SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2019, Vol. 62 022202:1–022202:16
https://doi.org/10.1007/s11432-017-9347-5

Remaining useful life prediction for multi-component systems with hidden dependencies

Xiaopeng XI[1], Maoyin CHEN[1*] & Donghua ZHOU[2,1*]

[1]*Department of Automation, Tsinghua University, Beijing* 100084, *China;*
[2]*College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao* 266590, *China*

Received 25 September 2017/Revised 27 November 2017/Accepted 31 January 2018/Published online 27 December 2018

Abstract How can we predict the remaining useful life (RUL) of a dynamic system subject to multiple dependent degradations? Is it possible to address the above problem when the degradation information is obtained indirectly? According to a new type of state space-based model, we mainly develop an online RUL prediction method for the above system. In this model, the dependencies among different degradations can be reflected in a diffusion coefficient matrix. Considering that some industrial systems like blast furnaces are usually equipped with multi-sensors, an efficient information fusion strategy also plays an important role in predicting the RUL. Based on multi-dimensional observations, the hidden degradation states are identified through the sequential Kalman filtering. Meanwhile, the unknown parameters in the model are updated iteratively by the expectation maximization (EM) algorithm. At last, the RUL distributions are simulated through the Monte Carlo method, in which three types of failure structures with regard to the degradations are considered. The effectiveness of the proposed method is fully verified by a numerical example as well as a case study about the blast furnace.

Keywords remaining useful life, hidden dependencies, state space model, multi-source information fusion, failure structure

Citation Xi X P, Chen M Y, Zhou D H. Remaining useful life prediction for multi-component systems with hidden dependencies. Sci China Inf Sci, 2019, 62(2): 022202, https://doi.org/10.1007/s11432-017-9347-5

1 Introduction

Remaining useful life (RUL) prediction through condition monitoring (CM) has gained much interest in recent years [1–3]. The term "RUL prediction" means to find the probability density function (PDF) or the expectation of the RUL [4]. However, owing to the uncertainty, a PDF has attracted more attention. Ahmadzadeh and Lundberg [5] summarized that the RUL prediction methods can be classified into physical, experimental, data-driven and hybrid approaches. Si et al. [6] reviewed the statistical data-driven approaches based on direct and indirect CM data. Because data-driven approaches rely only on the observation data, they are more commonly used to predict the RUL. For data-driven methods, most studies mainly deal with one-dimensional degradation processes, no matter how complex the systems are. It implies all the performance characteristics are assumed to be totally represented by one single degradation. Wei et al. [7] utilized the decentralized Kalman filtering to estimate the hidden degradation states, and then provided a two-stage method for predicting the RUL. Niu and Yang [8] assessed the RUL of the operating machines by using a data-fusion strategy. Li and Ryan [9] proposed a dynamic

* Corresponding author (email: mychen@mail.tsinghua.edu.cn, zdh@mail.tsinghua.edu.cn)

base-stock policy in a Bayesian manner, where the degradation is modeled according to a Wiener process. Tang et al. [10] used a truncated normal distribution to predict the RUL for lithium-ion batteries.

Note that a practical system usually contains some different kinds of operating units. For example, a large blast furnace is mainly composed of throat, stack, bosh, belly, and hearth. Different units are better characterized by different degradations. Because of the interactions of the units, there often exists certain dependencies among multiple degradations, such as the hearth erosion and tuyere burnout. Therefore, an important problem is to predict the RUL when multiple dependent degradations are considered. Over the past few years, some efforts have been put into the RUL prediction for multiple dependent degradation processes. Bian and Gebraeel [11] considered the component-to-component interaction, and proposed a basic framework by using stochastic differential equations. However, this method depends heavily on the directly measurable degradation signals, and may not be applied to hidden dependent degradations. Wang et al. [12] presented a method to predict the RUL of a product with the Wiener process-based degradations, in which the dependency was characterized by the Frank copula function. Xi et al. [13] developed a copula-based sampling method for data-driven prognostics, and tried to verify its validity by predicting the RUL of an electric cooling fan. Against the issue of stochastic dependence in multi-component systems, Shi and Zeng [14] first modeled the effect of one degradation on the RUL of the others, and then adopted the stochastic filtering theory for a real-time RUL prediction along with an opportunistic maintenance. Khorasgani et al. [15] divided the system-level prognostics into estimation and prediction procedures, and incorporated two methods including stochastic simulation and inverse first order reliability method to derive the RUL of the system. Based on multiple component-level health factors, Rodrigues [16] constructed a comprehensive performance indicator to implement the RUL prediction by a simple extrapolation. Furthermore, Prakash et al. [17] put forward a model-based prognostics framework with respect to hybrid multi-component systems, in which the directions of degradations can be analyzed using the dynamic sensitivity signature matrix.

In these studies, copula functions are often used to predict the RUL for dependent degradations. This is because copula functions are flexible to construct multivariate distributions in the condition that the concrete multivariate model is unknown. However, using copula functions to predict the RUL has several significant limitations. First, the dependencies among multiple degradations are not evident. Copula functions only consider the statistical property by the Kendall's rank correlation coefficient, and thus cannot give a closed mathematical form of real dependencies. Second, copula functions aim at translating the multivariate function into a one-dimensional function. But reducing dimensionality may lose some important information such as the cross terms in the covariance matrix. Third, how to choose an appropriate copula function is also a tough problem for various actual dependent structures. Different types of copula functions lead to completely different predicting results.

Here we mainly develop a new type of degradation model, and provide an online algorithm to predict the RUL of the system with multiple dependent degradations. Because the degradation processes are generally uncertain over time, one single sensor may be insufficient to measure the performance characteristic, especially for the core component [18, 19]. Hence some practical systems collect multi-source information to ensure the reliability of online monitoring. Unlike traditional degradation models, our model can deal with the hidden dependencies according to a class of multi-sensor measurement framework. Once the new sensor observations are available, the hidden degradation states will be estimated by a sequential Kalman filtering method. At the same time, the unknown parameters in the model are also updated by the expectation maximization (EM) algorithm. The reliability function as well as the PDF of the RUL can be simulated simultaneously using the Monte Carlo method, in which three types of failure structures are considered.

Based on the combination of a nolinear drift vector, a series of independent Wiener processes with a simple coefficient matrix, and the Gaussian measurement noises, an integrated state space model is provided for the complicated systems with multiple dependent degradations. As previously mentioned, Refs. [12, 13] also presented their unique degradation modeling methods to describe the hidden dependencies. Nevertheless, both of them rely heavily on the copula functions, and cannot directly address the multi-source information from the multi-sensor array. Aiming at the information fusion of multiple

sensors, Wei [20] gave a feasible solution by the traditional Kalman filtering method, but the dependencies remain invisible subject to four different copula functions. In detail, there exist three major differences between the proposed method and the existing degradation modeling methods considering hidden dependencies. First of all, our method does not only represent the statistical properties of the degradation paths, but it also reveals the specific dependency degree among different items quantitatively. The hidden dependencies lie in the diffusion coefficient matrix, which is more flexible than the fixed copula functions. In the second place, the dimensionality of the degradation data is not reduced, which means the entire data are available to the degradation modeling, without losing some important information for the RUL prediction. Different from [12, 13, 20], the multi-dimensional data set can be handled simultaneously by our method. At last, we can simply utilize the EM algorithm to identify the coefficient matrix according to the real degradation data set, and hence our method does not require idealistic and impractical prior knowledge like the assumption of appropriate copula functions.

There are two major contributions of this study. The first is to develop a new model to characterize the hidden dependencies among multiple degradations. This model is more convincing than those utilizing copula functions to characterize the virtual statistical dependencies. In particular, the detailed dependencies in our model can be easily estimated based on the indirect CM data, which cannot be used for most of the existing multi-component degradation models. The second is to predict the RUL of a system under the multi-sensor measurement framework, rather than the one-to-one degradation signals. Moreover, we take into account three failure structures for a whole system to ensure the validity of the proposed system-level method.

The remaining parts of the paper are organized as follows. A new state space model corresponding to multiple dependent degradations is developed in Section 2. The online algorithm with respect to the estimation of the hidden degradation states and unknown parameters is proposed in Section 3. To obtain the PDF of the RUL for systems with different failure structures, a simulation method is given in Section 4. Two examples are provided in Section 5 to verify the effectiveness of the proposed method. Finally, the conclusion is briefly drawn in Section 6.

2 A state space model with dependent hidden degradations

In the literature [21–25], Wiener process has been widely used in modeling the degradations such as light emitting diode and fatigue crack growth. As stated in Section 1, some complex systems may be characterized by multiple dependent degradations, and thus the traditional one-dimensional Wiener process is no longer applicable. Hence how to model the dependencies among multiple degradations accurately becomes a significant problem.

Although copula functions attempt to simplify the statistical properties among different degradations, they have no idea about the actual dependency structure of the system. For direct degradation data, in our previous study [26], we proposed a new kind of Wiener process-based degradation model, which describes the dependencies without using copula functions. Assume that there exist N degradations in the system, and then the model is given by

$$\boldsymbol{x}(t) = \boldsymbol{\eta} t + (\sigma \boldsymbol{I} + \boldsymbol{\Lambda})\boldsymbol{B}(t), \tag{1}$$

where $\boldsymbol{x}(t) = [x_1(t), \ldots, x_N(t)]^{\mathrm{T}}$, $x_i(t)$ is the ith degradation of the system at time t, $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_N]^{\mathrm{T}}$, η_i is the corresponding drift coefficient of $x_i(t)$, $\boldsymbol{B}(t)$ contains N independent standard Brownian motions, that is, $\boldsymbol{B}(t) = [B_1(t), \ldots, B_N(t)]^{\mathrm{T}}$, and different $B_i(t)$ expresses different diffusion process. In (1), σ is a scalar representing the general characteristics of all the degradations, while $\boldsymbol{\Lambda}$ is a real symmetric matrix denoted by $\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}$, which mainly determines the dependency structure. Here, $x_i(t)$ may not only be affected by $B_i(t)$, but also by $B_1(t), \ldots, B_{i-1}(t), B_{i+1}(t), \ldots, B_N(t)$.

