SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

January 2021, Vol. 64 112205:1–112205:16
https://doi.org/10.1007/s11432-019-2724-5

Online remaining-useful-life estimation with a Bayesian-updated expectation-conditional-maximization algorithm and a modified Bayesian-model-averaging method

Yong YU[1], Xiaosheng SI[1,2*], Changhua HU[1], Jianfei ZHENG[1] & Jianxun 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* 710025*, China*

Received 29 August 2019/Accepted 1 November 2019/Published online 18 December 2020

Abstract Online remaining-useful-life (RUL) estimation is an effective method with respect to ensuring the safety of complex-huge systems. Generally, current methods assume a specific degradation model when degradation values are observed in the initial degradation phase. However, this assumption may not always be robust enough owing to the often-ambiguous inherent incipient-degradation characteristic. Therefore, besides model-parameter uncertainty, the uncertainty of the degradation model is worth examining in online RUL estimations. In this paper, a Bayesian-updated expectation-conditional-maximization (ECM) algorithm is adopted to address the uncertainty of prior parameters, and a modified Bayesian-model-averaging method is used to deal with the uncertainty of the degradation model. Then, simulation studies are conducted to analyze the effectiveness of the proposed fusion algorithm. Results suggest that the Bayesian-updated ECM algorithm and modified Bayesian-model-averaging method effectively address the associated uncertainties of model parameters and the degradation model itself. Finally, we apply the proposed fusion algorithm to predict the RUL of a gyroscope.

Keywords online RUL estimation, parameter uncertainty, model uncertainty, Bayesian method, ECM algorithm, Bayesian model averaging

Citation 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(1): 112205, https://doi.org/10.1007/s11432-019-2724-5

1 Introduction

Complex-huge systems play significant roles in a number of industries, such as manufacturing, aerospace engineering, the military, and transportation. Failure in any one of these systems can have catastrophic consequences with profound impacts and potential safety hazards [1–3]. Remaining-useful-life (RUL) prediction is an important method with respect to avoiding serious consequences as well as ensuring operational safety in various industries [4–6]. Unlike other common systems, complex equipment, such as high-speed rails and aircraft carriers, is newly manufactured; accordingly, historical-degradation data are nonexistent. This is problematic because common approaches that require historical-degradation data to accomplish RUL predictions are rendered useless. To overcome this, online RUL prediction utilizes online monitoring values; therefore, it is the optimal choice for complex-huge systems [7].

Usually, online RUL prognostics mainly comprise two aspects: model selection and model-parameter estimation [8]. In the model-selection process, a specific degradation model is often assumed when the degradation signals are observed in the initial degradation phase. However, the degradation characteristics are often indistinct and usually lead to model uncertainty; for instance, the characteristics datasets comprising the fatigue-crack growth data from Hudak and Saxena [9] and laser-device degradation data from Meeker and Escobar [10]. Indeed, both of these degradation datasets have been modeled using

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

different degradation paths: the Gamma-stochastic process [11] and the regression model [12] for the fatigue-crack-growth data, and the Wiener process with both normal distribution [13] and skew-normal distribution [14] for the laser-device data. However, owing to indistinct degradation characteristics, these widely adopted degradation models should be carefully selected for particular application purposes. Besides indistinct degradation characteristics, imprecise model-selection criteria can also result in model uncertainty. Although various kinds of model-selection criteria have been proposed, such as minimum description length [15], Akaike information criterion [16], Bayesian information criterion [12], and empirical average log-likelihood, current model-selection criteria are not comprehensive. Usually, they primarily consider the sampling errors in parameter estimation [17]. Indeed, owing to indistinct degradation characteristics and insufficient selection criteria, a fixed degradation model may be unscientific, especially for real-time RUL prediction with incomplete degradation trajectory. Therefore, model uncertainty should be considered when dealing with online degradation data. Once the degradation model is selected, model-parameter estimation becomes the primary focus. In online RUL prognosis, Bayesian parameter updating is an effective method with respect to accomplishing parameter estimation. Usually, the accuracy of the estimated posterior Bayesian parameters is affected by the prior assumed parameters. In current Bayesian parameter updating, constant prior-parameter distribution is often assumed [18]. However, the prior-parameter distribution of newly manufactured devices is usually unknown. Thus, the uncertainty of the assumed prior parameters should be considered in online parameter estimating, and new parameter-updating methods should be adopted in order to reduce the influence of prior-parameter uncertainty.

In online RUL estimation, prior-parameter uncertainty affects the estimated-model posterior parameters, and model uncertainty influences the parameter form. Indeed, posterior parameters and parameter form both directly influence the prediction accuracy of RUL estimation. Thus, in order to simultaneously address prior-parameter uncertainty and model uncertainty, this paper proposes a novel fusion algorithm in order to predict the online RUL of newly manufactured devices. This fusion algorithm adopts a Bayesian-updated expectation-conditional-maximization (ECM) algorithm to address the prior-parameter uncertainty; each time a new degradation value is observed, the prior parameters can be updated so as to ensure estimation accuracy. Furthermore, by combining parameter-updating processes, a modified Bayesian-model-averaging method is proposed to address the model uncertainty of online RUL. Based on the modified Bayesian-model-averaging method, the degradation probability of each candidate model can be updated once a new degradation value is observed. Simulation studies and a gyro case study indicate that the proposed fusion algorithm can obtain precise RUL-estimation results when addressing prior-parameter uncertainty and degradation-model uncertainty.

This paper is organized as follows: Section 2 estimates the parameters of two candidate models using the Bayesian-updated ECM algorithm; Section 3 calculates the RUL of the target system with a modified Bayesian-model-averaging method; numerical experiments are provided in Section 4; and in Section 5, a conclusion is provided.

2 Parameter estimation with Bayesian-updated ECM algorithm

2.1 Candidate degradation models

The estimated posterior parameters of the traditional Bayesian parameter-updating method are directly affected by the assumed prior-parameter distributions, which are uncertain for newly manufactured devices. Usually, the collaboration between Bayesian updating and expectation maximization (EM) algorithms can be adopted in order to address this problem [19]. However, when the associated maximum likelihood estimation is complicated, the M-step of the EM algorithm is computationally unattractive and usually cannot be used to obtain analytical results [20]. Thus, in order to address prior-parameter uncertainty, we adopt a Bayesian-updated ECM algorithm to estimate the posterior parameters. As for candidate degradation-model selection, the selected models should possess similar characteristics, such as monotonicity, concavity, and convexity. Thus, considering the fact that exponential models are widely used in the prognostics of mechanical and electrical systems [21–23], we adopt two typical exponential models in order to demonstrate the effectiveness of the Bayesian-updated ECM algorithm with respect to estimating model parameters.

In the first model, the error term is described as an independent and identically distributed (IID) random variable. This means that the error terms, $\varepsilon(t_1), \dots, \varepsilon(t_k)$, are IID random variables. Letting $X(t)$ denote the degradation, the model can be expressed as

$$X(t) = \varphi_x + \theta_x \exp\left(\beta_x t_i + \varepsilon_x(t) - \frac{\sigma_x^2}{2}\right) = \varphi_x + \theta_x \exp(\beta_x t) \exp\left(\varepsilon_x(t) - \frac{\sigma_x^2}{2}\right), \tag{1}$$

where φ_x is the fixed intercept representing initial degradation, $\beta_x \sim N(\mu_{x1}, \sigma_{x1}^2)$ is a normal random variable, θ_x is a lognormal random variable that follows $\ln\theta_x \sim N(\mu_{x0}, \sigma_{x0}^2)$, and $\varepsilon_x(t) \sim N(0, \sigma_x^2)$ is a normal distribution. For simplicity, this model could be rewritten as

$$S_x(t) = \theta_x' + \beta_x' t + \varepsilon_x(t), \tag{2}$$

where $S_x(t) = \ln(X(t) - \varphi_x)$, $\theta_x' = \ln\theta_x - \sigma_x^2/2$, $\beta_x' = \beta_x$. Then $\theta_x' \sim N(\mu_{x0}', \sigma_{x0}^2)$ and $\beta_x' \sim N(\mu_{x1}', \sigma_{x1}^2)$, where $\mu_{x0}' = \mu_{x0} - \sigma_x^2/2$, $\mu_{x1}' = \mu_{x1}$.

In the second model, the error term is assumed as a Brownian motion [23]. In this case, the error increments, which are denoted as $\varepsilon(t_1), \varepsilon(t_2) - \varepsilon(t_1), \dots, \varepsilon(t_k) - \varepsilon(t_{k-1})$, are independent random variables. $Y(t)$ can denote this kind of degradation process model, represented as

$$Y(t) = \varphi_y + \theta_y \exp\left(\beta_y t_i + \varepsilon_y(t) - \frac{\sigma_y^2 t}{2}\right) = \varphi_y + \theta_y \exp(\beta_y t) \exp\left(\varepsilon_y(t) - \frac{\sigma_y^2 t}{2}\right), \tag{3}$$

where φ_y, β_y, and θ_y are the same as the parameters in (1), but $\varepsilon_y(t) = \sigma_y W(t)$ is a Brownian motion. Similarly, we assume that θ_y, β_y, and $\varepsilon_y(t)$ are mutually independent, and that they obey distributions as follows: $\ln\theta_y \sim N(\mu_{y0}, \sigma_{y0}^2)$, $\beta_y \sim N(\mu_{y1}, \sigma_{y1}^2)$, and $\varepsilon_y(t) \sim N(0, \sigma_y^2 t)$. For simplicity, this model can be re-expressed as

$$S_y(t) = \theta_y' + \beta_y' t + \varepsilon_y(t), \tag{4}$$

where $S_y(t) = \ln(Y(t) - \varphi_y)$, $\theta_y' = \ln\theta_y$ and $\beta_y' = \beta_y - \sigma_y^2/2$. Accordingly, $\theta_y' \sim N(\mu_{y0}', \sigma_{y0}^2)$ and $\beta_y' \sim N(\mu_{y1}', \sigma_{y1}^2)$, where $\mu_{y0}' = \mu_{y0}$, $\mu_{y1}' = \mu_{y1} - \sigma_y^2/2$.

2.2 Bayesian parameter-updating method

Because our aim is to address real-time monitoring information, we adopt a Bayesian method to update the model parameters. Assuming that the prior distributions of the candidate model parameters are known, we can obtain the joint posterior distribution of stochastic parameters.

