Multivariate Reparameterized Inverse Gaussian Processes with Common Effects for Degradation-based Reliability Prediction

Liangliang Zhuang[a], Ancha Xu[a,b,*], Guanqi Fang[a,b], Yincai Tang[c]

[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, China*
[c] *The KLATASDS-MOE, School of Statistics, East China Normal University, Shanghai 200241, China*

Abstract

In industry, many highly reliable products possess multiple performance characteristics (PCs) and they typically degrade simultaneously. When such PCs are governed by a common failure mechanism or influenced by a shared operating environmental condition, interdependence between these PCs arises. To model such dependence, this article proposes a novel multivariate reparameterized inverse Gaussian (rIG) process model. It utilizes an additive structure; that is, the degradation of each marginal PC is considered as the result of the sum of two independent rIG processes with one capturing the shared common effects across all PCs and the other describing the intrinsic randomness specific to that PC. The model has some nice statistical properties, and the system lifetime distribution can be conveniently approximated. An expectation-maximization algorithm is proposed for estimating the model parameters, and a parametric bootstrap method is designed to derive the confidence intervals. Comprehensive numerical simulations are conducted to validate the performance of the inference method. Two case studies are thoroughly investigated to demonstrate the applicability of the proposed methodology. Supplementary materials for this article are available online.

Keywords: Common effects; reliability; inverse Gaussian process; multiple performance characteristics; interdependence

*Corresponding author: xuancha@mail.zjgsu.edu.cn

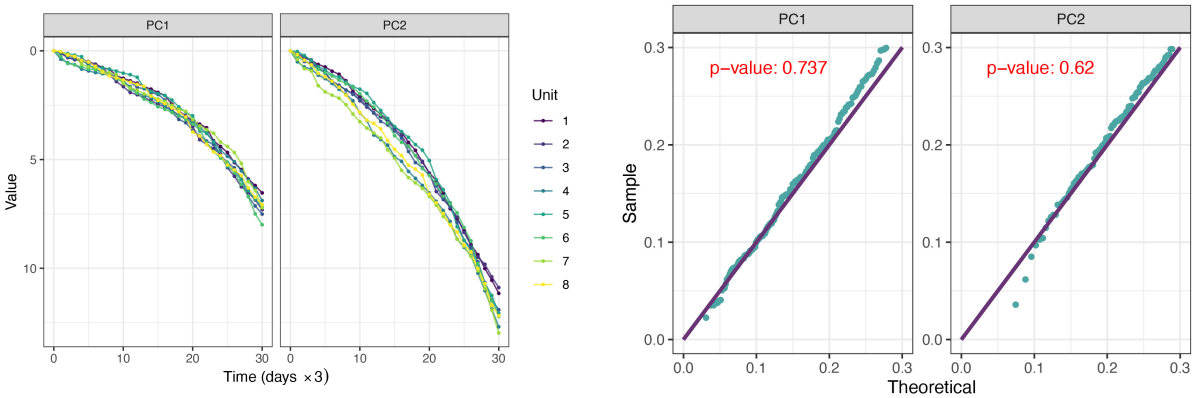
1. Introduction

1.1. Motivation

In industry, life testing has long been a conventional method for predicting the reliability of various engineering products. However, with the expansion of highly reliable products, this undertaking faces growing challenges, such as the scarcity of failure observations. An alternative strategy to overcome these challenges is to utilize degradation tests, which continuously monitor performance characteristics (PCs) directly related to the quality of products (Hong et al., 2020; Hajiha et al., 2021). Compared with the traditional life testing approach, leveraging the degradation data from these PCs proves to be an effective method for assessing system reliability in terms of improved accuracy and efficiency (Ye et al., 2019; Chen and Ye, 2018). When such PCs are governed by a common failure mechanism or influenced by a shared operating environmental condition, interdependence between these PCs arises. Zhai and Ye (2023) provided an example of polymer coating materials that undergo photooxidative degradation under ultraviolet radiation. Due to the correlated effects of photooxidation and the influence of common environmental conditions, the degradation behaviors of different chemical compositions within the material are interrelated. This interdependence leads to significant changes in the mechanical properties of the coating materials, such as modulus and hardness. Correctly capturing the interdependence, leading to the problem of multivariate degradation modeling, is crucial for an enhanced prediction of system reliability. Another insightful example below motivates our study.

Permanent magnet brakes (PMBs) utilize the magnetic force generated by permanent magnets to drive rotors and achieve braking by clamping a magnetic yoke. Renowned for their compact size, robust interference resistance, and high torque, PMBs are widely used in industrial production (Shin et al., 2013; Kou et al., 2021; Xu et al., 2024). To investigate the system reliability of PMBs, a degradation experiment was conducted, involving a test of eight samples in a typical high-temperature environment. In our experiment, these test units were employed as braking servo motors in injection molding machines. The malfunction of PMBs is the loss of magnetization, also known as demagnetization. This phenomenon is reflected by two PCs: braking torque (denoted as PC1), and response time (denoted as PC2). The degradation of these two PCs was measured in intervals of three days, and the experiment spanned three months. Figure 1(a) depicts the degradation paths of the PCs and Table 1 displays the Pearson correlation coefficients between the PCs. The results indicate that the sample correlations are all greater than 0.74, suggesting a strong interdependence. This

dependency arises from the shared underlying failure mechanism of demagnetization, and it can be referred to as a "common effect". Furthermore, we fit the degradation increments of each PC using inverse Gaussian (IG) distribution. Figure 1(b) presents the quantile-quantile (Q–Q) plots with the p-values from the Kolmogorov-Smirnov (KS) test. Apparently, the IG distribution provides a strong fit because most data points closely align along the theoretical straight lines, and the p-values from the KS tests are consistently greater than the chosen significance level (typically set as 0.05). Motivated by this example, we aim to establish a multivariate IG process model incorporating common effects.



(a) Degradation paths

(b) Q-Q plots using IG distribution

Figure 1: Summary of PMB data for two PCs: degradation paths and Q-Q Plots.

Table 1: Pearson correlation coefficients of two PCs across various units.

Unit	1	2	3	4	5	6	7	8
Correlation	0.819	0.749	0.806	0.840	0.779	0.749	0.765	0.800

1.2. Related literature

Over the past two decades, considerable research has been conducted on the problem of multivariate degradation modeling, for which the key is to account for the interdependence between multiple PCs (Hong et al., 2018; Kang et al., 2020). In the existing literature, three primary approaches to modeling multivariate degradation are identified.

i) **Copula-based method**: built upon some stochastic processes for modeling marginals, this method employs a designated copula function to capture the interdependence between PCs (Sun et al., 2021; Zhuang et al., 2021). For instance, Peng et al. (2016)

recommended modeling the degradation of each PC using a Wiener process or an IG process. They further advocated the incorporation of a multivariate copula function to represent the interdependence between different PCs. The effectiveness of this approach was substantiated through a comprehensive degradation analysis of heavy machine tools. Fang and Pan (2021) utilized three types of copula functions, namely elliptical copulas, exchangeable Archimedean copulas, and vine copulas, to characterize intricate dependence structures in correlated degradation processes. Jiang et al. (2023) performed a reliability analysis on a mechanical system involving two PCs. The model employed four copula functions to capture the dependencies between the PCs and constructed a predictive model for remaining useful life (RUL).

ii) **Multivariate distribution-based method**: this approach involves modeling the multi-dimensional degradation data using well-behaved multivariate stochastic processes or independent stochastic processes with multivariate random effects (Fang and Pan, 2023; Pan and Balakrishnan, 2011). For instance, Dong et al. (2020) proposed a correlated Wiener process to model a bivariate degradation process. They applied the proposed model to rail track geometry degradation data and conducted a reliability analysis. Yan et al. (2023) modeled the degradation processes of multiple PCs by using multivariate Wiener processes, considering the correlation between the degradation rate and volatility. Fang et al. (2022) utilized independent IG processes with correlated random effects to describe the degradation process of multiple dependent PCs. They further applied this model to both coating data and fatigue crack-size data.

iii) **Common-effect-based method**: this approach suggests that the subsystems within certain complex systems, operating in the same environment, are influenced by a common unobservable latent factor in their degradation paths. There are primarily two modeling approaches: a) Frailty model-based method: this approach incorporates frailty factors to account for unmeasured common influences and subsystem heterogeneity. Xu et al. (2018) first implemented this method in a two-dimensional degradation model, using a Wiener process and assuming a normally distributed frailty factor. Subsequently, this framework was further developed by Song and Cui (2022) and Barui et al. (2024). b) Stochastic process summation method: this method employs a specific stochastic process to depict the common environmental influences. By integrating the common process into the independent degradation process of each PC, a multi-PCs de-

pendent degradation model can be developed (Mercier and Pham, 2012; Huynh et al., 2022). Based on this method, Liu et al. (2021) proposed a bivariate gamma process for maintenance strategies, and Zhai and Ye (2023) introduced a multivariate Wiener process for reliability analysis.

Among the three mentioned methods, the copula-based approach encounters significant challenges in selecting an appropriate copula function or providing clear physical interpretations (Xu et al., 2018). The majority of methods based on multivariate random effects are confined to bivariate cases, and extending these models to encompass multivariate distributions remains a formidable challenge (Yan et al., 2023). The common-effect-based method is a prevalent technique in multivariate degradation modeling due to its degradation mechanism-driven and parsimonious nature. However, the frailty model-based method, which employs a single frailty factor in the multidimensional degradation model, restricts the model's universality. In contrast, the key advantage of the stochastic process summation method is that the number of model parameters increases only linearly with the variable dimension. This linear parameterization significantly simplifies the complexity of modeling high-dimensional performance degradation (Zhai and Ye, 2023). Nevertheless, existing work in multivariate degradation modeling primarily concentrates on Wiener processes due to their ideal physical interpretability and excellent mathematical properties for characterizing nonmonotonic degradation processes. A notable gap exists in the availability of a multivariate degradation modeling framework based on monotonic degradation processes, such as gamma or IG processes.

1.3. Contributions and key insights

For the multivariate gamma process, Huynh et al. (2022) proposed a common-effect-based multivariate degradation model, while their work only focused on maintenance assuming that the model parameters were known. To the best of our knowledge, there is scarce literature providing statistical inference and reliability analysis for monotonically dependent multivariate degradation processes subject to common environmental influences. To bridge this gap, we aim to propose a multivariate degradation model for IG processes based on the common-effect method, designed to model the monotonic degradation of multiple dependent PCs. Following the approach by Mercier and Pham (2012), we may use two IG processes to jointly characterize the degradation path of the k-th PC, i.e., $Y_k(t) = X_k(t) + Z(t)$, where $X_k(t)$ follows classical IG process $\mathcal{IG}(\vartheta_k t, \eta t^2)$ to describe the PC-specific degradation, and

$Z(t)$ follows classical IG process $\mathcal{IG}(\vartheta_0 t, \eta t^2)$ to capture the influence of the common environment on the PC. The classical IG process, defined by Ye and Chen (2014), is widely employed in modeling degradation data. Despite the prevalent use of the classical IG process, the summation of two independent classical IG processes does not yield an IG process. Consequently, deriving the stochastic properties of $Y_k(t)$ poses a significant challenge, making it difficult to obtain the lifetime distribution of the system. To address challenges associated with the traditional IG process, we propose a novel reparameterized IG (rIG) degradation model characterized by its inherent additivity. The main contributions and innovations of this paper are as follows:

- We construct a multivariate rIG process using the common-effect method and conduct analyses of its properties and the system's lifetime distribution.
- We utilize the Gauss-Legendre (GL) quadrature method for approximation due to the complexity of directly computing the integral in the lifetime distribution.
- The expectation-maximization (EM) algorithm is employed to estimate model parameters, supplemented by a parametric bootstrap approach to construct confidence intervals (CIs).
- Comprehensive numerical simulations and two case studies are conducted to demonstrate the practical applicability of the proposed approach.

