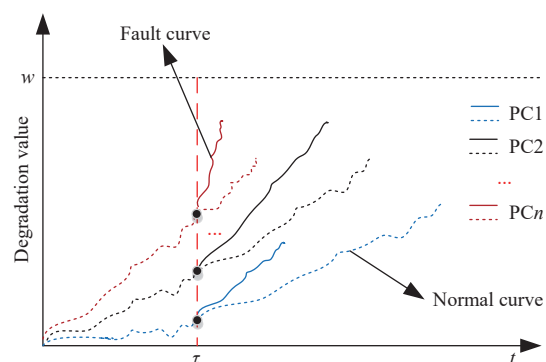


Fig. 1. The degradation process of multiple performance characteristics for two states.

In this paper, in order to solve the problem of the product RUL prediction with the fault effects, the proposed approach is divided into two parts. First, a degradation model based on the Wiener process is used to account for the changing of degradation process. In the presented model, the drift coefficient is utilized to describe the fault effects for the whole degradation path, and the diffusion coefficient characterizes the stability of the degradation process respectively. Then, the RUL distribution of the product based on the definition of first passage time is obtained, where it is assumed that FOM is a random variable and follows a certain distribution. In this work, the FOM is regarded as a missing-data in the observations since the FOM can not be detected directly, expectation maximization (EM) algorithm is employed to tackle the problem of parameter estimation. Then, the product RUL for single PC considering fault effects can be estimated. Second, as multiple PCs will be influenced by the fault at the same time, and may be dependent during the degradation process. A Copula function is used to describe the dependence between multiple PCs and derive the RUL joint distribution of the product based on the RUL distribution of single PC with the fault effects.

The main contributions of this paper are summarized as follows:

1) A novel product RUL prediction approach is proposed for the problem that degradation process is impacted by the fault effects, which considers the FOM as a random variable and follows a certain distribution. To the best of our knowledge, this is the first time to consider fault effects into the RUL prediction.

2) As FOM cannot be detected directly after observations, which causes the incomplete observations, FOM is regarded as a missing-data in problem of parameter estimation. EM algorithm is used to estimate parameters of degradation model and fault distribution simultaneously.

3) For the multiple PCs will be affected by the fault and may be dependent during the degradation process, Copula function is employed to describe the dependence of PCs and obtain the joint distribution function of the product's RUL.

The rest of this paper is organized as follows. In Section II, the degradation model based on the Wiener process is established considering fault effects. RUL distribution is discussed in Section III. Section IV focuses on the method of inference for the unknown parameters. The RUL prediction of product with multiple PCs that are influenced by fault is given in Section V. The experiments and results analyses are given in Section VI to illustrate the effectiveness accuracy of the proposed method. Finally, some concluding remarks and discussions are made in Section VII.

II. DEGRADATION MODEL WITH FOM FOR SINGLE PC

As a mathematical expression of Brown motion (BM), Wiener process has been widely used in degradation modeling. Consider single fault occurs in the degradation process, which can be described by changing the drift parameter of Wiener process. In this section, we focus on the degradation process models governed by the Wiener process considering fault effects. In general, a Wiener process based degradation model can be represented by

$$X(t) = X(0) + \lambda t + \sigma B(t) \quad (1)$$

where λ is a drift coefficient, σ is a diffusion coefficient and $B(t)$ is a standard BM, $B(t) \sim (0, t)$, $X(0)$ represents the initial value of the degradation process. Without loss of generality, $X(0)$ is zero, so $X(t) \sim N(\lambda t, \sigma^2 t)$.

1) Continuous Degradation Model Considering Fault Effects

Equation (1) demonstrates a normal degradation process, but through the above analysis in Section I, the fault may be generated during the degradation process, so the product suffers additional load, and this state has a higher degradation rate than the case in the normal state. The FOM separates the whole degradation process into two states, the former is a gradual degradation state expressed and the latter one is a rapid degradation state. Therefore, the whole process can be described by changing drift function, the diffusion coefficient σ is assumed to be the same during the both states. The degradation process considering fault effects is then indicated as


$$X(t) = \begin{cases} \lambda_1 \times t + \sigma B(t), & t < \tau \\ \lambda_2 \times (t - \tau) + \lambda_1 \times \tau + \sigma B(t), & t \geq \tau \end{cases} \quad (2)$$

where τ is the FOM, λ_1 and λ_2 are the degradation rates for $t < \tau$ and $t \geq \tau$, respectively. $X(t)$ is normally distributed as $X(t) \sim N(\lambda_1 t, \sigma^2 t)$ for $t \in [0, \tau)$ and $X(t) \sim N(\lambda_2(t - \tau) + \lambda_1 \tau, \sigma^2 t)$ for $t \in [\tau, +\infty)$.

2) Increment Degradation Model Considering Fault Effects


As the measurement time is discrete, we suppose that m similar products are tested, and $X_{i,j}$ denotes the degradation observation value of performance characteristics for product i at time $t_{i,j}$, $\{i = 1, \dots, m; j = 1, 2, \dots, n_i\}$, where n_i is the number of inspection time of each product. $\Delta X_{i,j} = X_{i,j+1} - X_{i,j}$ from $\Delta t_{i,j} = t_{i,j+1} - t_{i,j}$. The increment degradation model considering fault effects can be divided into three cases.

Case 1: ($\tau \geq t_{i,j+1}$)




$$\Delta X_{i,j} = \lambda_1 \Delta t_{i,j} + \sigma B(\Delta t_{i,j}) \quad (3)$$

Case 2: ($t_{i,j} < \tau < t_{i,j+1}$)



$$\Delta X_{i,j} = \lambda_1 (\tau - t_{i,j}) + \lambda_2 (t_{i,j+1} - \tau) + \sigma B(\Delta t_{i,j}) \quad (4)$$

Case 3: ($\tau \leq t_{i,j}$)



$$\Delta X_{i,j} = \lambda_2 \Delta t_{i,j} + \sigma B(\Delta t_{i,j}) \quad (5)$$

where three cases respectively represent the FOM occurs after the increment, FOM occurs in the increment, FOM occurs before the increment.

According to specifications above, the PDF of the increment $\Delta X_{i,j}$ is expressed as

$$f(\Delta X_{i,j}(k)) = \begin{cases} \frac{1}{\sqrt{2\pi\Delta t_{i,j}\sigma}} \exp\left(-\frac{(\Delta X_{i,j} - \lambda_1 \Delta t_{i,j})^2}{2\sigma^2 \Delta t_{i,j}}\right), & k = 1 \\ \frac{1}{\sqrt{2\pi\Delta t_{i,j}\sigma}} \exp\left(-\frac{(\Delta X_{i,j} - \lambda_1 (\tau - t_{i,j}) - \lambda_2 (t_{i,j+1} - \tau))^2}{2\sigma^2 \Delta t_{i,j}}\right), & k = 2 \\ \frac{1}{\sqrt{2\pi\Delta t_{i,j}\sigma}} \exp\left(-\frac{(\Delta X_{i,j} - \lambda_2 \Delta t_{i,j})^2}{2\sigma^2 \Delta t_{i,j}}\right), & k = 3 \end{cases} \quad (6)$$

where $k = 1, 2, 3$ denote the scenarios $\tau > t_{i,j+1}$, $t_{i,j} < \tau < t_{i,j+1}$, and $\tau < t_{i,j}$, respectively.

Here, our attention is paid on not only the degradation model for single PC considering fault effects, but also RUL distribution and parameter estimation. Therefore, two important parts in the following sections are described in detail.

III. RUL DISTRIBUTION CONSIDERING FAULT EFFECTS

In general, the prediction for product's RUL often means to find its PDF or cumulative distribution function (CDF) of RUL distribution and calculate the expectation value finally.