Let $S_{x,1:k} = \{S_{x,1}, S_{x,2}, \dots, S_{x,k}\}$, where $S_{x,k} = \ln(X(t_k) - \varphi_x)$. Then, for the exponential model with IID errors, the joint posterior distribution of the parameters can be expressed as $\theta_x', \beta_x' | S_{x,1:k} \sim N(\mu_{\theta_x',k}, \sigma_{\theta_x',k}^2, \mu_{\beta_x',k}, \sigma_{\beta_x',k}^2, \rho_{x,k})$ with [23]

$$\mu_{\theta_x',k} = \frac{(\sum_{i=1}^k S_{xi}\sigma_{x0}^2 + \mu_{x0}'\sigma_x^2)(\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2) - (\sum_{i=1}^k \sigma_{x0}^2 t_i)(\sum_{i=1}^k S_{xi} t_i \sigma_{x1}^2 + \mu_{x1}'\sigma_x^2)}{(k\sigma_{x0}^2 + \sigma_x^2)(\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2) - (\sum_{i=1}^k t_i \sigma_{x1}^2)(\sum_{i=1}^k t_i \sigma_{x0}^2)}, \tag{5}$$

$$\mu_{\beta_x',k} = \frac{(k\sigma_{x0}^2 + \sigma_x^2)(\sum_{i=1}^k S_{xi} t_i \sigma_{x1}^2 + \mu_{x1}'\sigma_x^2) - (\sum_{i=1}^k \sigma_{x1}^2 t_i)(\sum_{i=1}^k S_{xi}\sigma_{x0}^2 + \mu_{x0}'\sigma_x^2)}{(k\sigma_{x0}^2 + \sigma_x^2)(\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2) - (\sum_{i=1}^k t_i \sigma_{x1}^2)(\sum_{i=1}^k t_i \sigma_{x0}^2)}, \tag{6}$$

$$\sigma_{\theta_x',k}^2 = \frac{\sigma_{x0}^2 \sigma_x^2 (\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2)}{(k\sigma_{x0}^2 + \sigma_x^2)(\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2) - (\sum_{i=1}^k t_i)^2 \sigma_{x0}^2 \sigma_{x1}^2}, \tag{7}$$

$$\sigma_{\beta_x',k}^2 = \frac{\sigma_x^2 \sigma_{x1}^2 (k\sigma_{x0}^2 + \sigma_x^2)}{(k\sigma_{x0}^2 + \sigma_x^2)(\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2) - (\sum_{i=1}^k t_i)^2 \sigma_{x0}^2 \sigma_{x1}^2}, \tag{8}$$

$$\rho_{x,k} = \frac{-\sigma_{x0}\sigma_{x1} \sum_{i=1}^k t_i}{\sqrt{k\sigma_{x0}^2 + \sigma_x^2}\sqrt{\sum_{i=1}^k \sigma_{x1}^2 t_i^2 + \sigma_x^2}}. \tag{9}$$

Let $S_{y,1:k} = \{S_{y,1}, S_{y,2}, \ldots, S_{y,k}\}$, where $S_{y,k} = \ln(Y(t_k)-\varphi_y)-\ln(Y(t_{k-1})-\varphi_y)$ and $S_{y,1} = \ln(Y(t_1)-\varphi_y)$. Thus, for the exponential model with Brownian errors, the joint posterior distribution of the parameters can be expressed as $\theta'_y, \beta'_y|S_{y,1:k} \sim N(\mu_{\theta'_y,k}, \sigma^2_{\theta'_y,k}, \mu_{\beta'_y,k}, \sigma^2_{\beta'_y,k}, \rho_{y,k})$ with [19, 23]

$$\mu_{\theta'_y,k} = \frac{(S_{y1}\sigma^2_{y0} + \mu'_{y0}\sigma^2_y t_1)(\sigma^2_{y1}t_k + \sigma^2_y) - \sigma^2_{y0}t_1(\sigma^2_{y1}\sum_{i=1}^k S_{yi} + \mu'_{y1}\sigma^2_y)}{(\sigma^2_{y0} + \sigma^2_y t_1)(\sigma^2_{y1}t_k + \sigma^2_y) - \sigma^2_{y0}\sigma^2_{y1}t_1}, \tag{10}$$

$$\mu_{\beta'_y,k} = \frac{(\sigma^2_{y1}\sum_{i=1}^k S_{yi} + \mu'_{y1}\sigma^2_y)(\sigma^2_y t_1 + \sigma^2_{y0}) - \sigma^2_{y1}(S_{y1}\sigma^2_{y0} + \mu'_{y0}\sigma^2_y t_1)}{(\sigma^2_{y0} + \sigma^2_y t_1)(\sigma^2_{y1}t_k + \sigma^2_y) - \sigma^2_{y0}\sigma^2_{y1}t_1}, \tag{11}$$

$$\sigma^2_{\theta'_y,k} = \frac{\sigma^2_y\sigma^2_{y0}t_1(\sigma^2_{y1}t_k + \sigma^2_y)}{(\sigma^2_{y0} + \sigma^2_y t_1)(\sigma^2_{y1}t_k + \sigma^2_y) - \sigma^2_{y0}\sigma^2_{y1}t_1}, \tag{12}$$

$$\sigma^2_{\beta'_y,k} = \frac{\sigma^2_y\sigma^2_{y1}(\sigma^2_{y0} + \sigma^2_y t_1)}{(\sigma^2_{y0} + \sigma^2_y t_1)(\sigma^2_{y1}t_k + \sigma^2_y) - \sigma^2_{y0}\sigma^2_{y1}t_1}, \tag{13}$$

$$\rho_{y,k} = \frac{-\sigma_{y0}\sigma_{y1}\sqrt{t_1}}{\sqrt{(\sigma^2_{y0} + \sigma^2_y t_1)(\sigma^2_{y1}t_k + \sigma^2_y)}}. \tag{14}$$

Letting x and y denote the model with IID errors and the model with Brownian errors, respectively, we can use $S_{c,1:k} = \{S_{c,1}, S_{c,2}, \ldots, S_{c,k}\}$ to denote the observed degradation data of the corresponding degradation model, where $c \in \{x, y\}$. The joint posterior distribution of parameters with the observed degradation data $S_{c,1:k}$ can be expressed as $\theta'_c, \beta'_c|S_{c,1:k} \sim N(\mu_{\theta'_c,k}, \sigma^2_{\theta'_c,k}, \mu_{\beta'_c,k}, \sigma^2_{\beta'_c,k}, \rho_{c,k})$.

Based on precise priori-parameter distribution assumptions, the traditional Bayesian parameter-updating method can obtain accurate posterior distributions for these unknown model parameters. However, in online prognostics for newly manufactured systems, prior-parameter distribution is usually uncertain. Therefore, the accuracy of the posterior parameters with respect to the updated model is negatively affected, which, in turn, can result in inaccurate RUL prediction. In order to reduce the influence of uncertain prior information and computational complexity, we adopt a Bayesian-updated ECM algorithm to update the model parameters based on the observed degradation data $S_{c,1:k}$.

2.3 Bayesian-updated ECM algorithm

An ECM algorithm replaces the complicated M-step of the EM algorithm with several computationally simpler CM-steps [24, 25]. Therefore, the ECM algorithm is more attractive than the EM with respect to parameter estimation, particularly when the associated complete data maximum likelihood estimation is complicated. However, the ECM involves multiple iterative computations, which are time-consuming. In order to address this problem, a Bayesian-updated ECM algorithm is adopted, which can obtain optimal parameter estimation through an N-step calculating process for an N-dimensional parameter vector.

Let $\boldsymbol{\Theta}_c = [\sigma^2_c, \mu'_{c0}, \mu'_{c1}, \sigma^2_{c0}, \sigma^2_{c1}]$ denote the unknown prior parameters of candidate models, $\boldsymbol{\Theta}_{c,k} = [\sigma^2_{c,k}, \mu'_{c0,k}, \mu'_{c1,k}, \sigma^2_{c0,k}, \sigma^2_{c1,k}]$ denote the estimated parameters based on the observed degradation data $S_{c,1:k}$, and $\hat{\boldsymbol{\Theta}}^{(i)}_{c,k} = [\hat{\sigma}^{2(i)}_{c,k}, \hat{\mu}'^{(i)}_{c0,k}, \hat{\mu}'^{(i)}_{c1,k}, \hat{\sigma}^{2(i)}_{c0,k}, \hat{\sigma}^{2(i)}_{c1,k}]$ denote the i-th iteration result of $\boldsymbol{\Theta}_{c,k}$ in the ECM algorithm. Based on the candidate models, θ'_c and β'_c are selected as hidden variables and $\boldsymbol{\Theta}_{c,k}$ is treated as the objective parameter vector. Accordingly, the E-step of the ECM algorithm can be expressed as

$$\mathrm{l}(\boldsymbol{\Theta}_{c,k}|\hat{\boldsymbol{\Theta}}^{(i)}_{c,k}) = E_{\theta'_c,\beta'_c|S_{c,1:k},\hat{\boldsymbol{\Theta}}^{(i)}_{c,k}}\{\ln p(S_{c,1:k}, \theta'_c, \beta'_c|\boldsymbol{\Theta}_{c,k})\}, \tag{15}$$

where $E_{\theta'_c,\beta'_c|S_{c,1:k},\hat{\boldsymbol{\Theta}}^{(i)}_{c,k}}$ is the conditional expectation of hidden variables given $\hat{\boldsymbol{\Theta}}^{(i)}_{c,k}$ and $S_{c,1:k}$. Furthermore, the CM-step could be written as

