IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 11, 2019, accepted December 19, 2019, date of publication December 24, 2019,
date of current version January 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2961951

Remaining Useful Life Prediction Based on an Adaptive Inverse Gaussian Degradation Process With Measurement Errors

XUDAN CHEN[1], XINLI SUN[1], XIAOSHENG SI[2], AND GUODONG LI[1]
[1]College of Nuclear Engineering, Rocket Force University of Engineering, Xi'an 710025, China
[2]College of Missile Engineering, Rocket Force University of Engineering, Xi'an 710025, China

Corresponding author: Xiaosheng Si (sxs09@mails.tsinghua.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1306100, and in part by the National Nature Science Foundation of China under Grant 61922089, Grant 61773386, Grant 61673311, Grant 61573366, and Grant 61573076.

ABSTRACT Remaining useful life (RUL) prediction plays a crucial role in prognostics and health management (PHM). Recently, the adaptive model-based RUL prediction, which is proven effective and flexible, has gained considerable attention. Most research on adaptive degradation models focuses on the Wiener process. However, since the degradation process of some products is accumulated and irreversible, the inverse Gaussian (IG) process that can describe monotonic degradation paths is a natural choice for degradation modelling. This article proposes a nonlinear adaptive IG process along with the corresponding state space model considering measurement errors. Then, an improved particle filtering algorithm is presented to update the degradation parameter and estimate the underlying degradation state under the nonGaussian assumptions in the state space model. The RUL prediction depending on historical degradation data is derived based on the results of particle methods, which can avoid high-dimensional integration. In addition, the expectation-maximization (EM) algorithm combined with an improved particle smoother is developed to estimate and adaptively update the unknown model parameters once newly monitored degradation data become available. Finally, this article concludes with a simulation study and a case application to demonstrate the applicability and superiority of the proposed method.

INDEX TERMS Adaptive model, inverse Gaussian process, measurement errors, remaining useful life.

I. INTRODUCTION

With the improvement of the design levels and manufacturing techniques, products are getting more and more reliable. Accurately assessing the reliability of these highly reliable products has become an urgent need in industry. This also leads to an emerging concept called prognostics and health management (PHM) [1]. As a new engineering technology, PHM based on condition monitoring (CM) data can effectively reduce maintenance costs, improve the reliability, and mitigate system risk. The main purpose of PHM is to predict the remaining useful life (RUL) of a product accurately, and then the RUL will offer guidance for sequential management involving inspection schedule, maintenance, replacement, and spare parts ordering [2], [3]. The RUL prediction is typically characterized by the probability distribution function (PDF), which incorporates the expectation and the uncertainty of the predicted RUL simultaneously [4].

In order to predict the RUL accurately, it is necessary to construct an appropriate and applicable reliability model. Since stochastic processes have good properties in modelling the randomness of the degradation process of highly reliable products, the stochastic process-based degradation models have become a research focus [5]. The Wiener process, gamma process, and inverse Gaussian (IG) process are three popular stochastic processes used in degradation modelling. The variants of these stochastic degradation models that consider covariates, random effects, and measurement errors can be found in [6]–[12]. In recent years, the Ornstein-Uhlenbeck (OU) process has been introduced to describe the nonmonotonic degradation path by researchers [13], [14]. However, the OU process is still in the exploratory stage in degradation modelling without a widely used form [15].

The Wiener process that can describe the nonmonotonic degradation path has been well studied and gained

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Liu.

wide applications. Nevertheless, the conventional Wiener process model is based on a Markov assumption that the degradation evolution and the RUL prediction depend only on the current degradation observation, rather than the historical degradation data [4]. The RUL prediction based on incomplete degradation information is inaccurate. To overcome this limitation, Wang *et al.* [16] presented an adaptive Brownian motion-based model with an adaptive drift coefficient and explored to use historical degradation data for predicting RUL. In their model, the adaptation of the drift coefficient was performed by the Kalman filter (KF). However, the distribution of the drift coefficient was not incorporated into the RUL estimation. Motivated by this, Si *et al.* [17] developed a Wiener process-based degradation model with a recursive filtering algorithm. Also, they revealed that considering such distribution could reduce the uncertainty of the estimated RUL. Subsequently, this linear degradation model was extended to a nonlinear prognostic model by [1]. More research on nonlinear adaptive models can be found in [18]–[20]. In fact, the models presented in above literatures are all nonstandard state space models, which result in the inappropriateness for a KF or a strong tracking filter (STF) to estimate the degradation drift directly. Huang *et al.* [4] proposed an online filtering algorithm based on Bayes' theorem for the nonstandard state space model. Similar online filtering algorithm was investigated by [21].

Furthermore, Wang and Tsui [22] pointed that existing state space models used for RUL prediction assumed that the drift coefficient was constant during the given sampling interval until the next observation was available, which caused a contradiction with the drift coefficient evolution. To alleviate this assumption, they constructed a new space model and explained the main reasons why the new model could provided high RUL prediction accuracies. More details can be found in [23]. Zhai and Ye [24] had similar views and considered that the existing studies used an autoregressive model of order 1 for the adaptive drift. Accordingly, they introduced a new adaptive Wiener process model that modelled the adaptive drift by a continuous Brownian motion, and presented an analytical parameter estimation method without using the filtering algorithm. In addition, the error term was introduced in the adaptive models proposed by [25], [26], where the KF could be employed directly due to the introduction of two or more hidden states in the state space model. However, these two models still assumed that the drift coefficient was constant between adjacent measurements.

Since the degradation process is an accumulated and irreversible process for some products, the gamma process and IG process that can describe monotonic degradation paths become a more natural choice for degradation modelling [15], [27]. Compared with the Wiener process and gamma process, the IG process shows better goodness of fit for some degradation data [9], [10]. However, most IG process-based degradation models do not consider the error term, which limits the application of the IG process in degradation analysis, and regard the lifetime distribution as a reliability index, rather than the RUL based on historical degradation data [28], [29]. Although Peng *et al.* [30] derived the reliability functions depending on observed data under a general Bayesian framework, the corresponding degradation models did not incorporate measurement errors. Xu and Wang [31] developed a linear adaptive IG process to characterize the degradation process of condition monitored components, and to employ a general Bayesian filtering process, they assumed that the sampling interval was fixed. However, their model also did not incorporate measurement errors and the derived RUL did not make full use of historical degradation data.

From the above review of the related work, it can be found that there are three problems remaining to be addressed when utilizing the IG process for RUL prediction. The first is to construct an appropriate adaptive IG process model such that the historical degradation data can be incorporated into the PDF of the RUL with lower uncertainty. The second is to consider the error term in the degradation model to extend the application of IG process. The third is to solve the filtering problem when considering two hidden states in the state space model, and establish a parameter estimation procedure updated with the newly observed data. In this article, we propose a nonlinear IG process model to predict the RUL adaptively. To consider the measurement errors in degradation analysis and incorporate historical degradation data for RUL prediction, a state space model with two hidden states is constructed. Since the conventional KF cannot be employed to update the degradation parameter and estimate the underlying degradation state under nonGaussian assumptions, an improved particle filtering algorithm is presented. Based on the filtering results, we derive the particle representations of the RUL without using high-dimensional integration. In addition, the expectation-maximization (EM) algorithm combined with particle methods is developed to estimate and update the model parameters adaptively.

The remainder of this article is organized as follows. Section II presents an adaptive IG process model considering measurement errors. In Section III, we propose an improved filtering algorithm and derive the RUL grounded in particle representations. In Section IV, the parameter estimation and updating approach based on the EM algorithm is investigated. A simulation study and a case application are provided to demonstrate the performance of the proposed model in Section V. Section VI concludes this article.

II. ADAPTIVE IG PROCESS

A. BASIC IG PROCESS MODEL

Let $\{X(t), t \geq 0\}$ denote the underlying degradation process and follow an IG process with increment $X(t)-X(t-\Delta t) \sim \mathcal{IG}(\mu\Delta\Lambda(t), \eta(\Delta\Lambda(t))^2)$, $t > 0$, where $X(0) = 0$ and $\Delta\Lambda(t) = \Lambda(t) - \Lambda(t-\Delta t)$ with $\Lambda(0) = 0$. Here, $\Lambda(t|\boldsymbol{\theta}_\Lambda)$ is a given, monotone increasing shape function of time t

with unknown parameters $\boldsymbol{\theta}_\Lambda$. For notation simplicity, it is rewritten as $\Lambda(t)$. The IG density function is defined by

$$f(x|\mu,\eta,\boldsymbol{\theta}_\Lambda) = \sqrt{\frac{\eta\Lambda^2(t)}{2\pi x^3}}\exp\left[-\frac{\eta}{2x}\left(\frac{x}{\mu}-\Lambda(t)\right)^2\right], \quad x,\mu,\eta>0. \tag{1}$$

Then the IG process has mean $\mu\Lambda(t)$ and variance $\mu^3\Lambda(t)/\eta$. The parameter μ represents the degradation rate, and the parameter η has no direct physical meaning. The shape function $\Lambda(t)$ is usually decided by the particular failure-generating mechanism [10].

Let D denote the threshold level for the degradation path. Under the definition of the first passage time, the product's lifetime T is expressed as $T = \inf\{t|X(t)\geq D\}$. According to the monotonicity property of the IG process, the cumulative distribution function (CDF) of T is given by

$$\begin{aligned} F_T(t|\mu,\eta,\boldsymbol{\theta}_\Lambda) &= \Phi\left[-\sqrt{\frac{\eta}{D}}\left(\frac{D}{\mu}-\Lambda(t)\right)\right] \\ &\quad - \exp\left(\frac{2\eta\Lambda(t)}{\mu}\right)\Phi\left[-\sqrt{\frac{\eta}{D}}\left(\frac{D}{\mu}+\Lambda(t)\right)\right], \end{aligned} \tag{2}$$

where $\Phi(\cdot)$ is the standard normal CDF.

B. STATE SPACE MODEL WITH MEASUREMENT ERRORS

In practice, the CM data contain observation noise inevitably. Thus, the observed degradation process $\{Y(t), t\geq 0\}$ is given by

$$Y(t) = X(t) + \varepsilon, \tag{3}$$

where $Y(0)=0$ and ε denotes the measurement error, which is assumed to be s-independent and identically distributed with $\varepsilon\sim\mathcal{N}(0,\sigma_\varepsilon^2)$ at any time point t [6].