Because the data are generated by the factor model, a reasonable regularization method is to restrict the rank of $\boldsymbol{V}$ [27, 28]. The rank selection should be accordant with the practical background. For example, in a blast furnace, hearth erosion has a strong impact on the tuyere, but rarely affects the other

components. This implies $\boldsymbol{\Lambda}$ is relatively sparse for high dimensional data, with $\boldsymbol{V}$ corresponding to a low rank.

Note that there are several limitations in (1). On the one hand, as a scalar matrix, $\sigma\boldsymbol{I}$ only describes the uniform case, in which the unrelated parts of each diffusion process are completely identical. On the other hand, $\boldsymbol{\eta}t$ ignores the common nonlinearity in the practical industrial processes. Considering these problems, the model can be improved by

$$\boldsymbol{x}(t) = \boldsymbol{\lambda}t^{\gamma} + (\boldsymbol{\Sigma} + \boldsymbol{\Lambda})\boldsymbol{B}(t), \tag{2}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^{\mathrm{T}}$, λ_i is the ith drift coefficient of the system, γ is a nonlinear coefficient shared by all the degradations, representing the degree of the nonlinearity, and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ denotes the nonuniform residual diffusion coefficient matrix.

Generally speaking, the drift term is used to describe the major variation tendency of a degradation process. Regarding some practical industrial systems like the large blast furnaces, it is common to see the obvious nonlinearity in their degradation paths, attributing to the composition effects of complicated internal reactions and external environments. Although the drift term can be arbitrary time dependent function in theory, a reliable prior mathematical form seems quite necessary for the following procedures of parameter estimation and RUL prediction. As shown in (2), the power-law-type drift vector $\boldsymbol{\lambda}t^{\gamma}$ is selected in this paper. The combination of $\boldsymbol{\lambda}$ and γ corresponds to an infinite number of nonlinear paths. Specifically, when $\gamma = 1$, it degenerates into the traditional linear model. In fact, the power-law-type drifts exist extensively within our nature and society, such as the traffic enhancement of websites and the flowing curve of polymeric fluids. Compared with some other relevant degradation models, the power-law-type model is more natural and simple in revealing the nonlinearity from the real world. Besides, by creating an online identification algorithm of the parameters, this model can effectively track the degradation tendencies of multi-component systems.

From (1) and (2), it is obvious that the directly measurable degradation signals are required for the RUL prediction. The problem remains that these models may not be applied for hidden degradations. Different from the traditional studies, the degradations in this paper are assumed to be indirectly observed by the sensors. In other words, the degradation signal $\boldsymbol{x}(t)$ is actually hidden by the measurement data $\boldsymbol{y}(t)$. Here we first use $\boldsymbol{y}(t)$ to estimate $\boldsymbol{x}(t)$, and then predict the RUL of the system based on the estimation results.

Because one single sensor may be inaccurate to some extent, here, each performance degradation is monitored using multiple sensors. For simplicity, the sensors are divided into M groups, each group of which contains N sensors with the same sampling rate. Thus, the measurements of the jth group sensors can be described by

$$\boldsymbol{y}^{(j)}(t) = \boldsymbol{y}_0^{(j)} + \boldsymbol{g}^{(j)}\boldsymbol{x}(t) + \boldsymbol{\psi}^{(j)}, \tag{3}$$

where $\boldsymbol{y}^{(j)}(t) = [y_1^{(j)}(t), \ldots, y_N^{(j)}(t)]^{\mathrm{T}}$, $\boldsymbol{y}_0^{(j)}$ is the initial measurement deviation vector, $\boldsymbol{\psi}^{(j)}$ denotes the random measurement noise vector, $\boldsymbol{g}^{(j)}$ is a known diagonal matrix expressed as $\boldsymbol{g}^{(j)} = \mathrm{diag}(g_1^{(j)}, \ldots, g_N^{(j)})$, and the element $g_i^{(j)}$ corresponds to the ith degradation process $x_i(t)$.

Based on the discretization of the infinitely divisible model in (2) and the measurement equation in (3), a state space model with multiple dependent hidden degradations can be expressed by

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{\lambda}_k(t_{k+1}^{\gamma_k} - t_k^{\gamma_k}) + \boldsymbol{\phi}_k, \\ \boldsymbol{y}_k^{(j)} = \boldsymbol{y}_0^{(j)} + \boldsymbol{g}^{(j)}\boldsymbol{x}_k + \boldsymbol{\psi}_k^{(j)}, \end{cases} \tag{4}$$

where $\boldsymbol{x}_k$ and $\boldsymbol{y}_k^{(j)}$ denote the hidden degradation states and the measurement vector from the jth sensor group at time t_k, respectively. In particular, define τ_k as the kth monitoring interval, and then both $\boldsymbol{\phi}_k$ and $\boldsymbol{\psi}_k^{(j)}$ follow the multinormal distribution, i.e., $\boldsymbol{\phi}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$, $\boldsymbol{Q}_k = \tau_k(\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^{\mathrm{T}} + \boldsymbol{\Lambda}_k\boldsymbol{\Lambda}_k^{\mathrm{T}})$, $\boldsymbol{\psi}_k^{(j)} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Xi}_k^{(j)})$, and $\boldsymbol{\Xi}_k^{(j)} = \mathrm{diag}(\xi_{k,1}^{(j)}, \ldots, \xi_{k,N}^{(j)})$.

Compared with copula functions, Eq. (4) reveals the dependency without any statistical hypothesis. This is more suitable for the structure factors in practical systems. Moreover, our model outperforms the

existing multi-component degradation models from the perspective of data requirements. In other words, indirect CM data from multi-sensors can also be adopted here for the RUL prediction. Note that there exists an unknown parameter set $\boldsymbol{\Omega} = \{\boldsymbol{\lambda}, \gamma, \boldsymbol{Q}, \xi_i^{(j)}\}$ in (4). Therefore, the objective in the next section is to estimate these parameters, together with the degradation states $\boldsymbol{x}_k$.

3 Multiple dependent hidden degradations identification algorithm

In this section, we mainly introduce an online parameter identification algorithm for (4). First, the sequential Kalman fusion filtering [29] is used to estimate the hidden degradation states. Then, an EM algorithm is adopted for identifying the unknown model parameters.

3.1 Degradation state identification

Our goal is to estimate the fusion estimation of $\boldsymbol{x}_{k|k} = \mathrm{E}(\boldsymbol{x}_k|\boldsymbol{y}_0^{(1)}, \ldots, \boldsymbol{y}_k^{(1)}, \ldots, \boldsymbol{y}_0^{(M)}, \ldots, \boldsymbol{y}_{k_M}^{(M)}) = \mathrm{E}(\boldsymbol{x}_k|\boldsymbol{\Upsilon}_k)$. Then the covariance matrix can be expressed by $\boldsymbol{P}_{k|k} = \mathrm{E}[(\boldsymbol{x}_{k|k} - \boldsymbol{x}_k)(\boldsymbol{x}_{k|k} - \boldsymbol{x}_k)^{\mathrm{T}}|\boldsymbol{\Upsilon}_k]$. Because there exist multiple groups of sensors, we first define the time update of the hidden degradation state vector as $\boldsymbol{x}_{k+1|k}^{(j)} = \mathrm{E}(\boldsymbol{x}_{k+1}|\boldsymbol{\Upsilon}_k^{(j)})$, and the covariance matrix from the jth sensor group as $\boldsymbol{P}_{k+1|k}^{(j)} = \mathrm{E}[(\boldsymbol{x}_{k+1|k}^{(j)} - \boldsymbol{x}_{k+1})(\boldsymbol{x}_{k+1|k}^{(j)} - \boldsymbol{x}_{k+1})^{\mathrm{T}}|\boldsymbol{\Upsilon}_k^{(j)}]$. Then the measurement update and the covariance matrix can be represented as $\boldsymbol{x}_{k+1|k+1}^{(j)} = \mathrm{E}(\boldsymbol{x}_{k+1}|\boldsymbol{\Upsilon}_{k+1})$ and $\boldsymbol{P}_{k+1|k+1}^{(j)} = \mathrm{E}[(\boldsymbol{x}_{k+1|k+1}^{(j)} - \boldsymbol{x}_{k+1})(\boldsymbol{x}_{k+1|k+1}^{(j)} - \boldsymbol{x}_{k+1})^{\mathrm{T}}|\boldsymbol{\Upsilon}_{k+1}]$, respectively.

