SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

June 2022, Vol. 65 162206:1–162206:17
https://doi.org/10.1007/s11432-020-2980-9

Prognostics based on the generalized diffusion process with parameters updated by a sequential Bayesian method

Hong PEI[1], Xiaosheng SI[1,2*], Changhua HU[1], Jianxun ZHANG[1], Dangbo DU[1], Zhenan PANG[1] & Shengfei ZHANG[1]

[1]*School of Missile Engineering, Rocket Force University of Engineering, Xi'an* 710025*, China;*
[2]*School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an* 710049*, China*

Received 17 February 2020/Revised 6 May 2020/Accepted 29 May 2020/Published online 26 May 2022

Abstract The realistic degradation process for the engineering equipment is generally stochastic and complicated owing to the uncertain operational condition and multiple functional loading, exhibiting the absolute nonlinear distinction. Such a nonlinear degradation process is widely modeled as a generalized diffusion process. When utilizing the generalized diffusion process-based model, certain model parameters are considered as the random variables to characterize the unit-to-unit discrepancies. Hence, the estimation of these kinds of parameters usually resorts to the Bayesian method. However, owing to the complex pattern of the model parameters in the generalized diffusion process, computing the Bayesian updated parameters requires plenty of repeated calculation and integration operations once the new degradation monitoring information is available. This will inevitably lower the computing efficiency and real-time performance. Toward this end, this paper presents an adaptive prognostic method based on the generalized diffusion process to determine the remaining useful life (RUL) of degraded equipment. First, a generalized diffusion process-based degradation modeling framework is constructed to describe the health performance of stochastic degraded equipment under complex conditions. Then, we utilize the maximum likelihood estimation (MLE) method to estimate the initial model parameters by analyzing the historical degradation information. Furthermore, a sequential Bayesian method is proposed to recursively update the stochastic model parameters in the generalized diffusion process for particular equipment in service. Unlike the existing studies utilizing the Bayesian method, the primary contrast in the presented method lies in that there is no need to implement the calculation process with complicated integration repeatedly utilizing the whole degradation information obtained before the current time. Particularly, the current measured information is incorporated into the estimates of the stochastic parameters in the previous time to determine the corresponding posterior estimates at the current time. This can avoid repeated calculation and raise the efficiency to a certain extent. Thereafter, the RUL distribution is updated adaptively by incorporating the acquired posterior estimates. Finally, we provide two practical case studies associated with the gyroscope and 2017-T4 aluminum alloy to demonstrate the efficiency and advantage of the proposed sequential Bayesian method. The experimental results exhibit that the proposed method can increase the RUL prediction accuracy compared with the existing methods in the literature.

Keywords generalized diffusion process, stochastic model parameters, remaining useful life, maximum likelihood estimation, sequential Bayesian methods

Citation Pei H, Si X S, Hu C H, et al. Prognostics based on the generalized diffusion process with parameters updated by a sequential Bayesian method. Sci China Inf Sci, 2022, 65(6): 162206, https://doi.org/10.1007/s11432-020-2980-9

1 Introduction

In engineering practice, dependable, secure, and economical functioning during the life cycle for safety-critical equipment, including aviation control systems and nuclear power generators, is the basic and necessary requirement from the task completion dimension. To obtain such requirement, massive experts

* Corresponding author (email: sxs09@mails.tsinghua.edu.cn)

and scholars are dedicated mainly in the research on prognostics and health management (PHM) technique [1–3]. Being a fundamental and organized tool for determining the health status and creating the maintenance decision, the PHM can appraise the failure risk, predict the failure progression, and moderate the functional cost via management actions, like maintenance, replacement, etc. Over the past few decades, the PHM has attracted the major attention and been adopted in many industrial and military fields [4, 5].

It is distinguished that the foundation of the PHM is the remaining useful life (RUL) prediction. In general, the RUL of the particular equipment can be defined as the length of time from the present time to the end of useful life [6, 7]. The responsibility of RUL prediction is to evaluate the degradation, and to determine the RUL utilizing the in-situ condition monitoring (CM) data. Because a trouble that how long a monitored equipment can survive is frequently encountered in the industry [8–12], the necessity for executing the RUL prediction is transparent. In addition, RUL prediction can provide the vital information supporting for creating maintenance plans. Particularly, on the basis of RUL prediction, the time and way to take the management activities can be determined from both the cost-effective and safety perspectives [13–15]. Accordingly, an enormous volume of theoretical achievements on the RUL prediction have been created over continuous investigation and research.

With the rapid advancement of sensor technique, the mainstream methods for RUL prediction are data-driven methods. This primarily consists of the machine learning-based methods and the statistical data-driven methods [16]. The machine learning-based methods can establish the mapping relationship from the monitoring data to the RUL value without any proceeding specialization, such as the neural network (NN)-based methods [17] and the support vector machine (SVM)-based methods [18]. The machine learning-based methods can be considered as a black box model and it is applied in various fields owing to the strong universality. However, the deficiency for traditional machine learning-based methods is that the prediction results are just point estimations [19, 20]. The statistical data-driven methods have advantages in dealing with uncertainties for RUL prediction [21], whose typical representatives are diffusion process-based methods [22], Gamma process-based methods [23], and inverse Gaussian process-based methods [24]. As is well known, Gamma process-based methods and inverse Gaussian process-based methods are only suitable for the strictly monotonic degradation process. Owing to the complex operating environment and interaction of internal structure, the non-monotonic degradation is frequently identified in existence. When facing with the non-monotonic circumstances, it is transparently no longer suitable to utilize these two methods for degradation modeling and RUL prediction. In reverse, the diffusion process-based methods are efficient for describing the non-monotonic degradation, signifying such methods are closer to the actual situation and have the extensive application capacity. Moreover, such methods can exhibit the distinguished mathematical characteristics and obtain the parameters effortlessly. Additionally, temporal variability can be expressed by the diffusion process-based methods. Thus, the research in this paper is on the basis of diffusion process-based methods. It is noted that the calculation of the first hitting time (FHT) is not identical with that of non-first hitting time under the diffusion process-based methods. As a result, it is necessary to distinguish from these two calculations when implementing the RUL prediction.

As for diffusion process-based methods, Wiener process-based method is a special case. This method attracts the major attention for degradation modeling in recent years. The first attempt to apply Wiener process to degradation modeling and RUL prediction was implemented by Doksum and Hoyland [25]. Consequently, a large number of prognostic models based on Wiener process were developed to suit the actual situation finer and enhance the accuracy of RUL [26–29]. Among these prognostic models, Bayesian method is a typical method, which is usually utilized to update the stochastic model parameters, such as the work in [28,29]. Unfortunately, Wiener process can only conform to the linear degradation process but the various nonlinear degradation process may be identified in practice. When dealing with the nonlinear degradation process, the activity, including time scale transformation and log-transformation can be adopted to convert the nonlinear degradation process into the linear degradation process [30,31]. However, these transformations have certain limitations and are only applicable for particular cases. Therefore, Si et al. [32] presented the Brownian motion (BM) with a nonlinear drift to model the degradation process and derived the FHT distribution of the diffusion process crossing a constant threshold by the time-space transformation. Nevertheless, the model parameters are only estimated by maximum likelihood estimation (MLE) method and cannot be updated once the new monitoring data are available. To address this issue, the Bayesian method was adopted to update the stochastic parameters under a special nonlinear form [33]. Tang et al. [34] researched a two-stage parameter estimation method and derived an

approximate analytical RUL distribution in a closed-form of a diffusion process-based degradation process with measurement errors. From the iterative results by the Bayesian method, the posterior distribution of stochastic parameters can be updated once the new degradation monitoring information is available. In addition, Wang et al. [35] proposed an additive Wiener process model-based degradation modeling for hybrid deteriorating systems, which consists of a linear degradation part and a nonlinear part. The corresponding parameter estimation method similarly contained a two-stage procedure: offline estimating by MLE and online updating via the Bayesian method. Moreover, the Bayesian method can assure the accuracy of the stochastic parameters and RUL prediction.

Generally speaking, the estimation of stochastic parameters in diffusion process-based degradation model usually resorts to the Bayesian method. Owing to the complicated form of the model parameters in diffusion process, calculating the Bayesian updated parameters requires frequently repeated calculation and integration operations once the new degradation monitoring information is available. This will inevitably lower the computing efficiency and real-time performance. In particular, when the Wiener process is selected to describe the linear evolving process of the degraded equipment, the above problem may be non-existing owing to the special linear form of drift coefficient. Nevertheless, there will exists another transparent limitation that the Bayesian updated parameters can be obtained by only utilizing the current degradation measurement and cannot incorporate the whole degradation measurements. Thus, Bayesian updated mechanism for Wiener process cannot assure the prognosis accuracy of RUL. As a result, it is urgent to research a method to overcome the common problem for diffusion process and special problem for Wiener process.

To do so, this paper presents an adaptive prognostic method based on the generalized diffusion process to predict the RUL of degraded equipment. First, a generalized diffusion process-based degradation modeling framework is constructed to describe the health performance of stochastic degraded equipment under the complex conditions. Next, the MLE method is utilized to estimate the initial model parameters by analyzing the historical degradation information. Furthermore, a sequential Bayesian method is introduced to recursively update the stochastic model parameters in the generalized diffusion process for particular equipment in service. Thereafter, the RUL distribution is updated adaptively by incorporating the obtained posterior estimates. Overall, the main contribution of this study is that there is no need to implement the calculation process with complicated integration repeatedly utilizing the whole degradation information obtained before the current time compared with diffusion process-based model with the Bayesian method. Specifically, the current measured information is incorporated into the estimates of the stochastic parameters in the previous time to determine the corresponding posterior estimates at the current time. This can overcome the above mentioned common problem for diffusion process and raise the efficiency to some extent. In addition, the presented method can include the linear degradation model as a special case. Because the estimates of the stochastic parameters in the previous time can reflect the whole degradation information obtained before the current time, the presented method can avoid the problem that the Bayesian estimate for the stochastic parameter in the current time only depends on the current degradation measurement for Wiener process.