For a product in service, it is assumed that the degradation process is monitored at $t_1, t_2, \ldots, t_m$ with degradation observations $y_1, y_2, \ldots, y_m$, where m is the number of measurements. The corresponding underlying degradation state is expressed as $x_1, x_2, \ldots, x_m$. For simplicity, we further assume that $y_{1:m} = (y_1, y_2, \ldots, y_m)'$ and $x_{1:m} = (x_1, x_2, \ldots, x_m)'$. The conventional PDF of the RUL uses only the current degradation observation, rather than the historical data. In Wiener process-based degradation analysis, researcher investigated the RUL based on historical data by establishing a state space model. Si *et al.* [17] revealed that utilizing the historical data could reduce the uncertainty of RUL prediction. However, the lifetime distribution or the RUL of existing IG process models rarely incorporate historical data. To address this problem, we consider an updating process for degradation parameter μ by assuming $\xi_j = \xi_{j-1}$ with initial parameter ξ_0 following a truncated normal distribution $\mathcal{TN}(a_0,\sigma_0^2)$ to avoid negative values of ξ, where $\xi = 1/\mu$. Then, this IG process model with an adaptive degradation rate can be discretely expressed as a state space model, i.e.,

$$\begin{cases} \xi_j = \xi_{j-1} \\ x_j = x_{j-1} + \upsilon_j\,, \\ y_j = x_j + \varepsilon_j \end{cases} \tag{4}$$

where $\upsilon_j \sim \mathcal{IG}(\Delta\Lambda_j/\xi_{j-1}, \eta(\Delta\Lambda_j)^2)$, $\Delta\Lambda_j = \Lambda(t_j) - \Lambda(t_{j-1})$ with $\Lambda(t_0)=0$, and $\varepsilon_j\sim\mathcal{N}(0,\sigma_\varepsilon^2)$.

Based on (4) and particle methods, we can incorporate the historical degradation data $y_{1:m}$ into the PDF of the RUL. The details on particle methods are given in the following. In particular, if $\sigma_\varepsilon\to 0$, the model characterized by (4) reduces to the model without considering measurement errors, where $\varepsilon_j\sim\mathcal{N}(0,0)$ [24]. This situation can be viewed as a special case of (4).

C. MODEL COMPARISON AND DISCUSSION

In this section, two adaptive Wiener process models will be reviewed to illustrate the starting points of the proposed model constructed in (4). The first model is a nonlinear adaptive Wiener process model proposed by Wang *et al.* [26]. They assumed that there are multiple hidden parameters. Here, we only provide the state space model with one hidden parameter as follows

$$\begin{cases} \lambda_j = \lambda_{j-1} + \alpha \\ x_j = x_{j-1} + \lambda_{j-1}\Delta\Lambda_j + \sigma\Delta\mathcal{B}_j\,, \\ y_j = x_j + \varepsilon_j \end{cases} \tag{5}$$

where the drift coefficient λ is characterized by an updating procedure with $\alpha\sim\mathcal{N}(0,Q)$, $\Lambda(t)$ is the shape function with $\Delta\Lambda_j = \Lambda(t_j) - \Lambda(t_{j-1})$, $\mathcal{B}(t)$ is a standard Brownian motion with $\Delta\mathcal{B}_j = \mathcal{B}(t_j) - \mathcal{B}(t_{j-1})$, and σ is the volatility parameter. This model, considering measurement errors, extends the application of existing Wiener models, but adopts the assumption that the drift coefficient is constant between adjacent measurements, which can be viewed as an approximation of the time-varying drift.

The second adaptive model is proposed by Zhai and Ye [24] as follows

$$\begin{cases} \lambda_j = \lambda_{j-1} + k\Delta W_j \\ x_j = x_{j-1} + \lambda_{j-1}\Delta S_j \\ \quad + k\displaystyle\int_{t_{j-1}}^{t_j}\big(W(\tau) - W(t_{j-1})\big)dS(\tau) + \sigma\Delta\mathcal{B}_j \end{cases}, \tag{6}$$

where λ is an adaptive drift modelled by a Wiener process, $W(t)$ is a standard Brownian motion independent of $\mathcal{B}(t)$ with $\Delta W_j = W(t_j) - W(t_{j-1})$, and $S(t)$ is a monotonically increasing function of t with $\Delta S_j = S(t_j) - S(t_{j-1})$. The continuous form of the observation equation in (6) can be expressed as $X(t) = \int_0^t\lambda(\tau)dS(\tau) + \sigma\mathcal{B}(t)$. This means that the drift parameter is time-varying, and thus the contradiction in Wang's model [26] is overcome. However, the error term is not incorporated into this model, the applicability of which is worth further study.

The adaptive model established in this article can not only consider the influence of measurement errors, but also introduce the historical degradation data into the RUL prediction (see next section). Therefore, the model form is closer to the Wang's model, and at the same time, it overcomes the contradiction of the assumption in Wang's model. It is worth noting that in this article, the updating process for ξ is expressed as $\xi_j = \xi_{j-1}$, mainly for the following two reasons. First, the main purpose of establishing an adaptive model is to update the posterior distribution of the degradation parameter (μ or λ) and then incorporate the historical degradation data into RUL prediction. The updating state equation described in (4) can achieve the same effect and simplify the computation of statistical inferences simultaneously. In fact, we have also established updating equation $\xi_j = \xi_{j-1} + \alpha$ for ξ, which is similar to those in Wang's model and Zhai's model. However, the corresponding parameter estimation methods we built showed poor performance since they were sensitive to initial values. Second, we assume that the volatility of the degradation parameter is relatively small. The results in [24], [26] reveal that the volatility of the drift coefficient in both case applications is quite small, i.e., the estimated values of Q and k close to 0. Accordingly, the updating state equations in (5) and (6) actually reduce to the updating state equation proposed in (4). More case applications can be found in [1], [4], [16]. In addition, Peng *et al.* [32] presented a quantitative description of the approximated slope or steepness of the degradation path, named the degradation rate function. When the degradation slope changes dramatically, it is more appropriate to adopt this concept to investigate the evolving path of the degradation parameter. This is beyond the scope of this article, and thus more discussions are omitted here.

In many cases, the degradation process of products is accumulated and irreversible, such as fatigue crack growth and corrosion. The advantages of the IG process in modelling monotonic degradation paths have been well studied by Ye and Chen [9], Peng [10], and Chen *et al.* [15] respectively. This article aims to further study the IG process-based adaptive model, which will extend the application of the IG process.

III. RUL PREDICTION

A. BASIC RUL

This article focuses on the degradation of an individual product and uses the concept of the first passage time [33] to define RUL. Given $X(t_j) = x_j$, the RUL at time t_j is defined by

$$L_j = \inf\left\{l_j : X(t_j + l_j) \geq D | X(t_j) = x_j, \xi_j, \eta, \boldsymbol{\theta}_\Lambda\right\}. \quad (7)$$

According to the independent increment property, i.e., $X(t_j + l_j) - X(t_j) \sim \mathcal{IG}(\Delta\Lambda(l_j)/\xi_j, \eta(\Delta\Lambda(l_j))^2)$, the RUL L_j is equal to the first passage time of the process $\{Z(l_j), l_j \geq 0\}$ crossing threshold $\tilde{D} = D - x_j$, where $\Delta\Lambda(l_j) = \Lambda(t_j + l_j) - \Lambda(t_j)$, and $Z(l_j) = X(t_j + l_j) - X(t_j)$ with $Z(0) = 0$. The CDF of L_j is given by

$$\begin{aligned} &F_{L_j|x_j,\xi_j,\eta,\boldsymbol{\theta}_\Lambda}(l_j|x_j, \xi_j, \eta, \boldsymbol{\theta}_\Lambda) \\ &= \Phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\xi_j - \Delta\Lambda(l_j)\right)\right] \\ &\quad - \exp\left(2\eta\xi_j\Delta\Lambda(l_j)\right)\Phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\xi_j + \Delta\Lambda(l_j)\right)\right]. \quad (8) \end{aligned}$$

When $\Lambda(t)$ is differentiable, the PDF of L_j can be easily obtained as follows

$$\begin{aligned} &f_{L_j|x_j,\xi_j,\eta,\boldsymbol{\theta}_\Lambda}(l_j|x_j, \xi_j, \eta, \boldsymbol{\theta}_\Lambda) \\ &= 2\Lambda'(t_j + l_j)\sqrt{\frac{\eta}{\tilde{D}}}\phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\xi_j - \Delta\Lambda(l_j)\right)\right] \\ &\quad - 2\eta\xi_j\Lambda'(t_j + l_j)\exp\left(2\eta\xi_j\Delta\Lambda(l_j)\right) \\ &\quad \cdot \Phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\xi_j + \Delta\Lambda(l_j)\right)\right], \quad (9) \end{aligned}$$

where $\Lambda'(t_j + l_j) = d\Lambda(t_j + l_j)/dl_j$, and $\phi(\cdot)$ is the standard normal PDF.

B. IMPROVED PARTICLE FILTERING ALGORITHM

To incorporate historical degradation data, Wang *et al.* [26], and Zhai and Ye [24] respectively constructed the state space model with an adaptive drift and employed the KF to update the hidden state. However, the KF cannot be utilized in the proposed model due to its nonGaussian assumptions. Consequently, this article proposes an improved particle filter to address this problem. The filtering results can be used to calculated the PDF of the RUL in the following section. See [11], [34], [35] for more applications of particle filters in degradation analysis.

The particle filter is a sequential Monte Carlo method grounded in particle representations and can be viewed as the generalization of the well-known KF [36]. The main idea is to represent the PDF of x approximatively by using M particles and associated weights, i.e., $\{\mathrm{f}^{(k)}, \mathrm{w}^{(k)} | k = 1, \ldots, M\}$, where $\mathrm{w}^{(k)}$ denotes the weight of the kth particle $\mathrm{f}^{(k)}$. Since there are two hidden states in the proposed state space model, the conventional particle filter needs to be improved. The following is the improved particle filtering algorithm based on the sequential importance resampling (SIR) filter.

First, we assume symbols: (1) $(\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)})$ denotes the kth pair of particle filters of (x_j, ξ_j) at time t_j, i.e., $(\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)}) \sim f(x_j, \xi_j | y_{1:j})$; (2) $(\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)})$ denotes the kth pair of particle predictors of (x_j, ξ_j) at time t_j, i.e., $(\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) \sim f(x_j, \xi_j | y_{1:u})$, $(j > u)$. Given model parameters, the algorithm is summarized as follows.

Algorithm 1 (Improved SIR Filtering Algorithm):

Step 1. Generate $\mathrm{f}_0^{(k)} = 0$ and $\tilde{\mathrm{f}}_0^{(k)} \sim \mathcal{TN}\left(a_0, \sigma_0^2\right)$, $k = 1, \ldots, M$.