For each group of sensors, the time update equations can be formulated by

$$\hat{\boldsymbol{x}}_{k+1|k}^{(j)} = \hat{\boldsymbol{x}}_{k|k} + \boldsymbol{\lambda}_k(t_{k+1}^{\gamma_k} - t_k^{\gamma_k}), \tag{5}$$

$$\boldsymbol{P}_{k+1|k}^{(j)} = \boldsymbol{P}_{k|k} + \boldsymbol{Q}_k. \tag{6}$$

The measurement update equations are listed as follows:

$$\hat{\boldsymbol{x}}_{k+1|k+1}^{(j)} = \hat{\boldsymbol{x}}_{k+1|k}^{(j)} + \boldsymbol{C}_{k+1}^{(j)}\left(\boldsymbol{y}_{k+1}^{(j)} - \boldsymbol{y}_0^{(j)} - \boldsymbol{g}^{(j)}\hat{\boldsymbol{x}}_{k+1|k}^{(j)}\right), \tag{7}$$

$$\boldsymbol{P}_{k+1|k+1}^{(j)} = \left(\boldsymbol{I} - \boldsymbol{C}_{k+1}^{(j)}\boldsymbol{g}^{(j)}\right)\boldsymbol{P}_{k+1|k}^{(j)}, \tag{8}$$

where

$$\boldsymbol{C}_{k+1}^{(j)} = \boldsymbol{P}_{k+1|k}^{(j)}(\boldsymbol{g}^{(j)})^{\mathrm{T}}\left[\boldsymbol{g}^{(j)}\boldsymbol{P}_{k+1|k}^{(j)}(\boldsymbol{g}^{(j)})^{\mathrm{T}} + \boldsymbol{\Xi}_{k+1}^{(j)}\right]^{-1}. \tag{9}$$

With multi-dimensional observations, a suitable fusion strategy is necessary. Once multiple groups of sensor data are obtained at each monitoring time, the fusion estimation will be obtained by the sequential Kalman filtering method. In concrete terms, if $M = 2$, the fusion equations can be represented by

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k+1}^{(M)} + \boldsymbol{C}_{k+1}^{(1)}\left(\boldsymbol{y}_{k+1}^{(1)} - \boldsymbol{y}_0^{(1)} - \boldsymbol{g}^{(1)}\hat{\boldsymbol{x}}_{k+1|k+1}^{(M)}\right), \tag{10}$$

$$\boldsymbol{P}_{k+1|k+1} = (\boldsymbol{I} - \boldsymbol{C}_{k+1}^{(1)}\boldsymbol{g}^{(1)})\boldsymbol{P}_{k+1|k+1}^{(M)}, \tag{11}$$

where

$$\hat{\boldsymbol{x}}_{k+1|k+1}^{(M)} = \hat{\boldsymbol{x}}_{k+1|k}^{(M)} + \boldsymbol{C}_{k+1}^{(M)}\left(\boldsymbol{y}_{k+1}^{(M)} - \boldsymbol{y}_0^{(M)} - \boldsymbol{g}^{(M)}\hat{\boldsymbol{x}}_{k+1|k}^{(M)}\right), \tag{12}$$

$$\boldsymbol{P}_{k+1|k+1}^{(M)} = \left(\boldsymbol{I} - \boldsymbol{C}_{k+1}^{(M)}\boldsymbol{g}^{(M)}\right)\boldsymbol{P}_{k+1|k}^{(M)}, \tag{13}$$

$$\boldsymbol{C}_{k+1}^{(M)} = \boldsymbol{P}_{k+1|k}^{(M)}(\boldsymbol{g}^{(M)})^{\mathrm{T}}\left[\boldsymbol{g}^{(M)}\boldsymbol{P}_{k+1|k}^{(M)}(\boldsymbol{g}^{(M)})^{\mathrm{T}} + \boldsymbol{\Xi}_{k+1}^{(M)}\right]^{-1}, \tag{14}$$

$$\boldsymbol{C}_{k+1}^{(1)} = \boldsymbol{P}_{k+1|k+1}^{(M)}(\boldsymbol{g}^{(j)})^{\mathrm{T}}\left[\boldsymbol{g}^{(j)}\boldsymbol{P}_{k+1|k+1}^{(M)}(\boldsymbol{g}^{(j)})^{\mathrm{T}} + \boldsymbol{\Xi}_{k+1}^{(1)}\right]^{-1}. \tag{15}$$

For a more complicated situation, i.e., $M \geqslant 3$, Eqs. (10)–(15) should be repeated in sequence for $M-1$ times. In other words, based on the augmented measurement matrix, the batch processing operation can be written into

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k+1}^{(M)} + \boldsymbol{C}_{k+1}^{(1:M-1)} \left(\boldsymbol{y}_{k+1}^{(1:M-1)} - \boldsymbol{y}_{0}^{(1:M-1)} - \boldsymbol{g}^{(1:M-1)} \hat{\boldsymbol{x}}_{k+1|k+1}^{(M)} \right), \tag{16}$$

$$\boldsymbol{P}_{k+1|k+1} = \left(\boldsymbol{I} - \boldsymbol{C}_{k+1}^{(1:M-1)} \boldsymbol{g}^{(1:M-1)} \right) \boldsymbol{P}_{k+1|k+1}^{(M)}, \tag{17}$$

where

$$\boldsymbol{y}_{k+1}^{(1:M-1)} = \begin{bmatrix} \boldsymbol{y}_{k+1}^{(1)} \\ \boldsymbol{y}_{k+1}^{(2)} \\ \vdots \\ \boldsymbol{y}_{k+1}^{(M-1)} \end{bmatrix}, \quad \boldsymbol{y}_{0}^{(1:M-1)} = \begin{bmatrix} \boldsymbol{y}_{0}^{(1)} \\ \boldsymbol{y}_{0}^{(2)} \\ \vdots \\ \boldsymbol{y}_{0}^{(M-1)} \end{bmatrix}, \tag{18}$$

$$\boldsymbol{g}^{(1:M-1)} = \begin{bmatrix} \boldsymbol{g}^{(1)} \\ \boldsymbol{g}^{(2)} \\ \vdots \\ \boldsymbol{g}^{(M-1)} \end{bmatrix}, \quad \boldsymbol{\Xi}_{k+1}^{(1:M-1)} = \begin{bmatrix} \boldsymbol{\Xi}_{k+1}^{(1)} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Xi}_{k+1}^{(2)} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{\Xi}_{k+1}^{(M-1)} \end{bmatrix}, \tag{19}$$

$$\boldsymbol{C}_{k+1}^{(1:M-1)} = \boldsymbol{P}_{k+1|k+1}^{(M)} (\boldsymbol{g}^{(1:M-1)})^{\mathrm{T}} \left[\boldsymbol{g}^{(1:M-1)} \boldsymbol{P}_{k+1|k+1}^{(M)} (\boldsymbol{g}^{(1:M-1)})^{\mathrm{T}} + \boldsymbol{\Xi}_{k+1}^{(1:M-1)} \right]^{-1}. \tag{20}$$

3.2 Parameter identification

In this subsection, the unknown parameters in $\boldsymbol{\Omega}$ are identified by the EM algorithm. At the kth monitoring time, once the new sensor measurement data as well as the current state estimation results are available, the iterative estimation procedure for $\boldsymbol{\Omega}_k$ is started. Concretely, the qth iterative step is given by

$$\hat{\boldsymbol{\Omega}}_{k}^{(q+1)} = \underset{\boldsymbol{\Omega}}{\operatorname{argmax}} \, \mathrm{E}_{\boldsymbol{x}_k | \boldsymbol{\Upsilon}_k, \hat{\boldsymbol{\Omega}}_k^{(q)}} \left\{ \ln f(\boldsymbol{x}_k, \boldsymbol{\Upsilon}_k; \boldsymbol{\Omega}) \right\}. \tag{21}$$

3.2.1 *The E-step*

Based on the Bayesian rule, the qth E-step expectation function can be described by

$$\ell^{(q)}(\boldsymbol{\Omega}) = \mathrm{E}_{\boldsymbol{x}_k | \boldsymbol{\Upsilon}_k, \hat{\boldsymbol{\Omega}}_k^{(q)}} \left\{ \sum_{j=1}^{M} \ln f(\boldsymbol{y}^{(j)} | \boldsymbol{x}; \boldsymbol{\Omega}) + \ln f(\boldsymbol{x}; \boldsymbol{\Omega}) \right\} = \ell_1^{(q)}(\boldsymbol{\Omega}) + \ell_2^{(q)}(\boldsymbol{\Omega}), \tag{22}$$

where

$$\ell_1^{(q)}(\boldsymbol{\Omega}) = \sum_{j=1}^{M} \mathrm{E}_{\boldsymbol{x}_k | \boldsymbol{\Upsilon}_k, \hat{\boldsymbol{\Omega}}_k^{(q)}} \left\{ -\frac{k}{2} \ln \prod_{i=1}^{N} \xi_{k,i}^{(j)} - \sum_{i=1}^{N} \frac{1}{2\xi_{k,i}^{(j)}} \sum_{d=1}^{k} \left(y_{d,i}^{(j)} - y_{0,i}^{(j)} - g_i^{(j)} x_{d,i} \right)^2 \right\}, \tag{23}$$

$$\begin{aligned} \ell_2^{(q)}(\boldsymbol{\Omega}) = \mathrm{E}_{\boldsymbol{x}_k | \boldsymbol{\Upsilon}_k, \hat{\boldsymbol{\Omega}}_k^{(q)}} \Bigg\{ & -\frac{k}{2} \ln |\boldsymbol{Q}_k| - \frac{1}{2} \sum_{d=1}^{k} \mathrm{tr} \left(\boldsymbol{Q}_k^{-1} \boldsymbol{x}_d \boldsymbol{x}_d^{\mathrm{T}} \right) \\ & + \frac{1}{2} \sum_{d=1}^{k} \mathrm{tr} \left\{ \boldsymbol{Q}_k^{-1} \left[\left(\boldsymbol{x}_d \boldsymbol{x}_{d-1}^{\mathrm{T}} \right)^{\mathrm{T}} + \boldsymbol{x}_d \boldsymbol{x}_{d-1}^{\mathrm{T}} - \boldsymbol{x}_{d-1} \boldsymbol{x}_{d-1}^{\mathrm{T}} \right] \right\} \\ & - \frac{1}{2} \sum_{d=1}^{k} \mathrm{tr} \left\{ \left(t_d^{\gamma_k} - t_{d-1}^{\gamma_k} \right) \boldsymbol{Q}_k^{-1} \left[-\boldsymbol{\lambda}_k \left(\boldsymbol{x}_d - \boldsymbol{x}_{d-1} \right)^{\mathrm{T}} \right] \right\} \\ & + \frac{1}{2} \sum_{d=1}^{k} \mathrm{tr} \left\{ \left(t_d^{\gamma_k} - t_{d-1}^{\gamma_k} \right) \boldsymbol{Q}_k^{-1} \left[\left(\boldsymbol{x}_d - \boldsymbol{x}_{d-1} \right) \boldsymbol{\lambda}_k^{\mathrm{T}} \right] \right\} \end{aligned}$$

$$-\frac{1}{2}\sum_{d=1}^{k}\mathrm{tr}\left[\left(t_d^{\gamma_k}-t_{d-1}^{\gamma_k}\right)^2\boldsymbol{Q}_k^{-1}\boldsymbol{\lambda}_k\boldsymbol{\lambda}_k^{\mathrm{T}}\right]\Bigg\}. \tag{24}$$

