



Interfaces with Other Disciplines

Inverse Gaussian processes with correlated random effects for multivariate degradation modeling

Guanqi Fang^{a,b,*}, Rong Pan^c, Yukun Wang^d^a School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China^b Collaborative Innovation Center of Statistical Data Engineering, Technology & Application, Zhejiang Gongshang University, Hangzhou 310018, China^c School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85281, USA^d School of Economics and Management, Tianjin Chengjian University, Tianjin 300384, China

ARTICLE INFO

Article history:

Received 1 April 2021

Accepted 22 October 2021

Available online 29 October 2021

Keywords:

Reliability

Degradation process

Dependence modeling

EM algorithm

Lifetime distribution

Multivariate model

Random effects

ABSTRACT

Many engineering products have more than one failure mode and the evolution of each mode can be monitored by measuring a performance characteristic (PC). It is found that the underlying multi-dimensional degradation often occurs with inherent process stochasticity and heterogeneity across units, as well as dependency among PCs. To accommodate these features, in this paper, we propose a novel multivariate degradation model based on the inverse Gaussian process. The model incorporates random effects that are subject to a multivariate normal distribution to capture both the unit-wise variability and the PC-wise dependence. Built upon this structure, we obtain some mathematically tractable properties such as the joint and conditional distribution functions, which subsequently facilitate the future degradation prediction and lifetime estimation. An expectation-maximization algorithm is developed to infer the model parameters along with the validation tools for model checking. In addition, two simulation studies are performed to assess the performance of the inference method and to evaluate the effect of model misspecification. Finally, the application of the proposed methodology is demonstrated by two illustrative examples.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Background

Modern engineering products are often highly reliable. For example, according to the photovoltaic (PV) industry standard, solar panels are designed to work with a performance reduction of no more than 20% within the first 25 years if operated at the standard test condition (Lindig, Kaaya, Weiss, Moser, & Topic, 2018). To accurately assess the quality of these products, degradation tests that continuously monitor some performance characteristics (PCs) over time are often applied to provide timely failure information. The PCs such as output power, cell crack size, and the strength of encapsulant adhesion to glass for PV modules reveal some commonly-observed failure modes, including material fatigue and delamination, etc. Along with the occurrence of these failure modes, the stochastic nature of innate failure mechanisms makes the degradation process develop with a considerable

amount of uncertainty. In the meantime, heterogeneity across engineering products is often observed due to variations in raw materials and manufacturing processes even if these products are produced according to a common industry standard. This type of variability leads to the unit-to-unit variation among observed degradation processes. Besides this, it is further worth noting that any likely correlation among different failure modes will cause the PCs to behave not independently. Therefore, to analyze multivariate degradation data, we need to establish a statistical model that accounts for 1) the stochastic nature of each individual PC, 2) the heterogeneity among different units, and more importantly, 3) any possible dependence among these PCs. The main goal of this paper is to develop such a multivariate degradation model.

1.2. A motivating example

As a motivating example, we present a real dataset – the coating data that was initially provided by Lu, Wang, Hong, & Ye (2020). This dataset describes a 3-dimensional degradation process of a type of polymeric material over time. By exposing several test units to various environmental conditions of ultraviolet (UV) radiation, temperature, and relative humidity (RH), the three differ-

* Corresponding author at: School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China.

E-mail address: gfang5@asu.edu (G. Fang).

List of Notations

Indices:

- n number of units ($i \in \{1, 2, \dots, n\}$)
 p number of performance characteristics ($j \in \{1, 2, \dots, p\}$)
 m_i number of measurements ($k \in \{1, 2, \dots, m_i\}$)

Random variables:

- $\mathbf{Y}_i(t)$ degradation values on unit i at time t , where $\mathbf{Y}_i(t) = (Y_{i1}(t), Y_{i2}(t), \dots, Y_{ip}(t))'$
 δ_i element-wise inverse of the vector of the drift parameters (i.e. unobserved random effects), where $\delta_i = (\delta_{i1}, \delta_{i2}, \dots, \delta_{ip})' \equiv (\frac{1}{\mu_{i1}}, \frac{1}{\mu_{i2}}, \dots, \frac{1}{\mu_{ip}})'$

Data:

- t_{ik} the elapsed time when the k -th measurement on unit i is taken
 $\tilde{t}_{ik}$ measurement time interval with $\tilde{t}_{ik} = t_{ik} - t_{i,k-1}$
 y_{ijk} degradation measurement of process j on unit i at time point k
 $\tilde{y}_{ijk}$ degradation increment with $\tilde{y}_{ijk} = y_{ijk} - y_{i,j,k-1}$
 $\tilde{\mathbf{t}}$ collection of all the measurement time intervals, where $\tilde{\mathbf{t}} = (\tilde{t}_1', \tilde{t}_2', \dots, \tilde{t}_n')'$ and $\tilde{\mathbf{t}}_i = (\tilde{t}_{i1}, \tilde{t}_{i2}, \dots, \tilde{t}_{im_i})'$

 $\tilde{\mathbf{Y}}$

collection of all the degradation increments, where $\tilde{\mathbf{Y}} = (\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2, \dots, \tilde{\mathbf{Y}}_n)$ and $\tilde{\mathbf{Y}}_i = (\tilde{y}_{i1}, \tilde{y}_{i2}, \dots, \tilde{y}_{im_i})' = \begin{pmatrix} \tilde{y}_{i11} & \tilde{y}_{i12} & \dots & \tilde{y}_{i1m_i} \\ \tilde{y}_{i21} & \tilde{y}_{i22} & \dots & \tilde{y}_{i2m_i} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{ip1} & \tilde{y}_{ip2} & \dots & \tilde{y}_{ipm_i} \end{pmatrix}$

 $\mathbb{D}$

the total observed data with $\mathbb{D} = \{\tilde{\mathbf{Y}}, \tilde{\mathbf{t}}\}$

 $\mathcal{D}$

vector of all PC failure thresholds with $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_p)'$

Parameters:

- $\boldsymbol{\eta}$ means of the random effects, where $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_p)'$
 $\boldsymbol{\sigma}$ standard deviations of the random effects, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_p)'$
 $\boldsymbol{\rho}$ correlations of the random effects, where $\boldsymbol{\rho} = (\rho_{12}, \rho_{13}, \dots, \rho_{p-1,p})'$
 $\boldsymbol{\Sigma}$ variance-covariance matrix of the random effects, where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \dots & \sigma_p^2 \end{pmatrix}$$

- $\boldsymbol{\lambda}$ the shape parameters, where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)'$

- $\boldsymbol{\gamma}$ the time-scale transformation parameters, where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_p)'$

- $\boldsymbol{\theta}$ collection of all the unknown parameters, where $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}', \boldsymbol{\rho}', \boldsymbol{\lambda}')'$ or $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}', \boldsymbol{\rho}', \boldsymbol{\lambda}', \boldsymbol{\gamma}')$ if applicable

Transformed Quantities:

- τ_{ijk} transformed time scale with $\tau_{ijk} = \Lambda_j(t_{ik}) - \Lambda_j(t_{i,k-1})$

- $\mathbf{v}_i$ $\mathbf{v}_i = (t/y_{i1}, t/y_{i2}, \dots, t/y_{ip})'$
 $\mathbf{V}_i$ $\mathbf{V}_i = \begin{pmatrix} \lambda_1 y_{i1} & 0 & 0 & 0 \\ 0 & \lambda_2 y_{i2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p y_{ip} \end{pmatrix}$

- $\tilde{\mathbf{v}}_i$ $\tilde{\mathbf{v}}_i = (s/\tilde{y}_{i1}, s/\tilde{y}_{i2}, \dots, s/\tilde{y}_{ip})'$ or $(\tilde{t}_{ik}/\tilde{y}_{i1k}, \tilde{t}_{ik}/\tilde{y}_{i2k}, \dots, \tilde{t}_{ik}/\tilde{y}_{ipk})'$

- $\tilde{\mathbf{V}}_i$ $\tilde{\mathbf{V}}_i = \begin{pmatrix} \lambda_1 \tilde{y}_{i1} & 0 & 0 & 0 \\ 0 & \lambda_2 \tilde{y}_{i2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \tilde{y}_{ip} \end{pmatrix}$ or $\begin{pmatrix} \lambda_1 \tilde{y}_{i1k} & 0 & 0 & 0 \\ 0 & \lambda_2 \tilde{y}_{i2k} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \tilde{y}_{ipk} \end{pmatrix}$

ent PCs (change of chemical structures at the wavelength of 1250 cm^{-1} , 1510 cm^{-1} , 2925 cm^{-1}) were measured every few days using Fourier transform infrared (FTIR) spectroscopy. For illustrative purposes, we arbitrarily pick a subset of the data that was generated under the environmental setting – 60% UV intensity, 35°C, and 0% RH. The resulting degradation paths are depicted in Fig. 1, in which PC1, PC2, and PC3 represent the three PCs at the wavelength of 1250 cm^{-1} , 1510 cm^{-1} , and 2925 cm^{-1} , respectively. This sub-dataset consists of a sample size of four test units and forty-seven observations. As illustrated by Fig. 1, for each PC, disparate degradation paths are shown, which clearly indicates the heterogeneity among different test units. Furthermore, the relative ranking of the degradation paths among different units is preserved across the PCs. This implies the possible existence of dependence among the PCs. Thus, the desire for building a multivariate degradation model naturally arises when one is concerned with such monotonic degradation paths where both unit-wise variability and PC-wise dependence are present.

1.3. Related work

In the statistical degradation modeling literature, numerous studies have been conducted to model a single PC's degradation

process. Typical frameworks include general path models and stochastic process models (Ye & Xie, 2015), where a general path model is a time-based regression model that incorporates both fixed and random effects (Bae & Kvam, 2004; Fang, Rigdon, & Pan, 2018) and a stochastic process model assumes a random process for PC measurements over time. In the literature, there are three main classes of stochastic degradation processes – the Wiener process (Hong, Tan, & Ye, 2020; Ye, Wang, Tsui, & Pecht, 2013), the Gamma process (Lawless & Crowder, 2004; Ye, Xie, Tang, & Chen, 2014), and the inverse Gaussian (IG) process (Morita et al., 2020; Peng, 2015; Wang & Xu, 2010; Ye & Chen, 2014).

Built upon these two major modeling frameworks, several researchers have developed multivariate degradation models. Si, Yang, Wu, & Chen (2018) proposed a multivariate general path model to make reliability analysis for materials with deformation process. Xu, Shen, Wang, & Tang (2018a), Wang, Balakrishnan, & Guo (2015), and Liu, Al-Khalifa, Elsayed, Coit, & Hamouda (2014) constructed either bivariate or multivariate degradation models based on the Wiener process. More recent studies include the works by Sun, Ye, & Hong (2020b), Lu et al. (2020), Hajiha, Liu, & Hong (2020), and Liu, Pandey, Wang, & Zhao (2021). Sun et al. (2020b) proposed a multivariate model with two-layer block effects based on the Wiener process to carry out an in-depth study

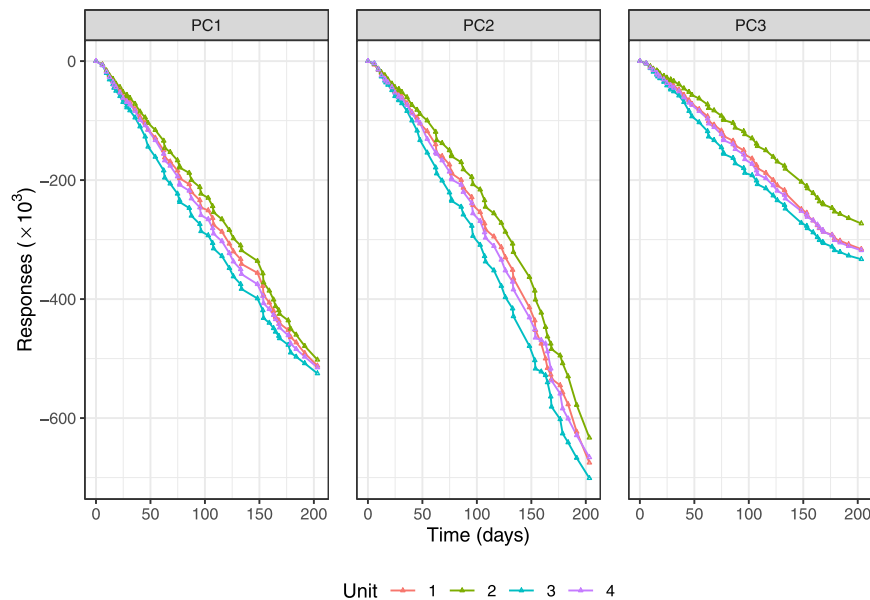


Fig. 1. Degradation Paths of Polymeric Materials.

on a destructive degradation testing dataset. Lu et al. (2020) proposed a multivariate general path model to analyze the aforementioned trivariate polymeric material's degradation process. Hajiha et al. (2020) used the Wiener process to analyze competing degradation processes under dynamic operation conditions. Liu et al. (2021) constructed a bivariate gamma process by trivariate reduction method to build condition-based maintenance policies. Besides these works, an alternative approach to constructing multivariate degradation models is to take advantage of the copula theory. Some developments in this regard have the articles based on regular copulas by Fang, Pan, & Hong (2020), Fang & Pan (2021), Xu et al. (2018b), Sun, Fu, Liao, & Xu (2020a), Palayangoda & Ng (2021). It also includes ways of constructing multi-dimensional Lévy processes with Lévy copulas, such as Shi, Feng, Shu, & Xiang (2020a), Mercier & Pham (2012) and Li, Deloux, & Dieulle (2016).

Despite the previous studies, there is a lack of IG process-based multivariate degradation modeling framework. This paper seeks to bridge the gap. Particularly, the proposed novel degradation model takes advantage of a structure of correlated random effects, which brings tractable properties and facilitates easy application for the future degradation prediction and lifetime estimation. Moreover, we provide an efficient inference method and model validation tools. With these nice features, our proposed methodology greatly complements the family of multivariate degradation models, especially when the degradation process is monotone.

1.4. Outline

The remainder of this paper is organized as follows. Section 2 elaborates on the IG process-based multivariate modeling framework and its derived lifetime distribution. Section 3 discusses the model parameters estimation and validation aspects. Section 4 provides two simulation studies and Section 5 demonstrates two illustrative examples. Finally, Section 6 concludes the paper with a discussion of techniques of building multivariate degradation models and remarks on future study. Some relevant mathematical proofs, algorithm details, and additional results of the illustrative examples are given in Appendices.

2. Modeling framework and lifetime distribution

2.1. Model description

Many degradation processes, as accumulation of additive and irreversible damages, demonstrate monotone paths (Ye & Chen, 2014). The IG process, as an attractive model to describe these processes, has been well studied. An IG process, $\{Y(t), t \geq 0\}$, satisfies the following properties:

- $Y(t)$ has independent increments, i.e. $Y(t + \Delta t) - Y(t)$ is independent of $Y(s)$, $\forall t \geq s > 0$ and $\Delta t \geq 0$;
- $Y(t + \Delta t) - Y(t)$ is subject to an IG distribution $IG(\mu \Delta t, \lambda \Delta t^2)$, $\forall t \geq 0$ and $\Delta t > 0$.

Thus, without loss of generality, let $Y(0) = 0$, then the probability density function (pdf) of $Y(t)$ is given by

$$f_{Y(t)}(y; \mu, \lambda, t) = \sqrt{\frac{\lambda t^2}{2\pi y^3}} \exp\left[-\frac{\lambda(y - \mu t)^2}{2\mu^2 y}\right], \quad y > 0, \quad (1)$$

where $\mu > 0$ is the drift parameter and $\lambda > 0$ is the shape parameter.