Step 2. For $j = 1, \ldots, m$,

- Generate a random number $\upsilon_j^{(k)} \sim \mathcal{IG}(\Delta\Lambda_j/\tilde{\mathrm{f}}_{j-1}^{(k)}, \eta(\Delta\Lambda_j)^2)$, $k = 1, \ldots, M$.

- Compute $\mathrm{p}_j^{(k)} = \mathrm{f}_{j-1}^{(k)} + \upsilon_j^{(k)}$ and $\tilde{\mathrm{p}}_j^{(k)} = \tilde{\mathrm{f}}_{j-1}^{(k)}$, $k = 1, \ldots, M$.
- Compute $\mathrm{w}_j^{(k)} = f(y_j|\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) \propto \exp\left[-(y_j - \mathrm{p}_j^{(k)})^2/2\sigma_\varepsilon^2\right]$, $k = 1, \ldots, M$.
- Generate $\{\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)}|k = 1, \ldots, M\}$ by resampling $\{\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}|k = 1, \ldots, M\}$ with weights $\{\mathrm{w}_j^{(k)}|k = 1, \ldots, M\}$.

At the end of this filtering algorithm, M pairs of samples from $f(x_j, \xi_j|y_{1:j})$ for $j = 0, \ldots, m$ are obtained. The proof of *Algorithm 1* is given in the Appendix A.

C. RUL BASED ON THE HISTORICAL DATA

Based on the results of the improved particle filter, we can obtain the particle representations of $f(x_j, \xi_j|y_{1:j})$. Then the PDF of the RUL depending on the historical degradation data, $f_{L_j|y_{1:j}}(l_j|y_{1:j})$, can be derived as follows

$$f_{L_j|y_{1:j}}(l_j|y_{1:j}) = \int_0^{+\infty}\int_0^{+\infty} f_{L_j|x_j,\xi_j}(l_j|x_j, \xi_j) f(\xi_j|x_j, y_{1:j}) \cdot f(x_j|y_{1:j}) d\xi_j dx_j. \tag{10}$$

For notation simplicity, $f_{L_j|x_j,\xi_j,\eta,\boldsymbol{\theta}_\Lambda}(l_j|x_j, \xi_j, \eta, \boldsymbol{\theta}_\Lambda)$ is replaced by $f_{L_j|x_j,\xi_j}(l_j|x_j, \xi_j)$ since some parameters are not random in (10). Once new degradation data are observed, the RUL based on the historical degradation data can be updated adaptively.

Assume that

$$f_{L_j|x_j,y_{1:j}}(l_j|x_j, y_{1:j}) = \int_0^{+\infty} f_{L_j|x_j,\xi_j}(l_j|x_j, \xi_j) f(\xi_j|x_j, y_{1:j}) d\xi_j, \tag{11}$$

and then the particle representations of $f_{L_j|y_{1:j}}(l_j|y_{1:j})$ can be expressed as

$$f_{L_j|y_{1:j}}(l_j|y_{1:j}) = \int_0^{+\infty} f_{L_j|x_j,y_{1:j}}(l_j|x_j, y_{1:j}) f(x_j|y_{1:j}) dx_j \approx \frac{1}{M}\sum_{k=1}^{M} f_{L_j|x_j,y_{1:j}}(l_j|\mathrm{f}_j^{(k)}, y_{1:j}). \tag{12}$$

From (12), it is easy to obtain numerical results, which avoids the high-dimensional integration in (10). It is worth noting that the key for estimating RUL is to derive the analytical expression of (11). The detailed expression is given in the Appendix B.

IV. PARAMETER ESTIMATION

Maximum likelihood estimation (MLE) is one of the most popular methods for parameter estimation [4]. The state space model proposed in this article considers two hidden states, i.e., the underlying degradation state and the adaptive degradation rate. Thus, it is difficult to directly employ MLE to estimate unknown parameters in the proposed model. The EM algorithm offers an alternative framework for parameter estimation in the model with hidden states. See [7], [10] for details. In the following, we present the parameter updating algorithm.

A. EM ALGORITHM

Let $\boldsymbol{\theta} = (\boldsymbol{\theta}_\Lambda, \eta, a_0, \sigma_0, \sigma_\varepsilon)'$ and $\xi_{0:m} = (\xi_0, \xi_1, \ldots, \xi_m)'$. Given CM data $y_{1:m}$, the complete log-likelihood about $y_{1:m}$, $x_{0:m}$, and $\xi_{0:m}$ can be formulated as

$$\mathcal{L}(y_{1:m}, x_{0:m}, \xi_{0:m}|\boldsymbol{\theta}) = \ln f(y_{1:m}|x_{0:m}, \boldsymbol{\theta}) + \ln f(x_{0:m}, \xi_{0:m}|\boldsymbol{\theta}). \tag{13}$$

The estimation procedure of the EM algorithm is given as follows.

- E-step: Calculate

$$\ell(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}_m^{(i)}) = E_{x_{0:m},\xi_{0:m}|y_{1:m},\hat{\boldsymbol{\theta}}_m^{(i)}}\left[\mathcal{L}(y_{1:m}, x_{0:m}, \xi_{0:m}|\boldsymbol{\theta})\right], \tag{14}$$

  where, $\hat{\boldsymbol{\theta}}_m^{(i)}$ denotes the estimate of $\boldsymbol{\theta}$ at the ith iterative step of the EM algorithm based on data $y_{1:m}$.
- M-step: Calculate

$$\hat{\boldsymbol{\theta}}_m^{(i+1)} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}_m^{(i)}). \tag{15}$$

Through iteratively calculating E-step and M-step, the parameter estimates converge to true values.

To be more precise, the expectation of the complete log-likelihood is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}_m^{(i)}) \propto & -\frac{m}{2}\ln\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{m}\left(y_j^2 - 2y_jE(x_j) + E(x_j^2)\right) \\ & + \frac{m}{2}\ln\eta + \sum_{j=1}^{m}\ln\Delta\Lambda_j - \frac{\eta}{2}\sum_{j=1}^{m}(E(\Delta x_j\xi_{j-1}^2) \\ & - 2\Delta\Lambda_jE(\xi_{j-1}) + \Delta\Lambda_j^2E(1/\Delta x_j)) \\ & - (m+1)\left[\frac{1}{2}\ln\sigma_0^2 + \ln\left(1 - \Phi\left(-\frac{a_0}{\sigma_0}\right)\right)\right] \\ & - \frac{1}{2\sigma_0^2}\sum_{j=0}^{m}(E(\xi_j^2) - 2a_0E(\xi_j) + a_0^2),\end{aligned} \tag{16}$$

where $\Delta x_j = x_j - x_{j-1}$ with $x_0 = 0$.

After deriving $\ell(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}_m^{(i)})$, we can obtain the $\hat{\boldsymbol{\theta}}_m^{(i+1)}$ by maximizing (16) with respect to $\boldsymbol{\theta}$. Since σ_ε in the second line of (16) is independent of other unknown parameters, σ_ε has the analytical results as follows

$$\hat{\sigma}_\varepsilon = \sqrt{\frac{\sum_{j=1}^{m}\left(y_j^2 - 2y_jE(x_j) + E(x_j^2)\right)}{m}}. \tag{17}$$

The remaining parameters can be obtained by using *fminsearch* function in MATLAB.

B. IMPROVED PARTICLE SMOOTHING ALGORITHM

It is worth noting that the expectations, i.e., $E(\cdot)$, in (16) and (17) are unknown. In this article, we use the results of particle smoother to calculate these expectations. Godsill *et al.* [37] suggested a particle smoother using backwards simulation that concerned the whole trajectory of states $(x_1, x_2 \ldots, x_m)'$. The state space model proposed in this article considers two hidden states, and thus it is necessary to improve this particle smoother.

We assume that $(\mathrm{s}_j^{(k)}, \tilde{\mathrm{s}}_j^{(k)})$ denotes the kth pair of particle smoothers of (x_j, ξ_j) at time t_j, i.e., $(\mathrm{s}_j^{(k)}, \tilde{\mathrm{s}}_j^{(k)}) \sim f(x_j, \xi_j|y_{1:v})$, $(j < v)$. Based on the filtering results in *Algorithm 1*, the improved particle smoothing algorithm is summarized in the following. The proof is given in the Appendix C.

Algorithm 2 (Improved Particle Smoothing Algorithm):

Step 1. Choose $[\mathrm{s}_m^{(k)}, \tilde{\mathrm{s}}_m^{(k)}] = [\mathrm{f}_m^{(l)}, \tilde{\mathrm{f}}_m^{(l)}]$ with probability $1/M$.

Step 2. For $j = m-1, \ldots, 1$

- Calculate

$$\mathrm{w}_{j|j+1}^{(l)} \propto f(\mathrm{s}_{j+1}^{(k)}|\mathrm{f}_j^{(l)}, \tilde{\mathrm{f}}_j^{(l)}) \propto \sqrt{\frac{(\Delta\Lambda_{j+1})^2}{(\mathrm{s}_{j+1}^{(k)} - \mathrm{f}_j^{(l)})^3}} \cdot \exp\left\{-\frac{\eta\left[\left(\mathrm{s}_{j+1}^{(k)} - \mathrm{f}_j^{(l)}\right)\tilde{\mathrm{f}}_j^{(l)} - \Delta\Lambda_{j+1}\right]^2}{2(\mathrm{s}_{j+1}^{(k)} - \mathrm{f}_j^{(l)})}\right\},$$

  $l = 1, \ldots, M$.
- Choose $[\mathrm{s}_j^{(k)}, \tilde{\mathrm{s}}_j^{(k)}] = [\mathrm{f}_j^{(l)}, \tilde{\mathrm{f}}_j^{(l)}]$ with probability $\mathrm{w}_{j|j+1}^{(l)}$.

Step 3. $\mathrm{s}_{1:m}^{(k)} = (\mathrm{s}_1^{(k)}, \ldots, \mathrm{s}_m^{(k)})'$ and $\tilde{\mathrm{s}}_{1:m}^{(k)} = (\tilde{\mathrm{s}}_1^{(k)}, \ldots, \tilde{\mathrm{s}}_m^{(k)})'$ are the random samples from $f(x_{1:m}, \xi_{1:m}|y_{1:m})$.

Step 4. Repeat Step 1-3 for $k = 1, \ldots, M$.

