

Research Article

Degradation Data-Driven Remaining Useful Life Estimation in the Absence of Prior Degradation Knowledge

Yong Yu,^{1,2} Changhua Hu,¹ Xiaosheng Si,¹ and Jianxun Zhang¹

¹Department of Automation, Xi'an Institute of High-Technology, Xi'an, Shaanxi 710025, China

²Institute No. 25, The Second Academy of China Aerospace Science and Industry Corporation, Beijing 100854, China

Correspondence should be addressed to Changhua Hu; hch_reu@sina.com

Received 6 August 2017; Accepted 23 October 2017; Published 3 December 2017

Academic Editor: Chunhui Zhao

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

Recent developments in prognostic and health management have been targeted at utilizing the observed degradation signals to estimate residual life distributions. Current degradation models mainly focus on a population of “identical” devices or an individual device with population information, not a single component in the absence of prior degradation knowledge. However, the fast development of science and technology provides us with many kinds of new systems, and we just have the real-time monitoring information to analyze the reliability for them. The fusion algorithm presented herein addresses this challenge by combining the excellent modeling ability of Bayesian updating method for the multilevel data and the prominent estimation ability of ECM algorithm for incomplete data. Residual life distributions and posterior distributions are first calculated through the Bayesian updating method based on random initial a priori distributions. Then the a priori distributions are revised and improved for future predictions by the ECM algorithm. Once a new signal is observed, we can reuse the fusion algorithm to improve the accuracy of residual life distributions. The applicability of this fusion algorithm is validated by a set of simulation experiments.

1. Introduction

Modern engineering systems are overwhelmingly complex because of increasing requirements on their functionalities and qualities. These systems often have a high standard of system reliability because a single failure can lead to catastrophic consequences with profound impacts, extreme costs, and potential safety hazards. It will take an exceedingly long time for a system to fail, so prognostics for systems have become extremely difficult, even if the actual operating conditions are severe and rigorous. Therefore, effective methods that can predict failure progression and evaluate the reliability of the system have long been sought. When there are sufficient monitoring data and efficient computational capability, prognostics for components based on observed degradation data is promising and effective [1–3].

The general path model is a typical method utilizing the observed degradation signals of these degraded systems. Lu and Meeker [4] introduced the model to the degradation literature in 1993, for the first time. In their model, the fixed-effects parameters affect the populations' characteristics

and the random-effects parameters describe an individual unit's characteristics. Once the parameters are known, the residual life is deterministic. Therefore, the core work is to estimate the unknown fixed and random parameters. Lu and Meeker used a two-stage method to estimate the unknown parameters. Lu et al. [5] extended the degradation model and suggested likelihood-based estimation methods. However, these are not suitable for all types of degradation data. Su et al. [6] considered random sample sizes and random repeated measurement times for each product unit. They discussed the advantages and disadvantages of two-stage least-squares (LS) estimation, maximum modified likelihood (MML) estimation, and maximum likelihood estimation (MLE). They showed that the LS estimators are not consistent in the case of random sample sizes. However, MLE can provide a consistent estimator and has smaller biases and variances compared to the LS and MML estimates. In the further study, Weaver et al. [7] also used MLE to estimate the unknown parameters. They extended the research and examined effects of sample size on the estimation precision. Under the mixed-effect path model, Wu and Shao [8] built the asymptotic properties of the

(weighted) least-squares estimators. They used these properties to calculate approximate confidence intervals and point estimates for percentiles of the failure-time distribution. They used the weighted least-squares estimators to predict the resistor of metal film and the metal fatigue crack length.

However, the above papers only focus on the estimation of the unknown parameters about population devices and need a fair amount of samples in the test. In order to solve this problem, Robinson and Crowder [9] described a fully Bayesian approach which allows a small sample size. They used a variety of simple prior distributions and observed that this aspect has little effect on the posterior distributions of these data, showing that the information in the degradation data dominates. So the Bayesian approach is more suitable for the parameters estimation of an individual device with population information compared with two-stage LS estimation and MLE. Gebraeel et al. [10] developed two different exponential degradation signal models. One model assumes that the error fluctuations follow an iid random error process; however, the second model considers that the error terms follow a Brownian motion process. In their paper, they used the Bayesian updating methods to combine the distribution information of the parameters across the population and the monitored degradation data from the individual device. The Bayesian updating methods can update the stochastic parameters of degradation models, every time a new degradation signal comes. Gebraeel [11] extended the Bayesian updating procedure by assuming that the stochastic parameters in the exponential degradation models follow a bivariate normal distribution. Chakraborty et al. [12] further extended the updating procedure and investigated the difference of the life time distributions when the stochastic parameters do not follow the normal distribution. They also built methods for calculating Remaining Useful Life when the stochastic parameters of the exponential model follow more general distributions. Chen and Tsui [13] adopted a piecewise log-linear degradation model and assumed the change time of the two different phases is random. This assumption explicitly accounts for the characteristics of different degradation phases. When new observations were available, they updated the a posteriori information of the model stochastic parameters including regression coefficients and the variance of the error term by using Bayesian methods. They also suggested a new method which took the correlations into consideration, among degradation predictions, to compute the RUL distribution with better accuracy. Their approach can be naturally extended to more general degradation models.

The above Bayesian approaches can be applied to predicting the RUL of an individual device with population information. However, the fast development of science and technology provides us with kinds of newly made systems, and we just have the observed monitoring degradation information to analyze the reliability of them. Traditional ways to predict system failures often fail because the domain knowledge and expert experience are limited and historical data is nonexistent.

Therefore, effective methods that can predict failures of these newly made systems in the absence of prior degradation

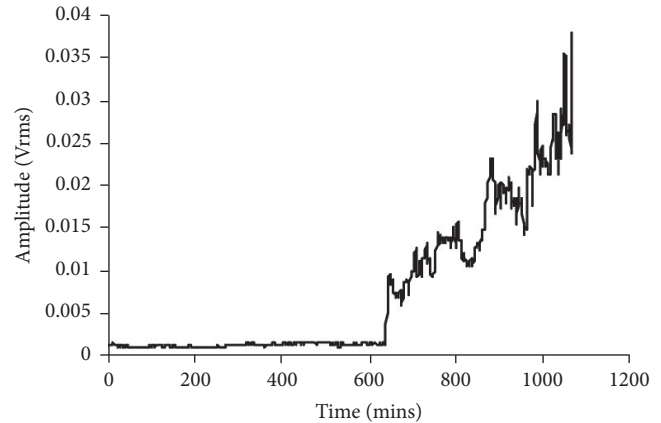


FIGURE 1: Vibration-based degradation signal.

knowledge have long been sought. Considering the excellent modeling ability of Bayesian updating method for the multilevel data and the prominent estimation ability of ECM algorithm for incomplete data, the goal of our paper is to develop a collaboration method between Bayesian updating method and ECM algorithm to estimate the Remaining Useful Life (RUL) of the newly made system just with real-time sensing data. To verify the applicability of this fusion algorithm, a set of simulation experiments are conducted.

The remainder of this paper is organized as follows. Section 2 develops a separate Bayesian updating method and calculates the RUL distributions for the exponential degradation model with a Brownian error term, under the assumption that the prior distributions are known. Section 3 explains the procedure of the fusion process in detail and estimates the prior information for the Bayesian updating method by our fusion algorithm. Section 4 illustrates the validity of the collaboration method by a set of simulation experiments. The paper concludes with some discussions and guidance for the estimation of RUL of the single component without prior degradation knowledge in Section 5.