In the above function, define $\hat{\boldsymbol{x}}_d^{(q)}=\mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\Upsilon}_k,\hat{\boldsymbol{\Omega}}_k^q}(\boldsymbol{x}_d)=\hat{\boldsymbol{x}}_{d|k}$, $(\hat{\boldsymbol{x}}_d\hat{\boldsymbol{x}}_d^{\mathrm{T}})^{(q)}=\mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\Upsilon}_k,\hat{\boldsymbol{\Omega}}_k^{(q)}}(\boldsymbol{x}_d\boldsymbol{x}_d^{\mathrm{T}})=\hat{\boldsymbol{x}}_{d|k}\hat{\boldsymbol{x}}_{d|k}^{\mathrm{T}}+P_{d|k}$, and $(\hat{\boldsymbol{x}}_d\hat{\boldsymbol{x}}_{d-1}^{\mathrm{T}})^{(q)}=\mathrm{E}_{\boldsymbol{x}_k|\boldsymbol{\Upsilon}_k,\hat{\boldsymbol{\Omega}}_k^{(q)}}(\boldsymbol{x}_d\boldsymbol{x}_{d-1}^{\mathrm{T}})=\hat{\boldsymbol{x}}_{d|k}\hat{\boldsymbol{x}}_{d-1|k}^{\mathrm{T}}+P_{d,d-1|k}$. The Kalman smoother, an optimal estimation method for linear systems, is utilized to identify the conditional expectations of hidden degradation states [30–32]. The backward iteration can be given by

$$\hat{\boldsymbol{x}}_{d|k}=\hat{\boldsymbol{x}}_{d|d}+\boldsymbol{P}_{d|d}\boldsymbol{P}_{d+1|d}^{-1}\left[\hat{\boldsymbol{x}}_{d+1|k}-\hat{\boldsymbol{x}}_{d|d}-\hat{\boldsymbol{\lambda}}_k^q(t_{d+1}^{\hat{\gamma}_k^q}-t_d^{\hat{\gamma}_k^q})\right], \tag{25}$$

$$\boldsymbol{P}_{d|k}=\boldsymbol{P}_{d|d}+(\boldsymbol{P}_{d|d}\boldsymbol{P}_{d+1|d}^{-1})^2(\boldsymbol{P}_{d+1|k}-\boldsymbol{P}_{d+1|d}), \tag{26}$$

$$\boldsymbol{P}_{d,d-1|k}=\boldsymbol{P}_{d|d}\boldsymbol{P}_{d-1|d-1}\boldsymbol{P}_{d|d-1}^{-1}+\boldsymbol{P}_{d|d}\boldsymbol{P}_{d+1|d}^{-1}(\boldsymbol{P}_{d+1,d|k}-\boldsymbol{P}_{d|d})\boldsymbol{P}_{d-1|d-1}\boldsymbol{P}_{d|d-1}^{-1}, \tag{27}$$

where $\boldsymbol{P}_{k,k-1|k}=\boldsymbol{P}_{k|k}\boldsymbol{P}_{k|k-1}^{-1}\boldsymbol{P}_{k-1|k-1}$.

3.2.2 *The M-step*

The parameter identification is represented by

$$\hat{\boldsymbol{\Omega}}_k^{(q+1)}=\underset{\boldsymbol{\Omega}}{\mathrm{argmax}}\,\ell^{(q)}(\boldsymbol{\Omega}). \tag{28}$$

Considering the particular coupling relationship, we divide the unknown parameters into two groups, i.e., γ_k and $\{\boldsymbol{\lambda}_k,\boldsymbol{Q}_k,\xi_{k,i}^{(j)}\}$. According to (22)–(28), these parameters can be simply calculated by maximizing the expectation function. In detail, we first take the partial derivative with respect to $\gamma_k^{(q+1)}$ and set the value to zero. The formula can be expressed by

$$\begin{aligned}&\sum_{d=1}^{k}\left\{(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}})\ln\gamma_k^{(q+1)}\mathrm{tr}\left[\left(\boldsymbol{Q}_k^{(q)}\right)^{-1}\boldsymbol{\lambda}_k^{(q)}\left(\hat{\boldsymbol{x}}_d^{(q)}-\hat{\boldsymbol{x}}_{d-1}^{(q)}\right)^{\mathrm{T}}\right]\right\}\\&+\sum_{d=1}^{k}\left\{(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}})\ln\gamma_k^{(q+1)}\mathrm{tr}\left[\left(\boldsymbol{Q}_k^{(q)}\right)^{-1}\left(\hat{\boldsymbol{x}}_d^{(q)}-\hat{\boldsymbol{x}}_{d-1}^{(q)}\right)\left(\boldsymbol{\lambda}_k^{(q)}\right)^{\mathrm{T}}\right]\right\}\\&-\sum_{d=1}^{k}\left\{2(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}})^2\ln\gamma_k^{(q+1)}\mathrm{tr}\left[\left(\boldsymbol{Q}_k^{(q)}\right)^{-1}\boldsymbol{\lambda}_k^{(q)}\left(\boldsymbol{\lambda}_k^{(q)}\right)^{\mathrm{T}}\right]\right\}=0.\end{aligned} \tag{29}$$

Owing to the complexity in (29), we cannot obtain the analytical solution of $\gamma_k^{(q+1)}$. Here, it is estimated by a numerical method (please refer to the "fsolve" function in MATLAB). Specially, such method only applies to the nonlinear drift situation, i.e., $\gamma_k\neq 1$.

Then, for the other parameters, we can directly obtain

$$\hat{\boldsymbol{\lambda}}_k^{(q+1)}=\frac{1}{\sum_{d=1}^{k}(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}})^2}\sum_{d=1}^{k}\left[\left(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}}\right)\left(\hat{\boldsymbol{x}}_d^{(q)}-\hat{\boldsymbol{x}}_{d-1}^{(q)}\right)\right], \tag{30}$$

$$\begin{aligned}\hat{\boldsymbol{Q}}_k^{(q+1)}=\frac{1}{k}\sum_{d=1}^{k}\Bigg[&\left(\hat{\boldsymbol{x}}_d\hat{\boldsymbol{x}}_d^{\mathrm{T}}\right)^{(q)}-\left(\left(\hat{\boldsymbol{x}}_d\hat{\boldsymbol{x}}_{d-1}^{\mathrm{T}}\right)^{(q)}\right)^{\mathrm{T}}-\left(\hat{\boldsymbol{x}}_d\hat{\boldsymbol{x}}_{d-1}^{\mathrm{T}}\right)^{(q)}+\left(\hat{\boldsymbol{x}}_{d-1}\hat{\boldsymbol{x}}_{d-1}^{\mathrm{T}}\right)^{(q)}\\&-\left(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}}\right)\boldsymbol{\lambda}_k^{(q+1)}\left(\hat{\boldsymbol{x}}_d^{(q)}-\hat{\boldsymbol{x}}_{d-1}^{(q)}\right)^{\mathrm{T}}\\&-\left(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}}\right)\left(\hat{\boldsymbol{x}}_d^{(q)}-\hat{\boldsymbol{x}}_{d-1}^{(q)}\right)\left(\boldsymbol{\lambda}_k^{(q+1)}\right)^{\mathrm{T}}\end{aligned}$$

$$+\left(t_d^{\gamma_k^{(q+1)}}-t_{d-1}^{\gamma_k^{(q+1)}}\right)^2\boldsymbol{\lambda}_k^{(q+1)}\left(\boldsymbol{\lambda}_k^{(q+1)}\right)^{\mathrm{T}}\Bigg], \tag{31}$$

$$(\hat{\xi}_{k,i}^{(j)})^{(q+1)}=\frac{1}{k}\sum_{d=1}^{k}\left[\left(y_{d,i}^{(j)}-y_{0,i}^{(j)}\right)^2-2\left(y_{d,i}^{(j)}-y_{0,i}^{(j)}\right)g_i^{(j)}\hat{x}_{d,i}^{(q)}+\left(g_i^{(j)}\right)^2\left(\hat{\boldsymbol{x}}_d\hat{\boldsymbol{x}}_d^{\mathrm{T}}\right)_{i,i}^{(q)}\right]. \tag{32}$$

After the qth iteration, we should check the convergence of the identified parameters by

$$\parallel\hat{\boldsymbol{\Omega}}_k^{(q+1)}-\hat{\boldsymbol{\Omega}}_k^{(q)}\parallel^2<\Delta, \tag{33}$$

where Δ is a small positive number.

The above online estimation procedure can be summarized as Algorithm 1.