According to the Appendix B, we assume that $E(\xi_j) = \varphi_1(x_m)$ and $E(\xi_j^2) = \varphi_2(x_m)$. The partial expectations can be calculated as follows

$$\begin{cases} E(x_j^2) \approx \dfrac{1}{M}\displaystyle\sum_{k=1}^{M}\left(\mathrm{s}_j^{(k)}\right)^2 \\ E(1/\Delta x_j) \approx \dfrac{1}{M}\displaystyle\sum_{k=1}^{M}\dfrac{1}{(\mathrm{s}_j^{(k)} - \mathrm{s}_{j-1}^{(k)})} \\ E(\Delta x_j \xi_{j-1}^2) \approx \dfrac{1}{M}\displaystyle\sum_{k=1}^{M}(\mathrm{s}_j^{(k)} - \mathrm{s}_{j-1}^{(k)})\varphi_2(\mathrm{s}_m^{(k)}). \end{cases} \tag{18}$$

C. CONVERGENCE ASSESSMENT

In the EM algorithm, the complete likelihood increases at each iteration and tends to a certain value, which is commonly used to assess the convergence of the EM algorithm. However, the proposed EM algorithm based on particle methods does not guarantee the strict monotone likelihood property. In this article, the relative likelihood, which is the ratio of the likelihoods at two adjacent iterations [36], is considered as a measure in assessing convergence. The relative likelihood function of the proposed model at the ith iteration is given by

$$\ln\left[\frac{f_{\boldsymbol{\theta}^{(i)}}(y_{1:m})}{f_{\boldsymbol{\theta}^{(i-1)}}(y_{1:m})}\right] = \ln E_{\boldsymbol{\theta}^{(i-1)}}\left[\frac{f_{\boldsymbol{\theta}^{(i)}}(x_{1:m}, y_{1:m})}{f_{\boldsymbol{\theta}^{(i-1)}}(x_{1:m}, y_{1:m})}\middle| y_{1:m}\right] \approx -\ln\left[\frac{1}{M}\sum_{k=1}^{M}\frac{f_{\boldsymbol{\theta}^{(i-1)}}(\mathrm{s}_{1:m}^{(k)}, y_{1:m})}{f_{\boldsymbol{\theta}^{(i)}}(\mathrm{s}_{1:m}^{(k)}, y_{1:m})}\right], \tag{19}$$

where $\mathrm{s}_{1:m}^{(k)}$ is the kth smoothing particle during the ith EM iteration, and $f_{\boldsymbol{\theta}^{(i)}}(x_{1:m}, y_{1:m})$ is the complete likelihood function about $x_{1:m}$ and $y_{1:m}$. See *Supplementary Materials* in [9] for further details of $f_{\boldsymbol{\theta}^{(i)}}(x_{1:m}, y_{1:m})$.

Generally, the EM algorithm is converged if the relative likelihood function is less than a predetermined tolerance. Theoretically, the relative likelihood function will converge to any predetermined tolerance eventually if the particle size M and the number of iterations are large enough. Practically, overlarge particle size and number of iterations will affect the computational efficiency. To strike a balance between the efficiency and accuracy, we provide a three-stage EM algorithm as follows:

- *Stage 1:* 200 particles with no tolerance or a very small tolerance are used for initial estimation. At this stage, the relative likelihood function decreases fast and the parameter estimates converge to reasonable small ranges.
- *Stage 2:* 500 particles with tolerance 0.001 are used. The parameter estimates tend to be stable within small number of iterations at this stage.
- *Stage 3:* 1000 particles with tolerance 0.001 are used. This stage further improves the calculation accuracy.

In fact, the accuracy of parameter estimates can be further improved by increasing the particle size and decreasing the tolerance. However, the proposed three-stage EM algorithm is able to balance the efficiency and accuracy.

D. PARAMETER UPDATING ALGORITHM

Once newly monitored degradation data become available, the model parameters and the RUL based on historical data can be adaptively estimated. Let $\hat{\boldsymbol{\theta}}_j^{(i)}$ denote the estimate of $\boldsymbol{\theta}$ at the ith iterative step of the EM algorithm based on CM data $y_{1:j}$. Given initial monitored data $y_{1:m_0}$, the parameter updating algorithm is summarized in *Algorithm 3*.

Algorithm 3 (Parameter Updating Algorithm):

Step 1. Estimate a set of reasonable initial parameters, $\hat{\boldsymbol{\theta}}_{m_0}^{(0)}$.

Step 2. For CM data $y_{1:j}$, employ the three-stage EM algorithm. Each stage is performed as follows:

- E-step: Substitute $\hat{\boldsymbol{\theta}}_j^{(i-1)}$ into (14) to obtain $\ell(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}_j^{(i-1)})$.
- M-step: Maximize $\ell(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}_j^{(i-1)})$ with respect to $\boldsymbol{\theta}$ to obtain new estimate $\hat{\boldsymbol{\theta}}_j^{(i)}$.
- Convergence assessment: Compare the relative likelihood with the predetermined tolerance, and decide whether to move to the next EM stage.

At the end of the third EM stage, regard the estimated $\hat{\boldsymbol{\theta}}_j^{(i_{\max})}$ as the parameter estimates with respect to $y_{1:j}$.

Step 3. Once observe new data, assume $\hat{\boldsymbol{\theta}}_{j+1}^{(0)} = \hat{\boldsymbol{\theta}}_j^{(i_{\max})}$ and repeat Step 2.

V. ILLUSTRATIVE EXAMPLES

In this section, a simulation study and a case application are provided to demonstrate the applicability and superiority of the proposed model. To compare the performance of the proposed model with two existing models, we introduce the mean square error (MSE) [1]. The MSE about the actual RUL at t_j is defined by

$$\mathrm{MSE}_j = \int_0^{\infty} \left(l_j - \tilde{l}_j\right)^2 f_{L_j|y_{1:j}}(l_j|y_{1:j}) dl_j, \tag{20}$$

where $\tilde{l}_j$ is the actual RUL obtained at time point t_j, and $f_{L_j|y_{1:j}}(l_j|y_{1:j})$ is the PDF of the RUL based on the historical degradation data.

A. SIMULATION STUDY

This simulation study is provided to demonstrate the applicability of the proposed model along with the parameter estimation algorithm. For illustration purpose, we assume that $\Lambda(t) = t^q$ and set $q = 1.7$, $\eta = 0.3$, $a_0 = 6$, $\sigma_0 = 0.1$, and $\sigma_\varepsilon = 0.2$. Then, the proposed model in (4) is used to generate a simulated degradation path, where $m = 150$ and sampling interval $\Delta t = 0.1$, as shown in Fig. 1. Here, the underlying degradation state x_{150} is set as the threshold, i.e., $D = 14.45$.

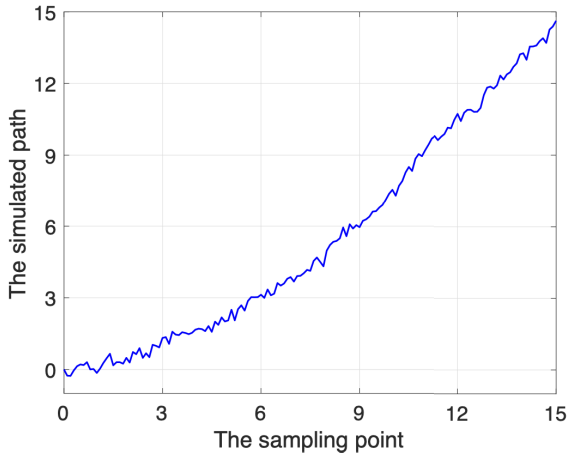


FIGURE 1. The simulated degradation path.

For the simulated degradation path, we apply the proposed model to estimate model parameters based on data $y_{1:j}$, where $j \geq 39$ such that there are sufficient degradation data for initial parameter estimation [24]. The initial parameters are set as $\hat{\boldsymbol{\theta}}_{39}^{(0)} = (1.77, 1.00, 6.54, 0.50, 0.50)$, where the estimates of q and a_0 are obtained by fitting data $y_{1:39}$ against nonlinear function t^q/a_0, and the remaining estimates are set as reasonable values. As the degradation data are accumulated, the model parameters are estimated adaptively. The evolving paths of parameter estimates are illustrated in Fig. 2. From Fig. 2, the parameter estimates are updated and tend to be stable, where the estimates of q and a_0 approach to the true values. For comparison, we also provide the evolving paths of parameter estimates based on initial parameters $\hat{\boldsymbol{\theta}}_{39}^{(02)} = (1.77, 5.00, 6.54, 5.00, 5.00)$ in Fig. 2, where q and a_0 are calculated by the same fitting method, and the remaining initial parameters are chosen as ten times more than true values. As shown, the evolving paths based on both sets of initial parameters are almost identical, except for η at the initial sampling points. These results verify the applicability of the proposed parameter estimation method. In the following, the results based on $\hat{\boldsymbol{\theta}}_{39}^{(0)}$ are used for further analysis.

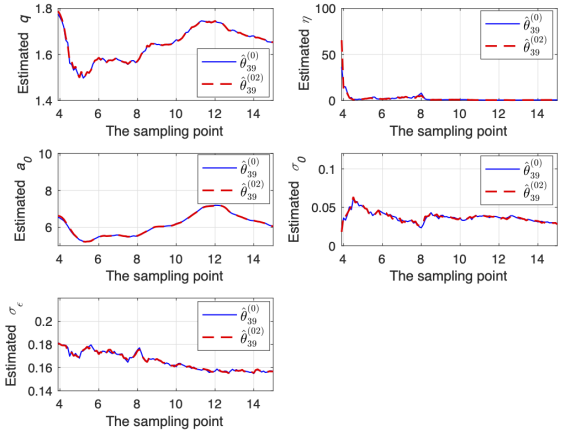


FIGURE 2. The evolving paths of parameter estimates.

Based on the updated model parameters, we can obtain the PDFs of the RUL at different sampling points. To compare the performance of predicting RUL, we also consider Wang's model in (5) and Zhai's model in (6). For simplicity, let M_1 denote the proposed model, M_W denote Wang's model considering measurement errors, and M_Z denote Zhai's model without considering measurement errors. Then the PDFs of the RUL of three models are plotted in Fig. 3 from the 85th sampling point to the 145th point in increments of 10 points.