Therefore, one of the main goals is to obtain RUL distribution for the product. Based on the concept of the first passage time, a product is regarded as to be failure, when its degradation value exceeds the preset failure threshold. Suppose w is the failure threshold, and $X_{i,j}$ has not exceeded the failure threshold w . The lifetime T under the concept of the first passage time is expressed as

$$T = \inf\{t : X(t) \geq w | X(0) < w\}. \quad (7)$$

Therefore, let $L_{i,j}$ denote the RUL at time $t_{i,j}$ as

$$L_{i,j} = \inf\{l : X(t_{i,j} + l) \geq w | X_{i,j} < w\} \quad (8)$$

where l is the time interval of a product that from normal state to failure state.

At first, consider the degradation process without the fault. According to properties of the Wiener process, the CDF of RUL in the normal state follows the inverse Gaussian distribution

$$\begin{aligned} F_{L_{i,j}}(t) &= P(L_{i,j} \leq t, Z_{i,j}) \\ &= \Phi\left(\frac{-w + \lambda t}{\sigma \sqrt{t}}\right) + \exp\left(\frac{\lambda t - Z_{i,j}}{\sigma^2 t} \times 2(w - Z_{i,j})\right) \\ &\quad \times \Phi\left(\frac{-w - \lambda t + 2Z_{i,j}}{\sigma \sqrt{t}}\right) \end{aligned} \quad (9)$$

where $\Phi(\cdot)$ is a standard normal distribution, $Z_{i,j}$ represents current observation value at observation time $t_{i,j}$.

The corresponding PDF is

$$f_{L_{i,j}}(t) = \frac{w - Z_{i,j}}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{-\frac{(w - Z_{i,j} - \lambda t)^2}{2\sigma^2 t}\right\}. \quad (10)$$

However, the degradation process in (2) is divided into two parts by FOM τ . The PDF of RUL distribution should also consist of two parts, and if the FOM is determined and its position is known in observations, the PDF can be expressed as

$$\begin{aligned} f_{L_{i,j}}(t) &= \\ &\begin{cases} \frac{w - Z_{i,j}}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{-\frac{(w - Z_{i,j} - \lambda_1 t)^2}{2\sigma^2 t}\right\}, & 0 < t < \tau \\ \frac{w - Z_{i,j}}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{-\frac{(w - Z_{i,j} - \lambda_2(t - \tau) - \lambda_1 \tau)^2}{2\sigma^2 t}\right\}, & t \geq \tau \end{cases} \end{aligned} \quad (11)$$

where $0 < t < \tau$ denotes the no fault occur in the product degradation process, $t \geq \tau$ denotes the fault will occur in the product degradation process.

Because the FOM is unobservable in the future, it is impossible to know its precise position, so form (11) is not applicable. To derive the RUL distribution, the FOM is regarded as a random variable which is assumed to follow a certain distribution, and the CDF and PDF of the FOM are $f(\tau; \theta_\tau)$, $F(\tau; \theta_\tau)$, respectively, where θ_τ is the set of unknown parameters of distribution function. The CDF of RUL distribution can be then expressed as

$$\begin{aligned} F_{L_{i,j}}(t) &= \int_t^{+\infty} \{F_{L_{i,j}}(t)(t|\lambda_1 t, \sigma^2, w)\} \times f_\tau(\tau, \theta_\tau) d\tau \\ &\quad + \int_0^t \{F_{L_{i,j}}(t)(t|\lambda_2(t - \tau) + \lambda_1 \tau, \sigma^2, w)\} \times f_\tau(\tau, \theta_\tau) d\tau \\ &= F_{L_{i,j}}(t)(t|\lambda_1 t, \sigma^2, w) \times (1 - F_\tau(t, \theta_\tau)) \\ &\quad + \int_0^t \{F_{L_{i,j}}(t)(t|\lambda_2(t - \tau) + \lambda_1 \tau, \sigma^2, w)\} \times f_\tau(\tau, \theta_\tau) d\tau. \end{aligned} \quad (12)$$

IV. PARAMETER ESTIMATION

According to above specification, the unknown parameters considering fault effects consist of $\theta_p = \{\lambda_1, \lambda_2, \sigma\}$ and θ_τ . In this section, we focus on the issue of parameter estimation under the assumption that the FOM is a random variable. FOM τ cannot be observed in the future span, this will cause the incomplete observations, and the FOM is regarded as a missing data in the observations. In a missing-data problem, EM algorithm is used to estimate the parameters θ_p , θ_τ . The EM algorithm consists of two steps performed iteratively. The expectation-step (E-step) computes the expectation of the log-likelihood with respect to the complete-data conditioned on the observed data. The maximization step (M-step) then finds the maximizer of this expected likelihood [31]. The two steps are repeated iteratively until satisfactory convergence is achieved.

For applying EM algorithm, observed data likelihood needs to be obtained firstly. Let $\{\tau_1, \tau_2, \dots, \tau_m\}$ be the FOM set corresponding to m products which are random variables during the degradation process of each product. Because the FOM is random in the degradation process, the joint PDF of increment for each product ΔX_i is then written in three cases:

1) Fault occurs before all observations ($\tau_i < t_{i,1}$),

$$A(\Delta X_i) = \prod_{j=1}^{n_i-1} f(\Delta X_{i,j} | (3)). \quad (13)$$

2) Fault occurs in observations ($t_{i,1} \leq \tau_i \leq t_{i,n_i}$),

$$\begin{aligned} B_j(\Delta X_i) &= \left\{ \prod_{j'=1}^{j-1} f(\Delta X_{i,j'} | (1)) \right\} \times f(\Delta X_{i,j} | (2)) \\ &\quad \times \left\{ \prod_{j'=j+1}^{n_i-1} f(\Delta X_{i,j'} | (3)) \right\}. \end{aligned} \quad (14)$$

3) Fault occurs after all observations ($t_{i,n_i} < \tau_i$),

$$C(\Delta X_i) = \prod_{j=1}^{n_i-1} f(\Delta X_{i,j} | (1)). \quad (15)$$

The joint PDF of ΔX_i conditioned on τ_i for i th product is then expressed as

$$\begin{aligned} f_{\Delta X_i}(\Delta X_i; \theta_p, \theta_\tau) &= \int_0^{+\infty} f(\Delta X_i; \theta_p) f_\tau(\tau_i; \theta_\tau) d\tau_i \\ &= \int_0^{t_{i,1}} A(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \\ &\quad + \int_{t_{i,1}}^{t_{i,n_i}} B_j(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \\ &\quad + \int_{t_{i,n_i}}^{+\infty} C(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \end{aligned}$$

$$\begin{aligned}
& + \sum_{j=1}^{n_i-1} \int_{t_{i,j}}^{t_{i,j+1}} B_j(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \\
& + \int_{t_{i,n_i}}^{+\infty} C(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i.
\end{aligned} \quad (16)$$

Observed data likelihood is then written as

$$\begin{aligned}
L_{obs} &= \prod_{i=1}^m \int_0^{+\infty} f(\Delta X_i; \theta_p) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \\
&= \prod_{i=1}^m \left\{ \int_0^{t_{i,1}} A(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \right. \\
&\quad + \sum_{j=1}^{n_i-1} \int_{t_{i,j}}^{t_{i,j+1}} B_j(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \\
&\quad \left. + \int_{t_{i,n_i}}^{+\infty} C(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau) d\tau_i \right\}. \quad (17)
\end{aligned}$$

The PDF of τ is given, and the complete data vector consists of $\{\Delta X, \Delta t, \tau\}$, where τ is a set of the FOMs $\{\tau_1, \tau_2, \dots, \tau_m\}$, the complete-data likelihood is

$$\begin{aligned}
L_c &= \prod_{i=1}^m \left[f(\tau_i; \theta_\tau) \right. \\
&\quad \times \prod_{j=1}^{n_i-1} \left(f(\Delta X_{i,j}|(1)) \times I\{\tau_i > t_{i,j+1}\} + f(\Delta X_{i,j}|(2)) \right. \\
&\quad \left. \times I\{t_{i,j} < \tau_i < t_{i,j+1}\} + f(\Delta X_{i,j}|(3)) \times I\{\tau_i < t_{i,j}\} \right) \left. \right] \\
&= \prod_{i=1}^m f_{\tau_i}(\tau_i, \theta_\tau) \\
&\quad \times \prod_{i=1}^m \prod_{j=1}^{n_i-1} \prod_{k=1}^3 f(\Delta X_{i,j}|(k); \lambda_1, \lambda_2, \tau_i, \sigma^2)^{\delta_{k,i,j}} \quad (18)
\end{aligned}$$

where $\delta_{k,i,j}$ is an indicator variable for τ and $\Delta t_{i,j}$ in three scenarios, such as when scenario $k=1, 2, 3$ occurs, it corresponds to the $\delta_{k,i,j} = 1$, $k \in 1, 2, 3$, $\delta_{k,i,j} = 0$ otherwise.