Algorithm 1 Multiple dependent hidden degradations identification

1: Initialization step: Initialize the degradation state vector $\boldsymbol{x}_0$ and the unknown parameter set $\hat{\boldsymbol{\Omega}}_0$;
2: Degradation state identification step: At the kth monitoring time, estimate the degradation states $\boldsymbol{x}_{k|k}$ by (5)–(20);
3: Parameter identification step: Update the parameter estimates $\hat{\boldsymbol{\Omega}}_k^{(q+1)}$ by (21)–(32);
4: Termination condition checking step: Check the convergence of the estimation. If the qth iteration satisfies (33), then go to the next monitoring time. Otherwise, set $q=q+1$, and then go to step 2.

4 RUL prediction

Taking the overall uptrend performance characteristic as an example, the RUL of the ith degradation process at time t_k is defined by

$$\mathrm{RUL}_k^{(i)}=\inf\left\{t:x_i(t+k\tau)\geqslant L_i|\hat{x}_{k|k}^{(i)},\Upsilon_k^{(i)}\right\}, \tag{34}$$

where L_i denotes the corresponding failure threshold.

Meanwhile, the reliability function of the ith degradation process can be characterized by

$$R_k^{(i)}(t)=P\left\{x_i(t+k\tau)<L_i,|\hat{x}_{k|k}^{(i)},\Upsilon_k^{(i)}\right\}. \tag{35}$$

4.1 Failure modes

We pay more attention to the prediction of the system-level RUL because one single performance characteristic may not represent the degradation degree of the whole system. It should be noted that the actual systems usually contain different failure modes or structures. Here, without loss of generality, we mainly consider three types of structures for multiple degradations.

4.1.1 *Case* 1

In this case, each degradation is critical, and the system is assumed to be failed as long as any degradation hits the failure threshold. The RUL of the series system is rewritten into

$$\mathrm{RUL}_k=\inf\{t:x_1(t+k\tau)\geqslant L_1,\mathrm{or}\ldots,\mathrm{or}\ x_N(t+k\tau)\geqslant L_N|\hat{\boldsymbol{x}}_{k|k},\boldsymbol{\Upsilon}_k\}. \tag{36}$$

The reliability function of the series system is given by

$$R_k(t)=P\{x_1(t+k\tau)<L_1,\ldots,x_N(t+k\tau)<L_N|\hat{\boldsymbol{x}}_{k|k},\boldsymbol{\Upsilon}_k\}. \tag{37}$$

4.1.2 *Case* 2

Compared with the series structure, one single characteristic falling out of the normal operating range cannot cause the failure of the parallel structure. Only if all of the degradations hit their thresholds, the system is assumed to be failed. The RUL of the parallel system is expressed by

$$\mathrm{RUL}_k = \inf\{t : x_1(t+k\tau) \geqslant L_1, \ldots, x_N(t+k\tau) \geqslant L_N | \hat{\boldsymbol{x}}_{k|k}, \boldsymbol{\Upsilon}_k\}. \tag{38}$$

The reliability function of the parallel system is

$$R_k(t) = P\{x_1(t+k\tau) < L_1, \text{or} \ldots, \text{or } x_N(t+k\tau) < L_N | \hat{\boldsymbol{x}}_{k|k}, \boldsymbol{\Upsilon}_k\}. \tag{39}$$

4.1.3 *Case* 3

In this case, the failure mode reflects a certain combination of multiple degradations. Taking summation structure as an example, the system is assumed to be failed if the sum of all the degradations hits the failure threshold. The RUL of the whole system is described by

$$\mathrm{RUL}_k = \inf\{t : x_1(t+k\tau) + \cdots + x_N(t+k\tau) \geqslant L | \hat{\boldsymbol{x}}_{k|k}, \boldsymbol{\Upsilon}_k\}. \tag{40}$$

The reliability function of this system is represented by

$$R_k(t) = P\{x_1(t+k\tau) + \cdots + x_N(t+k\tau) < L | \hat{\boldsymbol{x}}_{k|k}, \boldsymbol{\Upsilon}_k\}. \tag{41}$$

4.2 Monte Carlo simulation-based online RUL prediction

As stated above, copula functions are often used to deal with the dependency among multiple degradations. One superiority is that the PDF of the RUL can be obtained analytically. Totally different from copula functions, Eq. (4) directly describes the hidden dependency without any assumption.

But Eq. (4) results in the difficulty to obtain an analytical formula between the joint distribution of multiple degradations and their marginal distributions. This implies the analytical expressions of $R_k(t)$ for systems with the above three types of structures might not be available. Therefore, we utilize the Monte Carlo simulation method to predict the RUL, with the random sampling adopted for producing the observations [33]. Assume that ϱ denotes the sampling interval of the extrapolation. For any positive integer i, the simulation data at time t_k is firstly generated according to the joint distribution, given by

$$\boldsymbol{x}(t_k + i\varrho) = \boldsymbol{x}(t_k + (i-1)\varrho) + \boldsymbol{\lambda}_k \left[(t_k + i\varrho)^{\gamma_k} - (t_k + (i-1)\varrho)^{\gamma_k}\right] + \boldsymbol{\varepsilon}_k, \tag{42}$$

where $\boldsymbol{\varepsilon}_k \sim \mathcal{N}(\boldsymbol{0}, \varrho\hat{\boldsymbol{Q}}_k/\tau_k)$.

Secondly, we estimate the reliability function $R_k(t)$ from the generated simulation data. Thirdly, the corresponding cumulative distribution function (CDF) and the PDF of the RUL are calculated by $F_k(t) = 1 - R_k(t)$, and $f_k(t) = -\partial R_k(t)/\partial t$, respectively. The online RUL prediction method is described as Algorithm 2.

Algorithm 2 Monte Carlo simulation-based online RUL prediction

1: Initialization step: Initialize the sampling interval ϱ, and the discrete time index l;
2: Simulation data generating step: At the kth monitoring time, based on the identified parameter set $\hat{\boldsymbol{\Omega}}_k$, generate large amounts of samples $\hat{\boldsymbol{x}}(k\tau + l\varrho)$;
3: Reliability estimation step: According to (37), (39), and (41), calculate the reliability $R_k(l\varrho)$ from the generated samples;
4: RUL prediction step: If l is large enough, simulate the PDF of the RUL by taking the derivative of the discrete series $R_k(\varrho), R_k(2\varrho), \ldots, R_k(l\varrho)$. Otherwise, set $l = l + 1$, and then go to step 2.

5 Simulation examples

In this section, a numerical example and a case study are used to verify the validity of the proposed method.

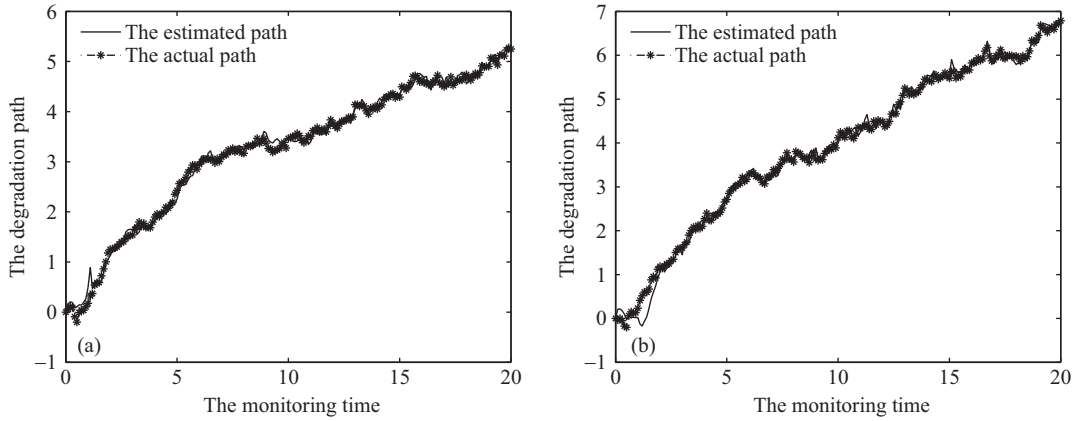

Figure 1 The actual and estimated degradation paths. (a) The first degradation; (b) the second degradation.

5.1 A numerical example

Because it is usually difficult to provide the physical model for the degradations in actual engineering systems, the data-driven model can be utilized to simulate the degradation processes, in which the hidden states and unknown parameters are identified by Algorithm 1. To test the availability of the parameter identification and the performance of the RUL prediction, the mean squared error (MSE) is used for the fitting model.

For convenience, assume $N = 2$, and $M = 2$. It means there exist two groups of sensors, and each group contains two sensors with the same sampling rate. The numerical data are generated by (4) with $\tau = 0.1, \gamma = 0.8, \boldsymbol{\lambda} = [0.5; 0.6], \boldsymbol{\Sigma} = [0.1, 0; 0, 0.2], \boldsymbol{V} = [0.4; 0.4], \boldsymbol{y}_0^{(1)} = \boldsymbol{y}_0^{(2)} = [0; 0], \boldsymbol{g}^{(1)} = [1.2, 0; 0, 1.1], \boldsymbol{g}^{(2)} = [1, 0; 0, 1], \boldsymbol{\Xi}^{(1)} = [0.01, 0; 0, 0.04], \boldsymbol{\Xi}^{(2)} = [0.02, 0; 0, 0.03]$. The thresholds of these two degradations are $L_1 = 5.2748$, and $L_2 = 6.7175$, therefore the first hitting time (FHT) is approximated at 200. Based on Algorithm 1, the estimated degradation paths are shown in Figure 1, together with the actual paths for comparison.