$$\begin{aligned}
\hat{\boldsymbol{\Theta}}^{(i+1/N)}_{c,k} &= \arg\max_{\boldsymbol{\Theta}_c} \mathrm{l}\left(\boldsymbol{\Theta}_{c,k}|\hat{\boldsymbol{\Theta}}^{(i)}_{c,k} = \left[\hat{\sigma}^{2(i)}_{c,k}, \hat{\mu}'^{(i)}_{c0,k}, \hat{\mu}'^{(i)}_{c1,k}, \hat{\sigma}^{2(i)}_{c0,k}, \hat{\sigma}^{2(i)}_{c1,k}\right]\right), \\
\hat{\boldsymbol{\Theta}}^{(i+2/N)}_{c,k} &= \arg\max_{\boldsymbol{\Theta}_c} \mathrm{l}\left(\boldsymbol{\Theta}_{c,k}|\hat{\boldsymbol{\Theta}}^{(i+1/N)}_{c,k} = \left[\hat{\sigma}^{2(i+1)}_{c,k}, \hat{\mu}'^{(i)}_{c0,k}, \hat{\mu}'^{(i)}_{c1,k}, \hat{\sigma}^{2(i)}_{c0,k}, \hat{\sigma}^{2(i)}_{c1,k}\right]\right), \\
\hat{\boldsymbol{\Theta}}^{(i+3/N)}_{c,k} &= \arg\max_{\boldsymbol{\Theta}_c} \mathrm{l}\left(\boldsymbol{\Theta}_{c,k}|\hat{\boldsymbol{\Theta}}^{(i+2/N)}_{c,k} = \left[\hat{\sigma}^{2(i+1)}_{c,k}, \hat{\mu}'^{(i+1)}_{c0,k}, \hat{\mu}'^{(i)}_{c1,k}, \hat{\sigma}^{2(i)}_{c0,k}, \hat{\sigma}^{2(i)}_{c1,k}\right]\right), \\
\hat{\boldsymbol{\Theta}}^{(i+4/N)}_{c,k} &= \arg\max_{\boldsymbol{\Theta}_c} \mathrm{l}\left(\boldsymbol{\Theta}_{c,k}|\hat{\boldsymbol{\Theta}}^{(i+3/N)}_{c,k} = \left[\hat{\sigma}^{2(i+1)}_{c,k}, \hat{\mu}'^{(i+1)}_{c0,k}, \hat{\mu}'^{(i+1)}_{c1,k}, \hat{\sigma}^{2(i)}_{c0,k}, \hat{\sigma}^{2(i)}_{c1,k}\right]\right), \\
\hat{\boldsymbol{\Theta}}^{(i+1)}_{c,k} &= \arg\max_{\boldsymbol{\Theta}_c} \mathrm{l}\left(\boldsymbol{\Theta}_{c,k}|\hat{\boldsymbol{\Theta}}^{(i+4/N)}_{c,k} = \left[\hat{\sigma}^{2(i+1)}_{c,k}, \hat{\mu}'^{(i+1)}_{c0,k}, \hat{\mu}'^{(i+1)}_{c1,k}, \hat{\sigma}^{2(i+1)}_{c0,k}, \hat{\sigma}^{2(i)}_{c1,k}\right]\right),
\end{aligned} \tag{16}$$

where $N=5$ is the dimension of vector $\boldsymbol{\Theta}_c$, and $\hat{\boldsymbol{\Theta}}_{c,k}^{(i+1)}$ is the maximum value of $\boldsymbol{\Theta}_c$ based on $\hat{\boldsymbol{\Theta}}_{c,k}^{(i+(N-1)/N)}$ from the derivative equation $\partial \mathrm{l}(\Theta_{c,k}|\hat{\Theta}_{c,k}^{(i+(N-1)/N)})/\partial\hat{\sigma}_{c1,k}^{2(i)}=0$.

In order to reduce computational costs, we modify the E-step in (15). The conditional distribution, $(\theta'_c,\beta'_c|S_{c,1:k},\hat{\boldsymbol{\Theta}}_{c,k}^{(i)})$ in (15) is replaced by the posterior distribution $(\theta'_c,\beta'_c|p(\theta'_c,\beta'_c|S_{c,1:k}))$ in (5)–(14). Accordingly, the modified E-step can be rewritten as

$$\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))=E_{\theta'_c,\beta'_c|p(\theta'_c,\beta'_c|S_{c,1:k})}\{\ln p(S_{c,1:k},\theta'_c,\beta'_c|\boldsymbol{\Theta}_{c,k})\}. \tag{17}$$

Furthermore, the corresponding CM-step can be modified as

$$\begin{aligned}
\hat{\boldsymbol{\Theta}}_{c,k}^{(1/5)}&=\arg\max_{\boldsymbol{\Theta}_c}\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))=[\sigma_{c,k}^2,\mu'_{c0,k},\mu'_{c1,k},\sigma_{c0,k}^2,\sigma_{c1,k}^2]),\\
\hat{\boldsymbol{\Theta}}_{c,k}^{(2/5)}&=\arg\max_{\boldsymbol{\Theta}_c}\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))^{(1/5)}=[\hat{\sigma}_{c,k}^2,\mu'_{c0,k},\mu'_{c1,k},\sigma_{c0,k},\sigma_{c1,k}]),\\
\hat{\boldsymbol{\Theta}}_{c,k}^{(3/5)}&=\arg\max_{\boldsymbol{\Theta}_c}\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))^{(2/5)}=[\hat{\sigma}_{c,k}^2,\hat{\mu}'_{c0,k},\mu'_{c1,k},\sigma_{c0,k},\sigma_{c1,k}]),\\
\hat{\boldsymbol{\Theta}}_{c,k}^{(4/5)}&=\arg\max_{\boldsymbol{\Theta}_c}\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))^{(3/5)}=[\hat{\sigma}_{c,k}^2,\hat{\mu}'_{c0,k},\hat{\mu}'_{c1,k},\sigma_{c0,k},\sigma_{c1,k}]),\\
\hat{\boldsymbol{\Theta}}_{c,k}&=\arg\max_{\boldsymbol{\Theta}_c}\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))^{(4/5)}=[\hat{\sigma}_{c,k}^2,\hat{\mu}'_{c0,k},\hat{\mu}'_{c1,k},\hat{\sigma}_{c0,k},\sigma_{c1,k}]).
\end{aligned} \tag{18}$$

Based on (18), with the observed degradation data $S_{c,1:k}$, the optimal prior parameters of the candidate models $\hat{\boldsymbol{\Theta}}_{c,k}=[\hat{\sigma}_{c,k}^2,\hat{\mu}'_{c0,k},\hat{\mu}'_{c1,k},\hat{\sigma}_{c0,k},\hat{\sigma}_{c1,k}]$ can be calculated. The specific calculation steps for each candidate model are shown below.

(a) Calculation for model M_x. The log-likelihood function of observed $S_{x,1:k}$ should be expressed as

$$\begin{aligned}
&\ln p(S_{x,1:k},\theta'_x,\beta'_x|\boldsymbol{\Theta}_{x,k})\\
&=\ln p(S_{x,1:k}|\theta'_x,\beta'_x,\boldsymbol{\Theta}_{x,k})+\ln p(\theta'_x,\beta'_x|\boldsymbol{\Theta}_{x,k})\\
&=-\frac{k+2}{2}\ln 2\pi-\frac{k}{2}\ln\sigma_{x,k}^2-\sum_{j=1}^{k}\frac{(S_{xj}-\theta'_x-\beta'_x t_j)^2}{2\sigma_{xk}^2}-\frac{1}{2}\ln\sigma_{x0,k}^2\\
&\quad-\frac{1}{2}\ln\sigma_{x1,k}^2-\frac{(\theta'_x-\mu'_{x0,k})^2}{2\sigma_{x0,k}^2}-\frac{(\beta'_x-\mu'_{x1,k})^2}{2\sigma_{x1,k}^2}.
\end{aligned} \tag{19}$$

Accordingly, the modified E-step is

$$\begin{aligned}
&\mathrm{l}(\boldsymbol{\Theta}_{x,k}|p(\theta'_x,\beta'_x|S_{x,1:k}))\\
&=E_{\theta'_x,\beta'_x|p(\theta'_x,\beta'_x|S_{x,1:k})}\{\ln p(S_{x,1:k},\theta'_x,\beta'_x|\boldsymbol{\Theta}_{x,k})\}\\
&=-\frac{k+2}{2}\ln 2\pi-\frac{k}{2}\ln\sigma_{x,k}^2-\frac{1}{2}\ln\sigma_{x0,k}^2-\frac{1}{2}\ln\sigma_{x1,k}^2\\
&-\frac{\mu_{\theta'_x,k}^2+\sigma_{\theta'_x,k}^2-2\mu_{\theta'_x,k}\mu'_{x0,k}+\mu'^2_{x0,k}}{2\sigma_{x0,k}^2}-\frac{\mu_{\beta'_x,k}^2+\sigma_{\beta'_x,k}^2-2\mu_{\beta'_x,k}\mu'_{x1,k}+\mu'^2_{x1,k}}{2\sigma_{x1,k}^2}\\
&-\sum_{j=1}^{k}\frac{S_{x,j}^2-2S_{x,j}(\mu_{\theta'_x,k}+\mu_{\beta'_x,k}t_j)+\mu_{\theta'_x,k}^2+\sigma_{\theta'_x,k}^2+2t_j(\rho_{x,k}\sigma_{\theta'_x,k}\sigma_{\beta'_x,k}+\mu_{\theta'_x,k}\mu_{\beta'_x,k})+t_j^2(\mu_{\beta'_x,k}^2+\sigma_{\beta'_x,k}^2)}{2\sigma_{x,k}^2}.
\end{aligned} \tag{20}$$

Based on the CM-step in (18), we can compute the optimal prior parameters $\hat{\boldsymbol{\Theta}}_{x,k}$ as

$$\begin{aligned}
&\hat{\sigma}_{x,k}^2=\frac{1}{k}\sum_{j=1}^{k}\left[\begin{array}{l}S_{x,j}^2-2S_{x,j}(\mu_{\theta'_x,k}+\mu_{\beta'_x,k}t_j)+\mu_{\theta'_x,k}^2+\sigma_{\theta'_x,k}^2\\+2t_j(\rho_{x,k}\sigma_{\theta'_x,k}\sigma_{\beta'_x,k}+\mu_{\theta'_x,k}\mu_{\beta'_x,k})+t_j^2(\mu_{\beta'_x,k}^2+\sigma_{\beta'_x,k}^2)\end{array}\right],\\
&\hat{\mu}'_{x0,k}=\mu_{\theta'_x,k},\quad \hat{\sigma}_{x0,k}^2=\sigma_{\theta'_x,k}^2,\quad \hat{\mu}'_{x1,k}=\mu_{\beta'_x,k},\quad \hat{\sigma}_{x1,k}^2=\sigma_{\beta'_x,k}^2.
\end{aligned} \tag{21}$$

(b) Calculation for model M_y. The log-likelihood function of the total observed data $S_{y,1:k}$ can be written as

$$\begin{aligned}
&\ln p(S_{y,1:k},\theta'_y,\beta'_y|\boldsymbol{\Theta}_{y,k})\\
&\quad=\ln p(S_{y,1:k}|\theta'_y,\beta'_y,\boldsymbol{\Theta}_{y,k})+\ln p(\theta'_y,\beta'_y|\boldsymbol{\Theta}_{y,k})\\
&\quad=-\frac{k+2}{2}\ln 2\pi-\frac{k}{2}\ln\sigma^2_{y,k}-\frac{1}{2}\ln\sigma^2_{y0,k}-\frac{1}{2}\ln\sigma^2_{y1,k}\\
&\quad\quad-\frac{(S_{y,1}-\theta'_y-\beta'_yt_1)^2}{2\sigma^2_{y,k}t_1}-\sum_{i=2}^{k}\frac{(S_{y,i}-\beta'_y(t_i-t_{i-1}))^2}{2\sigma^2_{y,k}(t_i-t_{i-1})}-\frac{(\theta'_y-\mu'_{y0,k})^2}{2\sigma^2_{y0,k}}-\frac{(\beta'_y-\mu'_{y1,k})^2}{2\sigma^2_{y1,k}}.
\end{aligned}\tag{22}$$

