Modeling Two-Scale Degradation with Heterogeneity: A Unified Random-Effects Inverse Gaussian Framework

Liangliang Zhuang[a], Yizhong Ma[a], Guanqi Fang[b,c], Ancha Xu[b,c,d,*]

[a]*School of Economics and Management, Department of Management Science and Engineering, Nanjing University of Science and Technology, Nanjing, 210094, China*
[b]*School of Statistics and Data Science, Zhejiang Gongshang University, Hangzhou 310018, China*
[c]*Collaborative Innovation Center of Statistical Data Engineering, Technology & Application Zhejiang Gongshang University, Hangzhou, China*
[d]*Laboratory for Statistical Monitoring and Intelligent Governance of Common Prosperity*

Abstract

Accurate modeling of product degradation is essential for reliable life prediction and maintenance planning. In many applications, degradation is jointly driven by age (calendar time) and usage (cumulative operation), and further complicated by unit-to-unit heterogeneity due to manufacturing or environmental variations. While existing models often focus on single-scale processes or assume homogeneous systems, few frameworks can flexibly accommodate both two scales and population-level variability. This paper proposes a unified two-scale degradation model based on additive reparameterized inverse Gaussian processes with random effects. The model captures monotonic degradation under the joint influence of age and usage, and introduces two forms of random effects to flexibly capture varying degrees of heterogeneity across units. It provides closed-form expressions for failure time distributions, enabling efficient reliability analysis. Two inference methods are developed: (i) a maximum likelihood estimator via an expectation-maximization algorithm with bootstrap intervals, and (ii) a Bayesian approach using Hamiltonian Monte Carlo sampling. Simulation studies confirm the accuracy and robustness of the proposed estimators and highlight the risks of model misspecification in reliability assessment. A real-world case study on outdoor coating degradation further demonstrates the model's practical applicability. An open-source R package is provided to support implementation, with additional materials available online.

Keywords: Inverse Gaussian process; Two scales; Degradation; Random effects

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

1. Introduction

1.1. Background

With the increasing complexity of modern operating environments, the factors driving product performance degradation have become increasingly diverse. Over a product's life-cycle, degradation may result from both calendar time (age-based aging) and cumulative operational usage (usage-based aging). Without timely intervention, such degradation can lead to reduced efficiency, suboptimal performance, or even catastrophic failure. Therefore, accurate degradation modeling and predictive maintenance strategies are critical to ensuring system reliability and safety (Wang and Tang, 2025; Zhao et al., 2025).

In practice, both *age-based* and *usage-based* degradation mechanisms often coexist and interact (Wang et al., 2020; Peng et al., 2024). Age-based degradation refers to natural deterioration over time due to material fatigue, corrosion, or environmental exposure. Usage-based degradation, by contrast, results from cumulative loads, repeated stress, or operational cycles. For instance, electric vehicle batteries experience age-based capacity fade as well as usage-driven increases in internal resistance. Capturing such dual degradation pathways requires models that account for both temporal and operational scales simultaneously.

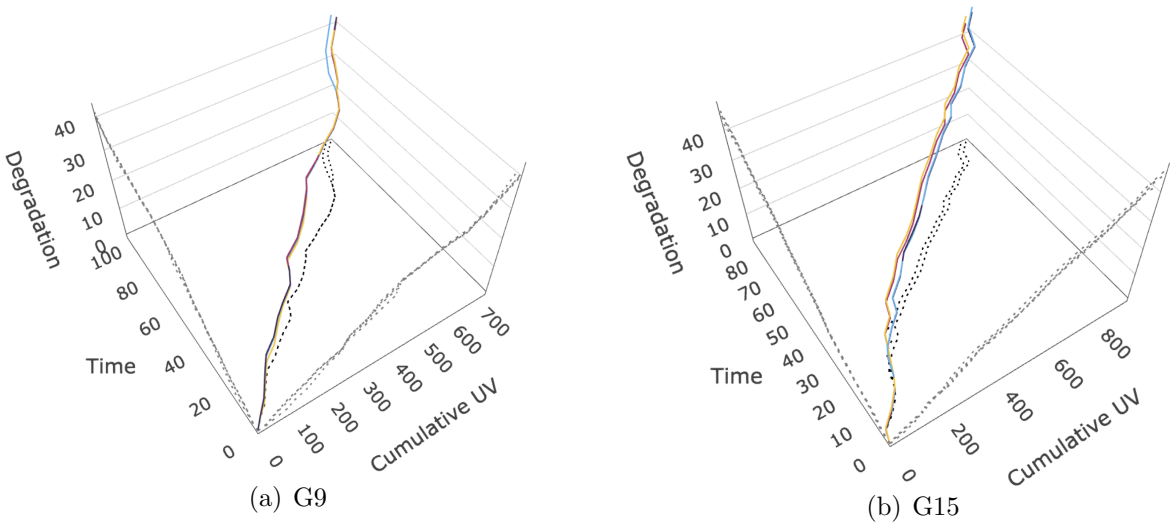
Moreover, real-world systems often exhibit unit-to-unit variability caused by manufacturing imperfections, operating conditions, or environmental heterogeneity (Xu and Wang, 2025). Ignoring such differences may lead to biased reliability estimates and misleading maintenance decisions. As a result, degradation models must not only incorporate multiple time scales but also account for cross-unit heterogeneity in degradation behavior. The following example further illustrates this research context.

1.2. Motivation

This example comes from a five-year outdoor experiment conducted by the national institute of standards and technology (Gu et al., 2009). In this study, researchers placed coating material samples in outdoor chambers to observe their degradation under environmental factors such as ultraviolet (UV) radiation exposure. By monitoring the attenuation of the 1250 cm^{-1} peak in Fourier transform infrared spectroscopy, they assessed the material's degradation (Hong et al., 2015). In this process, degradation accumulated gradually over time (time-based aging), while UV radiation accelerated polymer chain breakdown, molecular weight reduction, and mechanical weakening. Thus, cumulative UV exposure serves as a key indicator of usage-based aging.

Figure 1 shows the degradation paths of two sample groups across calendar time and cumulative UV exposure (both scaled by $\times 10^2$ for UV exposure and degradation values), along with projections of each pair of axes. The figure reveals that degradation increases monotonically with calendar time, while cumulative UV exposure grows nonlinearly. Although the degradation paths within each group (e.g., G9) show limited variation, small inter-unit differences remain, reflecting potential heterogeneity in practical settings.

These observations raise key questions: Can a model be developed to jointly capture degradation driven by both age and usage? Can it account for the combined and scale-dependent effects of calendar time and cumulative UV exposure while remaining flexible enough to incorporate unit-to-unit variability? To address these challenges, this article proposes a two-scale monotonic degradation model with random effects, along with methods for parameter estimation and reliability analysis.



(a) G9 (b) G15

Figure 1: Degradation of the coating material over calendar time and cumulative UV exposure.