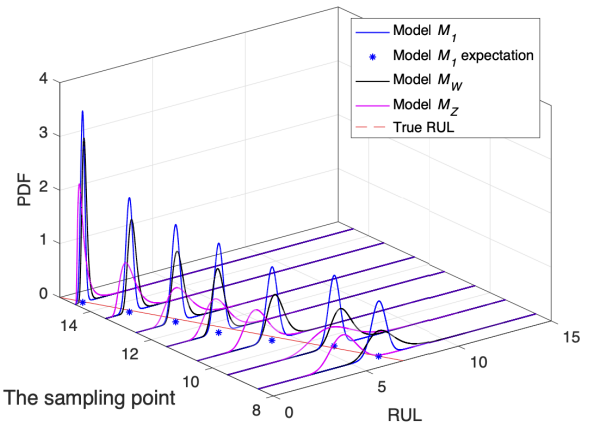


FIGURE 3. Predicted RUL based on M_1, M_W, and M_Z.

TABLE 1. MRUL, 95% CIs, and REs of the three models.

	80th sampling point			110th sampling point			140th sampling point		
	MRUL	95% CI	RE	MRUL	95% CI	RE	MRUL	95% CI	RE
M_1	8.50	[8.49,8.51]	19.76%	3.32	[3.30,3.34]	17.05%	0.76	[0.75,0.77]	15.89%
M_W	9.06	[9.01,9.10]	27.59%	3.12	[3.09,3.14]	22.10%	0.75	[0.74 0.76]	16.89%
M_Z	10.48	[10.34,10.60]	47.55%	2.76	[2.72,2.79]	31.13%	0.71	[0.68,0.73]	21.44%

It can be observed from Fig. 3 that compared with models M_W and M_Z, model M_1 provides a better prediction for the RUL, especially at the early sampling points. Meanwhile, the PDF curves of model M_1 are narrower than that of models M_W and M_Z at each sampling point, which means that the former has lower uncertainty. Additionally, the models considering measurement errors, i.e., M_1 and M_W, have better predictive performance than the model without considering measurement errors, i.e., M_Z. Therefore, it is necessary to consider the error term in the state space model for the degradation data with measurement errors.

To quantitatively compare the performance, the mediums of the RUL (MRUL) along with the associated 95% confidence intervals (CIs), and the relative errors (REs) of the estimated RUL of three models are calculated at the 80th, 110th, and 140th sampling point, as shown in Table 1. The concepts of MRUL and RE are defined in [1]. Table 1 shows that model M_1 provides a better prediction for the RUL with lower uncertainty, which is confirmed by Fig. 3 qualitatively.

In addition, the MSE about the actual RUL at each sampling point for all models are shown in Fig. 4. This figure indicates that model M_1 has a lower MSE as a whole. Before time t_{50}, the MSE of models M_W and M_Z evolves dramatically, while the MSE of model M_1 avoids the abrupt change. Thus, model M_1 is obviously a better choice for RUL prediction. As the degradation data are accumulated, the MSE of model M_W quickly reduces and approaches the MSE of model M_1. One reasonable explanation for this is that both models consider measurement errors.

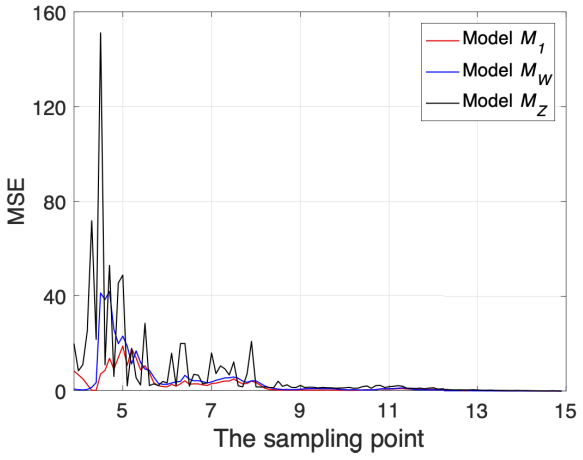


FIGURE 4. MSE of the RUL based on M_1, M_W, and M_Z.

Furthermore, the two filtering methods used in models M_1 and M_W are compared. In this article, the improved SIR filter is presented to obtain the particle representations of the underlying degradation state, i.e., $\hat{x}_j \approx \frac{1}{M}\sum_{k=1}^{M} \mathrm{f}_j^{(k)}$. In model M_W, Wang *et al.* [26] employed KF to obtain the distribution of the underlying degradation state, i.e., $x_j|y_{1:j} \sim \mathcal{N}(\hat{x}_{j|j}, P_{j|j})$. Thus, the estimated underlying degradation state can be expressed as $\hat{x}_j \approx \hat{x}_{j|j}$. Fig. 5 shows the absolute error between the actual underlying degradation state and the estimated underlying degradation state, $|x_j - \hat{x}_j|$, at each sampling point based on two filtering methods. Accordingly, the mean absolute errors (MAEs) [24] between x_j and $\hat{x}_j$ of models M_1 and M_W are calculated, i.e., 0.0871 and 0.1484. These results indicate that the proposed particle filter provides better filtering accuracy for the underlying degradation process, and also demonstrate the applicability of the proposed algorithm.

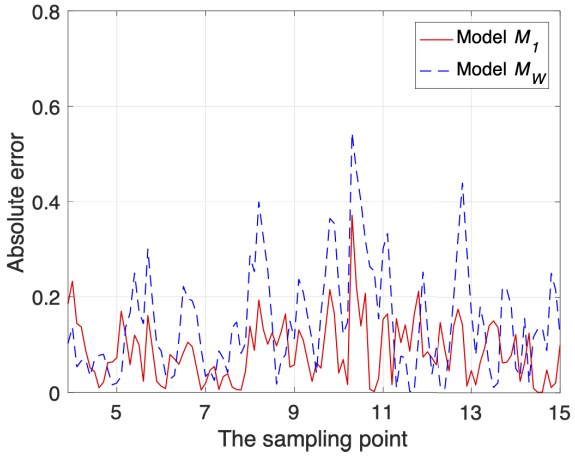


FIGURE 5. Comparison of two filtering methods based on MAE.

B. APPLICATION TO GAAS LASER DATA

In this section, we consider the degradation data of GaAs laser devices described by Meeker and Escobar [38] (Example 13.10). The quality characteristic of a laser device is its operating current. The lasers contain a feedback mechanism that can maintain nearly constant light output by increasing operating current as the laser degrades. The laser device is considered to have failed when the operating current exceeds a threshold level, $D = 10\%$. Fig. 6 gives 15 degradation paths of laser devices, the operating current of which is recorded every 250 hours. Details about this degradation dataset can be found in Table C.17 of [38].

Some degradation models [9], [10], [39] have been provided to demonstrate the effectiveness of the IG process in modelling the degradation dataset mentioned above. These studies assumed that $\Lambda(t) = t^q$, and thus it is adopted here. In the following, the first degradation path in Table C.17 of [38] is used for illustration purpose, i.e., the dark dotted line

TABLE 2. MRUL, 95% CIs, and REs of the three models.

	9th operating time			11th operating time			13th operating time		
	MRUL	95% CI	RE	MRUL	95% CI	RE	MRUL	95% CI	RE
M_1	1.73	[1.71,1.74]	1.43%	1.43	[1.42,1.44]	14.40%	0.70	[0.69,0.70]	7.20%
M_W	1.87	[1.85,1.88]	6.69%	1.53	[1.52,1.55]	22.48%	0.98	[0.97,0.99]	30.13%
M_Z	1.72	[1.71,1.73]	1.71%	1.45	[1.44,1.46]	16.01%	0.66	[0.65,0.67]	12.27%
M_2	1.73	[1.71,1.74]	1.43%	1.44	[1.43,1.45]	14.96%	0.70	[0.69,0.70]	7.33%

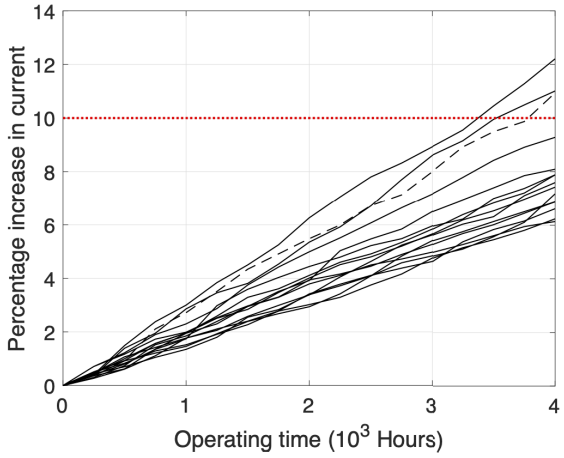


FIGURE 6. Degradation paths of the GaAs laser current.

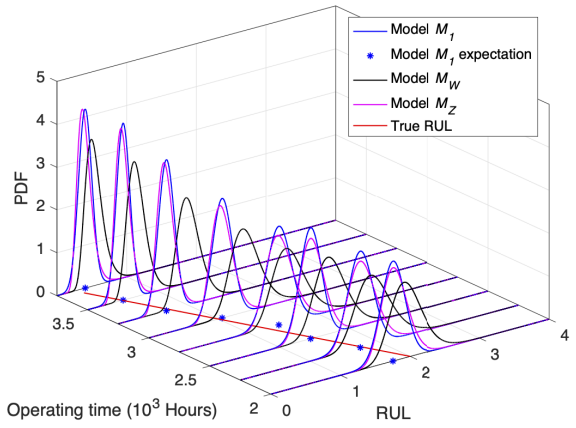


FIGURE 7. Predicted RUL based on M_1, M_W, and M_Z.

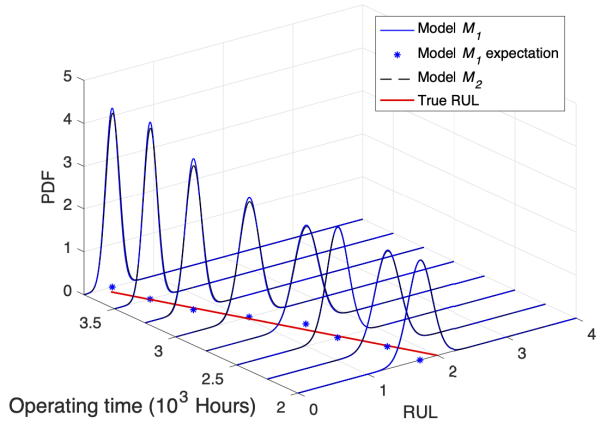


FIGURE 8. Predicted RUL based on M_1 and M_2.