Accordingly, the E-step can be modified as

$$\begin{aligned}
&\mathrm{l}(\boldsymbol{\Theta}_{y,k}|p(\theta'_y,\beta'_y|S_{y,1:k}))\\
&\quad=E_{\theta'_y,\beta'_y}|p(\theta'_y,\beta'_y|S_{y,1:k})\{\ln p(S_{y,1:k},\theta'_y,\beta'_y|\boldsymbol{\Theta}_{y,k})\}\\
&\quad=-\frac{k+2}{2}\ln 2\pi-\frac{k}{2}\ln\sigma^2_{y,k}-\frac{1}{2}\ln\sigma^2_{y0,k}-\frac{1}{2}\ln\sigma^2_{y1,k}\\
&\quad\quad-\frac{\mu^2_{\theta'_y,k}+\sigma^2_{\theta'_y,k}-2\mu_{\theta'_y,k}\mu'_{y0,k}+\mu'^2_{y0,k}}{2\sigma^2_{y0,k}}-\frac{\mu^2_{\beta'_y,k}+\sigma^2_{\beta'_y,k}-2\mu_{\beta'_y,k}\mu'_{y1,k}+\mu'^2_{y1,k}}{2\sigma^2_{y1,k}}\\
&\quad\quad-\frac{S^2_{y,1}-2S_{y,1}(\mu_{\theta'_y,k}+\mu_{\beta'_y,k}t_1)+\mu^2_{\theta'_y,k}+\sigma^2_{\theta'_y,k}+2t_1(\rho_{y,k}\sigma_{\theta'_y,k}\sigma_{\beta'_y,k}+\mu_{\theta'_y,k}\mu_{\beta'_y,k})+t_1^2(\mu^2_{\beta'_y,k}+\sigma^2_{\beta'_y,k})}{2\sigma^2_{y,k}t_1}\\
&\quad\quad-\sum_{i=2}^{k}\frac{S^2_{y,i}-2S_{y,i}\mu_{\beta'_y,k}(t_i-t_{i-1})+(\mu^2_{\beta'_y,k}+\sigma^2_{\beta'_y,k})(t_i-t_{i-1})^2}{2\sigma^2_{y,k}(t_i-t_{i-1})}.
\end{aligned}\tag{23}$$

Based on (18), we obtain the optimal $\hat{\boldsymbol{\Theta}}_{y,k}$ as

$$\begin{aligned}
&\hat\sigma^2_{x,k}=\frac{1}{k}\left[\begin{array}{l}\frac{S^2_{y,1}-2S_{y,1}(\mu_{\theta'_y,k}+\mu_{\beta'_y,k}t_1)+\mu^2_{\theta'_y,k}+\sigma^2_{\theta'_y,k}+2t_1(\rho_{y,k}\sigma_{\theta'_y,k}\sigma_{\beta'_y,k}+\mu_{\theta'_y,k}\mu_{\beta'_y,k})+t_1^2(\mu^2_{\beta'_y,k}+\sigma^2_{\beta'_y,k})}{t_1}\\ +\sum_{i=2}^{k}\frac{S^2_{y,i}-2S_{y,i}\mu_{\beta'_y,k}(t_i-t_{i-1})+(\mu^2_{\beta'_y,k}+\sigma^2_{\beta'_y,k})(t_i-t_{i-1})^2}{(t_i-t_{i-1})}\end{array}\right],\\
&\hat\mu'_{y0,k}=\mu_{\theta'_y,k},\quad \hat\sigma^2_{y0,k}=\sigma^2_{\theta'_y,k},\quad \hat\mu'_{y1,k}=\mu_{\beta'_y,k},\quad \hat\sigma^2_{y1,k}=\sigma^2_{\beta'_y,k}.
\end{aligned}\tag{24}$$

Theorem 1. According to the results of (21) and (24), the estimated optimal prior parameters $\hat{\boldsymbol{\Theta}}_{c,k}$ for model M_x or model M_y, consist of the only maximum point of the derivative equation $\partial\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k})/\partial\boldsymbol{\Theta}_{c,k}=0$.

Proof. First, from the E-step in (20) or (23), we can infer that $\hat{\boldsymbol{\Theta}}_{c,k}$ is the only solution of the derivative equation $\partial\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k})/\partial\boldsymbol{\Theta}_{c,k}=0$. Next, we prove that the estimated $\hat{\boldsymbol{\Theta}}_{c,k}$ in (21) or (24) maximizes the function $\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k})$. Based on the expectation function in (20) or (23), the second derivative on $\boldsymbol{\Theta}_{c,k}$ can be expressed as

$$\frac{\partial^2\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta'_c,\beta'_c|S_{c,1:k}))}{\partial\boldsymbol{\Theta}_{c,k}\partial\boldsymbol{\Theta}^{\mathrm{T}}_{c,k}}=\begin{bmatrix}\frac{k}{2\sigma^4_{c,k}}-\frac{\tau_c}{\sigma^6_{c,k}} & 0 & 0 & 0 & 0\\ & -\frac{1}{\sigma^2_{c0,k}} & 0 & \frac{\mu'_{c0,k}-\mu_{\theta'_c,k}}{\sigma^4_{c0,k}} & 0\\ & & -\frac{1}{\sigma^2_{c1,k}} & 0 & \frac{\mu'_{c1,k}-\mu_{\beta'_c,k}}{\sigma^4_{c1,k}}\\ & \text{symmetrical} & & \frac{1}{2\sigma^4_{c0,k}}-\frac{\psi_{c1}}{\sigma^6_{0c,k}} & 0\\ & & & & \frac{1}{2\sigma^4_{c1,k}}-\frac{\psi_{c2}}{\sigma^6_{c1,k}}\end{bmatrix},\tag{25}$$

where

$$\tau_{c=x}=\sum_{j=1}^{k}\left[\begin{array}{l}S_{x,j}^2-2S_{x,j}(\mu_{\theta_x',k}+\mu_{\beta_x',k}t_j)+\mu_{\theta_x',k}^2+\sigma_{\theta_x',k}^2\\+2t_j(\rho_{x,k}\sigma_{\theta_x',k}\sigma_{\beta_x',k}+\mu_{\theta_x',k}\mu_{\beta_x',k})+t_j^2(\mu_{\beta_x',k}^2+\sigma_{\beta_x',k}^2)\end{array}\right],$$

$$\tau_{c=y}=\frac{S_{y,1}^2-2S_{y,1}(\mu_{\theta_y',k}+\mu_{\beta_y',k}t_1)+\mu_{\theta_y',k}^2+\sigma_{\theta_y',k}^2+2t_1(\rho_{y,k}\sigma_{\theta_y',k}\sigma_{\beta_y',k}+\mu_{\theta_y',k}\mu_{\beta_y',k})+t_1^2(\mu_{\beta_y',k}^2+\sigma_{\beta_y',k}^2)}{t_1}$$
$$+\sum_{i=2}^{k}\frac{S_{y,i}^2-2S_{y,i}\mu_{\beta_y',k}(t_i-t_{i-1})+(\mu_{\beta_y',k}^2+\sigma_{\beta_y',k}^2)(t_i-t_{i-1})^2}{(t_i-t_{i-1})},$$

$$\psi_{c1}=\mu_{\theta_c',k}^2+\sigma_{\theta_c',k}^2-2\mu_{\theta_c',k}\mu_{c0,k}'+{\mu'}_{c0,k}^2,$$

$$\psi_{c2}=\mu_{\beta_c',k}^2+\sigma_{\beta_c',k}^2-2\mu_{\beta_c',k}\mu_{c1,k}'+{\mu'}_{c1,k}^2.$$

Based on the obtained optimal value of $\hat{\boldsymbol{\Theta}}_{c,k}=[\hat{\sigma}_{c,k}^2,\hat{\mu}_{c0,k}',\hat{\mu}_{c1,k}',\hat{\sigma}_{c0,k},\hat{\sigma}_{c1,k}]$ in (21) or (24), the order of principal minor determinants with respect to the second derivative matrix in (25) can be obtained as follows:

$$\Delta_1|_{\boldsymbol{\Theta}_{c,k}=\hat{\boldsymbol{\Theta}}_{c,k}}<0,\quad \Delta_2|_{\boldsymbol{\Theta}_{c,k}=\hat{\boldsymbol{\Theta}}_{c,k}}>0,\quad \Delta_3|_{\boldsymbol{\Theta}_{c,k}=\hat{\boldsymbol{\Theta}}_{c,k}}<0,\quad \Delta_4|_{\boldsymbol{\Theta}_{c,k}=\hat{\boldsymbol{\Theta}}_{c,k}}>0,\quad \Delta_5|_{\boldsymbol{\Theta}_{c,k}=\hat{\boldsymbol{\Theta}}_{c,k}}<0.$$

The results show that Eq. (25) is a negative-definite matrix when $\boldsymbol{\Theta}_{c,k}=\hat{\boldsymbol{\Theta}}_{c,k}$. Therefore, $\hat{\boldsymbol{\Theta}}_{c,k}=[\hat{\sigma}_{c,k}^2,\hat{\mu}_{c0,k}',\hat{\mu}_{c1,k}',\hat{\sigma}_{c0,k},\hat{\sigma}_{c1,k}]$ in (21) or (24) is the only maximum point of $\mathrm{l}(\boldsymbol{\Theta}_{c,k}|p(\theta_c',\beta_c'|S_{c,1:k}))$.

3 RUL prediction with modified Bayesian-model-averaging method

Based on the obtained parameter distributions using the Bayesian-updated ECM algorithm, we can predict the RUL of the target system. However, owing to indistinct degradation characteristics and insufficient model-selection criteria, the optimal model for the observed degradation data may be ambiguous. Thus, this paper proposes a modified Bayesian-model-averaging method in order to address model uncertainty. Accordingly, based on the probability of each candidate model and the estimated model parameters, the RUL can be obtained.