The complete-data log-likelihood is then written as

$$\begin{aligned}
\ln(L_c) &= \sum_{i=1}^m \ln(f_{\tau_i}(\tau_i, \theta_\tau)) \\
&+ \sum_{i=1}^m \sum_{j=1}^{n_i-1} \sum_{k=1}^3 \delta_{k,i,j} \times \ln f(\Delta X_{i,j}|(k); \lambda_1, \lambda_2, \tau_i, \sigma^2). \quad (19)
\end{aligned}$$

Use the following notation for simplifying the above log-likelihood expression

$$l_i^*(\theta_\tau) = \ln(f_{\tau_i}(\tau_i, \theta_\tau)) \quad (20)$$

$$l_{i,j,k}^{**}(\theta_p) = \ln f(\Delta X_{i,j}|(k); \lambda_1, \lambda_2, \tau_i, \sigma^2). \quad (21)$$

The complete-data log-likelihood can then be written as

$$L_c^* = \sum_{i=1}^m l_i^*(\theta_\tau) + \sum_{i=1}^m \sum_{j=1}^{n_i-1} \sum_{k=1}^3 \delta_{k,i,j} \times l_{i,j,k}^{**}(\theta_p). \quad (22)$$

In order to apply the EM algorithm to solve the maximum likelihood estimates of $(\lambda_1, \lambda_2, \sigma^2)$ for the case that the FOM is not observed, the problem is firstly cast as a missing-data problem, where τ is regarded as the missing data in observations.

1) *E-step*: The goal of the E-step is to compute the expectation of the complete-data log-likelihood of the missing-data conditioned on the observed data $\Delta X, \Delta t$.

$$\begin{aligned}
Q &= E(L_c^* | \Delta X, \theta_p^b, \theta_\tau^b) = Q_1 + Q_2 \\
&= \sum_{i=1}^m E\{l_i^*(\theta_\tau^b | \Delta X_i)\} \\
&\quad + \sum_{i=1}^m \sum_{j=1}^{n_i-1} \sum_{k=1}^3 E\{\delta_{k,i,j} \times l_{i,j,k}^{**}(\theta_p^b) | \Delta X_i\} \quad (23)
\end{aligned}$$

where θ_τ^b and θ_p^b , ($b = 0, 1, 2, \dots, n$) denote the currently iterated estimates, b is number of iterations.

The EM algorithm is only effective for the case that the complete-data likelihood is linearly separable in the set of missing-data statistics, and form (23) is not analytical-form. Therefore, the missing data τ from the function needs to be separated, log-likelihood expression can be written in the form of $L_c^* = V^T \cdot m^*$, where the vector V consists of only functions of the missing data τ , and m^* consists of only θ_p , θ_τ and observe data. In the forms, the components of the log-likelihood expressions are $l_i^* = V_i^{*T} \cdot m_i^*$, and $l_{i,j,k}^{**} = V_{i,j,k}^{**T} \cdot m_{i,j,k}^{**}$, where

$$V_{i,j}^{**T} = [\delta_{1,i,j} \times V_{i,j,k=1}^{**}, \delta_{2,i,j} \times V_{i,j,k=2}^{**}, \delta_{3,i,j} \times V_{i,j,k=3}^{**}]^T \quad (24)$$

$$m_{i,j}^{**} = [m_{i,j,k=1}^{**}, m_{i,j,k=2}^{**}, m_{i,j,k=3}^{**}]. \quad (25)$$

So, m^* consists of m_i^* and $m_{i,j}^{**}$, V consists of l_i^* and $l_{i,j}^{**}$. When linear separability can be established, it is then necessary to compute the expectations of V conditioned on observed data. Form Q_1 and Q_2 given in (23) can be expressed as respectively

$$Q_1 = \sum_{i=1}^m E(V_i^{*T} | \Delta X_i, \theta_p^b, \theta_\tau^b) \times m_i^* \quad (26)$$

$$Q_2 = \sum_{i=1}^m \sum_{j=1}^{n_i-1} E(V_{i,j}^{**T} | \Delta X_i, \theta_p^b, \theta_\tau^b) \times m_{i,j}^{**}. \quad (27)$$

The conditional expectation for V_i^* can be evaluated as follows:

$$\begin{aligned}
E(V_i^* | \Delta X_i, \theta_p^b, \theta_\tau^b) &= \frac{1}{f_{\Delta X_i}(\Delta X_i; \theta_p^b, \theta_\tau^b)} \int_0^{+\infty} V_i^*(\tau_i) \times f_{\Delta X_i}(\Delta X_i; \tau_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i
\end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{f_{\Delta X_i}(\Delta X_i; \theta_\tau^b, \theta_p^b)} \times \left\{ \int_0^{t_{i,1}} V_i^*(\tau_i) \times A(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i \right. \\
 &+ \sum_{j=1}^{n_i-1} \int_{t_{i,j}}^{t_{i,j+1}} V_i^*(\tau_i) \times B_j(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i \\
 &\left. + \int_{t_{i,n_i}}^{+\infty} V_i^*(\tau_i) \times C(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i \right\}. \quad (28)
 \end{aligned}$$

Next, the conditional expectation under the second summation series given in (23) is computed. However, it involves conditional expectation of $\delta_{k,i,j} \cdot l_{i,j,k}^{**}$, which needs to be discussed in three cases:

a) $\tau \geq t_{i,j+1} = (t_{i,j+1} \leq \tau \leq t_{i,n_i}) + (\tau > t_{i,n_i})$

$$\begin{aligned}
 &E(\delta_{1,i,j} \times V_{i,j,1}^{**} | \Delta X_i, \theta_p^b, \theta_\tau^b) \\
 &= \frac{1}{f_{\Delta X_i}(\Delta X_i; \theta_\tau^b, \theta_p^b)} \left[\sum_{j'=j+1}^{n_i-1} \int_{t_{i,j'}}^{t_{i,j'+1}} V_{i,j,1}^{**}(\tau_i) \right. \\
 &\times B_{j'}(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i + \int_{t_{i,n_i}}^{+\infty} V_{i,j,1}^{**}(\tau_i) \\
 &\times C(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i \left. \right]. \quad (29)
 \end{aligned}$$

b) $t_{i,j} < \tau < t_{i,j+1}$

$$\begin{aligned}
 &E(\delta_{2,i,j} \times V_{i,j,2}^{**} | \Delta X_i, \theta_p^b, \theta_\tau^b) \\
 &= \frac{1}{f_{\Delta X_i}(\Delta X_i; \theta_\tau^b, \theta_p^b)} \int_{t_{i,j}}^{t_{i,j+1}} V_{i,j,2}^{**}(\tau_i) \\
 &\times B(\Delta X_{i,j}) f_\tau(\tau_i; \theta_\tau^b) d\tau_i. \quad (30)
 \end{aligned}$$

c) $\tau \leq t_{i,j} = (\tau < t_{i,1}) + (t_{i,1} \leq \tau \leq t_{i,j})$

$$\begin{aligned}
 &E(\delta_{3,i,j} \times V_{i,j,3}^{**} | \Delta X_i, \theta_p^b, \theta_\tau^b) \\
 &= \frac{1}{f_{\Delta X_i}(\Delta X_i; \theta_\tau^b, \theta_p^b)} \left[\int_0^{t_{i,1}} V_{i,j,3}^{**}(\tau_i) \times A(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i \right. \\
 &+ \sum_{j'=1}^{j-1} \int_{t_{i,j'}}^{t_{i,j'+1}} V_{i,j,3}^{**}(\tau_i) \times B_{j'}(\Delta X_i) \times f_\tau(\tau_i; \theta_\tau^b) d\tau_i \left. \right]. \quad (31)
 \end{aligned}$$

2) *M-step*: Based on E-step, update $b = b + 1$, new estimates for θ^b can be generated by solving for the complete-data MLE to (23). As Q_1 only contains θ_τ and Q_2 only contains θ_p , the estimator vectors θ_τ^b and θ_p^b can be calculated by (32), (33), respectively.