2. Bayesian Updating Method and Residual Life Distribution

In order to develop our collaboration method, we first introduce the Bayesian updating method and its estimation result of the Remaining Useful Life in this section. We will adopt the exponential degradation signal model developed by Gebraeel et al. [10, 11], Kaiser and Gebraeel [14], and Elwany and Gebraeel [15] and assume that the error fluctuation of the degradation model follows a Brownian motion process. As our objective is to compute the RUL of a single system without prior distribution knowledge, we believe that the adopted exponential degradation model and the error fluctuation are adequate for the given degradation path in Figure 1. For a further discussion of model selection and evaluation, see Li et al. [16].

Under the above assumption, we could use the Bayesian updating procedure to compute the unknown random parameters of the exponential degradation model. Once

we have got the calculated posterior distributions of these random parameters in the exponential degradation model, we can derive the residual life distribution of the component. However, in our paper we only have the real-time monitoring degradation signal for the newly made component, and prior distributions are nonexistent. In order to utilize the Bayesian updating procedure, we first assume we have got the accurate and informative a priori information. The method of estimating these prior distributions will be illustrated and detailed in Section 3.

2.1. The Degradation Signal Model. First, we review the general definition of the Brownian motion process.

Definition 1. A standard Brownian motion process, $W(t)$, $t \geq 0$, possesses the following properties:

1. If $t_0 < t_1 < \dots < t_n$, then $W(t_0), W(t_1) - W(t_0), \dots, W(t_n) - W(t_{n-1})$ are mutually independent.
2. If $s, t \geq 0$, then

$$P(W(s+t) - W(s) \in A) = \int_A (2\pi t)^{-1/2} e^{-x^2/2t} dx. \quad (1)$$

3. With probability one $t \rightarrow W(t)$ is continuous.

The first part of the definition describes the fact that the process $W(t)$ has independent increments. The second part means that the increment $W(s+t) - W(s)$ follows a normal distribution with mean zero and variance t . The third part describes the fact that $W(t)$, $t \geq 0$, almost certainly has continuous paths. For an in-depth discussion about the Brownian motion process and its properties, see Durrett [17].

Then, we let $X(t)$ denote the real-time monitoring signal as a continuous stochastic process, with respect to time t . We define the functional form of $X(t)$ as

$$\begin{aligned} X(t) &= \varphi + \theta \exp\left(\beta t + \varepsilon(t) - \frac{\sigma^2 t}{2}\right) \\ &= \varphi + \theta \exp(\beta t) \exp\left(\varepsilon(t) - \frac{\sigma^2 t}{2}\right), \end{aligned} \quad (2)$$

where φ is the fixed intercept and represents the initial degradation, and β is a normal random variable such that the mean of β is μ_1 and the variance of β is σ_1^2 . θ is a lognormal random variable with mean μ_0 and variance σ_0^2 , and $\varepsilon(t) = \sigma W(t)$ is a Brownian motion with mean zero and variance $\sigma^2 t$. Under the assumption that θ , β , and $\varepsilon(t)$ are mutually independent, it is obvious to obtain that $E[\exp(\varepsilon(t) - (\sigma^2 t/2))] = 1$, and thus $E(S(t) | \theta, \beta) = \varphi + \theta \exp(\beta t)$.

Furthermore, we find that it is easy to calculate with the logarithmic degradation data. Thus, we define $S(t)$ as follows:

$$S(t) = \ln(X(t) - \varphi) = \ln \theta + \beta t + \varepsilon(t) - \frac{\sigma^2 t}{2}. \quad (3)$$

By defining $\theta' = \ln \theta$, $\beta' = \beta - \sigma^2/2$, we can further simplify $S(t)$ as follows:

$$S(t) = \theta' + \beta' t + \varepsilon(t). \quad (4)$$

2.2. Bayesian Updating Process of Stochastic Parameters. We let $S_k = S(t_k) - S(t_{k-1})$, $k = 2, 3, \dots$ denote the difference value of the monitored logged degradation signal at times t_k and t_{k-1} , with $S_1 = S(t_1)$.

Then, suppose we have obtained the logged difference value $S_{1:k} = \{S_1, S_2, \dots, S_k\}$ at times t_1, \dots, t_k . And the error increments, $\varepsilon(t_i) - \varepsilon(t_{i-1})$, $i = 2, \dots, k$, are independent normal random variables. If the stochastic parameters, θ' and β' , are given, we can define the conditional joint density function of $S_{1:k}$ as

$$\begin{aligned} p(S_{1:k} | \theta', \beta') &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^k \\ &\cdot \exp\left(-\sum_{i=2}^k \left(\frac{(S_i - \theta' - \beta' t_i)^2}{2\sigma^2 t_i} + \frac{(S_i - \beta' (t_i - t_{i-1}))^2}{2\sigma^2 (t_i - t_{i-1})}\right)\right). \end{aligned} \quad (5)$$

Generally, however, θ' and β' will be unknown. Based on the former assumption, we suppose we have got accurate and informative priors. And we let $\pi_0(\theta')$ and $\pi_1(\beta')$ denote the prior distributions on θ' and β' , respectively, where $\pi_0(\theta') = N(\mu_0, \sigma_0^2)$ and $\pi_1(\beta') = N(\mu_1', \sigma_1'^2)$, $\mu_1' = \mu_1 - \sigma^2/2$. Then, given the logged difference data, $S_{1:k}$, obtained at times t_1, \dots, t_k , the posterior joint distribution of (θ', β') can be expressed as follows.

Theorem 2. Given the obtained logged degradation signal, $S_{1:k}$, the conditional posterior joint distribution of (θ', β') is $\theta', \beta' | S_{1:k} \sim N(\mu_{\theta',k}, \sigma_{\theta',k}^2, \mu_{\beta',k}, \sigma_{\beta',k}^2, \rho_k)$, where

$$\begin{aligned} \mu_{\theta',k} &= \frac{(S_1 \sigma_0^2 + \mu_0 \sigma^2 t_1)(\sigma_1^2 t_k + \sigma^2) - \sigma_0^2 t_1 (\sigma_1^2 \sum_{i=1}^k S_i + \mu_1' \sigma^2)}{(\sigma_0^2 + \sigma^2 t_1)(\sigma_1^2 t_k + \sigma^2) - \sigma_0^2 \sigma_1^2 t_1}, \\ \mu_{\beta',k} &= \frac{(\sigma_1^2 \sum_{i=1}^k S_i + \mu_1' \sigma^2)(\sigma^2 t_1 + \sigma_0^2) - \sigma_1^2 (S_1 \sigma_0^2 + \mu_0 \sigma^2 t_1)}{(\sigma_0^2 + \sigma^2 t_1)(\sigma_1^2 t_k + \sigma^2) - \sigma_0^2 \sigma_1^2 t_1}, \\ \sigma_{\theta',k}^2 &= \frac{\sigma^2 \sigma_0^2 t_1 (\sigma_1^2 t_k + \sigma^2)}{(\sigma_0^2 + \sigma^2 t_1)(\sigma_1^2 t_k + \sigma^2) - \sigma_0^2 \sigma_1^2 t_1}, \\ \sigma_{\beta',k}^2 &= \frac{\sigma^2 \sigma_1^2 (\sigma_0^2 + \sigma^2 t_1)}{(\sigma_0^2 + \sigma^2 t_1)(\sigma_1^2 t_k + \sigma^2) - \sigma_0^2 \sigma_1^2 t_1}, \\ \rho_k &= \frac{-\sigma_0 \sigma_1 \sqrt{t_1}}{\sqrt{(\sigma_0^2 + \sigma^2 t_1)(\sigma_1^2 t_k + \sigma^2)}}. \end{aligned} \quad (6)$$