1.3. Related literature

Traditional degradation modeling primarily relies on general path models (Si et al., 2013; Lu et al., 2020). While effective in certain contexts, these models treat degradation as a deterministic process, ignoring inherent randomness. This simplification limits their ability to capture the true degradation paths of products, reducing their suitability for complex degradation behaviors (Zhai et al., 2023). To overcome these limitations, stochastic process

models, classified as either non-monotonic or monotonic based on degradation increment properties, have been widely adopted in degradation modeling.

i) **Non-monotonic degradation processes**: Common models include the Wiener process (Wang et al., 2023a; Fang and Pan, 2024), fractional Lévy motion (Zhang et al., 2021; Asgari et al., 2024), and the Student-t process (Peng and Cheng, 2020; Xu et al., 2024), which are suitable for degradation processes with fluctuations or non-monotonic behavior (Zhang et al., 2018). These models have been extensively studied and applied in engineering applications. For instance: Hu et al. (2020) developed a Wiener process model with random effects for online RUL prediction, using data from lead-acid batteries and digital communication systems. Wang et al. (2023b) applied an autoregressive model with a Wiener process to simulate stochastic degradation rates in rechargeable battery and bearing data. Liu et al. (2022) introduced a fractional Lévy motion-based method for RUL prediction, capturing long-range dependence in lithium-ion battery degradation. Li et al. (2023) further extended this with adaptive diffusion coefficients and nonlinear drift, applying it to truck rear axle data. Peng and Cheng (2020) proposed a robust Student-t process model for lifetime assessment of highly reliable products, illustrated by resistance, train wheel, and light-emitting diode (LED) degradation data.

ii) **Monotonic degradation processes**: Common models include the gamma process (Zhou et al., 2023; Song and Cui, 2022) and the inverse Gaussian (IG) process (Wang and Xu, 2010; Ye and Chen, 2014; Zhuang et al., 2024), which are effective in characterizing cumulative degradation. These models are also widely applied in practical scenarios. For instance: Wang et al. (2021) proposed a random-effects gamma degradation model and applied it to laser data and Device-B data. Both Ling et al. (2019) and Lin et al. (2021) developed gamma process models for two-phase degradation—Ling et al. (2019) applied the model to LED data, while Lin et al. (2021) used it to cycle aging data from lithium batteries. Peng et al. (2014) introduced a Bayesian framework using the IG process for degradation analysis, validating it with monotonic datasets. Extending this model, Fang et al. (2022) developed a multivariate IG model that accounts for performance dependencies and product heterogeneity, applied to coating and fatigue crack-size data.

However, the aforementioned studies primarily address degradation indexed by a single

scale. To avoid confusion, multivariate degradation models are not equivalent to multi-scale degradation models: although multivariate models are statistically multidimensional, they still operate on a single physical scale—typically time—because all degradation signals evolve along a common temporal axis. Few considering a two-scale approach. For instance, Zhang et al. (2017) introduced a linear Wiener process model to capture degradation across both age and usage scales, which Pei et al. (2019) later extended to a nonlinear form for gyroscope data. Nevertheless, these studies assume a constant usage rate and overlook dependencies between the two scales. To address these issues, Zhai et al. (2023) proposed an innovative two-scale model that combines two independent Wiener processes, $\mathcal{X}(t)$ and $\mathcal{Z}(u)$, representing degradation caused by the age scale t and the usage scale u, respectively. The degradation from these two scales combines into the observed degradation path, $\mathcal{Y}(t,u) = \mathcal{X}(t) + \mathcal{Z}(u)$, which follows a normal distribution given fixed t and u due to the property of Wiener process. This model provides two key advantages: (i) a clear physical interpretation of how both scales jointly affect product lifetime, and (ii) tractable properties, including a closed-form lifetime distribution, facilitating application in reliability prediction. Despite advancements, monotonic process-based two-scale models may be particularly suited for analyzing systems with monotonic degradation, such as wear in power shift steering transmissions (Song and Cui, 2022), operational currents of GaAs lasers (Zhou et al., 2023), and the example in Section 1.2. Monotonic models are generally more effective at capturing cumulative degradation patterns and enhancing reliability estimates, yet research in this area remains limited. Moreover, under the two-scale setting, unit-to-unit heterogeneity has not been sufficiently addressed. While recent studies have incorporated random effects to capture such heterogeneity in both monotonic and nonmonotonic degradation processes (Zhai et al., 2025; Fan et al., 2023; Zhou et al., 2023), most of these works remain within single-scale frameworks and rarely consider cross-scale dependence. Neglecting heterogeneity—arising from variations in material properties, manufacturing processes, or environmental exposure—may lead to biased inference and reduced prediction accuracy. This motivates the development of a unified two-scale model with shared random effects, which jointly characterizes age–usage dependence and unit-level variability within a coherent probabilistic framework.

Table 1: Comparison of representative degradation modeling frameworks.

Model	Scale dependency	Random effects	Path monotonicity	References
Single-scale	-	No	Monotonic	• Lin et al. (2021)
		Yes	Monotonic	• Fang et al. (2022)
		Yes	Non-monotonic	• Zhai et al. (2025)
Two-scale	Independent	No	Non-monotonic	• Zhang et al. (2017)
	Dependent	No	Non-monotonic	• Zhai et al. (2023)
	Dependent	Yes	Monotonic	• This paper

1.4. Contribution and overview

Table 1 summarizes representative degradation modeling frameworks in terms of scale dependency, random effects, and degradation-path monotonicity. As shown, single-scale models are indexed by a single scale (typically calendar time) and partially account for unit-level heterogeneity. Two-scale models have been developed to incorporate both calendar time and usage, but most existing studies either assume independence between the two scales, consider nonmonotonic degradation paths, or fail to capture unit-to-unit variability. In contrast, this study proposes a dependent two-scale framework that integrates shared random effects to represent cross-scale interactions and individual heterogeneity, while preserving the monotonic characteristics of physical degradation. The main contributions of this study can be summarized as follows:

- A unified dependent two-scale degradation framework is developed based on the reparameterized inverse Gaussian (rIG) process, capturing monotonic degradation jointly driven by time and usage while accounting for unit-level heterogeneity. Existing single-scale and independent two-scale models are encompassed as special cases.
- Statistical properties are derived, yielding a closed-form lifetime distribution that ensures computational efficiency for reliability assessment.
- Two complementary estimation approaches are proposed: maximum likelihood estimation via an expectation–maximization (EM) algorithm and Bayesian inference.

- An open-source `R` package, `r2IGP`, is developed to facilitate its implementation and reliability analysis in practice.

The remainder of this article is organized as follows. Section 2 discusses the model formulation and reliability analysis. Section 3.1 introduces the data and likelihood function, while Sections 3.2 and 3.3 detail the EM algorithm and the bootstrap method, respectively. Section 4 evaluates inference accuracy and the impact of model misspecification through simulations. Section 5 applies the proposed methodology to coating material data, and Section 6 concludes the study.

2. Model formulation

This section begins by defining the rIG process and applying it to the two-scale degradation model with random effects, along with other extended models (see Section 2.1). It then presents the theoretical properties and failure time distribution of the proposed model (see Section 2.2), followed by a reliability analysis (see Section 2.3).