From Figure 1, it can be observed that the predicted paths and the actual ones match well with each other, which demonstrates the effectiveness of the proposed model. Correspondingly, the updated parameters in the degradation model including $\{\boldsymbol{\lambda}, \gamma, \boldsymbol{Q}\}$ are illustrated in Figure 2. All the parameters converge quickly to their actual values as the observations are accumulated. Obviously, the multi-sensor observations can play an important role in improving the rate of convergence.

Based on Algorithms 1 and 2, we achieve an online prediction including both the reliability and the RUL for a dynamic system, as soon as the parameters are updated in real time, using the newly obtained degradation data. Considering the diversity of failure modes, three different kinds of structures (namely, series, parallel, and summation structures) with respect to the above bivariate degradation processes are simulated in this numerical example, with the purpose of analyzing the possible influences on the RUL prediction. As previously mentioned, most existing predictions of multi-performance variables or multi-component systems utilize various copula functions to simulate the statistical dependencies without the direct and visible multivariate degradation modeling. Here, to reveal the superiorities of our method, we further use the same data to provide a comparative simulation, according to the copula-based methods presented by Wei [20]. It should be noted that the copula functions can only handle the issue of competing failures, corresponding to the series structure in this paper. Without loss of generality, two typical Archimedean copulas, i.e., the Frank copula and the Gumbel copula, are adopted for the simulation. The hidden dependencies of these two copulas all rely on one simple parameter, which can be estimated by the maximum likelihood (ML) algorithm. Taking the 170th and 190th monitoring time for example, the comparison of the RUL prediction results between the proposed method and the copula-based methods is shown in Figure 3, with the local magnification plotted in the small figures for clarity.

As illustrated in Figure 3, the actual RULs fall well within the range of all the predicted PDFs with our method, no matter which structure the system belongs to. It is obvious that the prediction

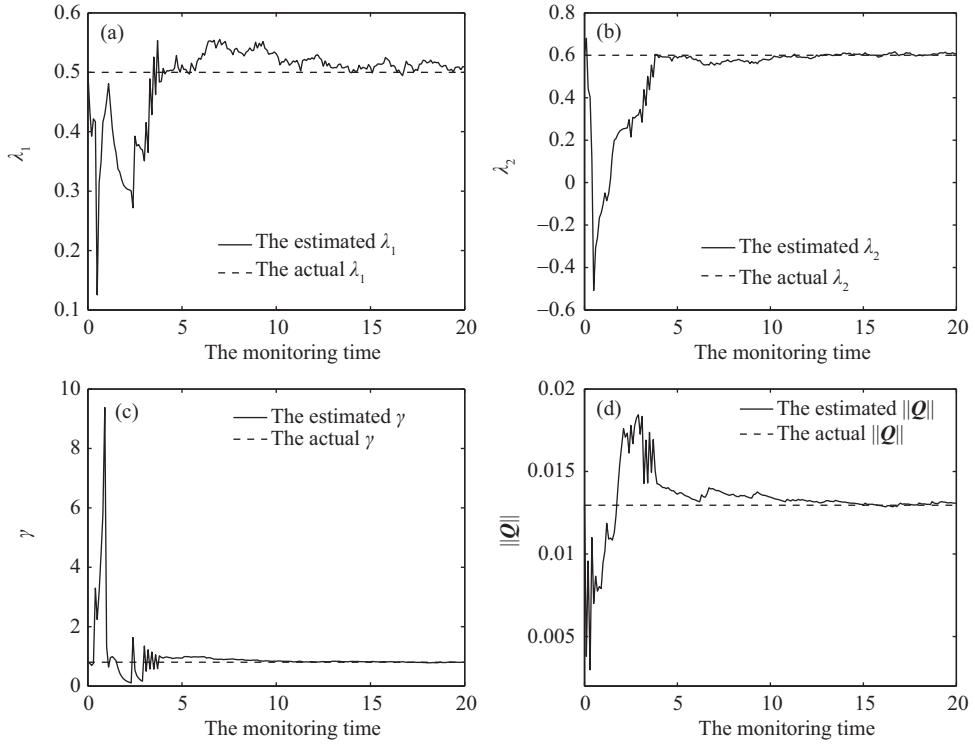

Figure 2 The updated parameters. (a) λ_1; (b) λ_2; (c) γ; (d) $\|\boldsymbol{Q}\|$.

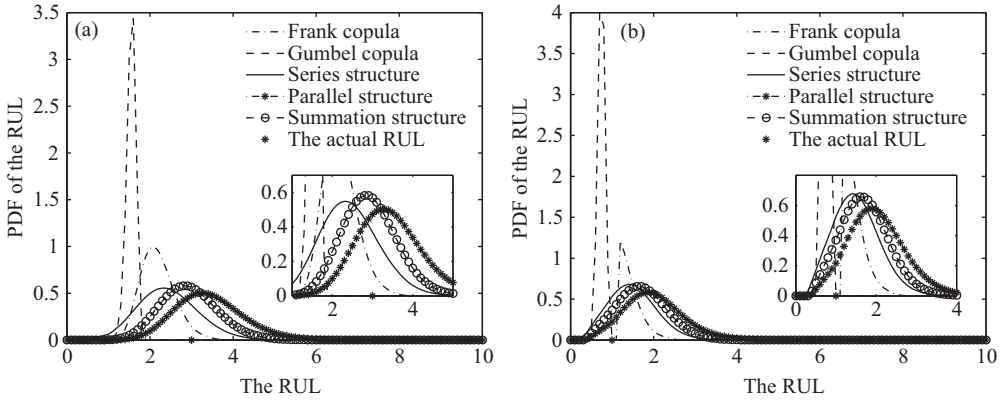

Figure 3 Comparison of the RUL prediction results with simulation data. (a) The 170th monitoring time; (b) the 190th monitoring time.

results for the series system are the most conservative, whereas for the parallel one. In other words, the competing failure results in the minimal expectations of the RULs. This is mainly because the failure condition is more rigorous for the series system. Meanwhile, the summation structure shows a neutral situation, for the reason that a linear combination of the two degradation processes is considered and a certain balance for the failure judgement is naturally made. Owing to some unknown uncertainties in the prediction procedure, the actual RULs deviate a bit from the peak of the predicted PDFs under these three structures. In particular, the shape of the PDFs at the 190th monitoring time is sharper than that at the 170th monitoring time. This means the obtained results tend to be more and more exact along with the accumulation of the sensor observations.

Moreover, the prediction results using the copula-based methods are also given in Figure 3 for comparison. We observe that the variances of these PDFs are relatively small, especially for the Gumbel copula. However, there are great differences between the corresponding expectations of the RULs and their truth values. Because the predicted PDFs are too thin and inaccurate, such results might go against

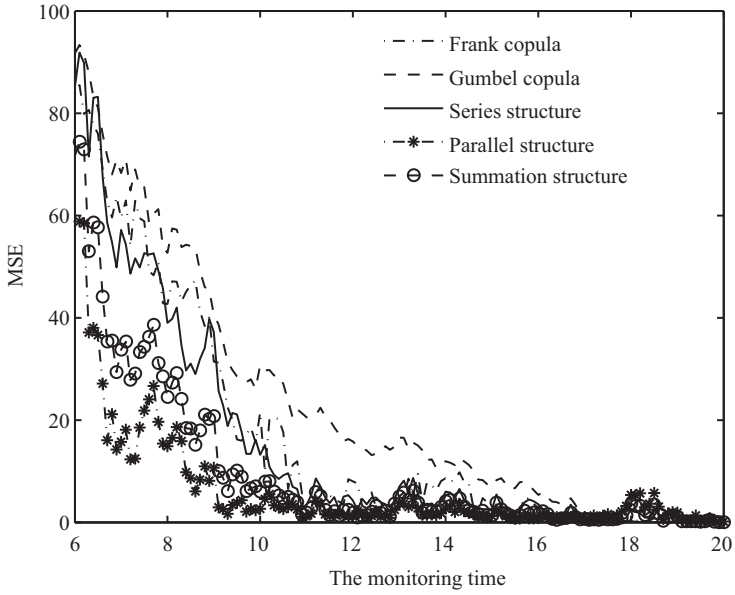

Figure 4 The MSEs of the RULs with simulation data.

the decision-making of the maintenances. In fact, these copula functions still employ the traditional one-dimensional Wiener process-based degradation model for the prediction, and hence cannot identify the real dependency structure established by the proposed multivariate model. After that, we utilize the MSE for quantitative analysis of the prediction precision with different methods. Specifically, the MSE at the kth monitoring time can be defined by

$$\mathrm{MSE}_k = \int_0^\infty (t - \widetilde{t}_k)^2 f_k(t)\mathrm{d}t, \tag{43}$$

where $\widetilde{t}_k$ denotes the actual RUL at the kth monitoring time, and $f_k(t)$ is the corresponding predicted PDF.

Considering that the PDFs obtained by Algorithm 2 are indeed discrete, here the MSEs are calculated approximatively through the numerical integration. For both the proposed method and the copula-based methods, the computation results under different failure modes are shown in Figure 4 (the Frank copula and the Gumbel copula belong to the series structure). It can be observed that the MSE curves of all the methods are relatively close. One major reason is that the simulation data appear simple enough without strong nonlinearity and measurement uncertainty. To be precise, most of the time our method obtains smaller MSEs compared with the other two copula functions, while the parallel structure usually shows the highest precision at the early stage owing to the mild failure criterion. In addition, although the generated degradation paths contain some fluctuations, the MSE value of our method decreases rapidly and eventually tends to zero with the prolongation of the monitoring time, which also implies the predicted PDFs of the RULs are robust, stable, and accurate to the moderate variation of observed data.