Proof. Given the prior distributions of θ' and β' , $\pi_0(\theta') = N(\mu_0, \sigma_0^2)$ and $\pi_1(\beta') = N(\mu_1', \sigma_1'^2)$, we can express the conditional posterior joint distribution of (θ', β') , $p(\theta', \beta' | S_{1:k})$, as follows:

$$\begin{aligned}
p(\theta', \beta' | S_{1:k}) &\propto p(S_{1:k} | \theta', \beta') \pi_0(\theta') \pi_1(\beta') \\
&\propto \frac{1}{2\pi\sigma_{\theta',k}\sigma_{\beta',k}\sqrt{1-\rho_k^2}} \cdot \exp \left[-\frac{1}{2(1-\rho_k^2)} \right. \\
&\quad \cdot \left. \begin{pmatrix} \frac{(\theta' - \mu_{\theta',k})^2}{\sigma_{\theta',k}^2} \\ -2\rho_k \frac{(\theta' - \mu_{\theta',k})(\beta' - \mu_{\beta',k})}{\sigma_{\theta',k}\sigma_{\beta',k}} \\ + \frac{(\beta' - \mu_{\beta',k})^2}{\sigma_{\beta',k}^2} \end{pmatrix} \right]. \quad (7)
\end{aligned}$$

As our focus is on the fusion algorithm, for a further detailed calculation of the posterior distribution of (θ', β') , see Gebraeel et al. [10]. \square

2.3. Estimation of the Residual Life Distribution. Every time a new degradation signal comes, we can compute a new posterior distribution of (θ', β') . As the objective of our paper is to estimate the distribution of the RUL of the monitored system, we suppose that the system's failure occurs when the observed degradation signal reaches the failure threshold, w , and thus we need to estimate the time until the degradation signal reaches w . In our paper, we assume that the threshold value is a constant value.

The objective of prognostics is to compute the distribution of the failure time until the degradation signal reaches the threshold w . To achieve this goal, we first calculate the posterior distribution of (θ', β') . Then, we let the random variable $S(t + t_k)$ denote the logged degradation signal value obtained at time $t + t_k$, $t > 0$, given $S_{1:k}$ obtained at times t_1, \dots, t_k . Under the above assumption, the distribution of $S(t + t_k)$ given $S_{1:k}$ can be expressed as follows.

Theorem 3. *Given the observed difference value of the logged degradation data, $S_{1:k}$, the distribution of $S(t + t_k)$ is $S(t + t_k) | S_{1:k} \sim N(\tilde{\mu}(t + t_k), \tilde{\sigma}^2(t + t_k))$, where*

$$\begin{aligned}
\tilde{\mu}(t + t_k) &\triangleq \sum_{i=1}^k S_i + \mu_{\beta',k}t = S(t_k) + \mu_{\beta',k}t, \\
\tilde{\sigma}^2(t + t_k) &\triangleq \sigma_{\beta',k}^2 t^2 + \sigma^2 t. \quad (8)
\end{aligned}$$

Proof. First, note that $S(t) = \theta' + \beta't + \varepsilon(t)$ and we can express $S(t + t_k) = S(t_k) + \beta't + \varepsilon(t + t_k) - \varepsilon(t_k)$, where $S(t_k) = \sum_{i=1}^k S_i$. Therefore, given $S_{1:k}$, $S(t + t_k)$ follows a normal distribution with mean $\tilde{\mu}(t + t_k) = S(t_k) + E[\beta']t = S(t_k) + \mu_{\beta',k}t$ and

variance $\tilde{\sigma}^2(t + t_k) = t^2 V[\beta'] + V[\varepsilon(t + t_k) - \varepsilon(t_k)] = \sigma_{\beta',k}^2 t^2 + \sigma^2 t$. \square

Then, we let T denote the RUL of the monitored system, and we know that T meets $S(T + t_k) = \ln w$. Given $S_{1:k}$, the conditional cumulative distribution function (CDF) of T , $F_{T|S_{1:k}}(t) = P\{T \leq t | S_{1:k}\}$, can be expressed as

$$\begin{aligned}
P(T \leq t | S_{1:k}) &= 1 - P(S(t + t_k) \leq \ln w | S_{1:k}) \\
&= 1 - P\left(Z < \frac{\ln w - \tilde{\mu}(t + t_k)}{\sqrt{\tilde{\sigma}^2(t + t_k)}}\right) \\
&= P\left(Z \geq \frac{\ln w - \tilde{\mu}(t + t_k)}{\sqrt{\tilde{\sigma}^2(t + t_k)}}\right) \\
&= \Phi(g(t)), \quad (9)
\end{aligned}$$

where Z represents a standard normal variable, $\Phi(\bullet)$ denotes the CDF of the standard normal variable, and $g(t) = (\tilde{\mu}(t + t_k) - \ln w) / \sqrt{\tilde{\sigma}^2(t + t_k)}$.

We compute the residual life distribution of the system at time t_k , under the condition that the observed degradation signal does not reach the threshold w ; that is, $\sum_{i=1}^k S_i = S(t_k) < \ln w$. Thus, we get

$$\lim_{t \rightarrow 0} g(t) = \lim_{t \rightarrow 0} \frac{\sum_{i=1}^k S_i + \mu_{\beta',k}t - \ln w}{\sqrt{\sigma_{\beta',k}^2 t^2 + \sigma^2 t}} = -\infty. \quad (10)$$

Therefore, we get $\lim_{t \rightarrow 0} F_{T|S_{1:k}}(t) = 0$, which means that the domain of the RUL, T , is $(0, \infty)$. We can express the conditional probability distribution function (PDF) of T , given $S_{1:k}$, as

$$f_{T|S_{1:k}, T \geq 0}(t) = \phi(g(t)) g'(t), \quad (11)$$

where $\phi(\bullet)$ denotes the PDF of the standard normal variable.

Given the conditional PDF of T , we can write the expectation of RUL, at time t_k , as

$$E_{T|S_{1:k}, T \geq 0}(t) = \int_0^{\infty} f_{T|S_{1:k}, T \geq 0}(t) dt. \quad (12)$$

In order to simplify the integral of expectation of RUL, we use the failure equation $S(T + t_k) = \ln w$ to compute the RUL, at time t_k , approximately. And according to Theorem 3, given the difference value of logged degradation signal, $S_{1:k}$, the distribution of $S(t + t_k)$ follows $S(t + t_k) | S_{1:k} \sim N(\tilde{\mu}(t + t_k), \tilde{\sigma}^2(t + t_k))$, and the mean of this distribution, $\tilde{\mu}(t + t_k)$, closely approximates $S(t + t_k)$. Thus, we can write $\tilde{\mu}(T + t_k) = \ln w$ and express the RUL, at time t_k , as

$$\text{RUL}_k = \frac{\ln w - S(t_k)}{\mu_{\beta',k}}, \quad (13)$$

where RUL_k is closely approximated to the expectation of RUL.

In other literature, the prior distribution parameters, σ^2 , μ_0 , σ_0^2 , μ_1' , σ_1^2 , are computed from historical monitoring data or derived from domain knowledge and expert experience. However, in this section we only have the real-time condition degradation information and do not know the accurate prior distribution parameters. What is more, the priors are fixed in the whole Bayesian updating procedure in the previous articles, and if the prior distribution parameters are inaccurate, the posterior distribution for these unknown parameters would have great errors. So our paper will solve these problems in the next section.

3. Estimating the Prior Information by Fusion Algorithm

In Section 2, we have estimated the RUL of the single component by Bayesian updating procedure under the assumption that a priori distributions were known. However, we only have real-time observations and the prior distributions needed in Bayesian updating process are nonexistent. So our paper will develop a collaboration algorithm between Bayesian updating and ECM algorithm to estimate these a priori parameters.