3.1 Modified Bayesian-model-averaging method

Bayesian model averaging is an effective method of addressing model uncertainties [26, 27]. Therefore, in order to obtain an adaptive model probability for each candidate model for the online prognostics, a modified Bayesian-model-averaging method is proposed. We use M_x and M_y to denote the candidate models with IID errors and Brownian errors, respectively. Let RUL_k denote the RUL of the target system with the observed data $S_{c,1:k}$, at time t_k. Considering Bayesian model averaging, RUL_k can be calculated as

$$\mathrm{RUL}_k=\sum\nolimits_c P(M_c|S_{c,1:k})\mathrm{RUL}_{c,k},\tag{26}$$

where $P(M_c|S_{c,1:k})$ is the posterior probability of model M_c, and $\mathrm{RUL}_{c,k}$ is the RUL of the target system when M_c is the correct degradation model. In original Bayesian model averaging, $P(M_c|S_{c,1:k})$ is the weight of the model average when $S_{c,1:k}$ is obtained, which can be calculated using

$$P(M_c|S_{c,1:k})=\frac{P(S_{c,1:k}|M_c)P(M_c)}{\sum P(S_{c,1:k}|M_c)P(M_c)},\tag{27}$$

where $P(S_{c,1:k}|M_c)$ is the integrated likelihood of model M_c, and $P(M_c)$ is the assumed prior probability of model M_c. Letting $\boldsymbol{\Phi}_c=[\theta_c',\beta_c']$, the integrated likelihood $P(S_{c,1:k}|M_c)$ can be expressed as

$$P(S_{c,1:k}|M_c)=\int L(S_{c,1:k}|\boldsymbol{\Phi}_c,M_c)P(\boldsymbol{\Phi}_c|M_c)\mathrm{d}\boldsymbol{\Phi}_c,\tag{28}$$

where $L(S_{c,1:k}|\boldsymbol{\Phi}_c,M_c)$ denotes the likelihood function of model M_c, and $P(\boldsymbol{\Phi}_c|M_c)$ is the prior-probability density of $\boldsymbol{\Phi}_c$ when M_c is the true model.

However, in online prognosis, the parameters in (26)–(28) need to be updated once $S_{c,k}$ is observed. In order to solve this problem, a modified Bayesian-model-averaging method is proposed, and Eq. (27) is modified as

$$P_A(M_c|S_{c,1:k}) = \frac{P(S_{c,1:k}|M_c)P_A(M_c|S_{c,1:k-1})}{\sum P(S_{c,1:k}|M_c)P_A(M_c|S_{c,1:k-1})}, \tag{29}$$

where $P_A(M_c|S_{c,1:k})$ is the adaptive model probability of model M_c at time $t = k$ with the degradation data $S_{c,1:k}$. At time $t = 1$, letting the initial value $P_A(M_c|S_{c,1}) = P(M_c)$, $P_A(M_c|S_{c,1:k})$ can be updated when $S_{c,k}$ is observed. Let $\boldsymbol{\Phi}_{c,k} = [\theta'_{c,k}, \beta'_{c,k}]$ denote the corresponding model parameters at time $t = k$ with the data $S_{c,1:k}$. In the same way, the integrated likelihood $P(S_{c,1:k}|M_c)$ of model M_c is modified as

$$P(S_{c,1:k}|M_c) = \int L(S_{c,1:k}|\boldsymbol{\Phi}_{c,k-1}, M_c)P(\boldsymbol{\Phi}_{c,k-1}|M_c)\mathrm{d}\boldsymbol{\Phi}_{c,k-1}, \tag{30}$$

where $P(\boldsymbol{\Phi}_{c,k-1}|M_c)$ is the conditional prior density of $\boldsymbol{\Phi}_{c,k-1}$ when M_c is the true model, which can be obtained using the Bayesian-updated ECM algorithm in (21) and (24). Based on (2) and (4), the likelihood functions $L(S_{x,1:k}|\boldsymbol{\Phi}_{x,k-1}, M_x)$ and $L(S_{y,1:k}|\boldsymbol{\Phi}_{y,k-1}, M_y)$ can be written as

$$L(S_{x,1:k}|\theta'_{x,k-1}, \beta'_{x,k-1}) = \left(\frac{1}{\sqrt{2\pi\sigma^2_{x,k-1}}}\right)^k \exp\left(-\sum_{i=1}^{k}\frac{(S_{x,i} - \theta'_{x,k-1} - \beta'_{x,k-1}t_i)^2}{2\sigma^2_{x,k-1}}\right), \tag{31}$$

$$L(S_{y,1:k}|\theta'_{y,k-1}, \beta'_{y,k-1}) = \left(\frac{1}{\sqrt{2\pi\sigma^2_{y,k-1}}}\right)^k \exp\left(\begin{array}{l} -\frac{(S_{y,1}-\theta'_{y,k-1}-\beta'_{y,k-1}t_1)^2}{2\sigma^2_{y,k-1}t_1} \\ -\sum\limits_{i=2}^{k}\left(\frac{(S_{y,i}-\beta'_{y,k-1}(t_i-t_{i-1}))^2}{2\sigma^2_{y,k-1}(t_i-t_{i-1})}\right)\end{array}\right). \tag{32}$$

Owing to the complexities of (30), this paper adopts a large-sample approximation method in order to numerically calculate the adaptive model probability $P_A(M_c|S_{c,1:k})$ in (29) at time $t = k$. The detailed steps are outlined below.

Step 1. Based on $P(\boldsymbol{\Phi}_{c,k-1}|M_c)$, simulate $\boldsymbol{\Phi}^r_{c,k-1}$ from the corresponding distribution $\pi(\boldsymbol{\Phi}_{c,k-1})$ for W_c times, where $r = 1, 2, \ldots, W_c$ and $W_c \geqslant 500000$.

Step 2. According to (29) and (30), the adaptive model probability $P_A(M_c|S_{c,1:k})$ is proportional to the integrated likelihood $P(S_{c,1:k}|M_c)$, which can be approximated as

$$P(S_{c,1:k}|M_c) = \frac{1}{W_c}\sum_{r=1}^{W_c} L(S_{x,1:k}|\boldsymbol{\Phi}^r_{c,k-1}, M_c). \tag{33}$$

Owing to $P_A(M_c|S_{c,1:k-1}) \propto P(S_{c,1:k}|M_c) \cdot P_A(M_c|S_{c,1:k-1})$, we obtain

$$\mathrm{Sum}_c = \left(\frac{1}{W_c}\sum_{r=1}^{W_c} L(S_{x,1:k}|\boldsymbol{\Phi}^r_{c,k-1}, M_c)\right) \cdot P_A(M_c|S_{c,1:k-1}). \tag{34}$$

Accordingly, the adaptive model probability can be calculated by

$$P_A(M_c|S_{c,1:k}) = \frac{\mathrm{Sum}_c}{\sum_c \mathrm{Sum}_c}. \tag{35}$$

3.2 RUL calculation

Based on the posterior distribution of the parameters, $\mathrm{RUL}_{c,k}$ of model M_c can be computed. For model M_x, given $S_{x,1:k}$, the degradation $S_x(t_k + T)$ at time $t = t_k + T$ should follow the distribution:

$$\begin{aligned} &S_x(T + t_k)|S_{x,1:k} \sim N(\tilde{\mu}_x(T + t_k), \tilde{\sigma}^2_x(T + t_k)), \\ &\tilde{\mu}_x(T + t_k) = \mu_{\theta'_x,k} + \mu_{\beta'_x,k}(T + t_k) - \sigma^2_{x,k}/2, \\ &\tilde{\sigma}^2_x(T + t_k) = \sigma^2_{\theta'_x,k} + \sigma^2_{\beta'_x,k}(T + t_k)^2 + \sigma^2_{x,k} + 2\rho_{x,k}(T + t_k)\sigma_{\theta'_x,k}\sigma_{\beta'_x,k}. \end{aligned} \tag{36}$$

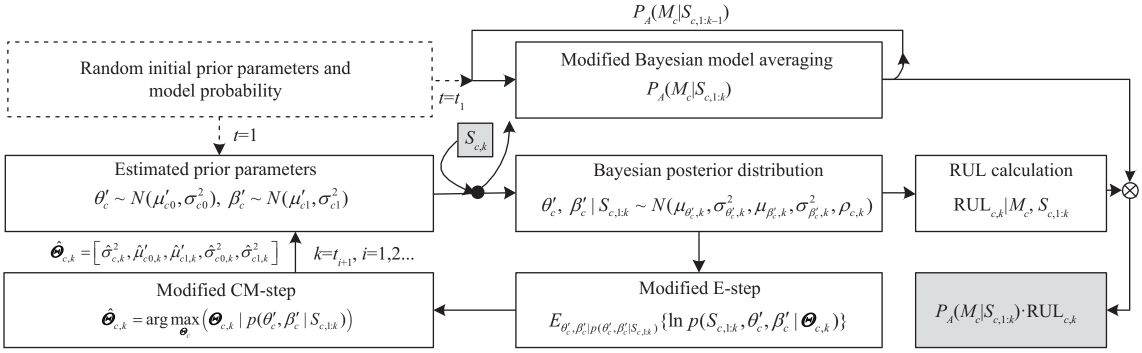

Figure 1 The algorithm of RUL calculation for model M_c.

For the model M_y, the conditional degradation $S_y(t_k + T)|S_{y,1:k}$ at time $t = t_k + T$ follows

$$
\begin{aligned}
&S_y(T + t_k)|S_{y,1:k} \sim N(\tilde{\mu}_y(T + t_k), \tilde{\sigma}^2_y(T + t_k)),\\
&\tilde{\mu}_y(T + t_k) = \sum_{i=1}^{k} S_{y,i} + \mu_{\beta'_y} T = S_y(t_k) + \mu_{\beta'_y,k} T,\\
&\tilde{\sigma}^2_y(T + t_k) = \sigma^2_{\beta'_y,k} T^2 + \sigma^2_{y,k} T.
\end{aligned}
\tag{37}
$$

Assuming that w is the failure threshold, and based on the failure equation $S_c(T + t_k) = \ln w$, $\mathrm{RUL}_{c,k}$ can be computed. For simplification, we replace the real degradation with the mean value in (36) and (37). Letting $\tilde{\mu}_c(T + t_k) = \ln w$, $\mathrm{RUL}_{c,k}$ can be approximated as

$$
\mathrm{RUL}_{x,k} = \frac{\ln w - \mu_{\theta'_x,k} + \sigma^2_{x,k}/2}{\mu_{\beta'_x,k}} - t_k, \tag{38}
$$

$$
\mathrm{RUL}_{y,k} = \frac{\ln w - S_y(t_k)}{\mu_{\beta'_y,k}}. \tag{39}
$$

Accordingly, based on the adaptive model probability obtained through the modified Bayesian-model-averaging method, we can calculate the RUL_k of the target system using

$$
\mathrm{RUL}_k = \sum\nolimits_c P_A(M_c|S_{c,1:k}) \cdot \mathrm{RUL}_{c,k} = \mathrm{RUL}_{x,k} \cdot P_A(M_x|S_{x,1:k}) + \mathrm{RUL}_{y,k} \cdot P_A(M_y|S_{y,1:k}). \tag{40}
$$

Figure 1 shows the algorithm of the RUL calculation for each candidate model $P_A(M_c|S_{c,1:k}) \cdot \mathrm{RUL}_{c,k}$.