Consider a degradation test with n test units and each unit presents a p -dimensional multivariate degradation process. Denote the degradation observations on unit i , $i = 1, 2, \dots, n$, at time t , $t \geq 0$, by $\mathbf{Y}_i(t) = (Y_{i1}(t), Y_{i2}(t), \dots, Y_{ip}(t))'$ with each individual observation $Y_{ij}(t)$, $\forall i = 1, 2, \dots, n, j = 1, 2, \dots, p$. We assume $Y_{ij}(t)$ is subject to the IG process as presented by Eq. (1). Meanwhile, it is assumed that correlated random effects exist among these degradation processes. In summary, we refer to the modeling framework as Model M_0 and it is represented by

$$M_0 : \begin{cases} Y_{ij}(t) \sim IG(t/\delta_{ij}, \lambda_j t^2) \\ \boldsymbol{\delta}_i \sim MVN(\boldsymbol{\eta}, \boldsymbol{\Sigma}) \end{cases},$$

where $\boldsymbol{\delta}_i = (\delta_{i1}, \delta_{i2}, \dots, \delta_{ip})' \equiv (\frac{1}{\mu_{i1}}, \frac{1}{\mu_{i2}}, \dots, \frac{1}{\mu_{ip}})'$ is the element-wise inverse of the vector of the drift parameters and it is subject to a multivariate normal (MVN) distribution with mean vector $\boldsymbol{\eta}$ and variance-covariance matrix $\boldsymbol{\Sigma} > 0$, i.e. a positive definite (pd)

matrix. We denote the elements in $\boldsymbol{\eta}$ and $\boldsymbol{\Sigma}$ by

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_p \end{pmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_p^2 \end{pmatrix},$$

respectively, where $\eta_j > 0$ and $\sigma_j^2 \geq 0$, $\forall j = 1, 2, \dots, p$, are the mean and variance, respectively. Furthermore, $\sigma_{jj'} = \rho_{jj'}\sigma_j\sigma_{j'}$, $\forall 1 \leq j < j' \leq p$, and it is the covariance with the Pearson correlation coefficient $\rho_{jj'} \in [-1, 1]$. Under the extreme scenario of $\boldsymbol{\eta}$ being fixed and $\sigma_j \rightarrow 0$, $\forall j = 1, 2, \dots, p$, the model reduces to the routine IG process. We denote all the unknown parameters by $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}', \boldsymbol{\rho}', \boldsymbol{\lambda}')$, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_p)'$, $\boldsymbol{\rho} = (\rho_{12}, \rho_{13}, \dots, \rho_{p-1,p})'$, and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)'$. The selection of the joint distribution of $\boldsymbol{\delta}_i$ is because of its natural conjugate property of the IG distribution. This feature not only provides a good fit for degradation data as illustrated by the subsequent illustrative examples in Section 5, but also leads the proposed model to be mathematically tractable.

Note that, for convenience, we directly make use of the original time scale t in the formulas above. Instead, a more general notation is $\Lambda(t)$, where $\Lambda(\cdot)$ is a function that transforms the original time scale so as to model any possible nonlinear degradation processes. Two popular choices of $\Lambda(\cdot)$ are the power law function and the exponential law function (Whitmore & Schenkelberg, 1997). We will illustrate how to incorporate $\Lambda(\cdot)$ in a later section. Moreover, similar to the assumption made by Lu & Meeker (1993) and Peng (2015), we assume that the probability of a negative μ_{ij} is negligible to avoid obtaining an infeasible degradation rate. Concretely, if $\eta_j \gg 0$ and σ_j is small, this problem is not a big concern.

By constructing such a hierarchical modeling framework, the aforementioned three features of variability among multivariate degradation measurements have been accommodated. First, at the top level, given μ_{ij} , a well-defined IG process is proposed to explain the evolution of each individual degradation process over time with randomness. This type of randomness is usually produced by the inherent failure mechanism. Second, at the second level, $\boldsymbol{\delta}_i$, a collection of the inverse of all μ_{ij} 's, is assumed to be stochastically distributed. This type of randomness manifests the latent variation of the drift parameters across units. Note that the IG process $Y_{ij}(t)$ has mean $\mu_{ij}t$ and variance $\mu_{ij}^3 t / \lambda_j$, where the parameter μ_{ij} and the combined parameter μ_{ij}^3 / λ_j represent the degradation rate and volatility, respectively. Therefore, through incorporating such random effects, a parsimonious model that could explain the unit-to-unit variability of both the two attributes is obtained. Finally, it is noted that, by modeling the random effects through a MVN distribution, a plausible way to incorporate the dependence among PCs is also formed. This type of dependence is associated with the drift parameters μ_{ij} 's that underpin the underlying degradation mechanism. Meanwhile, given μ_{ij} 's, the marginal degradation processes are conditionally independent, which brings many nice properties and facilitates the associated inference as illustrated in later sections. Therefore, the proposed modeling framework is applicable to the analysis of multivariate degradation processes with both unit-wise heterogeneity and PC-wise dependence.

2.2. Derived properties

Next, we present some theoretical properties that can be derived from Model M_0 in terms of the marginal, joint, and conditional pdf. Furthermore, the conditional distributions of both the random-effects terms and the future degradation predictions given the history of observations are specified, too.

First, note that under such a modeling framework, the marginal distribution of δ_{ij} is a normal distribution with mean η_j and variance σ_j^2 . Thus, the marginal pdf of $Y_{ij}(t)$, after integrating out δ_{ij} , is given by

$$f_{Y_{ij}(t)}(y_{ij}; \eta_j, \sigma_j, \lambda_j, t) = \sqrt{\frac{\lambda_j t^2}{2\pi y_{ij}^3 (\lambda_j \sigma_j^2 y_{ij} + 1)}} \times \exp \left[-\frac{\lambda_j (t - \eta_j y_{ij})^2}{2y_{ij} (\lambda_j \sigma_j^2 y_{ij} + 1)} \right]. \quad (2)$$

The derivation of Eq. (2) is given in Appendix A.

Next, the unconditional joint pdf of $\mathbf{Y}_i(t)$ is given by

$$f_{\mathbf{Y}_i(t)}(\mathbf{y}_i; \boldsymbol{\theta}, t) = (2\pi)^{-\frac{p}{2}} t^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{y_{ij}^3}} \right) |\mathbf{I} + \mathbf{V}_i \boldsymbol{\Sigma}|^{-1/2} \times \exp \left\{ -\frac{1}{2} (\boldsymbol{\eta} - \mathbf{v}_i)' [\mathbf{I} - (\mathbf{I} + \mathbf{V}_i \boldsymbol{\Sigma})^{-1}] \boldsymbol{\Sigma}^{-1} (\boldsymbol{\eta} - \mathbf{v}_i) \right\}, \quad (3)$$

where $\mathbf{v}_i = (t/y_{i1}, t/y_{i2}, \dots, t/y_{ip})'$ and

$$\mathbf{V}_i = \begin{pmatrix} \lambda_1 y_{i1} & 0 & 0 & 0 \\ 0 & \lambda_2 y_{i2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p y_{ip} \end{pmatrix}.$$

The derivation of Eq. (3) is given in Appendix B.

Suppose $\mathbf{Y}_i(t)$ is divided into two mutually exclusive subsets – $\mathbf{Y}_{iA}(t)$ and $\mathbf{Y}_{iB}(t)$. And we denote the indices associated with the two subsets by $A = \{j : Y_{ij}(t) \in \mathbf{Y}_{iA}(t)\}$ and $B = \{j : Y_{ij}(t) \in \mathbf{Y}_{iB}(t)\}$, respectively. Then, the conditional pdf of $\mathbf{Y}_{iA}(t) | \mathbf{Y}_{iB}(t)$ after integrating out $\boldsymbol{\delta}$ is given by

$$\begin{aligned} f_{\mathbf{Y}_{iA}(t) | \mathbf{Y}_{iB}(t)}(\mathbf{y}_{iA} | \mathbf{y}_{iB}; \boldsymbol{\theta}, t) &= \frac{f_{\mathbf{Y}_i(t)}(\mathbf{y}_i; \boldsymbol{\theta}, t)}{f_{\mathbf{Y}_{iB}(t)}(\mathbf{y}_{iB}; \boldsymbol{\theta}, t)} \\ &= \left(\prod_{j \in A} \sqrt{\frac{\lambda_j t^2}{2\pi y_{ij}^3}} \right) \frac{|\mathbf{I} + \mathbf{V}_i \boldsymbol{\Sigma}|^{-1/2}}{|\mathbf{I} + \mathbf{V}_{iB} \boldsymbol{\Sigma}_B|^{-1/2}} \\ &\quad \times \frac{\exp \left\{ -\frac{1}{2} (\boldsymbol{\eta} - \mathbf{v}_i)' [\mathbf{I} - (\mathbf{I} + \mathbf{V}_i \boldsymbol{\Sigma})^{-1}] \boldsymbol{\Sigma}^{-1} (\boldsymbol{\eta} - \mathbf{v}_i) \right\}}{\exp \left\{ -\frac{1}{2} (\boldsymbol{\eta}_B - \mathbf{v}_{iB})' [\mathbf{I} - (\mathbf{I} + \mathbf{V}_{iB} \boldsymbol{\Sigma}_B)^{-1}] \boldsymbol{\Sigma}_B^{-1} (\boldsymbol{\eta}_B - \mathbf{v}_{iB}) \right\}}, \quad (4) \end{aligned}$$

where $\boldsymbol{\eta}_B$, $\mathbf{v}_{iB}$, $\boldsymbol{\Sigma}_B$, and $\mathbf{V}_{iB}$ are the partitioned vector and matrices for relevant elements in the subset B from $\boldsymbol{\eta}$, $\mathbf{v}_i$, $\boldsymbol{\Sigma}$, and $\mathbf{V}_i$, respectively.

With the observation $\mathbf{y}_i(t) = (y_{i1}, y_{i2}, \dots, y_{ip})'$, a vector of degradation measurements at time t , the conditional distribution of $\boldsymbol{\delta}_i$ given $\mathbf{y}_i(t)$ can be computed as

$$\begin{aligned} f(\boldsymbol{\delta}_i | \mathbf{y}_i(t)) &\propto f(\mathbf{y}_i(t) | \boldsymbol{\delta}_i) f(\boldsymbol{\delta}_i) \\ &\propto \exp \left[-\frac{1}{2} \boldsymbol{\delta}_i' (\boldsymbol{\Sigma}^{-1} + \mathbf{V}_i) \boldsymbol{\delta}_i + \boldsymbol{\delta}_i' (\boldsymbol{\Sigma}^{-1} \boldsymbol{\eta} + \mathbf{V}_i \mathbf{v}_i) \right]. \quad (5) \end{aligned}$$

The derivation of Eq. (5) is given in Appendix C. This result implies that $\boldsymbol{\delta}_i | \mathbf{y}_i(t)$ must be subject to a MVN distribution with mean vector $\tilde{\boldsymbol{\eta}}_i \equiv (\boldsymbol{\Sigma}^{-1} + \mathbf{V}_i)^{-1} (\boldsymbol{\Sigma}^{-1} \boldsymbol{\eta} + \mathbf{V}_i \mathbf{v}_i)$ and variance-covariance matrix $\tilde{\boldsymbol{\Sigma}}_i \equiv (\boldsymbol{\Sigma}^{-1} + \mathbf{V}_i)^{-1}$. In fact, one could easily verify that the updated distribution of $\boldsymbol{\delta}_i$ is only dependent on the last measurement $\mathbf{y}_i(t)$ regardless of any intermediate measurements. This indicates that the Markovian property of the process is preserved.

Finally, by making use of the current information of observed degradation measurements $\mathbf{y}_i(t)$ at time t , the conditional distribution of the future degradation increments $\tilde{\mathbf{y}}_i(s) = (\tilde{y}_{i1}, \tilde{y}_{i2}, \dots, \tilde{y}_{ip})'$ after an interval of time s is given by (note that $\tilde{\mathbf{y}}_i(s) = \mathbf{y}_i(t+s) -$

$\mathbf{y}_i(t)$

$$f(\tilde{\mathbf{y}}_i(s)|\mathbf{y}_i(t)) = (2\pi)^{-\frac{p}{2}} s^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{\tilde{y}_{ij}^3}} \right) |\mathbf{I} + \tilde{\mathbf{V}}_i \tilde{\Sigma}_i|^{-1/2} \\ \times \exp \left\{ -\frac{1}{2} (\tilde{\boldsymbol{\eta}}_i - \tilde{\mathbf{v}}_i)' [\mathbf{I} - (\mathbf{I} + \tilde{\mathbf{V}}_i \tilde{\Sigma}_i)^{-1}] \tilde{\Sigma}_i^{-1} (\tilde{\boldsymbol{\eta}}_i - \tilde{\mathbf{v}}_i) \right\}, \quad (6)$$

where $\tilde{\mathbf{v}}_i = (s/\tilde{y}_{i1}, s/\tilde{y}_{i2}, \dots, s/\tilde{y}_{ip})'$ and

$$\tilde{\mathbf{V}}_i = \begin{pmatrix} \lambda_1 \tilde{y}_{i1} & 0 & 0 & 0 \\ 0 & \lambda_2 \tilde{y}_{i2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \tilde{y}_{ip} \end{pmatrix}.$$

The derivation of Eq. (6) is given in Appendix D. In this paper, we use $y, \mathbf{y}, t, \mathbf{t}, \mathbf{v}$, and $\mathbf{V}$, in normal mathematical fonts, to represent notations associated with degradation measurements, while $\tilde{y}, \tilde{\mathbf{y}}, \tilde{t}, \tilde{\mathbf{t}}, \tilde{\mathbf{v}}$, and $\tilde{\mathbf{V}}$, in serif mathematical fonts with the tilde character, for notations associated with degradation increments. We distinguish these two concepts mainly for the convenience of computation because degradation increments are the input for model inference as illustrated in Section 3. In fact, degradation measurements are interchangeable with degradation increments since one can always treat a measurement as an increment that is accumulated from the commencement of degradation to the present assuming the initial measurement is zero.

2.3. Lifetime distribution

In engineering practice, product lifetime is of much interest. For an individual degradation process j , the failure time T_{D_j} is defined as the first passage time of the process reaching a pre-defined failure threshold D_j . Then, the failure time is denoted by

$$T_{D_j} = \inf\{t : Y_j(t) \geq D_j\}.$$

Here, we drop the subscript i to indicate that it is a population characteristic; that is, $Y_j(t)$ denotes the population degradation process for the j -th PC. Similar changes apply to other relevant notations in this section.

Since the IG process produces monotone degradation paths, the distribution of the failure time is $P(T_{D_j} < t) = P(Y_j(t) \geq D_j)$. Then, given δ_j , the cumulative distribution function (cdf) of the failure time for an individual degradation process can be derived as

$$F_{T_{D_j}}(t|\delta_j; \lambda_j, D_j) = \Phi \left[-\sqrt{\frac{\lambda_j}{D_j}} (D_j \delta_j - t) \right] \\ - \exp(2\lambda_j \delta_j t) \Phi \left[-\sqrt{\frac{\lambda_j}{D_j}} (D_j \delta_j + t) \right], \quad (7)$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution.

With consideration of the random effects that have been introduced to δ_j , the marginal cdf of the failure time for an individual degradation process is derived as

$$F_{T_{D_j}}(t; \eta_j, \sigma_j, \lambda_j, D_j) = \Phi \left(\sqrt{\frac{\lambda_j}{D_j}} \frac{t - \eta_j D_j}{\sqrt{1 + \lambda_j \sigma_j^2 D_j}} \right) \\ - \exp[2\lambda_j t(\eta_j + \lambda_j \sigma_j^2 t)] \\ \times \Phi \left(-\sqrt{\frac{\lambda_j}{D_j}} \frac{t + \eta_j D_j + 2\lambda_j \sigma_j^2 D_j t}{\sqrt{1 + \lambda_j \sigma_j^2 D_j}} \right). \quad (8)$$

The derivation of Eq. (8) is given in Appendix E.

As a system has multiple competing risks (i.e. PCs) and each of them degrades over time, the system is considered to be failed

when any of these PCs passes its threshold, so the system failure time is defined by

$$T_D = \inf\{t : Y_1(t) \geq D_1 \text{ or } \dots \text{ or } Y_p(t) \geq D_p\},$$

where $\mathbf{D} = (D_1, D_2, \dots, D_p)'$ is a vector storing all PC failure thresholds.