The remaining parts of the paper are organized as follows. Section 2 formulates the problems and constructs the degradation models. We estimate the unknown parameters in our model and update the stochastic parameters by a sequential Bayesian method in Section 3. The probability density function (PDF) and expectation of the RUL are derived in Section 4. Section 5 provides two practical case studies associated with the gyroscope and 2017-T4 aluminum alloy to illustrate the effectiveness and superiority of the developed method. We conclude this paper and discuss the future direction in Section 6.

2 Problem formulation

2.1 A generalized diffusion-based degradation modeling

Since the modern engineering equipment is heavily affected by the internal structure wear and external environmental loading, the practical degradation trajectory reflecting the health status of the concerned equipment is changeable and stochastic. Generally speaking, there are certain disadvantages in using a single linear or nonlinear degradation model, which is difficult to accurately describe the real situation of degradation performance and guarantee the prediction accuracy. Therefore, it is natural to adopt a generalized diffusion process that is the combination of multiple linear or nonlinear degradation processes

to construct the degradation model. Specifically, the generalized mathematical model can be formulated as [36]

$$X(t) = x_0 + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{f}(t, \boldsymbol{\theta}) + \sigma_B B(t), \tag{1}$$

where $X(t)$ denotes the degradation level of the concerned equipment at time t, x_0 denotes the corresponding initial degradation level. $\boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{f}(t, \boldsymbol{\theta})$ represents the drift coefficient controlling the degradation speed during the life cycle, where $\boldsymbol{f}(t, \boldsymbol{\theta})$ with a parameter vector $\boldsymbol{\theta}$ represents an n-dimension vector composed of a series of fundamental functions, and $\boldsymbol{\lambda}$ is the parameter vectors composed of the coefficient of the fundamental functions, i.e., $\boldsymbol{\lambda} \in \mathbb{R}^{n \times 1}$. $[\cdot]^{\mathrm{T}}$ denotes the transposition operation, σ_B represents the diffusion coefficient, and $B(t)$ is the standard BM. Note that σ_B and $B(t)$ are constituted together to describe the time-varying dynamics of the degraded equipment.

Remark 1. The degradation models defined in (1) can include most of the existing linear and nonlinear degradation models as the special cases. That is to say, when $\boldsymbol{f}(t, \boldsymbol{\theta}) = t$, the generalized degradation model can be simplified as the simple linear Wiener process-based model in [29, 37]; when $\boldsymbol{f}(t, \boldsymbol{\theta}) = \int_0^t \mu(\tau, \theta)\, \mathrm{d}\tau$, the generalized degradation model can be simplified as the nonlinear diffusion process-based model in [32, 34]. Additionally, when $\boldsymbol{f}(t, \boldsymbol{\theta}) = [t, \int_0^t \mu(\tau, \theta)\, \mathrm{d}\tau]^{\mathrm{T}}$, we can obtain the additive hybrid degradation model in [36]. The specific forms of $\boldsymbol{f}(t, \boldsymbol{\theta})$ should be determined according to the practical performance of the concerned degraded equipment. Thus, the above equation can be used to describe any degradation trajectory once selecting the proper fundamental functions.

To characterize the unit-to-unit heterogeneity among a batch of equipment, the parameter vectors are generally treated as the stochastic vectors following a certain distribution. Note that $\boldsymbol{\lambda}$ and $B(t)$ need to be defined in the same probability space. It is assumed that $\boldsymbol{\lambda}$ is statistically independent of $B(t)$. The most commonly utilized distribution is the multidimensional normal distribution with unknown parameters $(\boldsymbol{\mu}_{\boldsymbol{\lambda}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$, that is, $\boldsymbol{\lambda} \sim \mathrm{MVN}(\boldsymbol{\mu}_{\boldsymbol{\lambda}}, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}})$. And the diffusion coefficient is conventionally considered as the fix parameter for describing the similar properties. The basic idea on these parameter estimation methods in most literature is to fully utilize the historical equipment measurement information for offline estimation and the monitoring data of equipment in service for online updating $\boldsymbol{\lambda}$.

Remark 2. The reasons that the parameter vectors $\boldsymbol{\lambda}$ are assumed to follow the multidimensional normal distribution are to facilitate the derivation process and to obtain the analytical solutions. Additionally, such assumption has been adopted in the existing literature. When $\boldsymbol{\lambda}$ follows other multivariate distributions, it may be difficult to derive the analytical expressions of RUL distribution. Thus, numerical simulation algorithms are needed to obtain the numerical solutions of RUL distribution, which are time-consuming and poorly operational.

2.2 RUL prediction for particular equipment conditional on the CM data

The main task is to predict the RUL adaptively for particular equipment conditional on the CM data. We let $\boldsymbol{X}_{1:k} = \{X(t_1), X(t_2), \ldots, X(t_k)\}$ represent the total monitoring data of the concerned equipment up to time t_k. According to most literature in the field of RUL prediction, the RUL at time t_k can be defined as the following equation from the point of FHT:

$$L_k = \inf\{l_k : X(t_k + l_k) \geqslant w \,|\boldsymbol{X}_{1:k}\}, \tag{2}$$

where $X(t_k + l_k)$ represents the degradation measurement at time $t_k + l_k$, l_k represents the realization value of L_k. w denotes the preset failure threshold specified by expert knowledge, engineering experiences and industrial standard. Note that inf indicates the abbreviation of the infimum operation.

For the sake of deriving and updating the PDF $f_{L_k|\boldsymbol{X}_{1:k}}(l_k\,|\boldsymbol{X}_{1:k})$ of the RUL at time t_k, the important information in $\boldsymbol{X}_{1:k}$ should be fully utilized. When considering the stochastic effect of the parameter vectors $\boldsymbol{\lambda}$, the above PDF can be formulated as follows based on the law of total probability:

$$f_{L_k|\boldsymbol{X}_{1:k}}(l_k\,|\boldsymbol{X}_{1:k}) = \int\int\cdots\int f_{L_k|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}(l_k\,|\boldsymbol{\lambda}, \boldsymbol{X}_{1:k}) f(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k})\, \mathrm{d}\boldsymbol{\lambda}, \tag{3}$$

where $f_{L_k|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}(l_k\,|\boldsymbol{\lambda}, \boldsymbol{X}_{1:k})$ represents the conditional PDF of the RUL at time t_k, and $f(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k})$ is the PDF of $\boldsymbol{\lambda}$ conditional on $\boldsymbol{X}_{1:k}$. The unknown model parameters in (1) include the stochastic parameter vectors $\boldsymbol{\lambda}$, the fixed parameter vector $\boldsymbol{\theta}$ of the fundamental functions and the fixed diffusion coefficient σ_B. Usually, the Bayesian method is utilized to update $\boldsymbol{\lambda}$ and guarantee the accuracy of RUL prediction.

However, owing to the complicated form of the model parameters in generalized diffusion process, the Bayesian updated parameters require plenty of repeated calculation and integration operations once the new degradation monitoring information is available, resulting in the problems of the low computing efficiency and poor real-time performance. Additionally, the PDF of the RUL defined in (3) contains complex multiple integrals, which is difficult to directly solve (3). Thus, the following problems will be investigated in the paper.

(1) How to determine the fixed parameters including $\boldsymbol{\theta}$ and σ_B, and the hyper-parameters in the prior distribution of $\boldsymbol{\lambda}$?

(2) How to update the posterior distribution $f(\boldsymbol{\lambda}|\boldsymbol{X}_{1:k})$ of stochastic parameter vectors recursively with the sequential Bayesian framework without plenty of repeated calculation?

(3) How to obtain the explicit form of the PDF of the RUL for the concerned equipment according to the defined degradation model and update it utilizing the estimated parameters?

3 Parameter estimation

3.1 Estimating the fixed parameters and hyper-parameters in the prior distribution

As described previously, the parameters to be estimated offline in the generalized diffusion process-based model can be classified as the fixed parameters and the hyper-parameters in the prior distribution, which are denoted as $\boldsymbol{\Theta}=[\boldsymbol{\mu}_{\boldsymbol{\lambda}}, \Sigma_{\boldsymbol{\lambda}}, \boldsymbol{\theta}, \sigma_{\mathrm{B}}]$. Let $\{X_i(t_{i,j}) = x_{i,j}, i = 1, \ldots, N, j = 1, \ldots, m_i\}$ represent the monitored historical degradation data from other analogous tested equipment, where N denotes the number of the tested equipment, m_i denotes the number of degradation measurements for the group of tested equipment, and $t_{i,j}$ represents the jth monitoring time for the ith group of tested equipment. Based on the degradation model in (1), the degradation observation at the jth monitoring time for the ith group of tested equipment can be expressed as

$$X_i(t_{i,j}) = x_{0,i} + \boldsymbol{\lambda}_i^{\mathrm{T}} \boldsymbol{f}(t_{i,j}, \boldsymbol{\theta}) + \sigma_B B(t_{i,j}). \tag{4}$$

To simplify the calculation of the likelihood function, we define $\boldsymbol{T}_i = [\boldsymbol{T}_{i,1}, \boldsymbol{T}_{i,2}, \ldots, \boldsymbol{T}_{i,m_i}]$, $\boldsymbol{T}_{i,j} = \boldsymbol{f}(t_{i,j}, \boldsymbol{\theta})$, $\boldsymbol{T}_i \in \mathbb{R}^{n \times m_i}$, $\boldsymbol{x}_i = [x_{i,1}, \ldots, x_{i,m_i}]^{\mathrm{T}}$. Without loss of generality, it is assumed that the degradation observations for the different equipment are statistically independent with each other. According to the independent increments property of diffusion process, it can be observed that $\boldsymbol{x}_i$ follows the multivariate normal distribution, i.e., $\boldsymbol{x}_i \sim \mathrm{MVN}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. The corresponding mean and covariance can be represented as

$$\boldsymbol{\mu}_i = x_{0,i}\boldsymbol{I}_1 + \boldsymbol{T}_i^{\mathrm{T}} \boldsymbol{\mu}_{\boldsymbol{\lambda}}, \tag{5}$$