4 Numerical experiment

4.1 Simulation study

The objective of the proposed fusion algorithm is to address both parameter uncertainty and model uncertainty. Therefore, we discuss the effectiveness of the Bayesian-updated ECM algorithm with respect to parameter estimation and the effectiveness of the modified Bayesian-model-averaging method with respect to model-probability computing.

4.1.1 *Effectiveness of Bayesian-updated ECM algorithm*

In order to verify the effectiveness of the Bayesian-updated ECM algorithm, we adopt a method to estimate the parameters of specific degradation models. Based on statistical data, we select a set of parameters for the candidate models. The distributions of the selected parameters are given in Table 1.

Based on Table 1, degradation trajectories with IID errors and Brownian errors can be obtained, all of which possess similar monotonicity, concavity, and convexity. In order to demonstrate the Bayesian-updated ECM algorithm, we adopt the distributions in Table 2 to estimate model parameters, and, for the purposes of comparison, we adopt the Bayesian method in [28].

Table 1 Distributions of the selected simulation parameters

	Simulation parameter	Simulation distribution
Exponential degradation with IID errors	θ'_x	$N(0.02, 1\times10^{-6})$
	β'_x	$N(0.01, 1\times10^{-6})$
	σ_x^2	0.1
Exponential degradation with Brownian errors	θ'_y	$N(0.02, 1\times10^{-6})$
	β'_y	$N(0.01, 1\times10^{-6})$
	σ_y^2	0.0004

Table 2 Prior-distribution assumption for Bayesian-updated ECM algorithm

Model	Parameter	Bayesian method [28]	Our method		
			Distribution 1	Distribution 2	Distribution 3
M_x	θ'_x	$N(0.04, 1\times10^{-6})$	$N(0.04, 1\times10^{-6})$	$N(0.1, 1\times10^{-6})$	$N(0.2, 1\times10^{-6})$
	β'_x	$N(0.02, 1\times10^{-6})$	$N(0.02, 1\times10^{-6})$	$N(0.05, 1\times10^{-6})$	$N(0.1, 1\times10^{-6})$
	σ_x^2	0.01	0.01	0.01	0.01
M_y	θ'_y	$N(0.04, 1\times10^{-6})$	$N(0.04, 1\times10^{-6})$	$N(0.1, 1\times10^{-6})$	$N(0.2, 1\times10^{-6})$
	β'_y	$N(0.02, 1\times10^{-6})$	$N(0.02, 1\times10^{-6})$	$N(0.05, 1\times10^{-6})$	$N(0.1, 1\times10^{-6})$
	σ_y^2	0.004	0.004	0.004	0.004

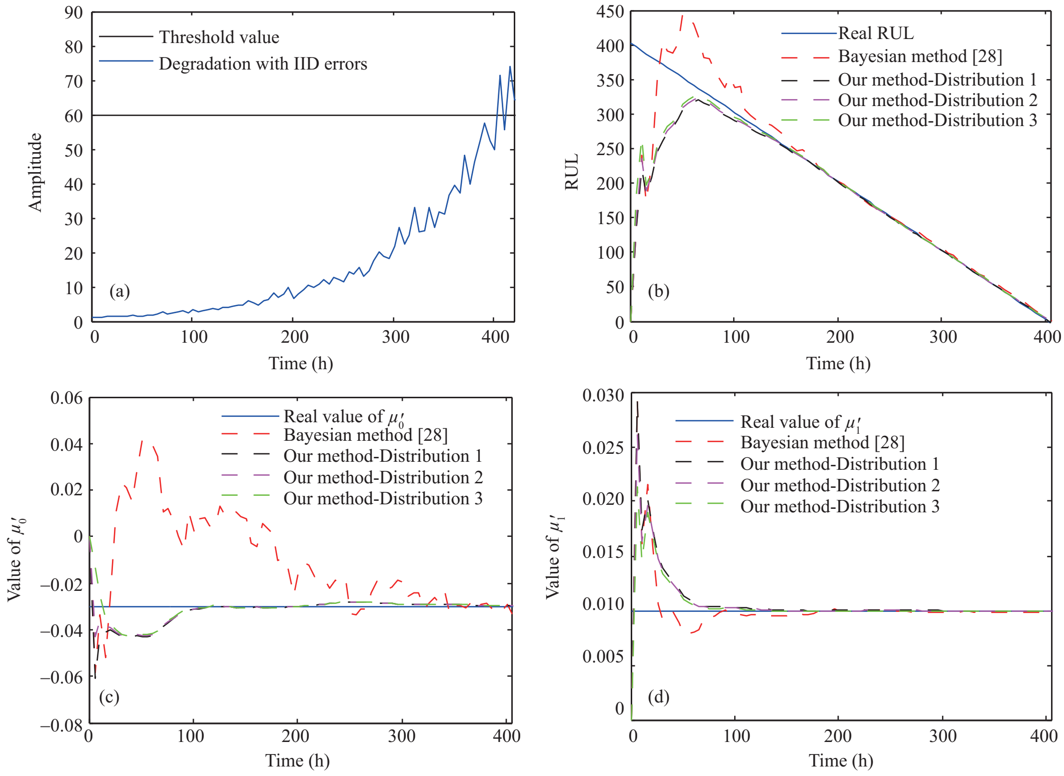

Figure 2 (Color online) Results of Bayesian-updated ECM algorithm for model M_x. (a) Simulated trajectory with IID errors; (b) predicted RUL of model M_x; (c) estimated μ'_{x0} of model M_x; (d) estimated μ'_{x1} of model M_x.

Figures 2 and 3 show the results of model M_x for the degradation trajectory with IID errors and the results of model M_y for the degradation trajectory with Brownian errors, respectively.

Figure 2(a) shows one simulated degradation trajectory with IID errors; the system fails at roughly $t = 402$. In Figure 2(b), for this group of degradation trajectory IID errors, model M_x can obtain an accurate RUL at roughly $t = 175$ using the Bayesian method. Indeed, compared to the Bayesian method, the Bayesian-updated ECM algorithm performs better with different prior-parameter distributions and can obtain an accurate RUL at roughly $t = 65$. Figures 2(c) and (d) show that the estimated μ'_{x0} and μ'_{x1} of the Bayesian-updated ECM algorithm converge to real values at roughly $t = 90$ and $t = 65$, respectively.

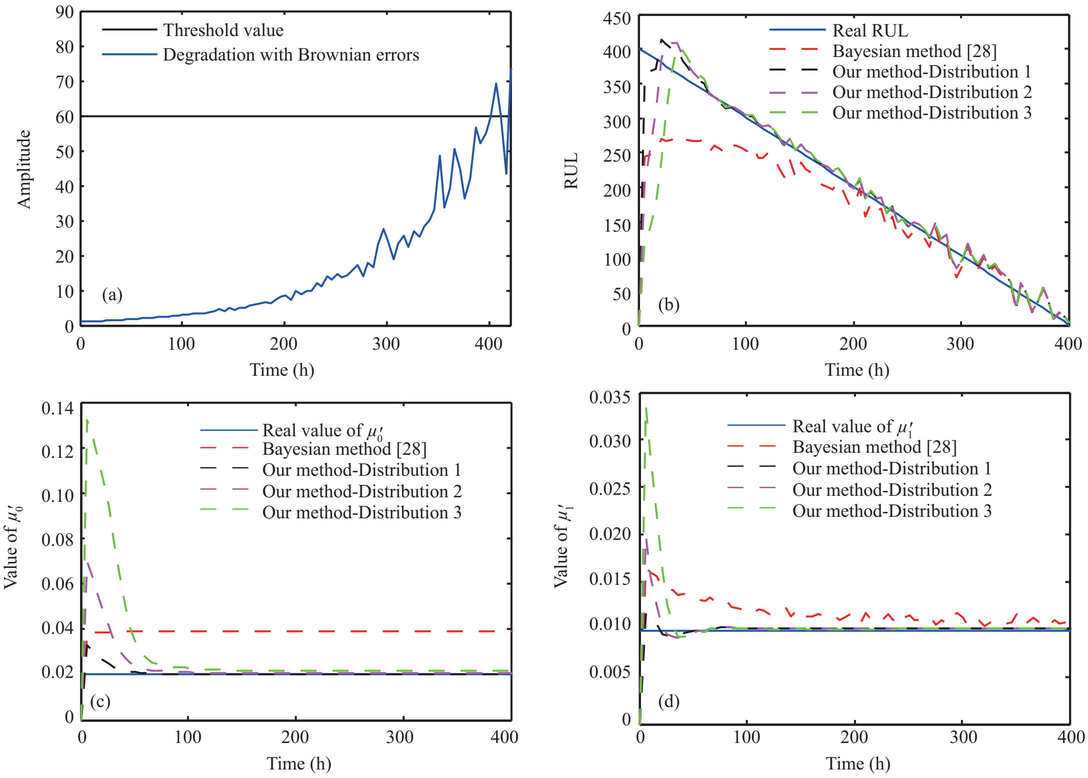

Figure 3 (Color online) Results of Bayesian-updated ECM algorithm for model M_y. (a) Simulated trajectory with Brownian errors; (b) predicted RUL of model M_y; (c) estimated μ'_{y0} of model M_y; (d) estimated μ'_{y1} of model M_y.

Figure 3(a) shows the simulated degradation trajectory with Brownian errors; the system fails at roughly $t = 400$. Figure 3(b) suggests that the Bayesian-updated ECM algorithm performs well with different prior-parameter distributions and the estimated RULs converge to the real RUL at roughly $t = 60$, whereas the estimated results of the Bayesian method converge to the real RUL at roughly $t = 200$. Figures 3(c) and (d) show that the estimated μ'_{y0} and μ'_{y1} of the Bayesian-updated ECM algorithm converge to the real values at roughly $t = 75$ and $t = 30$, respectively.

Based on the above results, the conclusion outlined below can be drawn.

(1) When adopting the proper degradation model to predict the RUL of the target system, both the Bayesian-updated ECM algorithm and Bayesian method give accurate results based on sufficient degradation data. This means that both algorithms are effective in processing online monitoring data.

(2) Compared to the Bayesian method, the Bayesian-updated ECM algorithm has better convergence with real RUL under the same initial prior-parameter distribution. This indicates that the Bayesian-updated ECM algorithm can obtain more precise degradation parameters in the initial degradation period.