One iteration is completed through M steps, repeat the above two steps so that when the parameters tend to converge or the number of iterations is reached, the iteration will be

stopped.

$$\frac{\partial Q_1}{\partial \theta_\tau^b} = \frac{\partial \left[\sum_{i=1}^m E \left\{ V_i^* | \Delta X \right\}^T \times m_i^* (\theta_\tau^b) \right]}{\partial \theta_\tau^b} = 0 \quad (32)$$

$$\frac{\partial Q_2}{\partial \theta_p^b} = \frac{\partial \left[\sum_{i=1}^m \sum_{j=1}^{n_i-1} \sum_{k=1}^3 E \{ \delta_{k,i,j} \times V_{i,j,k}^{**} | \Delta X \}^T m_{i,j,k}^{**} (\theta_p^b) \right]}{\partial \theta_p^b} = 0. \quad (33)$$

Algorithm 1 The Computational Procedure for the EM Algorithm

- 1: Begin with an initial set of estimates for $\theta_p^0, \theta_\tau^0$, initialize the iteration $b = 0$.
 - 2: Obtain Q in (23), the complete data likelihood conditioned on ΔX by using (28)–(31).
 - 3: Update $b = b + 1$. Compute the MLE $\theta_p^b, \theta_\tau^b$ by using (32) and (33).
 - 4: If satisfactory convergence of unknown parameter or expectation value is achieved, the procedure terminates. Otherwise go to Step 2.
-

V. RUL PREDICTION OF PRODUCT WITH MULTIPLE PCs

The RUL prediction of the product with single PC considering fault effects can be calculated by Sections II-IV. As the occurrence of the fault can change the degradation trajectories of multiple PCs, the final goal is to obtain RUL prediction of product with multiple PCs. Suppose that a product has n PCs, and these PCs are observed at the same measurement time. The RUL distribution $F_{i,j}^{(c)}(t)$, $\{c = 1, \dots, n\}$ for c th degradation feature at $t_{i,j}$ can be calculated by Sections II-IV considering fault effects. At first, when multiple PCs of the product are independent, the RUL of the target product at time $t_{i,j}$ can be expressed as $L_{i,j} = \min\{L_{i,j}^{(1)}, \dots, L_{i,j}^{(n)}\}$. Therefore, the joint CDF of product's RUL can be calculated by

$$\begin{aligned}
 F_{i,j}(t) &= P(L_{i,j} \leq t) = P(\min(L_{i,j}^{(1)}, \dots, L_{i,j}^{(n)}) \leq t) \\
 &= 1 - P(\min(L_{i,j}^{(1)}, \dots, L_{i,j}^{(n)}) > t) \\
 &= 1 - P(L_{i,j}^{(1)} > t, \dots, L_{i,j}^{(n)} > t) \\
 &= 1 - \prod_{c=1}^n P(L_{i,j}^{(c)} > t) \\
 &= 1 - \prod_{c=1}^n (1 - P(L_{i,j}^{(c)} \leq t)) \\
 &= 1 - \prod_{c=1}^n (1 - F_{i,j}^{(c)}(t)). \quad (34)
 \end{aligned}$$

Then, the joint PDF of product's RUL $f_{i,j}(t) = F'_{i,j}(t)$. It is easy to get the RUL distribution, if PCs are independent during the degradation process. However, as one of our main issues in this paper, it is not appropriate to ignore the dependence among multiple performance characteristics. To capture the dependence among multiple degradation features, the Copula function is employed to obtain the product's joint RUL distribution. The Copula function has been widely

adopted to modeling the dependency between PCs, due to its flexibility to link the relationship between the univariate marginal distributions and the joint distribution. According to Sklar's theory [34], suppose that $F(y_1), F(y_2), \dots, F(y_n)$ are CDFs of random variables y_1, y_2, \dots, y_n , respectively, a joint distribution function $H(y_1, y_2, \dots, y_n)$ for y_1, y_2, \dots, y_n is constructed through a multivariate Copula function as

$$H(y_1, y_2, \dots, y_n) = C(F(y_1), F(y_2), \dots, F(y_n); \alpha) \quad (35)$$

where $C(\cdot)$ is the Copula function, α denotes Copula parameter, it can be calculated by Kendall's tau shown in Table I.

TABLE I
FIVE COMMON COPULA FUNCTIONS

Copulas	$C(u_1, u_2, \dots, u_n; \alpha)$	Relationship
Gaussian	$\Phi(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$	$\text{tau} = \frac{2}{\pi} \arcsin(\alpha)$
t	$t(t^{-1}(u_1), \dots, t^{-1}(u_n))$	$\text{tau} = \frac{2}{\pi} \arcsin(\alpha)$
Clayton	$\left(\sum_{i=1}^n u_i^{-\alpha} - n + 1 \right)^{-1/\alpha}$	$\text{tau} = \frac{\alpha}{2+\alpha}$
Gumbel	$\exp \left\{ - \left[\sum_{i=1}^n (-\ln u_i)^\alpha \right]^{1/\alpha} \right\}$	$\text{tau} = 1 - \frac{1}{\alpha}$
Frank	$-\frac{1}{\alpha} \ln \left(1 + \frac{\prod_{i=1}^n (e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{n-1}} \right)$	$\text{tau} = 1 + 4 \frac{D_1(\alpha) - 1}{\alpha}$

where $\Phi^{-1}(\cdot)$ is the inverse of standard normal distribution, $t(\cdot)$ denotes student t distribution, $t^{-1}(\cdot)$ is the inverse of student t distribution.

The dependence of multiple variables is characterized by (35). By using different Copula functions, different types of multivariate distribution functions can be constructed. Popular multivariate Copula functions include: the Gaussian Copula, the Frank Copula, the Gumbel Copula, and the Clayton Copula. The corresponding five forms are shown in Table I, and more detail about Copula function can refer to [34].

Based on the RUL distribution for each PC has been predicted, and according to the Copula theory, the joint CDF of RUL distributions can be calculated as

$$F_{t_{i,j}}(t) = C(F_{t_{i,j}}^{(1)}(t), F_{t_{i,j}}^{(2)}(t), \dots, F_{t_{i,j}}^{(n)}(t); \alpha). \quad (36)$$

Then, the joint PDF of the product's RUL distribution can be calculated as

$$f_{t_{i,j}}(t) = c(F_{t_{i,j}}^{(1)}(t), F_{t_{i,j}}^{(2)}(t), \dots, F_{t_{i,j}}^{(n)}(t); \alpha) \prod_{c=1}^n f^{(c)}(t) \quad (37)$$

where $c(\cdot)$ is a derivative function.

The Akaike information criterion (AIC) is used to select the most suitable Copula function to link the marginal distribution function, and the smaller value of AIC denotes better, AIC is defined as

$$AIC = 2q - 2\ln(L) \quad (38)$$

where q denotes the number of unknown parameters in the model, $\ln(L)$ is the maximum value of the log-likelihood function.