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Omega}_i + \boldsymbol{T}_i^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{\lambda}} \boldsymbol{T}_i, \tag{6}$$

where $\boldsymbol{I}_1 = [1, 1, \ldots, 1]^{\mathrm{T}}$, $\boldsymbol{I} \in \mathbb{R}^{m_i \times 1}$, $\boldsymbol{\Omega}_i = \sigma_B^2 \boldsymbol{Q}_i$,

$$\boldsymbol{Q}_i = \begin{bmatrix} t_{i,1} & t_{i,1} & \cdots & t_{i,1} \\ t_{i,1} & t_{i,2} & \cdots & t_{i,2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{i,1} & t_{i,2} & \cdots & t_{i,m_i} \end{bmatrix}.$$

The degradation data for each equipment can together constitute all the monitored historical degradation data, i.e., $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]^{\mathrm{T}}$. Thus, from the characteristic of the multivariate normal distribution, the log-likelihood function of all historical degradation data $\boldsymbol{X}$ over parameters vector $\boldsymbol{\Theta}$ can be written as [32]

$$\ell(\boldsymbol{\Theta}|X) = -\frac{\ln(2\pi)}{2} \sum_{i=1}^{N} m_i - \frac{1}{2} \sum_{i=1}^{N} \ln|\boldsymbol{\Sigma}_i| - \frac{1}{2} \sum_{i=1}^{N} \Big[(\boldsymbol{x}_i - \boldsymbol{\mu}_i)^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{x}_i - \boldsymbol{\mu}_i) \Big]. \tag{7}$$

To maximize the log-likelihood function in (7), taking the first partial derivative of the log-likelihood function with respect to $\boldsymbol{\mu}_{\boldsymbol{\lambda}}$ can yield

$$\frac{\partial \ell(\boldsymbol{\Theta}|X)}{\partial \mu_\lambda} = \sum_{i=1}^{N} \big[\boldsymbol{T}_i \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{x}_i - x_{0,i}\boldsymbol{I}_1) - \boldsymbol{T}_i \boldsymbol{\Sigma}_i^{-1} \boldsymbol{T}_i^{\mathrm{T}} \boldsymbol{\mu}_\lambda \big]. \tag{8}$$

For the specific values of $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$, $\boldsymbol{\theta}$, and σ_B, let the derivative in (8) equal zero, and thus the estimated results of $\boldsymbol{\mu}_{\boldsymbol{\lambda}}$ can be formulated as

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{\lambda}} = \left[\sum_{i=1}^{N}\left(\boldsymbol{T}_i\boldsymbol{\Sigma}_i^{-1}\boldsymbol{T}_i^{\mathrm{T}}\right)\right]^{-1}\sum_{i=1}^{N}\left[\boldsymbol{T}_i\boldsymbol{\Sigma}_i^{-1}\left(\boldsymbol{x}_i - x_{0,i}\boldsymbol{I}_1\right)\right]. \tag{9}$$

After obtaining the results $\hat{\boldsymbol{\mu}}_{\boldsymbol{\lambda}}$ of MLE, we can further determine the profile log-likelihood function over $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$, $\boldsymbol{\theta}$, and σ_B by substituting (9) into (7), which can be formulated as

$$\begin{aligned}\ell\left(\boldsymbol{\Theta}\,|X\right) = &-\frac{\ln\left(2\pi\right)}{2}\sum_{i=1}^{N}m_i - \frac{1}{2}\sum_{i=1}^{N}\ln\left|\boldsymbol{\Sigma}_i\right| - \frac{1}{2}\sum_{i=1}^{N}\left(\boldsymbol{x}_i - x_{0,i}\boldsymbol{I}_1\right)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\left(\boldsymbol{x}_i - x_{0,i}\boldsymbol{I}_1\right)\\ &- 2\sum_{i=1}^{N}\left[\left(\boldsymbol{x}_i - x_{0,i}\boldsymbol{I}_1\right)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{T}_i^{\mathrm{T}}\right]\hat{\boldsymbol{\mu}}_{\boldsymbol{\lambda}} + \hat{\boldsymbol{\mu}}_{\boldsymbol{\lambda}}^{\mathrm{T}}\left[\sum_{i=1}^{N}\left(\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{T}_i\right)\right]\hat{\boldsymbol{\mu}}_{\boldsymbol{\lambda}}.\end{aligned} \tag{10}$$

As a result, the results of MLE for $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$, $\boldsymbol{\theta}$, and σ_B in $\boldsymbol{\Theta}$, can be obtained by maximizing the above profile log-likelihood function. Subsequently, the corresponding estimated results can be recorded as $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\lambda}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\sigma}_B$. The commonly utilized method to maximize the profile log-likelihood function is multi-dimensional search, which can be achieved by the "Fminsearch" function in MATLAB as described in [32, 35]. Furthermore, the estimated result of $(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\lambda}}, \hat{\boldsymbol{\theta}}, \hat{\sigma}_B)$ obtained by maximizing (10) is substituted into (9). Until now, the fixed parameters and the hyper-parameters in the prior distribution of the stochastic vector $\boldsymbol{\lambda}$ have been estimated utilizing the historical degradation measurement for all the tested equipment. Owing to the fact that the stochastic vector $\boldsymbol{\lambda}$ describes the unit-to-unit heterogeneity, the distribution parameters of $\boldsymbol{\lambda}$ are not identical for different equipment, and thus we will research how to update the distribution parameters of $\boldsymbol{\lambda}$ based on the real-time degradation measurement for particular equipment under the sequential Bayesian framework.

3.2 Stochastic parameter online updating using a sequential Bayesian method

The Bayesian estimation method can be served as the common method for parameters updating and has been widely used in existing literature, such as [38, 39]. The Bayesian estimation method in [38] belongs to the traditional Bayesian methods, indicating that the updated result in the last time is not fully utilized during the updating process when the newly measured data are available. The work in [39] offers a systematic introduction to the Bayesian state estimation framework. The main idea of such framework is that the updated information in the last time is fully utilized to update the parameters at the moment, which is consistent with the sequential Bayesian method. The differences between [39] and the sequential Bayesian method lie in the complex recursive mechanism of stochastic parameters in the state space equation. The complex recursive mechanism of stochastic parameters in [39] may not be conducive to the derivation of analytical expressions for stochastic parameters. As a result, we adopt the sequential Bayesian method to update the stochastic vector $\boldsymbol{\lambda}$ in the defined degradation model given the monitored degradation information up to time t_k for particular equipment in service.

As described before, the parameter vector $\boldsymbol{\lambda}$ follows the multidimensional normal distribution whose initial mean and covariance can be determined by MLE method in Subsection 3.1. To facilitate the illustration, the monitored degradation information up to time t_k for particular equipment in service can be denoted as $\boldsymbol{X}_{1:k}$. The main idea of sequential Bayesian method is that posterior estimates of $\boldsymbol{\lambda}$ in the last time are regarded as the prior information in the current time. Hence, once the new degradation data at time t_k are available, we can fully utilize the current degradation information and posterior estimates of $\boldsymbol{\lambda}$ at time t_{k-1} to compute the posterior estimates of $\boldsymbol{\lambda}$ at time t_k. Based on this, the posterior distribution of $\boldsymbol{\lambda}$ at time t_k can be written as

$$\begin{aligned}p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k}\right) &= \frac{p\left(\boldsymbol{X}_{1:k}\,|\boldsymbol{\lambda}\right)\cdot p\left(\boldsymbol{\lambda}\right)}{p\left(\boldsymbol{X}_{1:k}\right)} = \frac{p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)\cdot p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right)\cdot p\left(\boldsymbol{X}_{1:k-1}\right)}{p\left(\boldsymbol{X}_{1:k}\right)}\\ &= \frac{p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)\cdot p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right)}{p\left(x_k\,|\boldsymbol{X}_{1:k-1}\right)} \propto p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)\cdot p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right),\end{aligned} \tag{11}$$

where $p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)$ is the probability distribution of x_k given $\boldsymbol{X}_{1:k-1}$ and $\boldsymbol{\lambda}$, $p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right)$ represents the posterior distribution of $\boldsymbol{\lambda}$ at time t_{k-1} conditional on $\boldsymbol{X}_{1:k-1}$. The above equation describes the recur-

rence relation of the posterior estimates at two consecutive moments in essence, which can effectively simplify the repeated calculations. From the degradation model in (4), it can be observed that $x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}$ is a random variable following the normal distribution with the mean $x_{k-1}+\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{f}(t_k,\boldsymbol{\theta})-\boldsymbol{f}(t_{k-1},\boldsymbol{\theta}))$ and variance $\sigma_B^2\left(t_k-t_{k-1}\right)$, that is,

$$x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\sim N\left(x_{k-1}+\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right),\sigma_B^2\left(t_k-t_{k-1}\right)\right). \tag{12}$$

Hence, the corresponding PDF of $x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}$ can be formulated as

$$p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)=\frac{1}{\sqrt{2\pi\sigma_B^2\left(t_k-t_{k-1}\right)}}\cdot\exp\left[-\frac{\left(x_k-x_{k-1}-\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\right)^2}{2\sigma_B^2\left(t_k-t_{k-1}\right)}\right]. \tag{13}$$

Because of the conjugate prior distribution in Bayesian framework, the posterior distribution of $\boldsymbol{\lambda}$ at any time can be considered as the multidimensional normal distribution. That is to say, the mean and covariance of $\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}$ at time t_{k-1} conditional on $\boldsymbol{X}_{1:k-1}$ can be denoted as $\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}$ respectively. Accordingly, the mean and covariance of $\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k}$ at time t_k conditional on $\boldsymbol{X}_{1:k}$ can be also denoted as $\boldsymbol{\mu}_{\boldsymbol{\lambda},k}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}$ respectively. The posterior distribution of $\boldsymbol{\lambda}$ at time t_k can be provided by the following theorem.