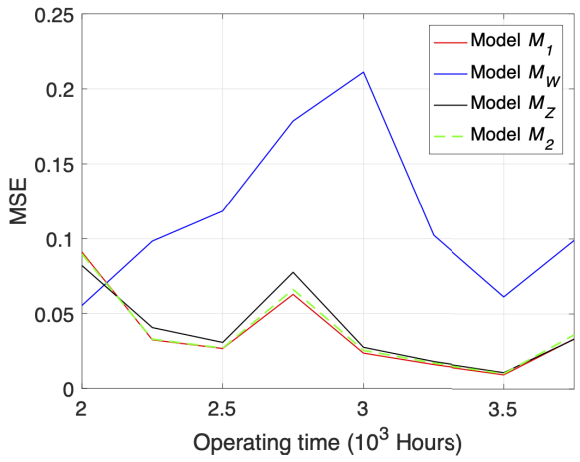


FIGURE 9. MSE of the RUL based on M_1, M_W, and M_Z.

in Fig. 6. It is noted that there are two other degradation paths that reach the threshold. Following the procedure below, we can draw similar conclusions from these degradation paths. Thus the detailed discussions about them are omitted.

Similarly, models M_1, M_W, and M_Z are used for RUL prediction. The parameter estimation is assumed to start from $t_8 = 2 \times 10^3$ hours. The initial parameters are set as $\hat{\boldsymbol{\theta}}_8^{(0)} = (1.10, 100.00, 0.39, 0.10, 0.10)$. Fig. 7 shows the PDFs of the RUL of three models at each sampling point. This figure indicates that model M_1 provides a more precise RUL prediction with lower uncertainty, while model M_W has relatively poor performance. In addition, the PDF curves of models M_1 and M_Z coincide with each other as the degradation data are accumulated. For further comparison, the MRUL along with the associated 95% CIs, and the REs of the estimated RUL of three models at the 9th, 11th, and 13th operating time are summarized in Table 2. From Table 2, it is found that the results of model M_1 are superior to the results of models M_W and M_Z. It is worth noting that although model M_W considers the error term, its predictive performance is not as good as that of model M_Z. It may be related to the assumption of model M_W that the drift coefficient is constant over the given sampling interval. In addition, due to small sample size, the RE of the same model does not necessary decrease as the degradation data are accumulated.

From Fig. 7 and Table 2, it is found that the performance of model M_Z without considering error term is close to that of model M_1. Therefore, we consider a special case of model M_1, i.e. the model characterized by (4) without considering error term, where $\varepsilon_j \sim \mathcal{N}(0, 0)$. Similar idea can be found in [24]. Correspondingly, *Algorithm 1* and *Algorithm 2* can be simplified directly according to Appendix A and Appendix C. Let M_2 denote this special case. Then, the results based on model M_2 are demonstrated in Table 2 and Fig. 8. As shown,

the error term has less influence on the RUL prediction, and thus model M_1 has reduced to model M_2.

To further compare the accuracy of RUL prediction, Fig. 9 shows the MSE about the actual RUL at each operating time for all models. As shown in Fig. 9, model M_1 and model M_2 have lower MSEs as a whole, except at operating time t_8. This is because the parameter estimates of both models have not yet converged at t_8 due to insufficient initial degradation data. In practice, increasing degradation data can avoid this problem.

From the qualitative and quantitative results in Fig. 7, Fig. 8, Fig. 9, and Table 2, we can conclude that the model M_1 and model M_2 provide better prediction accuracy for the laser data. Similar conclusions can be drawn from other degradation paths in Fig. 6, and thus the details are omitted here.

VI. CONCLUSION

In this article, an adaptive IG process model is presented for RUL prediction. To incorporate measurement errors and derive the RUL based on the historical degradation data, we construct a state space model accordingly. Then an improved SIR filter is proposed to overcome the nonGaussian assumptions in the state space model and obtain the particle representations of the estimated RUL, which can avoid high-dimensional integration. In addition, we investigate the EM algorithm combined with particle methods, based on which a parameter updating algorithm is provided to update the unknown parameters along with the RUL adaptively. To balance the efficiency and accuracy of the computation, we propose a multistage convergence assessment approach. Furthermore, a simulation study and a case application are used to demonstrate the applicability of the proposed method. Meanwhile, compared with the adaptive Wiener process models presented by [24], [26], the proposed model shows the superiority in describing monotonic degradation paths, which extends the application of IG process in degradation analysis.

Finally, there are two issues needed to be further studied. First, this article focuses on the situation that the volatility of the degradation parameter is relatively small. To address a more general case, it is necessary to change the structure of the proposed state space model and propose a new parameter estimation method. Second, we do not consider the effect of the error term on the PDF of RUL. For example, it can occur that the underlying degradation level exceeds the threshold although the observed degradation level does not exceed the threshold, or vice verse (see [40]). This situation deserves further investigation.

APPENDIXES

APPENDIX A
PROOF OF ALGORITHM 1

The proof of the SIR filter with only one hidden state can be found in [36], [41]. In the following, we provide the proof of the improved SIR particle filter with two hidden states.

Proof: The predictive distribution, $f(x_j, \xi_j | y_{1:j-1})$, can be expressed as

$$\begin{aligned} f(x_j, \xi_j | y_{1:j-1}) &= \iint f(x_j, \xi_j | x_{j-1}, \xi_{j-1}, y_{1:j-1}) \\ &\quad \cdot f(x_{j-1}, \xi_{j-1} | y_{1:j-1}) dx_{j-1} d\xi_{j-1} \\ &= \iint f(x_j | x_{j-1}, \xi_{j-1}) f(\xi_j | \xi_{j-1}) \\ &\quad \cdot f(x_{j-1}, \xi_{j-1} | y_{1:j-1}) dx_{j-1} d\xi_{j-1}. \end{aligned} \tag{21}$$

Given the particle representations of $f(x_{j-1}, \xi_{j-1} | y_{1:j-1})$, $\mathrm{p}_j^{(k)} = \mathrm{f}_{j-1}^{(k)} + \upsilon_j^{(k)}$ and $\tilde{\mathrm{p}}_j^{(k)} = \tilde{\mathrm{f}}_{j-1}^{(k)}$ can be regarded as random samples from $f(x_j, \xi_j | y_{1:j-1})$. Therefore, we obtain $\{\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)} | k = 1, \ldots, M\}$ by repeating above procedure M times.

If the new degradation data y_j is observed, we get $\{\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)} | k = 1, \ldots, M\}$ as follows

$$\begin{aligned} &\Pr\left((x_j, \xi_j) = (\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) | y_{1:j}\right) \\ &= \Pr\left((x_j, \xi_j) = (\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) | y_{1:j-1}, y_j\right) \\ &= \frac{p(y_j | \mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}, y_{1:j-1}) \Pr\left((x_j, \xi_j) = (\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) | y_{1:j-1}\right)}{p(y_j | y_{1:j-1})} \\ &= \frac{p(y_j | \mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) \Pr\left((x_j, \xi_j) = (\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) | y_{1:j-1}\right)}{\sum_{z=1}^{M} p(y_j | \mathrm{p}_j^{(z)}, \tilde{\mathrm{p}}_j^{(z)}) \Pr\left((x_j, \xi_j) = (\mathrm{p}_j^{(z)}, \tilde{\mathrm{p}}_j^{(z)}) | y_{1:j-1}\right)}. \end{aligned} \tag{22}$$

Here, $\Pr((x_j, \xi_j) = (\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)}) | y_{1:j-1}) = \frac{1}{M}$. If we sample from $\{\mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)} | k = 1, \ldots, M\}$ with weight $\mathrm{w}_j^{(k)} = p(y_j | \mathrm{p}_j^{(k)}, \tilde{\mathrm{p}}_j^{(k)})$, we get $(\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)})$.

This proof is complete.

APPENDIX B
ANALYTICAL EXPRESSION OF (11)

According to [9], we know that $\xi_j | x_{1:j} \sim \mathcal{TN}(\tilde{a}_0, \tilde{\sigma}_0^2)$ with parameters $\tilde{a}_0 = (\eta\sigma_0^2 \Lambda(t_j) + a_0)/(\eta\sigma_0^2 x_j + 1)$ and $\tilde{\sigma}_0^2 = \sigma_0^2/(\eta\sigma_0^2 x_j + 1)$. Then, the $f_{L_j | x_j, y_{1:j}}(l_j | x_j, y_{1:j})$ is further expressed in (23), as shown at the bottom of the next page. This equation can be regarded as the sum of two integrals, i.e., I_1 and I_2.

For deriving I_1, we need the following theorem, which can be obtained directly by using the formula 110 in [42].

Theorem 1 If $Z \sim \mathcal{N}(0, 1)$, and $a, b \in \mathbb{R}$, then

$$\begin{aligned} &\int_0^{+\infty} \phi(z)\phi(a + bz) dz \\ &\quad = \frac{1}{\sqrt{1 + b^2}} \phi\left(\frac{a}{\sqrt{1 + b^2}}\right) \Phi\left(-\frac{ab}{\sqrt{1 + b^2}}\right). \end{aligned} \tag{24}$$

Making the change of variable $m = (\xi_j - \tilde{a}_0)/\tilde{\sigma}_0$, and using Theorem 1, we obtain (25), as shown at the bottom of this page.

For deriving I_2, we provide Theorem 2, which can be derived by using the formulas 10,011.1 and 10,011.3 in [42].