3.1. The ECM Algorithm. We let $\Theta = [\sigma^2, \mu_0, \sigma_0^2, \mu_1', \sigma_1^2]$ denote the unknown prior distribution parameters. Given the difference value of logged degradation signal $S_{1:k}$, we can express the log-likelihood function of $S_{1:k}$ as

$$l_k(\Theta) = \ln [p(S_{1:k} | \Theta)] = \sum_{j=2}^k \ln [p(S_j | S_{1:j-1}, \Theta)], \quad (14)$$

where $p(S_{1:k} | \Theta)$ is the joint PDF of $S_{1:k}$. Thus, we can write the MLE of Θ at time t_k as

$$\hat{\Theta} = \arg \max_{\Theta} l_k(\Theta), \quad (15)$$

where $\hat{\Theta}$ is the variable value of Θ corresponding to maximum of $l_k(\Theta)$. Our goal is to estimate an appropriate prior distribution parameter $\hat{\Theta}$. However, in formula (4) θ' and β' are stochastic and unobserved, so the calculation of formula (15) is hard to complete.

In order to avoid the above problem, we propose a collaboration algorithm between Bayesian updating and ECM algorithm and use the fusion algorithm to estimate prior distributions based on degradation signals. We let $\Theta_k = [\sigma_k^2, \mu_{0,k}, \mu_{1,k}', \sigma_{0,k}^2, \sigma_{1,k}^2]$ denote the estimated parameters based on degradation signals, $S_{1:k}$, and $\hat{\Theta}_k^{(i)} = [\hat{\sigma}_k^{2(i)}, \hat{\mu}_{0,k}^{(i)}, \hat{\mu}_{1,k}^{(i)}, \hat{\sigma}_{0,k}^{2(i)}, \hat{\sigma}_{1,k}^{2(i)}]$ denote the estimated variable value of Θ_k of the i th iteration result in the ECM algorithm. We treat θ' and β' as the hidden variables, because θ' and β' are stochastic and unobserved in the whole degradation process. Our objective is to calculate the variable value of Θ_k , so the E -step of ECM algorithm can be written as

$$l(\Theta_k | \hat{\Theta}_k^{(i)}) = E_{\theta', \beta' | S_{1:k}, \hat{\Theta}_k^{(i)}} \{ \ln p(S_{1:k}, \theta', \beta' | \Theta_k) \}, \quad (16)$$

where $E_{\theta', \beta' | S_{1:k}, \hat{\Theta}_k^{(i)}}$ is the conditional expectation of hidden variables, θ' and β' , given $\hat{\Theta}_k^{(i)}$ and $S_{1:k}$.

Next we express the CM-step of ECM algorithm as

$$\begin{aligned} \hat{\Theta}_k^{(i+1/S)} &= \arg \max_{\Theta} l(\Theta_k | \hat{\Theta}_k^{(i)}) \\ &= [\hat{\sigma}_k^{2(i)}, \hat{\mu}_{0,k}^{(i)}, \hat{\mu}_{1,k}^{(i)}, \hat{\sigma}_{0,k}^{2(i)}, \hat{\sigma}_{1,k}^{2(i)}], \\ \hat{\Theta}_k^{(i+2/S)} &= \arg \max_{\Theta} l(\Theta_k | \hat{\Theta}_k^{(i+1/S)}) \\ &= [\hat{\sigma}_k^{2(i+1)}, \hat{\mu}_{0,k}^{(i)}, \hat{\mu}_{1,k}^{(i)}, \hat{\sigma}_{0,k}^{2(i)}, \hat{\sigma}_{1,k}^{2(i)}], \\ \hat{\Theta}_k^{(i+3/S)} &= \arg \max_{\Theta} l(\Theta_k | \hat{\Theta}_k^{(i+2/S)}) \\ &= [\hat{\sigma}_k^{2(i+1)}, \hat{\mu}_{0,k}^{(i+1)}, \hat{\mu}_{1,k}^{(i)}, \hat{\sigma}_{0,k}^{2(i)}, \hat{\sigma}_{1,k}^{2(i)}], \\ \hat{\Theta}_k^{(i+4/S)} &= \arg \max_{\Theta} l(\Theta_k | \hat{\Theta}_k^{(i+3/S)}) \\ &= [\hat{\sigma}_k^{2(i+1)}, \hat{\mu}_{0,k}^{(i+1)}, \hat{\mu}_{1,k}^{(i+1)}, \hat{\sigma}_{0,k}^{2(i)}, \hat{\sigma}_{1,k}^{2(i)}] \\ &\vdots \\ \hat{\Theta}_k^{(i+1)} &= \arg \max_{\Theta} l(\Theta_k | \hat{\Theta}_k^{(i+(S-1)/S)}) \\ &= [\hat{\sigma}_k^{2(i+1)}, \hat{\mu}_{0,k}^{(i+1)}, \hat{\mu}_{1,k}^{(i+1)}, \hat{\sigma}_{0,k}^{2(i+1)}, \hat{\sigma}_{1,k}^{2(i)}], \end{aligned} \quad (17)$$

where S ($S = 5$) is the dimension of vector Θ , and $\hat{\Theta}_k^{(i+1)}$ is the maximum value of Θ based on $\hat{\Theta}_k^{(i+(S-1)/S)}$ from the derivative equation $\partial l(\Theta_k | \hat{\Theta}_k^{(i+(S-1)/S)}) / \partial \hat{\sigma}_{1,k}^{2(i)} = 0$. As our focus is on the algorithm fusion process, for the theory of ECM algorithm and its convergence analysis, see Meng and Rubin [18], Van et al. [19], and Liu and Rubin [20].

3.2. The Collaboration between Bayesian Updating and ECM Algorithm. After we have reviewed the ECM algorithm, we begin our fusion process. Figure 2 shows the procedure of the fusion process.

First, we use random initial a priori distributions to start the fusion algorithm, when we collect the degradation signal $S(t_1)$ at time t_1 . Then, we get the posterior distribution of (θ', β') , $p(\theta', \beta' | S_{1:k})$, and a residual life distribution by Bayesian updating method. Of course, the results are inaccurate.

Then, we use the posterior distribution of (θ', β') , $p(\theta', \beta' | S_{1:k})$, to substitute the distribution, given $\hat{\Theta}_k^{(i)}$, of hidden variables in the E -step. So, the E -step of the ECM algorithm can be rewritten as

$$\begin{aligned} l(\Theta_k | p(\theta', \beta' | S_{1:k})) \\ = E_{\theta', \beta' | p(\theta', \beta' | S_{1:k})} \{ \ln p(S_{1:k}, \theta', \beta' | \Theta_k) \}. \end{aligned} \quad (18)$$

Because of the different E -step, we get a rewritten CM-step which is different from the one in the ECM algorithm. In