Theorem 1. For the specific equipment in service, the mean and covariance of $\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k}$ at time t_k conditional on degradation data $\boldsymbol{X}_{1:k}$ can be updated as

$$\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}=\left[\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)^{\mathrm{T}}}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\right]^{-1}, \tag{14}$$

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{\lambda},k}=&\left[\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)^{\mathrm{T}}}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\right]^{-1}\\&\cdot\left(\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(x_k-x_{k-1}\right)}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right).\end{aligned} \tag{15}$$

Proof. As illustrated previously, it can be found that $\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\sim\mathrm{MVN}\left(\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1},\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}\right)$, and thus its PDF can be formulated as

$$p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right)=\frac{1}{(2\pi)^{\frac{n}{2}}\left|\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}\right|^{\frac{1}{2}}}\cdot\exp\left[-\frac{1}{2}(\boldsymbol{\lambda}-\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1})^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\left(\boldsymbol{\lambda}-\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right)\right]. \tag{16}$$

Substituting $p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)$ in (13) and $p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right)$ in (16) into (11), we can further obtain the posterior distribution of $\boldsymbol{\lambda}$ at time t_k, which can be expressed as

$$\begin{aligned}&p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k}\right)\propto p\left(x_k\,|\boldsymbol{X}_{1:k-1},\boldsymbol{\lambda}\right)\cdot p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k-1}\right)\\&\propto\exp\left[-\frac{\left(x_k-x_{k-1}-\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\right)^2}{2\sigma_B^2\left(t_k-t_{k-1}\right)}\right]\cdot\exp\left[-\frac{1}{2}(\boldsymbol{\lambda}-\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1})^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\left(\boldsymbol{\lambda}-\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right)\right]\\&=\exp\left[-\frac{\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)^{\mathrm{T}}\boldsymbol{\lambda}}{2\sigma_B^2\left(t_k-t_{k-1}\right)}\right.\\&\quad+\frac{\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(x_k-x_{k-1}\right)}{\sigma_B^2\left(t_k-t_{k-1}\right)}\\&\quad\left.-\frac{\left(x_k-x_{k-1}\right)^2}{2\sigma_B^2\left(t_k-t_{k-1}\right)}-\frac{1}{2}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\lambda}+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}-\frac{1}{2}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right]\\&=\exp\left[-\left(\frac{\left(x_k-x_{k-1}\right)^2}{2\sigma_B^2\left(t_k-t_{k-1}\right)}+\frac{1}{2}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right)\right.\\&\quad+\boldsymbol{\lambda}^{\mathrm{T}}\left(\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(x_k-x_{k-1}\right)}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right)\\&\quad\left.-\boldsymbol{\lambda}^{\mathrm{T}}\left(\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)^{\mathrm{T}}}{2\sigma_B^2\left(t_k-t_{k-1}\right)}+\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\right)\boldsymbol{\lambda}\right].\end{aligned} \tag{17}$$

Moreover, from the multidimensional normal property of $\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k}$ at time t_k conditional on $\boldsymbol{X}_{1:k}$, we have

$$\begin{aligned} p\left(\boldsymbol{\lambda}\,|\boldsymbol{X}_{1:k}\right) &= \frac{1}{(2\pi)^{\frac{n}{2}}\left|\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}\right|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\boldsymbol{\lambda}-\boldsymbol{\mu}_{\boldsymbol{\lambda},k})^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}^{-1}\left(\boldsymbol{\lambda}-\boldsymbol{\mu}_{\boldsymbol{\lambda},k}\right)\right] \\ &\propto \exp\left[-\frac{1}{2}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}^{-1}\boldsymbol{\lambda}+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k}-\frac{1}{2}\boldsymbol{\mu}_{\boldsymbol{\lambda},k}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k}\right]. \end{aligned} \tag{18}$$

As a result, the posterior parameters can be determined by comparing (17) and (18). Specifically, the third term of the exponential in (17) corresponds to the first term of the exponential in (18), and the second term of the exponential in (17) corresponds to the second term of the exponential in (18). Under these two corresponding relationships, the constant terms of the exponential in (17) and (18) seem to be not important. Thus, $\boldsymbol{\mu}_{\boldsymbol{\lambda},k}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k}$ can be formulated as

$$\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k} = \left[\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)^{\mathrm{T}}}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\right]^{-1},$$

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\lambda},k} = &\left[\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)^{\mathrm{T}}}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\right]^{-1} \\ &\cdot\left(\frac{\left(\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_{k-1},\boldsymbol{\theta}\right)\right)\left(x_k-x_{k-1}\right)}{\sigma_B^2\left(t_k-t_{k-1}\right)}+\boldsymbol{\Sigma}_{\boldsymbol{\lambda},k-1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\lambda},k-1}\right). \end{aligned}$$

This completes the proof of Theorem 1.

On the basis of Theorem 1, the stochastic parameter vectors $\boldsymbol{\lambda}$ can be updated for the equipment once new degradation data are obtained. The posterior estimators of $\boldsymbol{\lambda}$ at time t_k are associated with the current degradation data and posterior estimators of $\boldsymbol{\lambda}$ at time t_{k-1}. Such update mechanism can avoid plenty of repeated calculation and integration operations for traditional Bayesian method during the update process, which is beneficial for estimating model parameters quickly and guaranteeing the accuracy of RUL prediction effectively.

4 RUL prediction

In Section 2, a generalized diffusion-based degradation model is established and the concept of the RUL at time t_k is defined. Nevertheless, the life distribution should be derived first as the basis of RUL prediction. Thus, the life of the equipment can be defined as the time when the performance indicator firstly crosses the failure threshold, which is formulated as

$$T = \inf\left\{t : X\left(t\right) \geqslant w\,|X\left(0\right) < w\right\}, \tag{19}$$

where t represents the realization value of T. The PDFs of life and RUL should be derived from (19) and (2). To implement the corresponding derivation, the following two lemmas should be given at first.

Lemma 1. When ignoring the stochastic nature of $\boldsymbol{\lambda}$, the conditional PDF of the life for the equipment with the defined degradation process can be approximated as [32]

$$f_{T|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(t\,|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}\right) = \frac{w-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)+t\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}'\left(t,\boldsymbol{\theta}\right)}{\sqrt{2\pi t^3\sigma_B^2}}\cdot\exp\left[-\frac{\left(w-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)\right)^2}{2t\sigma_B^2}\right], \tag{20}$$

where $\boldsymbol{f}'\left(t,\boldsymbol{\theta}\right)$ represents the partial derivation of $\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)$ with respect to t. The work in [32] provides the detailed proof process, and thus the proof is omitted in the paper.

Lemma 2. Ignoring the unit-to-unit variability, the conditional PDF of the RUL at the time t_k for the equipment with the defined degradation process can be approximately expressed as

$$f_{L_k|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(l_k\,|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}\right) = \frac{w_k-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)+l_k\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}'\left(t_k+l_k,\boldsymbol{\theta}\right)}{\sqrt{2\pi l_k^3\sigma_B^2}}\cdot\exp\left[-\frac{\left(w_k-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\right)^2}{2l_k\sigma_B^2}\right], \tag{21}$$

where $\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right) = \boldsymbol{f}\left(t_k+l_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)$, $w_k = w-X\left(t_k\right)$.

Proof. According to the general diffusion process-based model, the degradation level at time t_k can be formulated as

$$X\left(t_k\right)=x_0+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)+\sigma_B B\left(t_k\right). \tag{22}$$

And the degradation level at any time $t\,(t>t_k)$ can be also formulated as

$$X\left(t\right)=x_0+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)+\sigma_B B\left(t\right). \tag{23}$$

Combining these two equations, we have

$$X\left(t\right)-X\left(t_k\right)=\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)\right)+\sigma_B\left(B\left(t\right)-B\left(t_k\right)\right). \tag{24}$$

Since the time t and t_k satisfy the relationship $t=t_k+l_k$, thus the above equation can be further represented as

$$\begin{aligned}U\left(l_k\right)&=\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{f}\left(t_k+l_k,\boldsymbol{\theta}\right)-\boldsymbol{f}\left(t_k,\boldsymbol{\theta}\right)\right)+\sigma_B B\left(l_k\right)\\&=\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}^{*}\left(l_k,\boldsymbol{\theta}\right)+\sigma_B B\left(l_k\right),\end{aligned} \tag{25}$$

where $U\left(l_k\right)=X\left(t\right)-X\left(t_k\right)$. It can be observed that the RUL at time t_k can be considered as the time when the new stochastic process $\left\{U\left(l_k\right),l_k\geqslant 0\right\}$ crosses the threshold $w_k=w-X\left(t_k\right)$.

For the transformed stochastic process $\left\{U\left(l_k\right),l_k\geqslant 0\right\}$, according to the theorem in [32], we can easily obtain the PDF of time when $\left\{U\left(l_k\right),l_k\geqslant 0\right\}$ crosses the threshold w_k. Therefore, when ignoring the unit-to-unit variability, the conditional RUL at time t_k for the concerned equipment can be represented as

$$f_{L_k|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(l_k\left|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}\right.\right)=\frac{w_k-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}^{*}\left(l_k,\boldsymbol{\theta}\right)+l_k\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}'\left(t_k+l_k,\boldsymbol{\theta}\right)}{\sqrt{2\pi l_k^3\sigma_B^2}}\cdot\exp\left[-\frac{\left(w_k-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}^{*}\left(l_k,\boldsymbol{\theta}\right)\right)^2}{2l_k\sigma_B^2}\right]. \tag{26}$$

This completes the proof of Lemma 2.