VI. EXPERIMENTS AND RESULTS ANALYSES

In this section, an open source degradation dataset of

turbofan engines denoted by "FD001" (please refer to [35]) is used to verify the proposed approach. The dataset includes 26 columns of data, the 1st to the 3rd columns represent the engine's serial number, the number of engine flight cycles, the engine's settings under different operating conditions, respectively. Then, the 6th column to the 26th columns represent 21 sensors, each sensor corresponds to a PC of the engine, and these sensors have collected sufficient degradation data from the engine. The parameters measured by sensors include total temperature at fan inlet, total temperature at low-pressure compressor (LPC) outlet, total temperature at high-pressure compressor (HPC) outlet, total temperature at low-pressure turbine (LPT) outlet, pressure at fan inlet, total pressure in bypass-duct, total pressure at HPC outlet, physical fan speed, etc. Some measurements are constant value, such as the 6th, the 10th, and the 21st. Some measurements vary over time, such as the 12th, the 14th, and the 19th. The dataset includes one training set and one testing set of 100 engines respectively. The engine operates normally at the start of each time series, and develops a fault at some point during the series. Each unit in the training set runs from initial degradation until the system fails, there are 20631 sets of monitoring data in the training set. The longest life span in the training set is 378 cycles, and the shortest life span is 128 cycles. In the test set, the time series is before the system fails, and there are 13096 sets of monitoring data. In the training dataset, the engine operates normally at the start of each time series, and develops a fault at a moment during the series, when the fault occurs, PCs begin to degrade rapidly until failure.

At first, the monotonicity of each performance characteristic is evaluated by the Spearman coefficient, which is expressed as

$$\rho^i = \frac{|\sum_{k=1}^K (T_k - T_{1:k})(X_k^i - X_{1:k}^i)|}{\sqrt{\sum_{k=1}^K (T_k - T_{1:k})^2 \sum_{k=1}^K (X_k^i - X_{1:k}^i)^2}} \quad (39)$$

where $T_{1:k}$ and $X_{1:k}^i$ are the means of $\{T_1, T_2 \dots T_k\}$ and $\{X_1^i, X_2^i, \dots, X_k^i\}$ of the i th feature.

When the degradation value of a PC monotonously increases or decreases over time, the trendability value is not equal to zero. In contrast, if the degradation value of a PC is constant or varies randomly over time, the value is zero. So, a higher trendability means a better monotonicity. The trendability values of 21 sensors are calculated by Spearman coefficients and the results are shown in Table II. The values of Sensor 9 and Sensor 14 are bigger than others, so they are selected as two key PCs. In the training dataset, some units do not have an increasing trend, 65 units that have an increasing trend used to verify the proposed method are selected as training data. For simplicity, Sensor 9 is named as PC1, Sensor 14 is named as PC2.

The dataset contained unknown noises to a great extent. Because the noise sources are complicated and indeed not available, it is difficult to directly use the original observed data. Therefore, moving average filtering is applied to eliminate the high frequency fluctuations from the degradation paths. The denoising results of two degradation features in

TABLE II
THE TRENDABILITY VALUES OF 21 SENSORS

Sensors (n)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Trendability values (ρ)	0	0.63	0.62	0.84	0	0.93	0.84	0.90	0.93	0	0.88	0.87	0.90	0.94	0.76	0	0.64	0	0	0.77	0.73

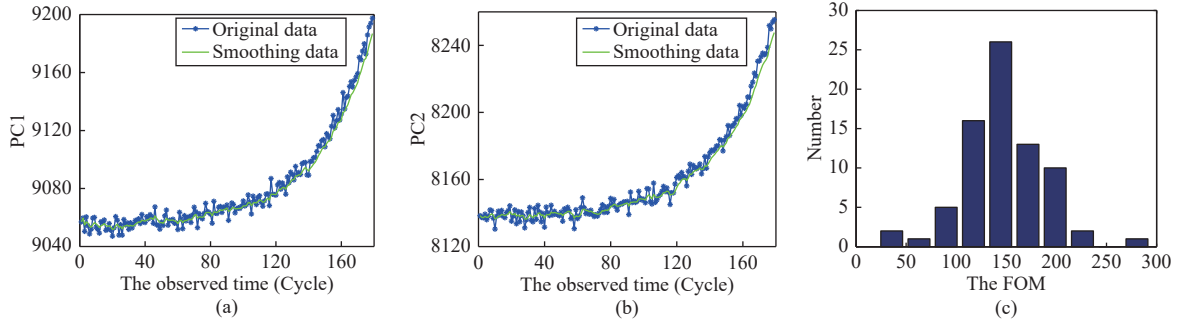


Fig. 2. The denoising results of two degradation features in Engine 3 for (a) PC1 and (b) PC2; (c) Histogram of FOM for 100 engines.

Engine 3 are shown in Figs. 2(a) and 2(b), respectively.

A. The Experiment for Single PC Considering Fault Effects

The RUL predictions of product with PC1 and PC2 considering fault effects are discussed separately. For a comparison purpose, traditional Wiener process that considers no fault effects is used to this dataset. As FOM is a random variable and assumed to follow a certain distribution, the distribution τ followed that should be determined under the training data. The method of combination of cumulative sum charts (the detailed principles can refer to [36]) is used to roughly discriminate the position of the fault occurrence for 100 engines, and the bar graph is shown in Fig. 2(c). Through the distribution histogram for 100 engines, the FOM τ can be assumed to follow a normal distribution, $\tau \sim N(u_\tau, \sigma_\tau^2)$. The PDF is then written as

$$f(\tau_i; \theta_\tau) = \frac{1}{\sqrt{2\pi}\sigma_\tau} \exp\left(-\frac{(\tau_i - u_\tau)^2}{2\sigma_\tau^2}\right). \quad (40)$$

Then, $l_i^*(\theta_\tau)$ and $l_{i,j,k}^{**}(\theta_p)$ in (20) and (21) can be written as

$$l_i^*(\theta_\tau) = \left(-\ln \sqrt{2\pi}\sigma_\tau - \frac{\sigma_\tau^2}{2\sigma_\tau^2}\right) + \frac{u_\tau \times \tau_i}{\sigma_\tau^2} - \frac{\tau_i^2}{2\sigma_\tau^2} \quad (41)$$

$$l_{i,j,k}^{**}(\theta_p) = \begin{cases} a_{i,j}, & k=1 \\ b_{i,j} + c_{i,j} \times \tau_i + d_{i,j} \times \tau_i^2, & k=2 \\ e_{i,j}, & k=3 \end{cases} \quad (42)$$

$$a_{i,j} = -\ln \sqrt{2\pi\Delta t_{i,j}\sigma^2} - \frac{(\Delta X_{i,j} - \lambda_1 \Delta t_{i,j})^2}{2\sigma^2 \Delta t_{i,j}}$$

$$b_{i,j} = -\ln \sqrt{2\pi\Delta t_{i,j}\sigma^2} - \frac{(\Delta X_{i,j} - \lambda_2 \Delta t_{i,j+1} + \lambda_1 \Delta t_{i,j})^2}{2\sigma^2 \Delta t_{i,j}}$$

$$c_{i,j} = \frac{2(\Delta X_{i,j} - \lambda_2 \Delta t_{i,j+1} + \lambda_1 \Delta t_{i,j})(\lambda_1 - \lambda_2)}{2\sigma^2 \Delta t_{i,j}}$$

$$d_{i,j} = \frac{-(\lambda_1 - \lambda_2)^2}{2\sigma^2 \Delta t_{i,j}}$$

$$e_{i,j} = -\ln \sqrt{2\pi\Delta t_{i,j}\sigma^2} - \frac{(\Delta X_{i,j} - \lambda_2 \Delta t_{i,j})^2}{2\sigma^2 \Delta t_{i,j}}.$$

Because the EM algorithm is effective only for the case that the complete-data likelihood is linearly separable in the set of missing-data statistics, the conditional expectation needs to be linearly separated, so $l_i^*(\theta_\tau)$ and $l_{i,j}^{**}(\theta_p, \theta_\tau)$ are expressed as

$$\begin{aligned} l_i^*(\theta_\tau) &= V_i^T \times m_i^* \\ &= [1, \tau_i, \tau_i^2]^T \times \left[-\ln \sqrt{2\pi}\sigma_\tau - \frac{u_\tau^2}{2\sigma_\tau^2}, \frac{u_\tau}{\sigma_\tau^2}, -\frac{1}{2\sigma_\tau^2} \right] \\ l_{i,j}^{**}(\theta_p, \theta_\tau) &= [V_{i,j,1}^{**}, V_{i,j,2}^{**}, V_{i,j,3}^{**}]^T \times [m_{i,j,1}^{**}, m_{i,j,2}^{**}, m_{i,j,3}^{**}] \\ &= [1, 1, \tau_i, \tau_i^2, 1][a_{i,j}, b_{i,j}, c_{i,j}, d_{i,j}, e_{i,j}]^T. \end{aligned}$$

Therefore, when we assume that fault point τ_i is normally distributed, the complete-data log-likelihood function is separable. Then, the conditional expectation can be evaluated by (28)–(31), for Q_1 and Q_2 , the maximum likelihood estimators for $\theta_p = \{u_\tau, \sigma_\tau\}$, $\theta_p = \{\lambda_1, \lambda_2, \sigma\}$ can be updated by (32), (33). Finally, the unknown parameters can be obtained, when the iteration is terminated.