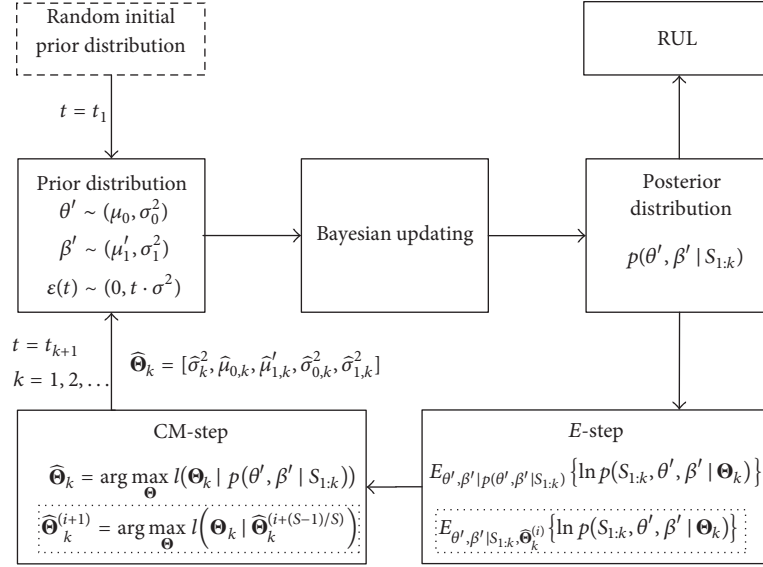


FIGURE 2: The collaboration algorithm between Bayesian updating and ECM algorithm.

order to get a more accurate estimated value, $\widehat{\Theta}_k$, in ECM algorithm we need multiple iterations. However, we can get the optimal solution of Θ_k only through S ($S = 5$) steps calculation, in our fusion algorithm. This will be proved in Theorem 4. In our fusion algorithm the rewritten CM-step can be expressed as

$$\begin{aligned}
 \widehat{\Theta}_k^{(1/5)} &= \arg \max_{\Theta_k} l(\Theta_k | p(\theta', \beta' | S_{1:k})) \\
 &= [\sigma_k^2, \mu_{0,k}, \mu'_{1,k}, \sigma_{0,k}^2, \sigma_{1,k}^2], \\
 \widehat{\Theta}_k^{(2/5)} &= \arg \max_{\Theta_k} l(\Theta_k | p(\theta', \beta' | S_{1:k}))^{(1/5)} \\
 &= [\widehat{\sigma}_k^2, \mu_{0,k}, \mu'_{1,k}, \sigma_{0,k}^2, \sigma_{1,k}^2], \\
 \widehat{\Theta}_k^{(3/5)} &= \arg \max_{\Theta_k} l(\Theta_k | p(\theta', \beta' | S_{1:k}))^{(2/5)} \\
 &= [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \mu'_{1,k}, \sigma_{0,k}^2, \sigma_{1,k}^2], \\
 \widehat{\Theta}_k^{(4/5)} &= \arg \max_{\Theta_k} l(\Theta_k | p(\theta', \beta' | S_{1:k}))^{(3/5)} \\
 &= [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \widehat{\mu}'_{1,k}, \sigma_{0,k}^2, \sigma_{1,k}^2], \\
 \widehat{\Theta}_k &= \arg \max_{\Theta_k} l(\Theta_k | p(\theta', \beta' | S_{1:k}))^{(4/5)} \\
 &= [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \widehat{\mu}'_{1,k}, \widehat{\sigma}_{0,k}^2, \sigma_{1,k}^2].
 \end{aligned} \tag{19}$$

In this five-step calculation of formula (19), we can get the optimal estimations $\sigma_k^2 = \widehat{\sigma}_k^2$, $\mu_{0,k} = \widehat{\mu}_{0,k}$, $\mu'_{1,k} = \widehat{\mu}'_{1,k}$, $\sigma_{0,k}^2 = \widehat{\sigma}_{0,k}^2$, $\sigma_{1,k}^2 = \widehat{\sigma}_{1,k}^2$, respectively. The results of the rewritten CM-step are the a priori distributions of the next Bayesian updating procedure. Each time we collect a new degradation signal $S(t_k)$ at time t_k , we can recalculate the a priori distributions and the residual life distribution.

3.3. Estimation of Prior Information. After we have finished the fusion process, we begin to calculate the optimal variable value of Θ_k by our fusion algorithm. The log-likelihood function of complete data can be expressed as

$$\begin{aligned}
 \ln p(S_{1:k}, \theta', \beta' | \Theta_k) &= \ln p(S_{1:k} | \theta', \beta', \Theta_k) + \ln p(\theta', \beta' | \Theta_k) \\
 &= -\frac{k+2}{2} \ln 2\pi - \frac{k}{2} \ln \sigma_k^2 - \frac{(S_1 - \theta' - \beta' t_1)^2}{2\sigma_k^2 t_1} \\
 &\quad - \sum_{i=2}^k \frac{(S_i - \beta' (t_i - t_{i-1}))^2}{2\sigma_k^2 (t_i - t_{i-1})} - \frac{1}{2} \ln \sigma_{0,k}^2 - \frac{1}{2} \ln \sigma_{1,k}^2 \\
 &\quad - \frac{(\theta' - \mu_{0,k})^2}{2\sigma_{0,k}^2} - \frac{(\beta' - \mu'_{1,k})^2}{2\sigma_{1,k}^2}.
 \end{aligned} \tag{20}$$

According to formulas (18) and (19), we can compute the rewritten E-step as follows:

$$\begin{aligned}
 l(\Theta_k | p(\theta', \beta' | S_{1:k})) &= E_{\theta', \beta' | p(\theta', \beta' | S_{1:k})} \{ \ln p(S_{1:k}, \theta', \beta' | \Theta_k) \} \\
 &= -\frac{k+2}{2} \ln 2\pi - \frac{k}{2} \ln \sigma_k^2
 \end{aligned}$$

$$\begin{aligned}
& \frac{S_1^2 - 2S_1(\mu_{\theta',k} + \mu_{\beta',k}t_1) + \mu_{\theta',k}^2 + \sigma_{\theta',k}^2 + 2t_1(\rho_k\sigma_{\theta',k}\sigma_{\beta',k} + \mu_{\theta',k}\mu_{\beta',k}) + t_1^2(\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)}{2\sigma_k^2 t_1} \\
& - \sum_{i=2}^k \frac{S_i^2 - 2S_i\mu_{\beta',k}(t_i - t_{i-1}) + (\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)(t_i - t_{i-1})^2}{2\sigma_k^2(t_i - t_{i-1})} - \frac{1}{2} \ln \sigma_{0,k}^2 - \frac{1}{2} \ln \sigma_{1,k}^2 \\
& - \frac{\mu_{\theta',k}^2 + \sigma_{\theta',k}^2 - 2\mu_{\theta',k}\mu_{0,k} + \mu_{0,k}^2}{2\sigma_{0,k}^2} - \frac{\mu_{\beta',k}^2 + \sigma_{\beta',k}^2 - 2\mu_{\beta',k}\mu_{1,k} + \mu_{1,k}^2}{2\sigma_{1,k}^2}.
\end{aligned} \tag{21}$$

Then, based on the rewritten CM-step, we can get the optimal $\hat{\sigma}_k^2, \hat{\mu}_{0,k}, \hat{\mu}'_{1,k}, \hat{\sigma}_{0,k}^2, \hat{\sigma}_{1,k}^2$ by the following steps, respectively.

Step 1. By calculating $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k}) / \partial \sigma_k^2 = 0$, we can get the optimal

$$\hat{\sigma}_k^2 = \frac{1}{k} \left[\frac{S_1^2 - 2S_1(\mu_{\theta',k} + \mu_{\beta',k}t_1) + \mu_{\theta',k}^2 + \sigma_{\theta',k}^2 + 2t_1(\rho_k\sigma_{\theta',k}\sigma_{\beta',k} + \mu_{\theta',k}\mu_{\beta',k}) + t_1^2(\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)}{t_1} + \sum_{i=2}^k \frac{S_i^2 - 2S_i\mu_{\beta',k}(t_i - t_{i-1}) + (\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)(t_i - t_{i-1})^2}{(t_i - t_{i-1})} \right]. \tag{22}$$

Step 2. By calculating $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k})^{(1/5)} / \partial \mu_{0,k} = 0$, we can get the optimal $\hat{\mu}_{0,k} = \mu_{\theta',k}$.