2.1. Two-scale rIG processes

Consider a stochastic process $\mathcal{J}(u)$, defined as an rIG process $r\mathcal{IG}(\Lambda(u), \gamma)$, if it satisfies the following properties: (i) $\mathcal{J}(0) = 0$ with probability 1; (ii) $\mathcal{J}(u)$ has independent increments, that is, for any $u_4 > u_3 \geq u_2 > u_1 \geq 0$, the increments $\mathcal{J}(u_4) - \mathcal{J}(u_3)$ and $\mathcal{J}(u_2) - \mathcal{J}(u_1)$ are independent; (iii) For all $u > v \geq 0$, the increment $\mathcal{J}(u) - \mathcal{J}(v)$ follows an rIG distribution $rIG(\Lambda(u) - \Lambda(v), \gamma)$. In this context, $\Lambda(\cdot)$ is a non-decreasing function satisfying $\Lambda(0) = 0$, representing the drift parameter, and γ is the dispersion parameter. The probability density function (PDF) of $rIG(\theta, \gamma)$ is defined as

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

Note that the rIG distribution, $rIG(\theta, \gamma)$, relates to the traditional IG distribution, $IG(a, b)$, through the parameter transformations $a = \theta/\gamma$ and $b = \theta^2$. Specifically, if $\mathcal{Y}_1 \sim rIG(\theta_1, \gamma)$ and $\mathcal{Y}_2 \sim rIG(\theta_2, \gamma)$ are two independent random variables, then their summation, $\mathcal{Y} = \mathcal{Y}_1 + \mathcal{Y}_2$, also follows an rIG distribution $rIG(\theta_1 + \theta_2, \gamma)$ (Zhuang et al., 2024).

Motivated by this additive closure, we construct an additive two-scale rIG model for monotonic degradation. For a unit $i, i = 1, 2, \ldots, n$, the degradation process $\mathcal{Y}_i(t, u)$ is represented by two scales $t > 0$ and $u > 0$, and is formulated as

$$\mathcal{Y}_i(t, u) = \mathcal{X}_i(t) + \mathcal{Z}_i(u), \tag{2}$$

with $\mathcal{X}_i(t) \sim r\mathcal{IG}\left(\Lambda^t(t), \gamma\right)$, and $\mathcal{Z}_i(u) \sim r\mathcal{IG}\left(\Lambda^u(u), \gamma\right)$. Here, $\mathcal{X}_i(t)$ and $\mathcal{Z}_i(u)$ are independent rIG processes that both follow the form described in (1) and share the same dispersion parameter γ, assumed to reflect the underlying failure mechanism of the product. The functions $\Lambda^t(t)$ and $\Lambda^u(u)$ act as the drift parameters for $\mathcal{X}_i(t)$ and $\mathcal{Z}_i(u)$, respectively. In practical applications, the specific functional form of $\Lambda(\cdot)$ can be identified based on engineering knowledge or empirical analysis.

Note that the proposed two-scale model adopts an additive structure, which represents the cumulative contributions of calendar time and usage to degradation. In many physicochemical aging mechanisms (e.g., corrosion progression, coating photodegradation, and fatigue damage accumulation), different degradation drivers accumulate through distinct damage pathways (Wang, 2023). This makes the additive formulation a reasonable approximation for a broad class of degradation processes. In addition, the additive form leads to closed-form reliability expressions and improves parameter identifiability by avoiding the confounding that often arises in multiplicative or interaction-based models. Nevertheless, when strong interactions exist between the two scales, the additive form may become inadequate, and multiplicative or interaction models can serve as potential extensions. A brief discussion of such possible extensions is provided in the Section 6.

Due to the diversity in manufacturing processes, material properties, and operating environments, products often exhibit unique degradation characteristics. To capture this heterogeneity among products, we model γ as a random variable with a normal distribution, expressed as $\gamma_i \sim \mathcal{N}(\kappa, \sigma^2)$. The probability of γ_i being negative becomes negligibly small when $\kappa/\sigma^2 \gg 0$, which is typically a reasonable assumption in practice (Peng and Cheng, 2020). In summary, the proposed model labeled as M_0 is formulated as follows:

$$M_0 : \begin{cases} \mathcal{Y}_i(t, u) = \mathcal{X}_i(t) + \mathcal{Z}_i(u), \\ \mathcal{X}_i(t) \mid \gamma_i \sim r\mathcal{IG}\left(\Lambda^t(t), \gamma_i\right), \\ \mathcal{Z}_i(u) \mid \gamma_i \sim r\mathcal{IG}\left(\Lambda^u(u), \gamma_i\right), \\ \gamma_i \sim \mathcal{N}(\kappa, \sigma^2). \end{cases} \tag{M_0}$$

Model $[M_0]$ draws inspiration from the degradation mechanisms observed in physical systems, such as coating materials and batteries. Here, $\mathcal{X}_i(t)$ represents age-based degradation, capturing processes like material fatigue that progress over calendar time t, while $\mathcal{Z}_i(u)$ captures usage-based degradation, reflecting wear due to cumulative operational cycles. This two-scale framework enables independent modeling of each degradation path while capturing their combined influence on product reliability. The structure effectively addresses the complexity of real-world product degradation mechanisms, making it particularly suitable for predicting failures with both age- and usage-based deterioration.

To accommodate the practical needs of degradation modeling in various engineering scenarios, the proposed model $[M_0]$ is formulated as a unified two-scale stochastic degradation framework. It not only provides full parameter flexibility but also encompasses four meaningful model variants derived from real-world considerations. These variants are constructed by adjusting two key dimensions: (i) the degradation scale involved (single or dual), and (ii) the presence or absence of unit-to-unit heterogeneity (i.e., random effects). Such structural differences correspond to typical degradation settings observed in practice, including systems affected by multiple degradation mechanisms or those exhibiting variability due to manufacturing inconsistencies. Table 2 summarizes the proposed unified framework and its representative variants, including both newly introduced models and existing formulations from the literature. To our knowledge, this is the first work to systematically organize these models under a unified rIG-based framework.

Specifically, model M_1 assumes no random effects across units, implying a common dispersion parameter ($\gamma_1 = \cdots = \gamma_n$). In this case, the variability in degradation can be sufficiently explained by the scale-specific processes $\mathcal{X}_i(t)$ and $\mathcal{Z}_i(u)$ alone. This model is suited for highly standardized systems with minimal unit-level differences, such as precision electronic components or transistor arrays tested under controlled laboratory conditions, where environmental and manufacturing variability can be considered negligible. Notably, the two-scale Wiener process model proposed in Zhai et al. (2023) adopts a similar additive structure to describe degradation across age and usage scales. Their formulation models $\mathcal{X}(t)$ and $\mathcal{Z}(u)$ as independent Wiener processes with fixed parameters shared across all units. In contrast, our model builds upon this idea by extending it to a broader class of inverse Gaussian processes and further incorporating random effects to capture unit-level heterogeneity.

Model M_2 adopts the identity-trick approach to capture mild heterogeneity among

units, where variations fluctuate around a shared baseline. The scaling parameter κ controls the overall variability, while the random effect $\gamma_i \sim \mathcal{N}(1, \sigma^2)$ models deviations from the common center. This formulation is appropriate for systems with moderate batch-level variability, such as LED modules or lithium-ion battery packs across different production runs. Although similar strategies have been widely applied in single-scale gamma processes (Song and Cui, 2022; Barui et al., 2024; Zhou et al., 2023), this study is the first to extend them to the rIG-based setting.