1.4. Overview

The structure of the remainder of this paper is as follows. Section 2 introduces the multivariate modeling framework based on the rIG process and thoroughly examines the properties of the proposed model. Following that, Section 3 delves into the reliability analysis, while Section 4 develops an EM algorithm for acquiring maximum likelihood (ML) estimates. In this section, CIs for model parameters are established using parametric bootstrap. To validate the performance of the proposed model, Section 5 conducts simulation studies. In Section 6, the proposed model is applied to two datasets. Finally, the conclusion is presented in Section 7.

2. Model formulation

2.1. The rIG degradation model

A stochastic process $\{Z(t), t \geq 0\}$ is termed an rIG process, denoted as $r\mathcal{IG}(\Lambda(t), \gamma)$, if it satisfies the following properties. (i) $Z(0) = 0$ with probability one; (ii) The increments of $Z(t)$ are independent. Specifically, for all $t_2 > t_1 \geq s_2 > s_1 \geq 0$, the increments $Z(t_2) - Z(t_1)$ and $Z(s_2) - Z(s_1)$ are mutually independent; (iii) For all $t > s \geq 0$, the increment $Z(t) - Z(s)$ follows the rIG distribution denoted as $rIG(\Lambda(t) - \Lambda(s), \gamma)$. Here, $\Lambda(t)$ is a monotone increasing function with the initial condition $\Lambda(0) = 0$, δ represents the drift parameter and γ denotes the dispersion parameter. The cumulative distribution function (CDF) and probability density function (PDF) of $rIG(\delta, \gamma)$ are defined as

$$F_{rIG}(y; \delta, \gamma) = \Phi\left[\sqrt{y}\gamma - \frac{\delta}{\sqrt{y}}\right] + e^{2\delta\gamma}\Phi\left[-\sqrt{y}\gamma - \frac{\delta}{\sqrt{y}}\right], \tag{1}$$

and

$$f_{rIG}(y; \delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} e^{\delta\gamma} y^{-3/2} e^{-\left(\delta^2 y^{-1} + \gamma^2 y\right)/2}, \ y > 0, \ \delta > 0, \ \gamma > 0, \tag{2}$$

respectively, where $\Phi(\cdot)$ is the CDF of standard normal distribution. The parameter relationship between the rIG distribution $rIG(\delta, \gamma)$ and the traditional IG distribution $IG(a, b)$ is $a = \delta/\gamma$ and $b = \delta^2$. The motivation for the reparameterization approach lies in the specific additive property of the rIG distribution. Specifically, if a random variable $Y \sim rIG(\delta, \gamma)$, then its moment generating function is given by

$$M_Y(t) = E(e^{ty}) = e^{\delta\gamma\left(1 - \sqrt{1 - \frac{2t}{\gamma^2}}\right)}. \tag{3}$$

Based on (3), it can be found that if $Y_1 \sim rIG(\delta_1, \gamma)$ and $Y_2 \sim rIG(\delta_2, \gamma)$, then $Y = Y_1 + Y_2 \sim rIG(\delta_1 + \delta_2, \gamma)$. Utilizing (3), we can obtain that the mean and variance of $\{Y(t), t \geq 0\}$ are $(\delta_1 + \delta_2)/\gamma$ and $(\delta_1 + \delta_2)/\gamma^3$, respectively.

2.2. Multivariate rIG process with common effects

We consider a system with K PCs. By convention, we posit that the initial degradation of each PC is zero at time $t = 0$, and the degradation process of the k-th PC can be formulated as

$$Y_k(t) = X_k(t) + Z(t), k = 1, \ldots, K \tag{4}$$

with

$$Z(t) \sim r\mathcal{IG}(\Lambda_0(t), \gamma) \quad \text{and} \quad X_k(t) \sim r\mathcal{IG}(\Lambda_k(t), \gamma), k = 1, \ldots, K, \tag{5}$$

where $Z(t)$ and $X_k(t)$, $k = 1, \ldots, K$ are independent of each other, $\Lambda_k(t)$ and $\Lambda_0(t)$ represent monotonically increasing functions of t. In practical applications, the specific functional form of $\Lambda_k(t)$ can be determined through engineering expertise or empirical investigation. Model (4) is inspired by the observed degradation mechanism in physical systems (e.g., PMBs, heavy machine tools). We formulate the degradation for each PC as the sum of two independent rIG processes. One rIG process represents the common effects affecting all PCs, reflecting the overall health status of the system that impacts each PC uniformly. The other rIG process describes degradation arising from distinctive randomness specific to each PC, capturing the accumulation of damage that is independent of the systemic effect. By combining the internal degradation effects and the overall functional dependence, model (4) offers a comprehensive framework that allows for a more holistic consideration of the degradation and dependence among various PCs within the system. This approach brings the model closer to real-world scenarios and facilitates a more accurate understanding of the system's operational and maintenance requirements.

Based on the additive property of the rIG distribution induced by (3), $Y_k(t)$ is also a rIG process satisfying

$$Y_k(t) \sim rIG\left(\Lambda_k(t) + \Lambda_0(t), \gamma\right), k = 1, \ldots, K. \tag{6}$$

For the proposed model, the following two propositions are established.

Proposition 1. *The mean and variance of the degradation process $Y_k(t)$ are*

$$\mathrm{E}\left[Y_k(t)\right] = \frac{\Lambda_0(t) + \Lambda_k(t)}{\gamma}, \quad \textit{and} \quad \mathrm{Var}\left[Y_k(t)\right] = \frac{\Lambda_0(t) + \Lambda_k(t)}{\gamma^3}, \tag{7}$$

respectively. Meanwhile, the common effect $Z(t)$ introduces dependence among the multiple degradation processes

$$\mathrm{Cov}\left[Y_{k_1}\left(t_1\right), Y_{k_2}\left(t_2\right)\right] = \frac{\min\left(\Lambda_0\left(t_1\right), \Lambda_0\left(t_2\right)\right)}{\gamma^3}, k_1 \neq k_2. \tag{8}$$

Consequently, at any given time t, the Pearson correlation coefficient between $Y_{k_1}(t)$ and $Y_{k_2}(t)$ is

$$\rho\left[Y_{k_1}(t), Y_{k_2}(t)\right] = \frac{\Lambda_0(t)}{\sqrt{\left(\Lambda_0(t) + \Lambda_{k_1}(t)\right)\left(\Lambda_0(t) + \Lambda_{k_2}(t)\right)}}, k_1 \neq k_2. \tag{9}$$

The proof is provided in Supplementary Section S1.1. Proposition 1 reveals that the correlation between any two PCs primarily arises from the common effect $Z(t)$. When

$\Lambda_0(t)$ significantly surpasses $\Lambda_k(t)$, the Pearson correlation coefficient (9) tends towards 1. Conversely, if the common effect $Z(t)$ vanishes (i.e., $\Lambda_0(t) = 0$), the Pearson correlation coefficient will be 0.

Proposition 2. *The joint PDF of* $Y_1(t), \ldots, Y_K(t)$ *can be obtained as*

$$f_{Y(t)}(y_1, \ldots, y_K) = \int_0^{\tilde{y}} f_{rIG}(z; \Lambda_0(t), \gamma) \prod_{k=1}^{K} f_{rIG}(y_k - z; \Lambda_k(t), \gamma)\, \mathrm{d}z, \tag{10}$$

where $\tilde{y} = \min\{y_1, \ldots, y_K\}$, *where* $y_1, \ldots, y_K$ *are the observed dagradation values. And* $f_{rIG}(\cdot)$ *is given by* (2). *The corresponding CDF is expressed as*

$$F_{Y(t)}(y_1, \ldots, y_K) = \int_0^{\tilde{y}} f_{rIG}(z; \Lambda_0(t), \gamma) \prod_{k=1}^{K} F_{rIG}(y_k - z; \Lambda_k(t), \gamma)\, \mathrm{d}z, \tag{11}$$

where $F_{rIG}(\cdot)$ *is given by* (1).

Proposition 2 gives the joint PDF and CDF of $Y_1(t), \ldots, Y_K(t)$, which can be used for deriving the ML estimates of the model parameters.

3. Reliability analysis

According to engineering practice, the failure time of the system is of critical concern for engineers. A common method of failure definition for the k-th PC is associated with a failure threshold $\mathcal{D}_k$. That is, if any of the K PCs exceeds its respective threshold, the system is considered to have failed (Lu et al., 2021; Zhuang et al., 2024). Following this concept, we derive the system failure time for the proposed model. Firstly, let the failure time of the k-th PC be represented as follows

$$T_{\mathcal{D}_k} = \inf\{t : Y_k(t) \geq \mathcal{D}_k\}.$$

As a rIG process has monotone degradation paths, the distribution of the failure time can be expressed as $P(T_{\mathcal{D}_k} < t) = P(Y_k(t) \geq \mathcal{D}_k)$. Then, the CDF of the failure time for an individual degradation process can be derived as follows

$$\begin{aligned} &F_{T_{\mathcal{D}_k}}(t \mid \Lambda_0(t), \Lambda_k(t), \gamma, \mathcal{D}_k) \\ &\quad = 1 - \left\{ \Phi\left[\sqrt{\mathcal{D}_k}\gamma - \frac{\Lambda_0(t) + \Lambda_k(t)}{\sqrt{\mathcal{D}_k}}\right] + e^{2[\Lambda_0(t)+\Lambda_k(t)]\gamma} \Phi\left[-\sqrt{\mathcal{D}_k}\gamma - \frac{\Lambda_0(t) + \Lambda_k(t)}{\sqrt{\mathcal{D}_k}}\right] \right\}, \\ &\quad \triangleq 1 - F_{rIG}(\mathcal{D}_k; \Lambda_0(t) + \Lambda_k(t), \gamma). \end{aligned} \tag{12}$$

Then, the system failure time can be defined by

$$T_{\boldsymbol{\mathcal{D}}} = \inf\left\{t : Y_1(t) \geq \mathcal{D}_1 \text{ or } \cdots \text{ or } Y_K(t) \geq \mathcal{D}_K\right\},$$

where $\boldsymbol{\mathcal{D}} = (\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_K)'$ is a vector storing all PC failure thresholds. We have the following result for the CDF of system failure time.