The dataset is given min-max normalized treatment, and the failure threshold of each engine is the value at the end of trajectory, so the failure threshold of each unit is 1. Then, the unknown parameters of the degradation model for each PCs can be estimated by EM algorithm, and the distribution parameters of the FOM are also estimated $u_\tau = 143.1$, $\sigma_\tau^2 = 58.2$, the parameters of degradation model for two PCs are presented in Table III. Finally, each PC's RUL distribution considering fault effects can be calculated by (12). For comparative purposes, the Wiener process based conventional

TABLE III
PARAMETER SPECIFICATION

Situation	λ_1	λ_2	σ^2
Proposed model for PC1	0.0014	0.0141	1.1144E-04
Proposed model for PC2	0.0016	0.0136	1.1173E-04
Conventional method	0.0055	0.0055	1.1234E-04

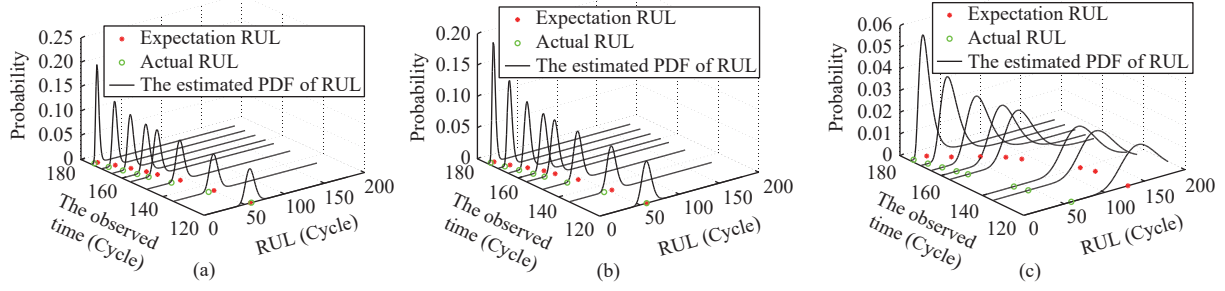


Fig. 3. The prediction results of proposed model: (a) PC1; (b) PC2; (c) Conventional model.

method without considering fault effects is used to predict engine's RUL, where the unknown parameters are estimated by maximum likelihood estimation (shown in Table III), and the existing machine learning algorithms: convolutional neural network (CNN) and multilayer perceptron (MLP) are employed to compare the proposed methods in this paper [37].

The prediction results of three methods are shown in Figs. 3(a)–3(c), and the comparison results of three models at different observed time are shown in Table IV. The mean absolute error (MAE) and root mean square error (RMSE) average error percentage are used to evaluate the accuracy of the prediction results. When the values of MAE and RMSE are smaller, the accuracy of prediction is higher.

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i^* - P_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i^* - P_i)^2}$$

where P_i^* is the actual life and P_i is predicted life at the i th prediction moment.

TABLE IV
COMPARISON OF RUL PREDICTION

Actual value (cycles)	52	47	42	37	32
Conventional method	126.63	120.12	115.62	113.76	100.54
Proposed model for PC1	55.03	52.06	50.06	48.34	42.38
Proposed model for PC2	57.01	54.10	53.42	47.91	45.54
Independent	54.01	50.50	48.03	45.12	40.73

Table V shows the accuracy analysis from the prediction results of conventional method and proposed method for single PC, the maximum error is 76.76 cycles for conventional method, proposed method for PC1 and PC2 is 11.34 cycles and 13.91 cycles, respectively. Moreover, compared to conventional method, MAE and RMSE of the proposed method for single PC are reduced significantly, which demonstrates that the proposed model will be quite useful for the observations with the fault effects. However, compared to CNN and MLP, MAE and RMSE of the proposed method for single PC are higher, it shows that the accuracy of prediction for proposed method for single PC is not the highest.

TABLE V
ACCURACY ANALYSIS OF THE PREDICTION RESULTS

Method	Maximum error (cycles)	MAE	RMSE
Conventional method	76.76	73.33	73.38
Proposed model for PC1	11.34	7.57	8.20
Proposed model for PC2	13.91	9.60	10.08
Independent	8.73	5.68	6.24
CNN	7.72	4.12	5.42
MLP	9.46	5.46	6.81

B. The Experiment for Multiple PCs Considering Fault Effects

As above analysis, multiple PCs' degradation trajectories will be changed by the FOM, and PCs may be independent or dependent during the degradation process. When the two PCs are considered being independent firstly, it can be calculated by (34), the prediction results are shown in Fig. 4, and the expectation value is calculated in Table IV. Meanwhile, as shown in Table V, MAE and RMSE of the prediction method with multi-PC independent considering fault effects have reduced to a certain extent, which shows that the prediction accuracy of remaining life considering multiple parameters is higher.

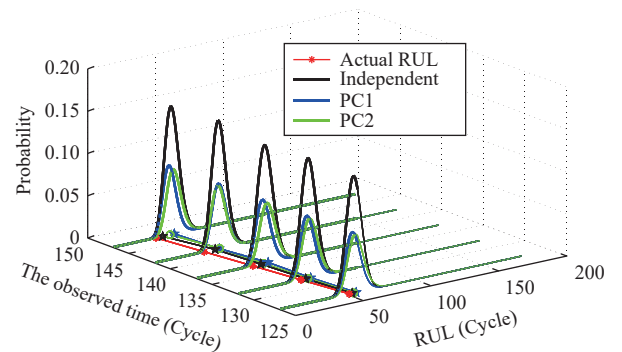


Fig. 4. The prediction results of proposed model of PC1, PC2, Independent.

However, PCs are usually dependent during the degradation process. The Copula function is employed to describe their dependence and obtain the joint PDF of RUL distribution. As shown in Table II, there are five commonly used Copula functions. The AIC of different Copula functions is calculated by (38), and the AIC of Clayton function is smaller among these five Copula functions. Therefore, Clayton function is selected to link the marginal distribution function, where the

TABLE VI
ACCURACY ANALYSIS OF THE PREDICTION RESULTS

Engine 3	Actual value (cycles)	52	47	42	37	32	MAE	RMSE
	Conventional method	49.52	44.48	44.31	38.26	30.74		
	The observed time	127	132	137	142	147	1.97	2.05
	α	0.4821	0.4834	0.4163	0.4426	0.4698		
Engine 4	Actual value (cycles)	49	44	39	34	29	MAE	RMSE
	Predicted RUL	48.25	44.97	41.71	37.48	32.61		
	The observed time	140	145	150	155	160	2.30	2.61
	α	0.3435	0.3775	0.4118	0.4457	0.4852		
Engine 11	Actual value (cycles)	60	55	50	45	40	MAE	RMSE
	Predicted RUL	59.15	55.72	52.04	48.58	41.91		
	The observed time	180	185	190	195	200	1.82	2.09
	α	0.3610	0.3497	0.3160	0.3165	0.2546		
Engine 17	Actual value (cycles)	210	215	220	225	230	MAE	RMSE
	Conventional method	67.14	62.35	58.30	52.82	48.64		
	The observed time	66	61	56	51	46	1.85	1.93
	α	0.3574	0.3510	0.3564	0.3575	0.3592		