We also introduce two single-scale degradation models, with or without random effects, denoted as M_3 and M_4. In these cases, the scale function $\Lambda(t)$ refers to either calendar time or accumulated usage. Such IG-based models are frequently used in practice (Ye and Chen, 2014; Peng, 2015), and are well-suited for systems where degradation is predominantly driven by a single mechanism. For instance, the capacity fade of lithium-ion batteries during storage is primarily governed by calendar aging, while wear degradation in components such as bearings or cutting tools is mainly usage-driven. In these scenarios, single-scale models offer a simpler structure that facilitates estimation and maintenance decision-making.

Overall, models M_1–M_4 are not merely structural reductions of the general model $[M_0]$, but are purposefully designed to reflect representative degradation behaviors encountered in engineering applications. By systematically presenting these model variants under a unified framework, we enhance modeling flexibility and provide practitioners with clear guidance for model selection and interpretation.

2.2. Theoretical properties

Based on the definition, the mean and variance of the rIG distribution are θ/γ and θ/γ^3, respectively. Utilizing the summation property of the rIG distribution, the properties of $\mathcal{Y}_i(t, u)$ in model $[M_0]$, given γ_i, are summarized in Lemma 1.

Lemma 1. *Conditioned on γ_i, $\mathcal{Y}_i(t, u) \mid \gamma_i \sim r\mathcal{IG}\left(\Lambda^t(t) + \Lambda^u(u), \gamma_i\right)$. The expectation and variance are given by*

$$\mathbb{E}\left[\mathcal{Y}_i(t, u) \mid \gamma_i\right] = \frac{\Lambda^t(t) + \Lambda^u(u)}{\gamma_i} \quad \textit{and} \quad \mathrm{Var}\left(\mathcal{Y}_i(t, u) \mid \gamma_i\right) = \frac{\Lambda^t(t) + \Lambda^u(u)}{\gamma_i^3}.$$

To derive the unconditional distribution of $\mathcal{Y}_i(t, u)$, γ_i is integrated out. The result is presented in Proposition 1, with the proof provided in supplementary Section S1.1.

Table 2: Representative models under the unified two-scale framework.

Model	Scales	Random-effects	Formulation	Reference
M_0	$Y_i(t,u)$	$\gamma_i \sim \mathcal{N}(\kappa, \sigma^2)$	$X_i(t) \mid \gamma_i \sim r\mathcal{IG}\big(\Lambda^t(t), \gamma_i\big)$ $Z_i(u) \mid \gamma_i \sim r\mathcal{IG}(\Lambda^u(u), \gamma_i)$	–
M_1	$Y_i(t,u)$	–	$X_i(t) \sim r\mathcal{IG}\big(\Lambda^t(t), \gamma\big)$ $Z_i(u) \sim r\mathcal{IG}(\Lambda^u(u), \gamma)$	• Zhai et al. (2023)
M_2	$Y_i(t,u)$	$\gamma_i \sim \mathcal{N}(1, \sigma^2)$	$X_i(t) \mid \gamma_i \sim r\mathcal{IG}\big(\Lambda^t(t), \kappa\gamma_i\big)$ $Z_i(u) \mid \gamma_i \sim r\mathcal{IG}(\Lambda^u(u), \kappa\gamma_i)$	–
M_3	$Y_i(t)$	$\gamma_i \sim \mathcal{N}(\kappa, \sigma^2)$	$Y_i(t) \sim r\mathcal{IG}(\Lambda(t), \gamma_i)$	• Song and Cui (2022) • Barui et al. (2024) • Zhou et al. (2023)
M_4	$Y_i(t)$	–	$Y_i(t) \sim r\mathcal{IG}(\Lambda(t), \gamma)$	• Ye and Chen (2014) • Peng (2015)

* All models are formulated under the unified two-scale framework. Models M_0 and M_2 are newly proposed in this study, while M_1, M_3, and M_4 correspond to related forms previously discussed in the literature.

Proposition 1. *The unconditional distribution of $\mathcal{Y}_i(t,u)$ is given by*

$$f_{\mathcal{Y}_i(t,u)}(y) = y^{-3/2}\frac{\Lambda_i(t,u)}{\sqrt{2\pi(1+y\sigma^2)}}\exp\left[-\frac{\kappa^2 y - 2\kappa\Lambda_i(t,u) + \Lambda_i(t,u)^2/y}{2(y\sigma^2+1)}\right], \tag{3}$$

where $\Lambda_i(t,u) = \Lambda_i^t(t) + \Lambda_i^u(u)$. Based on Lemma 1 and by applying the laws of total expectation and total variance, the unconditional mean and variance of $\mathcal{Y}_i(t,u)$ are given as follows:

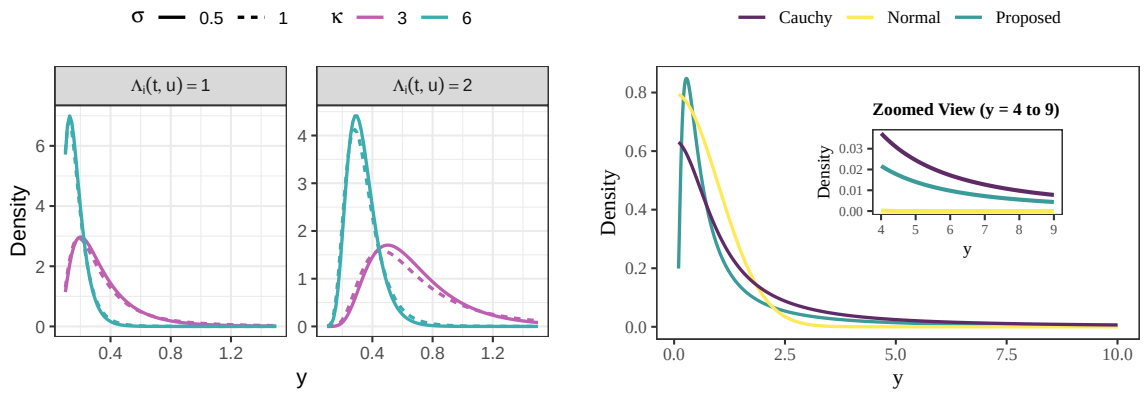
$$\begin{aligned}\mathbb{E}[\mathcal{Y}_i(t,u)] &= \Lambda(t,u)\,\mathbb{E}(\gamma_i^{-1}),\\ \mathrm{Var}[\mathcal{Y}_i(t,u)] &= \Lambda(t,u)\,\mathbb{E}(\gamma_i^{-3}) + \Lambda(t,u)^2\Big\{\mathbb{E}(\gamma_i^{-2}) - [\mathbb{E}(\gamma_i^{-1})]^2\Big\}.\end{aligned} \tag{4}$$

The shared random effect γ_i induces a nonzero cross-scale covariance and hence statistical dependence between the two components.

$$\begin{aligned}\mathrm{Cov}\big(\mathcal{X}_i(t), \mathcal{Z}_i(u)\big) &= \Lambda^t(t)\Lambda^u(u)\ \mathrm{Var}\big(\gamma_i^{-1}\big),\\ \mathrm{Corr}\big(\mathcal{X}_i(t), \mathcal{Z}_i(u)\big) &= \frac{\sqrt{\Lambda^t(t)\Lambda^u(u)}\ \mathrm{Var}(\gamma_i^{-1})}{\sqrt{\big[\mathbb{E}(\gamma_i^{-3}) + \Lambda^t(t)\,\mathrm{Var}(\gamma_i^{-1})\big]\big[\mathbb{E}(\gamma_i^{-3}) + \Lambda^u(u)\,\mathrm{Var}(\gamma_i^{-1})\big]}}.\end{aligned} \tag{5}$$

According to Proposition 1, the introduction of the random effect γ_i explicitly captures unit-level heterogeneity and mediates the interaction between the two degradation scales.