Given δ , the conditional cdf of system failure time is

$$F_{T_D}(t|\delta; \lambda, \mathbf{D}) = 1 - P(Y_1(t) < D_1, Y_2(t) < D_2, \dots, Y_p(t) < D_p) \\ = 1 - \prod_{j=1}^p P(Y_j(t) < D_j) \\ = 1 - \prod_{j=1}^p \left(1 - F_{T_{D_j}}(t|\delta_j; \lambda_j, D_j) \right). \quad (9)$$

Again, considering the random effects, the unconditional cdf of the system failure time is provided by

$$F_{T_D}(t; \boldsymbol{\theta}, \mathbf{D}) = 1 - \int_0^{D_1} \dots \int_0^{D_p} f_{\mathbf{Y}(t)}(\mathbf{y}; \boldsymbol{\theta}, t) dy_1 \dots dy_p \\ = 1 - \int_0^{D_1} \dots \int_0^{D_p} (2\pi)^{-\frac{p}{2}} t^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{y_j^3}} \right) |\mathbf{I} + \mathbf{V}\Sigma|^{-1/2} \\ \times \exp \left\{ -\frac{1}{2} (\boldsymbol{\eta} - \mathbf{v})' [\mathbf{I} - (\mathbf{I} + \mathbf{V}\Sigma)^{-1}] \Sigma^{-1} (\boldsymbol{\eta} - \mathbf{v}) \right\} dy_1 \dots dy_p. \quad (10)$$

The multivariate integral involved in Eq. (10) can be evaluated numerically.

3. Model estimation and validation

This section aims at providing the statistical inference procedure for model parameters estimation and the tools for model validation.

3.1. Point estimation

For a total of $p \times m_i$ observations generated from the i -th test unit, we denote the dataset of the corresponding degradation increments by

$$\tilde{\mathbf{Y}}_i = (\tilde{\mathbf{y}}_{i1}, \tilde{\mathbf{y}}_{i2}, \dots, \tilde{\mathbf{y}}_{im_i}) \\ = \begin{pmatrix} \tilde{y}_{i11} & \tilde{y}_{i12} & \dots & \tilde{y}_{i1m_i} \\ \tilde{y}_{i21} & \tilde{y}_{i22} & \dots & \tilde{y}_{i2m_i} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{ip1} & \tilde{y}_{ip2} & \dots & \tilde{y}_{ipm_i} \end{pmatrix}_{p \times m_i},$$

where the degradation measurement and increment of process j on unit i at time point k are y_{ijk} and $\tilde{y}_{ijk} = y_{ijk} - y_{ij,k-1}$, $\forall i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$, $k = 1, 2, \dots, m_i$, respectively. Without loss of generality, the initial measurement y_{ij0} is assumed to be known. Then, the whole dataset of degradation increments across all test units can be represented by $\tilde{\mathbf{Y}} = (\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2, \dots, \tilde{\mathbf{Y}}_n)$.

Meanwhile, we denote the vector of measuring time intervals on unit i by

$$\tilde{\mathbf{t}}_i = (\tilde{t}_{i1}, \tilde{t}_{i2}, \dots, \tilde{t}_{im_i})'.$$

Here, the interval time between two consecutive measurements on unit i is denoted as $\tilde{t}_{ik}$, where $\tilde{t}_{ik} = t_{ik} - t_{i,k-1}$, $\forall i = 1, 2, \dots, n$, $k = 1, 2, \dots, m_i$. t_{ik} is the elapsed time when the k -th measurement on unit i is taken. Obviously, we suppose $t_{i0} = 0$, $\forall i = 1, 2, \dots, n$. Note that the number of measurements may vary unit by unit. Without loss of generality, we assume that measurements are conducted over all individual degradation processes at a time. Thus, the whole dataset of measurement time intervals across all test units can be

represented by $\tilde{\mathbf{t}} = (\tilde{\mathbf{t}}'_1, \tilde{\mathbf{t}}'_2, \dots, \tilde{\mathbf{t}}'_n)'$. Furthermore, we let $\mathbb{D} = \{\tilde{\mathbf{Y}}, \tilde{\mathbf{t}}\}$ denote the total observed data.

Hence, based on Eq. (3), the total log-likelihood function is provided by

$$\ell(\boldsymbol{\theta}|\mathbb{D}) = \sum_{i=1}^n \sum_{k=1}^{m_i} \left\{ -\frac{p}{2} \ln(2\pi) + p \ln(\tilde{t}_{ik}) + \frac{1}{2} \sum_{j=1}^p \ln \frac{\lambda_j}{\tilde{y}_{ijk}^3} - \frac{1}{2} \ln |\mathbf{I} + \tilde{\mathbf{v}}_{ik} \boldsymbol{\Sigma}| - \frac{1}{2} (\boldsymbol{\eta} - \tilde{\mathbf{v}}_{ik})' [\mathbf{I} - (\mathbf{I} + \tilde{\mathbf{v}}_{ik} \boldsymbol{\Sigma})^{-1}] \boldsymbol{\Sigma}^{-1} (\boldsymbol{\eta} - \tilde{\mathbf{v}}_{ik}) \right\}, \quad (11)$$

where $\tilde{\mathbf{v}}_{ik} = (\tilde{t}_{ik}/\tilde{y}_{i1k}, \tilde{t}_{ik}/\tilde{y}_{i2k}, \dots, \tilde{t}_{ik}/\tilde{y}_{ipk})'$ and

$$\tilde{\mathbf{v}}_{ik} = \begin{pmatrix} \lambda_1 \tilde{y}_{i1k} & 0 & 0 & 0 \\ 0 & \lambda_2 \tilde{y}_{i2k} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \tilde{y}_{ipk} \end{pmatrix}.$$

Due to the high dimensionality of $\boldsymbol{\theta}$ and the complexity of Eq. (11), it is difficult to directly apply maximum likelihood estimation (MLE). Therefore, we seek for an expectation maximization (EM) algorithm to infer all the unknown parameters.

To carry out the EM algorithm, we treat the independently and identically distributed (i.i.d.) random effects $\boldsymbol{\delta} = (\delta'_1, \delta'_2, \dots, \delta'_n)'$ as missing data or latent variables. In such case, the complete data are assumed to be $\{\mathbb{D}, \boldsymbol{\delta}\}$ and the total log-likelihood function is given by

$$\ell(\boldsymbol{\theta}|\mathbb{D}, \boldsymbol{\delta}) = \ell(\boldsymbol{\lambda}, \boldsymbol{\delta}|\mathbb{D}) + \ell(\boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\rho}|\boldsymbol{\delta}),$$

where

$$\ell(\boldsymbol{\lambda}, \boldsymbol{\delta}|\mathbb{D}) = \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left[\frac{1}{2} \ln \left(\frac{\lambda_j \tilde{t}_{ik}^2}{2\pi \tilde{y}_{ijk}^3} \right) - \frac{1}{2} \frac{(\tilde{t}_{ik} - \tilde{y}_{ijk} \delta_{ij})^2}{\tilde{y}_{ijk}/\lambda_j} \right] \text{ and}$$

$$\ell(\boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\rho}|\boldsymbol{\delta}) = \sum_{i=1}^n \left[-\frac{p}{2} \ln(2\pi) - \frac{1}{2} \ln |\boldsymbol{\Sigma}| - \frac{1}{2} (\boldsymbol{\delta}_i - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\delta}_i - \boldsymbol{\eta}) \right].$$

Then, if denoting $\boldsymbol{\theta}^{(s)}$ as the EM estimates of $\boldsymbol{\theta}$ at the s -th iteration, the EM algorithm is an iterative process consisting of two steps, which are described by Algorithm 1 in brief. We denote $\boldsymbol{\theta}^{(s)}$ by $\hat{\boldsymbol{\theta}}$ when the iteration stops (i.e. the algorithm converges). Technical details about the EM algorithm are provided in Appendix F.1. To expedite the convergence of the algorithm, it is important to start with a good guess of starting parameters. The initial values can be obtained from a summary of data, while treating each degradation path as an independent realization of a simple IG process. The instructions for finding starting parameter values are given in Appendix F.2. Additionally, if a time scale transformation is needed, inferring the unknown parameter in the transformation function $\Lambda(\cdot)$ has to be carried out too. Denote the transformed time interval by $\tau_{ijk} = \Lambda_j(t_{ik}) - \Lambda_j(t_{i,k-1})$, $\forall i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$, $k = 1, 2, \dots, m_i$, with an unknown parameter γ_j . In this setting, we assume the time scale transformation is a fixed effect so that for a certain PC j , γ_j is identical across all test units. The adaptation of the EM algorithm with this consideration is given in Appendix F.3. Note that when the time scale transformation is necessary, the set of unknown parameters becomes $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}', \boldsymbol{\rho}', \boldsymbol{\lambda}', \boldsymbol{\gamma}')'$, where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_p)'$.

3.2. Interval estimation

In addition to the point estimation, the interval estimation of some population properties, such as the failure time probability, or a general function of model parameters, $g(\boldsymbol{\theta})$, are often of interest as well. To construct such confidence intervals (C.I.s), a common routine is to make use of asymptotic theories. However, for

Algorithm 1: The EM Algorithm (in brief) for Parameters Estimation.

Data: The dataset $\mathbb{D} = \{\tilde{\mathbf{Y}}, \tilde{\mathbf{t}}\}$ in terms of degradation increments and measurement time intervals.

Input: The current EM estimates $\boldsymbol{\theta}^{(s)}$ (i.e. $\boldsymbol{\lambda}^{(s)}$, $\boldsymbol{\eta}^{(s)}$, and $\boldsymbol{\Sigma}^{(s)}$).

Output: The updated EM estimates $\boldsymbol{\theta}^{(s+1)}$ (i.e. $\boldsymbol{\lambda}^{(s+1)}$, $\boldsymbol{\eta}^{(s+1)}$, and $\boldsymbol{\Sigma}^{(s+1)}$).

Expectation step (E-step): Define a function

$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)}) = E_{\boldsymbol{\delta}|\mathbb{D}, \boldsymbol{\theta}^{(s)}}[\ell(\boldsymbol{\theta}|\mathbb{D}, \boldsymbol{\delta})]$, which is the expected value of $\ell(\boldsymbol{\theta}|\mathbb{D}, \boldsymbol{\delta})$ with respect to the current conditional distribution of $\boldsymbol{\delta}$ given $\mathbb{D}$ and $\boldsymbol{\theta}^{(s)}$. In summary, the function is given by

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)}) = \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left\{ \frac{1}{2} \ln \left(\frac{\lambda_j \tilde{t}_{ik}^2}{2\pi \tilde{y}_{ijk}^3} \right) - \frac{1}{2} \frac{\left\{ \tilde{t}_{ik}^2 - 2\tilde{t}_{ik} \tilde{y}_{ijk} \tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{j(i)}^{(s)} \right)^2 \right] \right\}}{\tilde{y}_{ijk}/\lambda_j} \right\}$$

$$+ \sum_{i=1}^n \left\{ -\frac{p}{2} \ln(2\pi) - \frac{1}{2} \ln |\boldsymbol{\Sigma}| - \frac{1}{2} \left[\text{tr}(\boldsymbol{\Sigma}^{-1} \tilde{\boldsymbol{\Sigma}}_i^{(s)}) + (\tilde{\boldsymbol{\eta}}_i^{(s)} - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\tilde{\boldsymbol{\eta}}_i^{(s)} - \boldsymbol{\eta}) \right] \right\}.$$

Maximization step (M-step): Find the parameters that maximize the quantity $\boldsymbol{\theta}^{(s+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)})$. In summary, the result is given by

$$\lambda_j^{(s+1)} = \frac{\sum_{i=1}^n m_i}{\sum_{i=1}^n \sum_{k=1}^{m_i} \frac{\left\{ \tilde{t}_{ik}^2 - 2\tilde{t}_{ik} \tilde{y}_{ijk} \tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{j(i)}^{(s)} \right)^2 \right] \right\}}{\tilde{y}_{ijk}}}},$$

$$\boldsymbol{\eta}^{(s+1)} = \frac{1}{n} \sum_{i=1}^n \tilde{\boldsymbol{\eta}}_i^{(s)},$$

$$\boldsymbol{\Sigma}^{(s+1)} = \frac{\sum_{i=1}^n \left[\tilde{\boldsymbol{\Sigma}}_i^{(s)} + (\tilde{\boldsymbol{\eta}}_i^{(s)} - \boldsymbol{\eta}^{(s+1)}) (\tilde{\boldsymbol{\eta}}_i^{(s)} - \boldsymbol{\eta}^{(s+1)})' \right]}{n}.$$

the proposed model, it is difficult to evaluate its Fisher information matrix. Instead, we adopt the bias-corrected percentile (BCp) bootstrap method (Efron & Tibshirani, 1994; Meeker & Escobar, 2014). The details about this method is given in Appendix G.

3.3. Model validation

In order to verify the goodness of fit (GOF) of Model M_0 , we extend the model validation techniques for univariate IG processes to the multivariate case. According to Wang & Xu (2010) and Ye & Chen (2014), if $X \sim IG(a, b)$, then $b(X - a)^2 / (a^2 X) \sim \chi_1^2$. Thus, for an individual PC in Model M_0 , the estimated quantity $\hat{\lambda}_j (\tilde{\eta}_{ij} \tilde{y}_{ijk} - \tilde{t}_{ik})^2 / \tilde{y}_{ijk}$ are approximately i.i.d. χ_1^2 , $\forall i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$, $k = 1, 2, \dots, m_i$. Here, $\tilde{\eta}_{ij}$ is the estimated sample inverse drift parameter for process j on unit i . The resulting χ_1^2 quantile-quantile (Q-Q) plot can be used to visualize the GOF of each PC's IG process model. In addition, to test the independence among all random effects (i.e. $H_0: \sigma_{j'j} = 0$, $\forall 1 \leq j' < j \leq p$), define $U' = -\left[v - \frac{1}{6}(2p+5) \right] \ln U$, where U is the determinant of the estimated correlation matrix $\hat{\mathbf{R}}$ (calculated from $\hat{\boldsymbol{\Sigma}}$) with degrees of freedom $v = n(m-1)$. According to Rencher & Christensen (2012), we can use U' as a test statistic for testing the null hypothesis of independent PCs. H_0 is rejected if the calculated value u' is greater than an upper Chi-square percentile value $\chi_{\alpha, d}^2$, where the degrees of freedom are as $d = \frac{1}{2}p(p-1)$. Lastly, to compare the proposed model with other models such as a model of multiple univariate IG processes without random effects, the Akaike information criterion (AIC), $\text{AIC} = 2|\boldsymbol{\theta}| - 2\ell$, is

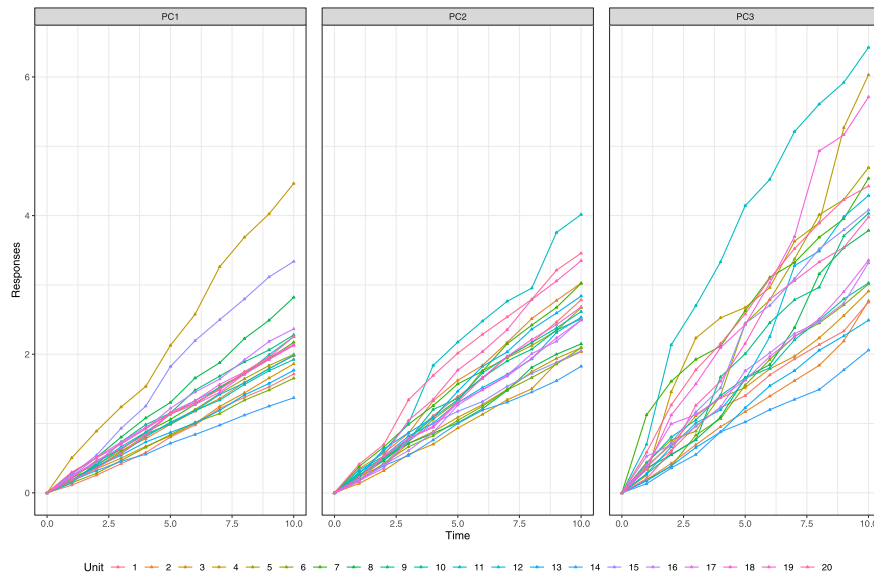


Fig. 2. Simulated Degradation Paths with $n = 20$ and $m = 10$.

adopted, where $|\theta|$ is the total number of model parameters and ℓ is the corresponding log-likelihood.