Step 5. By calculating $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k})^{(4/5)} / \partial \sigma_{1,k}^2 = 0$, we can get the optimal $\hat{\sigma}_{1,k}^2 = \sigma_{\beta',k}^2$.

Step 3. By calculating $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k})^{(2/5)} / \partial \mu'_{1,k} = 0$, we can get the optimal $\hat{\mu}'_{1,k} = \mu_{\beta',k}$.

According to the above five steps, $\hat{\Theta}_k = [\hat{\sigma}_k^2, \hat{\mu}_{0,k}, \hat{\mu}'_{1,k}, \hat{\sigma}_{0,k}^2, \hat{\sigma}_{1,k}^2]$ can be expressed as

Step 4. By calculating $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k})^{(3/5)} / \partial \sigma_{0,k}^2 = 0$, we can get the optimal $\hat{\sigma}_{0,k}^2 = \sigma_{\theta',k}^2$.

$$\hat{\sigma}_k^2 = \frac{1}{k} \left[\frac{S_1^2 - 2S_1(\mu_{\theta',k} + \mu_{\beta',k}t_1) + \mu_{\theta',k}^2 + \sigma_{\theta',k}^2 + 2t_1(\rho_k\sigma_{\theta',k}\sigma_{\beta',k} + \mu_{\theta',k}\mu_{\beta',k}) + t_1^2(\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)}{t_1} + \sum_{i=2}^k \frac{S_i^2 - 2S_i\mu_{\beta',k}(t_i - t_{i-1}) + (\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)(t_i - t_{i-1})^2}{(t_i - t_{i-1})} \right], \tag{23}$$

$$\hat{\mu}_{0,k} = \mu_{\theta',k},$$

$$\hat{\sigma}_{0,k}^2 = \sigma_{\theta',k}^2,$$

$$\hat{\mu}'_{1,k} = \mu_{\beta',k},$$

$$\hat{\sigma}_{1,k}^2 = \sigma_{\beta',k}^2.$$

We get the optimal $\hat{\sigma}_k^2, \hat{\mu}_{0,k}, \hat{\mu}'_{1,k}, \hat{\sigma}_{0,k}^2, \hat{\sigma}_{1,k}^2$ in each step, respectively. However, it does not mean that $\hat{\Theta}_k =$

$[\hat{\sigma}_k^2, \hat{\mu}_{0,k}, \hat{\mu}'_{1,k}, \hat{\sigma}_{0,k}^2, \hat{\sigma}_{1,k}^2]$ is the only maximum point of $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k}) / \partial \Theta_k = 0$. So, we need a further discussion.

Theorem 4. According to the result of formula (23), the estimated $\widehat{\Theta}_k$ is the only maximum point of $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k}) / \partial \Theta_k = 0$.

Proof. First, from formula (21) we can get $\widehat{\Theta}_k$ which is the only solution of $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k}) / \partial \Theta_k = 0$. Next, we

prove that $\widehat{\Theta}_k$ is corresponding to maximum $l(\Theta_k | p(\theta', \beta' | S_{1:k}))$. We can perform the second derivative of log-likelihood function on vector Θ_k as

$$\frac{\partial^2 l(\Theta_k | p(\theta', \beta' | S_{1:k}))}{\partial \Theta_k \partial \Theta_k^T} = \begin{bmatrix} \frac{k}{2\sigma_k^4} - \frac{\tau}{\sigma_k^6} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{\sigma_{0,k}^2} & 0 & \frac{\mu_{0,k} - \mu_{\theta',k}}{\sigma_{0,k}^4} & 0 \\ 0 & 0 & -\frac{1}{\sigma_{1,k}^2} & 0 & \frac{\mu'_{1,k} - \mu_{\beta',k}}{\sigma_{1,k}^4} \\ 0 & \frac{\mu_{0,k} - \mu_{\theta',k}}{\sigma_{0,k}^4} & 0 & \frac{1}{2\sigma_{0,k}^4} - \frac{\psi_1}{\sigma_{0,k}^6} & 0 \\ 0 & 0 & \frac{\mu'_{1,k} - \mu_{\beta',k}}{\sigma_{1,k}^4} & 0 & \frac{1}{2\sigma_{1,k}^4} - \frac{\psi_2}{\sigma_{1,k}^6} \end{bmatrix}, \quad (24)$$

where

$$\begin{aligned} \tau &= \frac{S_1^2 - 2S_1(\mu_{\theta',k} + \mu_{\beta',k}t_1) + \mu_{\theta',k}^2 + \sigma_{\theta',k}^2 + 2t_1(\rho_k \sigma_{\theta',k} \sigma_{\beta',k} + \mu_{\theta',k} \mu_{\beta',k}) + t_1^2(\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)}{t_1} \\ &+ \sum_{i=2}^k \frac{S_i^2 - 2S_i \mu_{\beta',k}(t_i - t_{i-1}) + (\mu_{\beta',k}^2 + \sigma_{\beta',k}^2)(t_i - t_{i-1})^2}{(t_i - t_{i-1})}, \\ \psi_1 &= \mu_{\theta',k}^2 + \sigma_{\theta',k}^2 - 2\mu_{\theta',k} \mu_{0,k} + \mu_{0,k}^2, \\ \psi_2 &= \mu_{\beta',k}^2 + \sigma_{\beta',k}^2 - 2\mu_{\beta',k} \mu'_{1,k} + \mu_{1,k}^2. \end{aligned} \quad (25)$$

Then we use the estimated variable value $\widehat{\Theta}_k = [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \widehat{\mu}'_{1,k}, \widehat{\sigma}_{0,k}^2, \widehat{\sigma}_{1,k}^2]$ in (23), and we can find the order principal minor of the matrix in (24) as follows:

$$\begin{aligned} \Delta_1 |_{\Theta_k = \widehat{\Theta}_k} &< 0, \\ \Delta_2 |_{\Theta_k = \widehat{\Theta}_k} &> 0, \\ \Delta_3 |_{\Theta_k = \widehat{\Theta}_k} &< 0, \\ \Delta_4 |_{\Theta_k = \widehat{\Theta}_k} &> 0, \\ \Delta_5 |_{\Theta_k = \widehat{\Theta}_k} &< 0. \end{aligned} \quad (26)$$

Thus, we can conclude that (24) is a negative-definite matrix when $\Theta_k = \widehat{\Theta}_k$, which means that $\widehat{\Theta}_k = [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \widehat{\mu}'_{1,k}, \widehat{\sigma}_{0,k}^2, \widehat{\sigma}_{1,k}^2]$ is the only maximum point. \square

This fusion procedure between Bayesian updating and ECM algorithm can be performed each time a new degradation signal is observed. That is to say, each time the

degradation signal $S(t_k)$ is observed, we can recalculate $\widehat{\Theta}_k = [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \widehat{\mu}'_{1,k}, \widehat{\sigma}_{0,k}^2, \widehat{\sigma}_{1,k}^2]$ and obtain new estimates of residual life for the newly made system. What is more, the initial values of priors in the Bayesian updating for the first time are unrestricted, and once we get a new $\widehat{\Theta}_k$, it would be used in the next Bayesian updating as the priors. In Section 4, we will use simulation method to further evaluate the performance of our collaboration algorithm and illustrate that the unrestricted initial values of priors have little influence on the estimated vector $\widehat{\Theta}_k = [\widehat{\sigma}_k^2, \widehat{\mu}_{0,k}, \widehat{\mu}'_{1,k}, \widehat{\sigma}_{0,k}^2, \widehat{\sigma}_{1,k}^2]$.