Note that the conditional PDFs of the life and RUL ignoring the stochastic nature of $\boldsymbol{\lambda}$ are provided in (20) and (21), and it is natural to consider that the law of total probability should be adopted to derive the PDFs of the life and RUL. The specific expressions can be formulated as

$$\begin{aligned}f_{T|\boldsymbol{X}_{1:k}}\left(t\left|\boldsymbol{X}_{1:k}\right.\right)&=\int\cdots\int f_{T|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(t\left|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}\right.\right)f\left(\boldsymbol{\lambda}\left|\boldsymbol{X}_{1:k}\right.\right)\mathrm{d}\boldsymbol{\lambda}\\&=E_{\boldsymbol{\lambda}|\boldsymbol{X}_{1:k}}\left[f_{T|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(t\left|\boldsymbol{\lambda},X_{1:k}\right.\right)\right],\end{aligned} \tag{27}$$

$$\begin{aligned}f_{L_k|\boldsymbol{X}_{1:k}}\left(l_k\left|\boldsymbol{X}_{1:k}\right.\right)&=\int\cdots\int f_{L_k|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(l_k\left|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}\right.\right)f\left(\boldsymbol{\lambda}\left|\boldsymbol{X}_{1:k}\right.\right)\mathrm{d}\boldsymbol{\lambda}\\&=E_{\boldsymbol{\lambda}|\boldsymbol{X}_{1:k}}\left[f_{L_k|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}}\left(l_k\left|\boldsymbol{\lambda},\boldsymbol{X}_{1:k}\right.\right)\right].\end{aligned} \tag{28}$$

From the previous description, it can be found that $\boldsymbol{\lambda}\sim\mathrm{MVN}\left(\boldsymbol{\mu}_{\boldsymbol{\lambda}},\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\right)$. However, the complex multiple integrals exist in the process of implementing the law of total probability. It may be especially difficult to derive the corresponding theoretical expressions. To address this problem, we give the following lemma for facilitating the derivation of the PDFs of the life and RUL.

Lemma 3. It is assumed that the stochastic vector $\boldsymbol{\rho}$ follows an n-dimensional normal distribution with corresponding mean $\boldsymbol{\mu}_{\boldsymbol{\rho}}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{\rho}}$, where the covariance matrix is a positive definite matrix. υ_1, υ_2 and c are the given constants, and $\boldsymbol{a}$, $\boldsymbol{b}$ are n-dimensional vectors. Then we can obtain the following formulation:

$$\begin{aligned}&E_{\boldsymbol{\rho}}\left[\left(\upsilon_1-\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\rho}\right)\exp\left(-\frac{\left(\upsilon_2-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\rho}\right)^2}{2c}\right)\right]\\&=\sqrt{\frac{c^n}{\left|\boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\rho}}+c\boldsymbol{I}\right|}}\left(\upsilon_1-\frac{\upsilon_1\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\rho}}\boldsymbol{b}+c\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\rho}}}{c+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\rho}}\boldsymbol{b}}\right)\exp\left[-\frac{\left(\upsilon_2-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\rho}}\right)^2}{2\left(c+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\rho}}\boldsymbol{b}\right)}\right],\end{aligned} \tag{29}$$

where $\boldsymbol{I}$ is an identity matrix and $\boldsymbol{I}\in\mathbb{R}^{n\times n}$. The proof of Lemma 3 was provided in [36]. This lemma is so critical to derive the theoretical expressions of the PDFs of life and RUL. On the basis of Lemma 3,

the PDF of the life in (27) and the PDF of the RUL in (28) can be further expressed as

$$f_{T|\boldsymbol{X}_{1:k}}\left(t\,|\boldsymbol{X}_{1:k}\right)=\left[\frac{w\left(t\sigma_B^2+\boldsymbol{f}^{\mathrm{T}}\left(t,\boldsymbol{\theta}\right)\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)-\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)\right)+t\sigma_B^2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\lambda}}}{t\sigma_B^2+\boldsymbol{f}^{\mathrm{T}}\left(t,\boldsymbol{\theta}\right)\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)}\right]\times\sqrt{\frac{\left(t\sigma_B^2\right)^{n-1}}{2\pi t^2\left|\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)\boldsymbol{f}^{\mathrm{T}}\left(t,\boldsymbol{\theta}\right)\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}+t\sigma_B^2\boldsymbol{I}\right|}}\times\exp\left[-\frac{\left(w-\boldsymbol{f}^{\mathrm{T}}\left(t,\boldsymbol{\theta}\right)\boldsymbol{\mu}_{\boldsymbol{\lambda}}\right)^2}{2\left(t\sigma_B^2+\boldsymbol{f}^{\mathrm{T}}\left(t,\boldsymbol{\theta}\right)\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)\right)}\right],\tag{30}$$

where $\boldsymbol{a}^{\mathrm{T}}=\left[\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)-t\boldsymbol{f}'\left(t,\boldsymbol{\theta}\right)\right]^{\mathrm{T}}$.

$$f_{L_k|\boldsymbol{X}_{1:k}}\left(l_k\,|\boldsymbol{X}_{1:k}\right)=\left[\frac{w_k(l_k\sigma_B^2+\left(\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)-\boldsymbol{a}_L^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right))+l_k\sigma_B^2\boldsymbol{a}_L^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\lambda}}}{l_k\sigma_B^2+\left(\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)}\right]\times\sqrt{\frac{\left(l_k\sigma_B^2\right)^{n-1}}{2\pi l_k^2|\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\left(\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}+l_k\sigma_B^2\boldsymbol{I}|}}\times\exp\left[-\frac{\left(w-\left(\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\lambda}}\right)^2}{2(l_k\sigma_B^2+\left(\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right))}\right],\tag{31}$$

where $\boldsymbol{a}_L^{\mathrm{T}}=\left[\boldsymbol{f}^*\left(l_k,\boldsymbol{\theta}\right)-l_k\boldsymbol{f}'\left(t_k+l_k,\boldsymbol{\theta}\right)\right]^{\mathrm{T}}$. When we obtain the offline estimated results for $\boldsymbol{\theta}$ and σ_B, and the online updated results for $\boldsymbol{\mu}_{\boldsymbol{\lambda}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$ at t_k by the methods in Section 3, the PDFs of the life and RUL can be adaptively determined once new CM data are gained, which is beneficial to make the scheduled maintenance activities scientifically. Additionally, the expectations of the life and RUL can be further derived on the basis of the relevant content in the probability statistics. The specific expressions can be formulated as

$$E\left(T\right)=\int_0^\infty tf_{T|\boldsymbol{X}_{1:k}}\left(t\,|\boldsymbol{X}_{1:k}\right)\mathrm{d}t,\tag{32}$$

$$E\left(L_k\right)=\int_0^\infty l_kf_{L_k|\boldsymbol{X}_{1:k}}\left(l_k\,|\boldsymbol{X}_{1:k}\right)\mathrm{d}l_k.\tag{33}$$

Owing to the complicated form of $f_{T|\boldsymbol{X}_{1:k}}\left(t\,|\boldsymbol{X}_{1:k}\right)$ and $f_{L_k|\boldsymbol{X}_{1:k}}\left(l_k\,|\boldsymbol{X}_{1:k}\right)$, it is especially difficult to directly acquire the theoretical explicit solutions by the integral operation in (32) and (33). Nevertheless, the numerical solution can be easily obtained by converting integral operation into a large number of summations with the help of MATLAB.

To facilitate the understanding and application, the main procedure of the sequential Bayesian updated diffusion process model for adaptive RUL prediction is summarized in Algorithm 1.

Algorithm 1 Sequential Bayesian updated diffusion process model for adaptive RUL prediction

Input: Historical degradation data $\boldsymbol{X}$, and online degradation data $\boldsymbol{X}_{1:k}$;
Output: PDFs and expectations of life and RUL;
Step1: Construct the generalized diffusion process in (1) to describe the degradation model for the engineering equipment;
Step2: Estimate the fixed parameters and the hyper-parameters in the prior distribution based on $\boldsymbol{X}$ by MLE, namely $\hat{\Theta}$;
Step3: Initialize the hyper-parameters of stochastic vector $\boldsymbol{\lambda}$ utilizing the offline estimated results;
Step4: Update the hyper-parameters of stochastic vector $\boldsymbol{\lambda}$ at next time by (14) and (15);
Step5: $k=k+1$, and go to Step4. When stochastic vector $\boldsymbol{\lambda}$ at any time is updated, go to Step6;
Step6: Predict the RUL and quantify uncertainty based on (30)–(33).

5 Experimental studies

For the verification purpose, this section provides two experimental studies related with the gyropic drift data of inertial navigation systems and the fatigue-crack growth data of 2017-T4 aluminum alloy to implement the research on RUL prediction. The state-of-the-art prognostic methods in [32, 35], which are the nonlinear Wiener process-based model without the updating mechanism, and additive Wiener process-based model updated by traditional Bayesian method respectively, will be introduced to compare the prediction performance. To better illustrate the comparison results, the nonlinear Wiener process-based model without the updating mechanism, and additive Wiener process-based model updated by traditional Bayesian method can be recorded as Methods 1 and 2. Note that $\boldsymbol{f}\left(t,\boldsymbol{\theta}\right)$ of the proposed method can be selected as a relatively simple form $\left[t,\exp\left(\theta t\right)\right]^{\mathrm{T}}$, and the nonlinear drift coefficient of Methods 1 and 2 are considered as the exponential function form.

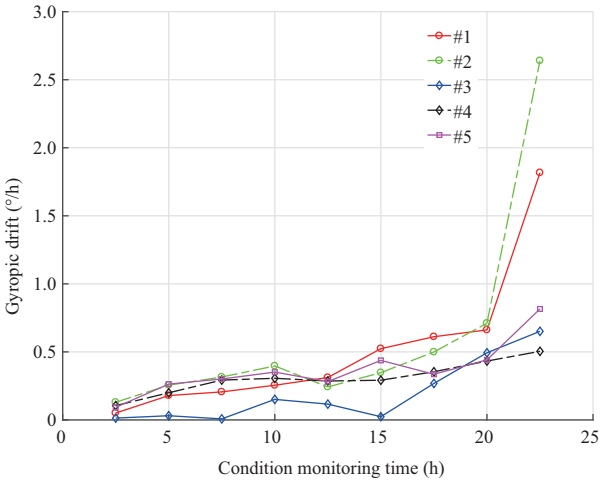

Figure 1 (Color online) The drift data of five gyroscopes.