4. Simulation studies

4.1. Performance of the inference method

In this section, we carry out a Monte Carlo simulation study to evaluate the performance of the EM algorithm proposed in Section 3.1. The model is a 3-dimensional degradation process model given by

$$\begin{cases} Y_{i1}(t) \sim IG(t/\delta_{i1}, 6t^2) \\ Y_{i2}(t) \sim IG(t/\delta_{i2}, 4t^2) \\ Y_{i3}(t) \sim IG(t/\delta_{i3}, 2t^2) \\ \delta_i \sim MVN\left(\begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 & 0.2 & 0.8 \\ 0.2 & 1 & 0.5 \\ 0.8 & 0.5 & 1 \end{pmatrix}\right) \end{cases},$$

where we assume the measurement is taken every one unit of time and it is the same to each PC on all units. Meanwhile, we suppose $n = 20, 40$, and 60 along with $m = 10, 30$, and 50 . It results in a total of nine combinations of sample size. Fig. 2 demonstrates a sample of simulated degradation paths that is generated from the model above with $n = 20$ and $m = 10$. To carry out statistical inference, the starting parameters are obtained according to the instruction provided by Appendix F.2 and the error tolerance ϵ used in the EM algorithm is 10^{-5} . It turns out the computation time for the inference to meet the convergence criterion is within one quarter of a minute under any sample size combination on a personal computer with an Intel Core i7 2.9GHz CPU. We think it is a satisfactory performance for practical use. In total, for each case, we generate 1000 replications of data from its simulation model and fit Model M_0 to the data. The root mean squared errors (RMSEs) of model parameter estimators are given in Table 1.

According to these results, in general, the RMSE decreases with the increase of sample size. It is found that among all the parameters, ρ_{13} with much strong correlation is better estimated than the others. This may be due to the phenomenon that stronger correlation results in more “similar” data that could help improve the estimation accuracy. It is also noticed that the improvement of estimation accuracy for η and Σ is much sensitive to n . This is not surprising because, according to the M-step in the EM algorithm, the

estimation accuracy of these two estimators heavily relies on the number of test units. Therefore, the inference method proposed in Section 3.1 performs more effectively when sample size is large.

4.2. Effect of model misspecification

In this section, we conduct another simulation study to assess the effect of model misspecification. In degradation analysis, two other IG process-based models, listed as models M_1 and M_2 below, can be viewed as alternatives to M_0 . Model M_1 assumes the randomness completely originates from the stochastic nature of the degradation process and each PC is governed by an individual IG process. Model M_2 is a collection of classical IG process models with random effects as proposed by Ye & Chen (2014). Both models do not involve any dependence structure among PCs.

$$M_1 : Y_{ij}(t) \sim IG(t/\delta_{ij}, \lambda_j t^2),$$

$$M_2 : \begin{cases} Y_{ij}(t) \sim IG(t/\delta_{ij}, \lambda_j t^2) \\ \delta_{ij} \sim N(\eta_j, \sigma_j^2) \end{cases}.$$

For simplicity, the simulation model we make use of is either a 2-dimensional or 3-dimensional degradation process (i.e. $p = 2$ or 3) that is given by

$$\begin{cases} Y_{ij}(t) \sim IG(t/\delta_{ij}, 6t^2) \\ \delta_i \sim MVN\left(\begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right) \text{ or } MVN\left(\begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}\right) \end{cases}$$

Three different levels of correlation – $\rho = 0.2, 0.5, 0.8$ – are considered. Thus, we produce six different scenarios and for each of them we fit four models – M_0 , M_1 , and M_2 . The sample size of the simulation is chosen to be $n = 60$ and $m = 50$. A simulated degradation observation is generated every one unit of time. We replicate the simulation 1,000 times for each scenario. With the simulated data in each replication, we perform model inferences and calculate the model-based reliability prediction at the time point where the reliability is equal to 0.5 implied by the true model. The failure threshold is assumed to be 1.5 for each marginal process. Eqs. (7), (8), and (10) are involved in the calculation of reliability, where the “hcubature” function in R package – cubature is utilized to perform multi-dimensional integration of integrands. Note that the inference model for Model M_1 is a simple MLE process. For Model M_2 , similar to the work by Ye & Chen (2014), it is

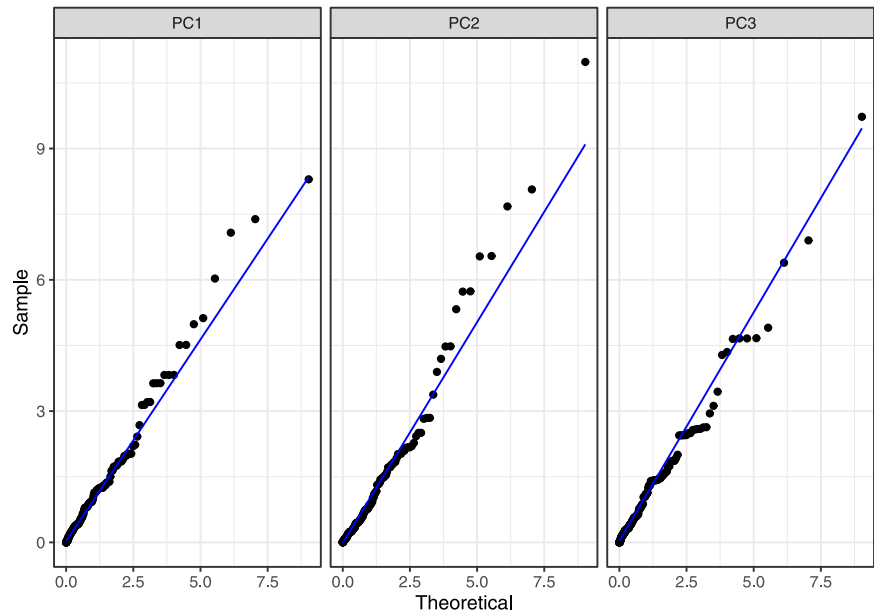


Fig. 3. Q-Q Plots for the Coating Data.

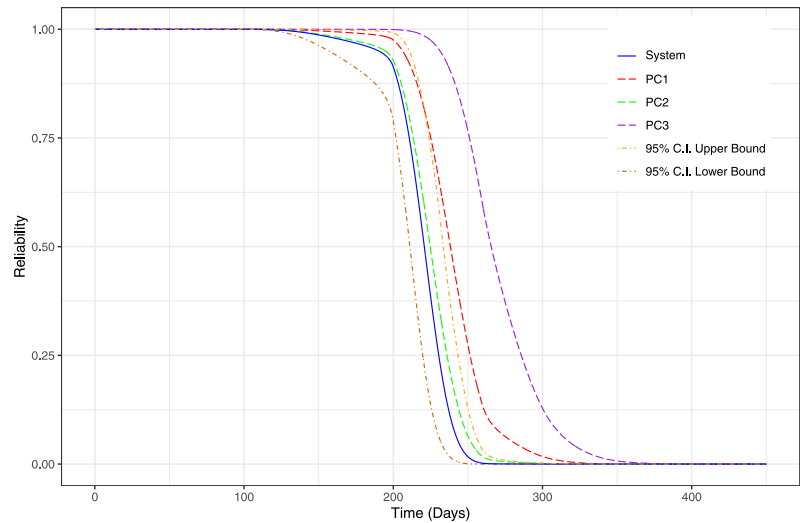


Fig. 4. Reliability Curves for the Coating Data.

Table 1
RMSE ($\times 10$) of the EM Estimate for Each Parameter.

n	m	$\lambda_1 = 6$	$\lambda_2 = 4$	$\lambda_3 = 2$	$\eta_1 = 5$	$\eta_2 = 4$	$\eta_3 = 3$	$\sigma_1 = 1$	$\sigma_2 = 1$	$\sigma_3 = 1$	$\rho_{12} = 0.2$	$\rho_{13} = 0.8$	$\rho_{23} = 0.5$
20	10	6.501	4.432	2.185	2.337	2.319	2.378	1.742	1.688	1.901	2.482	1.237	2.048
	30	3.663	2.286	1.138	2.278	2.227	2.284	1.653	1.587	1.713	2.308	1.033	1.911
	50	2.810	1.851	0.943	2.276	2.211	2.256	1.628	1.582	1.685	2.240	0.989	1.849
40	10	4.521	3.039	1.501	1.630	1.597	1.637	1.215	1.251	1.256	1.667	0.806	1.419
	30	2.388	1.696	0.824	1.603	1.552	1.559	1.140	1.165	1.147	1.584	0.648	1.244
	50	1.944	1.263	0.644	1.591	1.537	1.529	1.138	1.159	1.122	1.549	0.625	1.209
60	10	3.663	2.454	1.250	1.297	1.353	1.350	0.966	1.012	1.044	1.311	0.639	1.147
	30	2.004	1.396	0.683	1.254	1.289	1.278	0.916	0.969	0.960	1.222	0.526	1.040
	50	1.607	1.078	0.528	1.249	1.285	1.256	0.893	0.951	0.941	1.201	0.498	1.013

just a little variation of the aforementioned EM algorithm by treating the marginal degradation process independently. The adaptation of the EM algorithm to fit Model M_2 is given in [Appendix F.4](#). [Table 2](#) reports the absolute value of reliability gap on average from the 1,000 replications of each simulation scenario. The metric is a measure of the accuracy of the estimated reliability value

to its true value. As indicated by the result, compared with Model M_0 , the reliability estimate by the two alternatives is poor. The bad performance comes from two sources: 1) the biases in modeling the randomness existing in the data and 2) the overlooked PC dependence structure that will lead to a skewed reliability calculation. Thus, correctly identifying random effects and any possible

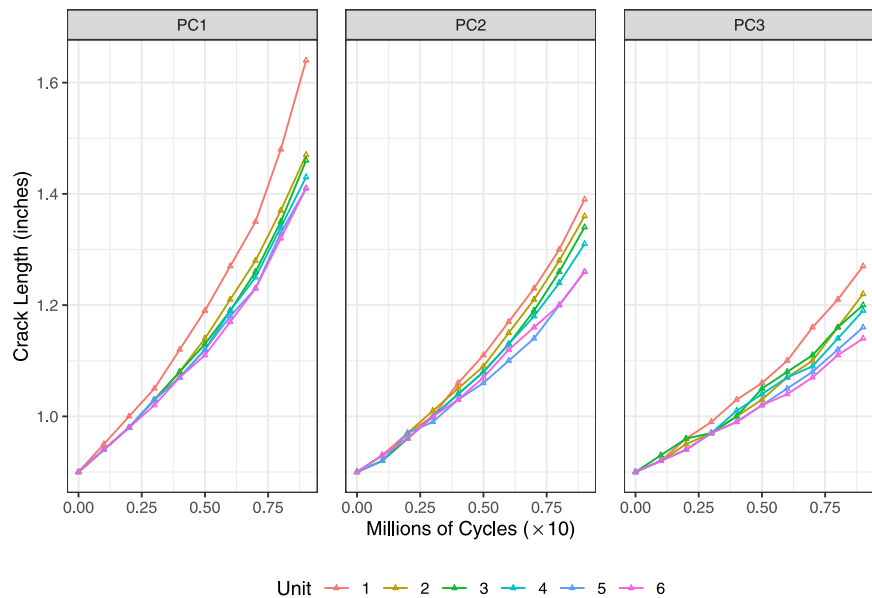


Fig. 5. Fatigue Crack-size Growth.

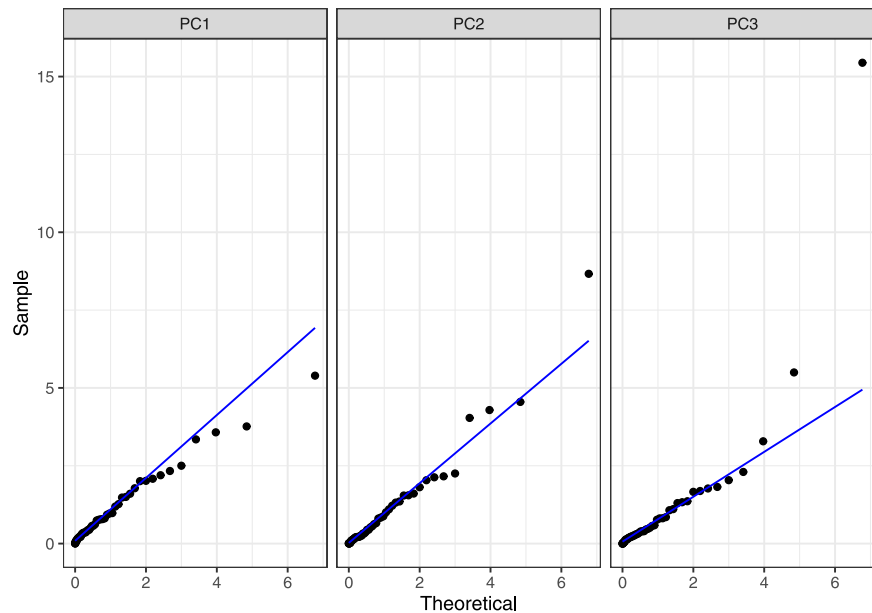


Fig. 6. Q-Q Plots for the Fatigue Crack-size Data.

Table 2
Absolute Value of Reliability Gap (it is $|\hat{R} - 0.5| \times 10^3$) on Average for Various Fitted Models.

p	ρ	M_0	M_1	M_2
2	0.2	39.169	127.476	42.118
	0.5	39.956	103.918	66.222
	0.8	40.776	136.935	112.703
3	0.2	40.790	249.763	45.334
	0.5	42.533	159.492	93.027
	0.8	45.192	132.369	175.305

underlying PC dependence is critical to assessing system reliability when a multivariate degradation process is present.

5. Two illustrative examples

In this section, we provide two illustrative examples to demonstrate the implementation of the proposed multivariate analysis method.

5.1. Coating data

First, we illustrate how the proposed methodology can be applied for the motivating example – the coating data. To accommodate the monotone increasing assumption implied by the IG process negative increments (i.e. $-\tilde{y}_{ijk}$) are computed and treated as the raw data for analysis. After fitting the proposed model to the

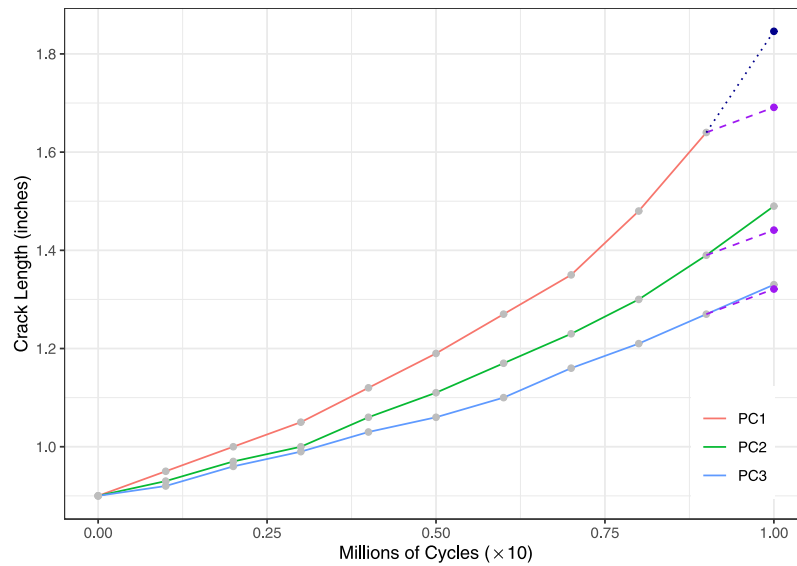


Fig. 7. One-period Ahead Degradation Predictions.