4. Simulation and Analysis

In this section, we will adopt the simulation method to further evaluate the performance of our collaboration algorithm. First, in order to represent a degradation process, we create a set of simulation data based on the exponential degradation model in Section 2. In the simulation, we assume that $\varphi = 0$, $\theta' \sim N(0.02, 2 \times 10^{-6})$, $\beta' \sim N(0.01, 1 \times 10^{-6})$, $\varepsilon(t) \sim N(0, t \cdot 4 \times 10^{-4})$. In order to observe enough degradation signals and

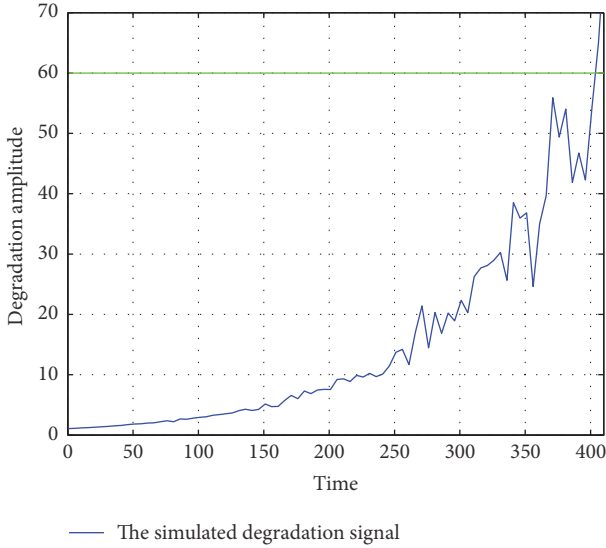


FIGURE 3: Simulation of degradation path.

TABLE 1: Different initial prior distributions.

Data	θ'	β'	$\varepsilon(t)$
(1)	$N(0.1, 2 \times 10^{-4})$	$N(0.05, 1 \times 10^{-4})$	$N(0, t \cdot 4 \times 10^{-3})$
(2)	$N(0.2, 2 \times 10^{-4})$	$N(0.1, 1 \times 10^{-4})$	$N(0, t \cdot 4 \times 10^{-3})$
(3)	$N(0.3, 2 \times 10^{-4})$	$N(0.15, 1 \times 10^{-4})$	$N(0, t \cdot 4 \times 10^{-3})$

obvious degradation trend, we let the threshold $w = 60$ and the sampling interval $(t_i - t_{i-1}) = 10$, $i = 2, 3, 4, \dots$ Figure 3 shows the trajectory of the simulated degradation signal. We obtain 41 degradation samples, and the degradation reaches the standard threshold, $w = 60$, at time $t = 403.5$.

We know that $\theta' \sim N(\mu_0, \sigma_0^2)$ and $\beta' \sim N(\mu_1 - \sigma^2/2, \sigma_1^2)$, and μ_0 and $\mu_1 - \sigma^2/2$ dominate the degradation and RUL. We will use the estimated results of them to prove that these unrestricted initial priors have little influence on the estimated accuracy of the RUL. The different initial prior distributions are as shown in Table 1.

On the other hand, in order to prove that our collaboration algorithm can get a more accurate RUL than other separate Bayesian updating methods, we will compare our fusion algorithm with the method in Gebrael et al. [10]. by using the 2nd set of initial priors in Table 1. Figure 4 represents the estimated results of μ_0 . Figure 5 represents the estimated results of $\mu_1 - \sigma^2/2$.

From Figures 4 and 5, we know that the collaboration algorithm can estimate the mean of θ' and β' accurately even with inaccurate prior information. Although we have used a variety of inaccurate prior distributions, it can still be observed that this aspect has little effect on the estimation of the means of θ' and β' , showing that the degradation measurements dominate and the collaboration algorithm is effective. We can know that the estimated means of θ' and β' gradually approximate to the simulated value, 0.02 and 0.01, at time about 50, in spite of different initial priors.

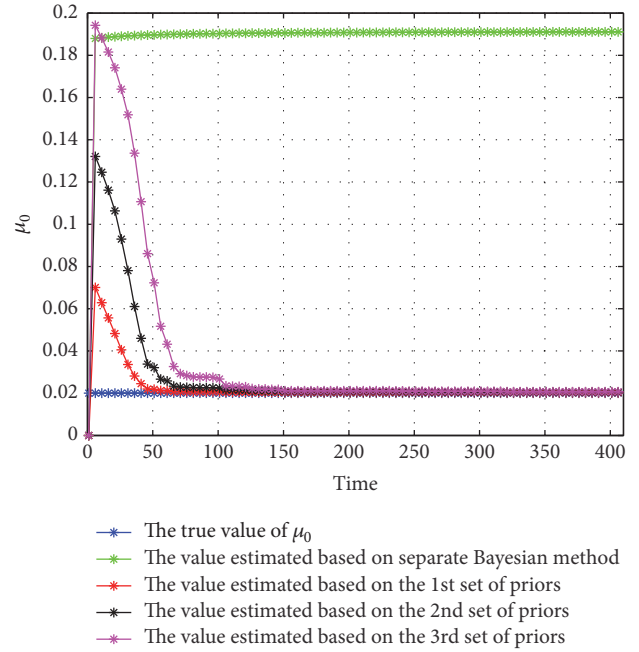


FIGURE 4: Estimated results of μ_0 .

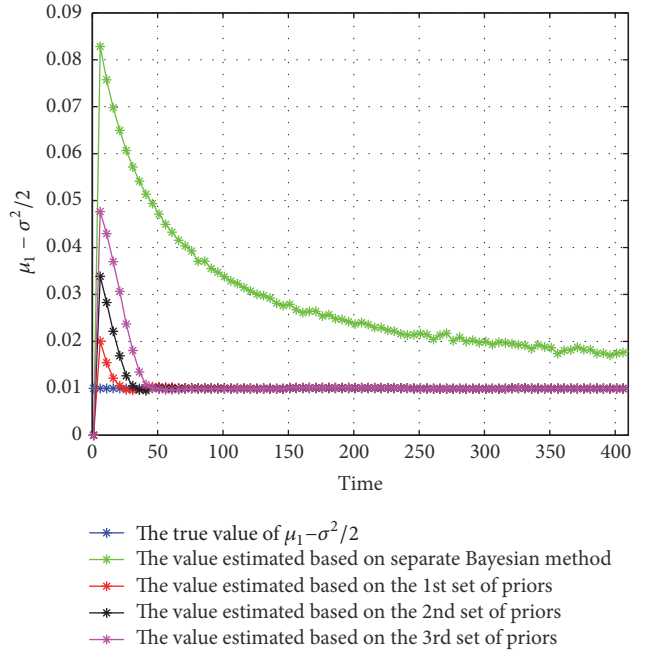


FIGURE 5: Estimated results of $\mu_1 - \sigma^2/2$.

What is more, our collaboration algorithm can get more accurate estimations compared with the separate Bayesian updating by using the 2nd set of initial priors in Table 1.

Given the estimated prior distributions we calculate the point estimations of RUL of the newly made component. Figure 6 shows the point estimations of RUL by our algorithm and separate Bayesian updating method.