Table 1 The estimated values of all the parameters for the three methods

	μ_{λ_1}	μ_{λ_2}	$\sigma^2_{\lambda_1}$	$\sigma^2_{\lambda_2}$	ς	θ	σ_B
Our model	1.13E−11	6.35E−9	8.53E−11	8.12E−11	8.35E−11	0.820	0.063
Method 1	9.34E−9	–	8.66E−9	–	–	0.815	0.066
Method 2	1.37E−11	6.63E−9	8.57E−11	8.14E−11	–	0.823	0.061

5.1 The RUL prediction for gyroscope

As a class of typical mechanical and electrical products, the gyroscope is mainly employed to measure the relevant information related with the angular velocity and plays an irreplaceable role in various fields including missiles, rockets, aircraft, and warships. Owing to the complicated structure and diverse environment, the gyroscope exhibits the degradation characteristic of the combination of multiple linear and nonlinear degradation processes instead of a single degradation process. In reality, such degradation may accumulate as time passes by, which can lead to bigger and bigger drift coefficient. In other words, the observed drift coefficient through the precision testing procedure can reflect the health condition of the gyroscope. Furthermore, we can implement the research on RUL based on the observed drift coefficient and the given technical specifications (such as the preset failure threshold). Note that once the drift coefficient exceeds the failure threshold, the gyroscope requires to carry out maintenance or replacement to guarantee the precision of the inertial navigation system.

Five groups of historical drift data of gyroscopes provided in [32] are utilized for degradation modeling and RUL prediction, which are depicted in Figure 1. The drift data of all the gyroscopes show an upward trend in the life cycle as a whole, indicating the performance degradation of gyroscopes to some extent. In this experiment, the monitoring interval for the concerned gyroscopes is 2.5 h. According to Figure 1, it can be observed that the mean time to failure (MTTF) is around 21.5 h when experiencing the continuous operation. As for every gyroscope, 9 suits of drift coefficient data are collected by CM technique. As described previously, we can estimate the fixed parameters and hyper-parameters in our model by the proposed MLE method utilizing these drift data of five gyroscopes, and the corresponding parameters can be summarized in Table 1. It is worth mentioning that because $\boldsymbol{f}(t,\boldsymbol{\theta})$ equals $[t,\exp(\theta t)]^{\mathrm{T}}$, the mean and covariance of $\boldsymbol{\lambda}=[\lambda_1,\lambda_2]^{\mathrm{T}}$ can be represented as $\boldsymbol{\mu}_{\boldsymbol{\lambda}}=[\mu_{\lambda_1},\mu_{\lambda_2}]^{\mathrm{T}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}=\left[\begin{smallmatrix}\sigma^2_{\lambda_1} & \varsigma\\ \varsigma & \sigma^2_{\lambda_2}\end{smallmatrix}\right]$, where ς represents the covariance of λ_1 and λ_2. Analogously, the estimated parameters for Methods 1 and 2 can also obtained by analyzing the drift data of five gyroscopes, which are listed in the same table.

After acquiring the offline estimated values of all the parameters in the proposed model, it is necessary to select a specific gyroscope for parameters updating and RUL prediction. Hence, gyroscope #4 can be selected as a candidates and the corresponding failure threshold is defined as the last drift data, which is different from the work in [32], i.e., $w=0.5042$. That is to say, its true lifetime is 22.5 h. Consequently, we can update the parameters of the stochastic vector via the sequential Bayesian method for the monitored gyroscope at each CM time and substitute the updated results into the PDF and expectation of the RUL in (31) and (33). To investigate the prediction capability of these three methods in the RUL prediction,

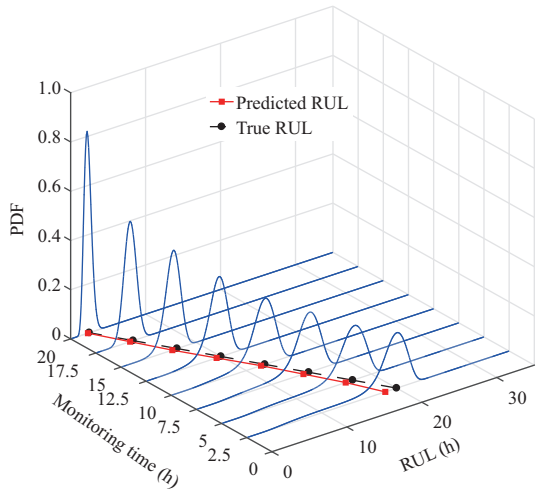

Figure 2 (Color online) The PDFs of the RULs under the proposed method with the drift data of the gyroscopes.

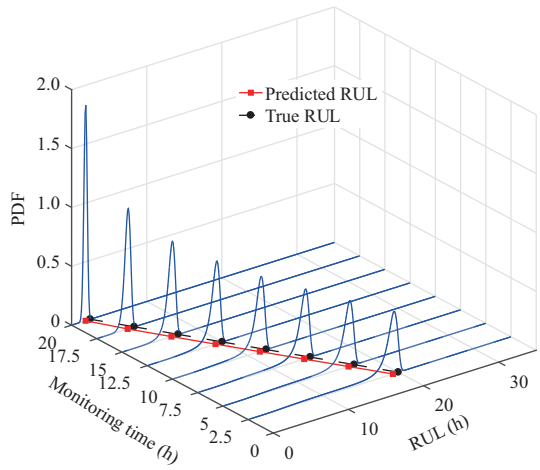

Figure 3 (Color online) The PDFs of the RULs under the Method 1 with the drift data of the gyroscopes.

we depict the PDF curves and the corresponding predicted values from the first monitoring point to the eighth monitoring point for these three methods in Figures 2–4. It can be seen from these three figures that all the PDF curves can cover the predicted results of the RUL well over all the CM times. Moreover, as more and more drift data are acquired, the PDF curves for these three methods will become higher and tighter, which implies that the uncertainties of RUL prediction can be decreased since more useful information is incorporated into the results of the RUL prediction. However, it can be easily observed from Figures 2–4 that the predicted RULs by Method 1 deviate greatly from the true RUL, especially at the last stage of the life cycle. This is because the parameter estimation of Method 1 is just MLE and cannot update the model parameters. Besides, the degradation model for Method 1 only contains the nonlinear component and can be regarded as the special case of the proposed method and Method 2, and thus the fitting capacity by Method 1 is inferior to that by other two methods. As for the proposed method and Method 2, there is little difference in prediction effect of RUL intuitively.

For the quantitative comparison, two generally utilized indicators including relative error (RE) and mean-squared error (MSE) are selected to evaluate the prediction performance of RUL. The former mainly reflects the prediction accuracy, while the latter takes both prediction accuracy and uncertainty into consideration. Specifically, RE and MSE of RUL prediction can be formulated as

$$\mathrm{RE}_k = \frac{|E\left(L_k\right) - \tilde{l}_k|}{\tilde{l}_k} \times 100\%, \tag{34}$$

$$\mathrm{MSE}_k = \int_0^{\infty} \left(l_k - \tilde{l}_k\right)^2 f_{L_k|\boldsymbol{X}_{1:k}}\left(l_k \,|\boldsymbol{X}_{1:k}\right) \mathrm{d}l_k, \tag{35}$$

where $E\left(L_k\right)$ and $\tilde{l}_k$ represent the predicted RUL and actual RUL at the observation time respectively. $f_{L_k|\boldsymbol{X}_{1:k}}\left(l_k \,|\boldsymbol{X}_{1:k}\right)$ denotes the PDF of the RUL utilizing the available degradation data up to time t_k based on (31). Based on the above illustration, Table 2 provides the REs of the predicted RULs for these three methods. From Table 2, it can be found that all the REs of the predicted RULs for the proposed model are below 10%, which indicates the effectiveness of our model in RUL prediction of the gyroscope. At the first and second monitoring points, the REs of the predicted RULs by Method 1 are less than those by other two methods. This is due to the fact the offline parameter estimation results by MLE are accurate at this time. However, there is no updating mechanism for Method 1, and thus the parameter estimation results will not change with new data acquisition. As a result, the REs of the predicted RULs for the proposed model are less than those by Method 1 after the second monitoring points. In addition, since our model and Method 2 utilize the updating mechanism, the prediction effects of these two methods are satisfied as a whole. Specifically, the prediction effects by our model are slightly better than those by Method 2. The average REs of the predicted RULs for these three methods are 6.5825%, 10.6337%, and 7.9225%. Moreover, the calculating times for our model and Method 2 are respectively 16.39 s and

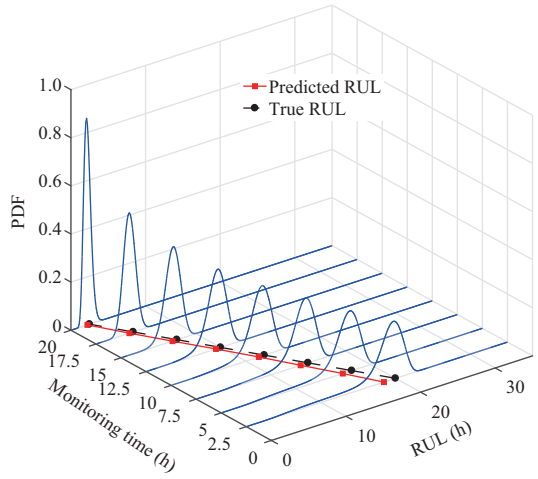

Figure 4 (Color online) The PDFs of the RULs under the Method 2 with the drift data of the gyroscopes.