However, since the inverse moments $\mathbb{E}(\gamma_i^{-r})$ $(r = 1, 2, 3)$ diverge at $\gamma_i = 0$, these expectations are not mathematically defined when $\gamma_i \sim \mathcal{N}(\kappa, \sigma^2)$. To address this, we approximate these expectations using Monte Carlo (MC) simulation. Specifically, when κ/σ is large, most γ_i samples are positive, making this approach viable. A large number of positive γ_i samples are generated from $\mathcal{N}(\kappa, \sigma^2)$, and their reciprocal sample means are used as approximations.



(a) PDF of $\mathcal{Y}_i(t, u)$ under different parameters (b) Comparison among PDFs

Figure 2: Properties of the unconditional PDF of $\mathcal{Y}_i(t, u)$, and comparison with other baseline distributions.

Figure 2 illustrates the properties of the unconditional PDF of $\mathcal{Y}_i(t, u)$ under various parameter settings. In Figure 2(a), the standard deviation parameter σ is represented by line type, the random-effect mean κ by color, and each panel corresponds to a different value of the drift function $\Lambda_i(t, u)$. These plots demonstrate how the proposed distribution adapts flexibly to different degradation patterns by adjusting the scale and heterogeneity parameters. To further examine the tail behavior, Figure 2(b) compares the proposed distribution with the standard truncated normal and truncated Cauchy distributions, whose tails decay as $\exp(-y^2)$ and y^{-2}, respectively. The proposed distribution exhibits a tail decay rate between these two extremes, reflecting a moderate heavy-tail property. As shown in the middle panel $(y = 4 - 9)$, the proposed density exceeds that of the truncated normal, allowing it to capture larger degradation values. Beyond this range, it decays faster than the truncated Cauchy, effectively preventing unrealistically large degradation. This balance enables the model to represent occasional large degradation events while maintaining physically reasonable tail behavior for reliability analysis.

2.3. Reliability analysis

In engineering applications, predicting a product's failure time is crucial for reliability analysis. We denote the failure threshold by $\mathcal{H}$, with the failure time distribution expressed as $P\left(T_{\mathcal{H}} < t\right) = F\left(\mathcal{Y}(t,u) \geq \mathcal{H}\right)$. To focus on population-level behavior, we omit the subscript i. In practice, engineers typically perform periodic inspections based on a product's service age t, while recording cumulative usage as $u = \varrho(t)$, a time-dependent function. This function may be linear, such as $u = \xi t$, where ξ represents a fixed usage rate (Wang et al., 2020). Alternatively, $\varrho(t)$ can be defined with ξ as a random variable to capture usage rate variability (Zhai et al., 2023). Specifying $\varrho(t)$ allows the two-scale degradation model to be reduced to a single-scale form parameterized by t. Given γ, the conditional cumulative distribution function (CDF) of failure time is:

$$F_{T_{\mathcal{H}}}\left(t \mid \gamma; \mathcal{H}, \Upsilon(t)\right) = 1 - \left\{\Phi\left[\sqrt{\mathcal{H}}\gamma - \frac{\Upsilon(t)}{\sqrt{\mathcal{H}}}\right] + e^{2\Upsilon(t)\gamma}\Phi\left[-\sqrt{\mathcal{H}}\gamma - \frac{\Upsilon(t)}{\sqrt{\mathcal{H}}}\right]\right\}, \tag{6}$$

where $\Upsilon(t) = \Lambda^t(t) + \Lambda^u(\varrho(t))$. Subsequently, by integrating out γ from 6, Proposition 2 provides the CDF of the unconditional failure time.

Proposition 2. *The uncondtional failure time distribution is given by*

$$F_{T_{\mathcal{H}}}(t; \mathcal{H}, \Upsilon(t), \kappa, \sigma^2)$$
$$= 1 - \Phi\left(\frac{-\Upsilon(t)/\sqrt{\mathcal{H}} + \kappa\sqrt{\mathcal{H}}}{\sqrt{1+\mathcal{H}\sigma^2}}\right) - \exp\left[2\Upsilon(t)\kappa + 2\Upsilon(t)^2\sigma^2\right] \times \Phi\left(\frac{-\frac{\Upsilon(t)}{\sqrt{\mathcal{H}}} - \sqrt{\mathcal{H}}\kappa - 2\Upsilon(t)\sqrt{\mathcal{H}}\sigma^2}{\sqrt{1+\mathcal{H}\sigma^2}}\right). \tag{7}$$

Given (7), the mean-time-to-failure (MTTF) can be computed as:

$$\mathcal{MTTF} = \mathbb{E}(T_{\mathcal{H}}) = \int_0^\infty 1 - F_{T_{\mathcal{H}}}(t; \Upsilon(t), \kappa, \sigma^2, \mathcal{H})dt. \tag{8}$$

The proof of (7) is provided in supplementary Section S1.2. It is worth noting that the failure time distribution for models M_2 and M_3 can be derived in a similar manner, as γ_i continues to serve as a random-effects term. In contrast, for models M_1 and M_4, the failure time distribution corresponds to a standard fixed-effects rIG process, following the structure given in (6).

3. Statistical inference

3.1. Data and likelihood

Assume that each unit i is inspected at time points $\boldsymbol{t}_i = (t_{i0}, \ldots, t_{im_i})'$, with the corresponding cumulative usage levels recorded as $\boldsymbol{u}_i = (u_{i0}, \ldots, u_{im_i})'$. The degradation mea-

surement for unit i at time t_{ij} is represented as y_{ij} for $i = 1, 2, \ldots, n$ and $j = 0, 1, \ldots, m_i$, with the initial condition $y_{i0} = 0$ at $t_{i0} = 0$. According to model $[M_0]$, y_{ij} consists of degradation from the age scale, x_{ij}, and degradation from the usage scale, z_{ij}. For simplicity, we denote the drift parameters as $\Lambda^t_{ij} = \Lambda^t(t_{ij}; \alpha_t, \beta_t)$ and $\Lambda^u_{ij} = \Lambda^u(u_{ij}; \alpha_u, \beta_u)$, with $\boldsymbol{\phi}_t = (\alpha_t, \beta_t)'$ and $\boldsymbol{\phi}_u = (\alpha_u, \beta_u)'$ as the respective parameter vectors.