5.2 A case study

To demonstrate the proposed multivariate degradation model as well as the effectiveness of the corresponding RUL prediction method, a practical case study with respect to the large blast furnace is provided in this subsection. Because of the simple technological process and the huge production capacity, the large blast furnace has been regarded as an efficient smelting equipment since modern times, and it even plays a major role in the global ironmaking industry. Generally speaking, the effective body volume of a large blast furnace can achieve 1000 m^3. Such a scale implies an operation risk and the high maintenance costs. Because of the complicated internal physicochemical reactions, the furnace wall or lining will be gradually eroded by high temperatures and hot metals. In other words, they will become more and more

thin. Once the hearth burns through, the hot metals might flow out and cause significant losses of life and property. In fact, there are frequent accidents of blast furnaces around the world each year. Therefore, an accurate RUL prediction of the blast furnaces is very important for the sake of long-term safety.

As mentioned in Section 1, the large blast furnace is a complicated multi-component system containing the throat, stack, bosh, belly, and hearth. Different components may correspond to various degradation modes and have some uncertain dependency structures. Note that the thickness of the furnace wall is difficult to measure owing to technical limitations, but is usually conjectured through the variation tendency of the related temperature measurements, which means the temperature can be simply considered as a performance degradation for the RUL prediction. In view of the circumstance that the thermocouples are laid in advance and cannot be repaired when problems occur, the degradation processes of the system are monitored in real time by hundreds of sensors throughout the whole blast furnace. Specifically, because any burn through component results in a complete failure of the whole system, the failure type of blast furnaces can be regarded as a series structure.

Based on the above analysis, the issue of predicting the RUL of the large blast furnace can be solved by the proposed method. Here, the temperature measurements from a certain type of blast furnace (1000 m^3) in Jiayuguan, China, are utilized to test the practicability of (4) as well as our RUL prediction method. For simplicity, we select the sensor {10-1,10-2} as group one and sensor {11-1,11-2} as group two to analyze the dependencies between the bivariate degradation processes at different radial directions under a specific layout, e.g., the sensor 10-1 locates at the elevation of 7.896 m and the depth of 0.15 m, and the number represents the first sensor on the 10th floor. The sampling interval of all the sensors is 8 h, that is, one third of a day. To avoid confusion, we still set $\tau = 1$ in the simulation, which can be easily converted to other units. It should be noted that the mentioned blast furnace is still in service, and hence we cannot obtain the degradation paths from the entire life cycle. In this case, we only make use of the available data to give a trend analysis as a reference of the future work. All the adopted sensor measurements are shown in Figure 5. Obviously, the degradation data present a strong positive correlation, for the reason that high similarity really exists between each degradation path.

In this case study, we also use Eq. (4) to model the hidden degradation processes for the blast furnace, and then verify the RUL prediction method on the basis of the above data set. Before that, we simply subtract the initial value for all the data as a zero setting preprocessing procedure, i.e., $\boldsymbol{y}_0^{(1)} = \boldsymbol{y}_0^{(2)} = [0; 0]$. The unknown model parameters are estimated recursively through Algorithm 1, in which the measurement matrixes are set to $\boldsymbol{g}^{(1)} = \boldsymbol{g}^{(2)} = [1, 0; 0, 1]$ for convenience. Because the actual model of such data set is indeed unknowable, we have no valid prior knowledge on whether the drift term is linear or nonlinear. By comparing several different initial values of the nonlinear coefficient, we find that $\gamma = 1$ corresponds to a relatively satisfied convergence for all the parameters, and hence assume the drift term as the linear case for the degradation model. In this way, the drift coefficient and the process noise covariance are finally estimated at $\hat{\boldsymbol{\lambda}} = [1.5669; 0.0938]$, and $\hat{\boldsymbol{Q}} = [4.6317, 4.6285; 4.6285, 4.6284]$, respectively. As seen from $\hat{\boldsymbol{Q}}$, the value of each element is relatively large. This is mainly because the used degradation data fluctuate fiercely with considerable measurement noises. Besides, the elements on the primary diagonal are quite similar to the others, which implies the dependency between the two degradation processes is comparatively strong, conforming to our previously conjecture from Figure 5.

Similarly, the representative copula-based methods including the Frank copula and the Gumbel copula in [20] are also implemented as comparisons. The model parameters are estimated iteratively by the EM algorithm together with the ML algorithm. However, most of the parameters are difficult to converge under these methods, subject to the traditional univariate Wiener process-based degradation model. As a matter of fact, the copula-based methods only apply a simple statistical indicator called Kendall's rank correlation coefficient to describe the variation trends of different variables, which cannot well reflect the real dependencies from the model itself. In contrast, Eq. (4) represents the hidden dependencies more directly, and seems more strict and practical for the RUL prediction of complicated systems. Then, on the basis of the identified models and their correlation coefficients, the joint distributions of the RULs can be derived from the well-developed Inverse Gaussian distribution along with the used copulas.

At some monitoring time before the failure occurs, Figures 6(a) and (b) illustrate the RUL prediction

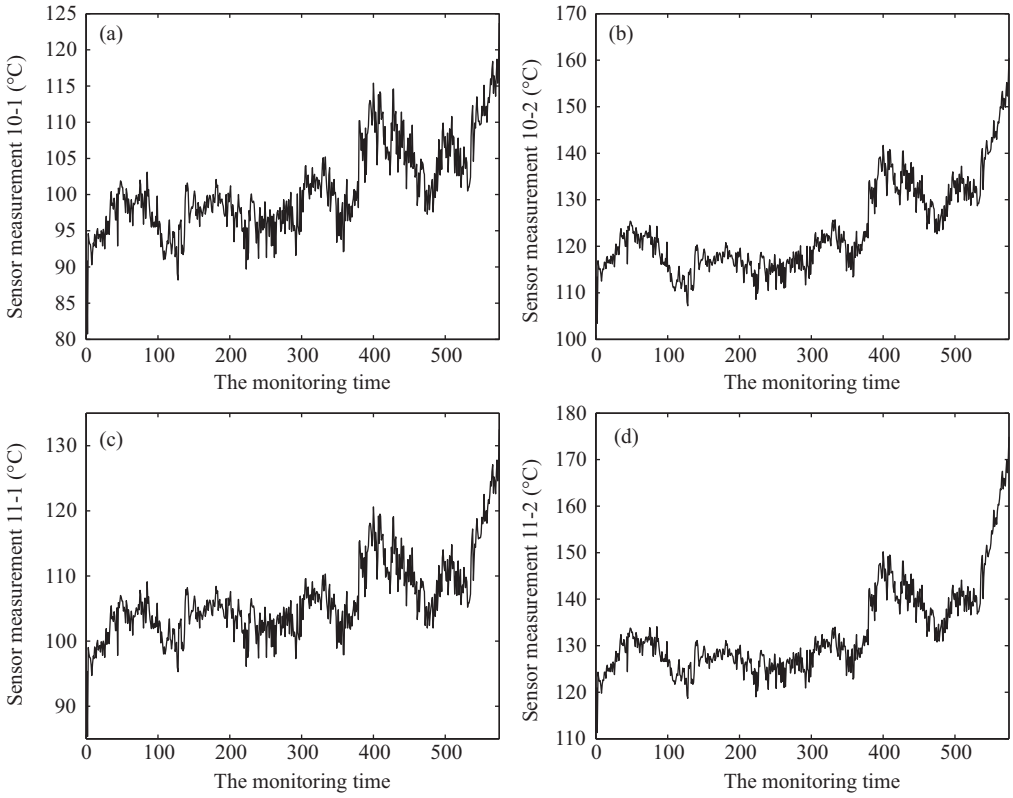

Figure 5 Temperature measurements of the blast furnace. (a) Sensor measurement 10-1; (b) sensor measurement 10-2; (c) sensor measurement 11-1; (d) sensor measurement 11-2.

results using the Frank copula and the Gumbel copula, respectively. Using Algorithm 2, the predicted PDFs of the RULs under (4) are also shown in Figure 6(c), where the thresholds are set to $L_1 = 895.6716$ and $L_2 = 50.0965$ from some prior information with respect to the historical degradation data, and the FHT can be approximated at the 575th monitoring time, i.e., 4600 h. Compared with the actual RULs, it is quite clear that only our method presents the relatively accurate prediction results. The PDFs in Figures 6(a) and (b) completely deviate from their truth expectations especially at the early stage. Particularly, the shape of the PDFs in Figure 6(c) becomes sharper and sharper as the time goes on, which means the progressively smaller variance or the more exact prediction results. For further comparative analysis, the MSEs of the RULs with the blast furnace degradation data are illustrated in Figure 7. Benefitting from the excavation of dependency structure in the diffusion term, the value of the MSE under our method is always the smallest, and decreases rapidly until close to zero, while the copula-based methods show irregular fluctuations, which might be less stable for the online monitoring and prognostics. In general, this case study fully verifies the effectiveness of the proposed method.

6 Conclusion

In this paper, we consider the problem of the RUL prediction in the presence of multiple dependent hidden degradations. We introduce a new type of state space model, in which the dependencies among multiple degradations are represented by a diffusion coefficient matrix. One superiority is that the dependency structure can be directly estimated based on the sensor measurements. Although there exist some realistic problems such as multi-source information fusion, we can still obtain the estimation of unknown parameters by using the sequential Kalman filtering and EM algorithm. We also use the Monte Carlo simulation method to predict the RUL for the systems of different structures. Simulation results fully verify the effectiveness of the proposed approach.