Proposition 3. *The CDF of system failure time $T_{\boldsymbol{\mathcal{D}}}$ is*

$$F_{T_{\boldsymbol{\mathcal{D}}}}(t \mid \boldsymbol{\Lambda}(t), \gamma, \boldsymbol{\mathcal{D}}) = \int_0^{\tilde{y}} \left[1 - \prod_{k=1}^{K} \left(F_{rIG}(\mathcal{D}_k - z; \Lambda_k(t), \gamma)\right)\right] f_{rIG}\left(z; \Lambda_0(t), \gamma\right) \mathrm{d}z, \tag{13}$$

where $\boldsymbol{\Lambda}(t) = (\Lambda_0(t), \ldots, \Lambda_K(t))'$, and $\tilde{y} = \min\{y_1, \ldots, y_K\}$.

The proof is provided in Supplementary Section S1.2. Due to the complexity of directly computing the integral in (13), we employ the GL quadrature method to approximate this integral. GL quadrature is a specific form of Gaussian quadrature designed for the accurate approximation of challenging integrals (Swarztrauber, 2003). In comparison to Monte Carlo (MC) integration, GL quadrature provides precise estimates of the integral while requiring a significantly reduced computational budget (Babolian et al., 2005). After applying the GL integral approximation, (13) can be approximated as

$$\begin{aligned} &F_{T_{\boldsymbol{\mathcal{D}}}}(t \mid \boldsymbol{\Lambda}(t), \gamma, \boldsymbol{\mathcal{D}}) \\ &\quad \approx \frac{\tilde{y}}{2} \sum_{q=1}^{l} w_q \left[1 - \prod_{k=1}^{K} \left(F_{rIG}(\mathcal{D}_k - \frac{\tilde{y}(u_q+1)}{2}; \Lambda_k(t), \gamma)\right)\right] f_{rIG}\left(\frac{\tilde{y}(u_q+1)}{2}; \Lambda_0(t), \gamma\right). \end{aligned} \tag{14}$$

where l is a given order, u_q is the root of the Legendre polynomial and w_q is the corresponding weight. More details on integration approximation can be found in Supplementary Section S2. To validate the accuracy of the GL approximation, we plot the CDF of system failure time for four scenarios from Table 2 using the GL formula in (14), alongside empirical CDFs and MC-based approximation method with 1000 samples for comparison (see Figure 2). The results demonstrate that the GL-based approximation outperforms the MC-based method in all scenarios with $l = 10$ and $l = 15$. In terms of computational efficiency, the MC method averages 1.837 seconds per execution, while the GL method only takes about 0.020 and 0.029 seconds for $l = 10$ and $l = 15$, respectively. Consequently, to balance accuracy and efficiency, this study opts for the GL method with $l = 10$ to approximate the complex integral.

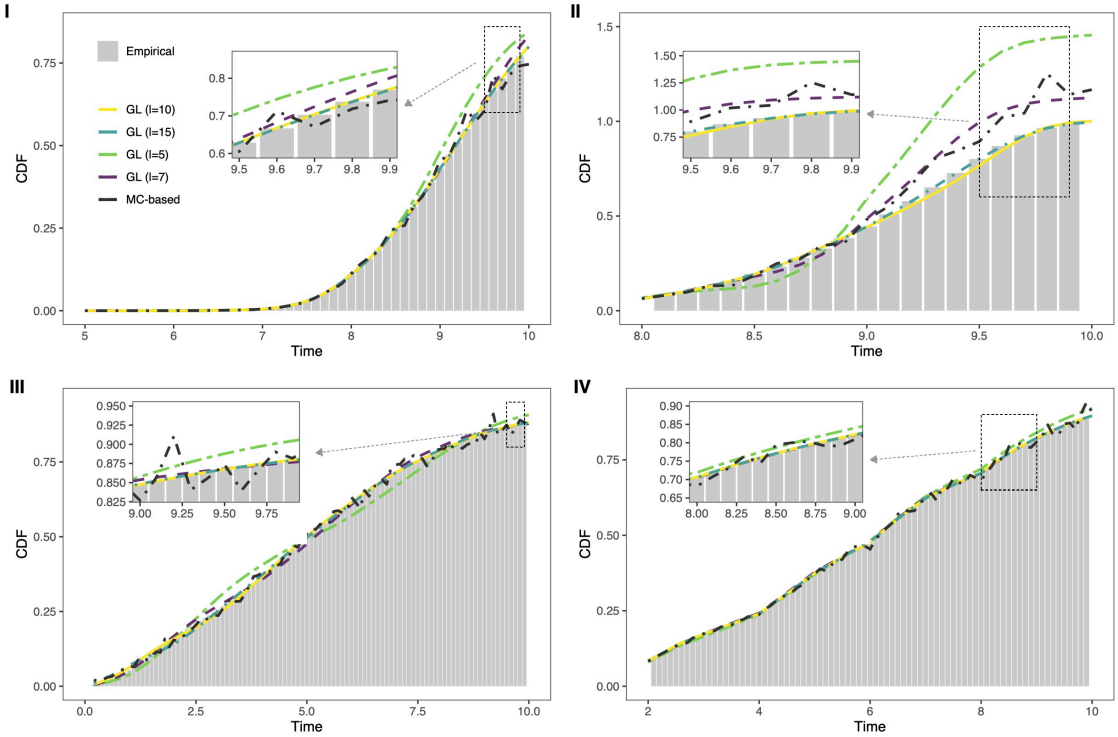


Figure 2: CDF of system failure time across four scenarios.

4. Parameter estimation

This section focuses on the parametric multivariate rIG model when $\Lambda_k(t)$ has a specific function form with unknown parameters. We assume that $\Lambda_k(t) = \Lambda_k(t; \alpha_k, \beta_k)$ involves unknown parameters α_k and β_k, where $k = 0, \ldots, K$. Commonly used functions include power-law form $\Lambda_k(t; \alpha_k, \beta_k) = \beta_k t^{\alpha_k}$ and log-linear form $\Lambda_k(t; \alpha_k, \beta_k) = \beta_k \left[\exp(\alpha_k t) - 1\right]$. However, randomly choosing function forms may cause the issue of parameter nonidentifiability. For instance, when $\Lambda_k(t) = \beta_k t^{\alpha}$ and $\Lambda_0(t) = \beta_0 t^{\alpha}$, we derive that $Y_k(t) \sim rIG\left((\beta_0 + \beta_k)t^{\alpha}, \gamma\right)$. In such a case, β_0 and β_k can not be identified, and we can only estimate $\beta_0 + \beta_k$ from the observed data. To avoid this issue, we assume $\Lambda_0(t) = \Lambda_k(t; \alpha_0)$. Let $\boldsymbol{\Lambda}(t) = (\Lambda_0(t), \ldots, \Lambda_K(t))'$. The model parameters are represented as $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma\}$, with $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)'$ and $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_K)'$.

Assume that there are n systems tested in an experiment. The degradation of the K PCs in the i-th system are measured at m_i time points, denoted as $\boldsymbol{T}_i = (t_{i,1}, \ldots, t_{i,m_i})'$, and the degradation values are $\boldsymbol{Y}_{i,k} = (Y_{i,k,1}, \ldots, Y_{i,k,m_i})'$ for $k = 1, \ldots, K$, $i = 1, \ldots, n$. Define the degradation increments of the k-th PC between $(t_{i,j-1}, t_{i,j}]$ as $\Delta Y_{i,k,j} \triangleq Y_{i,k,j} - Y_{i,k,j-1}$ for $j = 1, \ldots, m_i$, where we set $t_{i,0} = 0$ and $Y_{i,k,0} = 0$. Let $\Delta \boldsymbol{Y}_{i,:,j} = (\Delta Y_{i,1,j}, \ldots, \Delta Y_{i,K,j})'$.

Based on (10), we have

$$p\left(\Delta \boldsymbol{Y}_{i,:,j} \mid \boldsymbol{\theta}\right) = \int_0^{\Delta \tilde{y}_{i,j}} f_{rIG}\left(\Delta z_{i,j}; \Delta\Lambda_0(t_{i,j}), \gamma\right) \times \prod_{k=1}^{K} f_{rIG}\left(\Delta y_{i,k,j} - \Delta z_{i,j}; \Delta\Lambda_k(t_{i,j}), \gamma\right) \mathrm{d}\Delta z_{i,j}, \tag{15}$$

where $\Delta\tilde{y}_{i,j} = \min\{\Delta Y_{i,1,j}, \ldots, \Delta Y_{i,K,j}\}$. The log-likelihood function can be easily derived (Shao and Wang, 2023), and the model parameters can be estimated by maximizing the given log-likelihood function

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \ln p\left(\Delta \boldsymbol{Y}_{i,:,j} \mid \boldsymbol{\theta}\right).$$

However, since $p(\Delta \boldsymbol{Y}_{i,:,j}|\boldsymbol{\theta})$ is a complex function of $\boldsymbol{\theta}$, obtaining the analytical form of $\hat{\boldsymbol{\theta}}$ is challenging. Instead, we develop an EM algorithm to derive the ML estimates.

4.1. EM algorithm for point estimation

We consider the unobserved common effect $Z_{i,j} = Z_i\left(t_{i,j}\right)$ for $i = 1, \ldots, n, j = 1, \ldots, m_i$ as the missing data, and define $\boldsymbol{Z}_i = (Z_{i,1}, \ldots, Z_{i,m_i})'$, $\Delta\Lambda_{i,k,j} = \Delta\Lambda_k(t_{i,j}), k = 1, \ldots, K$. Given $\boldsymbol{Z}_i$, we have that

$$\Delta Y_{i,k,j} - \Delta Z_{i,j} \mid \boldsymbol{Z}_i \sim r\mathcal{IG}\left(\Delta\Lambda_{i,k,j}, \gamma\right), \tag{16}$$

with $0 \leq \Delta Z_{i,j} \leq \Delta\tilde{y}_{i,j}$. Denote $\mathbb{Y} = \{\Delta \boldsymbol{Y}_{i,:,j}, i = 1, \ldots, n, j = 1, \ldots, m_i\}$, and $\mathbb{Z} = \{\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n\}$. The log-likelihood with the complete data is

$$\ell(\boldsymbol{\theta} \mid \mathbb{Y}, \mathbb{Z}) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \left\{ \sum_{k=1}^{K} \ln p\left(\Delta Y_{i,k,j} - \Delta Z_{i,j} \mid \Delta Z_{i,j}\right) + \ln p\left(\Delta Z_{i,j}\right) \right\}, \tag{17}$$

where

$$\ln p\left(\Delta Y_{i,k,j} - \Delta Z_{i,j} \mid \Delta Z_{i,j}\right) = -\frac{1}{2}\ln(2\pi) + \ln \Delta\Lambda_{i,k,j} - \frac{3}{2}\ln\left(\Delta Y_{i,k,j} - \Delta Z_{i,j}\right) + \gamma\Delta\Lambda_{i,k,j} - \frac{\Delta\Lambda_{i,k,j}^2}{2\left(\Delta Y_{i,k,j} - \Delta Z_{i,j}\right)} - \frac{\gamma^2\left(\Delta Y_{i,k,j} - \Delta Z_{i,j}\right)}{2},$$

and

$$\ln p\left(\Delta Z_{i,j}\right) = -\frac{1}{2}\ln(2\pi) + \ln \Delta\Lambda_{i,0,j} + \gamma\Delta\Lambda_{i,0,j} - \frac{3}{2}\ln \Delta Z_{i,j} - \frac{\Delta\Lambda_{i,0,j}^2}{2\Delta Z_{i,j}} - \frac{\gamma^2 \Delta Z_{i,j}}{2}.$$

