



A Class of Hierarchical Multivariate Wiener Processes for Modeling Dependent Degradation Data

Guanqi Fang & Rong Pan

To cite this article: Guanqi Fang & Rong Pan (31 Jul 2023): A Class of Hierarchical Multivariate Wiener Processes for Modeling Dependent Degradation Data, Technometrics, DOI: [10.1080/00401706.2023.2242413](https://doi.org/10.1080/00401706.2023.2242413)

To link to this article: <https://doi.org/10.1080/00401706.2023.2242413>

 View supplementary material 

 Published online: 31 Jul 2023.

 Submit your article to this journal 

 Article views: 218

 View related articles 

 View Crossmark data 



A Class of Hierarchical Multivariate Wiener Processes for Modeling Dependent Degradation Data

Guanqi Fang ^{a,b} and Rong Pan ^c

^aCollaborative Innovation Center of Statistical Data Engineering Technology & Application, Zhejiang Gongshang University, Hangzhou, China; ^bSchool of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou, China; ^cSchool of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ

ABSTRACT

In engineering practice, many products exhibit multiple and dependent degrading performance characteristics (PCs). It is common to observe that these PCs' initial measurements are nonconstant and sometimes correlated with the subsequent degradation rate, which typically varies from one unit to another. To accommodate the unit-wise heterogeneity, PC-wise dependency, and "initiation-growth" correlation, this article proposes a broad class of multi-dimensional degradation models under a framework of hierarchical multivariate Wiener processes. These models incorporate dual multi-normally distributed random effects concerning the initial values and degradation rates. To infer model parameters, expectation-maximization (EM) algorithms and several tools for model validation and selection are developed. Various simulation studies are carried out to assess the performance of the inference method and to compare different models. Two case studies are conducted to demonstrate the applicability of the proposed methodology. The online supplementary materials of this article contain derivations of EM estimators, additional numerical results, and R codes.

ARTICLE HISTORY

Received August 2022
Accepted July 2023

KEYWORDS

Degradation data; EM algorithm; Hierarchical model; Multivariate degradation; Random effects; Wiener process

1. Introduction

1.1. Background

Degradation is a major cause of failure for many highly reliable consumer products and industrial systems, including inkjet printer heads (Weaver et al. 2013), Micro-ElectroMechanical Systems (Peng, Feng, and Coit 2009), and pipeline circuits (Tian, Liu, and Meeker 2019), among others. The overall quality of such products or systems is often evaluated using multiple degrading performance characteristics (PCs), and the stochastic nature of these degradation processes typically results in strongly correlated PCs. For instance, Si et al. (2018) introduced a three-dimensional deformation process leading to plate failure due to cracks on three local points. Gu et al. (2009) presented a multi-dimensional photodegradation process in polymeric materials. Meanwhile, an eye-catching observation of some degradation processes shows that the initial performance level is random; and this variability may be caused by unavoidable manufacturing defects or some unknown usage, such as burn-in screening (Shen et al. 2019). In addition, another interesting phenomenon is that the initial observations are sometimes correlated with the subsequent degradation rate. This type of correlation can be referred to as the "initiation-growth" correlation. For example, Ye, Hu, and Yu (2019) described infrared sensors where a unit with a higher initial noise level (i.e., the value of a PC) tends to possess a higher degradation rate. Therefore, to account for (a) the unit-wise heterogeneity, (b) the PC-wise dependency, and

perhaps (c) the initiation-growth correlation, a class of flexible multivariate degradation models is much desired. The example below further elaborates on these concerns and motivates our study.

1.2. A Motivating Example

A transceiver is a device that transmits and receives different signals and is widely used in fiber-optic systems. Affected by its operational environment, one of this product's important performance metrics, Receiver (Rx) Sensitivity, gradually degrades over time. This metric is defined as the minimum signal optical power level required at the receiver to achieve a certain Bit Error Ratio (BER) level. Figure 1 shows an experimental result of nine samples that underwent a degradation test. Each specimen was tested by capturing two types of light signals with different wavelengths, referred to as two separate channels. The Rx Sensitivity to these two channels is measured in decibels (dB) relative to the product design specification. Hence, each unit has two performance measurements, denoted as PC1 and PC2.

From Figure 1, diverse degradation paths can be observed among these samples, and the orderings of these paths are relatively consistent for the two PCs. They suggest unit-to-unit variation and process dependency. Moreover, the initial status of these samples doesn't hold at a constant level. This variability may result from unavoidable manufacturing variation or nonuniform raw materials. Then, the following questions would

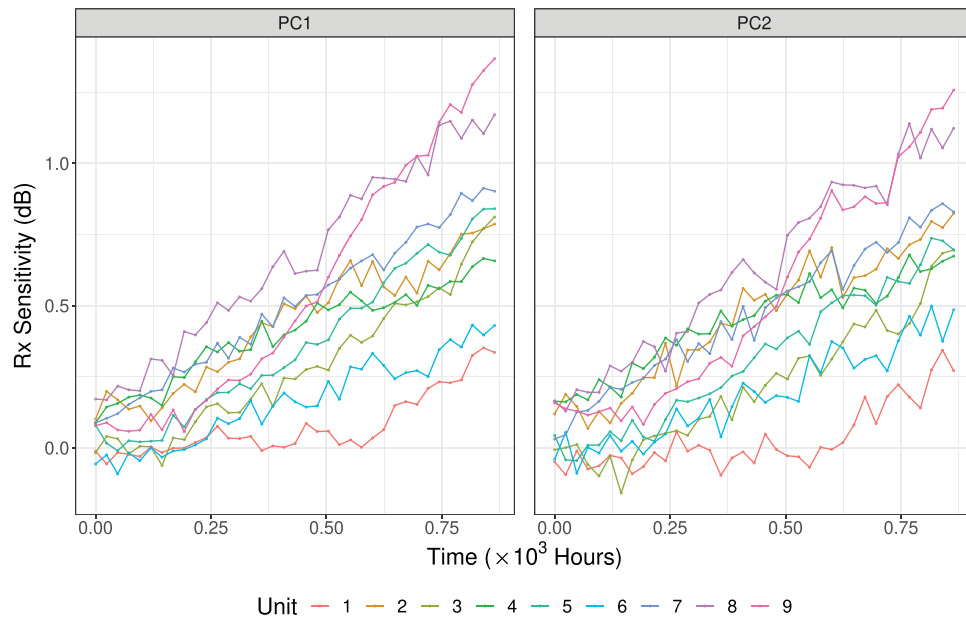


Figure 1. Degradation paths of transceivers.

naturally arise: Is there a proper model that can account for the unit-to-unit variation, process dependency, and randomness of the initial observations? Do the initial observations provide some partial information about the subsequent degradation rates? To address these inquiries, this article proposes a class of multivariate Wiener process-based degradation models.

1.3. Related Literature

The literature on degradation modeling is extensive, so this section only highlights a few select articles relevant to this article. For a univariate degradation process, a natural choice is to apply regression analysis such that the degradation measure and the measurement time are treated as a response variable and a regressor, respectively. This modeling technique, known as the general path model, has gained much attention since the early 1990s (Lu and Meeker 1993). Built upon the basic structure, some model variants addressing additional complexities, including the unit-to-unit variation (Bae and Kvam 2004; Fang, Rigdon, and Pan 2018) and multi-phase degradation (Bae et al. 2015; Wen, Wu, and Yuan 2017), have been developed. Another approach is to view the degradation path as a result of an accumulation of infinitesimal segments, leading to the stochastic process model. It encompasses three commonly used ones—the Wiener process (Wang 2010; Hong, Tan, and Ye 2020), the gamma process (Lawless and Crowder 2004; Ye et al. 2014), and the inverse Gaussian process (Ye and Chen 2014; Peng 2015). For a detailed comparison of these models, readers are referred to Ye and Xie (2015), Zhang et al. (2018), and Kang, Gong, and Chen (2020).

When extending to multi-dimensional scenarios, a key consideration is how to account for the dependency between different PCs. One research stream in the literature is the copula-based approach, which builds the dependency between degradation increments of two/several PCs through a particular copula

function. Some developments in this regard can be found in the articles by Fang and Pan (2021), Sun et al. (2020a), Fang, Pan, and Hong (2020), Palayangoda and Ng (2021), and Hu et al. (2021). Even though this approach is flexible to cover many types of tail dependency, it often suffers from missing clear physical interpretations and lacking desirable mathematical properties (Hong, Zhang, and Meeker 2018b). An alternative is to build multi-dimensional distributions by extending the aforementioned univariate models. For instance, Si et al. (2018) and Lu et al. (2021) constructed multivariate general path models to analyze the material deformation and the photodegradation process, respectively. Hong, Ye, and Ling (2018a) used a two-dimensional Wiener process to analyze the degradation behavior of different emerging contaminants. Sun, Ye, and Hong (2020b) further developed a multivariate Wiener process that could account for both the rig-layer and gauge-layer block effect, and Wang et al. (2020) proposed a similar model to analyze the impact of time-variant covariates and imperfect preventative maintenance actions. Under a multi-dimensional Wiener process framework, Hao et al. (2017) designed a Bayesian method to dynamically update the remaining useful life distribution of multistage manufacturing processes with the interaction between tool wear and production quality degradation. More recently, Song and Cui (2022) proposed a bivariate gamma process with a common random effect for remaining useful life prediction, and Yan, Wang, and Ma (2022) made use of the similar idea of assigning a common random-effects term to several Wiener processes. Zhai and Ye (2023) introduced a common stochastic time scale to model the dependence from the dynamic operating environment. Fang, Pan, and Wang (2022) built a multivariate inverse Gaussian process with correlated random effects.

Despite these previous studies, to the best of our knowledge, there is still a dearth of discussion on the multivariate degradation models that can jointly capture the three aspects of the variability described earlier, namely the unit-wise heterogeneity,

the PC-wise dependency, and the initiation-growth correlation. To fill the gap, in this article, we introduce a class of hierarchical multivariate Wiener process models that can be applied to a broad range of practical scenarios. The proposed methodology provides a significant enhancement to the existing family of multivariate degradation models.

1.4. Overview

The remainder of this article is organized as follows. Section 2.1 presents the formulation of the proposed multivariate Wiener process model, which incorporates structured and correlated random effects. Built upon this model, a number of theoretically tractable results (See Sections 2.2 and 2.3) are produced. They greatly facilitate model fitting and validation/selection (See Section 3). By simulation studies (See Section 4) and case studies (See Section 5), the effectiveness and flexibility of the proposed methodology are demonstrated. Finally, Section 6 concludes this article with a summary and discussion of the future study.

2. A Hierarchical Multivariate Wiener Process Model

2.1. Model Formulation

Let $Y_{ij}(t)$ denote the degradation measurement of process j , $j = 1, 2, \dots, p$, on unit i , $i = 1, 2, \dots, n$, at the elapsed time t , $t \geq 0$. It is assumed that the evolution of $Y_{ij}(t)$ is subject to a Wiener process with a random initial value $Y_{ij}(0)$ and a drift rate b_{ij} . Meanwhile, the multiple degradation processes corresponding to a unit's multiple PCs are statistically dependent. By letting $\mathbf{Y}_i(t) = (Y_{i1}(t), Y_{i2}(t), \dots, Y_{ip}(t))'$ be the vector of all process observations, the following hierarchical multivariate Wiener process model is proposed:

$$\begin{cases} \mathbf{Y}_i(t) = \mathbf{Y}_i(0) + t\mathbf{b}_i + \mathbf{D}^{1/2}\mathbf{B}_p(t) \\ (\mathbf{Y}_i(0)', \mathbf{b}_i')' \sim \mathcal{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \end{cases},$$

where $\mathbf{Y}_i(0) = (Y_{i1}(0), Y_{i2}(0), \dots, Y_{ip}(0))'$ and $\mathbf{b}_i = (b_{i1}, b_{i2}, \dots, b_{ip})'$ are the vector of the i th unit's initial degradation values and rates, respectively. The stochastic process $\mathbf{B}_p(t)$ is a p -dimensional standard Brownian motion and $\mathbf{D}$ is a diagonal matrix with the form $\mathbf{D} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$, where $\sigma_j > 0$ is the diffusion parameter. The joint vector $(\mathbf{Y}_i(0)', \mathbf{b}_i')'$ is assumed to follow a multivariate normal (MVN) distribution with the mean vector $\boldsymbol{\mu}$ and the variance-covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{B}_p(t)$ is regarded independently from both $\mathbf{Y}_i(0)$ and $\mathbf{b}_i$. A more concise model form, which is referred to as model M_0 , is given by

$$M_0 : \begin{cases} \mathbf{Y}_i(t) | (\mathbf{Y}_i(0)', \mathbf{b}_i')' \sim \mathcal{MVN}(\mathbf{Y}_i(0) + t\mathbf{b}_i, t\mathbf{D}) \\ \begin{pmatrix} \mathbf{Y}_i(0) \\ \mathbf{b}_i \end{pmatrix} \sim \mathcal{MVN}\left(\begin{bmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_a & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_b \end{bmatrix}\right), \end{cases}$$