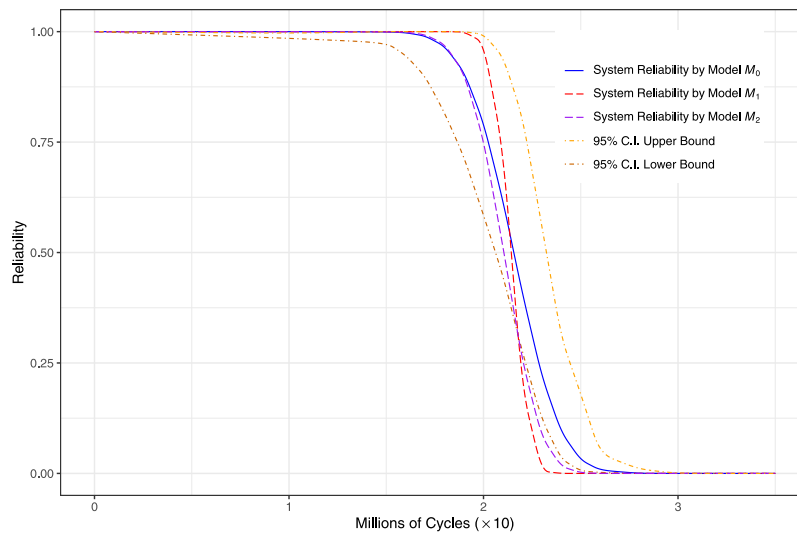


Fig. 8. Reliability Curves for the Fatigue Crack-size Data.

data, the result of parameters estimation is obtained as follows:

$$\hat{\lambda} = (2.09998, 0.39288, 1.61180)',$$

$$\hat{\gamma} = (0.97996, 1.15386, 0.96347)',$$

$$\hat{\eta} = (0.35590, 0.68947, 0.54000)',$$

$$\text{and } \hat{\Sigma} = \begin{pmatrix} 1.59788 & 1.93898 & 3.07767 \\ 1.93898 & 13.19950 & 13.32560 \\ 3.07767 & 13.32560 & 23.86160 \end{pmatrix} \times 10^{-6},$$

where $\hat{\gamma}$ is the estimate of the parameter γ in the time scale transformation function $\Lambda(t) = t^\gamma$. Based on the result, the estimated correlations are $\hat{\rho}_{12} = 0.42$, $\hat{\rho}_{13} = 0.50$, and $\hat{\rho}_{23} = 0.76$. The estimated correlations represent the dependent relationship existing in the innate degradation mechanisms among the multiple PCs. The convergence of the EM algorithm is indicated by plots of parameter estimates and log-likelihood over iteration, which are presented by Fig. 9a in Appendix H. The appendix also presents the interval estimation of the model parameters in Table 4.

To validate the fitted model, we proceed with the χ^2 Q-Q plot and the test of independence as discussed in Section 3.3. From the Q-Q plots in Fig. 3, one can see that most data points fall

close to a straight line suggesting that the proposed model provides a roughly good fit to this dataset. Moreover, the independence test indicates that $u' = 204.4376$, which is greater than the critical value $\chi^2_{0.01,3} = 11.34487$. This result suggests that the correlation among the degradation processes is significant.

Lastly, we carry out the reliability analysis by evaluating Eq. (10) over a certain time period. For illustration, the failure thresholds for the three wavelengths are identified at -0.60 , -0.75 , and -0.40 , respectively. Fig. 4 indicates both system reliability and marginal reliability of the three PCs. Meanwhile, the pointwise 95% C.I. for the system reliability through the BCp bootstrap method is also shown.

5.2. Fatigue crack-size data

To facilitate more general applications, in this example, we utilize a subset of the fatigue crack-size data, which was initially given in Appendix Table C.14 of Meeker & Escobar's book (2014). This dataset describes an alloy's crack growth over time. Similar to the work of Wang et al. (2015), this dataset is split to three parts so that a 3-dimensional degradation process is created with a sample

size of six test units (i.e. $n = 6$ and $p = 3$). The resulting degradation paths are depicted in Fig. 5, in which nine observations (i.e. $m = 9$) for each PC were taken on an individual test unit since the start of the experiment. The corresponding data table is provided in Appendix H.

After fitting the proposed model to this dataset, the result of parameters estimation is obtained as follows:

$$\begin{aligned}\hat{\lambda} &= (141.47632, 118.08734, 43.74568)', \\ \hat{\gamma} &= (1.32673, 1.32303, 1.24242)', \\ \hat{\eta} &= (1.54561, 2.09412, 3.00609)', \\ \text{and } \hat{\Sigma} &= \begin{pmatrix} 0.02859 & 0.03665 & 0.06335 \\ 0.03665 & 0.04712 & 0.08135 \\ 0.06335 & 0.08135 & 0.14072 \end{pmatrix},\end{aligned}$$

where $\hat{\gamma}$ is the estimate of the parameter γ in the time scale transformation function $\Lambda(t) = t^\gamma$. The convergence of the EM algorithm is indicated by plots of parameters estimation and log-likelihood over iteration, which are presented by Fig. 9b in Appendix H. The appendix also presents the interval estimation of the model parameters in Table 5.

By examining the Q-Q plots in Fig. 6, one can see that most data points fall close to a straight line except a few ones locate a little far away. This may result from the remaining uncertainty in the mean estimator. Nevertheless, the proposed model provides a reasonably good fit to this dataset. Moreover, the independence test indicates that $u' = 567.45880$, which is much greater than the critical value $\chi_{0.01,3}^2 = 11.34487$. This result suggests that the correlation among the degradation processes is significant. In addition to Model M_0 , we also fit models M_1 and M_2 to this dataset. The estimated parameters of the two alternatives are given in Appendix H. It turns out the AIC values for models M_0 , M_1 and M_2 are -1074.18600 , -976.25580 and -1002.40500 , respectively. By comparison, Model M_0 is chosen.

In engineering practice, making a prediction of future degradation values is often desired. To demonstrate this point, we carry out one-period ahead degradation predictions of the three PCs for Unit 1. Note that Eq. (6) provides the conditional distribution of future degradation increments given the current observation. Thus, based on the estimated model, it is easy to compute the expected degradation increments for the next inspection interval. Accordingly, Fig. 7 demonstrates the predicted degradation paths for the three PCs using purple dashed lines for the next 0.01 millions of cycles, where in contrast the real degradation paths of PC2 and PC3 indicated by the original data table are depicted in solid lines.

Moreover, when the degradation values of one or several PCs are absent due to certain reasons such as lack of measuring instruments, estimating the missing data is possible via borrowing the information of the rest PCs by leveraging the dependence structure. Our proposed model provides an exact solution to this task. Note that Eq. (4) is the conditional pdf of a subset of degradation measures given the rest observations. Based on this equation, for Unit 1, the expected value of PC1 at 0.1 millions of cycles can be calculated given the observed PC2 and PC3 values. The darkblue dot of PC1 in Fig. 7 shows the result, and one can see it is higher than the previous prediction. This is because both observed values of PC2 and PC3 are greater than the previous predictions and these three PCs have strong positive correlations.

Finally, we carry out the reliability analysis by plugging the estimated parameters into Eqs (7), (8) and (10). Through evaluating them over a certain time period, the estimated system reliability functions of the three models – M_0 , M_1 and M_2 – are depicted in Fig. 8. The failure thresholds of the three PCs are assumed to be 1.8, 1.4, and 1.3 inches, respectively. It can be seen that making independence assumption by either model, M_1 or M_2 (i.e. PCs are independent with or without random effects), could greatly alter the

predicted reliability function. Thus, care must be taken to examine the underlying dependence among PCs. In Fig. 8, the point-wise 95% C.I. through the BCp bootstrap method is also shown.

6. Concluding remarks

In this paper, a novel multivariate degradation model is proposed. The model is built upon multiple IG processes, where the element-wise inverse of the vector of the drift parameters is assumed to be subject to a MVN distribution. This approach facilitates the applicability and effectiveness of the model in accommodating the three common variability features in multivariate degradation observations; they are the randomness of degradation process, heterogeneity among units, and dependence between PCs. It also brings mathematically tractable properties to assist with both lifetime estimation and degradation prediction. Furthermore, we provide the EM algorithm to help estimate the unknown model parameters and the validation tools to check model adequacy. Through case studies, we have demonstrated the applications of our proposed framework including the prediction of future degradation values, the inference of missing degradation data by leveraging the dependence structure, and the reliability function evaluation.

It is further worth noting that unlike the conventional way of modeling the PC-wise dependency by assuming correlations between degradation observations directly (i.e. $\text{Cov}(Y_{ij}(t), Y_{ij'}(t))$), our approach incorporates the dependency in the drift parameters as random effects. These unobserved frailties provide specific physical interpretations – the dependency originates from the correlation existing in the innate degradation mechanisms among multiple PCs. This is because the drift parameters, which reflect the degradation rate, have close relation with degradation physics as shown by Ye & Chen (2014) and Peng (2015). As a comparison, the conventional modeling technique containing the models based on the copula theory (Fang & Pan, 2021; Palayangoda & Ng, 2021; Sun et al., 2020a) and the Wiener process (Hong, Ye, & Ling, 2018a; Sun et al., 2020b; Wang, Gaudoin, Doyen, Bérenguer, & Xie, 2020) is mainly a data-driven approach, as it suggests a comprehensive measure of dependency that may be contributed from various sources, including correlated measurement errors and shared environmental conditions, etc. And as we have seen in the previous sections, the tractability and flexibility in handling both heterogeneity and dependency further underpin the correlated random-effects modeling technique as an attractive approach in modeling multivariate degradation processes.

Beyond the current research effort, the following research directions are worth of a future study:

- In this paper, calculating multivariate integrals is involved in the reliability assessment. However, when high dimensionality is present, numerical evaluation will become slow. Thus, it is desired to develop more efficient ways to solve this problem.
- Evidently, incorporating explanatory variables is a necessary extension for the methodology to have broader applicability. Two recent articles by Lu et al. (2020) and Xu, Zhou, & Tang (2021) discussed a similar issue. Also, including measurement errors is another aspect to be considered. See Hao, Yang, & Berenguer (2019) for an example.
- Besides of the unit-to-unit variability and dependency between PCs, some other sources of random effects including the inspection effect and the block effect are of interest to be considered. Incorporating these features into our model is desired. Relevant research includes the works by Zhai & Ye (2018), Sun et al. (2020b), and Zhao, Chen, Gaudoin, & Doyen (2021). Applying the similar idea on accelerated life tests that consider multiple

sources of random effects has been recently explored in [Seo & Pan \(2020\)](#).

- It is found in this paper quantifying the interval estimation of a random-effects model's parameters is of much difficulty. There is a need to develop a method that could make the interval computation more efficiently. The relevant methodologies have been proposed by [Chen & Ye \(2018\)](#), [Hong, Ye, & Sari \(2018b\)](#), and [Wang, Wang, Hong, & Jiang \(2021\)](#).
- Furthermore, a more comprehensive investigation is desired to study the applications of the proposed model in certain areas, including the design of a degradation test ([Fang, Pan, & Stufken, 2021](#); [Shi, Xiang, Liao, Zhu, & Hong, 2020b](#)), on-line monitoring ([Hu, Sun, Ye, & Zhou, 2020](#)), and maintenance policy ([Keizer, Flapper, & Teunter, 2017](#); [Liu et al., 2021](#); [Mercier & Pham, 2012](#); [Wu & Castro, 2020](#)), etc.
- Finally, it is also of our interest to develop other types of multivariate degradation model, including those improved ones based on multi-dimensional Lévy processes with Lévy copulas and non-normal random effects, etc.

Acknowledgment

The work by Fang was partially supported by the Characteristic & Preponderant Discipline of Key Construction Universities in Zhejiang Province (Zhejiang Gongshang University - Statistics). The work by Pan was partially supported by the [National Science Foundation](#) (NSF) grant 1726445. The work by Wang was partially supported by the [National Natural Science Foundation of China](#) (NSFC) grant 71801171. We would like to thank the editor and three anonymous referees for their constructive comments and suggestions that have considerably improved the article.

Appendix A. Derivation of Eq. (2)

To derive the unconditional marginal pdf of $Y_{ij}(t)$, we utilize an existing conclusion (shown in [Lemma 1](#) below) provided by [Si & Zhou \(2014\)](#).

Lemma 1. If $X \sim N(\mu, \sigma^2)$, and $A, C, D, G \in \mathbb{R}$, $H \in \mathbb{R}^+$, then the following holds:

$$\begin{aligned} & E_X \left[(A - DX) \exp \left[-\frac{(C - GX)^2}{2H} \right] \right] \\ &= \sqrt{\frac{H}{G^2\sigma^2 + H}} \left(A - D \frac{GC\sigma^2 + H\mu}{G^2\sigma^2 + H} \right) \exp \left[-\frac{(C - G\mu)^2}{2(G^2\sigma^2 + H)} \right]. \end{aligned}$$

Recall that the conditional pdf of $Y_{ij}(t)$ given δ_{ij} implied by [Eq. \(1\)](#) can be reparameterized as

$$f_{Y_{ij}(t)}(y_{ij}|\delta_{ij}; \lambda_j, t) = \sqrt{\frac{\lambda_j t^2}{2\pi y_{ij}^3}} \exp \left[-\frac{(t - y_{ij}\delta_{ij})^2}{2y_{ij}/\lambda_j} \right].$$

Given [Lemma 1](#), if setting $A = 1$, $C = t$, $D = 0$, $G = y_{ij}$, and $H = y_{ij}/\lambda_j$, then the unconditional pdf of $Y_{ij}(t)$ is equivalent to $E_{\delta_{ij}}[f_{Y_{ij}(t)}(y_{ij}|\delta_{ij}; \lambda_j, t)]$, which is given by

$$f_{Y_{ij}(t)}(y_{ij}; \eta_j, \sigma_j, \lambda_j, t) = \sqrt{\frac{\lambda_j t^2}{2\pi y_{ij}^3 (\lambda_j \sigma_j^2 y_{ij} + 1)}} \exp \left[-\frac{\lambda_j (t - \eta_j y_{ij})^2}{2y_{ij} (\lambda_j \sigma_j^2 y_{ij} + 1)} \right].$$

Appendix B. Derivation of Eq. (3)

To derive the unconditional joint pdf of $\mathbf{Y}_i(t)$, we utilize a conclusion (shown in [Lemma 2](#) below) indicated on page 108 of [Rencher & Schaafje's book \(2008\)](#).

Lemma 2. If $\mathbf{x} \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the moment generating function (MGF) of $\mathbf{x}'\mathbf{A}\mathbf{x}$ is

$$\begin{aligned} M_{\mathbf{x}'\mathbf{A}\mathbf{x}}(t) &= E_{\mathbf{x}}[\exp(t\mathbf{x}'\mathbf{A}\mathbf{x})] \\ &= |\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma}|^{-1/2} \exp[-\boldsymbol{\mu}'[\mathbf{I} - (\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma})^{-1}]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}/2], \quad t \in \mathbb{R}. \end{aligned}$$

Note that in this lemma, t is a real parameter in the MGF, not the time in the degradation modeling. Then, based on [Lemma 2](#), the unconditional joint pdf of $\mathbf{Y}_i(t)$ is provided by

$$\begin{aligned} & f_{\mathbf{Y}_i(t)}(\mathbf{y}_i; \boldsymbol{\theta}, t) \\ &= \int \cdots \int f_{Y_{i1}(t)}(\mathbf{y}_i|\delta_i; \boldsymbol{\lambda}, t) f(\delta_i; \boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\rho}) d\delta_1 \cdots d\delta_p \\ &= E_{\delta_i}[f_{Y_{i1}(t)}(\mathbf{y}_i|\delta_i; \boldsymbol{\lambda}, t)] \\ &= \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j t^2}{2\pi y_{ij}^3}} \right) E_{\delta_i} \left[\exp \left[-\sum_{j=1}^p \frac{(t - y_{ij}\delta_{ij})^2}{2y_{ij}/\lambda_j} \right] \right] \\ &= (2\pi)^{-\frac{p}{2}} t^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{y_{ij}^3}} \right) E_{\delta_i} \left[\exp \left[-\sum_{j=1}^p \frac{\lambda_j y_{ij} (\delta_{ij} - t/y_{ij})^2}{2} \right] \right] \\ &= (2\pi)^{-\frac{p}{2}} t^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{y_{ij}^3}} \right) E_{\delta_i} \left[\exp \left[-\frac{1}{2} (\delta_i - \mathbf{v}_i)' \mathbf{V}_i (\delta_i - \mathbf{v}_i) \right] \right] \\ &= (2\pi)^{-\frac{p}{2}} t^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{y_{ij}^3}} \right) |\mathbf{I} + \mathbf{V}_i \boldsymbol{\Sigma}|^{-1/2} \\ &\quad \times \exp \left\{ -\frac{1}{2} (\boldsymbol{\eta} - \mathbf{v}_i)' [\mathbf{I} - (\mathbf{I} + \mathbf{V}_i \boldsymbol{\Sigma})^{-1}] \boldsymbol{\Sigma}^{-1} (\boldsymbol{\eta} - \mathbf{v}_i) \right\}, \end{aligned}$$

where $\mathbf{v}_i = (t/y_{i1}, t/y_{i2}, \dots, t/y_{ip})'$ and

$$\mathbf{V}_i = \begin{pmatrix} \lambda_1 y_{i1} & 0 & 0 & 0 \\ 0 & \lambda_2 y_{i2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p y_{ip} \end{pmatrix}.$$

This conclusion is based on the result that $(\delta_i - \mathbf{v}_i) \sim MVN(\boldsymbol{\eta} - \mathbf{v}_i, \boldsymbol{\Sigma})$.