Here, $\Delta\Lambda_{i,0,j} = \Delta\Lambda_0(t_{i,j})$, $\Delta\Lambda_{i,0,j} = \Lambda_{i,0,j} - \Lambda_{i,0,j-1}$. The main idea of the EM algorithm is to obtain an estimate of $\boldsymbol{\theta}$ by iteratively executing E-steps and M-steps until convergence is achieved. We assume that $\boldsymbol{\theta}^{(s)}$ is the optimal solution in the M-step during the s-th iteration. At the $(s+1)$-th iteration, we need to calculate the following Q-function

$$Q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(s)}\right) = \mathrm{E}\left[\ell(\boldsymbol{\theta} \mid \mathbb{Y}, \mathbb{Z}) \mid \mathbb{Y}, \boldsymbol{\theta}^{(s)}\right], \tag{18}$$

which is the expectation of the complete log-likelihood with respect to $p\left(\mathbb{Z} \mid \mathbb{Y}, \boldsymbol{\theta}^{(s)}\right)$. The detailed derivation of the conditional distribution $p\left(\mathbb{Z} \mid \mathbb{Y}, \boldsymbol{\theta}^{(s)}\right)$, and the conditional expectations involved, $\mathrm{E}\left[(\Delta Y_{i,k,j} - \Delta Z_{i,j})^{-1} \mid \Delta \boldsymbol{Y}_{i,:,j}\right]$, $\mathrm{E}\left[\Delta Z_{i,j} \mid \Delta \boldsymbol{Y}_{i,:,j}\right]$, and $\mathrm{E}\left[\Delta Z_{i,j}^{-1} \mid \Delta \boldsymbol{Y}_{i,:,j}\right]$, can be found in Supplementary Section 3.1. When we obtain the Q-function in (18), the update of the optimal solution in the M-step is carried out as follows

$$\boldsymbol{\theta}^{(s+1)} = \arg\max_{\boldsymbol{\theta}} Q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(s)}\right). \tag{19}$$

This can be accomplished by computing the partial derivatives of the Q-function with respect to $\boldsymbol{\theta}$ and setting these derivatives equal to zero (refer to Supplementary Section 3.2 for detailed explanations). Finally, the ML estimation for $\boldsymbol{\theta}$ is obtained until convergence according to a specified criterion. The proposed EM algorithm can be implemented with the following key steps.

- **Initialization**: Start with initial values $\boldsymbol{\theta}^{(\mathbf{0})}$ for the parameters $\boldsymbol{\theta}$ (see Section 4.2), and set the tolerance error ω.
- **E-step**: Calculate $Q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(s)}\right)$ by (18), based on the s-th iteration of parameters estimation $\boldsymbol{\theta}^{(s)}$.
- **M-step**: Compute the $(s+1)$-th parameter estimation $\boldsymbol{\theta}^{(s+1)}$ using (19).
- **Iteration**: Iterate through the E-step and M-step until $\left\|\boldsymbol{\theta}^{(s+1)} - \boldsymbol{\theta}^{(s)}\right\| < \omega$, where $\|\cdot\|$ denotes the Euclidean distance.
- **Output**: Obtain the ML estimates of $\boldsymbol{\theta}$.

4.2. Determine initial parameter estimation values

In the execution of the EM algorithm, the choice of initial parameter values is pivotal, as it substantially impacts both the convergence time and the algorithm's effectiveness in

discovering globally optimal solutions. To determine the initial values of $\boldsymbol{\theta}$, a two-step method is proposed. Initially, we calculate the sample mean and variance of the degradation paths for k-th PC at time j as follows

$$\Delta\bar{Y}_{:,k,j} = \frac{1}{n}\sum_{i=1}^{n}\Delta Y_{i,k,j}, \quad \text{and} \quad \Delta s^2_{:,k,j} = \frac{\sum_{i=1}^{n}\left(\Delta Y_{i,k,j} - \Delta\bar{Y}_{:,k,j}\right)^2}{n-1}.$$

According to (7) and the principles of moment estimation, we calculate the estimate for γ,

$$\hat{\gamma} = \sqrt{\frac{\sum_{k=1}^{K}\sum_{j=1}^{m_i}\Delta\bar{Y}_{:,k,j}}{\sum_{k=1}^{K}\sum_{j=1}^{m_i}\Delta s^2_{:,k,j}}}.$$

As $\mathrm{E}\left[\Delta Y_{i,k,j}\right]$ represents the average degradation increment of the k-th PC of unit i at time j, a preliminary estimate of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ can be derived by minimizing the weighted mean square error between $\mathrm{E}\left[\Delta Y_{i,k,j}\right]$ and $\Delta Y_{i,k,j}$. Assuming the estimate of γ is known, we optimize the following formula to calculate the estimated values of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$.

$$\begin{aligned}\Psi &= \arg\min_{\boldsymbol{\beta},\boldsymbol{\alpha}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=1}^{K}\frac{\left(\mathrm{E}\left[\Delta Y_{i,k,j}\right]-\Delta Y_{i,k,j}\right)^2}{\mathrm{var}\left[\Delta Y_{i,k,j}\right]}\\ &= \arg\min_{\boldsymbol{\beta},\boldsymbol{\alpha}}\hat{\gamma}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=1}^{K}\left[\frac{\hat{\gamma}^2\Delta Y^2_{i,k,j}}{\Delta\Lambda_{i,0,j}+\Delta\Lambda_{i,k,j}}+\Delta\Lambda_{i,0,j}+\Delta\Lambda_{i,k,j}\right].\end{aligned} \tag{20}$$

4.3. Bootstrap for interval estimation

Our current emphasis lies in establishing interval estimates for $\boldsymbol{\theta}$ through the utilization of the parametric bootstrap method (Zhai and Ye, 2023; Lamu and Yan, 2024). The detailed bootstrap procedure is delineated in Algorithm 1. To derive an approximate $100(1-\varsigma)\%$ bootstrap CIs for a function of the parameters, denoted as $\mathcal{X}(\boldsymbol{\theta})$, we follow two steps as follows.

- **Bootstrap sampling**: implement Algorithm 1 to generate B bootstrap estimates $\left\{\hat{\boldsymbol{\theta}}^*_1,\ldots,\hat{\boldsymbol{\theta}}^*_B\right\}$.
- **Interval construction**: utilize the obtained bootstrap estimates to construct the bootstrap CIs. Specifically, for a given significance level ς, the interval is computed as follows

$$\left[\mathcal{X}\left(\hat{\boldsymbol{\theta}}^*\right)_{(\varsigma B/2)}, \mathcal{X}\left(\hat{\boldsymbol{\theta}}^*\right)_{((1-\varsigma/2)B)}\right],$$

where $\mathcal{X}\left(\hat{\boldsymbol{\theta}}^*\right)_{(p)}$ denotes the p-th empirical quantile of the bootstrap estimates.

Algorithm 1: Bootstrap algorithm procedure.

Input: Point estimate $\hat{\boldsymbol{\theta}}$.

Output: B replicates of estimates, $\left\{\hat{\boldsymbol{\theta}}_1^*, \ldots, \hat{\boldsymbol{\theta}}_B^*\right\}$.

```
for b = 1 to B do
    for i = 1 to n do
        for j = 1 to m_i do
            Generate samples ΔZ̃_{i,j} from rIG(Δλ̂_{i,0,j}, γ̂);
            for k = 1 to K do
                Generate samples ΔX̃_{i,k,j} from rIG(Δλ̂_{i,k,j}, γ̂);
                Obtain samples ΔỸ_{i,k,j} = ΔX̃_{i,k,j} + ΔZ̃_{i,j}.
            end
        end
    end
    Obtain pseudo degradation increment data ΔỸ.
    Obtain θ̂*_b based on ΔỸ using the proposed EM algorithm.
end
```

This approach furnishes an effective means of approximating the interval within which the function $\mathcal{X}(\boldsymbol{\theta})$ is expected to lie with confidence level $100(1-\varsigma)\%$.

4.4. Model validation

In this section, we concentrate on verifying the goodness of fit (GOF) of the proposed model. Following the approach described in Fang et al. (2022), we adapt model validation techniques from univariate IG processes to the multivariate case. Based on the properties of the IG distribution as detailed in Wang and Xu (2010), we extend these properties to the rIG distribution. Specifically, for a random variable $Y \sim rIG(\delta, \gamma)$, the transformation $(\gamma Y - \delta)^2/Y$ follows a χ^2 distribution with one degree of freedom. Hence, for the k-th PC, where $\Delta Y_{i,j,k} \sim rIG\left(\Delta\Lambda_k(t_{i,j}) + \Delta\Lambda_0(t_{i,j}), \gamma\right)$, the statistic $[\hat{\gamma}\Delta Y_{i,j,k} - \Delta\hat{\Lambda}_k(t_{i,j}) - \Delta\hat{\Lambda}_0(t_{i,j})]^2/\Delta Y_{i,j,k}$ approximates an i.i.d. χ_1^2 distribution. The χ_1^2 Q-Q plots serve to evaluate the GOF for each PC's rIG process model, supplemented by the KS test for statistical GOF assessment. If the p-value is greater than the significance level, we cannot reject the null hypothesis that the data follows χ_1^2 distribution. To investigate whether different PCs share a common dispersion

parameter γ, a likelihood ratio test is implemented. Under the assumption of independence among PCs, two models are examined: i) a unified model where all PCs share a common γ, and ii) a heterogeneous model where each PC is characterized by a distinct dispersion parameter $\gamma_i, i = 1, \ldots, K$. The parameters of these models are estimated using the MLE method, with their log-likelihoods denoted as ℓ_1 and ℓ_2 respectively. The test statistic $\tau = -2(\ell_1 - \ell_2)$ follows a chi-square distribution with $K - 1$ degrees of freedom under the null hypothesis that asserts a uniform γ across all PCs. A p-value below the significance level leads to the rejection of this hypothesis, thereby suggesting significant variations in γ among the PCs. Finally, we employ the Akaike information criterion (AIC) to select the optimal model, calculated as $\text{AIC} = 2\kappa - 2\ell$, where ℓ represents the corresponding log-likelihood and κ represents the total number of model parameters.