where $\boldsymbol{\mu}$ is decomposed into two parts – $\boldsymbol{\mu}_a = (\mu_{a1}, \mu_{a2}, \dots, \mu_{ap})'$ and $\boldsymbol{\mu}_b = (\mu_{b1}, \mu_{b2}, \dots, \mu_{bp})'$ – with respect to $\mathbf{Y}_i(0)$ and $\mathbf{b}_i$, respectively. Correspondingly, $\boldsymbol{\Sigma}_a$ and $\boldsymbol{\Sigma}_b$ are the variance-covariance matrices for $\mathbf{Y}_i(0)$ and $\mathbf{b}_i$, respectively. By symmetry, $\boldsymbol{\Sigma}_{ab} = \boldsymbol{\Sigma}_{ba}'$ and they are the

cross-covariance matrices between $\mathbf{Y}_i(0)$ and $\mathbf{b}_i$. Specifically, $\boldsymbol{\Sigma}_a$, $\boldsymbol{\Sigma}_b$, and $\boldsymbol{\Sigma}_{ab}$ are defined as

$$\begin{aligned} \boldsymbol{\Sigma}_a &= \begin{pmatrix} \sigma_{a1}^2 & \sigma_{a12}^a & \cdots & \sigma_{a1p}^a \\ \sigma_{a12}^a & \sigma_{a2}^2 & \cdots & \sigma_{a2p}^a \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{a1p}^a & \sigma_{a2p}^a & \cdots & \sigma_{ap}^2 \end{pmatrix}, \\ \boldsymbol{\Sigma}_b &= \begin{pmatrix} \sigma_{b1}^2 & \sigma_{b12}^b & \cdots & \sigma_{b1p}^b \\ \sigma_{b12}^b & \sigma_{b2}^2 & \cdots & \sigma_{b2p}^b \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{b1p}^b & \sigma_{b2p}^b & \cdots & \sigma_{bp}^2 \end{pmatrix}, \text{ and} \\ \boldsymbol{\Sigma}_{ab} &= \begin{pmatrix} \sigma_1^{ab} & \sigma_{12}^{ab} & \cdots & \sigma_{1p}^{ab} \\ \sigma_{21}^{ab} & \sigma_2^{ab} & \cdots & \sigma_{2p}^{ab} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}^{ab} & \sigma_{p2}^{ab} & \cdots & \sigma_p^{ab} \end{pmatrix}. \end{aligned}$$

In these matrices, σ_{aj} and σ_{bj} are the standard deviation of $Y_{ij}(0)$ and b_{ij} , respectively. The covariances of the initial degradation value and the rate, between the j th and j' th PC, $\forall 1 \leq j < j' \leq p$, are given by $\sigma_{jj'}^a = \rho_{jj'}^a \sigma_{aj} \sigma_{aj'}$ and $\sigma_{jj'}^b = \rho_{jj'}^b \sigma_{bj} \sigma_{bj'}$, respectively. The covariance between $Y_{ij}(0)$ and b_{ij} , $\forall j = 1, 2, \dots, p$, is $\sigma_j^{ab} = \rho_j^{ab} \sigma_{aj} \sigma_{bj}$ and the cross-covariance between $Y_{ij}(0)$ and $b_{ij'}$, $\forall 1 \leq j \neq j' \leq p$, is $\sigma_{jj'}^{ab} = \rho_{jj'}^{ab} \sigma_{aj} \sigma_{bj'}$. Each of these Pearson correlation coefficients – $\rho_{jj'}^a$, $\rho_{jj'}^b$, ρ_j^{ab} , and $\rho_{jj'}^{ab}$ – is within the range of $[-1, 1]$. Note that $\boldsymbol{\Sigma}_{ab}$ is generally a dense matrix, but it can be reduced to a sparse one if most or even all elements in $\boldsymbol{\Sigma}_{ab}$ are 0's, such as a null matrix or a diagonal matrix. These circumstances correspond to the special model variants to be discussed later. In this article, we denote all the unknown parameters by a vector $\boldsymbol{\theta} = (\boldsymbol{\mu}_a', \boldsymbol{\mu}_b', \boldsymbol{\sigma}_a^{2'}, \boldsymbol{\rho}_a', \boldsymbol{\sigma}_b^{2'}, \boldsymbol{\rho}_b', \boldsymbol{\rho}_{ab}', \boldsymbol{\sigma}^{2'})'$, where $\boldsymbol{\sigma}_a^2 = (\sigma_{a1}^2, \sigma_{a2}^2, \dots, \sigma_{ap}^2)'$, $\boldsymbol{\rho}_a = (\rho_{12}^a, \rho_{13}^a, \dots, \rho_{p-1,p}^a)'$, $\boldsymbol{\sigma}_b^2 = (\sigma_{b1}^2, \sigma_{b2}^2, \dots, \sigma_{bp}^2)'$, $\boldsymbol{\rho}_b = (\rho_{12}^b, \rho_{13}^b, \dots, \rho_{p-1,p}^b)'$, $\boldsymbol{\rho}_{ab} = (\rho_1^{ab}, \rho_2^{ab}, \dots, \rho_p^{ab})'$, $\boldsymbol{\rho}_{ab} = (\rho_{12}^{ab}, \rho_{13}^{ab}, \dots, \rho_{p-1,p}^{ab})'$, and $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)'$. In short, an abbreviated representation is $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{D}\} = \{\boldsymbol{\mu}_a, \boldsymbol{\mu}_b, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b, \boldsymbol{\Sigma}_{ab}, \mathbf{D}\}$. For notational convenience, the preceding model assumes the original time scale t , implying linear degradation paths. A more general representation is $\Lambda(t)$, a transformation function of t . Choices for $\Lambda(\cdot)$ include the power law function and exponential law function (Whitmore and Schenkelberg 1997). Later, we will illustrate how to handle this time transformation. In the case of positive degradation values and rates, a conventional assumption of negligible nonpositive $Y_{ij}(0)$ and b_{ij} is made (Lu and Meeker 1993; Peng 2015).

Apparently, the proposed model creates a hierarchical structure for multiple Wiener processes, and this structure provides several meaningful interpretations. First, a well-defined multivariate Wiener process with unit-to-unit variation is constructed. This unit-wise heterogeneity is engineered into the model by incorporating the randomness of initial levels and degradation rates. Conditioning on the random effects, the property of mutually independent PCs is achieved. This nice property would largely facilitate the characterization of lifetime distribution and parameter estimation, as illustrated in later

Table 1. Structure of random effects of model M_0 and its various special variants.

Model	Variation sources	
	$Y_i(0)$	b_i
M_0	Fixed effect	$\begin{pmatrix} Y_i(0) \\ b_i \end{pmatrix} \sim \mathcal{MVN}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \begin{bmatrix} \Sigma_a & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_b \end{bmatrix}\right)$
M_1		$\begin{pmatrix} Y_i(0) \\ b_i \end{pmatrix} \sim \mathcal{MVN}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \begin{bmatrix} \Sigma_a & 0 \\ 0 & \Sigma_b \end{bmatrix}\right)$
M_2		$b_i \sim \mathcal{MVN}(\mu_b, \Sigma_b)$
M_3		$\begin{pmatrix} Y_i(0) \\ b_i \end{pmatrix} \sim \mathcal{MVN}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \begin{bmatrix} \text{diag}(\sigma_a^2) & \text{diag}(\sigma_{ab}) \\ \text{diag}(\sigma_{ab}) & \text{diag}(\sigma_b^2) \end{bmatrix}\right)$
M_4		$\begin{pmatrix} Y_i(0) \\ b_i \end{pmatrix} \sim \mathcal{MVN}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \begin{bmatrix} \sigma_a^2(\mu_a \otimes \mu_a) & \rho_{ab}\sigma_a\sigma_b \text{diag}(\mu_a \circ \mu_b) \\ \rho_{ab}\sigma_a\sigma_b \text{diag}(\mu_a \circ \mu_b) & \sigma_b^2(\mu_b \otimes \mu_b) \end{bmatrix}\right)$

Table 2. Characteristics of models M_0 – M_4 .

Model	Random initial levels	Random degradation rates	Initiation-growth correlation	Model attribute
M_0	✓	✓	✓	Generalized form
M_1	✓	✓	✗	Initiation-growth independence
M_2	✗	✓	✗	Fixed initial levels
M_3	✓($\rho_{jj'}^a = 0$)	✓($\rho_{jj'}^b = 0$)	✓	PC-wise independence
M_4	✓($\rho_{jj'}^a = 1$)	✓($\rho_{jj'}^b = 1$)	✓	Common random effects

sections. Second, any possible dependency that remains in PCs is accounted for by the unobserved random-effects terms – $Y_i(0)$ and b_i – modeled by a MVN distribution. This implies that the PC-wise dependency originates from the correlation among intrinsic degradation mechanisms (Fang, Pan, and Wang 2022). Lastly, it is noted that another type of dependency—the correlation between the initial degradation level and the degradation rate (i.e., the initiation-growth correlation), often neglected but ubiquitous in engineering practice (Ye, Hu, and Yu 2019), is taken into account by assuming the cross-covariance component in the MVN distribution.

The proposed model M_0 is designed to be sufficiently general to cover four special model variants, referred to as models M_1 , M_2 , M_3 , and M_4 in Table 1. These variants are produced by varying the structure of the two variation sources – $Y_i(0)$ and b_i . Specifically, in model M_1 , the cross-covariance matrix Σ_{ab} is a null matrix, which indicates that the initial levels and degradation rates behave independently. It can be seen as a multivariate extension of the model proposed by Xiao and Ye (2016). Model M_2 assumes that $Y_i(0)$ is fixed, resulting in a multivariate Wiener process with random degradation rates only. As an example, in Lu et al. (2021), an identical zero-damage level across all test units at the beginning is present for the photodegradation process. This model extends the random-effects Wiener process model proposed by Li, Pan, and Chen (2015) to the multivariate setting. Next, when the off-diagonal elements in Σ_a , Σ_b , and Σ_{ab} are all left out, the resulting matrices become diagonal ones as indicated by model M_3 . Consequently, the model consists of p independent sub-models with the j th one defined as follows:

$$M_3 : \begin{cases} Y_{ij}(t) = Y_{ij}(0) + b_{ij}t + \sigma_j \mathcal{B}(t) \\ \begin{pmatrix} Y_{ij}(0) \\ b_{ij} \end{pmatrix} \sim \mathcal{BVN}\left(\begin{bmatrix} \mu_{aj} \\ \mu_{bj} \end{bmatrix}, \begin{bmatrix} \sigma_{aj}^2 & \rho_j^{ab}\sigma_{aj}\sigma_{bj} \\ \rho_j^{ab}\sigma_{aj}\sigma_{bj} & \sigma_{bj}^2 \end{bmatrix}\right) \end{cases}$$

Here, $Y_{ij}(0)$ and b_{ij} are subject to a bivariate normal (BVN) distribution. Lastly, by denoting a random vector $(a_i, b_i)'$ and letting $\mu_{aj}a_i$ and $\mu_{bj}b_i$ be the initial levels and degradation rates

of the j th PC, respectively, we define the following j th sub-model:

$$M_4 : \begin{cases} Y_{ij}(t) = \mu_{aj}a_i + \mu_{bj}b_i t + \sigma_j \mathcal{B}(t) \\ \begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim \mathcal{BVN}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \rho_{ab}\sigma_a\sigma_b \\ \rho_{ab}\sigma_a\sigma_b & \sigma_b^2 \end{bmatrix}\right) \end{cases}$$

Since these PCs share the same stochastic terms (i.e., a_i and b_i), this model essentially mimics a modeling idea known as common random effects, which has been discussed in numerous previous publications, including Xu et al. (2018), Sun, Ye, and Hong (2020b), Song and Cui (2022), and Yan, Wang, and Ma (2022). If denoting $Y_i(0) = (\mu_{a1}a_i, \mu_{a2}a_i, \dots, \mu_{ap}a_i)'$ and $b_i = (\mu_{b1}b_i, \mu_{b2}b_i, \dots, \mu_{bp}b_i)'$, the structure of random effects in M_4 is actually a special case of our proposed modeling framework, as shown in Table 1. Here, $\circ$ and $\otimes$ denote the Hadamard and outer product, respectively. Note that models M_3 and M_4 demonstrate the two extreme scenarios of PC-wise dependency, where $\rho_{jj'}^a = \rho_{jj'}^b = 0$ for the former and $\rho_{jj'}^a = \rho_{jj'}^b = 1, \forall 1 \leq j < j' \leq p$, for the latter.