Define $\boldsymbol{\Delta y_i} = (\Delta y_{i1}, \ldots, \Delta y_{im_i})'$, where each increment $\Delta y_{ij} = y_{ij} - y_{i,j-1}$ represents the degradation increments for the i-th unit at inspection time t_{ij}. Each Δy_{ij} is composed of increments $\Delta x_{ij} = x_{ij} - x_{i,j-1}$ and $\Delta z_{ij} = z_{ij} - z_{i,j-1}$. Define the vectors $\boldsymbol{\Delta x_i} = (\Delta x_{i1}, \ldots, \Delta x_{im_i})'$ and $\boldsymbol{\Delta z_i} = (\Delta z_{i1}, \ldots, \Delta z_{im_i})'$ to represent these increments. Similarly, let $\boldsymbol{\Delta \tau}_i = (\Delta \tau_{i1}, \ldots, \Delta \tau_{im_i})'$ and $\boldsymbol{\Delta \nu}_i = (\Delta \nu_{i1}, \ldots, \Delta \nu_{im_i})'$, where $\Delta \tau_{ij} = \Lambda^t_{ij} - \Lambda^t_{i,j-1}$ and $\Delta \nu_{ij} = \Lambda^u_{ij} - \Lambda^u_{i,j-1}$ represent the changes in the drift parameters over time. The complete observed data set is denoted by $\mathbb{Q} = \{\mathbb{Q}_1, \ldots, \mathbb{Q}_n\}$, where $\mathbb{Q}_i = \{\boldsymbol{\Delta y}_i, \boldsymbol{t}_i, \boldsymbol{u}_i\}$. The model parameter vector is defined as $\boldsymbol{\Theta} = (\alpha_t, \beta_t, \alpha_u, \beta_u, \kappa, \sigma^2)'$. Based on Proposition 1, the log-likelihood function is given by

$$\begin{aligned} \ell(\boldsymbol{\Theta}; \mathbb{Q}) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \Bigg[& -\frac{1}{2}\log(2\pi) - \frac{3}{2}\log \Delta y_{ij} - \frac{1}{2}\log(1 + \Delta y_{ij}\sigma^2) + \log(\Delta \tau_{ij} + \Delta \nu_{ij}) \\ & - \frac{\kappa^2 \Delta y_{ij} - 2\kappa(\Delta \tau_{ij} + \Delta \nu_{ij}) + (\Delta \tau_{ij} + \Delta \nu_{ij})^2 / \Delta y_{ij}}{2(\Delta y_{ij}\sigma^2 + 1)} \Bigg]. \end{aligned} \tag{9}$$

Given the high dimensionality of the parameter vector $\boldsymbol{\Theta}$, directly maximizing (9) is challenging. Therefore, we employ two inference methods: an EM algorithm, detailed in Section 3.2, and a Bayesian approach, discussed in supplementary Section S5.

3.2. EM algorithm

The EM algorithm iteratively finds ML estimates when latent variables or incomplete data are involved. Here, we treat $\boldsymbol{\Delta x} = \{\boldsymbol{\Delta x}_1, \boldsymbol{\Delta x}_2, \ldots, \boldsymbol{\Delta x}_n\}$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_n)'$ as missing data, with $\{\mathbb{Q}, \boldsymbol{\Delta x}, \boldsymbol{\gamma}\}$ forming the complete dataset. Given that $\gamma_i \sim \mathcal{N}(\kappa, \sigma^2)$, $\Delta x_{ij} \mid \gamma_i \sim r\mathcal{IG}(\Delta \tau_{ij}, \gamma_i)$, and $(\Delta y_{ij} - \Delta x_{ij}) \mid (\Delta x_{ij}, \gamma_i) \sim r\mathcal{IG}(\Delta \nu_{ij}, \gamma_i)$, the complete log-likelihood function is provided by

$$\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\Delta x}, \boldsymbol{\gamma}) = \ell(\alpha_u, \beta_u; \mathbb{Q} \mid \boldsymbol{\Delta x}, \boldsymbol{\gamma}) + \ell(\alpha_t, \beta_t; \boldsymbol{\Delta x} \mid \boldsymbol{\gamma}) + \ell(\kappa, \sigma^2; \boldsymbol{\gamma}),$$

where

$$\begin{aligned}
\ell\left(\alpha_u, \beta_u; \mathbb{Q} \mid \boldsymbol{\Delta x}, \boldsymbol{\gamma}\right) = & \sum_{i=1}^{n} \sum_{j=1}^{m_i} \Bigg[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\left[\log(1-\Delta x_{ij}/\Delta y_{ij}) + \log \Delta y_{ij}\right] + \log \Delta \nu_{ij} \\
& -\frac{1}{2}\left(\gamma_i \sqrt{\Delta y_{ij} - \Delta x_{ij}} - \Delta \nu_{ij}/\sqrt{\Delta y_{ij} - \Delta x_{ij}}\right)^2 \Bigg], \\
\ell\left(\alpha_t, \beta_t; \boldsymbol{\Delta x} \mid \boldsymbol{\gamma}\right) = & \sum_{i=1}^{n} \sum_{j=1}^{m_i} \Bigg[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\log \Delta x_{ij} + \log \Delta \tau_{ij} \\
& -\frac{1}{2}\left(\gamma_i \sqrt{\Delta x_{ij}} - \Delta \tau_{ij}/\sqrt{\Delta x_{ij}}\right)^2 \Bigg], \\
\ell(\kappa, \sigma^2; \boldsymbol{\gamma}) = & \sum_{i=1}^{n} -\frac{1}{2}\left[\log(2\pi) + \log \sigma^2 + \left(\kappa - \gamma_i\right)^2/\sigma^2\right].
\end{aligned}$$

The EM algorithm alternates between the E-step and M-step to iteratively estimate $\boldsymbol{\Theta}$ until convergence. Starting with the current parameter estimates $\boldsymbol{\Theta}^{(s)}$ from the s-th iteration, we define $\Delta\varsigma_{ij} = \Delta x_{ij}/\Delta y_{ij}$ and calculate the Q-function for the $(s+1)$-th iteration as follows:

$$\begin{aligned}
& Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right) = \mathbb{E}\left[\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\Delta x}, \boldsymbol{\gamma}) \mid \mathbb{Q}, \boldsymbol{\Theta}^{(s)}\right] \\
& = \sum_{i=1}^{n} \Bigg\{ -\frac{1}{2}\left(\sum_{j=1}^{m_i} \Delta y_{ij} + \frac{1}{\sigma^2}\right) \mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i^2] + \left[\sum_{j=1}^{m_i} (\Delta \tau_{ij} + \Delta \nu_{ij}) + \frac{\kappa}{\sigma^2}\right] \mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i] \\
& \quad -\frac{3}{2}\sum_{j=1}^{m_i} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\log(1-\Delta\varsigma_{ij})] - \frac{1}{2}\sum_{j=1}^{m_i} \frac{\Delta \nu_{ij}^2}{\Delta y_{ij}} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[(1-\Delta\varsigma_{ij})^{-1}\right] \\
& \quad -\frac{3}{2}\sum_{j=1}^{m_i} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[\log(\Delta\varsigma_{ij})\right] - \frac{1}{2}\sum_{j=1}^{m_i} \frac{\Delta \tau_{ij}^2}{\Delta y_{ij}} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[{\Delta\varsigma_{ij}}^{-1}\right] \Bigg\} + \mathcal{C}_0,
\end{aligned} \tag{10}$$