5. Simulation study

In this section, we evaluate the performance of the proposed inference method through an MC simulation study. Specifically, we consider a multivariate rIG process with common effects with three PCs, i.e., $K = 3$. Four combinations of $\Lambda_k(t)$ and $\Lambda_0(t)$ are examined, as detailed in Table 2, along with their corresponding parameter values. Assuming the degradation of the n units is measured at the same time intervals, and all m_i are equal, i.e., $m_1 = \cdots = m_n \equiv m = 10$. To assess the influence of the sample size on inference, we consider three unit sizes n: $5, 8$, and 10. For each configuration, we perform 500 MC replications of data generated from the simulated model and fit them using the EM algorithm, on a laptop with an Apple M2 Pro CPU. Simulated degradation paths for three PCs are presented in Figure S1 of Supplementary Section S4, with additional results for higher dimensions also available in the same section.

Table 2: Four combinations of $\Lambda_k(t)$ and $\Lambda_0(t)$ with corresponding parameters.

Scen.	$\Lambda_k(t)$	$\Lambda_0(t)$	$\boldsymbol{\alpha}'$	$\boldsymbol{\beta}'$	γ	$\boldsymbol{\mathcal{D}}$
I	$\beta_k t^{\alpha_k}$	t^{α_0}	$(1, 0.8, 1, 1.2)$	$(0.8, 1, 1.2)$	4	$(3.6, 4.8, 7.2)$
II	$\beta_k t^{\alpha_k}$	$\exp(\alpha_0 t) - 1$	$(1, 0.25, 0.33, 0.37)$	$(0.8, 1, 1.2)$	4	$(4, 8, 13)$
III	$\beta_k\,[\exp(\alpha_k t) - 1]$	t^{α_0}	$(0.05, 0.3, 0.4, 0.4)$	$(0.8, 1, 1.2)$	4	$(0.44, 0.66, 0.88)$
IV	$\beta_k\,[\exp(\alpha_k t) - 1]$	$\exp(\alpha_0 t) - 1$	$(0.1, 0.1, 0.1, 0.1)$	$(0.8, 1, 1.2)$	4	$(0.42, 0.56, 0.42)$

5.1. Validation of model parameter estimators

The proposed EM algorithm is employed for point estimation on each synthetic dataset. It's noteworthy that the proposed EM algorithm typically converges in approximately 200 iterations (with $\left\|\boldsymbol{\theta}^{(s+1)} - \boldsymbol{\theta}^{(s)}\right\| < 10^{-4}$). The interval estimates are constructed based on 200 bootstrap samples for the EM algorithm. To assess the overall quality of the parameter estimation method, we compute the relative root mean square error (RRMSE), and the coverage probability (CP) of the 95% CIs for the three different n across the 100 replications. RRMSE for the parameter v based on 100 repetitions is defined as follows:

$$\text{RRMSE}(\hat{v}) = \left[\frac{1}{100}\sum_{i=1}^{100}\left(\frac{\hat{v}_i - v}{v}\right)^2\right]^{1/2}.$$

Figures 3 and 4 show the RRMSE and CP of estimators based on the proposed parameter estimation method. These results reveal that for each estimator, the RRMSE decreases as n increases. Conversely, the CP tends towards 95% with increasing n. While the performance of parameter estimation differs across scenarios, the overall effectiveness is satisfactory, particularly with moderately sized n.

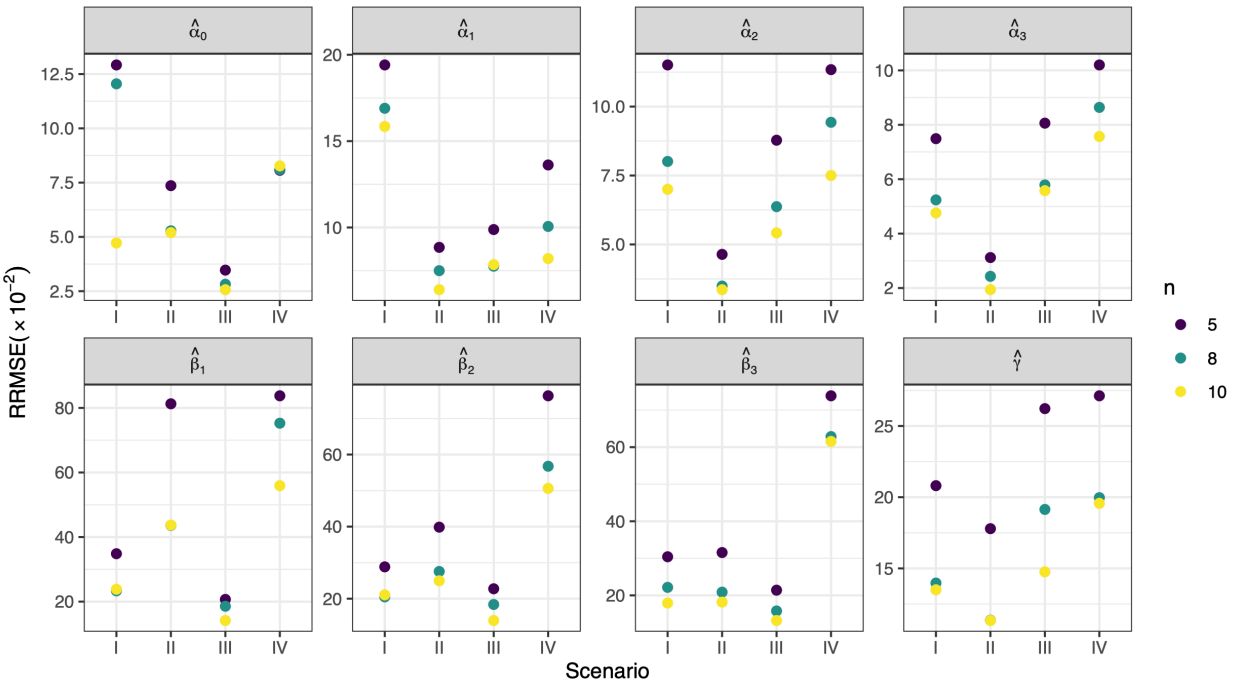


Figure 3: RRMSE ($\times 10^{-2}$) of estimators across different unit sizes and scenarios.

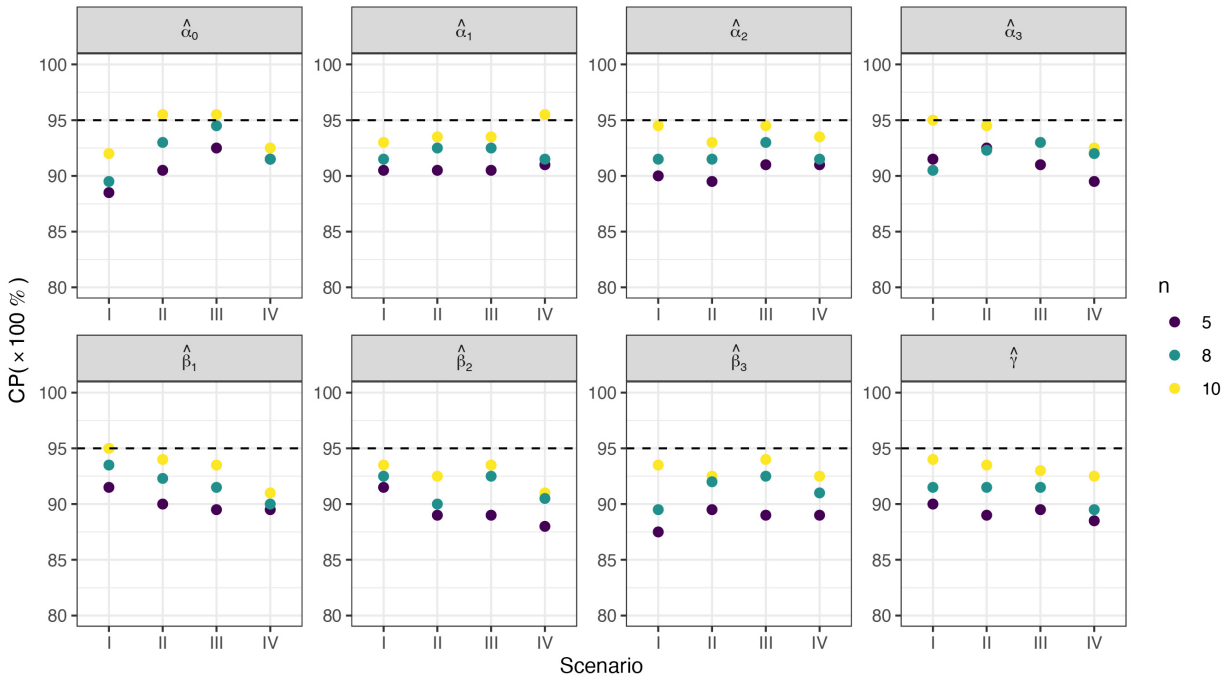


Figure 4: CP (×100%) of estimators across different unit sizes and scenarios.

5.2. Estimates of correlation coefficients between PCs

We calculate correlation coefficients for three pairs of PCs: $\rho(Y_1(t), Y_2(t))$, $\rho(Y_1(t), Y_3(t))$, and $\rho(Y_2(t), Y_3(t))$ based on (9). Due to space limitations, we place the RRMSE ($\times 10^{-2}$) of the correlation estimates for various unit sizes and scenarios over time in Figure S2 with Supplementary Section S4.1. Similar to the parameter estimation results, the RRMSE of the correlation estimates is relatively small and decreases as the sample size increases. Additionally, RRMSE exhibits different trends across scenarios. Scenario II shows an initial increase followed by a decrease, particularly in $\rho(Y_2(t), Y_3(t))$. In other scenarios, RRMSE gradually increases over time. This phenomenon can be attributed to the nonlinearity of the correlations over time.

5.3. Performance of reliability estimation

Given the parameter estimations under and failure thresholds, the mean time to failure (MTTF) of the system,

$$\text{MTTF} = E(T) = \int_0^\infty 1 - F_{T_{\boldsymbol{\mathcal{D}}}}(t \mid \boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma, \boldsymbol{\mathcal{D}}) \mathrm{d}t,$$

is computed to evaluate the performance of reliability estimation. Based on the results from previous parameter estimations, we calculate the MTTF for each simulated sample

and compare these calculations with the actual MTTF values. The boxplots of RRMSE of MTTF under different scenarios and unit sizes are presented in Figure 5. Consistent with previous results, our estimates exhibit relatively small RRMSE values, which gradually decreased with increasing sample sizes, and the occurrence of outliers also decreased. In addition, we compare the reliability of the proposed model with another model to assess the effect of model mis-specification. A benchmark model is considered, which assumes that PCs are independent of each other, i.e., $Y_k(t) \sim r\mathcal{IG}\left(\Lambda_k(t;\alpha_k,\beta_k),\gamma\right), k = 1,\ldots,K$. We utilize ML method to infer parameter estimates. After repeating the simulation 100 times for each scenario, the average RRMSE ($\times 10^{-2}$) of MTTF under the four scenarios based on the PC independence model is 4.629, compared with 2.215 for the proposed model. It can be seen that for simulation data with inter-PC dependencies, if we use the PC independence model for analysis, the reliability estimate obtained has a large deviation. Such inaccurate reliability estimates may cause unnecessary maintenance costs and increase the risk of production interruption.