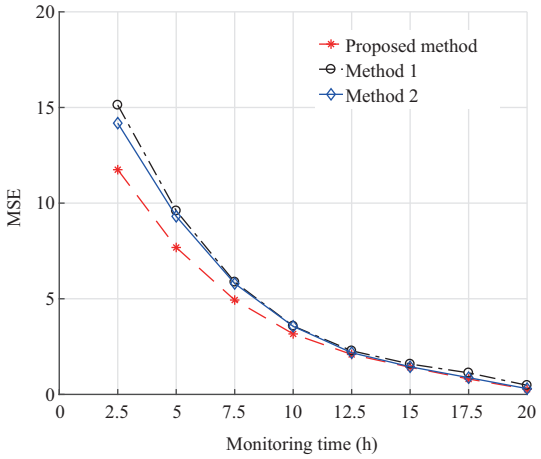

Figure 5 (Color online) The MSE comparison of the predicted RULs for three methods.

Table 2 The comparison of RE (%) for these three methods

	Time (h)							
	2.5	5	7.5	10	12.5	15	17.5	20
Our model	7.61	5.39	4.84	4.17	5.96	8.64	8.38	7.67
Method 1	3.64	4.42	5.48	6.96	9.06	12.3	17.85	25.36
Method 2	7.76	6.91	6.64	6.01	6.37	8.00	10.3	11.39

41.57 s, which implies that the proposed sequential Bayesian method can overcome the problem of plenty of repeated calculation and integration operations.

Furthermore, we provide the MSEs of the predicted RULs for these three methods in Figure 5. It can be observed that MSEs of the predicted RULs in the life cycle for these three methods have the downward trend as a whole. That is to say, as more drift data are available, MSEs of the predicted RULs can be decreased. Nevertheless, all the MSEs of the predicted RULs by our model are smaller than those by Methods 1 and 2. Note that at the late period in the life cycle (12.5 h, 15 h, and 20 h), the prediction effects by our model are slighter than those by other methods, which is inconsistent with the results of RE. Because the MSE takes not only the prediction accuracy, but also the uncertainty into consideration. Although the REs of the predicted RULs for Method 1 are so large at the late period, less uncertainties can lead to small MSE. The MSE values at 20 h for the proposed method, Method 1, and Method 2 are 0.2696, 0.4770, and 0.3022.

5.2 The RUL prediction for 2017-T4 aluminum alloy

As a vital metal material, 2017-T4 aluminum alloy is frequently applied in the military equipment and spacecraft. Engineering practice shows that the performance of such material is evaluated by the length of fatigue cracks. In general, when the crack length reaches or exceeds the preset threshold value, the mechanism structure will reach a critical state, which can be considered as the failure. The fatigue-crack growth data in four test specimens of 2017-T4 aluminum alloy provided in [32] are obtained under a stress level of 200 MPa. For each sample, ten crack levels are recorded, and the monitoring interval is 100000 cycles until the end of the experiment. Figure 6 depicts the evolving process of fatigue-crack growth data of 2017-T4. From Figure 6, it can be found that the degradation trajectories for these four test specimens of 2017-T4 aluminum alloy exhibit the nonlinearity to some extent.

Similarly, the offline parameter estimators for the above three methods can be acquired by MLE utilizing all the fatigue-crack growth data of 2017-T4, which are summarized in Table 3. After that, 2017-T4 aluminum alloy #3 is regarded as the research object for RUL prediction. Accordingly, the failure threshold is set to be the last fatigue crack data ($w = 5.6$). We utilize the proposed sequential Bayesian method in Section 3 to update the parameters of the stochastic vector $\boldsymbol{\lambda}$. Furthermore, the PDFs and the corresponding predicted values of the RUL from 1.5×10^5 cycles to 2.3×10^5 cycles for these

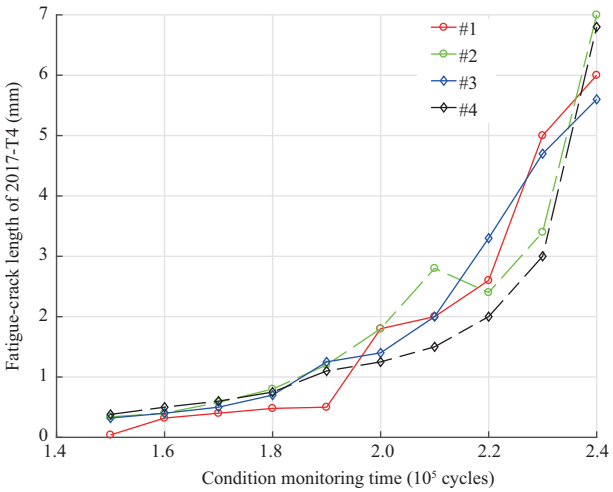

Figure 6 (Color online) Fatigue-crack growth data of 2017-T4.

Table 3 The estimated values of all the parameters for the three methods

	μ_{λ_1}	μ_{λ_2}	$\sigma^2_{\lambda_1}$	$\sigma^2_{\lambda_2}$	ς	θ	σ_B
Our model	1.15E−9	1.49E−4	7.13-11	3.43E−5	7.64E−11	4.4107	0.7535
Method 1	1.50E−4	–	3.531E−5	–	–	4.4402	0.7788
Method 2	1.42E−9	1.63E−4	7.83E−11	3.47E−5	–	4.3998	0.7623

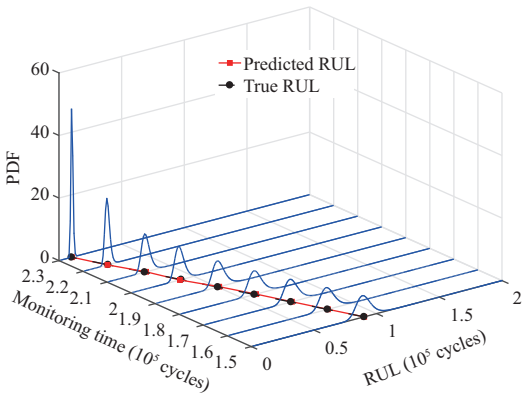

Figure 7 (Color online) The PDFs of the RULs under the proposed method with the fatigue crack data of 2017-T4 aluminum alloy.

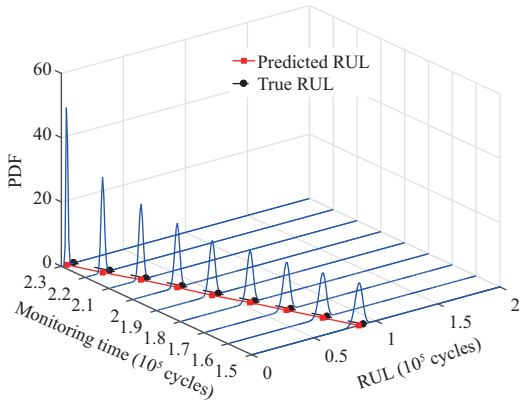

Figure 8 (Color online) The PDFs of the RULs under the Method 1 with the fatigue crack data of 2017-T4 aluminum alloy.

three methods are provided in Figures 7–9. It can be seen from these figures that the proposed method and Method 2 can predict RUL for 2017-T4 aluminum alloy more accurately compared with Method 1. Because the updating mechanism in the proposed method and Method 2 can guarantee the accuracy of the stochastic parameters utilizing the online CM data. Moreover, the prediction effect of the proposed method is slightly better than that of Method 2.

For sake of further investigating the prediction performance quantitatively, we provide the REs and MSEs of the RULs for these three methods in Figures 10 and 11. It can be observed that all the REs and MSEs of the RULs at each monitoring point by our proposed method are less than those by Methods 1 and 2. Especially at the 2.3×10^5 cycles, the RE and MSE of the predicted RUL for the proposed method are 3.84% and 8.41×10^{-4} respectively, which indicates the effectiveness of the proposed method in RUL prediction. What is more, the box plots of the predicted RUL at the 6th and 8th monitoring points for the three methods are depicted in Figure 12. From the figure, it can be found that the median of the RUL for the proposed method is closer to the true RUL compared with other methods. And it can be concluded that the dispersion degree of the proposed method is slightly less than that of Method 2,

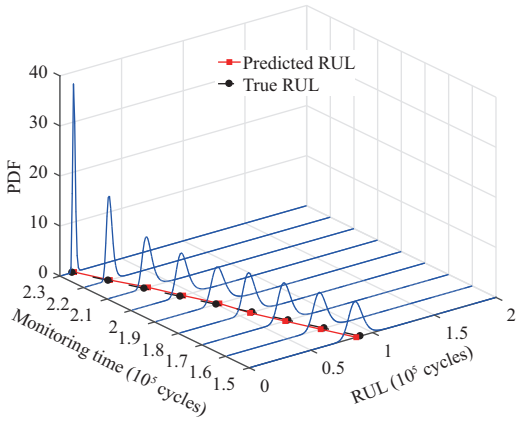

Figure 9 (Color online) The PDFs of the RULs under the Method 2 with the fatigue crack data of 2017-T4 aluminum alloy.

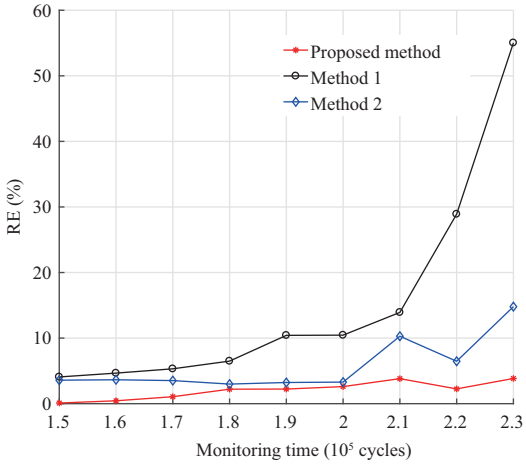

Figure 10 (Color online) The comparison of RE for these three methods.