Appendix C. Derivation of Eq. (5)

To derive the conditional distribution of δ_i given $\mathbf{y}_i(t)$, we utilize the Bayes' rule. That is

$$\begin{aligned} & f(\delta_i|\mathbf{y}_i(t)) \\ &= f(\mathbf{y}_i(t)|\delta_i) f(\delta_i) / f(\mathbf{y}_i(t)) \\ &\propto f(\mathbf{y}_i(t)|\delta_i) f(\delta_i) \\ &= (2\pi)^{-\frac{p}{2}} t^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{y_{ij}^3}} \right) \exp \left[-\sum_{j=1}^p \frac{\lambda_j y_{ij} (\delta_{ij} - t/y_{ij})^2}{2} \right] \\ &\quad \times (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\delta_i - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\delta_i - \boldsymbol{\eta}) \right] \\ &\propto \exp \left[-\frac{1}{2} (\delta_i - \mathbf{v}_i)' \mathbf{V}_i (\delta_i - \mathbf{v}_i) \right] \exp \left[-\frac{1}{2} (\delta_i - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\delta_i - \boldsymbol{\eta}) \right] \\ &\propto \exp \left[-\frac{1}{2} \delta_i' \mathbf{V}_i \delta_i + \delta_i' \mathbf{V}_i \mathbf{v}_i \right] \exp \left[-\frac{1}{2} \delta_i' \boldsymbol{\Sigma}^{-1} \delta_i + \delta_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\eta} \right] \\ &= \exp \left[-\frac{1}{2} \delta_i' (\boldsymbol{\Sigma}^{-1} + \mathbf{V}_i) \delta_i + \delta_i' (\boldsymbol{\Sigma}^{-1} \boldsymbol{\eta} + \mathbf{V}_i \mathbf{v}_i) \right]. \end{aligned}$$

The derivation is analogous to finding the posterior mean of a MVN distribution when a semiconjugate prior for the mean is given. For details, please refer to pages 107–108 of [Hoff's book \(2009\)](#).

Appendix D. Derivation of Eq. (6)

To derive the conditional distribution of the future degradation increments $\tilde{y}_i(s)$ given the current degradation measurements $\mathbf{y}_i(t)$, it is necessary to first condition on δ_i and then marginalize over $\delta_i|\mathbf{y}_i(t)$. Also based on Lemma 2, the result is given by

$$\begin{aligned} f(\tilde{\mathbf{y}}_i(s)|\mathbf{y}_i(t)) &= \int f(\tilde{\mathbf{y}}_i(s)|\delta_i)f(\delta_i|\mathbf{y}_i(t))d\delta_i \\ &= E_{\delta_i|\mathbf{y}_i(t)} \left[\prod_{j=1}^p \sqrt{\frac{\lambda_j s^2}{2\pi \tilde{y}_{ij}^3}} \exp \left[-\frac{(s - \tilde{y}_{ij}\delta_j)^2}{2\tilde{y}_{ij}/\lambda_j} \right] \right] \\ &= (2\pi)^{-\frac{p}{2}} s^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{\tilde{y}_{ij}^3}} \right) E_{\delta_i|\mathbf{y}_i(t)} \left[\exp \left[-\sum_{j=1}^p \frac{\lambda_j \tilde{y}_{ij} (\delta_j - s/\tilde{y}_{ij})^2}{2} \right] \right] \\ &= (2\pi)^{-\frac{p}{2}} s^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{\tilde{y}_{ij}^3}} \right) E_{\delta_i|\mathbf{y}_i(t)} \left[\exp \left[-\frac{1}{2} (\delta_i - \tilde{\mathbf{v}}_i)' \tilde{\Sigma}_i (\delta_i - \tilde{\mathbf{v}}_i) \right] \right] \\ &= (2\pi)^{-\frac{p}{2}} s^p \left(\prod_{j=1}^p \sqrt{\frac{\lambda_j}{\tilde{y}_{ij}^3}} \right) |\mathbf{I} + \tilde{\mathbf{v}}_i' \tilde{\Sigma}_i^{-1} \tilde{\mathbf{v}}_i|^{-1/2} \\ &\quad \times \exp \left\{ -\frac{1}{2} (\tilde{\eta}_i - \tilde{\mathbf{v}}_i)' [\mathbf{I} - (\mathbf{I} + \tilde{\mathbf{v}}_i' \tilde{\Sigma}_i^{-1})] \tilde{\Sigma}_i^{-1} (\tilde{\eta}_i - \tilde{\mathbf{v}}_i) \right\}, \end{aligned}$$

where $\tilde{\mathbf{v}}_i = (s/\tilde{y}_{i1}, s/\tilde{y}_{i2}, \dots, s/\tilde{y}_{ip})'$ and

$$\tilde{\mathbf{v}}_i = \begin{pmatrix} \lambda_1 \tilde{y}_{i1} & 0 & 0 & 0 \\ 0 & \lambda_2 \tilde{y}_{i2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \tilde{y}_{ip} \end{pmatrix}.$$

This conclusion is based on the result that $(\delta_i - \tilde{\mathbf{v}}_i)|\mathbf{y}_i(t) \sim MVN(\tilde{\eta}_i - \tilde{\mathbf{v}}_i, \tilde{\Sigma}_i)$.

Appendix E. Derivation of Eq. (8)

To derive the unconditional cdf of the failure time for an individual degradation process, we utilize an existing conclusion (shown in Lemma 3 below) provided by Si & Zhou (2014).

Lemma 3. If $X \sim N(\mu, \sigma^2)$, and $A, C, D \in \mathbb{R}$, then the following holds:

$$E_X[\exp(AX)\Phi(C+DX)] = \exp\left(A\mu + \frac{A^2}{2}\sigma^2\right)\Phi\left(\frac{C+D\mu+AD\sigma^2}{\sqrt{1+D^2\sigma^2}}\right).$$

Recall that the conditional cdf of the failure time for an individual degradation process implied by Eq. (7) is

$$\begin{aligned} F_{T_{D_j}}(t|\delta_j; \lambda_j, D_j) &= \Phi\left[-\sqrt{\frac{\lambda_j}{D_j}}(D_j\delta_j - t)\right] \\ &\quad - \exp(2\lambda_j\delta_j t)\Phi\left[-\sqrt{\frac{\lambda_j}{D_j}}(D_j\delta_j + t)\right], \end{aligned}$$

Given Lemma 3, if setting $A = 0$, $C = t\sqrt{\lambda_j/D_j}$, and $D = -\sqrt{\lambda_j/D_j}$ for the first part (i.e. $\Phi\left[-\sqrt{\frac{\lambda_j}{D_j}}(D_j\delta_j - t)\right]$) of $F_{T_{D_j}}(t; \delta_j, \lambda_j, D_j)$ and $A = 2\lambda_j t$, $C = -t\sqrt{\lambda_j/D_j}$, and $D = -\sqrt{\lambda_j/D_j}$ for the remaining part, then the unconditional cdf of the failure time is equivalent to $E_{\delta_j}[F_{T_{D_j}}(t|\delta_j; \lambda_j, D_j)]$, which is given by

$$F_{T_{D_j}}(t; \eta_j, \sigma_j, \lambda_j, D_j) = \Phi\left(\sqrt{\frac{\lambda_j}{D_j}} \frac{t - \eta_j D_j}{\sqrt{1 + \lambda_j \sigma_j^2 D_j}}\right)$$

$$- \exp[2\lambda_j t(\eta_j + \lambda_j \sigma_j^2 t)] \Phi\left(-\sqrt{\frac{\lambda_j}{D_j}} \frac{t + \eta_j D_j + 2\lambda_j \sigma_j^2 D_j t}{\sqrt{1 + \lambda_j \sigma_j^2 D_j}}\right).$$

Appendix F. The EM algorithm for statistical inference

F1. Main technical details

To carry out the EM algorithm, we utilize a conclusion (shown in Lemma 4 below) indicated on page 107 of Rencher & Schaale's book (2008).

Lemma 4. If $\mathbf{x}$ is a random vector with mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$ and if $\mathbf{A}$ is a symmetric matrix of constants, then

$$E_{\mathbf{x}}[\mathbf{x}'\mathbf{A}\mathbf{x}] = \text{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}.$$

Recall that the complete data are $\{\mathbb{D}, \delta\}$, of which the total log-likelihood function is

$$\ell(\boldsymbol{\theta}|\mathbb{D}, \delta) = \ell(\boldsymbol{\lambda}, \delta|\mathbb{D}) + \ell(\boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\rho}|\delta),$$

where

$$\begin{aligned} \ell(\boldsymbol{\lambda}, \delta|\mathbb{D}) &= \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left[\frac{1}{2} \ln \left(\frac{\lambda_j \tilde{\tau}_{ik}^2}{2\pi \tilde{y}_{ijk}^3} \right) - \frac{1}{2} \frac{(\tilde{\tau}_{ik} - \tilde{y}_{ijk}\delta_j)^2}{\tilde{y}_{ijk}/\lambda_j} \right] \text{ and} \\ \ell(\boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\rho}|\delta) &= \sum_{i=1}^n \left[-\frac{p}{2} \ln(2\pi) - \frac{1}{2} \ln |\boldsymbol{\Sigma}| - \frac{1}{2} (\delta_i - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\delta_i - \boldsymbol{\eta}) \right]. \end{aligned}$$

Given the current EM estimates $\boldsymbol{\theta}^{(s)}$, which consists of $\boldsymbol{\eta}^{(s)}$, $\boldsymbol{\Sigma}^{(s)}$ (i.e. $\boldsymbol{\sigma}^{(s)}$ and $\boldsymbol{\rho}^{(s)}$), and $\boldsymbol{\lambda}^{(s)}$. The EM algorithm is performed according to the followings:

- E-step: First, according to the conclusion implied by Eq. (5), $\delta_i|\mathbb{D}, \boldsymbol{\theta}^{(s)}$ is subject to a MVN distribution with mean vector

$$\begin{aligned} \tilde{\eta}_i^{(s)} &\equiv (\tilde{\eta}_{i1}^{(s)}, \tilde{\eta}_{i2}^{(s)}, \dots, \tilde{\eta}_{ip}^{(s)})' \\ &= \left(\boldsymbol{\Sigma}^{(s)-1} + \sum_{k=1}^{m_i} \tilde{\mathbf{v}}_{ik}^{(s)} \right)^{-1} \left(\boldsymbol{\Sigma}^{(s)-1} \boldsymbol{\eta}^{(s)} + \sum_{k=1}^{m_i} \tilde{\mathbf{v}}_{ik}^{(s)} \tilde{y}_{ik} \right) \end{aligned}$$

and variance-covariance matrix

$$\begin{aligned} \tilde{\Sigma}_i^{(s)} &\equiv \begin{pmatrix} (\tilde{\sigma}_{1(i)}^{(s)})^2 & \tilde{\sigma}_{12(i)}^{(s)} & \dots & \tilde{\sigma}_{1p(i)}^{(s)} \\ \tilde{\sigma}_{12(i)}^{(s)} & (\tilde{\sigma}_{2(i)}^{(s)})^2 & \dots & \tilde{\sigma}_{2p(i)}^{(s)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\sigma}_{1p(i)}^{(s)} & \tilde{\sigma}_{2p(i)}^{(s)} & \dots & (\tilde{\sigma}_{p(i)}^{(s)})^2 \end{pmatrix} \\ &= (\boldsymbol{\Sigma}^{(s)-1} + \sum_{k=1}^{m_i} \tilde{\mathbf{v}}_{ik}^{(s)})^{-1}, \end{aligned}$$

where $\tilde{\mathbf{v}}_{ik} = (\tilde{\tau}_{ik}/\tilde{y}_{i1k}, \tilde{\tau}_{ik}/\tilde{y}_{i2k}, \dots, \tilde{\tau}_{ik}/\tilde{y}_{ipk})'$ and

$$\tilde{\mathbf{v}}_{ik}^{(s)} = \begin{pmatrix} \lambda_1^{(s)} \tilde{y}_{i1k} & 0 & 0 & 0 \\ 0 & \lambda_2^{(s)} \tilde{y}_{i2k} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p^{(s)} \tilde{y}_{ipk} \end{pmatrix}.$$

Here, the conclusion implied by Eq. (5) is generalized to the case that multiple inspections are taken. Then, we know that

$$\begin{aligned} E_{\delta_i|\mathbb{D}, \boldsymbol{\theta}^{(s)}}[\delta_{ij}] &= \tilde{\eta}_{ij}^{(s)}, \quad E_{\delta_i|\mathbb{D}, \boldsymbol{\theta}^{(s)}}[\delta_{ij}^2] = (\tilde{\eta}_{ij}^{(s)})^2 + (\tilde{\sigma}_{j(i)}^{(s)})^2, \text{ and} \\ E_{\delta_i|\mathbb{D}, \boldsymbol{\theta}^{(s)}}[(\delta_i - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\delta_i - \boldsymbol{\eta})] &= \text{tr}(\boldsymbol{\Sigma}^{-1} \tilde{\Sigma}_i^{(s)}) \\ &\quad + (\tilde{\eta}_i^{(s)} - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\tilde{\eta}_i^{(s)} - \boldsymbol{\eta}), \end{aligned}$$

where the last equation is obtained through applying Lemma 4. Therefore, $Q(\theta|\theta^{(s)})$ is given by

$$\begin{aligned} Q(\theta|\theta^{(s)}) &= \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left\{ \frac{1}{2} \ln \left(\frac{\lambda_j \tau_{ijk}^2}{2\pi \tilde{y}_{ijk}^3} \right) \right. \\ &\quad \left. - \frac{1}{2} \frac{\left[\tau_{ijk}^2 - 2\tilde{t}_{ik}\tilde{y}_{ijk}\tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{j(i)}^{(s)} \right)^2 \right] \right]}{\tilde{y}_{ijk}/\lambda_j} \right\} \\ &\quad + \sum_{i=1}^n \left\{ -\frac{p}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} \left[\text{tr} \left(\Sigma^{-1} \tilde{\Sigma}_i^{(s)} \right) \right. \right. \\ &\quad \left. \left. + \left(\tilde{\eta}_i^{(s)} - \eta \right)' \Sigma^{-1} \left(\tilde{\eta}_i^{(s)} - \eta \right) \right] \right\}. \end{aligned}$$

- M-step: In this step, we update the estimated parameters by solving the equations that set the first derivative of $Q(\theta|\theta^{(s)})$ equal to 0. And, based on some existing conclusions about matrix differentiation, the results are given by

$$\begin{aligned} \lambda_j^{(s+1)} &= \frac{\sum_{i=1}^n m_i}{\sum_{i=1}^n \sum_{k=1}^{m_i} \frac{\left[\tau_{ijk}^2 - 2\tilde{t}_{ik}\tilde{y}_{ijk}\tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{j(i)}^{(s)} \right)^2 \right] \right]}{\tilde{y}_{ijk}}}, \\ \eta^{(s+1)} &= \frac{1}{n} \sum_{i=1}^n \tilde{\eta}_i^{(s)}, \\ \Sigma^{(s+1)} &= \frac{\sum_{i=1}^n \left[\tilde{\Sigma}_i^{(s)} + \left(\tilde{\eta}_i^{(s)} - \eta^{(s+1)} \right) \left(\tilde{\eta}_i^{(s)} - \eta^{(s+1)} \right)' \right]}{n}. \end{aligned}$$

These two steps iterate until achieving convergence. The criterion for convergence is usually built upon $\left| \ell(\theta^{(s+1)}|\mathbb{D}, \delta) - \ell(\theta^{(s)}|\mathbb{D}, \delta) \right| < \epsilon$ or $\max \left| \theta^{(s+1)} - \theta^{(s)} \right| < \epsilon$, where $\max$ is the element-wise maximum of a vector and ϵ is the error tolerance.