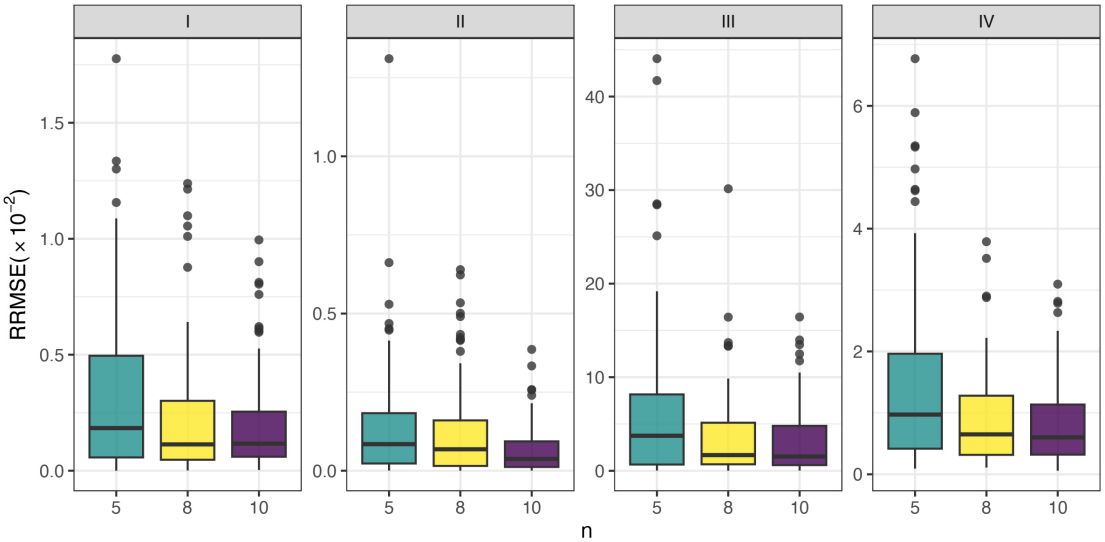


Figure 5: RRMSE ($\times 10^{-2}$) of MTTF estimators across different unit sizes and scenarios.

6. Case studies

This section demonstrates the application of the proposed methodology through analysis of the PMB degradation data that motivated this study (see Section 6.1), and the fatigue crack-size data (see Section 6.2).

6.1. PMB degradation data

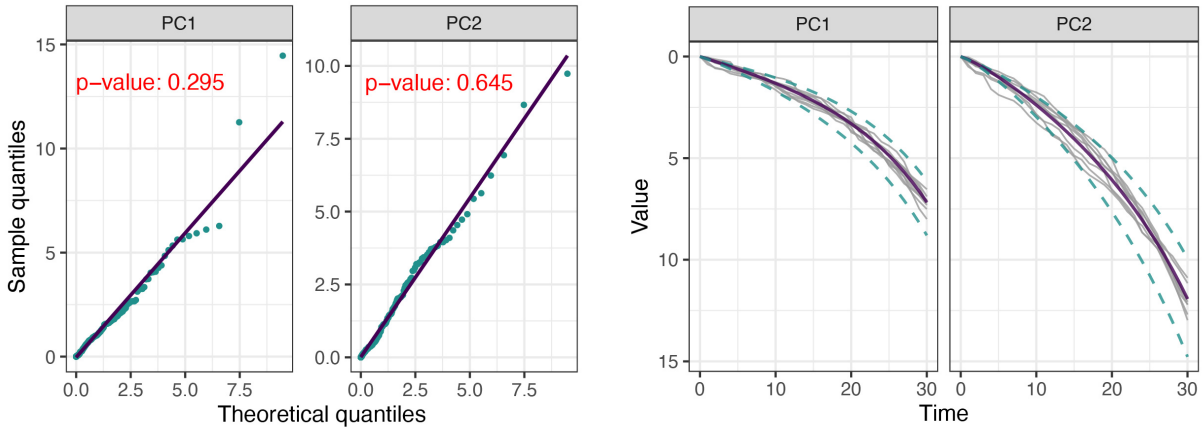
We apply the proposed model for data fitting and consider four different scenarios. Model parameters are estimated using the EM algorithm, and the convergence of model inference is assessed through trace plots of parameter estimation, as presented in Figure S4 within Supplementary Section S5. We employ the bootstrap approach with $B = 500$ replicates to obtain interval estimation. To compare the superiority of the proposed models, we introduce a model assuming independence between all PCs, i.e., $\Lambda_0(t) = 0$. The model parameters are estimated using ML methods. Table 3 presents the parameter estimation results for different models. Additionally, the AIC values for each model are provided. It is evident from these results that the AIC values for the proposed models are consistently lower than those of the independent model. Among the four scenarios, Scenario III exhibits the lowest AIC, which we regard as the optimal model for PMB data. The estimated results of α from the table indicate that the degradation of PC2 is faster than that of PC1. This is consistent with what we observe in the degradation paths in Figure 1.

Table 3: Parameter point estimates regarding the PMB data.

Model	Scen.	α_0	α_1	α_2	β_1	β_2	γ	AIC
Proposed	I	0.866	1.296	1.463	0.028	0.124	3.030	-2219.427
	II	0.724	0.104	0.068	0.942	6.182	4.375	-2494.123
	III	0.100	0.994	1.205	0.395	0.531	4.263	**-2603.588**
	IV	0.098	0.009	0.025	42.099	30.476	4.299	-2594.714
Independent	Power	-	1.518	1.456	0.151	0.317	3.703	-637.266
	Log-linear	-	0.056	0.054	6.830	11.944	4.079	-744.887

To validate the fitted model, we utilize the χ^2_1 Q-Q plot and conduct tests to determine whether different PCs share the same dispersion parameter γ, as outlined in Section 4.4. The Q-Q plot presented in Figure 6(a) illustrates that most data points closely align with the theoretical straight line. Furthermore, the KS test yields p-values greater than 0.05 for both PCs, supporting the adequacy of the model fit to the dataset. Additionally, based on the optimal model for PMB data (Scenario III where $\Lambda_k(t) = \beta_k[\exp(\alpha_k t) - 1]$), we conduct a likelihood ratio test. The p-value is 0.233, which exceeds 0.05, indicating no significant differences in the γ values across the PCs, thereby affirming their homogeneity. In Figure

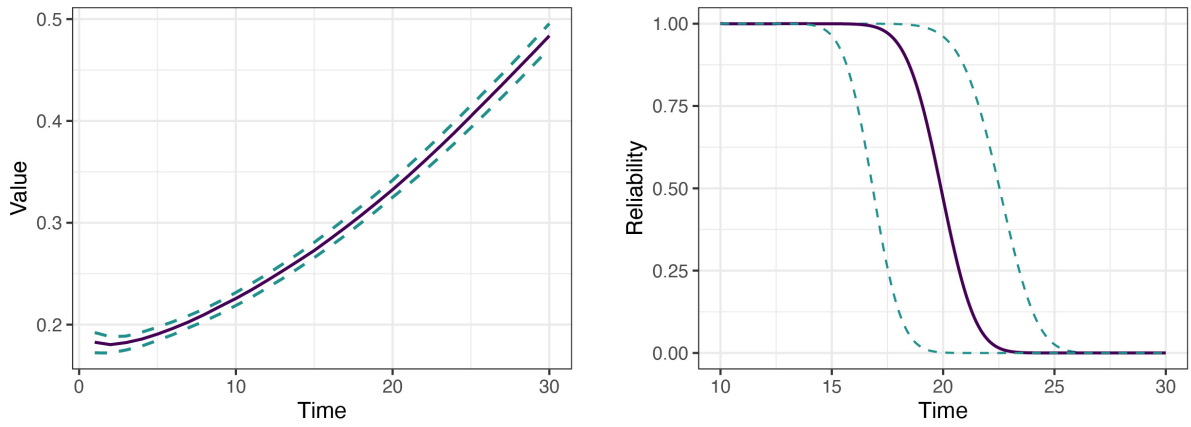
6(b), the estimated mean degradation path (purple line) and 90% confidence band (green dashed line) for two PC degradation paths (gray line) for PMB data are presented. The figure illustrates that the 90% confidence bands adequately encompass the actual degradation paths for both PCs. This observation substantiates the validity of the proposed model. Figure 7(a) shows the correlation coefficient estimation results between the two PC degradation paths. It can be seen from the figure that $\rho[Y_1(t), Y_2(t)]$ is minimized when the initial value is 0.172. As time increases, the correlation between the two PCs gradually increases, reaching 0.484 at $t = 30$. This implies that over time, the impact of magnetic degradation and common environmental effects on both PCs (braking torque and armature response time) is increasing relative to their individual degradation effects.



(a) Q-Q plots under scenario III model (b) Estimated mean degradation path

Figure 6: Summary of PMB data analysis results: Q-Q plots under scenario III model and the estimated mean degradation path.

Based on the degradation path fitting results, we conduct a thorough analysis of the system's reliability by applying (13) over a specified period. To demonstrate the predictive performance of the proposed model for failure times, we set low failure thresholds for two PCs: 3.087 and 5.733, ensuring that system failures occurred before the final observation period. Figure 7(b) illustrates the reliability curves of the system's failure time. Additionally, the pointwise 90% confidence band for the system reliability, determined through the parametric bootstrap method, is also depicted. Given the parameter estimations and failure thresholds, we compute the system's MTTF. The actual average failure time is determined to be 19, while the predicted average MTTF value is 19.647. The average RRMSE between the actual and predicted values is 0.034. This comprehensive evaluation suggests that our



(a) Correlation coefficients (b) Reliability curves

Figure 7: Correlation coefficients and reliability curves for PMB data.

prediction performance is relatively accurate. The close alignment between predicted and actual results provides a solid foundation for future reliability maintenance strategies.

6.2. Fatigue crack-size data

For broader applications, we utilize a subset of fatigue crack size data that describes the progression of cracks in an alloy over time. Similar to the approach by Wang et al. (2014) and Fang et al. (2022), the dataset is partitioned into three segments, forming a 3-dimensional degradation process with a sample size of six test units (i.e., $K = 3$, and $n = 6$). The initial crack size is 0.9; for convenience, we subtract 0.9 from all data and perform a scale transformation by multiplying by 10. The degradation paths are illustrated in Figure 8, with nine observations conducted on each PC (i.e., $m = 9$). The original dataset is initially presented in Meeker et al. (2022), and the processed data can be found in Appendix H of Fang et al. (2022).