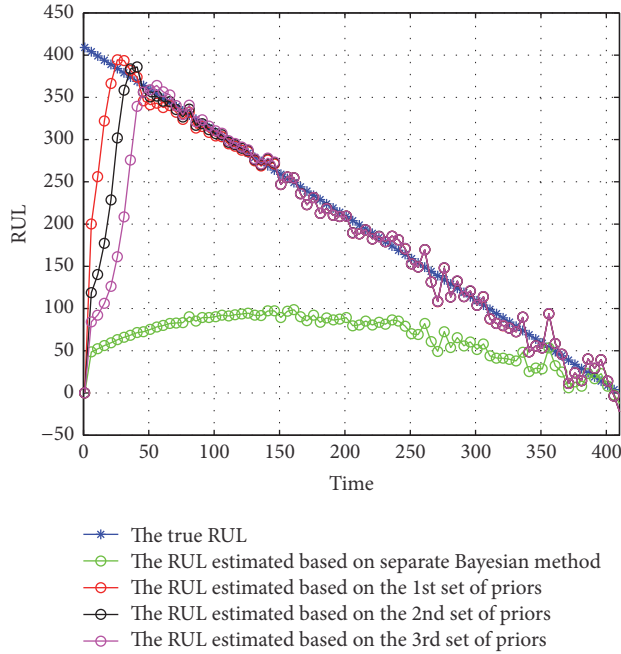


FIGURE 6: Point estimations of RUL by collaboration algorithm and separate Bayesian Updating method.

From Figure 6, we can know that our collaboration algorithm can get a more accurate RUL comparing with the separate Bayesian method. And inaccurate a priori distributions have little effect on the estimation of the RUL. Furthermore, the point estimations of RUL by our collaboration algorithm can also reflect the fluctuation of degradation caused by Brownian motion error, and this can be known from Figure 3.

5. Conclusion

In this paper, we presented a collaboration algorithm that contains the characteristic of Bayesian updating and ECM algorithm. The difficulties of the fusion process mainly consist of two parts; the first is building the connection between ECM algorithm and Bayesian updating. It is not easy to find another substitute for the i th iteration distribution in the rewritten E-step, and the posterior distribution of the Bayesian updating is an optimal one. The second is proving the optimal estimating result about the prior information. We can get a maximum estimating result, $\hat{\Theta}_k = [\hat{\sigma}_k^2, \hat{\mu}_{0,k}, \hat{\mu}'_{1,k}, \hat{\sigma}_{0,k}^2, \hat{\sigma}_{1,k}^2]$, by the rewritten CM-step. However, it does not mean that the estimating result is the only maximum point of $\partial l(\Theta_k | p(\theta', \beta' | S_{1:k}) / \partial \Theta_k = 0$. In this paper, we use the second derivative of log-likelihood function on vector Θ_k and the order principal minor of the matrix in (24) to prove it.

Our fusion algorithm can predict failures of newly made systems in the absence of prior degradation knowledge. Although our fusion algorithm started with random initial a priori distributions, the simulation experiments show that the inaccurate a priori distributions have little effect on the estimation of the RUL, and our fusion algorithm can

get a better prediction than the separate Bayesian method. Nevertheless, there are still some issues that needed further investigation for the estimation of the RUL on the single component without prior degradation knowledge.

First, we assume that the exponential degradation signal model and the error fluctuation are adequate for the given degradation path. Actually, we can not guarantee which model is the best one by visual judgment. So we also need to focus on model selection and evaluate the goodness-of-fit of various degradation path models.

Second, we assume that the stochastic parameters of the exponential degradation model are normally distributed. However, the stochastic parameters may follow a Gamma distribution or other distributions. Therefore, we also should investigate the performance of our collaboration algorithm when the underlying normal distribution assumptions are not satisfied.

Third, our work started the RUL calculation in the degradation processes assuming that the exact point-in-time of the initial degradation is known. However, the functioning system would be stable within a period of time, and the point-in-time for the initial degradation is unknown and stochastic. By considering the distribution of the initial degradation point, we may be able to predict the RUL right after its installment.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study is financially supported by the National Natural Science Foundation of China (61573365, 61374120, and 61473094) and the National Science Fund for Distinguished Youth Scholars of China (61025014).

References

- [1] Z.-S. Ye and M. Xie, "Stochastic modelling and analysis of degradation for highly reliable products," *Applied Stochastic Models in Business and Industry*, vol. 31, no. 1, pp. 16–32, 2015.
- [2] M. A. Freitas, M. L. G. De Toledo, E. A. Colosimo, and M. C. Pires, "Using degradation data to assess reliability: a case study on train wheel degradation," *Quality and Reliability Engineering International*, vol. 25, no. 5, pp. 607–629, 2009.
- [3] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation—a review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, 2011.
- [4] C. J. Lu and W. O. Meeker, "Using degradation measures to estimate a time-to-failure distribution," *Technometrics*, vol. 35, no. 2, pp. 161–174, 1993.
- [5] J.-C. Lu, J. Park, and Q. Yang, "Statistical inference of a time-to-failure distribution derived from linear degradation data," *Technometrics*, vol. 39, no. 4, pp. 391–400, 1997.
- [6] C. Su, J.-C. Lu, D. Chen, and J. M. Hughes-Oliver, "A random coefficient degradation model with random sample size," *Lifetime Data Analysis*, vol. 5, no. 2, pp. 173–183, 1999.

- [7] B. P. Weaver, W. Q. Meeker, L. A. Escobar, and J. Wendelberger, "Methods for planning repeated measures degradation studies," *Technometrics*, vol. 55, no. 2, pp. 122–134, 2013.
- [8] S.-J. Wu and J. Shao, "Reliability analysis using the least squares method in nonlinear mixed-effect degradation models," *Statistica Sinica*, vol. 9, no. 3, pp. 855–877, 1999.
- [9] M. E. Robinson and M. J. Crowder, "Bayesian methods for a growth-curve degradation model with repeated measures," *Lifetime Data Analysis*, vol. 6, no. 4, pp. 357–374, 2000.
- [10] N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, "Residual-life distributions from component degradation signals: a Bayesian approach," *IIE Transactions*, vol. 37, no. 6, pp. 543–557, 2005.
- [11] N. Gebraeel, "Sensory-updated residual life distributions for components with exponential degradation patterns," *IEEE Transactions on Automation Science and Engineering*, vol. 3, no. 4, pp. 382–393, 2006.
- [12] S. Chakraborty, N. Gebraeel, M. Lawley, and H. Wan, "Residual-life estimation for components with non-symmetric priors," *IIE Transactions*, vol. 41, no. 4, pp. 372–387, 2009.
- [13] N. Chen and K. L. Tsui, "Condition monitoring and remaining useful life prediction using degradation signals: revisited," *IIE Transactions*, vol. 45, no. 9, pp. 939–952, 2013.
- [14] K. A. Kaiser and N. Z. Gebraeel, "Predictive maintenance management using sensor-based degradation models," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 39, no. 4, pp. 840–849, 2009.
- [15] A. Elwany and N. Gebraeel, "Real-time estimation of mean remaining life using sensor-based degradation models," *Journal of Manufacturing Science and Engineering*, vol. 131, no. 5, pp. 0510051–0510059, 2009.
- [16] Z. Li, Y. Deng, and C. Mastrangelo, "Model selection for degradation-based Bayesian reliability analysis," *Journal of Manufacturing Systems*, vol. 37, pp. 72–82, 2015.
- [17] R. Durrett, *Probability: Theory and Examples*, Cambridge University Press, Cambridge, UK, 2010.
- [18] X.-L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: a general framework," *Biometrika*, vol. 80, no. 2, pp. 267–278, 1993.
- [19] D. A. Van, X.-L. Meng, and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: Computing the asymptotic variance," *Statistica Sinica*, vol. 5, no. 1, pp. 55–75, 1995.
- [20] C. Liu and D. B. Rubin, "The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence," *Biometrika*, vol. 81, no. 4, pp. 633–648, 1994.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