Table 2 summarizes the characteristics of these models (M_0 – M_4) from three perspectives—random initial levels, random degradation rates, and the initiation-growth correlation. In the table, a check mark indicates that a particular variability feature is considered, while a cross mark indicates its absence. It is evident that models M_1 – M_4 represent various reduced forms of M_0 . They capture several practical scenarios, including initiation-growth independence, fixed initial levels, PC-wise independence, and common random effects. The generalized form, M_0 , is essentially a broad class of these multivariate degradation models.

2.2. Theoretical Properties

In model M_0 , one can observe that in a Bayesian manner, the joint MVN distribution governing $Y_i(0)$ and b_i serves as the conjugate distribution for multi-normally distributed data implied by the Wiener process. This assumption of conjugacy

is employed for technical convenience since it releases beneficial theoretical properties with closed-form expressions for the marginal, joint, and conditional functions. Moreover, it allows for the specification of conditional distributions for future degradation predictions given existing observations.

First, under model M_0 , the marginal $(Y_{ij}(0), b_{ij})'$ is subject to a BVN distribution, that is,

$$\begin{pmatrix} Y_{ij}(0) \\ b_{ij} \end{pmatrix} \sim \mathcal{BVN} \left(\begin{bmatrix} \mu_{aj} \\ \mu_{bj} \end{bmatrix}, \begin{bmatrix} \sigma_{aj}^2 & \rho_j^{ab} \sigma_{aj} \sigma_{bj} \\ \rho_j^{ab} \sigma_{aj} \sigma_{bj} & \sigma_{bj}^2 \end{bmatrix} \right).$$

Then, it is easy to see that the marginal process $Y_{ij}(t)$ after integrating out $(Y_{ij}(0), b_{ij})'$ is subject to a normal distribution such that

$$Y_{ij}(t) \sim \mathcal{N}(\mu_{aj} + t\mu_{bj}, \sigma_{aj}^2 + 2t\rho_j^{ab}\sigma_{aj}\sigma_{bj} + t^2\sigma_{bj}^2 + t\sigma_j^2). \quad (1)$$

Similarly, the unconditional joint distribution of $Y_i(t)$ is

$$Y_i(t) \sim \mathcal{MVN}(\mu_a + t\mu_b, \Sigma_a + t\Sigma_{ab} + t\Sigma_{ba} + t^2\Sigma_b + tD). \quad (2)$$

The derivation of (1) and (2) is based on the basic property of an affine transformation applied to a multi-normally distributed random vector (See chap. 3.10.2 of Rencher and Christensen 2012). Now, suppose all the individual PCs are divided into two mutually exclusive subsets— $Y_{i\mathbb{A}}(t)$ and $Y_{i\mathbb{B}}(t)$, in which the collections of indices are denoted by $\mathbb{A} = \{j : Y_{ij}(t) \in Y_{i\mathbb{A}}(t)\}$ and $\mathbb{B} = \{j : Y_{ij}(t) \in Y_{i\mathbb{B}}(t)\}$. Then, if denoting $\mu_t = \mu_a + t\mu_b$ and $\Sigma_t = \Sigma_a + t\Sigma_{ab} + t\Sigma_{ba} + t^2\Sigma_b + tD$, the conditional distribution of $Y_{i\mathbb{A}}(t)$ given $Y_{i\mathbb{B}}(t) = y_{i\mathbb{B}}(t)$ satisfies that

$$Y_{i\mathbb{A}}(t)|Y_{i\mathbb{B}}(t) = y_{i\mathbb{B}}(t) \sim \mathcal{MVN} \left(\mu_{t\mathbb{A}} + \Sigma_{t\mathbb{A}\mathbb{B}} \Sigma_{t\mathbb{B}}^{-1} (y_{i\mathbb{B}}(t) - \mu_{t\mathbb{B}}), \Sigma_{t\mathbb{A}} - \Sigma_{t\mathbb{A}\mathbb{B}} \Sigma_{t\mathbb{B}}^{-1} \Sigma_{t\mathbb{B}\mathbb{A}} \right), \quad (3)$$

where $\mu_{t\mathbb{A}}$ and $\mu_{t\mathbb{B}}$ are the partitioned vectors from μ_t for $\mathbb{A}$ and $\mathbb{B}$, respectively. Correspondingly, partitioned from Σ_t , $\Sigma_{t\mathbb{A}}$ and $\Sigma_{t\mathbb{B}}$ are the resulting block matrices, along with the off-diagonal cross-covariance matrices $\Sigma_{t\mathbb{A}\mathbb{B}} = \Sigma'_{t\mathbb{B}\mathbb{A}}$.

Furthermore, provided that a measurement is taken at certain time t , $t > 0$, it produces the degradation measure $y_i(t) = (y_{i1}(t), y_{i2}(t), \dots, y_{ip}(t))'$. Combining with the initial measurement $y_i(0) = (y_{i1}(0), y_{i2}(0), \dots, y_{ip}(0))'$, we denote the vector of degradation increments $y_i(t) - y_i(0)$ for this duration from 0 to t by $\tilde{y}_i(t, 0)$. (Remark: In this article, notations related to degradation measurements, such as $y_{ij}(t)$ and $y_i(t)$, are expressed in normal mathematical fonts. The symbols, including $\tilde{y}_i(t, 0)$, in sans-serif mathematical fonts with the tilde character on the top represent a variety of notations associated with degradation increments.)

Evidently, $y_i(0)$ is just the realization of $Y_i(0)$, that is

$$Y_i(0) \sim \mathcal{MVN}(\mu_a, \Sigma_a). \quad (4)$$

And $\tilde{y}_i(t, 0)$ is the observed outcome of the random vector $t\mathbf{b}_i + D^{1/2}\mathcal{B}_p(t)$, that is,

$$\tilde{Y}_i(t, 0) \sim \mathcal{MVN}(t\mu_b, t^2\Sigma_b + tD). \quad (5)$$

Hence, $(Y_i(0)', b_i', \tilde{Y}_i(t, 0)')$ is jointly subject to a MVN distribution such that

$$\begin{bmatrix} Y_i(0) \\ b_i \\ \tilde{Y}_i(t, 0) \end{bmatrix} \sim \mathcal{MVN} \left(\begin{bmatrix} \mu_a \\ \mu_b \\ t\mu_b \end{bmatrix}, \begin{bmatrix} \Sigma_a & \Sigma_{ab} & t\Sigma_{ab} \\ -\Sigma_{ba} & \Sigma_b & t\Sigma_b \\ t\Sigma_{ba} & t\Sigma_b & t^2\Sigma_b + tD \end{bmatrix} \right).$$

Then, the conditional (posterior) distribution of b_i given that $Y_i(0) = y_i(0)$ is

$$b_i|y_i(0) \sim \mathcal{MVN}(\mu_{b_i|y_i(0)}, \Sigma_{b_i|y_i(0)}), \quad (6)$$

where

$$\begin{aligned} \mu_{b_i|y_i(0)} &= \mu_b + \Sigma_{ba} \Sigma_a^{-1} (y_i(0) - \mu_a) \text{ and} \\ \Sigma_{b_i|y_i(0)} &= \Sigma_b - \Sigma_{ba} \Sigma_a^{-1} \Sigma_{ab}. \end{aligned}$$

And the conditional (posterior) distribution of b_i given that $\tilde{Y}_i(t, 0) = \tilde{y}_i(t, 0)$ is

$$b_i|\tilde{y}_i(t, 0) \sim \mathcal{MVN}(\mu_{b_i|\tilde{y}_i(t, 0)}, \Sigma_{b_i|\tilde{y}_i(t, 0)}), \quad (7)$$

where

$$\begin{aligned} \mu_{b_i|\tilde{y}_i(t, 0)} &= \mu_b + \Sigma_b (\Sigma_b + D/t)^{-1} (\tilde{y}_i(t, 0)/t - \mu_b) \\ &= (\Sigma_b^{-1} + tD^{-1})^{-1} (\Sigma_b^{-1} \mu_b + D^{-1} \tilde{y}_i(t, 0)) \text{ and} \\ \Sigma_{b_i|\tilde{y}_i(t, 0)} &= \Sigma_b - \Sigma_b (\Sigma_b + D/t)^{-1} \Sigma_b \\ &= (\Sigma_b^{-1} + tD^{-1})^{-1}. \end{aligned}$$

Note that the first line in the equation for $\mu_{b_i|\tilde{y}_i(t, 0)}$ or $\Sigma_{b_i|\tilde{y}_i(t, 0)}$ directly results from the conditional mean or variance of a MVN distribution. The equivalent expression defined in the second line is obtained by applying the Bayes' rule. The derivation is given in Supplementary Section A. The latter representation mirrors the result of the univariate conditional (posterior) distribution of b_{ij} given $\tilde{y}_{ij}(t, 0)$ as presented by eq. (14) in Li, Pan, and Chen (2015).

If $y_i(0)$ and $\tilde{y}_i(t, 0)$ are both known, the conditional (posterior) distribution of b_i is

$$b_i|y_i(0), \tilde{y}_i(t, 0) \sim \mathcal{MVN}(\mu_{b_i|y_i(0), \tilde{y}_i(t, 0)}, \Sigma_{b_i|y_i(0), \tilde{y}_i(t, 0)}), \quad (8)$$

where

$$\begin{aligned} \mu_{b_i|y_i(0), \tilde{y}_i(t, 0)} &= \mu_b + \begin{bmatrix} \Sigma_{ba} & t\Sigma_b \end{bmatrix} \begin{bmatrix} \Sigma_a & t\Sigma_{ab} \\ t\Sigma_{ba} & t^2\Sigma_b + tD \end{bmatrix}^{-1} \\ &\quad \begin{bmatrix} y_i(0) - \mu_a \\ \tilde{y}_i(t, 0) - t\mu_b \end{bmatrix} \\ &= \left(\Sigma_b^{-1} + tD^{-1} + \Sigma_b^{-1} \Sigma_{ba} \Sigma_{Y_i(0)|b_i}^{-1} \Sigma_{ab} \Sigma_b^{-1} \right)^{-1} \\ &\quad \left[\Sigma_b^{-1} \mu_b + D^{-1} \tilde{y}_i(t, 0) + \Sigma_b^{-1} \Sigma_{ba} \Sigma_{Y_i(0)|b_i}^{-1} \right. \\ &\quad \left. \left(y_i(0) - \mu_a + \Sigma_{ab} \Sigma_b^{-1} \mu_b \right) \right] \text{ and} \\ \Sigma_{b_i|y_i(0), \tilde{y}_i(t, 0)} &= \Sigma_b - \begin{bmatrix} \Sigma_{ba} & t\Sigma_b \end{bmatrix} \\ &\quad \begin{bmatrix} \Sigma_a & t\Sigma_{ab} \\ t\Sigma_{ba} & t^2\Sigma_b + tD \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{ab} \\ t\Sigma_b \end{bmatrix} \\ &= \left(\Sigma_b^{-1} + tD^{-1} + \Sigma_b^{-1} \Sigma_{ba} \Sigma_{Y_i(0)|b_i}^{-1} \Sigma_{ab} \Sigma_b^{-1} \right)^{-1}. \end{aligned}$$

In the equations above, $\Sigma_{Y_i(0)|b_i} = \Sigma_a - \Sigma_{ab} \Sigma_b^{-1} \Sigma_{ba}$. Similar to (7), $\mu_{b_i|y_i(0), \tilde{y}_i(t, 0)}$ and $\Sigma_{b_i|y_i(0), \tilde{y}_i(t, 0)}$ can also be expressed

in two ways. Their derivations are provided in Supplementary Section B. It is interesting to observe that $\mathbf{b}_i|\tilde{\mathbf{y}}_i(t, 0)$ and $\mathbf{b}_i|y_i(0), \tilde{\mathbf{y}}_i(t, 0)$ are only dependent on the up-to-date information $\tilde{\mathbf{y}}_i(t, 0)$ and $\{y_i(0), \tilde{\mathbf{y}}_i(t, 0)\}$, respectively, regardless of any intermediate information. This suggests that the Markovian property holds for this proposed model. In addition, the univariate conditional (posterior) distribution of $b_{ij}|y_{ij}(0), \tilde{y}_{ij}(t, 0)$ and $b_i|y_i(0), \tilde{\mathbf{y}}_i(t, 0)$ can be derived, which will serve for the inference issue of models M_3 and M_4 , respectively. Supplementary Section C includes the details.