We apply the proposed model to this dataset. Convergence of the EM algorithm is monitored using trace plots of parameter estimates across iterations, as depicted in Figure S5 within Supplementary Section S5. The parameter estimation results for four scenarios are summarized in Table 4. It can be observed that Scenario IV has the lowest AIC, leading us to select it as the optimal model for further analysis. Next, we conduct a GOF assessment for the model. Examination of the Q-Q plots in Figure 9(a) reveals that the majority of data points align closely with the theoretical straight line, although a few are notably distant. These outliers may be attributed to uncertainties in the mean estimator (Fang et al., 2022).

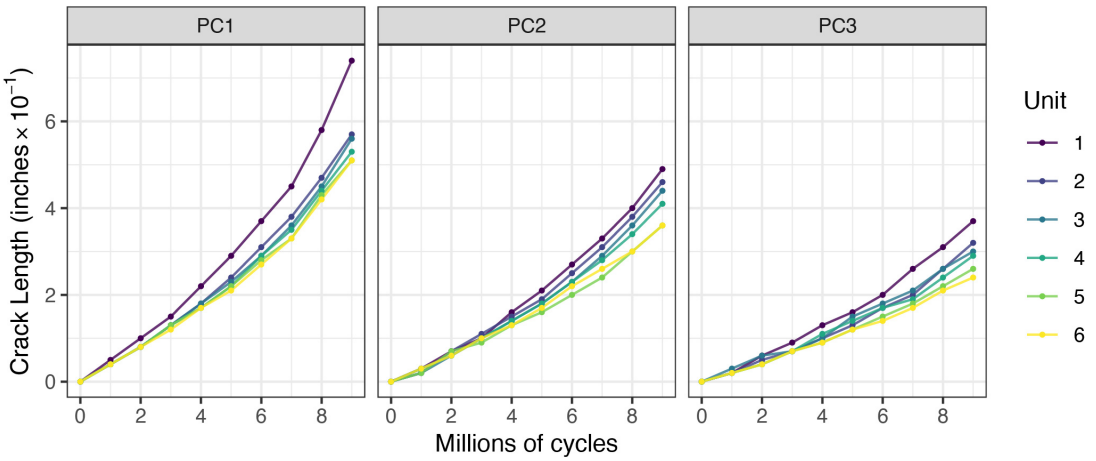


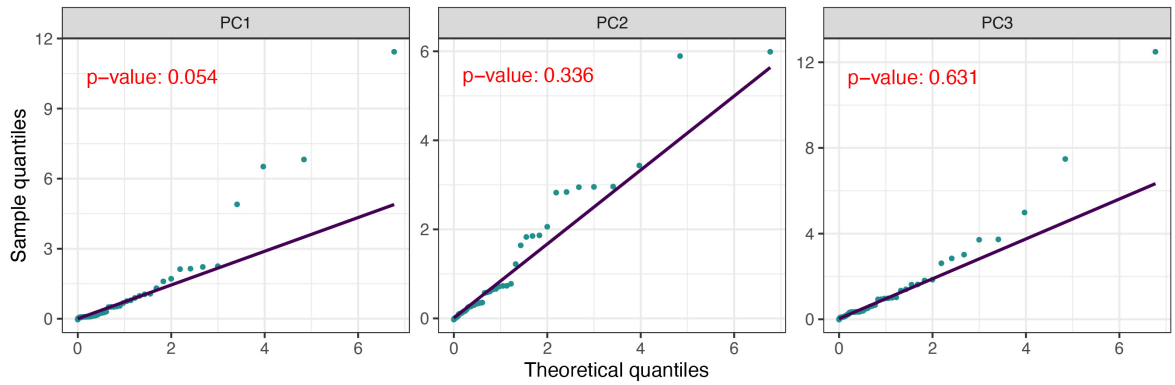
Figure 8: Degradation paths for fatigue crack-size growth data.

Furthermore, the KS test yields p-values greater than 0.05 for all three PCs, indicating that the proposed model achieved a reasonable fit to the dataset. The likelihood ratio test is performed based on the optimal model. The p-value is 0.385, which exceeds 0.05, indicating no significant differences in the γ values across the PCs. This suggests that our model is suitable for this dataset. Figure 9(b) presents the average path fitting based on the optimal model and the corresponding CIs for the degradation paths of the three PCs. From the figure, it is evident that the 90% confidence bands appropriately cover the actual degradation paths for all three PCs, demonstrating the correctness of the proposed model.

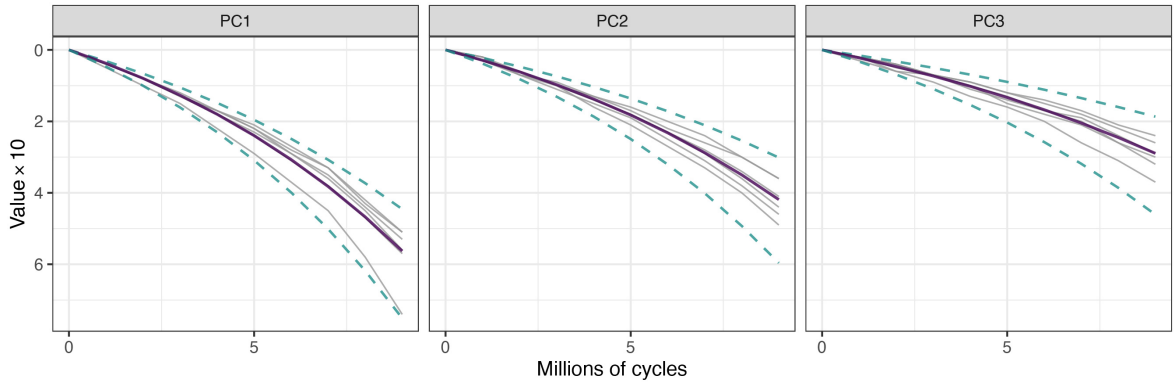
Table 4: Parameter point estimates regarding the fatigue crack-size data.

Model	Scen.	α_0	α_1	α_2	α_3	β_1	β_2	β_3	γ	AIC
Dependent	I	1.178	1.327	1.332	0.736	0.796	0.415	0.249	4.836	-717.838
	II	0.957	0.155	0.161	0.162	9.828	6.094	3.429	6.648	-804.636
	III	0.249	1.201	1.153	0.946	1.999	1.490	1.310	6.412	-822.131
	IV	0.067	0.119	0.111	0.090	19.683	16.286	15.236	6.789	**-1410.667**
Independent	Power	-	1.479	1.359	1.206	1.129	1.119	1.107	5.254	-221.197
	Log-linear	-	0.126	0.105	0.081	17.880	17.698	17.978	6.602	-281.749

To carry out the reliability analysis, the estimated parameters are plugged into (13). Figure 10 gives the PDF and reliability function of the system's failure time, where we assume three PCs have relatively low failure thresholds: 3.15, 2.45, and 1.40 (in inches



(a) Q-Q plots under scenario IV model



(b) Estimated mean degradation path

Figure 9: Summary of fatigue crack-size data analysis results: Q-Q plots under scenario IV and the estimated mean degradation path.

$\times 10^{-1}$), respectively. In addition, the corresponding 90% confidence bands are also provided. Based on the failure thresholds, the actual average failure time is determined to be 5.833, while the predicted average MTTF value is 5.087. The calculated average RRMSE between the actual and predicted values is 0.147.

7. Conclusion

In this study, we introduce a novel multivariate reparameterized IG process model. It adopts the common-effect approach, wherein the degradation of the marginal process is decomposed into two independent components—an individual intrinsic part and a separate common-effect part that applies to all the marginals uniformly. In a physical sense, this common-effect term has a meaningful interpretation, suggesting that the interdependence among PCs stems from a shared failure mechanism or is influenced by common operating

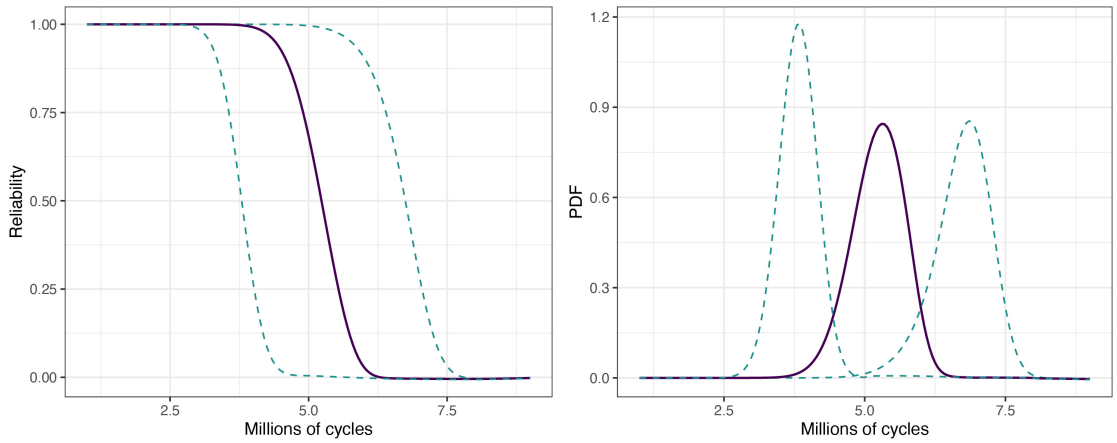


Figure 10: Reliability function and PDF for fatigue crack-size data.

environmental conditions. Statistically, the quantity of model parameters scales in a linear fashion with the number of PCs, offering flexibility for expanding the model into high-dimensional scenarios. To facilitate applicability, we provide the GL quadrature method to approximate the system reliability function, the EM algorithm for point estimates, as well as the parametric bootstrap approach for interval estimates. Through the simulation and case studies, the efficacy of our proposed methodology has been demonstrated. The main findings demonstrate that: a) Increasing the sample size significantly enhances the accuracy of point estimates, leading to more precise parameter values; b) Interval estimates based on the Bootstrap method can achieve reasonable CPs across all combinations; c) Failing to account for the interdependence among PCs within the model introduces substantial biases in the estimation of the MTTFs.

In practice, PC-specific effects in certain systems may display non-monotonic behavior. Examples include the capacity of lithium batteries (Zhang et al., 2023) and the light intensity of light emitting diodes (Ye et al., 2013). Given these observations, we can consider extending the proposed model. Specifically, $X_k(t)$ could be modeled by a Wiener process, and the common effect by an rIG process. This configuration leads to the degradation process of the k-th PC being represented by the following convolution integral:

$$f_{Y_k}(y_k) = f_{rIG}(z) * f_W(y_k) = \int_0^{y_k} f_{rIG}(z)\, f_W(y_k - z)\, \mathrm{d}z,$$

where $f_W(\cdot)$ represents the PDF for degradation values derived from the Wiener process. Due to the absence of an explicit solution for the distribution of Y_k, approximations such as the GL quadrature method become essential. Addressing the challenges in parameter estimation,

the complexity of the model increases as the log-likelihood function, represented in (17), may not always be expressed in closed form. This complexity makes it difficult to directly apply methods like the EM algorithm. To solve these complexities, we may employ likelihood-free methods for statistical inference, such as approximate Bayesian computation (Blum, 2010) or synthetic likelihood approaches (Frazier et al., 2023). Moreover, the framework presented can be further extended to applications in system design, maintenance, and optimization.