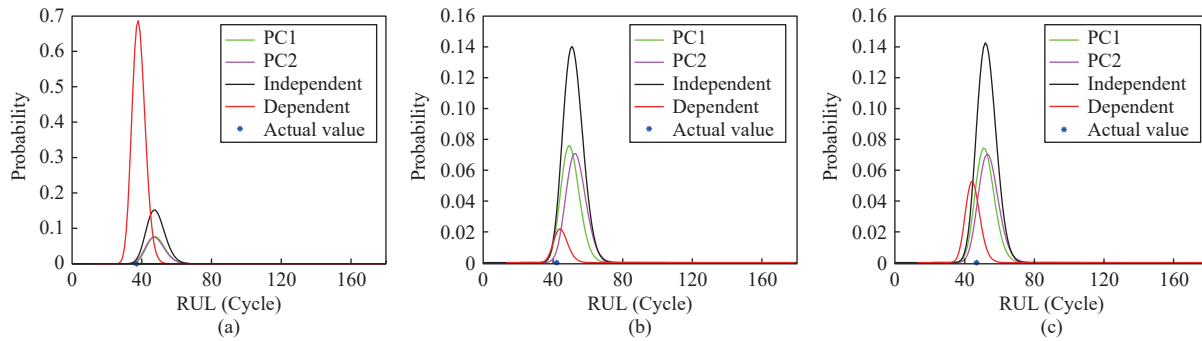


Fig. 5. Three prediction results at different observed time: (a) The observed time 142; (b) The observed time 137; (c) The observed time 1132.

corresponding Copula parameter α can be calculated by Kendall's tau. Therefore, the product's joint PDF of RUL of the 3rd engine can be calculated by (37), and the corresponding expectation value is calculated in Table VI. In order to compare and analyze the prediction results better, the RUL distributions of the three prediction times of the proposed method under three different situations (only one PC, independent, dependent) are selected to be depicted on the two-dimensional plane, shown in Figs. 5(a)–5(c).

From Table V and Table VI, compared with CNN and MLP, it is clearly to see that the MAE and RMSE of the remaining life prediction method under the multi-parameter dependent considering fault effects are minimum, it proves that the prediction accuracy of this method is the highest, and the stability of the prediction is also the best. In other words, compared to considering fault effects with single-PC or multi-PC independent, considering fault effects with multi-PC is more reasonable. In addition, in order to reflect the generality, the RUL prediction method under multi-parameter dependent considering fault effects is applied to Engine 4, Engine 11 and Engine 17, respectively, and predicted RUL distribution curves are shown in Figs. 6(b)–6(d), the predicted results are shown in Table VI. It can be obtained that the MAE and

RMSE of each engine RUL prediction results are very small, and they are about same as Engine 3, which proves the effectiveness of the method.

VII. CONCLUSIONS AND DISCUSSIONS

This paper concerns a problem of predicting RUL considering fault effects. Wiener process is employed to construct the degradation model of single PC with the fault effects. The RUL distribution is discussed based on the definition of first passage time, where the FOM is considered as a random variable. EM algorithm for the problem of parameter estimation is presented. Copula function is employed to describe the dependence between PCs and obtain the joint RUL distribution. The effectiveness of the proposed approach is demonstrated with experimental results.

The current study only considers single fault occurring in the product degradation process, but there may be multiple faults occurring in the degradation process. Therefore, the proposed approach may be extended to accurately predict product RUL where multiple faults may occur in the degradation process. Moreover, the degradation process with the fault effects of this paper is established by the linear Wiener process. In practice, the product degradation process

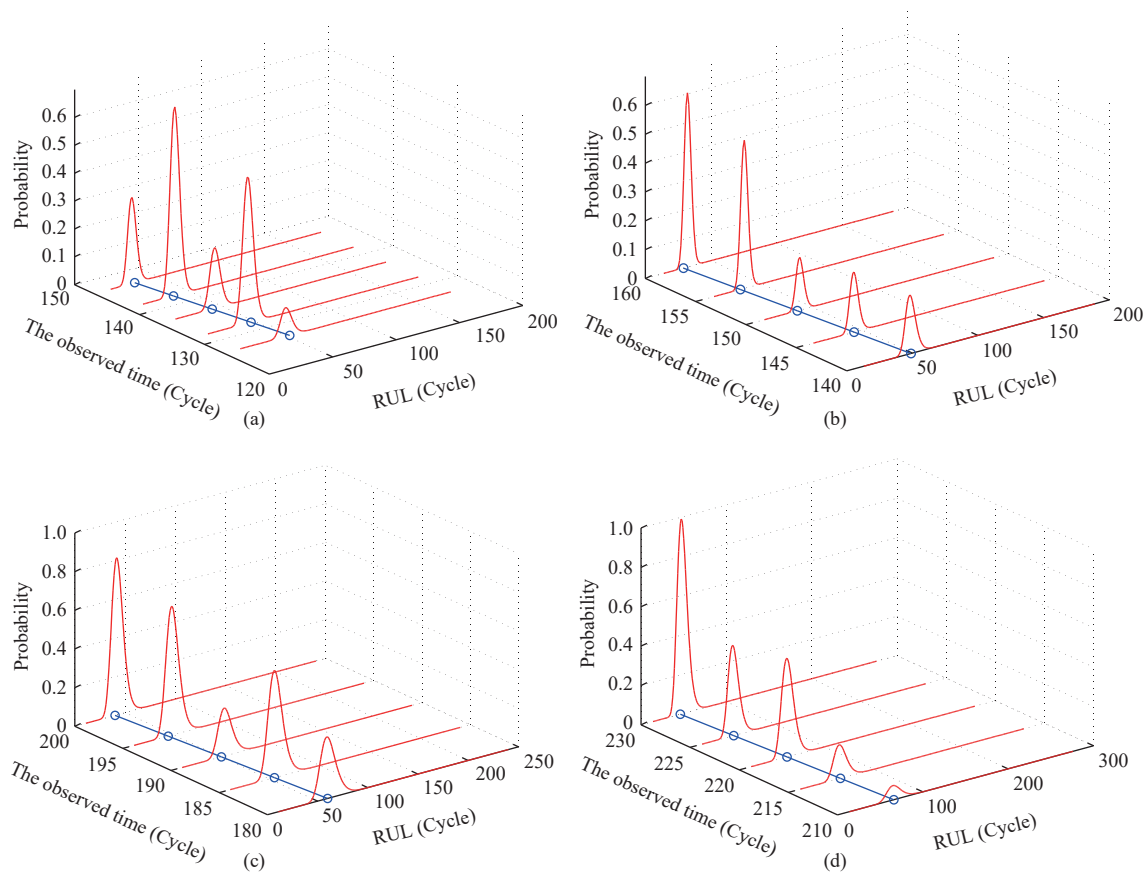


Fig. 6. The prediction results of the proposed method is applied to four engines: (a) Engine 3; (b) Engine 4; (c) Engine 11; (d) Engine 17.

often exhibits nonlinear characteristics, which limits the application of the linear model, degradation model with the fault effects based on the nonlinear Wiener process is worthy of further exploration.