Last, conditioning on $\{y_i(0), \tilde{\mathbf{y}}_i(t, 0)\}$, the distribution of the subsequent degradation increments $\tilde{\mathbf{y}}_i(t + s, t)$ for an interval of time s is

$$\begin{aligned} & \tilde{\mathbf{y}}_i(t + s, t)|y_i(0), \tilde{\mathbf{y}}_i(t, 0) \\ & \sim \mathcal{MVN}\left(s\boldsymbol{\mu}_{b_i|y_i(0), \tilde{\mathbf{y}}_i(t, 0)}, s^2\boldsymbol{\Sigma}_{b_i|y_i(0), \tilde{\mathbf{y}}_i(t, 0)} + s\mathbf{D}\right). \end{aligned} \quad (9)$$

2.3. Lifetime Distribution

It is well known that the first passage time of reaching a pre-specified threshold for an individual Wiener process is subject to an inverse Gaussian distribution (Whitmore and Seshadri 1987). Concretely, given the initial degradation level $Y_j(0) = y_j(0)$, the degradation rate b_j , and the failure threshold $\mathcal{D}_j$ for the particular j th degradation process, the lifetime $T_{\mathcal{D}_j} \sim \mathcal{IG}\left(\frac{\mathcal{D}_j - y_j(0)}{b_j}, \left(\frac{\mathcal{D}_j - y_j(0)}{\sigma_j}\right)^2\right)$. Thus, the cdf of the lifetime is given by

$$\begin{aligned} & F_{T_{\mathcal{D}_j}}(t|y_j(0), b_j; \sigma_j^2, \mathcal{D}_j) \\ & = \Phi\left[\frac{tb_j - (\mathcal{D}_j - y_j(0))}{\sigma_j\sqrt{t}}\right] \\ & \quad + \exp\left[\frac{2b_j(\mathcal{D}_j - y_j(0))}{\sigma_j^2}\right] \Phi\left[-\frac{tb_j + \mathcal{D}_j - y_j(0)}{\sigma_j\sqrt{t}}\right], \end{aligned} \quad (10)$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. In this section, the subscript i is dropped from $y_{ij}(0)$ and b_{ij} , which indicates that population characteristics (i.e., the average values of the variables of interest after integrating out the random effects) are concerned with. Similar changes apply to other relevant notations.

Then, the unconditional cdf of the lifetime for PC j is provided by

$$\begin{aligned} & F_{T_{\mathcal{D}_j}}(t; \mu_{aj}, \mu_{bj}, \sigma_{aj}^2, \sigma_{bj}^2, \rho_j^{ab}, \sigma_j^2, \mathcal{D}_j) \\ & = \iint_{\mathbb{R}^2} F_{T_{\mathcal{D}_j}}(t|y_j(0), b_j; \sigma_j^2, \mathcal{D}_j) \\ & \quad \times f(b_j|y_j(0); \mu_{aj}, \mu_{bj}, \sigma_{aj}^2, \sigma_{bj}^2, \rho_j^{ab}) \\ & \quad \times f(y_j(0); \mu_{aj}, \sigma_{aj}^2) db_j dy_j(0). \end{aligned} \quad (11)$$

When multiple risks (i.e., degradation processes) threaten a product/system, the system failure time is the working period until any one of these processes reaches its failure

threshold (i.e., a competing-risk model). Mathematically, it is defined by

$$T_{\mathcal{D}} = \inf\{t : Y_1(t) \geq \mathcal{D}_1 \text{ or } \dots \text{ or } Y_p(t) \geq \mathcal{D}_p\},$$

where $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_p)'$ is the vector of all PCs' failure thresholds. Then, given $\{y(0), \mathbf{b}\}$, the conditional cdf of system lifetime is

$$\begin{aligned} F_{T_{\mathcal{D}}}(t|y(0), \mathbf{b}; \sigma^2, \mathcal{D}) & = 1 - \prod_{j=1}^p \left[1 - P(T_{\mathcal{D}_j} < t)\right] \\ & = 1 - \prod_{j=1}^p \left[1 - F_{T_{\mathcal{D}_j}}(t|y_j(0), b_j; \sigma_j^2, \mathcal{D}_j)\right]. \end{aligned} \quad (12)$$

This conclusion is based on the conditional independence of the marginal degradation processes given the random effects. The unconditional cdf of the system lifetime is obtained by integrating out the random effects. It is given by

$$F_{T_{\mathcal{D}}}(t; \boldsymbol{\theta}, \mathcal{D}) = \iint_{\mathbb{R}^{2p}} F_{T_{\mathcal{D}}}(t|y(0), \mathbf{b}; \sigma^2, \mathcal{D}) f(y(0), \mathbf{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) d^p y(0) d^p \mathbf{b}, \quad (13)$$

where $f(y(0), \mathbf{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents the joint pdf of the MVN distribution regarding $(Y(0)', \mathbf{b}')'$. The multivariate integral in (11) and (13) can be evaluated numerically.

3. Statistical Inference

3.1. EM Algorithm

For degradation data modeled by stochastic processes, using the property of independent increments is the typical approach to derive the log-likelihood function and estimate parameters. However, as our model involves three additional variabilities—the heterogeneous initial values, the varying degradation rates, and the initiation-growth correlation, it results in a high dimensional $\boldsymbol{\theta}$ to be inferred. Consequently, a simple maximum likelihood estimation (MLE) process may struggle to provide accurate parameter estimates. To address this challenge, we resort to an expectation-maximization (EM) algorithm. It tackles the estimation problem by separating the initial observations from the subsequent degradation increments, allowing for efficient parameter estimation. To proceed, we introduce several notations to describe the observed data.

Suppose a degradation test has m_i measurements on test unit i , $i = 1, 2, \dots, n$, after the initial inspection; thus, a total of $(m_i + 1) \times p$ observations are generated from this unit. Denote an individual observation by y_{ijk} , $\forall i = 1, 2, \dots, n, j = 1, 2, \dots, p, k = 0, 1, \dots, m_i$, and the corresponding degradation increment by $\tilde{y}_{ijk}$ (i.e., $y_{ijk} - y_{ijk-1}$), $\forall k = 1, \dots, m_i$. When an inspection on the i th unit is taken, the elapsed time t_{ik} , $\forall k = 0, 1, \dots, m_i$, is kept noted, and the time interval since the last measurement is recorded as $\tilde{t}_{ik}$ (i.e., $t_{ik} - t_{ik-1}$), $\forall k = 1, \dots, m_i$. Then, y_{ij0} would be the initial degradation observation at the test commencement time $t_{i0} = 0$. And we denote the initial observations of the i th unit by $\mathbf{y}_{i0} = (y_{i10}, y_{i20}, \dots, y_{ip0})'$ and the degradation increments of the i th unit among the interval from the $(k - 1)$ th to k th time point by $\tilde{\mathbf{y}}_{ik} = (\tilde{y}_{i1k}, \tilde{y}_{i2k}, \dots, \tilde{y}_{ipk})'$. Further, the collection of $\mathbf{y}_{i0}$'s across all units, the degradation

increments, and the measurement time intervals of the i th unit are respectively written as

$$\mathbb{Y}_0 = \begin{pmatrix} y'_{10} \\ y'_{20} \\ \vdots \\ y'_{n0} \end{pmatrix} = \begin{pmatrix} y_{110} & y_{120} & \cdots & y_{1p0} \\ y_{210} & y_{220} & \cdots & y_{2p0} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n10} & y_{n20} & \cdots & y_{np0} \end{pmatrix}_{n \times p},$$

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

$$\tilde{\mathbf{t}}_i = (\tilde{t}_{i1}, \tilde{t}_{i2}, \dots, \tilde{t}_{im_i})'_{m_i \times 1}.$$

Together, the total observed data across n units are defined to be $\{\mathbb{Y}_0, \mathbb{D}\}$, where $\mathbb{D} = \{\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_n, \tilde{\mathbf{t}}_1, \tilde{\mathbf{t}}_2, \dots, \tilde{\mathbf{t}}_n\}$.

Recall that $\mathbf{y}_{i0}$, previously referred to as $\mathbf{y}_i(0)$, is a realization of $\mathbf{Y}_i(0)$. Thus, the estimates of $\boldsymbol{\mu}_a$ and $\boldsymbol{\Sigma}_a$ can be obtained by simply maximizing the log-likelihood function of $\mathbb{Y}_0$. But it is difficult to obtain estimates of the other parameters because of the complicated form of the total log-likelihood function and its high dimensionality in terms of these parameters. To resolve this issue, in addition to the observed data, the remaining iid random-effects terms, $\mathbf{b}' = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_n)'$, can be treated as missing data or latent variables. Jointly, $\{\mathbb{Y}_0, \mathbb{D}, \mathbf{b}\}$ are viewed as the complete data, and their total log-likelihood function (up to a constant) is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbb{Y}_0, \mathbb{D}, \mathbf{b}) \\ = \ell(\mathbf{D}; \mathbb{D} \mid \mathbf{b}) + \ell(\boldsymbol{\mu}_a, \boldsymbol{\mu}_b, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b, \boldsymbol{\Sigma}_{ab}; \mathbf{b} \mid \mathbb{Y}_0) \\ + \ell(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a; \mathbb{Y}_0), \end{aligned}$$

where

$$\begin{aligned} \ell(\mathbf{D}; \mathbb{D} \mid \mathbf{b}) \\ = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left[\ln(\tilde{t}_{ik} \sigma_j^2) + \frac{(\tilde{y}_{ijk} - \tilde{t}_{ik} b_{ij})^2}{\tilde{t}_{ik} \sigma_j^2} \right], \end{aligned}$$

$$\begin{aligned} \ell(\boldsymbol{\mu}_a, \boldsymbol{\mu}_b, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b, \boldsymbol{\Sigma}_{ab}; \mathbf{b} \mid \mathbb{Y}_0) \\ = -\frac{1}{2} \sum_{i=1}^n \left[\ln |\boldsymbol{\Sigma}_{b_i|y_{i0}}| + (\mathbf{b}_i - \boldsymbol{\mu}_{b_i|y_{i0}})' \boldsymbol{\Sigma}_{b_i|y_{i0}}^{-1} (\mathbf{b}_i - \boldsymbol{\mu}_{b_i|y_{i0}}) \right], \text{ and} \end{aligned}$$

$$\begin{aligned} \ell(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a; \mathbb{Y}_0) \\ = -\frac{1}{2} \sum_{i=1}^n \left[\ln |\boldsymbol{\Sigma}_a| + (\mathbf{y}_{i0} - \boldsymbol{\mu}_a)' \boldsymbol{\Sigma}_a^{-1} (\mathbf{y}_{i0} - \boldsymbol{\mu}_a) \right]. \end{aligned}$$

In $\ell(\boldsymbol{\mu}_a, \boldsymbol{\mu}_b, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b, \boldsymbol{\Sigma}_{ab}; \mathbf{b} \mid \mathbb{Y}_0)$, $\boldsymbol{\mu}_{b_i|y_{i0}}$ and $\boldsymbol{\Sigma}_{b_i|y_{i0}}$ are the conditional mean vector and variance-covariance matrix, respectively, according to (6).

The EM algorithm alternates between an expectation step (E-step) and a maximization step (M-step) until convergence. The convergence criterion is built upon $\max |\boldsymbol{\theta}^{(s+1)} - \boldsymbol{\theta}^{(s)}| < \epsilon$, where $|\cdot|$ is the element-wise absolute value with the maximum

one returning by max, and ϵ is the error tolerance (e.g., 10^{-5}). In the process of iterative computation, we denote the estimate of $\boldsymbol{\theta}$ at the s th iteration by $\boldsymbol{\theta}^{(s)}$. When the algorithm terminates, $\hat{\boldsymbol{\theta}}$ is employed to denote the ultimate parameter estimates. Algorithm 1 briefly summarizes the algorithm. Please refer to Supplementary Section D.1 for more technical details. Note that the four special model variants in Table 1 are the simplified versions of the original model M_0 , so their inference steps can be adjusted accordingly. The adapted procedure is included in Supplementary Section D.2. When nonlinear degradation processes are present, transforming the original time scale via the aforementioned function $\Lambda(\cdot)$ is necessary. We assume that the transformation is a fixed effect and varies process by process, thus, leading to the transformed time interval denoted by $\tau_{ijk} = \Lambda_j(t_{ik}) - \Lambda_j(t_{i,k-1})$. Supplementary Section D.3 illustrates the consideration of this point. To achieve a fast convergence of this algorithm, finding a helpful starting point (i.e., values of parameter estimates) is necessary. This can be handled by summarizing the ad-hoc results that treat each degradation path as an independent realization from a simple (non)linear Wiener process. The detailed instructions are presented in Supplementary Section D.4. In addition, the R codes for implementing our proposed EM algorithms are available on the GitHub repository https://github.com/hnasu-code/multi_wiener.

3.2. Interval Estimation

In addition to the point estimate $\hat{\boldsymbol{\theta}}$, the confidence intervals (CIs) of either $\boldsymbol{\theta}$ or a general function $g(\boldsymbol{\theta})$ are often of interest too. Typically, an interval estimator is established based on the asymptotic theory, and it requires a large sample size to achieve sufficiently accurate results. But, due to the difficulty of evaluating the Fisher's information matrix of the proposed model, accomplishing this task is not easy. Instead, the bias-corrected percentile (BCp) bootstrap method (Efron and Tibshirani 1994; Meeker, Escobar, and Pascual 2022) is adopted for interval estimation. Supplementary Section E includes the details of this method.