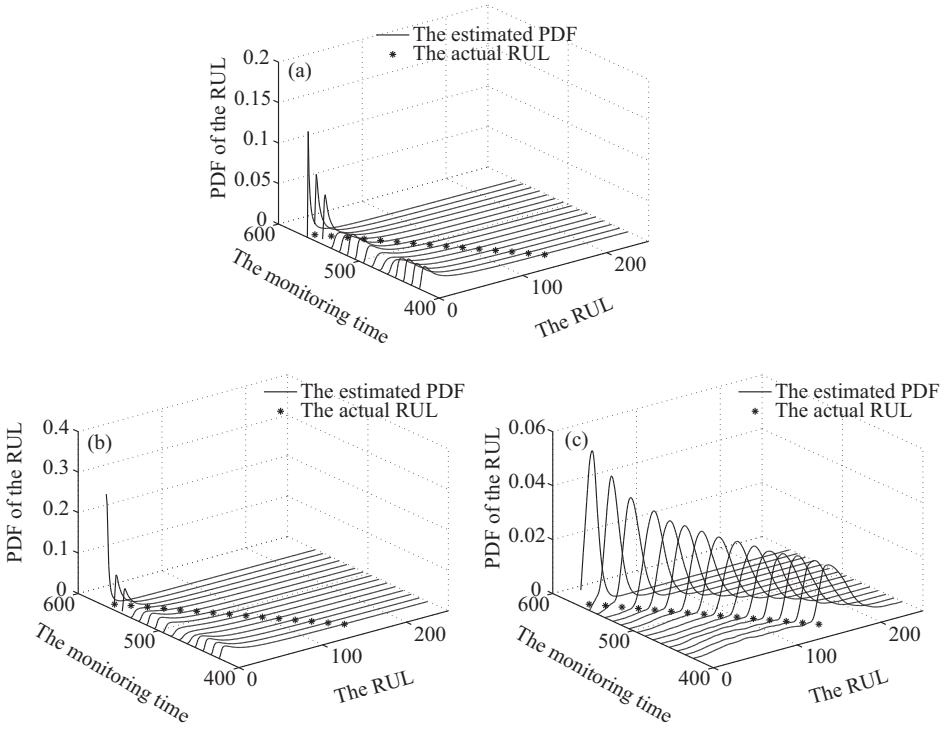

Figure 6 Comparison of the RUL prediction results with the blast furnace degradation data. (a) Frank copula; (b) Gumbel copula; (c) our method.

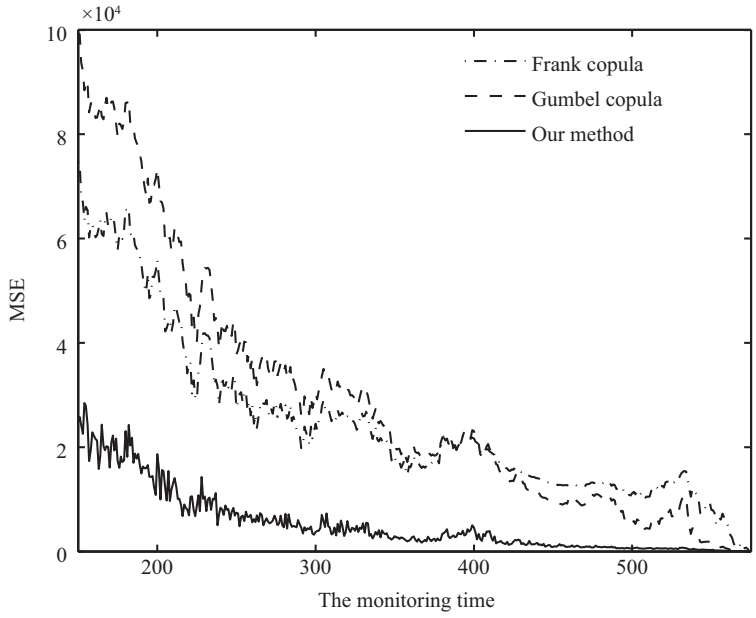

Figure 7 The MSEs of the RULs with the blast furnace degradation data.

Although the examples show that the proposed model is useful for the RUL prediction, there are still many challenges to be addressed. First, we only focus on the Wiener process-based degradation modeling, but various types of degradation processes may exist in actual systems. Second, regarding the phenomenon of losing data, the accuracy of the parameter identification is seriously affected. In addition, the fusion strategy of multi-sensor observations is preliminary in this paper. How to utilize the available information more effectively for the RUL prediction still remains an open problem.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61490701, 61290324, 61473164) and Research Fund for the Taishan Scholar Project of Shandong Province of China.

References

1 Sikorska J Z, Hodkiewicz M, Ma L. Prognostic modelling options for remaining useful life estimation by industry. Mech Syst Signal Process, 2011, 25: 1803–1836
2 Okoh C, Roy R, Mehnen J, et al. Overview of remaining useful life prediction techniques in through-life engineering services. Procedia CIRP, 2014, 16: 158–163
3 Xi X P, Chen M Y, Zhou D H. Remaining useful life prediction for degradation processes with memory effects. IEEE Trans Rel, 2017, 66: 751–760
4 Goebel K F. Management of uncertainty in sensor validation, sensor fusion, and diagnosis of mechanical systems using soft computing techniques. Dissertation for Ph.D. Degree. Berkeley: University of California, 1996
5 Ahmadzadeh F, Lundberg J. Remaining useful life estimation: review. Int J Syst Assur Eng Manag, 2014, 5: 461–474
6 Si X, Wang W, Hu C, et al. Remaining useful life estimation-a review on the statistical data driven approaches. Eur J Oper Res, 2011, 213: 1–14
7 Wei M H, Chen M Y, Zhou D H. Multi-sensor information based remaining useful life prediction with anticipated performance. IEEE Trans Rel, 2013, 62: 183–198
8 Niu G, Yang B S. Intelligent condition monitoring and prognostics system based on data-fusion strategy. Expert Syst Appl, 2010, 37: 8831–8840
9 Li R, Ryan J K. A Bayesian inventory model using real-time condition monitoring information. Prod Oper Manage, 2011, 20: 754–771
10 Tang S, Yu C, Wang X, et al. Remaining useful life prediction of lithium-ion batteries based on the wiener process with measurement error. Energies, 2014, 7: 520–547
11 Bian L, Gebraeel N. Stochastic framework for partially degradation systems with continuous component degradation-rate-interactions. Naval Res Log, 2014, 61: 286–303
12 Wang X, Guo B, Cheng Z. Residual life estimation based on bivariate Wiener degradation process with measurement errors. J Cent South Univ, 2012, 20: 1844–1851
13 Xi Z M, Jing R, Wang P F, et al. A copula-based sampling method for data-driven prognostics. Reliability Eng Syst Saf, 2014, 132: 72–82
14 Shi A H, Zeng J C. Real-time prediction of remaining useful life and preventive opportunistic maintenance strategy for multi-component systems considering stochastic dependence. Comput Industrial Eng, 2016, 93: 192–204
15 Khorasgani H, Biswas G, Sankararaman S. Methodologies for system-level remaining useful life prediction. Reliab Eng Syst Saf, 2016, 154: 8–18
16 Rodrigues L R. Remaining useful life prediction for multiple-component systems based on a system-level performance indicator. IEEE/ASME Trans Mechatron, 2018, 23: 141–150
17 Prakash O, Samantaray A K, Bhattacharyya R. Model-based multi-component adaptive prognosis for hybrid dynamical systems. Control Eng Practice, 2018, 72: 1–18
18 Mercier S, Pham H H. A preventive maintenance policy for a continuously monitored system with correlated wear indicators. Eur J Oper Res, 2012, 222: 263–272
19 Wang X L, Guo B, Cheng Z J. Residual life estimation based on bivariate Wiener degradation process with time-scale transformations. J Stat Comput Simul, 2014, 84: 545–563
20 Wei M H. Multi-sensor monitoring information based remaining useful life prediction for industrial equipments. Dissertation for Ph.D. Degree. Beijing: Tsinghua University, 2013
21 Liao H, Elsayed E A. Reliability inference for field conditions from accelerated degradation testing. Naval Res Log, 2006, 53: 576–587
22 Trevisanello L, Meneghini M, Mura G, et al. Accelerated life test of high brightness light emitting diodes. IEEE Trans Device Mater Relib, 2008, 8: 304–311
23 Ye Z S, Wang Y, Tsui K L, et al. Degradation data analysis using wiener processes with measurement errors. IEEE Trans Rel, 2013, 62: 772–780
24 Wang X, Balakrishnan N, Guo B. Residual life estimation based on a generalized Wiener degradation process. Reliability Eng Syst Saf, 2014, 124: 13–23
25 Ye Z S, Chen N, Shen Y. A new class of Wiener process models for degradation analysis. Reliability Eng Syst Saf, 2015, 139: 58–67
26 Xi X P, Chen M Y, Zhou D H. Online prognostics based on multiple dependent degradation processes. In: Proceedings of Prognostics and System Health Management Conference (PHM), Harbin, 2017. 1–6
27 Kao Y H, van Roy B. Directed principal component analysis. Oper Res, 2014, 62: 957–972
28 Tipping M E, Bishop C M. Probabilistic principal component analysis. J R Stat Soc B, 1999, 61: 611–622
29 Chui C K, Chen G. Kalman Filtering. Berlin: Springer, 2009
30 Shumway R H, Stoffer D S. Time Series Analysis and Its Applications. New York: Springer, 2000
31 Houtekamer P L, Mitchell H L. Ensemble Kalman filtering. Q J R Meteorol Soc, 2005, 131: 3269–3289
32 Sun S L. Multi-sensor optimal fusion fixed-interval Kalman smoothers. Inf Fusion, 2008, 9: 293–299
33 Lemieux C. Monte Carlo and Quasi-Monte Carlo Sampling. New York: Springer, 2009