F2. Good guess of starting parameters

To provide a good set of starting parameters, it is performed by the following steps:

1. Treat the degradation increments of each degradation path, i.e. $\tilde{y}_{ij} = (\tilde{y}_{ij1}, \tilde{y}_{ij2}, \dots, \tilde{y}_{ijm_i})'$, $\forall i = 1, 2, \dots, n, j = 1, 2, \dots, p$ as an independent realization from a simple IG process, $IG(t/\delta_{ij}, \lambda_{ij}t^2)$. Fit the model to the data and obtain estimated parameters $\hat{\delta}_{ij}$ and $\hat{\lambda}_{ij}$, $\forall i = 1, 2, \dots, n, j = 1, 2, \dots, p$.
2. Set $\lambda_j^{(0)} = \frac{1}{n} \sum_{i=1}^n \hat{\lambda}_{ij}$, $\eta^{(0)} = \frac{1}{n} \sum_{i=1}^n \hat{\delta}_i$, and $\Sigma^{(0)} = \frac{1}{n} (\hat{\delta}_i - \eta^{(0)}) (\hat{\delta}_i - \eta^{(0)})'$ as the initial parameters feeding into the E-step, where $\hat{\delta}_i = (\hat{\delta}_{i1}, \hat{\delta}_{i2}, \dots, \hat{\delta}_{ip})'$, $\forall i = 1, 2, \dots, n$. Note that sometimes the resulting $\Sigma^{(0)}$ may not be pd, its nearest pd matrix (Higham, 2002) can be computed and fed into the E-step instead.

F3. Incorporation of time scale transformation function

If denoting the transformed time interval by $\tau_{ijk} = \Lambda_j(t_{ik}) - \Lambda_j(t_{i,k-1})$, $\forall i = 1, 2, \dots, n, j = 1, 2, \dots, p, k = 1, 2, \dots, m_i$ with an unknown parameter γ_j , then $Q(\theta|\theta^{(s)})$ in the E-step becomes

$$\begin{aligned} Q(\theta|\theta^{(s)}) &= \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left\{ \frac{1}{2} \ln \left(\frac{\lambda_j \tau_{ijk}^2}{2\pi \tilde{y}_{ijk}^3} \right) \right. \end{aligned}$$

$$\begin{aligned} &\quad \left. - \frac{1}{2} \frac{\left[\tau_{ijk}^2 - 2\tau_{ijk}\tilde{y}_{ijk}\tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{j(i)}^{(s)} \right)^2 \right] \right]}{\tilde{y}_{ijk}/\lambda_j} \right\} \\ &\quad + \sum_{i=1}^n \left\{ -\frac{p}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} \left[\text{tr} \left(\Sigma^{-1} \tilde{\Sigma}_i^{(s)} \right) \right. \right. \\ &\quad \left. \left. + \left(\tilde{\eta}_i^{(s)} - \eta \right)' \Sigma^{-1} \left(\tilde{\eta}_i^{(s)} - \eta \right) \right] \right\}. \end{aligned}$$

where $\tilde{\eta}_{ik} = (\tau_{i1k}/\tilde{y}_{i1k}, \tau_{i2k}/\tilde{y}_{i2k}, \dots, \tau_{ipk}/\tilde{y}_{ipk})'$.

And in the M-step, to obtain $\gamma_j^{(s+1)}$, we need to solve the following equation

$$\sum_{i=1}^n \sum_{k=1}^{m_i} \left\{ \frac{\tau'_{ijk}}{\tau_{ijk}} + \lambda_j^{(s+1)} \tilde{\eta}_{ij}^{(s)} \tau'_{ijk} - \lambda_j^{(s+1)} \tau_{ijk} \tau'_{ijk} / \tilde{y}_{ijk} \right\} = 0,$$

where τ'_{ijk} is the derivative of τ_{ijk} with respect to γ_j . Solutions to this equation can be found numerically.

In this paper, we make use of the power transformation, i.e. $\tau_{ijk} = t_{ik}^{\gamma_j} - t_{i,k-1}^{\gamma_j}$. Thus, τ'_{ijk} should be replaced with $t_{ik}^{\gamma_j} \ln t_{ik} - t_{i,k-1}^{\gamma_j} \ln t_{i,k-1}$ in the equation above. Particularly, when $t_{i,k-1} = 0$, we set $\tau_{ijk} = t_{ik}^{\gamma_j}$.

F4. Adaptation of the EM algorithm to fit model M_2

For Model M_2 ,

$$M_2 : \begin{cases} Y_{ij}(t) \sim IG(t/\delta_{ij}, \lambda_j t^2) \\ \delta_{ij} \sim N(\eta_j, \sigma_j^2) \end{cases},$$

we denote its parameters by $\theta = (\theta'_1, \theta'_2, \dots, \theta'_p)'$, where $\theta'_j = (\lambda_j, \eta_j, \sigma_j^2)'$, $j = 1, 2, \dots, p$. Also, we denote the complete data for process j by $\{\mathbb{D}_j, \delta_j\}$, where $\mathbb{D}_j$ is a subset of $\mathbb{D}$ associated with process j and $\delta_j = (\delta_{1j}, \delta_{2j}, \dots, \delta_{nj})'$. Then, the total log-likelihood function for process j is

$$\ell(\theta_j|\mathbb{D}_j, \delta_j) = \ell(\lambda_j, \delta_j|\mathbb{D}_j) + \ell(\eta_j, \sigma_j^2|\delta_j),$$

where

$$\begin{aligned} \ell(\lambda_j, \delta_j|\mathbb{D}_j) &= \sum_{i=1}^n \sum_{k=1}^{m_i} \left[\frac{1}{2} \ln \left(\frac{\lambda_j \tau_{ijk}^2}{2\pi \tilde{y}_{ijk}^3} \right) - \frac{1}{2} \frac{(\tau_{ijk} - \tilde{y}_{ijk} \delta_{ij})^2}{\tilde{y}_{ijk}/\lambda_j} \right] \text{ and} \\ \ell(\eta_j, \sigma_j^2|\delta_j) &= \sum_{i=1}^n \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \sigma_j^2 - \frac{1}{2\sigma_j^2} (\delta_{ij} - \eta_j)^2 \right]. \end{aligned}$$

Given the current EM estimates $\theta_j^{(s)}$, which consists of $\lambda_j^{(s)}$, $\eta_j^{(s)}$, and $\sigma_j^{(s)}$. The EM algorithm is performed according to the followings:

- E-step: First, analogous to the conclusion implied by Eq. (5), a routine calculation shows that $\delta_{ij}|\mathbb{D}_j, \theta_j^{(s)}$ is subject to a normal distribution with mean

$$\tilde{\eta}_{ij}^{(s)} = \frac{\lambda_j^{(s)} \sum_{k=1}^{m_i} \tilde{t}_{ik} + \eta_j^{(s)} / (\sigma_j^{(s)})^2}{1 / (\sigma_j^{(s)})^2}$$

and standard deviation

$$\tilde{\sigma}_{ij}^{(s)} = \sqrt{\frac{1}{\lambda_j^{(s)} \sum_{k=1}^{m_i} \tilde{y}_{ijk} + 1 / (\sigma_j^{(s)})^2}}.$$

Then, we know that

$$E_{\delta_{ij}|\mathbb{D}_j, \theta_j^{(s)}}[\delta_{ij}] = \tilde{\eta}_{ij}^{(s)} \text{ and } E_{\delta_{ij}|\mathbb{D}_j, \theta_j^{(s)}}[\delta_{ij}^2] = \left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{ij}^{(s)} \right)^2.$$

Therefore, $Q(\theta_j|\theta_j^{(s)})$ is given by

$$Q(\theta_j|\theta_j^{(s)}) = \sum_{i=1}^n \sum_{k=1}^{m_i} \left\{ \frac{1}{2} \ln \left(\frac{\lambda_j \tilde{\tau}_{ijk}^2}{2\pi \tilde{y}_{ijk}} \right) - \frac{1}{2} \frac{\left[\tilde{\tau}_{ijk}^2 - 2\tilde{\tau}_{ijk} \tilde{y}_{ijk} \tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{ij}^{(s)} \right)^2 \right] \right]}{\tilde{y}_{ijk} / \lambda_j} \right\} \\ + \sum_{i=1}^n \left\{ -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_j^2) - \frac{1}{2\sigma_j^2} \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{ij}^{(s)} \right)^2 - 2\eta_j \tilde{\eta}_{ij}^{(s)} + \eta_j^2 \right] \right\}.$$

- M-step: In this step, we update the estimated parameters by solving the equations that set the first derivative of $Q(\theta_j|\theta_j^{(s)})$ equal to 0. Therefore, the results are

$$\lambda_j^{(s+1)} = \frac{\sum_{i=1}^n \sum_{k=1}^{m_i} \tilde{y}_{ijk}}{\sum_{i=1}^n \sum_{k=1}^{m_i} \left[\frac{\tilde{\tau}_{ijk}^2 - 2\tilde{\tau}_{ijk} \tilde{y}_{ijk} \tilde{\eta}_{ij}^{(s)} + \tilde{y}_{ijk}^2 \left[\left(\tilde{\eta}_{ij}^{(s)} \right)^2 + \left(\tilde{\sigma}_{ij}^{(s)} \right)^2 \right]}{\tilde{y}_{ijk}} \right]}, \\ \eta_j^{(s+1)} = \frac{1}{n} \sum_{i=1}^n \tilde{\eta}_{ij}^{(s)}, \\ \sigma_j^{(s+1)} = \sqrt{\frac{\sum_{i=1}^n \left[\left(\tilde{\sigma}_{ij}^{(s)} \right)^2 + \left(\tilde{\eta}_{ij}^{(s)} - \eta_j^{(s+1)} \right)^2 \right]}{n}}.$$

To estimate all the unknown parameters, we carry out the EM algorithm for each individual process one at a time with a total of p implementations. The conclusion by Appendix F.2 is still applicable to provide a good guess of starting parameters except it is necessary to check $\sigma^{(0)}$ instead of $\Sigma^{(0)}$. In terms of incorporation of time scale transformation, the conclusion by Appendix F.3 remains the same. Thus, the inference for Model M_2 can be viewed as an analogously univariate version of the inference for Model M_0 .

Appendix G. The BCp bootstrap method

For illustrative purposes, we use the BCp bootstrap method to demonstrate the computation of the C.I. for $F_{T_D}(t; \theta, \mathcal{D})$ of the proposed model. It is performed according to the following steps:

1. Given the observed data $\mathbb{D}$, implement the EM algorithm to obtain the estimated parameters $\hat{\theta}$ and calculate the estimated failure time probability $\hat{F}_{T_D}(t; \hat{\theta}, \mathcal{D})$ (abbreviated $\hat{F}(t)$) at desired values of t .
2. Generate a large number B (say $B = 1,000$) of bootstrap samples that mimic the original sample and compute the

corresponding bootstrap estimates $\hat{F}_{T_D}^*(t; \hat{\theta}_b^*, \mathcal{D})_b$ (abbreviated $\hat{F}^*(t)_b$), $b = 1, 2, \dots, B$, according to the following steps:

- (a) Generate n simulated realizations of the random mean vector, i.e. $\delta_i^* \sim MVN(\hat{\eta}, \hat{\Sigma})$, $i = 1, 2, \dots, n$.
 - (b) For each δ_i^* , generate p simulated degradation paths based on $y_{ijk}^* = y_{ij,k-1}^* + y_{ijk}^*$, $\forall i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$, $k = 1, 2, \dots, m_i$, where $y_{ij0}^* = 0$ and y_{ijk}^* is sampled from $IG(\tau_{ik}/\delta_{ij}^*, \hat{\lambda}_j \tau_{ik}^2)$.
 - (c) Use the simulated degradation paths as inputs to produce the bootstrap estimates $\hat{\theta}_b^*$ and compute $\hat{F}^*(t)_b$ at desired values of t .
3. For each desired value of t , the bootstrap C.I. for $F_{T_D}(t; \theta, \mathcal{D})$ is constructed as below:
 - (a) Sort the bootstrap estimates $\hat{F}^*(t)_1, \dots, \hat{F}^*(t)_B$ in increasing order giving $\hat{F}^*(t)_{(b)}$, $b = 1, 2, \dots, B$.
 - (b) The lower and upper bounds of point-wise approximate $100(1 - \alpha)\%$ C.I. are $[\hat{F}^*(t)_{(L)}, \hat{F}^*(t)_{(U)}]$, where

$$L = B \times \Phi \left[2\Phi^{-1}(q) + \Phi^{-1}(\alpha/2) \right], \\ U = B \times \Phi \left[2\Phi^{-1}(q) + \Phi^{-1}(1 - \alpha/2) \right],$$

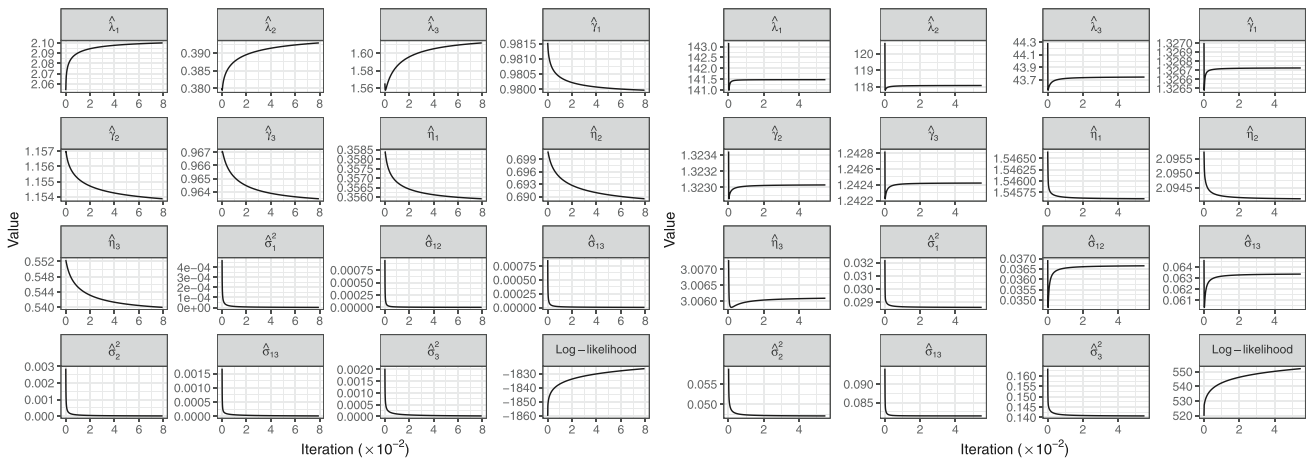
and q is the proportion of the bootstrap estimates $\{\hat{F}^*(t)_b, b = 1, 2, \dots, B\}$ that are less than $\hat{F}(t)$.

Appendix H. Additional results of the illustrative examples

Table 3 provides the data table of the fatigue crack-size data. Tables 4 and 5 show the point estimates of model parameters and their 95% bootstrap C.I.s (shown in the parentheses) for the coating data and the fatigue crack-size data, respectively. Fig. 9a and b provide plots of parameter estimates and log-likelihood over iteration for the coating data and the fatigue crack-size data, respectively.

For Model M_1 , the result of parameters estimation for the fatigue crack-size data is obtained as follows: $\hat{\delta} = (1.52670, 2.07223, 2.95884)'$, $\hat{\lambda} = (110.52359, 93.33662, 36.10819)'$, and $\hat{\gamma} = (1.31943, 1.31812, 1.23736)'$.

For Model M_2 , the result of parameters estimation for the fatigue crack-size data is obtained as follows: $\hat{\lambda} = (135.90509, 111.83610, 40.43586)'$, $\hat{\gamma} = (1.32563, 1.32199, 1.24042)'$, $\hat{\eta} = (1.54283, 2.08948, 2.98782)'$, and $\hat{\sigma} = (0.15363, 0.19554, 0.29746)'$.