(3) The Bayesian-updated ECM algorithm can obtain accurate μ'_{c0} and μ'_{c1}, which are the core parameters of RUL prediction.

(4) Although different initial prior distributions are adopted, the results estimated by the Bayesian-updated ECM algorithm converge to the real RUL in a relatively short period of time. This means that the Bayesian-updated ECM algorithm effectively addresses the uncertainty of initial prior-parameter distribution, and that it can obtain an accurate RUL when the selected degradation model is appropriate for the observed monitoring data.

4.1.2 *Effectiveness of modified Bayesian-model-averaging method*

In order to verify the effectiveness of the modified Bayesian-model-averaging method, we compare it with the original method. The simulated trajectories in Figures 2(a) and 3(a) are adopted as the objective degradation paths. Distribution 1 in Table 2 is adopted to estimate the model parameters, and the prior-model probabilities are set as follows: (1) $P(M_x) = 1/3$ and $P(M_y) = 2/3$, (2) $P(M_x) = 1/2$ and $P(M_y) = 1/2$, and (3) $P(M_x) = 2/3$ and $P(M_y) = 1/3$.

Figure 4 suggests that the results of model M_y cannot converge with the real RUL for the degradation

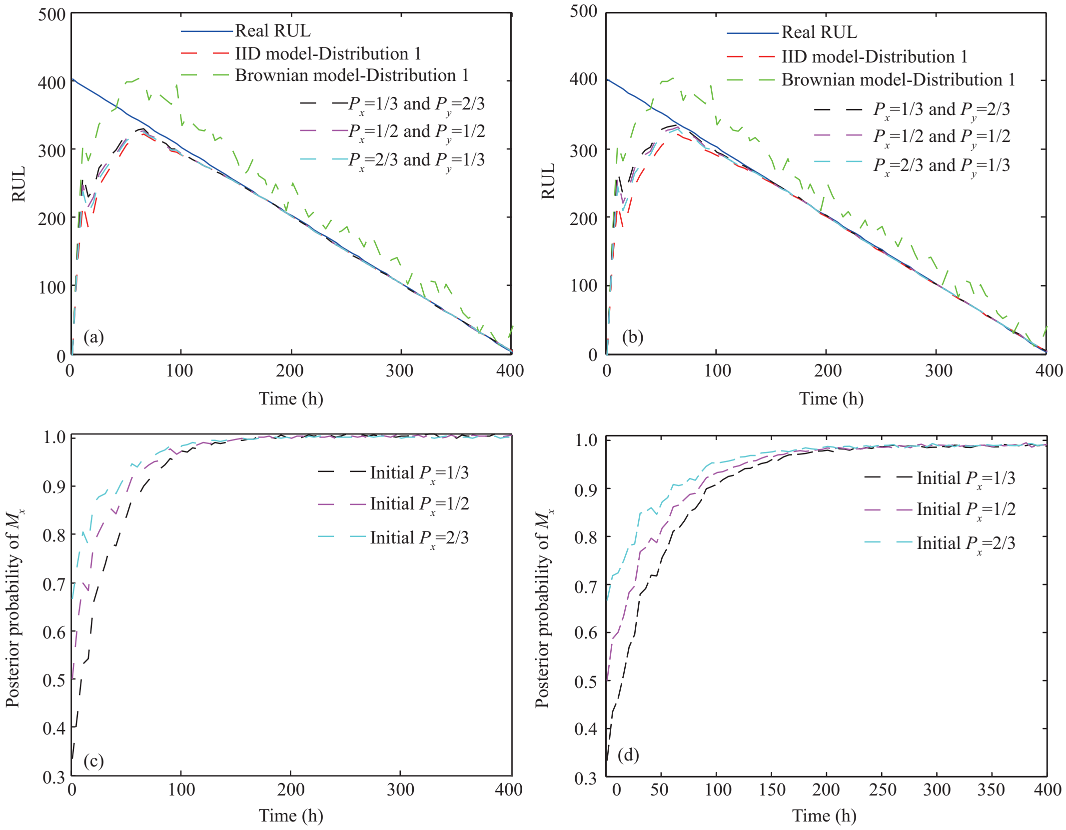

Figure 4 (Color online) Effectiveness of modified Bayesian-model-averaging method for trajectory with IID errors. (a) Modified Bayesian model averaging for trajectory with IID errors; (b) original Bayesian model averaging for trajectory with IID errors; (c) the obtained model probability of the modified Bayesian model averaging for model M_x; (d) the obtained model probability of the original Bayesian model averaging for model M_x.

trajectory with IID errors. Moreover, the RUL of the modified Bayesian-model-averaging method in Figure 4(a) converges with the predicted RUL of M_x at roughly $t = 125$; however, the original Bayesian model averaging in Figure 4(b) converges with predicted RUL of M_x at roughly $t = 150$. Figure 4 suggests that the obtained model probabilities of the modified Bayesian model averaging and the original Bayesian model averaging can converge, but the results of the former perform better than the latter.

Figure 5 shows that the RUL of model M_x cannot converge with the real RUL with respect to the degradation trajectory with Brownian errors. Indeed, both the modified Bayesian model averaging and the original Bayesian model averaging converge with the real RUL; however, the modified Bayesian model averaging has better convergence than the original. Figures 5(c) and (d) show that the obtained model probabilities of the modified Bayesian model averaging converge at roughly $t = 75$; however, the results of the original Bayesian model averaging converge at roughly $t = 130$.

Based on the above results, the conclusion outlined below can be drawn.

(1) Unfaithful models may lead to inaccurate RUL-prediction results. Therefore, in order to ensure a proper degradation model is obtained, it is essential to consider the model-selection process.

(2) Both the modified and original Bayesian-model-averaging methods effectively address model uncertainty; moreover, the inaccurate degradation model slightly effects the RUL-prediction results of the averaged results.

(3) Compared to the original Bayesian-model-averaging method, the modified Bayesian-model-averaging method has a higher convergence speed; moreover, it addresses the uncertainty of initial model probability better than the original with respect to the online RUL prediction.

4.2 Case study

Gyroscopes are vital devices with respect to the inertial platform of strategic missile inertial navigation systems (INSs). When the inertial platform is operating, high-speed rotating of the gyroscope can result in rotation-axis deterioration. Indeed, this will eventually result in the deformation of the gyroscope's

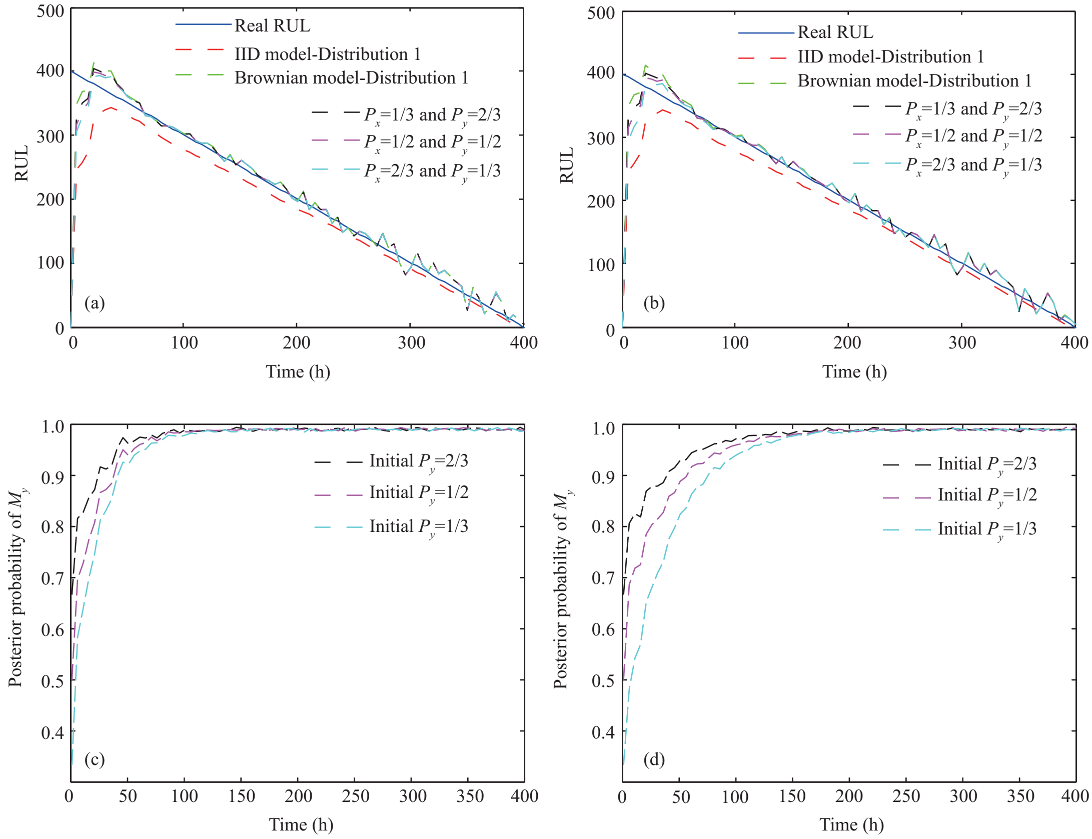

Figure 5 (Color online) Effectiveness of modified Bayesian model averaging for trajectory with Brownian errors. (a) Modified Bayesian model averaging for trajectory with Brownian errors; (b) original Bayesian model averaging for trajectory Brownian errors; (c) obtained model probability of modified Bayesian model averaging for model M_y; (d) obtained model probability of original Bayesian model averaging for model M_y.

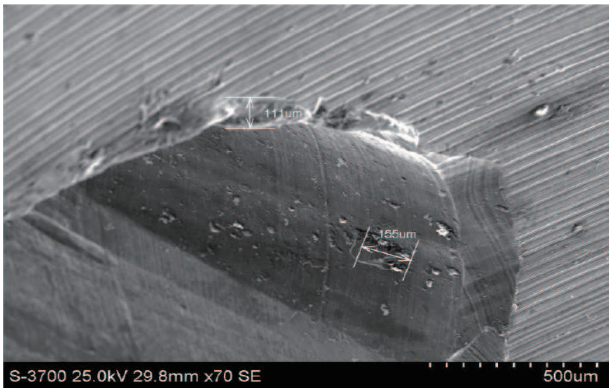

Figure 6 Illustration of a deformed motor bearing.

motor bearing, which will result in gyroscope drift as well as failure of the gyroscope and the inertial platform [29]. Statistical analysis shows that gyroscope failure is largely the result of the bearings; moreover, gyroscope drift is responsible for 70.