Theorem 2 If $Z \sim \mathcal{N}(0, 1)$, and $a, b, c, d \in \mathbb{R}$, then

$$\begin{aligned}&\int_r^{+\infty} (cz+d)\phi(z)\Phi(a+bz)\,dz \\ &= \frac{cb}{\sqrt{1+b^2}}\phi\left(\frac{a}{\sqrt{1+b^2}}\right)\Phi\left(-r\sqrt{1+b^2}-\frac{ab}{\sqrt{1+b^2}}\right) \\ &\quad + c\phi(r)\,\Phi(a+br) + d\cdot\left[\Phi(-r)\,\Phi\left(\frac{a}{\sqrt{1+b^2}}\right)\right. \\ &\quad - \mathrm{T}\left(r, \frac{a}{r\sqrt{1+b^2}}\right) - \mathrm{T}\left(\frac{a}{\sqrt{1+b^2}}, \frac{r\sqrt{1+b^2}}{a}\right) \\ &\quad \left. + \mathrm{T}\left(r, \frac{a+br}{r}\right) + \mathrm{T}\left(\frac{a}{\sqrt{1+b^2}}, \frac{ab+r(1+b^2)}{a}\right)\right],\end{aligned} \tag{26}$$

where the T-function is defined as

$$\mathrm{T}(u, v) = \int_0^v \frac{\phi(u)\phi(ux)}{1+x^2}dx.$$

For notation simplicity, we assume that

$$\begin{cases}\tilde{h} = \tilde{a}_0 + 2\eta\tilde{\sigma}_0^2\Delta\Lambda(l_j) \\ \tilde{A} = -\dfrac{2\eta\Lambda'(t_j+l_j)\exp(2\tilde{a}_0\eta\Delta\Lambda(l_j) - 2\eta^2\tilde{\sigma}_0^2\Delta\Lambda^2(l_j))}{1-\Phi\left(-\frac{\tilde{a}_0}{\tilde{\sigma}_0}\right)}.\end{cases}$$

Making the change of variable $m = (\xi_j - \tilde{h})/\tilde{\sigma}_0$, we have

$$\begin{aligned}I_2 &= \frac{\tilde{A}}{\tilde{\sigma}_0}\int_0^{+\infty}\xi_j\phi\left(\frac{\xi_j-\tilde{h}}{\tilde{\sigma}_0}\right)\Phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}(\tilde{D}\xi_j + \Delta\Lambda(l_j))\right]d\xi_j \\ &= \tilde{A}\int_{-\frac{\tilde{h}}{\tilde{\sigma}_0}}^{+\infty}(\tilde{\sigma}_0 m + \tilde{h})\phi(m)\Phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}(\tilde{D}\tilde{h} + \Delta\Lambda(l_j))\right. \\ &\quad \left. -\sqrt{\eta\tilde{D}}\tilde{\sigma}_0 m\right]dm.\end{aligned} \tag{27}$$

Further, we assume that

$$\begin{cases}a = -\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\tilde{h} + \Delta\Lambda(l_j)\right) \\ b = -\sqrt{\eta\tilde{D}}\tilde{\sigma}_0 \\ c = \tilde{\sigma}_0 \\ d = \tilde{h}\end{cases}$$

Then I_2 is derived by using Theorem 2. The numerical calculation of the T-function can refer to the algorithm in [43].

Finally, the result of (11) follows by combing I_1 and I_2.

APPENDIX C
PROOF OF ALGORITHM 2

The proof of the particle smoother using backward simulation with only one hidden state can be found in [36], [37]. In the following, we provide the proof of the improved particle smoother with two hidden states.

Proof: The $p(x_{1:m}, \xi_{1:m}|y_{1:m})$ can be expressed as

$$\begin{aligned}&p(x_{1:m}, \xi_{1:m}|y_{1:m}) \\ &= p(x_m, \xi_m|y_{1:m})\prod_{j=1}^{m-1} p\left(x_j, \xi_j|x_{j+1}, \ldots, x_m, \xi_{j+1}\ldots, \xi_m, y_{1:m}\right)\end{aligned} \tag{28}$$

and

$$\begin{aligned}&p\left(x_j, \xi_j|x_{j+1}, \ldots, x_m, \xi_{j+1}\ldots, \xi_m, y_{1:m}\right) \\ &= p\left(x_j, \xi_j|x_{j+1}, \xi_{j+1}, y_{1:j}\right) \\ &= \frac{p\left(x_{j+1}, \xi_{j+1}, x_j, \xi_j|y_{1:j}\right)}{p\left(x_{j+1}, \xi_{j+1}|y_{1:j}\right)} \\ &\propto p\left(x_{j+1}, \xi_{j+1}|x_j, \xi_j, y_{1:j}\right)p\left(x_j, \xi_j|y_{1:j}\right) \\ &= p\left(x_{j+1}|x_j, \xi_j\right)p\left(\xi_{j+1}|\xi_j\right)p\left(x_j, \xi_j|y_{1:j}\right).\end{aligned} \tag{29}$$

Given filtering results $\{\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)}|k = 1,\ldots,M\}$ and associated weights $\{\mathrm{w}_j^{(k)}|k = 1,\ldots,M\}$, the $p(x_j, \xi_j|x_{j+1}, \ldots,$

$$\begin{aligned}&f_{L_j|x_j,y_{1:j}}(l_j|x_j, y_{1:j}) \\ &= \int_0^{+\infty} 2\Lambda'(t_j+l_j)\sqrt{\frac{\eta}{\tilde{D}}}\phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\xi_j - \Delta\Lambda(l_j)\right)\right]\frac{\phi\left[\left(\xi_j-\tilde{a}_0\right)/\tilde{\sigma}_0\right]}{\tilde{\sigma}_0\left[1-\Phi\left(-\tilde{a}_0/\tilde{\sigma}_0\right)\right]}d\xi_j \\ &\quad + \int_0^{+\infty}\left\{-2\eta\xi_j\Lambda'(t_j+l_j)\exp\left(2\eta\xi_j\Delta\Lambda(l_j)\right)\Phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\left(\tilde{D}\xi_j + \Delta\Lambda(l_j)\right)\right]\right\}\frac{\phi\left[\left(\xi_j-\tilde{a}_0\right)/\tilde{\sigma}_0\right]}{\tilde{\sigma}_0\left[1-\Phi\left(-\tilde{a}_0/\tilde{\sigma}_0\right)\right]}d\xi_j \\ &= I_1 + I_2.\end{aligned} \tag{23}$$

$$\begin{aligned}I_1 &= \frac{2\Lambda'(t_j+l_j)}{1-\Phi(-\frac{\tilde{a}_0}{\tilde{\sigma}_0})}\sqrt{\frac{\eta}{\tilde{D}}}\int_{-\frac{\tilde{a}_0}{\tilde{\sigma}_0}}^{+\infty}\phi(m)\phi\left[-\sqrt{\eta\tilde{D}}\tilde{\sigma}_0 m - \sqrt{\frac{\eta}{\tilde{D}}}(\tilde{D}\tilde{a}_0 - \Delta\Lambda(l_j))\right]dm \\ &= \frac{2\Lambda'(t_j+l_j)}{\sqrt{1+\eta\tilde{D}\tilde{\sigma}_0^2}\left[1-\Phi(-\frac{\tilde{a}_0}{\tilde{\sigma}_0})\right]}\sqrt{\frac{\eta}{\tilde{D}}}\phi\left[-\sqrt{\frac{\eta}{\tilde{D}}}\frac{(\tilde{D}\tilde{a}_0 - \Delta\Lambda(l_j))}{\sqrt{1+\eta\tilde{D}\tilde{\sigma}_0^2}}\right]\Phi\left[\frac{\tilde{a}_0}{\tilde{\sigma}_0}\sqrt{1+\eta\tilde{D}\tilde{\sigma}_0^2} - \frac{\eta\tilde{\sigma}_0(\tilde{D}\tilde{a}_0 - \Delta\Lambda(l_j))}{\sqrt{1+\eta\tilde{D}\tilde{\sigma}_0^2}}\right].\end{aligned} \tag{25}$$

$x_m, \xi_{j+1} \ldots, \xi_m, y_{1:m})$ can be approximated as follows

$$p\left(x_j, \xi_j | x_{j+1}, \ldots, x_m, \xi_{j+1} \ldots, \xi_m, y_{1:m}\right) \approx \sum_{k=1}^{M} \mathrm{w}_{j|j+1}^{(k)} \delta\left((x_j, \xi_j) - (\mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)})\right), \quad (30)$$

where

$$\mathrm{w}_{j|j+1}^{(k)} = \frac{\mathrm{w}_j^{(k)} f(x_{j+1} | \mathrm{f}_j^{(k)}, \tilde{\mathrm{f}}_j^{(k)})}{\sum_{z=1}^{M} \mathrm{w}_j^{(z)} f(x_{j+1} | \mathrm{f}_j^{(z)}, \tilde{\mathrm{f}}_j^{(z)})},$$

and $\delta(x)$ denotes an indicator function whose value is 1 if $x = 0$, 0 otherwise.

The idea of the particle smoothing algorithm is: given random samples $(\mathrm{s}_{j+1}, \ldots, \mathrm{s}_m)'$ and $(\tilde{\mathrm{s}}_{j+1}, \ldots, \tilde{\mathrm{s}}_m)'$ approximately from $p\left(x_{j+1}, \ldots, x_m, \xi_{j+1} \ldots, \xi_m | y_{1:m}\right)$, it is possible to obtain samples s_j and $\tilde{\mathrm{s}}_j$ from $p(x_j, \xi_j | \mathrm{s}_{j+1}, \ldots, \mathrm{s}_m, \tilde{\mathrm{s}}_{j+1} \ldots, \tilde{\mathrm{s}}_m, y_{1:m})$, and thus $(\mathrm{s}_j, \mathrm{s}_{j+1}, \ldots, \mathrm{s}_m)'$ and $(\tilde{\mathrm{s}}_j, \tilde{\mathrm{s}}_{j+1}, \ldots, \tilde{\mathrm{s}}_m)'$ are random samples from $p(x_j, \ldots, x_m, \xi_j \ldots, \xi_m | y_{1:m})$. By repeating this procedure sequentially back over time, the results of the particle smoother can be obtained.

This proof is complete.

ACKNOWLEDGMENT

The authors are grateful to the editor, associate editor and anonymous referees for their insightful suggestions and constructive comments, which improved the quality of the article. They would particularly like to thank QingQing Zhai and Weiwen Peng for their helpful suggestions.

REFERENCES

[1] X. S. Si, "An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data," *IEEE Trans. Ind. Electron.*, vol. 62, no. 8, pp. 5082–5096, Aug. 2015.

[2] Z. Zhang, X. Si, C. Hu, and X. Kong, "Degradation modeling-based remaining useful life estimation: A review on approaches for systems with heterogeneity," *Proc. Inst. Mech. Eng. O-J. Risk Rel.*, vol. 229, no. 4, pp. 343–355, Aug. 2015.

[3] C.-H. Hu, H. Pei, X.-S. Si, D.-B. Du, Z.-N. Pang, and X. Wang, "A prognostic model based on DBN and diffusion process for degrading bearing," *IEEE Trans. Ind. Electron.*, to be published.

[4] Z. Huang, Z. Xu, W. Wang, and Y. Sun, "Remaining useful life prediction for a nonlinear heterogeneous Wiener process model with an adaptive drift," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 687–700, Jun. 2015.

[5] Z.-S. Ye and M. Xie, "Stochastic modelling and analysis of degradation for highly reliable products," *Appl. Stochastic Models Bus. Ind.*, vol. 31, no. 1, pp. 16–32, 2015, doi: 10.1002/asmb.2063.

[6] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Estimating remaining useful life with three-source variability in degradation modeling," *IEEE Trans. Rel.*, vol. 63, no. 1, pp. 167–190, Mar. 2013.

[7] Z.-S. Ye, N. Chen, and Y. Shen, "A new class of Wiener process models for degradation analysis," *Rel. Eng. Syst. Saf.*, vol. 139, pp. 58–67, Jul. 2015.