(a) Coating Data

(b) Fatigue Crack-size Data

Fig. 9. Parameter Estimates and Log-likelihood over Iteration.

Table 3
Fatigue Crack-size Data.

Unit	Millions of Cycles ($\times 10$)									
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>PC1</i>										
1	0.90	0.95	1.00	1.05	1.12	1.19	1.27	1.35	1.48	1.64
2	0.90	0.94	0.98	1.03	1.08	1.14	1.21	1.28	1.37	1.47
3	0.90	0.94	0.98	1.03	1.08	1.13	1.19	1.26	1.35	1.46
4	0.90	0.94	0.98	1.03	1.07	1.12	1.19	1.25	1.34	1.43
5	0.90	0.94	0.98	1.03	1.07	1.12	1.18	1.23	1.33	1.41
6	0.90	0.94	0.98	1.02	1.07	1.11	1.17	1.23	1.32	1.41
<i>PC2</i>										
1	0.90	0.93	0.97	1.00	1.06	1.11	1.17	1.23	1.30	1.39
2	0.90	0.92	0.97	1.01	1.05	1.09	1.15	1.21	1.28	1.36
3	0.90	0.92	0.96	1.00	1.04	1.08	1.13	1.19	1.26	1.34
4	0.90	0.93	0.96	1.00	1.04	1.08	1.13	1.18	1.24	1.31
5	0.90	0.92	0.97	0.99	1.03	1.06	1.10	1.14	1.20	1.26
6	0.90	0.93	0.96	1.00	1.03	1.07	1.12	1.16	1.20	1.26
<i>PC3</i>										
1	0.90	0.92	0.96	0.99	1.03	1.06	1.10	1.16	1.21	1.27
2	0.90	0.92	0.95	0.97	1.00	1.03	1.07	1.10	1.16	1.22
3	0.90	0.93	0.96	0.97	1.00	1.05	1.08	1.11	1.16	1.20
4	0.90	0.92	0.94	0.97	1.01	1.04	1.07	1.09	1.14	1.19
5	0.90	0.92	0.94	0.97	0.99	1.02	1.05	1.08	1.12	1.16
6	0.90	0.92	0.94	0.97	0.99	1.02	1.04	1.07	1.11	1.14

Table 4
Results of Point and Interval Estimates of Model Parameters for the Coating Data.

Parameter	Estimation	Estimation	Estimation
λ	$\hat{\lambda}_1 = 2.09998$ (1.00035, 4.19730)	$\hat{\lambda}_2 = 0.39288$ (0.19208, 0.93007)	$\hat{\lambda}_3 = 1.61180$ (0.76006, 3.18489)
γ	$\hat{\gamma}_1 = 0.97996$ (0.91746, 1.04835)	$\hat{\gamma}_2 = 1.15386$ (1.06981, 1.21906)	$\hat{\gamma}_3 = 0.96347$ (0.90100, 1.03169)
η	$\hat{\eta}_1 = 0.35590$ (0.25360, 0.51402)	$\hat{\eta}_2 = 0.68947$ (0.43220, 0.98513)	$\hat{\eta}_3 = 0.54000$ (0.38455, 0.77688)
σ_j^2	$\hat{\sigma}_1^2 = 1.59788 \times 10^{-6}$ (8.43673×10^{-8} , 3.48284×10^{-6})	$\hat{\sigma}_2^2 = 13.19950 \times 10^{-6}$ (1.36569×10^{-6} , 2.86679×10^{-4})	$\hat{\sigma}_3^2 = 23.86160 \times 10^{-6}$ (7.42015×10^{-7} , 1.10614×10^{-3})
$\sigma_{jj'}$	$\hat{\sigma}_{12} = 1.93898 \times 10^{-6}$ (-9.40336×10^{-4} , 19.93434×10^{-4})	$\hat{\sigma}_{23} = 13.32560 \times 10^{-6}$ (-9.82983×10^{-4} , 3.21900×10^{-3})	$\hat{\sigma}_{13} = 3.07767 \times 10^{-6}$ (-4.50918×10^{-4} , 1.50048×10^{-3})

Table 5
Results of Point and Interval Estimates of Model Parameters for the Fatigue Crack-size Data.

Parameter	Estimation	Estimation	Estimation
λ	$\hat{\lambda}_1 = 141.47632$ (94.70044, 199.23721)	$\hat{\lambda}_2 = 118.08734$ (76.95887, 165.63843)	$\hat{\lambda}_3 = 43.74568$ (28.09148, 61.03539)
γ	$\hat{\gamma}_1 = 1.32673$ (1.24922, 1.40003)	$\hat{\gamma}_2 = 1.32303$ (1.25886, 1.38871)	$\hat{\gamma}_3 = 1.24242$ (1.15558, 1.32581)
η	$\hat{\eta}_1 = 1.54561$ (1.37547, 1.70408)	$\hat{\eta}_2 = 2.09412$ (1.88137, 2.29041)	$\hat{\eta}_3 = 3.00609$ (2.62293, 3.37090)
σ_j^2	$\hat{\sigma}_1^2 = 0.02859$ (37.94254×10^{-4} , 10.64927×10^{-2})	$\hat{\sigma}_2^2 = 0.04712$ (60.37184×10^{-4} , 17.17543×10^{-2})	$\hat{\sigma}_3^2 = 0.14072$ (17.28414×10^{-3} , 51.81899×10^{-2})
$\sigma_{jj'}$	$\hat{\sigma}_{12} = 0.03665$ (63.46901×10^{-4} , 11.93211×10^{-2})	$\hat{\sigma}_{23} = 0.08135$ (19.78837×10^{-3} , 31.27065×10^{-2})	$\hat{\sigma}_{13} = 0.06335$ (12.76516×10^{-3} , 23.09770×10^{-2})

References

Bae, S. J., & Kvam, P. H. (2004). A nonlinear random-coefficients model for degradation testing. *Technometrics*, 46(4), 460–469.

Chen, P., & Ye, Z.-S. (2018). Uncertainty quantification for monotone stochastic degradation models. *Journal of Quality Technology*, 50(2), 207–219.

Efron, B., & Tibshirani, R. (1994). *An introduction to the bootstrap* (1st ed.). New York: CRC Press.

Fang, G., & Pan, R. (2021). On multivariate copula modeling of dependent degradation processes. *Computers & Industrial Engineering*, 159, 107450.

Fang, G., Pan, R., & Hong, Y. (2020). Copula-based reliability analysis of degrading systems with dependent failures. *Reliability Engineering & System Safety*, 193, 106618.

Fang, G., Pan, R., & Stufken, J. (2021). Optimal setting of test conditions and allocation of test units for accelerated degradation tests with two stress variables. *IEEE Transactions on Reliability*, 70(3), 1096–1111.

Fang, G., Rigdon, S. E., & Pan, R. (2018). Predicting lifetime by degradation tests: A case study of ISO 10995. *Quality and Reliability Engineering International*, 34(6), 1228–1237.

Hajiha, M., Liu, X., & Hong, Y. (2020). Degradation under dynamic operating conditions: Modeling, competing processes and applications. *Journal of Quality Technology*, 0(0), 1–22.

Hao, S., Yang, J., & Berenguer, C. (2019). Degradation analysis based on an extended inverse Gaussian process model with skew-normal random effects and measurement errors. *Reliability Engineering & System Safety*, 189, 261–270. <https://doi.org/10.1016/j.res.2019.04.031>. <https://www.sciencedirect.com/science/article/pii/S095183201831442X>.

Higham, N. (2002). Computing the nearest correlation matrix – a problem from finance. *IMA Journal of Numerical Analysis*, 22, 329–343.

Hoff, P. D. (2009). *A first course in Bayesian statistical methods*. Springer.

Hong, L., Tan, M. H. Y., & Ye, Z.-S. (2020). Nonparametric link functions with shape constraints in stochastic degradation processes: Application to emerging contaminants. *Journal of Quality Technology*, 52(4), 370–384.

Hong, L., Ye, Z.-S., & Ling, R. (2018a). Environmental risk assessment of emerging contaminants using degradation data. *Journal of Agricultural, Biological and Environmental Statistics*, 23(3), 390–409.

Hong, L., Ye, Z.-S., & Sari, J. K. (2018b). Interval estimation for wiener processes based on accelerated degradation test data. *IIE Transactions*, 50(12), 1043–1057.

- Hu, J., Sun, Q., Ye, Z., & Zhou, Q. (2020). Joint modeling of degradation and lifetime data for RUL prediction of deteriorating products. *IEEE Transactions on Industrial Informatics*, 1–1.
- Keizer, M. C. A. O., Flapper, S. D. P., & Teunter, R. H. (2017). Condition-based maintenance policies for systems with multiple dependent components: A review. *European Journal of Operational Research*, 261(2), 405–420.
- Lawless, J., & Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, 10(3), 213–227.
- Li, H., Deloux, E., & Dieulle, L. (2016). A condition-based maintenance policy for multi-component systems with Lévy copulas dependence. *Reliability Engineering & System Safety*, 149, 44–55. <https://doi.org/10.1016/j.res.2015.12.011>.
- Lindig, S., Kaaya, I., Weiss, K., Moser, D., & Topic, M. (2018). Review of statistical and analytical degradation models for photovoltaic modules and systems as well as related improvements. *IEEE Journal of Photovoltaics*, 8(6), 1773–1786.
- Liu, B., Pandey, M. D., Wang, X., & Zhao, X. (2021). A finite-horizon condition-based maintenance policy for a two-unit system with dependent degradation processes. *European Journal of Operational Research*. <https://doi.org/10.1016/j.ejor.2021.03.010>. <https://www.sciencedirect.com/science/article/pii/S0377221721002009>.
- Liu, X., Al-Khalifa, K. N., Elsayed, E. A., Coit, D. W., & Hamouda, A. S. (2014). Criticality measures for components with multi-dimensional degradation. *IIE Transactions*, 46(10), 987–998.
- Lu, C. J., & Meeker, W. O. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 35(2), 161–174.
- Lu, L., Wang, B., Hong, Y., & Ye, Z. (2020). General path models for degradation data with multiple characteristics and covariates. *Technometrics*, 0(0), 1–16.
- Meeker, W. Q., & Escobar, L. A. (2014). *Statistical methods for reliability data*. John Wiley & Sons.
- Mercier, S., & Pham, H. H. (2012). A preventive maintenance policy for a continuously monitored system with correlated wear indicators. *European Journal of Operational Research*, 222(2), 263–272.
- Morita, L. H. M., Tomazella, V. L., Balakrishnan, N., Ramos, P. L., Ferreira, P. H., & Louzada, F. (2020). Inverse Gaussian process model with frailty term in reliability analysis. *Quality and Reliability Engineering International*, 37(2), 763–784.
- Palayangoda, L. K., & Ng, H. K. T. (2021). Semiparametric and nonparametric evaluation of first-passage distribution of bivariate degradation processes. *Reliability Engineering & System Safety*, 205, 107230. <https://doi.org/10.1016/j.res.2020.107230>. <https://www.sciencedirect.com/science/article/pii/S0951832020307304>.
- Peng, C.-Y. (2015). Inverse Gaussian processes with random effects and explanatory variables for degradation data. *Technometrics*, 57(1), 100–111.
- Rencher, A. C., & Christensen, W. F. (2012). *Methods of multivariate analysis*. John Wiley & Sons.
- Rencher, A. C., & Schaalje, G. B. (2008). *Linear models in statistics*. John Wiley & Sons.
- Seo, K., & Pan, R. (2020). Planning accelerated life tests with multiple sources of random effects. *Journal of Quality Technology*, 0(0), 1–22.
- Shi, Y., Feng, Q., Shu, Y., & Xiang, Y. (2020a). Multi-dimensional Lévy processes with Lévy copulas for multiple dependent degradation processes in lifetime analysis. *Quality Engineering*, 32(3), 434–448.
- Shi, Y., Xiang, Y., Liao, Y., Zhu, Z., & Hong, Y. (2020b). Optimal burn-in policies for multiple dependent degradation processes. *IIE Transactions*, 53(11), 1281–1293.
- Si, W., Yang, Q., Wu, X., & Chen, Y. (2018). Reliability analysis considering dynamic material local deformation. *Journal of Quality Technology*, 50(2), 183–197.
- Si, X.-S., & Zhou, D. (2014). A generalized result for degradation model-based reliability estimation. *IEEE Transactions on Automation Science and Engineering*, 11(2), 632–637.
- Sun, F., Fu, F., Liao, H., & Xu, D. (2020a). Analysis of multivariate dependent accelerated degradation data using a random-effect general Wiener process and d-vine copula. *Reliability Engineering & System Safety*, 204, 107168.
- Sun, Q., Ye, Z.-S., & Hong, Y. (2020b). Statistical modeling of multivariate destructive degradation tests with blocking. *Technometrics*, 62(4), 536–548.
- Wang, X., Balakrishnan, N., & Guo, B. (2015). Residual life estimation based on non-linear-multivariate Wiener processes. *Journal of Statistical Computation and Simulation*, 85(9), 1742–1764.
- Wang, X., Gaudoin, O., Doyen, L., Bérenguer, C., & Xie, M. (2020). Modeling multivariate degradation processes with time-variant covariates and imperfect maintenance effects. *Applied Stochastic Models in Business and Industry*, 37(3), 592–611.
- Wang, X., Wang, B. X., Hong, Y., & Jiang, P. H. (2021). Degradation data analysis based on gamma process with random effects. *European Journal of Operational Research*, 292(3), 1200–1208. <https://doi.org/10.1016/j.ejor.2020.11.036>. <https://www.sciencedirect.com/science/article/pii/S0377221720309826>.
- Wang, X., & Xu, D. (2010). An inverse Gaussian process model for degradation data. *Technometrics*, 52(2), 188–197.
- Whitmore, G. A., & Schenkelberg, F. (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, 3(1), 27–45.
- Wu, S., & Castro, I. T. (2020). Maintenance policy for a system with a weighted linear combination of degradation processes. *European Journal of Operational Research*, 280(1), 124–133. <https://doi.org/10.1016/j.ejor.2019.06.048>. <https://www.sciencedirect.com/science/article/pii/S0377221719305417>.
- Xu, A., Shen, L., Wang, B., & Tang, Y. (2018a). On modeling bivariate Wiener degradation process. *IEEE Transactions on Reliability*, 67(3), 897–906.
- Xu, A., Zhou, S., & Tang, Y. (2021). A unified model for system reliability evaluation under dynamic operating conditions. *IEEE Transactions on Reliability*, 70(1), 65–72. <https://doi.org/10.1109/tr.2019.2948173>.
- Xu, D., Xing, M., Wei, Q., Qin, Y., Xu, J., Chen, Y., & Kang, R. (2018b). Failure behavior modeling and reliability estimation of product based on Vine-copula and accelerated degradation data. *Mechanical Systems and Signal Processing*, 113, 50–64.
- Ye, Z.-S., & Chen, N. (2014). The inverse Gaussian process as a degradation model. *Technometrics*, 56(3), 302–311.
- Ye, Z.-S., Wang, Y., Tsui, K.-L., & Pecht, M. (2013). Degradation data analysis using Wiener processes with measurement errors. *IEEE Transactions on Reliability*, 62(4), 772–780.
- Ye, Z.-S., & Xie, M. (2015). Stochastic modelling and analysis of degradation for highly reliable products. *Applied Stochastic Models in Business and Industry*, 31(1), 16–32.
- Ye, Z.-S., Xie, M., Tang, L.-C., & Chen, N. (2014). Semiparametric estimation of Gamma processes for deteriorating products. *Technometrics*, 56(4), 504–513.
- Zhai, Q., & Ye, Z.-S. (2018). Degradation in common dynamic environments. *Technometrics*, 60(4), 461–471.
- Zhao, X., Chen, P., Gaudoin, O., & Doyen, L. (2021). Accelerated degradation tests with inspection effects. *European Journal of Operational Research*, 292(3), 1099–1114. <https://doi.org/10.1016/j.ejor.2020.11.041>. <https://www.sciencedirect.com/science/article/pii/S0377221720310006>.