where $\mathcal{C}_0 = -(M+n/2)\log(2\pi) + \sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[\log\left(\Delta\nu_{ij}\Delta\tau_{ij}\right) - 3\log\Delta y_{ij}\right] - n\log\sigma - (n\kappa^2)/(2\sigma^2)$, and $M = \sum_{i=1}^{n} m_i$. The Q-function involves six conditional expectations calculated with respect to the conditional distribution $f(\boldsymbol{\Delta x}, \boldsymbol{\gamma} \mid \mathbb{Q}, \boldsymbol{\Theta}^{(s)})$. These expectations are $\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i]$, $\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i^2]$, $\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\log(1-\Delta\varsigma_{ij})]$, $\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[(1-\Delta\varsigma_{ij})^{-1}\right]$, $\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[{\Delta\varsigma_{ij}}^{-1}\right]$, and $\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[\log(\Delta\varsigma_{ij})\right]$. We need the following results to derive the conditional expectations.

Theorem 1. *Given* $\mathbb{Q}_i$, $\gamma_i \mid \mathbb{Q}_i \sim \mathcal{N}\left(\frac{\kappa/\sigma^2 + \sum_{j=1}^{m_i}(\Delta\nu_{ij}+\Delta\tau_{ij})}{1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij}}, \frac{1}{1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij}}\right)$. *Hence, we have*

$$\begin{aligned}
\mathbb{E}_{\gamma_i|\mathbb{Q}_i}[\gamma_i] &= \frac{\kappa/\sigma^2 + \sum_{j=1}^{m_i}(\Delta\nu_{ij} + \Delta\tau_{ij})}{1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij}}, \\
\mathbb{E}_{\gamma_i|\mathbb{Q}_i}[\gamma_i^2] &= \frac{1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij} + \left(\kappa/\sigma^2 + \sum_{j=1}^{m_i}(\Delta\nu_{ij} + \Delta\tau_{ij})\right)^2}{\left[1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij}\right]^2}.
\end{aligned} \tag{11}$$

Given $\mathbb{Q}_i$, then we have:

$$f\left(\Delta\varsigma_{ij} \mid \mathbb{Q}_i\right) = \frac{\left(1-\varsigma_{i,j}\right)^{-3/2}\varsigma_{i,j}^{-3/2}\exp\left[-\frac{1}{2\Delta y_{i,j}}\left(\frac{\Delta\nu_{i,j}^2}{1-\varsigma_{i,j}}+\frac{\Delta\tau_{i,j}^2}{\varsigma_{i,j}}\right)\right]}{\int_0^1\left(1-\varsigma_{i,j}\right)^{-3/2}\varsigma_{i,j}^{-3/2}\exp\left[-\frac{1}{2\Delta y_{i,j}}\left(\frac{\Delta\nu_{i,j}^2}{1-\varsigma_{i,j}}+\frac{\Delta\tau_{i,j}^2}{\varsigma_{i,j}}\right)\right]d\varsigma_{ij}}. \tag{12}$$

We can obtain the remaining four conditional expectations related to the function $\Delta\varsigma_{ij}$, i.e., $g(\Delta\varsigma_{ij})$, as follows:

$$\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[g(\Delta\varsigma_{ij})] = \int_0^1 g(\Delta\varsigma_{ij})f(\Delta\varsigma_{ij} \mid \mathbb{Q}_i)\,\mathrm{d}\Delta\varsigma_{ij}. \tag{13}$$

The proof of Theorem 1 is provided in supplementary Section S2.1. To approximate the integral in (13), we present three methods: Trapezoidal approximation (TZ), MC integration, and Gauss-Legendre (GL) quadrature, with detailed theoretical explanations available in supplementary Section S2.3. In Section 4.3, these methods are compared, and we find that the GL quadrature method with an order of $l = 30$ performs best in balancing efficiency and accuracy. Therefore, this method is adopted to approximate the integral in this study.

Given Theorem 1, the Q-function is fully specified. The optimal solution for the $(s+1)$-th iteration in the M-step is updated as:

$$\boldsymbol{\Theta}^{(s+1)} = \arg\max_{\boldsymbol{\Theta}} Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right). \tag{14}$$

This involves computing the partial derivatives of $Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right)$ with respect to each parameter and solving the resulting equations. The parameter update formulas are provided in Theorem 2, with detailed derivations available in supplementary Section 2.2.

Theorem 2. *Starting from the M-step solution at the s-th iteration, denoted as $\boldsymbol{\Theta}^{(s)}$, the update for (14) in the next iteration is determined as follows:*

$$\kappa^{(s+1)} = \frac{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]}{n}, \text{ and } \sigma^{2(s+1)} = \frac{\sum_{i=1}^n\left(\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i^2] - 2\kappa^{(s+1)}\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]\right)}{n} + \kappa^{2(s+1)}.$$

Besides, the update of $\boldsymbol{\phi}_t^{(s+1)} = (\alpha_t^{(s+1)}, \beta_t^{(s+1)})'$, and $\boldsymbol{\phi}_u^{(s+1)} = (\alpha_u^{(s+1)}, \beta_u^{(s+1)})'$ can be implemented by solving the following functions:

$$\sum_{i=1}^n\sum_{j=1}^{m_i}\frac{\partial\Delta\tau_{ij}}{\partial\boldsymbol{\phi_t}}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i] - \frac{\Delta\tau_{ij}}{\Delta y_{ij}}\mathbb{E}_{ij|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\Delta\varsigma_{ij}^{-1}] + \frac{1}{\Delta\tau_{ij}}\right\} = \mathbf{0},$$

$$\sum_{i=1}^n\sum_{j=1}^{m_i}\frac{\partial\Delta\nu_{ij}}{\partial\boldsymbol{\phi_u}}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i] - \frac{\Delta\nu_{ij}}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[(1-\Delta\varsigma_{ij})^{-1}] + \frac{1}{\Delta\nu_{ij}}\right\} = \mathbf{0}.$$

Given initial values $\boldsymbol{\Theta}^{(0)}$, the EM algorithm iterates until a predefined convergence criterion is satisfied. Specifically, convergence is achieved when $|\boldsymbol{\Theta}^{(s+1)} - \boldsymbol{\Theta}^{(s)}| < \epsilon$, where $|\cdot|$ represents the absolute difference and ϵ is a preset tolerance. The final parameter estimates are denoted as $\hat{\boldsymbol{\Theta}}$. The steps of the EM algorithm used in this study are summarized in Algorithm 1. Details on implementing the EM algorithm for model variants are provided in supplementary Section S3. To improve convergence speed, selecting an effective starting point is crucial. Initial estimates can be obtained by treating each degradation path as an independent instance of a basic nonlinear rIG process. Guidelines for initializing the algorithm are discussed in supplementary Section S2.4.