3.3. Model Validation and Selection

Quantile-quantile (Q-Q) plots can be used to assess the goodness of fit (GOF) of the proposed model to a certain degradation dataset. It is noted that $(\tilde{y}_{ijk} - \tilde{t}_{ik} \hat{b}_{ij}) / \sqrt{\tilde{t}_{ik} \hat{\sigma}_j^2}$ is approximately iid as the standard normal distribution. Here, the estimate $\hat{b}_{ij} = E_{b_i|y_{i0}, \tilde{\mathbf{y}}_i, \tilde{\mathbf{t}}_i, \hat{\boldsymbol{\theta}}}[b_{ij}]$ can be calculated based on the conclusion implied by (8). If the significance of the initiation-growth correlation is of concern, one could use the likelihood-ratio (Wilks) test. Namely, to test if a degradation dataset favors model M_1 rather than the full model M_0 , we define the test statistic $\lambda_{LR} = 2[\ell_{M_0} - \ell_{M_1}]$, where ℓ_{M_0} and ℓ_{M_1} are the log-likelihood of models M_0 and M_1 , respectively. The null model (i.e., model M_1) is rejected if λ_{LR} is greater than the Chi-square upper percentile $\chi_{\alpha, p}^2$. Lastly, to compare various models, the Akaike information criterion (AIC), $AIC = 2 \dim(\boldsymbol{\theta}) - 2\ell$, is adopted, where $\dim(\boldsymbol{\theta})$ is the total number of model parameters, and ℓ is the corresponding log-likelihood.

Algorithm 1: An Outline of the EM Algorithm for Parameter Estimation.

Data: The dataset $\{\mathbb{Y}_0, \mathbb{D}\}$ in terms of the initial measurements, degradation increments, and measurement time intervals.

Input: The current EM estimate $\theta^{(s)}$ (i.e., $\mu_b^{(s)}, \Sigma_b^{(s)}, \Sigma_{ab}^{(s)}$, and $D^{(s)}$).

Output: The updated EM estimate $\theta^{(s+1)}$ (i.e., $\mu_b^{(s+1)}, \Sigma_b^{(s+1)}, \Sigma_{ab}^{(s+1)}$, and $D^{(s+1)}$).

while Convergence criterion is not met **do**

E-step: Define a function $Q(\theta|\theta^{(s)}) = E_{\mathbf{b}|\mathbb{Y}_0, \mathbb{D}, \theta^{(s)}} [\ell(\theta; \mathbb{Y}_0, \mathbb{D}, \mathbf{b})]$, which is the expected value of $\ell(\theta; \mathbb{Y}_0, \mathbb{D}, \mathbf{b})$ with respect to the current conditional distribution of $\mathbf{b}$ given $\mathbb{Y}_0, \mathbb{D}$ and $\theta^{(s)}$. In summary, the function is given by

$$Q(\theta|\theta^{(s)}) = -\frac{1}{2} \left\{ \sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^{m_i} \left[\ln(\tilde{t}_{ik} \sigma_j^2) + \frac{(\check{\mu}_{b_{i,j}}^{(s)})^2 + \check{\sigma}_{b_{i,j}}^{2(s)} - 2\check{y}_{ijk} \check{\mu}_{b_{i,j}}^{(s)} / \tilde{t}_{ik} + \check{y}_{ijk}^2 / \tilde{t}_{ik}^2}{\sigma_j^2 / \tilde{t}_{ik}} \right] + \right. \\ \sum_{i=1}^n \left[\ln |\Sigma_b - \Sigma_{ba} \hat{\Sigma}_a^{-1} \Sigma_{ab}| + \text{tr} \left((\Sigma_b - \Sigma_{ba} \hat{\Sigma}_a^{-1} \Sigma_{ab})^{-1} \check{\Sigma}_{b_i}^{(s)} \right) + \right. \\ \left. \left(\check{\mu}_{b_i}^{(s)} - \mu_b - \Sigma_{ba} \hat{\Sigma}_a^{-1} (y_{i0} - \hat{\mu}_a) \right)' (\Sigma_b - \Sigma_{ba} \hat{\Sigma}_a^{-1} \Sigma_{ab})^{-1} \left(\check{\mu}_{b_i}^{(s)} - \mu_b - \Sigma_{ba} \hat{\Sigma}_a^{-1} (y_{i0} - \hat{\mu}_a) \right) \right] + \\ \left. \sum_{i=1}^n \left[\ln |\hat{\Sigma}_a| + (y_{i0} - \hat{\mu}_a)' \hat{\Sigma}_a^{-1} (y_{i0} - \hat{\mu}_a) \right] \right\}.$$

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

$$\mu_b^{(s+1)} = \frac{1}{n} \sum_{i=1}^n \check{\mu}_{b_i}^{(s)}, \\ \Sigma_{ab}^{(s+1)} = \frac{1}{n-1} \sum_{i=1}^n (y_{i0} - \hat{\mu}_a) \left(\check{\mu}_{b_i}^{(s)} - \mu_b^{(s+1)} \right)', \\ \Sigma_b^{(s+1)} = \Sigma_{ba}^{(s+1)} \hat{\Sigma}_a^{-1} \Sigma_{ab}^{(s+1)} + \frac{1}{n} \sum_{i=1}^n \left\{ \check{\Sigma}_{b_i}^{(s)} + \right. \\ \left. \left[\check{\mu}_{b_i}^{(s)} - \mu_b^{(s+1)} - \Sigma_{ba}^{(s+1)} \hat{\Sigma}_a^{-1} (y_{i0} - \hat{\mu}_a) \right] \left[\check{\mu}_{b_i}^{(s)} - \mu_b^{(s+1)} - \Sigma_{ba}^{(s+1)} \hat{\Sigma}_a^{-1} (y_{i0} - \hat{\mu}_a) \right]' \right\}, \text{ and} \\ D^{(s+1)} = \text{diag} \left(\sigma_1^{2(s+1)}, \sigma_2^{2(s+1)}, \dots, \sigma_p^{2(s+1)} \right), \\ \text{where } \sigma_j^{2(s+1)} = \frac{\sum_{i=1}^n \sum_{k=1}^{m_i} \left[\tilde{t}_{ik} (\check{\mu}_{b_{i,j}}^{(s)})^2 + \tilde{t}_{ik} \check{\sigma}_{b_{i,j}}^{2(s)} - 2\check{y}_{ijk} \check{\mu}_{b_{i,j}}^{(s)} + \check{y}_{ijk}^2 / \tilde{t}_{ik} \right]}{\sum_{i=1}^n m_i}.$$

end

4. Simulation Studies

4.1. Performance of the Inference Method

To assess the performance of the proposed inference method, a Monte Carlo simulation study is carried out. In particular, a model of a three-dimensional degradation process is used. The model parameters are set as $\mu_a = (3, 4, 5)'$, $\mu_b = (5, 4, 3)'$, $\text{vech}(\Sigma_a) = (1, 0.5, 0.5, 1, 0.5, 1)'$, $\text{vech}(\Sigma_b) = (1, 0.8, 0.8, 1, 0.8, 1)'$, $\text{vec}(\Sigma_{ab}) = (0.3, 0.1, 0.1, 0.1, 0.3, 0.1, 0.1, 0.1, 0.3)'$, and $\text{vech}(D) = (4, 0, 0.5, 0, 6)'$, where vec is a vectorization operator that stacks the column vectors of a matrix into a single vector and vech is a half-vectorization operator that stacks the lower triangular block of a symmetric $p \times p$ matrix

into a single vector of length $p(p+1)/2$. It is assumed that the degradation measurement is taken every one unit of time. To understand the impact of sample size on the inference, three choices on the number of test units with $n = 10, 20, 30$ and two choices on the number of post-initial inspections with $m = 15, 30$ (i.e., identical across all units) are created. Thus, in total, six combinations of sample size are explored. For each setting, $B = 1000$ replications of data from the simulated model are generated and fitted by model M_0 using Algorithm 1. Figure 2 demonstrates some simulated degradation paths of the three PCs with $n = 10$ and $m = 15$.

The EM algorithm takes less than one-quarter of a minute to converge for each replication, regardless of the sample size

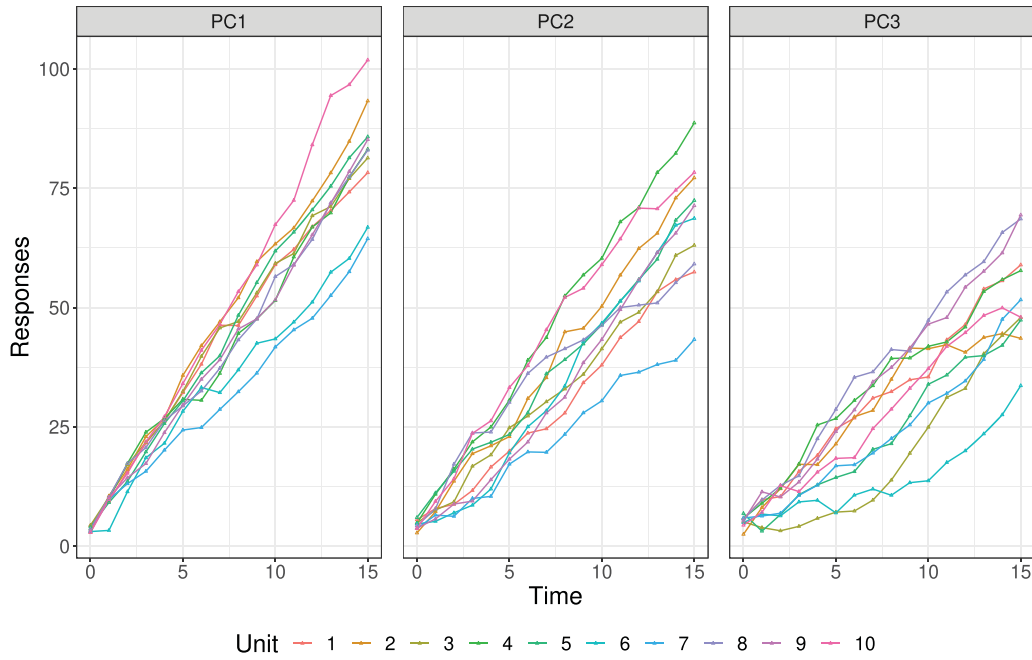


Figure 2. Simulated degradation paths with $n = 10$ and $m = 15$.

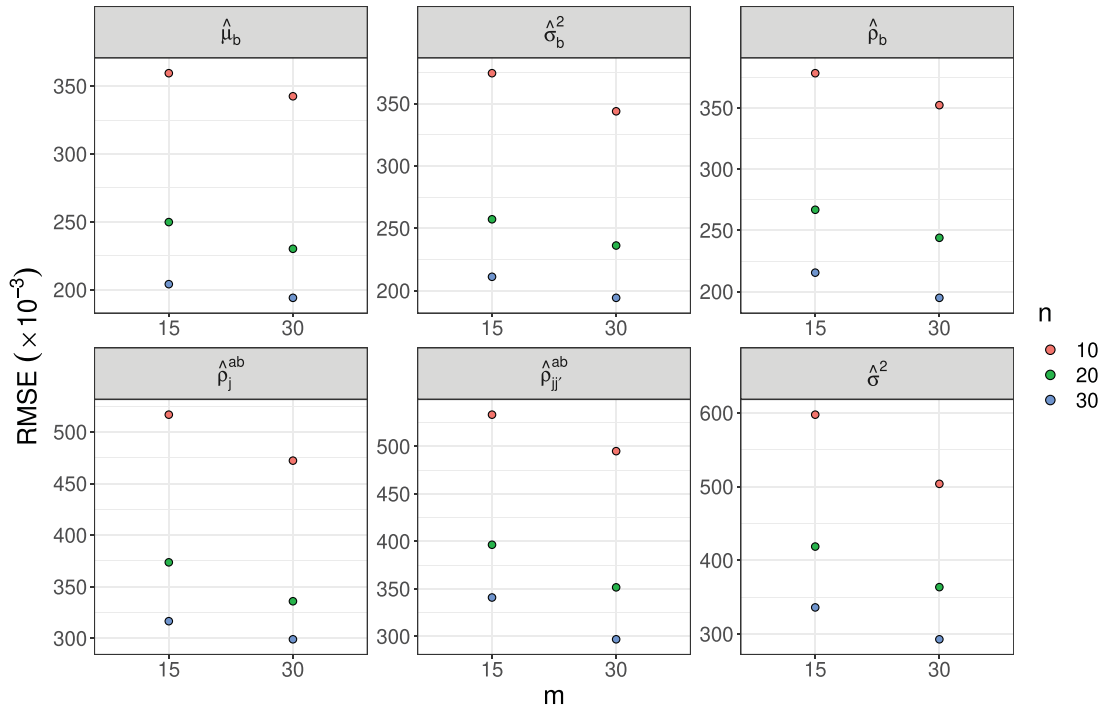


Figure 3. RMSE ($\times 10^{-3}$) of the EM estimators versus various sample sizes.