REFERENCES

- [1] M. Pecht and R. Jaai, "A prognostics and health management roadmap for information and electronics-rich systems," *Microelectron. Rel.*, vol. 50, no. 3, pp. 317–323, 2010.
- [2] E. Principi, D. Rossetti, S. Squartini, and F. Piazza, "Unsupervised electric motor fault detection by using deep autoencoders," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 2, pp. 441–451, Mar. 2019.
- [3] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [4] R. Li, Y. Huang, and J. Wang, "Long-term traffic volume prediction based on K-means Gaussian interval type-2 fuzzy sets," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 6, pp. 1344–1351, Nov. 2019.
- [5] Z. Chen and S. Zheng, "Lifetime distribution based degradation analysis," *IEEE Trans. Rel.*, vol. 54, no. 1, pp. 3–10, Mar. 2005.
- [6] N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, "Residual-life distributions from component degradation signals: A Bayesian approach," *IIE Trans.*, vol. 37, no. 6, pp. 543–557, 2005.
- [7] S. Gao, M. Zhou, Y. Wang, J. Chen, and H. Yachi, "Dendritic neuron model with effective learning algorithms for classification, approximation and prediction," *IEEE Trans. Neural Networks and Learning Systems*, vol. 30, no. 2, pp. 601–614, Feb. 2019.
- [8] T. D. Pham, K. Wardell, A. Eklund, and G. Salerud, "Classification of short time series in early Parkinson's disease with deep learning of fuzzy recurrence plots," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 6, pp. 1306–1317, Nov. 2019.
- [9] F. Ahmadzadeh and J. Lundberg, "Remaining useful life estimation: review," *Int. J. Syst. Assur. Eng. Manage.*, vol. 5, no. 4, pp. 461–474, 2013.
- [10] N. Z. Gebraeel and M. A. Lawley, "A neural network degradation model for computing and updating residual life distributions," *IEEE Trans. Auto. Sci. Eng.*, vol. 5, no. 1, pp. 154–163, 2008.
- [11] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, and L. Zhang, "Remaining useful life prediction for lithium-ion battery: A deep learning approach," *IEEE Access*, vol. 6, pp. 50587–50598, 2018.
- [12] J. Wu, C. Deng, X. Shao, and S. Q. Xie, "A reliability assessment method based on support vector machines for CNC equipment," *Sci. Chi. Ser. E: Tech. Sci.*, vol. 52, no. 7, pp. 1849–1857, 2009.
- [13] V. Tran, H. Pham, B. Yang, and T. Nguyen, "Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine," *Mech. Sys. and Sig. Pro.*, vol. 32, pp. 320–330, 2012.
- [14] X. Si, W. Wang, C. Hu, and D. Zhou, "Remaining useful life estimation – A review on the statistical data driven approaches," *Eur. J. Oper. Res.*, vol. 213, no. 1, pp. 1–14, 2011.
- [15] S. T. Tseng, J. Tang, and I. H. Ku, "Determination of burn-in parameters and residual life for highly reliable products," *Nav. Res. Logist.*, vol. 50, no. 1, pp. 1–14, 2003.
- [16] Q. Zhai and Z. Ye, "RUL prediction of deteriorating products using an adaptive Wiener process model," *IEEE Trans. Ind. Infor.*, vol. 13, no. 6, pp. 2911–2921, Dec. 2017.
- [17] Z. Ye and M. Xie, "Stochastic modeling and analysis of degradation for highly reliable products," *Appl. Stoch. Mod. Bus. Ind.*, vol. 31, no. 8, pp. 16–32, 2015.
- [18] S. Lu, H. Lu, and W. J. Kolarik, "Multivariate performance reliability prediction in real-time," *Rel. Eng. Syst. Saf.*, vol. 72, no. 1, pp. 39–45, 2001.
- [19] P. Wang and D. W. Coit, "Reliability prediction based on degradation modeling for systems with multiple degradation measures," in *Proc. Annual Sym. Rel. Maintain.*, Los Angeles, USA, 2004, pp. 302–307.
- [20] F. Sun, L. Liu, X. Li, and H. Liao, "Stochastic modeling and analysis of multiple nonlinear accelerated degradation processes through

information fusion,” *Sensors*, vol. 16, no. 8, Article No. 1242, 2016.

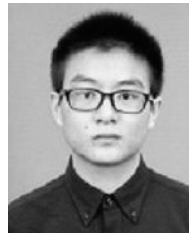
- [21] Z. Pan and N. Balakrishnan, “Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes,” *Rel. Eng. Syst. Saf.*, vol. 96, no. 8, pp. 949–957, 2011.
- [22] X. Wang, B. Guo, and J. Cheng, “Residual life estimation based on bivariate Wiener degradation process with time-scale transformations,” *J. Stat. Comput. Simul.*, vol. 84, no. 3, pp. 545–563, 2014.
- [23] X. Wang, N. Balakrishnan, and B. Guo, “Residual life estimation based on nonlinear-multivariate Wiener process,” *J. Stat. Comput. Simul.*, vol. 85, no. 9, pp. 1742–1764, 2015.
- [24] X. Zhang and A. Wilson, “System reliability and component importance under dependence: A Copula approach,” *Technometrics*, vol. 59, no. 2, pp. 215–224, 2017.
- [25] J. K. Sari, M. J. Newby, A. C. Brombacher, and L. C. Tang, “Bivariate constant stress degradation model: LED lighting system reliability estimation with two-stage modeling,” *Qual. Rel. Eng. Int.*, vol. 25, no. 8, pp. 1067–1084, 2009.
- [26] Z. Xi, R. Jing, P. Wang, and C. Hu, “A Copula-based sampling method for data-driven prognostics,” *Rel. Eng. Syst. Saf.*, vol. 132, pp. 72–82, 2014.
- [27] W. Peng, Y. Li, Y. Yang, S. Zhu, and H. Huang, “Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and Copulas,” *IEEE Trans. Rel.*, vol. 65, no. 2, pp. 624–639, Jun. 2016.
- [28] Y. Wang and H. Pham, “Modeling the dependent competing risks with multiple degradation processes and random shock using time-varying Copulas,” *IEEE Trans. Rel.*, vol. 61, no. 1, pp. 13–22, Mar. 2012.
- [29] C. C. Tsai, S. T. Tseng, and N. Balakrishnan, “Mis-specification analyses of gamma and Wiener degradation processes,” *J. Statist. Plan. Inference*, vol. 141, no. 12, pp. 3725–3735, 2011.
- [30] P. Wang, Y. Tang, S. J. Bae, and Y. He, “Bayesian analysis of two-phase degradation data based on change-point wiener process,” *Rel. Eng. Syst. Saf.*, vol. 170, pp. 244–256, 2018.
- [31] T. S. Ng, “An Application of the EM algorithm to degradation modeling,” *IEEE Trans. Rel.*, vol. 57, no. 1, pp. 2–13, Mar. 2008.
- [32] X. Si, C. Hu, X. Kong, and D. Zhou, “A residual storage life prediction approach for systems with operation state switches,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6304–6315, Nov. 2014.
- [33] L. Cui, J. Huang, and Y. Li, “Degradation models with Wiener diffusion processes under calibrations,” *IEEE Trans. Rel.*, vol. 65, no. 2, pp. 613–623, Jun. 2016.
- [34] R. B. Nelsen, *An Introduction to Copulas*, New York, USA: Springer, 2006.
- [35] A. Saxena, K. Goebel, D. Simon, and N. Eklund, “Damage propagation modeling for aircraft engine run-to-failure simulation,” in *Proc. Int. Conf. Prognostics and Health Management*, IEEE, Denver, USA, 2008, pp. 1–9.
- [36] I. K. Lin, “A concordance correlation coefficient to evaluate reproducibility,” *Biometrics*, vol. 45, no. 1, pp. 255–268, Mar. 1989.
- [37] G. Babu, P. Zhao, and X. Li, “Deep convolutional neural network based regression approach for estimation of remaining useful life,” in *Proc. Int. Conf. Database Syst. advanced Appl.*, Springer, Cham, 2016, pp. 214–228.



Jingdong Lin received the B.E. M.S. and Ph.D. degrees from Chongqing University, China in 1989, 1992 and 2002, respectively. He joined the College of Automation, Chongqing University in 1992. He is currently an Associate Dean of the College of Automation, Chongqing University. His current research interests mainly include industrial automation and remaining useful life prediction.



Zheng Lin received the M.S. degree from the School of Automation, Chongqing University, China in 2020. His current research interests include remaining useful prediction and fault diagnosis.



Guobo Liao received the B.E. degree from College of Automation, Chongqing University, China in 2019. He is currently a master student at the College of Automation, Chongqing University, China. His research interests include remaining useful life prediction and fault diagnosis.



Hongpeng Yin received the B.E. and Ph.D. degrees all from College of Automation, Chongqing University, China in 2003 and 2009, respectively. After that, he joined the College of Automation, Chongqing University. He is currently an Associate Dean of the College of Automation, Chongqing University. His current research interests mainly include machine vision, information fusion and data mining.