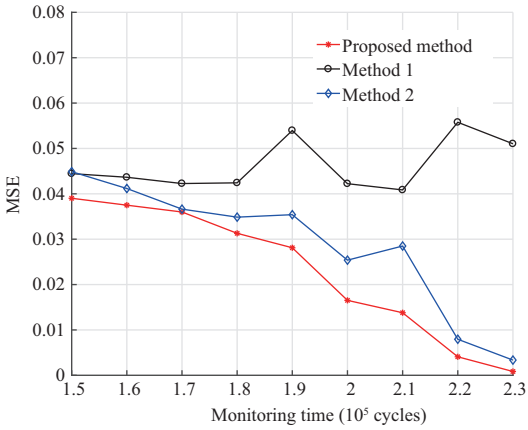

Figure 11 (Color online) The MSE comparison of the predicted RULs for three methods.

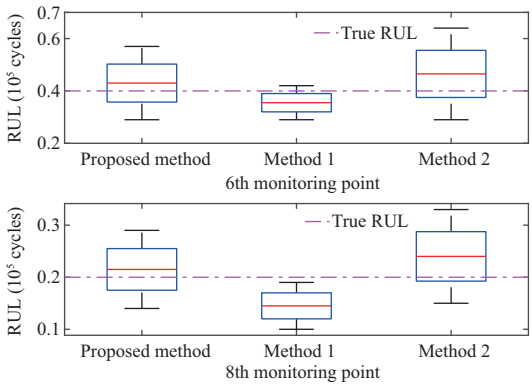

Figure 12 (Color online) Box plots of the predicted RUL at the 6th and 8th monitoring points.

which is consistent with the PDF curves in Figures 7 and 9.

Overall, the results in these two experimental studies imply that our proposed sequential Bayesian method is superior to Methods 1 and 2. Therefore, the proposed method has potential application prospects for the RUL estimation issues of the engineering equipment.

6 Conclusion

A sequential Bayesian updated diffusion process model is developed for adaptive RUL prediction. First, a generalized diffusion process-based degradation modeling framework is constructed to describe the health performance of stochastic degraded equipment under the complex conditions. Then, we utilize the MLE method to estimate the initial model parameters by analyzing the historical degradation information. Furthermore, the stochastic parameters in generalized diffusion process can be updated recursively based on the sequential Bayesian method for particular equipment in service. Hence, the RUL distribution is updated adaptively by incorporating the obtained posterior estimates. Finally, we provide two practical case studies associated with the gyroscope and 2017-T4 aluminum alloy to demonstrate the effectiveness and superiority of the proposed sequential Bayesian method. The improvements for degradation modeling and parameters updating are especially significant for decision making related activities including spare parts ordering and maintenance scheduling.

Despite the fact the generalized diffusion process-based degradation model for adaptive RUL prediction

via the sequential Bayesian method can achieve satisfactory results in the experimental studies and exhibit the superiority in RUL prediction performance, some issues are still available in this paper and require the additional research. First, the hidden or partially observed degradation cases may be encountered in the practical case, which deserves the further exploration. Second, when the stochastic vector does not follow multidimensional normal distribution, this will cause challenge for derivation. Thus, it is necessary to research the non-Gaussian case of the stochastic vector.

Acknowledgements The work was supported by National Natural Science Foundation of China (Grant Nos. 61833016, 61922089, 61773386, 61573365, 61573366, 61903376, 61773389, 61673311), Shaanxi National Science Foundation (Grant Nos. 2020JQ-489, 2020JM-360, 2020JQ-298), and National Key R&D Program of China (Grant No. 2018YFB1306100). The authors would like to thank the anonymous reviewers for their valuable suggestions.

References

1 Pecht M. Prognostics and Health Management of Electronics. Hoboken: Wiley, 2008
2 Jardine A K S, Lin D, Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Syst Signal Process, 2006, 20: 1483–1510
3 Lee J, Wu F, Zhao W, et al. Prognostics and health management design for rotary machinery systems-Reviews, methodology and applications. Mech Syst Signal Process, 2014, 42: 314–334
4 Tsui K L, Chen N, Zhou Q, et al. Prognostics and health management: a review on data driven approaches. Math Problems Eng, 2015, 2015: 1–17
5 Lei Y, Li N, Guo L, et al. Machinery health prognostics: a systematic review from data acquisition to RUL prediction. Mech Syst Signal Process, 2018, 104: 799–834
6 Si X S, Li T M, Zhang Q. A general stochastic degradation modeling approach for prognostics of degrading systems with surviving and uncertain measurements. IEEE Trans Rel, 2019, 68: 1080–1100
7 Zhai Q, Ye Z S. RUL prediction of deteriorating products using an adaptive Wiener process model. IEEE Trans Ind Inf, 2017, 13: 2911–2921
8 Lall P, Lowe R, Goebel K. Prognostics health management of electronic systems under mechanical shock and vibration using kalman filter models and metrics. IEEE Trans Ind Electron, 2012, 59: 4301–4314
9 Si X S, Li T M, Zhang Q, et al. Prognostics for linear stochastic degrading systems with survival measurements. IEEE Trans Ind Electron, 2020, 67: 3202–3215
10 Xi X P, Chen M Y, Zhou D H. Remaining useful life prediction for multi-component systems with hidden dependencies. Sci China Inf Sci, 2019, 62: 022202
11 Yu Y, Si X S, Hu C H, et al. Online remaining-useful-life estimation with a Bayesian-updated expectation-conditional-maximization algorithm and a modified Bayesian-model-averaging method. Sci China Inf Sci, 2021, 64: 112205
12 Jin X, Sun Y, Que Z, et al. Anomaly detection and fault prognosis for bearings. IEEE Trans Instrum Meas, 2016, 65: 2046–2054
13 Si X S, Li T M, Zhang Q. Optimal replacement of degrading components: a control-limit policy. Sci China Inf Sci, 2021, 64: 209205
14 Si X S, Hu C H, Li T M, et al. A joint order-replacement policy for deteriorating components with reliability constraint. Sci China Inf Sci, 2021, 64: 189203
15 Verbert K, de Schutter B, Babuska R. A multiple-model reliability prediction approach for condition-based maintenance. IEEE Trans Rel, 2018, 67: 1364–1376
16 Hu C H, Pei H, Si X S, et al. A prognostic model based on DBN and diffusion process for degrading bearing. IEEE Trans Ind Electron, 2020, 67: 8767–8777
17 Gebraeel N, Lawley M, Liu R, et al. Residual life predictions from vibration-based degradation signals: a neural network approach. IEEE Trans Ind Electron, 2004, 51: 694–700
18 García-Nieto P J, García-Gonzalo E, Sánchez-Lasheras F, et al. Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. Reliability Eng Syst Saf, 2015, 138: 219–231
19 Guo L, Lei Y, Li N, et al. Machinery health indicator construction based on convolutional neural networks considering trend burr. Neurocomputing, 2018, 292: 142–150
20 Deutsch J, He D. Using deep learning-based approach to predict remaining useful life of rotating components. IEEE Trans Syst Man Cybern Syst, 2018, 48: 11–20
21 Zhang J X, Hu C H, He X, et al. A novel lifetime estimation method for two-phase degrading systems. IEEE Trans Rel, 2019, 68: 689–709
22 Wang D, Zhao Y, Yang F, et al. Nonlinear-drifted Brownian motion with multiple hidden states for remaining useful life prediction of rechargeable batteries. Mech Syst Signal Process, 2017, 93: 531–544
23 Ling M H, Ng H K T, Tsui K L. Bayesian and likelihood inferences on remaining useful life in two-phase degradation models under gamma process. Reliability Eng Syst Saf, 2019, 184: 77–85
24 Ye Z S, Chen N. The inverse Gaussian process as a degradation model. Technometrics, 2014, 56: 302–311
25 Doksum K A, Hoyland A. Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse gaussian distribution. Theor Probab Appl, 1993, 37: 137–139
26 Tseng S T, Tang J, Ku I H. Determination of burn-in parameters and residual life for highly reliable products. Naval Res Logistics, 2003, 50: 1–14
27 Tseng S T, Peng C Y. Optimal burn-in policy by using an integrated Wiener process. IIE Trans, 2004, 36: 1161–1170
28 Elwany A, Gebraeel N. Real-time estimation of mean remaining life using sensor-based degradation models. J Manufacturing Sci Eng, 2009, 131: 051005
29 Si X S, Wang W, Chen M Y, et al. A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution. Eur J Operational Res, 2013, 226: 53–66
30 Gebraeel N Z, Lawley M A, Li R, et al. Residual-life distributions from component degradation signals: a Bayesian approach. IIE Trans, 2005, 37: 543–557
31 Whitmore G A, Schenkelberg F. Modeling accelerated degradation data using Wiener diffusion with a time scale transformation. Lifetime Data Anal, 1997, 3: 27–45

32 Si X S, Wang W, Hu C H, et al. Remaining useful life estimation based on a nonlinear diffusion degradation process. IEEE Trans Rel, 2012, 61: 50–67

33 Si X S, Ren Z Q, Hu X X, et al. A novel degradation modeling and prognostic framework for closed-loop systems with degrading actuator. IEEE Trans Ind Electron, 2020, 67: 9635–9647

34 Tang S J, Guo X S, Yu C Q, et al. Real time remaining useful life prediction based on nonlinear Wiener based degradation processes with measurement errors. J Cent South Univ, 2014, 21: 4509–4517

35 Wang Z Q, Hu C H, Wang W, et al. An additive Wiener process-based prognostic model for hybrid deteriorating systems. IEEE Trans Rel, 2014, 63: 208–222

36 Zhang Z X, Hu C H, Si X S, et al. Stochastic degradation process modeling and remaining useful life estimation with flexible random-effects. J Franklin Institute, 2017, 354: 2477–2499

37 Wang Y E, Ma W M, Chow T W S, et al. A two-step parametric method for failure prediction in hard disk drives. IEEE Trans Indust Inform, 2014, 10: 419–430

38 Ebenezer R H P, Susan E, Srinivasan R, et al. Template-based gait authentication through Bayesian thresholding. IEEE/CAA J Autom Sin, 2019, 6: 209–219

39 Fang H Z, Tian N, Wang Y B, et al. Nonlinear Bayesian estimation: from Kalman filtering to a broader horizon. IEEE/CAA J Autom Sin, 2018, 5: 401–417