combination, on a laptop with an Apple M1 Max CPU. This is a satisfactory performance for practical use. To quantify the overall quality of the EM estimators, we calculate the root mean square error (RMSE) for all parameter estimates over the 1000 replications. Supplementary Section F.1 includes the raw data. To better summarize the result, we report the respective average RMSE over the elements in $\hat{\mu}_b$, $\hat{\sigma}_b^2$, $\hat{\rho}_b$, $\hat{\rho}_j^{ab}$, $\hat{\rho}_{jj'}^{ab}$, and $\hat{\sigma}^2$, where $\hat{\rho}_j^{ab}$ and $\hat{\rho}_{jj'}^{ab}$ represent the initiation-growth correlation within and between PCs, respectively. Figure 3 presents the result graphically in dot plots. It shows that, for each estimator,

the RMSE decreases with the increase of n for a fixed m and vice versa. This means that model estimation would get improved by having more data. A further investigation shows that the improvement of estimation accuracy for elements in μ_b , Σ_b , and Σ_{ab} is more sensitive to the change of n than m . This is unsurprising as the M-step in Algorithm 1 uses all sample units' information to estimate these three components. Among the three correlation estimators ($\hat{\rho}_b$, $\hat{\rho}_j^{ab}$, and $\hat{\rho}_{jj'}^{ab}$) of correlations, $\hat{\rho}_b$ performs the best in terms of the smallest RMSE under each sample size combination. This can be understood intuitively

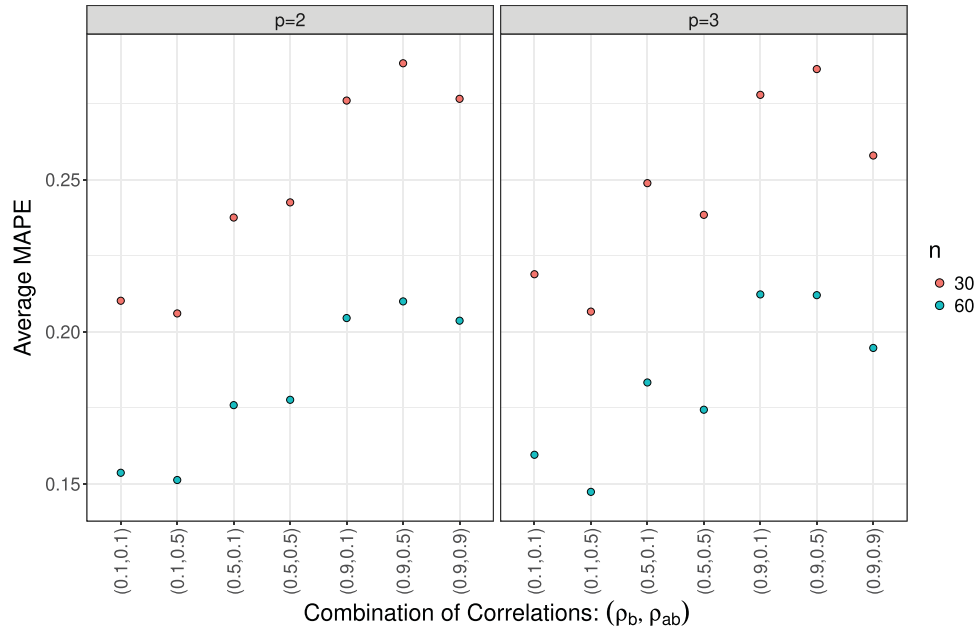


Figure 4. Performance of reliability estimation versus sample sizes under various model settings.

because more “similar” data, resulting from the strong correlation defined by ρ_b , help improve the estimation accuracy.

To examine the effect of degradation curvature on model estimation, we conduct another set of simulation studies with $\gamma = 0.7, 1, 1.3$, where the time scale is transformed by a power function t^γ . The three parameter values correspond to the concave, linear, and convex degradation trends. Similarly, the impacts of both the PC-wise and initiation-growth correlation at a low, medium, and high level are investigated. To explore the scalability of the inference method for higher dimensions, more simulation runs with three-, six-, and nine-dimensional models are implemented. Lastly, the estimation performance for models M_2 – M_4 is also validated. These additional studies indicate that our proposed inference method is effective and efficient. For more details, please refer to Supplementary Sections F.2–F.5.

4.2. Model Comparison in Reliability Estimation

To evaluate the performance of reliability estimation, we carry out an additional simulation study using two- or three-dimensional process models (i.e., $p = 2$ or $p = 3$). Data are generated from model M_0 under $m = 30$ with two options for the number of units ($n = 30$ or 60). We consider three different values (0.1, 0.5, or 0.9) for both the PC-wise correlation (ρ_b) and the initiation-growth correlation (ρ_{ab}). Switching among the three choices leads to a result of weak, medium, or strong pairwise correlation. To be precise, $\rho_b = \rho_b \mathbf{1}_{p(p-1)/2}$ and $\rho_{ab} = \rho_{ab} \mathbf{1}_{p^2}$, where $\mathbf{1}_p$ represents the p -dimensional vector of ones. The remaining parameters are set as $\mu_a = 3\mathbf{1}_p$, $\sigma_a^2 = \mathbf{1}_p$, $\rho_a = 0.9\mathbf{1}_{p(p-1)/2}$, $\mu_b = 3\mathbf{1}_p$, $\sigma_b^2 = \mathbf{1}_p$, and $\sigma^2 = 0.5\mathbf{1}_p$. For each model setting, $B = 1000$ replications of data are generated and fitted by model M_0 . Then, for each replication, the system reliability is computed for the core period, during which the true reliability gradually decays from 0.95 to 0.05. It is assumed that the failure threshold for each PC is 80. Finally,

we obtain the average mean absolute percentage error (MAPE) to evaluate the overall accuracy of the estimated reliability over the entire time window of interest. Figure 4 presents the result in dot plots. One can see that for all model settings, the reliability estimation accuracy will improve as the sample size increases. This conclusion stays consistent with the parameter estimation result.

After reviewing Table 2, it becomes apparent that model M_0 is a generalized form that comprehensively accommodates the features of random initial levels, random degradation rates, and initiation-growth correlation. In contrast, the special variants either incorporate only a part of these characteristics or hold some restrictive assumptions. To investigate the consequences of model misspecification, we extend the study above to cover models M_1 , M_3 , and M_4 . Precisely, in addition to model M_0 , we fit these alternative models to each simulated dataset and compute the average MAPE likewise. To quantify the relative performance of reliability estimation by the alternatives compared to the baseline model M_0 , we calculate a metric called the relative performance loss, defined as $100\% \times (\hat{\text{MAPE}}_{M_l} - \hat{\text{MAPE}}_{M_0}) / \hat{\text{MAPE}}_{M_0}$, where $\hat{\text{MAPE}}_{M_l}$ is the average MAPE of model M_l , $l = 1, 3, 4$.

Figure 5(a) shows the result for model M_1 . One can see that, as the value of ρ_{ab} increases, the metric also increases, indicating that the difference in the reliability estimation performance between models M_1 and M_0 becomes more significant. This matches our expectation because model M_1 fails to capture the feature of initiation-growth correlation. Additionally, a closer look at the figure reveals that the performance gap widens as the process dimension p increases. And the performance gap does not seem to have any obvious relationship with the sample size n .

Figure 5(b) shows the result for model M_3 . Since model M_3 overlooks the PC-wise correlation, the accuracy of reliability estimation would have a considerably larger loss with the increase of ρ_b . Moreover, having more data (i.e., larger n) or

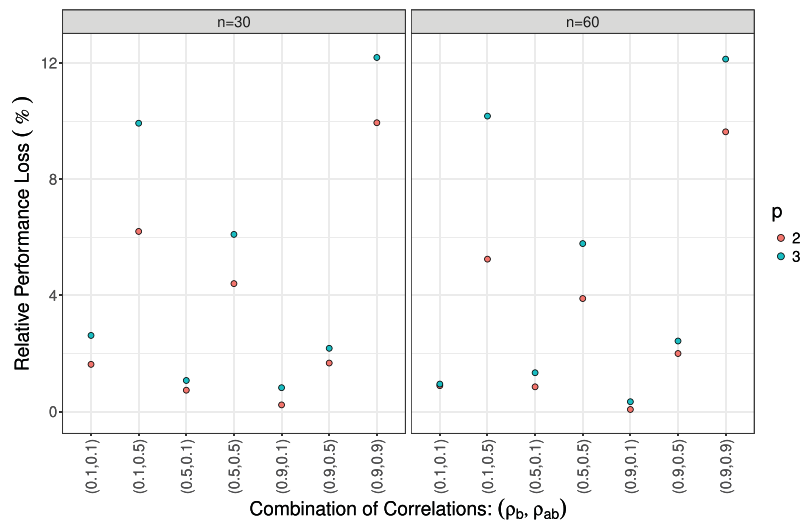
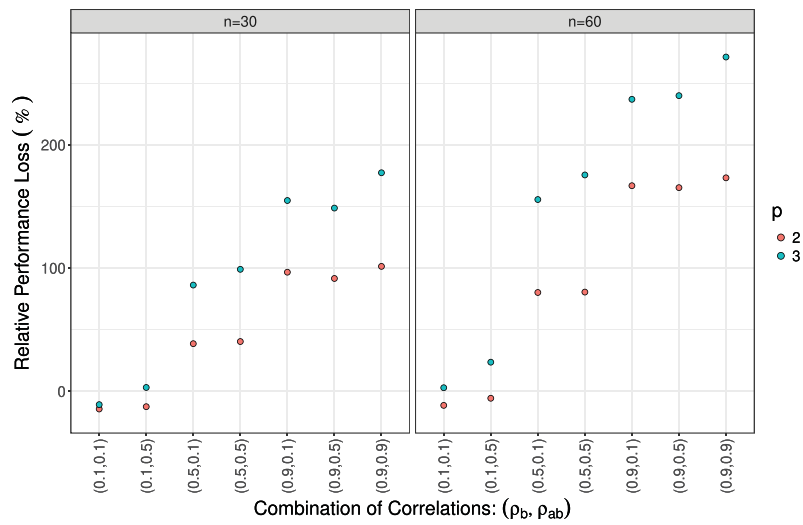
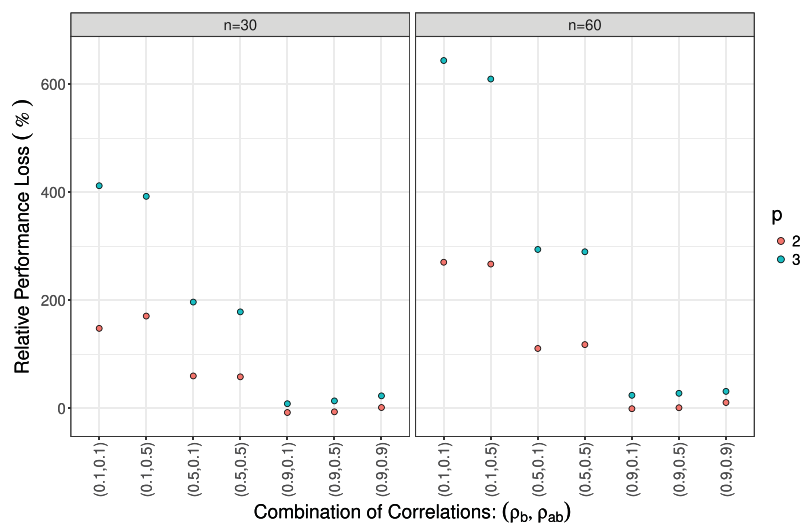
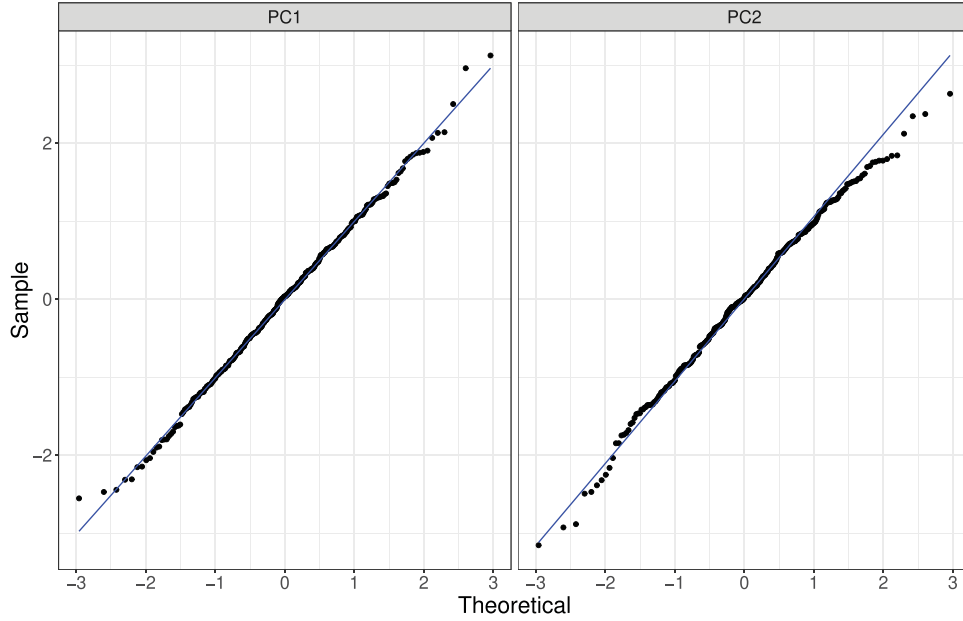
(a) Model M_1 .(b) Model M_3 .(c) Model M_4 .**Figure 5.** Relative performance loss under various model settings.