Funding

The research was supported by National Natural Science Foundation of China (12171432, 12271168, 72201242), Zhejiang Provincial Natural Science Foundation of China (LZ24A010002, LQ22G010003), the 111 Project of China (No. B14019), the Fundamental Research Funds for the Provincial Universities of Zhejiang, the China Scholarship Council program (202308330508), the Summit Advancement Disciplines of Zhejiang Province (Zhejiang Gongshang University—Statistics), and Collaborative Innovation Center of Statistical Data Engineering Technology & Application.

Supplementary Materials

Supplementary document: (a) proof of propositions; (b) technical details of the integral approximation; (c) technical details of the EM algorithm, including conditional expectations in the E-step, first-order partial derivatives of Q-function; (d) additional results of simulation and case studies. **Source Codes**: R codes for the GL method and EM algorithm are provided.

Disclosure of interest

No potential competing interest was reported by the authors.

About the authors

Liangliang Zhuang is currently pursuing a Ph.D. in Statistics at Zhejiang Gongshang University, China, and is also a visiting Ph.D. student at the National University of Singapore, Singapore. His research interests include reliability analysis, predictive maintenance, and optimization.

Ancha Xu received the B.S. degree in Mathematics from the Central China Normal University, Wuhan, China, in 2006, and the Ph.D. degree in Probability Theory and Mathematical Statistics from East China Normal University, Shanghai, China, in 2011. From 2011 to 2020, he was an Associate Professor of Statistics at Wenzhou University, Wenzhou, China. Since 2021, he has been a Professor at the Department of Statistics, Zhejiang Gongshang University, Hangzhou, China. He is the author of more than 50 articles. His research interests include Bayesian computation, data-driven degradation modeling, and industrial statistics.

Guanqi (Kevin) Fang is an Associate Professor in School of Statistics and Mathematics at Zhejiang Gongshang University, Hangzhou, China. He obtained his Ph.D. degree in Industrial Engineering and a concurrent M.S. degree in Statistics from Arizona State University. His research interests include reliability modeling, statistical inference, and multivariate data analysis, among others.

Yincai Tang is a professor of statistics in the College of Statistics, East China Normal University, Shanghai, China. He received his PhD degree from East China Normal University. His professional publications and research interests have focused on lifetime data analysis, degradation data analysis, big data analysis and Bayesian inference.

Data Availability Statement

The PMB degradation data that support the findings of this study are available on request from the corresponding author, and the fatigue crack-size data that support the findings of this study are openly available in Meeker et al. (2022).

References

Babolian, E., MasjedJamei, M., Eslahchi, M., 2005. On numerical improvement of Gauss–Legendre quadrature rules. Applied Mathematics and Computation 160, 779–789. doi:10.1016/j.amc.2003.11.031.

Barui, S., Mitra, D., Balakrishnan, N., 2024. Flexible modelling of a bivariate degradation process with a shared frailty and an application to fatigue crack data. Reliability Engineering & System Safety 242, 109722. doi:10.1016/j.ress.2023.109722.

Blum, M.G., 2010. Approximate Bayesian computation: a nonparametric perspective. Journal of the American Statistical Association 105, 1178–1187. doi:10.1198/jasa.2010.tm09448.

Chen, P., Ye, Z.S., 2018. Uncertainty quantification for monotone stochastic degradation models. Journal of Quality Technology 50, 207–219. doi:10.1080/00224065.2018.1436839.

Dong, Q., Cui, L., Si, S., 2020. Reliability and availability analysis of stochastic degradation systems based on bivariate Wiener processes. Applied Mathematical Modelling 79, 414–433. doi:10.1016/j.apm.2019.10.044.

Fang, G., Pan, R., 2021. On multivariate copula modeling of dependent degradation processes. Computers & Industrial Engineering 159, 107450. doi:10.1016/j.cie.2021.107450.

Fang, G., Pan, R., 2023. A class of hierarchical multivariate Wiener processes for modeling dependent degradation data. Technometrics 66, 141–156. doi:10.1080/00401706.2023.2242413.

Fang, G., Pan, R., Wang, Y., 2022. Inverse Gaussian processes with correlated random effects for multivariate degradation modeling. European Journal of Operational Research 300, 1177–1193. doi:10.1016/j.ejor.2021.10.049.

Frazier, D.T., Nott, D.J., Drovandi, C., Kohn, R., 2023. Bayesian inference using synthetic likelihood: asymptotics and adjustments. Journal of the American Statistical Association 118, 2821–2832. doi:10.1080/01621459.2022.2086132.

Hajiha, M., Liu, X., Hong, Y., 2021. Degradation under dynamic operating conditions: Modeling, competing processes and applications. Journal of Quality Technology 53, 347–368. doi:10.1080/00224065.2020.1757390.

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, 370–384. doi:10.1080/00224065.2019.1611353.

Hong, Y., Zhang, M., Meeker, W.Q., 2018. Big data and reliability applications: The complexity dimension. Journal of Quality Technology 50, 135–149. doi:10.1080/00224065.2018.1438007.

Huynh, K.T., Vu, H.C., Nguyen, T.D., Ho, A.C., 2022. A predictive maintenance model for k-out-of-n:F continuously deteriorating systems subject to stochastic and economic dependencies. Reliability Engineering & System Safety 226, 108671. doi:10.1016/j.ress.2022.108671.

Jiang, D., Chen, T., Xie, J., Cui, W., Song, B., 2023. A mechanical system reliability degradation analysis and remaining life estimation method——With the example of an aircraft hatch lock mechanism. Reliability Engineering & System Safety 230, 108922. doi:10.1016/j.ress.2022.108922.

Kang, R., Gong, W., Chen, Y., 2020. Model-driven degradation modeling approaches: Investigation and review. Chinese Journal of Aeronautics 33, 1137–1153. doi:10.1016/j.cja.2019.12.006.

Kou, B., Chen, W., Jin, Y., 2021. A novel cage-secondary permanent magnet linear eddy current brake with wide speed range and its analytical model. IEEE Transactions on Industrial Electronics 69, 7130–7139. doi:10.1109/TIE.2021.3097603.

Lamu, D., Yan, R., 2024. Reliability estimation of s-out-of-k system with Kumaraswamy distribution based on partially constant stress accelerated life tests. Statistical Theory and Related Fields , 1–20doi:10.1080/24754269.2024.2359826.

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 295, 705–717. doi:10.1016/j.ejor.2021.03.010.

Lu, L., Wang, B., Hong, Y., Ye, Z., 2021. General path models for degradation data with multiple characteristics and covariates. Technometrics 63, 354–369. doi:10.1080/00401706.2020.1796814.

Meeker, W.Q., Escobar, L.A., Pascual, F.G., 2022. 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, 263–272. doi:10.1016/j.ejor.2012.05.011.

Pan, Z., Balakrishnan, N., 2011. Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. Reliability Engineering & System Safety 96, 949–957. doi:10.1016/j.ress.2011.03.014.

Peng, W., Li, Y.F., Mi, J., Yu, L., Huang, H.Z., 2016. Reliability of complex systems under dynamic conditions: A Bayesian multivariate degradation perspective. Reliability Engineering & System Safety 153, 75–87. doi:10.1016/j.ress.2016.04.005.

Shao, J., Wang, X., 2023. MLE with datasets from populations having shared parameters. Statistical Theory and Related Fields 7, 213–222. doi:10.1080/24754269.2023.2180185.

Shin, H.J., Choi, J.Y., Cho, H.W., Jang, S.M., 2013. Analytical torque calculations and experimental testing of permanent magnet axial eddy current brake. IEEE Transactions on Magnetics 49, 4152–4155. doi:10.1109/TMAG.2013.2250932.

Song, K., Cui, L., 2022. A common random effect induced bivariate gamma degradation process with application to remaining useful life prediction. Reliability Engineering & System Safety 219, 108200. doi:10.1016/j.ress.2021.108200.

Sun, F., Li, H., Cheng, Y., Liao, H., 2021. Reliability analysis for a system experiencing dependent degradation processes and random shocks based on a nonlinear Wiener process model. Reliability Engineering & System Safety 215, 107906. doi:10.1016/j.ress.2021.107906.

Swarztrauber, P.N., 2003. On computing the points and weights for Gauss–Legendre quadrature. SIAM Journal on Scientific Computing 24, 945–954. doi:10.1137/S1064827500379690.

Wang, X., Balakrishnan, N., Guo, B., 2014. Residual life estimation based on nonlinear-multivariate Wiener processes. Journal of Statistical Computation and Simulation 85, 1742–1764. doi:10.1080/00949655.2014.898765.

Wang, X., Xu, D., 2010. An inverse Gaussian process model for degradation data. Technometrics 52, 188–197. doi:10.1198/TECH.2009.08197.

Xu, A., Fang, G., Zhuang, L., Gu, C., 2024. A multivariate Student-t process model for dependent tail-weighted degradation data. IISE Transactions doi:10.1080/24725854.2024.2389538.

Xu, A., Shen, L., Wang, B., Tang, Y., 2018. On modeling bivariate Wiener degradation process. IEEE Transactions on Reliability 67, 897–906. doi:10.1109/TR.2018.2791616.

Yan, B., Wang, H., Ma, X., 2023. Correlation-driven multivariate degradation modeling and RUL prediction based on Wiener process model. Quality and Reliability Engineering International 39, 3203–3229. doi:10.1002/qre.3105.

Ye, Z.S., Chen, N., 2014. The inverse Gaussian process as a degradation model. Technometrics 56, 302–311. doi:10.1080/00401706.2013.830074.

Ye, Z.S., Hu, Q., Yu, D., 2019. Strategic allocation of test units in an accelerated degradation test plan. Journal of Quality Technology 51, 64–80. doi:10.1080/00224065.2018.1545495.

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, 772–780. doi:10.1109/TR.2013.2284733.

Zhai, Q., Ye, Z.S., 2023. A multivariate stochastic degradation model for dependent performance characteristics. Technometrics 65, 315–327. doi:10.1080/00401706.2022.2157881.

Zhang, Y., Feng, F., Wang, S., Meng, J., Xie, J., Ling, R., Yin, H., Zhang, K., Chai, Y., 2023. Joint nonlinear-drift-driven Wiener process-Markov chain degradation switching model for adaptive online predicting lithium-ion battery remaining useful life. Applied Energy 341, 121043. doi:10.1016/j.apenergy.2023.121043.

Zhuang, L., Xu, A., Wang, Y., Tang, Y., 2024. Remaining useful life prediction for two-phase degradation model based on reparameterized inverse Gaussian process. European Journal of Operational Research 319, 877–890. doi:10.1016/j.ejor.2024.06.032.

Zhuang, X., Yu, T., Saraygord Afshari, S., Sun, Z., Song, K., Liang, X., 2021. Remaining useful life prediction of a mechanism considering wear correlation of multiple joints. Mechanical Systems and Signal Processing 149, 107328. doi:10.1016/j.ymssp.2020.107328.