[8] X. Si, T. Li, Q. Zhang, and C. Hu, "Prognostics for linear stochastic degrading systems with survival measurements," *IEEE Trans. Ind. Electron.*, vol. 67, no. 4, pp. 3202–3215, Apr. 2020.

[9] Z.-S. Ye and N. Chen, "The inverse Gaussian process as a degradation model," *Technometrics*, vol. 56, no. 3, pp. 302–311, Jul. 2014, doi: 10.1080/00401706.2013.830074.

[10] C.-Y. Peng, "Inverse Gaussian processes with random effects and explanatory variables for degradation data," *Technometrics*, vol. 57, no. 1, pp. 100–111, 2014.

[11] Y. Zhou, Y. Sun, J. Mathew, R. Wolff, and L. Ma, "Latent degradation indicators estimation and prediction: A Monte Carlo approach," *Mech. Syst. Sig. Process.*, vol. 25, no. 1, pp. 222–236, 2011.

[12] C.-C. Tsai, S.-T. Tseng, and N. Balakrishnan, "Optimal design for degradation tests based on gamma processes with random effects," *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 604–613, Jun. 2012.

[13] Y. Shu, Q. Feng, E. P. C. Kao, and H. Liu, "Lévy-driven non-Gaussian Ornstein–Uhlenbeck processes for degradation-based reliability analysis," *IIE Trans.*, vol. 48, no. 11, pp. 993–1003, 2016.

[14] Y. Deng, A. Barros, and A. Grall, "Degradation modeling based on a time-dependent Ornstein-Uhlenbeck process and residual useful lifetime estimation," *IEEE Trans. Rel.*, vol. 65, no. 1, pp. 126–140, Mar. 2016.

[15] X. Chen, G. Ji, X. Sun, and Z. Li, "Inverse Gaussian–based model with measurement errors for degradation analysis," *Proc. Inst. Mech. Eng. O, J. Risk Rel.*, vol. 233, no. 6, pp. 1086–1098, 2019, doi: 10.1177/1748006X19860682.

[16] W. Wang, M. Carr, W. Xu, and K. Kobbacy, "A model for residual life prediction based on Brownian motion with an adaptive drift," *Microelectron. Rel.*, vol. 51, no. 2, pp. 285–293, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0026271410005032

[17] X.-S. Si, W. Wang, C.-H. Hu, M.-Y. Chen, and D.-H. Zhou, "A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation," *Mech. Syst. Signal Process.*, vol. 35, nos. 1–2, pp. 219–237, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327012003226

[18] X. Wang, N. Balakrishnan, and B. Guo, "Residual life estimation based on a generalized Wiener degradation process," *Rel. Eng. Syst. Saf.*, vol. 124, pp. 13–23, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832013003086

[19] H. Wang, X. Ma, and Y. Zhao, "An improved Wiener process model with adaptive drift and diffusion for online remaining useful life prediction," *Mech. Syst. Signal Process.*, vol. 127, pp. 370–387, Jul. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327019301931

[20] Z.-X. Zhang, X.-S. Si, C.-H. Hu, X.-X. Hu, and G.-X. Sun, "An adaptive prognostic approach incorporating inspection influence for deteriorating systems," *IEEE Trans. Rel.*, vol. 68, no. 1, pp. 302–316, Mar. 2019.

[21] Z. Huang, Z. Xu, X. Ke, W. Wang, and Y. Sun, "Remaining useful life prediction for an adaptive skew-Wiener process model," *Mech. Syst. Signal Process.*, vol. 87, pp. 294–306, Mar. 2017.

[22] D. Wang and K.-L. Tsui, "Brownian motion with adaptive drift for remaining useful life prediction: Revisited," *Mech. Syst. Signal.*, vol. 99, pp. 691–701, Jan. 2018.

[23] D. Wang, Y. Zhao, F. F. Yang, and K.-L. Tsui, "Nonlinear-drifted Brownian motion with multiple hidden states for remaining useful life prediction of rechargeable batteries," *Mech. Syst. Signal Process.*, vol. 93, no. 1, pp. 531–544, Sep. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327017301012

[24] Q. Zhai and Z.-S. Ye, "RUL prediction of deteriorating products using an adaptive Wiener process model," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2911–2921, Dec. 2017.

[25] X. Wang, C. Hu, X. Si, J. Zheng, H. Pei, Y. Yu, and Z. Ren, "An adaptive prognostic approach for newly developed system with three-source variability," *IEEE Access*, vol. 7, pp. 53091–53102, 2019.

[26] X. Wang, C. Hu, X. Si, Z. Pang, and Z. Ren, "An adaptive remaining useful life estimation approach for newly developed system based on nonlinear degradation model," *IEEE Access*, vol. 7, pp. 82162–82173, 2019.

[27] Q. Wei and D. Xu, "Remaining useful life estimation based on gamma process considered with measurement error," in *Proc. 10th Int. Conf. Rel., Maintainability Saf. (ICRMS)*, Aug. 2014, pp. 645–649.

[28] S. Zhang, W. Zhou, and H. Qin, "Inverse Gaussian process-based corrosion growth model for energy pipelines considering the sizing error in inspection data," *Corros. Sci.*, vol. 73, pp. 309–320, Aug. 2013.

[29] H. Qin, S. Zhang, and W. Zhou, "Inverse Gaussian process-based corrosion growth modeling and its application in the reliability analysis for energy pipelines," *Frontiers Structural Civil Eng.*, vol. 7, no. 3, pp. 276–287, 2013.

[30] W. Peng, Y.-F. Li, Y.-J. Yang, H.-Z. Huang, and M. J. Zuo, "Inverse Gaussian process models for degradation analysis: A Bayesian perspective," *Rel. Eng. Syst. Saf.*, vol. 130, no. 1, pp. 175–189, Oct. 2014.

[31] W. Xu and W. Wang, "RUL estimation using an adaptive inverse Gaussian model," in *Proc. IEEE Prognostics Health Manage. Conf. (PHM)*, in Chemical Engineering Transactions, vol. 33. Milan, Italy: Aidic Servizi Srl, 2013, pp. 331–336.

[32] W. Peng, Y.-F. Li, Y.-J. Yang, J. Mi, and H.-Z. Huang, "Bayesian degradation analysis with inverse Gaussian process models under time-varying degradation rates," *IEEE Trans. Rel.*, vol. 66, no. 1, pp. 84–96, Mar. 2017.

[33] D. Pan, J.-B. Liu, and J. Cao, "Remaining useful life estimation using an inverse Gaussian degradation model," *Neurocomputing*, vol. 185, pp. 64–72, Apr. 2016.
[34] X. Si, T. Li, and Q. Zhang, "A general stochastic degradation modeling approach for prognostics of degrading systems with surviving and uncertain measurements," *IEEE Trans. Rel.*, vol. 68, no. 3, pp. 1080–1100, Sep. 2019.
[35] X.-S. Si, Z. Ren, X. Hu, C.-H. Hu, and Q. Shi, "A novel degradation modelling and prognostic framework for closed-loop systems with degrading actuator," *IEEE Trans. Ind. Electron.*, to be published.
[36] J. Kim, "Parameter estimation in stochastic volatility models with missing data using particle methods and the EM algorithm," Ph.D. dissertation, Dept. Statist., Univ. Pittsburgh, Pittsburgh, PA, USA, 2005.
[37] S. J. Godsill, A. Doucet, and M. West, "Monte Carlo smoothing for nonlinear time series," *J. Amer. Stat. Assoc.*, vol. 99, no. 465, pp. 156–168, 2004, doi: 10.1198/016214504000000151.
[38] W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*. New York, NY, USA: Wiley, 1998.
[39] X. Wang and D. Xu, "An inverse Gaussian process model for degradation data," *Technometrics*, vol. 52, no. 2, pp. 188–197, May 2010. . [Online]. Available: http://www.jstor.org/stable/27867223
[40] G. Pulcini, "A perturbed gamma process with statistically dependent measurement errors," *Rel. Eng. Syst. Saf.*, vol. 152, pp. 296–306, Aug. 2016.
[41] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *J. Comput. Graph. Statist.*, vol. 5, no. 1, pp. 1–25, 1996. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10618600.1996.10474692
[42] D. B. Owen, "A table of normal integrals," *Commun. Stat.-Simul. Comput.*, vol. 9, no. 4, pp. 389–419, 1980.
[43] J. C. Young and C. E. Minder, "Algorithm as 76: An integral useful in calculating non-central t and bivariate normal probabilities," *J. Roy. Stat. Soc. C (Appl. Statist.)*, vol. 23, no. 3, pp. 455–457, 1974. [Online]. Available: http://www.jstor.org/stable/2347148

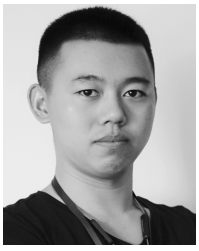
XUDAN CHEN received the B.Eng. and M.Eng. degrees from the Rocket Force University of Engineering, Xi'an, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree. His research interests include reliability estimation, safety assessment, and life estimation.

XINLI SUN received the B.Eng., M.Eng., and Ph.D. degrees from the Rocket Force University of Engineering, Xi'an, China, in 1985, 1988, and 2003, respectively. He is currently a Professor and a Ph.D. Supervisor with the Rocket Force University of Engineering. He has published 4 books and about 60 articles. His research interests include reliability engineering, safety assessment, and mechanics of explosion.

XIAOSHENG SI received the B.Eng., M.Eng., and Ph.D. degrees from the Department of Automation, Rocket Force University of Engineering, Xi'an, China, in 2006, 2009, and 2014, respectively.

He is currently a Professor with the Rocket Force University of Engineering. He has authored or coauthored more than 50 articles in several journals including the *European Journal of Operational Research*, the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, the IEEE TRANSACTIONS ON RELIABILITY, the IEEE TRANSACTIONS ON FUZZY SYSTEMS, the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS—PART A, the IEEE TRANSACTION ON AUTOMATION SCIENCE AND ENGINEERING, *Reliability Engineering and System Safety*, and *Mechanical Systems and Signal Processing*. His research interests include evidence theory, expert systems, prognostics and health management, reliability estimation, predictive maintenance, and lifetime estimation.

Dr. Si is an Associate Editor of IEEE ACCESS. He is an Active Reviewer for a number of international journals.



GUODONG LI received the B.Eng. degree from the Rocket Force University of Engineering, Xi'an, China, in 2017, where he is currently pursuing the M.Eng. degree. His research interests include reliability estimation and life estimation.