Table 3. Parameter point estimates and 90% C.I.s for model M_0 regarding the transceiver degradation data.

Parameter	Point estimate	90% C.I.	Parameter	Point estimate	90% C.I.
$\mu_{a1} \times 10^2$	5.794	[1.904, 9.990]	ρ_{12}^b	0.988	[0.979, 0.995]
$\mu_{a2} \times 10^2$	6.510	[1.637, 11.479]	ρ_1^{ab}	0.870	[0.375, 0.998]
$\sigma_{a1}^2 \times 10^3$	5.131	[2.120, 11.068]	ρ_{21}^{ab}	0.991	[0.980, 0.999]
$\sigma_{a2}^2 \times 10^3$	7.788	[3.388, 17.606]	ρ_{12}^{ab}	0.920	[0.640, 0.998]
ρ_{12}^a	0.840	[0.417, 0.943]	ρ_2^{ab}	0.975	[0.949, 0.997]
μ_{b1}	0.885	[0.669, 1.080]	$\sigma_1^2 \times 10^2$	8.581	[7.571, 9.932]
μ_{b2}	0.822	[0.571, 1.078]	$\sigma_2^2 \times 10$	1.570	[1.392, 1.792]
$\sigma_{b1}^2 \times 10^2$	4.024	[0.126, 11.985]	γ_1	1.098	[0.984, 1.210]
$\sigma_{b2}^2 \times 10^2$	3.201	[0.029, 9.213]	γ_2	1.125	[1.004, 1.250]

**Figure 6.** Normal Q-Q plots for model M_0 regarding the transceiver degradation data.

process dimension (i.e., larger p) will exacerbate the performance inferiority.

Figure 5(c) shows the result for model M_4 , where we observe a metric change pattern opposite to model M_3 . Specifically, as we increase ρ_b , a reduction in the performance loss is led to. This is because the common random-effects attribute of model M_4 implies that the PC-wise correlation always equals 1. Thus, as ρ_b increases, the true model becomes closer to model M_4 . However, similar to what we have observed above, model M_4 also exhibits an increasing pattern of performance loss as the process dimension or sample size expands.

Overall, these simulation results clearly demonstrate the importance of using a correct model that appropriately identifies the variability features in a degradation process in order to accurately estimate system reliability.

5. Case Studies

5.1. Application: Transceiver Degradation Data

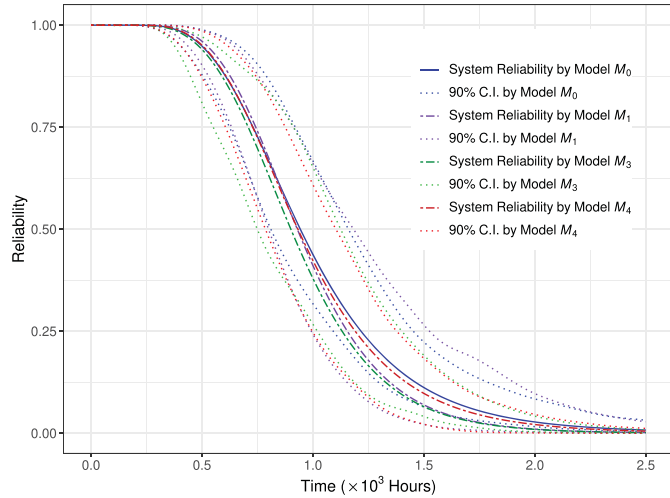
In this section, we demonstrate the implementation of the proposed methodology by analyzing the transceiver degradation data that motivated this study. As described in Section 1.2, the data exhibit three features of variability: unit-to-unit variation,

process dependency, and random initial observations, making the general model M_0 appropriate for fitting the data. After implementing the EM algorithm, we assess the convergence of model inference by examining trace plots of parameter estimates, presented in Figure G.1 in Supplementary Section G. Table 3 gives the result of parameter estimation. In the table, $\hat{\gamma}_j$ is the estimate of the parameter γ_j in the time scale transformation function $\Lambda_j(t) = t^{\gamma_j}$. The results indicate a very strong correlation between degradation rates for the two PCs. The positive values of the cross-correlation elements (i.e., $\hat{\rho}_1^{ab}$, $\hat{\rho}_{21}^{ab}$, $\hat{\rho}_{12}^{ab}$, and $\hat{\rho}_2^{ab}$) indicate that transceivers with higher degradation measurements at the beginning tend to have larger degradation rates.

The normal Q-Q plots, as depicted in Figure 6, show that the majority of data points closely align with two straight lines, indicating that the model offers a good fit to the dataset. We also verify the significance of the initiation-growth correlation by comparing the Wilks test statistic with the threshold associated with the Chi-square percentile. Additionally, as a comparison, we fit the data using alternative models listed in Table 2. The resulting log-likelihood and AIC values, presented in Table 4, indicate that Model M_0 is preferred as it yields the smallest AIC value.

Table 4. AIC values for model selection regarding the transceiver degradation data.

Candidate models	Log-likelihood	AIC
M_0	1065.518	−2095.036
M_1	1047.436	−2066.871
M_3	1058.945	−2089.889
M_4	1035.650	−2049.299

**Figure 7.** Reliability curves for transceivers.**Table 5.** AIC values for model selection regarding the IRLED degradation data.

Candidate models	Log-likelihood	AIC
M_0	−284.155	604.310
M_1	−286.920	601.840
M_3	−305.507	639.013
M_4	−303.874	629.748

To make reliability predictions, it is necessary to evaluate (13) over a certain period. For illustration purposes, failure thresholds of 1.2 dB and 1.0 dB are set for the two PCs. The reliability predictions are made using the four different models mentioned above. The evaluated reliability over time for these models is presented by the various curves in Figure 7, where point-wise C.I.s through the BCp bootstrap method are also shown. It can be seen that the predicted reliability between models M_0 and M_4 is very similar. This is not surprising since the default assumption of complete PC-wise dependence made by model M_4 almost matches the strong correlation estimated from model M_0 . In contrast, the PC-wise independence assumption causes the curve of model M_3 to deviate significantly from the curve of model M_0 . For model M_1 , the deviation of predicted reliability from the curve of model M_0 lies somewhere in between these two scenarios. This is because model M_1 , despite incorporating PC-wise dependence, overlooks the strong positive initiation-growth correlation estimated from model M_0 . Therefore, it is crucial to examine the underlying dependence structure carefully.

5.2. Additional Application: IRLED Degradation Data

In this section, we illustrate the use of our proposed methodology with another example—the degradation of GaAs/GaAs infrared light-emitting diodes (IRLEDs), which was originally

presented in Yang’s book (2007). This dataset, included in Table G.1 in Supplementary Section G, describes the change in the variation ratio of luminous power in percentage over time. Given that the original data have a single PC, we randomly split the data into two streams to mimic a multivariate flavor. The resulting artificially made data contain 6 units and 10 observations for each unit. Figure 8 shows the degradation paths of the two PCs.

Since the degradation starts from the status of zero damage, we adopt model M_2 to fit the data and estimate the parameters using the EM algorithm. The estimation result is included in Table G.2, where both log-likelihood and AIC values are also shown. The convergence of the EM algorithm is indicated by trace plots of parameter estimates over iteration, which is presented in Figure G.3a in Supplementary Section G. The normal Q–Q plots (shown in Figure G.2a) validate the fitted model.

Now, we consider a hypothetical scenario where a batch of manufactured products, including these six units, undergoes a 6-hr burn-in test to screen out those with infant mortality. As a result, the degradation paths of these units during the burn-in period are unavailable, and the measurements at the end of the burn-in duration become the initial observations when a new degradation test officially starts. In Figure 8, the resulting degradation paths are displayed, with the gray lines indicating the masked observations during the burn-in period. We fit four different models to the remaining data and compare their log-likelihood and AIC values, which are given in Table 5. Based on the lowest AIC value, model M_1 is chosen as the final model.

Lastly, we showcase how the proposed methodology can make unit-specific future degradation predictions. Recall that equation (9) provides the conditional distribution of the future degradation increments for a certain interval given the existing observations. Based on this property, we compute the expected degradation increments of both PCs for the next 160 hr for Unit 4. Figure 9 displays the predicted degradation paths for the two PCs using purple dashed lines, with gray dashed lines representing the 90% interval estimates.

6. Concluding Remarks

This article has systematically investigated a class of multivariate Wiener processes for degradation data modeling. This model is specified in a hierarchical fashion by incorporating multi-normally distributed random effects concerning the initial values and degradation rates. This characteristic produces a number of nice properties, including closed-form marginal, joint, and conditional (posterior) distributions. An EM algorithm has been developed to estimate the unknown parameters in the model, and the simulation and case studies show the effectiveness of the proposed methodology. It is worth mentioning that the model is flexible enough to accommodate different levels of variability, leading to various special model variants that can be selected to fit the degradation data in practice. Therefore, the proposed model is versatile and represents a novel class of multivariate degradation models. Overall, the proposed modeling framework is computationally tractable, physically meaningful, and practically applicable.

A different way of process-dependency modeling (see Hong, Ye, and Ling (2018a), Wang et al. (2020), and Sun, Ye, and

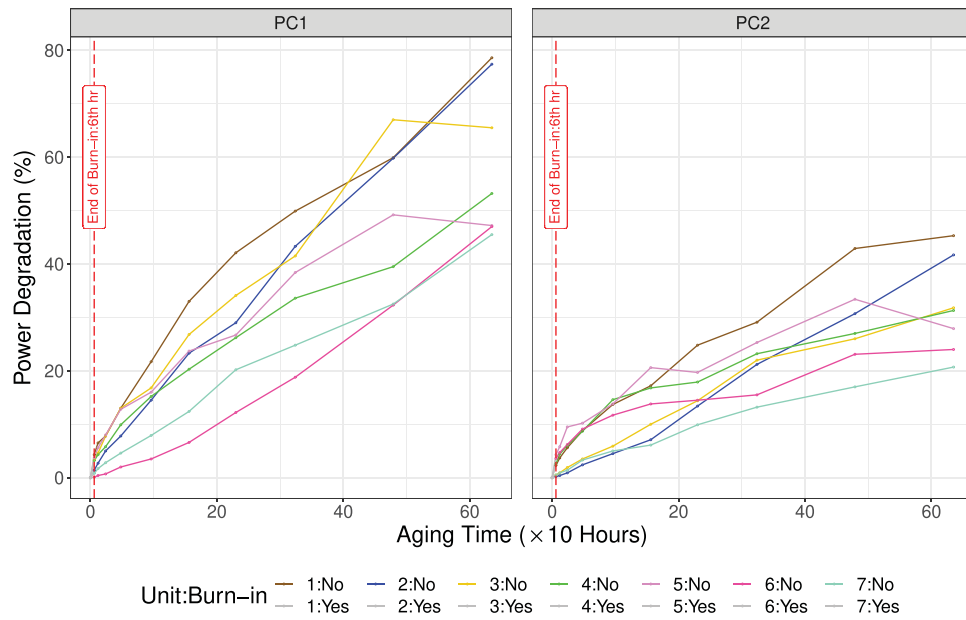


Figure 8. Degradation paths of IRLDs.