In Figure 6, a deformed motor bearing of a gyroscope is shown, which is obtained using a scanning electron microscopy S-3700N. The maximum length of the metal flake is 155 µm, which is the result of deformation. Indeed, deformations such as these are reflected by the gyroscope drift.

In our experiment, the monitoring interval is 2.5 h and the failure threshold is 0.37°/h. In practice, this threshold is strictly enforced because INS devices are critical in navigated weapon systems. In order to illustrate the validity of our method, we adopt different prior-model parameters and prior-model probabilities in order to predict the RUL of the observed gyroscope in Figure 7(a). With the prior-model

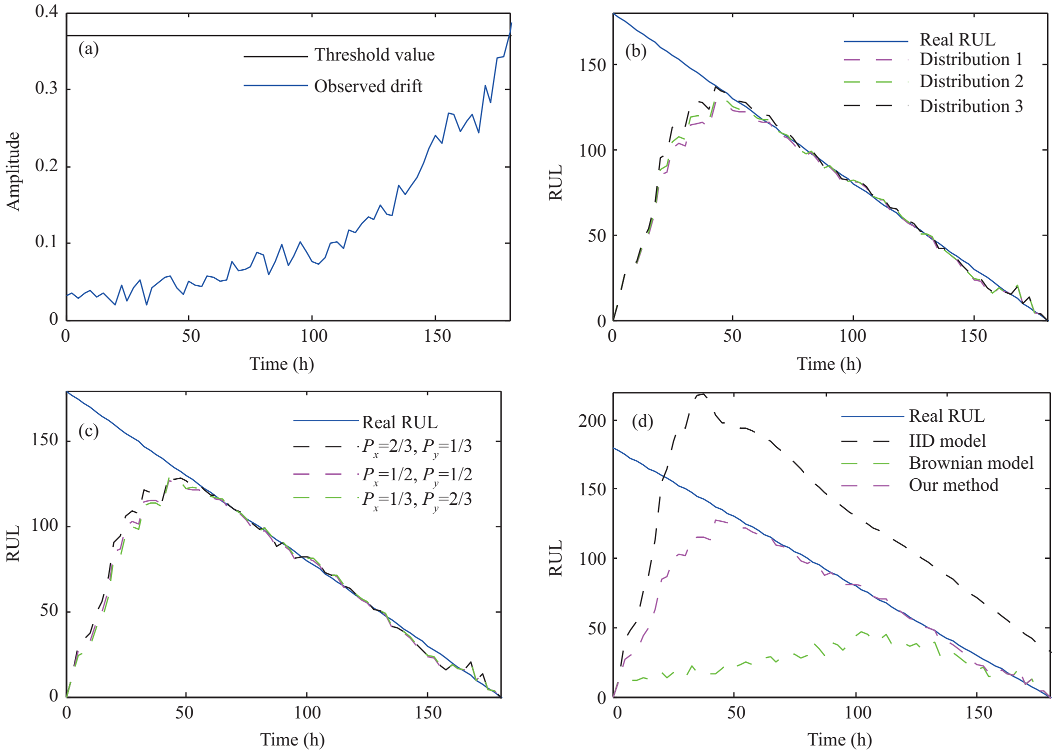

Figure 7 (Color online) Effectiveness of the proposed algorithm. (a) The observed drift data of the gyroscope; (b) estimated RUL of different prior parameters; (c) estimated RUL of different prior probabilities; (d) estimated RUL of our method.

Table 3 Prior-distribution assumption for observed drift data

Model	Parameter	Distribution 1	Distribution 2	Distribution 3
	θ'_x	$N(0.0025, 1\times10^{-6})$	$N(0.005, 1\times10^{-6})$	$N(0.0075, 1\times10^{-6})$
M_x	β'_x	$N(0.001, 1\times10^{-6})$	$N(0.002, 1\times10^{-6})$	$N(0.003, 1\times10^{-6})$
	σ_x^2	0.01	0.01	0.01
	θ'_y	$N(0.0025, 1\times10^{-6})$	$N(0.005, 1\times10^{-6})$	$N(0.0075, 1\times10^{-6})$
M_y	β'_y	$N(0.001, 1\times10^{-6})$	$N(0.002, 1\times10^{-6})$	$N(0.003, 1\times10^{-6})$
	σ_y^2	0.00004	0.00004	0.00004

probability assumption $P(M_x) = 1/2$ and $P(M_y) = 1/2$, Figure 7(b) shows the results of different prior-model parameters that are shown in Table 3. With Distribution 1 in Table 3, Figure 7(c) shows the results of different prior-model probabilities. With the prior-model probability assumption $P(M_x) = 1/2$ and $P(M_y) = 1/2$ and Distribution 1 in Table 3, Figure 7(d) shows the predicted RUL results of our method, the single model M_x, and the single model M_y.

Figure 7(a) shows the observed gyroscope drift; the inertial platform fails at $t = 180$ h. Figures 7(b) and (c) show that our method can effectively address prior-parameter uncertainty and model uncertainty: the estimated RULs of different prior-parameters and different prior-model probabilities converge to the real RUL at roughly $t = 60$. In Figure 7(d), although the single model M_x and the single model M_y could not effectively address the observed drift data of the gyroscope, the proposed method can obtain an accurate RUL based on the advantages of the Bayesian-updated ECM algorithm and the modified Bayesian-model-averaging method. Indeed, this indicates that the proposed fusion method should be adopted when model uncertainty and parameter uncertainty both exist.

5 Conclusion

In this paper, we studied the problem of online RUL prediction for newly manufactured systems that lack historical-degradation data. The method for RUL prediction was based on two typical exponential models. In particular, we examined posterior-parameter estimation with prior-parameter uncertainty in online RUL prognostics, as well as online RUL prediction with degradation-model uncertainty for newly manufactured systems using real-time degradation data. With respect to the former, a Bayesian-updated ECM algorithm was proposed by combining an ECM algorithm and the traditional Bayesian updating method. In our parameter-estimating method, the prior parameters were updated by the ECM algorithm each time a new degradation value was observed. This implies that our method can reduce prior-parameter uncertainty, thereby ensuring RUL-estimation accuracy. With respect to the latter, we used a modified Bayesian-model-averaging method that relies on the prior-parameter-updating process. Combining the updated prior parameters of the degradation model, the prior-probability term of the original Bayesian-model-averaging method was updated. This implies that the degradation-model uncertainty can be processed by updating the prior parameters. Indeed, combining the advantages of Bayesian-updated ECM and the modified Bayesian-model-averaging method, our fusion algorithm can obtain an accurate RUL result for newly manufactured systems that lack historical-degradation data. Overall, this paper analyzes the uncertainties of prior-parameter distribution and prior-model probability with two typical exponential models. Future work can and should extend this study to stochastic-process degradation models, such as Wiener, Gamma, and inverse Gaussian process models.

Acknowledgements This work was supported by National Key R&D Program of China (Grant No. 2018YFB1306100) and National Natural Science Foundation of China (Grant Nos. 61922089, 61773386, 61833016, 61903376, 61673311).

References

1 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
2 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
3 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
4 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
5 Xi J, Wang C, Liu H, et al. Completely distributed guaranteed-performance consensualization for high-order multiagent systems with switching topologies. IEEE Trans Syst Man Cybern Syst, 2019, 49: 1338–1348
6 Si X S, Li T M, Zhang Q, et al. An optimal condition-based replacement method for systems with observed degradation signals. IEEE Trans Rel, 2018, 67: 1281–1293
7 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, 2019. doi: 10.1109/TIE.2019.2947839
8 Zhang Z, Si X, Hu C, et al. Degradation data analysis and remaining useful life estimation: a review on Wiener-process-based methods. Eur J Oper Res, 2018, 271: 775–796
9 Hudak S J J, Saxena A, Bucci R J, et al. Development of Standard Methods of Testing and Analyzing Fatigue Crack Growth Rate Data. Westinghouse R&D Center Westinghouse Electric Corporation, 1978
10 Meeker W Q, Escobar L A. Statistical Methods for Reliability Data. New York: John Wiley & Sons, Inc., 1998
11 Ye Z S, Xie M, Tang L C, et al. Semiparametric estimation of Gamma processes for deteriorating products. Technometrics, 2014, 56: 504–513
12 Li Z J, Deng Y M, Mastrangelo C. Model selection for degradation-based Bayesian reliability analysis. J Manufacturing Syst, 2015, 37: 72–82
13 Peng C-Y, Tseng S-T. Mis-specification analysis of linear degradation models. IEEE Trans Rel, 2009, 58: 444–455
14 Peng C-Y, Tseng S-T. Statistical lifetime inference with Skew-Wiener linear degradation models. IEEE Trans Rel, 2013, 62: 338–350
15 Ramos A A. The minimum description length principle and model selection in spectropolarimetry. Astrophys J, 2006, 646: 1445–1451
16 Bozdogan H. Akaike's information criterion and recent developments in information complexity. J Math Psychol, 2000, 44: 62–91
17 Nguyen K T P, Fouladirad M, Grall A. Model selection for degradation modeling and prognosis with health monitoring data. Reliab Eng Syst Saf, 2018, 169: 105–116
18 Abdessalem A B, Dervilis N, Wagg D, et al. Model selection and parameter estimation in structural dynamics using approximate Bayesian computation. Mech Syst Signal Process, 2018, 99: 306–325
19 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 Oper Res, 2013, 226: 53–66
20 Meng X L, Rubin D B. Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika, 1993, 80: 267–278

21 Chen N, Tsui K L. Condition monitoring and remaining useful life prediction using degradation signals: revisited. IIE Trans, 2013, 45: 939–952

22 Chakraborty S, Gebraeel N, Lawley M, et al. Residual-life estimation for components with non-symmetric priors. IIE Trans, 2009, 41: 372–387

23 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

24 Nitithumbundit T, Chan J S K. An ECM algorithm for skewed multivariate variance Gamma distribution in normal mean-variance representation. 2015. arXiv: 1504.01239

25 van Dyk D A, Meng X, Rubin D B. Maximum likelihood estimation via the ECM algorithm: computing the asymptotic variance. Stat Sin, 1995, 5: 55–75

26 Zhang X, Yan X. A new statistical precipitation downscaling method with Bayesian model averaging: a case study in China. Clim Dyn, 2015, 45: 2541–2555

27 Hoeting J A, Madigan D, Raftery A E, et al. Bayesian model averaging: a tutorial. Stat Sci, 1999, 14: 382–401

28 Gebraeel N, Elwany A, Pan J. Residual life predictions in the absence of prior degradation knowledge. IEEE Trans Rel, 2009, 58: 106–117

29 Si X S, Chen M Y, Wang W, et al. Specifying measurement errors for required lifetime estimation performance. Eur J Oper Res, 2013, 231: 631–644