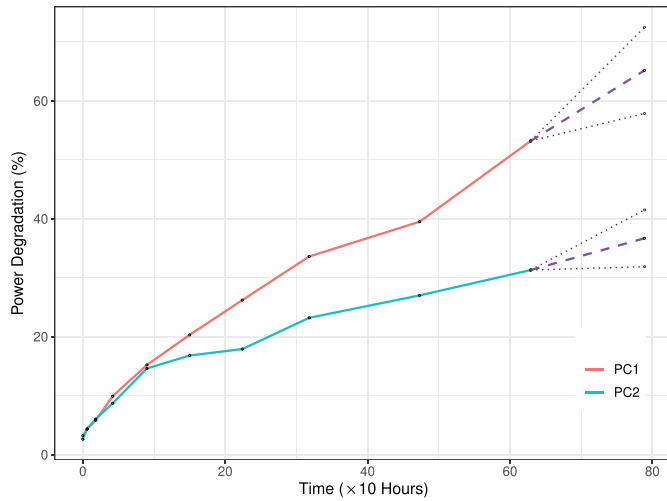


Figure 9. Future degradation predictions for Unit 4 of IRLDs.

Hong (2020b)) is to assume that the degradation observations of any PC pairs are correlated directly through the diffusion parameters in Wiener processes. To some extent, this approach suggests a measure of the temporal variability that may be attributed to various sources, including correlated measurement noises and shared environmental conditions. Our model takes a different approach by incorporating the dependency in the drift parameters as random effects. This design choice allows for a closer alignment with the inherent degradation physics implied by material properties (Ye and Xie 2015). Therefore, our approach conveys a specific physical interpretation, that is, the dependency originates from the interaction between the innate degradation mechanisms. Furthermore, the hierarchical approach used in this article makes the formulation and direct computation of multiple quantities of interest, such as conditional distributions and lifetime prediction, possible. But the conventional multi-dimensional Wiener process faces the problem of no explicit solutions to these quantities, especially the first

passage time distribution. Another direction of random effects modeling in degradation analysis is the Bayesian approach, as in Peng et al. (2016) and Fang, Pan, and Hong (2020). However, the Bayesian approach requires some subjective priors for the random-effects terms, whereas, in our hierarchical approach, they are estimated from the data directly (Demidenko 2013). Therefore, the proposed methodology provides solid physical ground and allows ample data-fitting flexibility.

Based on the current research result, some future studies can be pursued. For instance, the models for incorporating additional considerations such as nonperfect measurements, serial correlation, multi-phase degradation, and time-varying covariates can be explored. Meanwhile, some practical considerations warrant our attention. For example, in light of the generality and complexity of the proposed model, ensuring data sufficiency becomes the prerequisite for establishing a proper model; therefore, for practitioners, carefully defining product PCs and designing data collection schemes are critical steps. Furthermore, how to apply the proposed model in planning burn-in tests, scheduling preventative maintenance, and making residual life predictions would require a lot of research effort.

Supplementary Materials

In the online supplementary materials of this article, we provide a PDF file containing technical details (the miscellaneous proofs and inference methods), additional numerical results of the simulation and case studies, a list of notations, and a description of data and codes. Additionally, we provide a zip file containing the R codes for reproducing Table 5 in this article.

Acknowledgments

We would like to thank the editor, associate editor, and referees for their constructive comments and suggestions that helped us considerably improve the article. We gratefully acknowledge Dr. Wendai Wang for providing data resources for this study.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

G. Fang was partially supported by National Natural Science Foundation of China (Grant No. 72201242), Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ22G010003), and the Fundamental Research Funds for the Provincial Universities of Zhejiang (Grant No. XT202208). R. Pan was partially supported by National Science Foundation (Grant No. 2134409).

ORCID

Guanqi Fang  <https://orcid.org/0000-0002-3520-2986>

Rong Pan  <https://orcid.org/0000-0001-5171-8248>

References

- Bae, S. J., and Kvam, P. H. (2004), "A Nonlinear Random-Coefficients Model for Degradation Testing," *Technometrics*, 46, 460–469. [2]
- Bae, S. J., Yuan, T., Ning, S., and Kuo, W. (2015), "A Bayesian Approach to Modeling Two-Phase Degradation Using Change-Point Regression," *Reliability Engineering & System Safety*, 134, 66–74. [2]
- Demidenko, E. (2013), *Mixed Models: Theory and Applications with R*, Hoboken, NJ: Wiley. [14]
- Efron, B., and Tibshirani, R. (1994), *An introduction to the Bootstrap* (1st ed.), New York: CRC Press. [7]
- Fang, G., and Pan, R. (2021), "On Multivariate Copula Modeling of Dependent Degradation Processes," *Computers & Industrial Engineering*, 159, 107450. [2]
- Fang, G., Pan, R., and Hong, Y. (2020), "Copula-based Reliability Analysis of Degrading Systems with Dependent Failures," *Reliability Engineering & System Safety*, 193, 106618. [2,14]
- Fang, G., Pan, R., and Wang, Y. (2022), "Inverse Gaussian Processes with Correlated Random Effects for Multivariate Degradation Modeling," *European Journal of Operational Research*, 300, 1177–1193. [2,4]
- Fang, G., Rigdon, S. E., and Pan, R. (2018), "Predicting Lifetime by Degradation Tests: A Case Study of ISO 10995," *Quality and Reliability Engineering International*, 34, 1228–1237. [2]
- Gu, X., Jacobs, D., Byrd, W., Dickens, B., Vaca-Trigo, I., Meeker, W., Nguyen, T., and Martin, J. (2009), "Linking Accelerated Laboratory Test with Outdoor Performance Results for a Model Epoxy Coating System," in *Service Life Prediction of Polymeric Materials*, eds. J. W. Martin, R. A. Ryntz, J. Chin, and R. A. Dickie, pp. 3–28, Boston, MA: Springer US. [1]
- Hao, L., Bian, L., Gebraeel, N., and Shi, J. (2017), "Residual Life Prediction of Multistage Manufacturing Processes with Interaction between Tool Wear and Product Quality Degradation," *IEEE Transactions on Automation Science and Engineering*, 14, 1211–1224. [2]
- Hong, L., Tan, M. H. Y., and 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. [2]
- Hong, L., Ye, Z.-S., and Ling, R. (2018a), "Environmental Risk Assessment of Emerging Contaminants Using Degradation Data," *Journal of Agricultural, Biological and Environmental Statistics*, 23, 390–409. [2,13]
- Hong, Y., Zhang, M., and Meeker, W. Q. (2018b), "Big Data and Reliability Applications: The Complexity Dimension," *Journal of Quality Technology*, 50, 135–149. [2]
- Hu, J., Sun, Q., Ye, Z., and Zhou, Q. (2021), "Joint Modeling of Degradation and Lifetime Data for RUL Prediction of Deteriorating Products," *IEEE Transactions on Industrial Informatics*, 17, 4521–4531. [2]
- Kang, R., Gong, W., and Chen, Y. (2020), "Model-Driven Degradation Modeling Approaches: Investigation and Review," *Chinese Journal of Aeronautics*, 33, 1137–1153. [2]
- Lawless, J., and Crowder, M. (2004), "Covariates and Random Effects in a Gamma Process Model with Application to Degradation and Failure," *Lifetime Data Analysis*, 10, 213–227. [2]
- Li, H., Pan, D., and Chen, C. L. P. (2015), "Reliability Modeling and Life Estimation Using an Expectation Maximization based Wiener Degradation Model for Momentum Wheels," *IEEE Transactions on Cybernetics*, 45, 969–977. [4,5]
- Lu, C. J., and Meeker, W. O. (1993), "Using Degradation Measures to Estimate a Time-to-Failure Distribution," *Technometrics*, 35, 161–174. [2,3]
- Lu, L., Wang, B., Hong, Y., and Ye, Z. (2021), "General Path Models for Degradation Data with Multiple Characteristics and Covariates," *Technometrics*, 63, 354–369. [2,4]
- Meeker, W. Q., Escobar, L. A., and Pascual, F. G. (2022), *Statistical Methods for Reliability Data*, New York: Wiley. [7]
- Palayangoda, L. K., and Ng, H. K. T. (2021), "Semiparametric and Nonparametric Evaluation of First-Passage Distribution of Bivariate Degradation Processes," *Reliability Engineering & System Safety*, 205, 107230. [2]
- Peng, C.-Y. (2015), "Inverse Gaussian Processes with Random Effects and Explanatory Variables for Degradation Data," *Technometrics*, 57, 100–111. [2,3]
- Peng, H., Feng, Q., and Coit, D. (2009), "Simultaneous Quality and Reliability Optimization for Microengines Subject to Degradation," *IEEE Transactions on Reliability*, 58, 98–105. [1]
- Peng, W., Li, Y.-F., Mi, J., Yu, L., and Huang, H.-Z. (2016), "Reliability of Complex Systems Under Dynamic Conditions: A Bayesian Multivariate Degradation Perspective," *Reliability Engineering and System Safety*, 153, 75–87. [14]
- Rencher, A. C., and Christensen, W. F. (2012), *Methods of Multivariate Analysis*, Hoboken, NJ: Wiley. [5]
- Shen, L., Wang, Y., Zhai, Q., and Tang, Y. (2019), "Degradation Modeling Using Stochastic Processes with Random Initial Degradation," *IEEE Transactions on Reliability*, 68, 1320–1329. [1]
- Si, W., Yang, Q., Wu, X., and Chen, Y. (2018), "Reliability Analysis Considering Dynamic Material Local Deformation," *Journal of Quality Technology*, 50, 183–197. [1,2]
- Song, K., and 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. [2,4]
- Sun, F., Fu, F., Liao, H., and 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. [2]
- Sun, Q., Ye, Z.-S., and Hong, Y. (2020b), "Statistical Modeling of Multivariate Destructive Degradation Tests with Blocking," *Technometrics*, 62, 536–548. [2,4,14]
- Tian, Q., Liu, S., and Meeker, W. Q. (2019), "Using Degradation Models to Assess Pipeline Life," *Applied Stochastic Models in Business and Industry*, 35, 1411–1430. [1]
- Wang, X. (2010), "Wiener Processes with Random Effects for Degradation Data," *Journal of Multivariate Analysis*, 101, 340–351. [2]
- Wang, X., Gaudoin, O., Doyen, L., Bérenguer, C., and Xie, M. (2020), "Modeling Multivariate Degradation Processes with Time-Variant Covariates and Imperfect Maintenance Effects," *Applied Stochastic Models in Business and Industry*, 37, 592–611. [2,13]
- Weaver, B. P., Meeker, W. Q., Escobar, L. A., and Wendelberger, J. (2013), "Methods for Planning Repeated Measures Degradation Studies," *Technometrics*, 55, 122–134. [1]
- Wen, Y., Wu, J., and Yuan, Y. (2017), "Multiple-Phase Modeling of Degradation Signal for Condition Monitoring and Remaining Useful Life Prediction," *IEEE Transactions on Reliability*, 66, 924–938. [2]
- Whitmore, G., and Schenkelberg, F. (1997), "Modelling Accelerated Degradation Data Using Wiener Diffusion with a Time Scale Transformation," *Lifetime Data Analysis*, 3, 27–45. [3]
- Whitmore, G. A., and Seshadri, V. (1987), "A Heuristic Derivation of the Inverse Gaussian Distribution," *The American Statistician*, 41, 280–281. [6]
- Xiao, X., and Ye, Z. (2016), "Optimal Design for Destructive Degradation Tests with Random Initial Degradation Values using the Wiener Process," *IEEE Transactions on Reliability*, 65, 1327–1342. [4]

- Xu, A., Shen, L., Wang, B., and Tang, Y. (2018), "On Modeling Bivariate Wiener Degradation Process," *IEEE Transactions on Reliability*, 67, 897–906. [4]
- Yan, B., Wang, H., and Ma, X. (2022), "Correlation-Driven Multivariate Degradation Modeling and RUL Prediction based on Wiener Process Model," *Quality and Reliability Engineering International*. [2,4]
- Yang, G. (2007), *Life Cycle Reliability Engineering*, Hoboken, NJ: Wiley. [13]
- Ye, Z.-S., and Chen, N. (2014), "The Inverse Gaussian Process as a Degradation Model," *Technometrics*, 56, 302–311. [2]
- Ye, Z.-S., Hu, Q., and Yu, D. (2019), "Strategic Allocation of Test Units in an Accelerated Degradation Test Plan," *Journal of Quality Technology*, 51, 64–80. [1,4]
- Ye, Z.-S., and Xie, M. (2015), "Stochastic Modelling and Analysis of Degradation for Highly Reliable Products," *Applied Stochastic Models in Business and Industry*, 31, 16–32. [2,14]
- Ye, Z.-S., Xie, M., Tang, L.-C., and Chen, N. (2014), "Semiparametric Estimation of Gamma Processes for Deteriorating Products," *Technometrics*, 56, 504–513. [2]
- Zhai, Q., and Ye, Z.-S. (2023), "A Multivariate Stochastic Degradation Model for Dependent Performance Characteristics," *Technometrics*, 1–13. [2]
- Zhang, Z., Si, X., Hu, C., and Lei, Y. (2018), "Degradation Data Analysis and Remaining Useful Life Estimation: A Review on Wiener-Process-based Methods," *European Journal of Operational Research*, 271, 775–796. [2]