Algorithm 1: Implementation of the proposed EM algorithm.

Input: $\mathbb{Q}$, $\boldsymbol{\Theta}^{(0)}$, ϵ.

Output: $\hat{\boldsymbol{\Theta}} = (\hat{\alpha}_t, \hat{\gamma}_t, \hat{\alpha}_u, \hat{\gamma}_u, \hat{\kappa}, \hat{\sigma}^2)'$.

while $|\boldsymbol{\Theta}^{(s+1)} - \boldsymbol{\Theta}^{(s)}| \geq \epsilon$ **do**

E-step:

Compute $\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]$, $\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i^2]$, and $\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[g(\Delta\varsigma_{ij})]$ by Theorem 1;

M-step:

Update $\boldsymbol{\Theta}^{(s+1)}$ by Theorem 2.

end

3.3. Interval estimation

Beyond point estimation, establishing interval estimates for the model parameters $\boldsymbol{\Theta}$ is essential. These intervals account for uncertainty and variability in the estimates, offering a plausible range for the parameters, which helps practitioners evaluate the precision and reliability of the results. A confidence interval (CI) based on the bootstrap method can be used, as outlined in Algorithm S1 in supplementary Section S4. With B bootstrap estimates $\{\hat{\boldsymbol{\Theta}}_1^*, \ldots, \hat{\boldsymbol{\Theta}}_B^*\}$ in hand, an approximate $100(1-\alpha)\%$ bootstrap CI for a function of the parameters $h(\boldsymbol{\Theta})$ can be constructed. This interval is given by: $[h(\hat{\boldsymbol{\Theta}}^*)_{(\alpha B/2)}, h(\hat{\boldsymbol{\Psi}}^*)_{((1-\alpha/2)B)}]$ where $h(\hat{\boldsymbol{\Theta}}^*)_{(b)}$ denotes the b-th order statistic from $\{h(\hat{\boldsymbol{\Theta}}^*)_1, \ldots, h(\hat{\boldsymbol{\Theta}}^*)_B\}$.

4. Simulation studies

To evaluate the proposed methodology, we conduct MC simulations using degradation data generated from model $[M_0]$. Both the EM algorithm and Bayesian methods are applied

to estimate the model parameters. To examine the effects of sample size and measurement frequency, we consider configurations with $n = 10, 30, 50$ and $m = 20, 40$. The age-scale and usage-scale functions are defined as $\Lambda^t(t; \alpha_t, \beta_t) = \beta_t t^{\alpha_t}$ and $\Lambda^u(u; \alpha_u, \beta_u) = \beta_u u^{\alpha_u}$, respectively. The parameters are set to $\alpha_t = 1.2$, $\alpha_u = 0.8$, $\beta_t = 4$, $\beta_u = 1$, $\kappa = 5$, and $\sigma^2 = 0.5$.

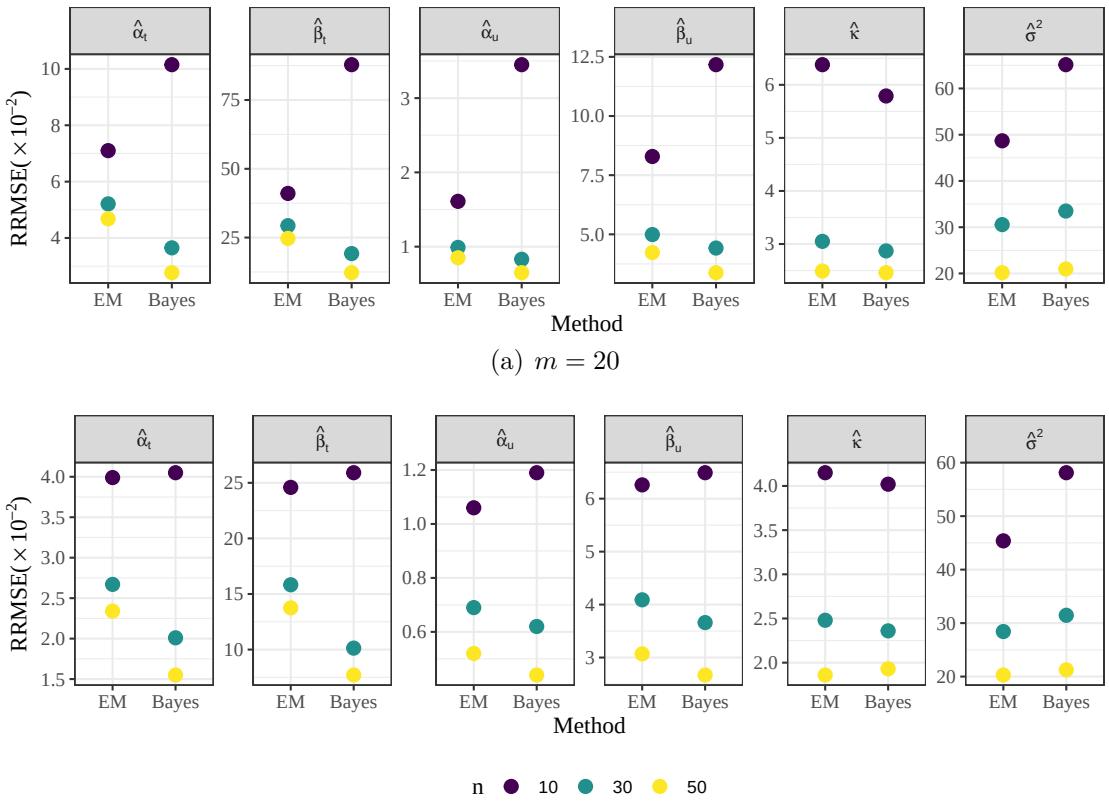
Degradation values are generated at uniform time intervals $\Delta t_{ij} = t_{ij} - t_{i,j-1} = 1$, while the usage scale increment is defined as $\Delta\varpi_{ij} = \varrho(t_{ij}) - \varrho(t_{i,j-1}) = \xi_j \Delta t_{ij}$. To reflect stochastic usage patterns observed in practice, the usage rate ξ_j follows an independent and identically distributed $r\mathcal{IG}(10, 1)$ distribution. Note that the model does not rely on equally spaced inspections nor on a fixed number of inspection points. Results for irregular inspection schedules are included in the supplementary material. For each configuration, we perform 200 replications, applying both inference methods on a laptop with an Apple M2 Pro CPU. The accuracy of parameter estimation is evaluated by calculating the relative bias (RB) and relative root mean square error (RRMSE) across combinations of n and m. These metrics are defined as follows:

$$\mathrm{RB}(\hat{\vartheta}) = \frac{1}{200}\sum_{i=1}^{200}\left(\frac{\hat{\vartheta}_i - \vartheta}{\vartheta}\right), \text{ and } \mathrm{RRMSE}(\hat{\vartheta}) = \left[\frac{1}{200}\sum_{i=1}^{200}\left(\frac{\hat{\vartheta}_i - \vartheta}{\vartheta}\right)^2\right]^{1/2},$$

where ϑ denotes the true parameter value, and $\hat{\vartheta}_i$ represents the estimate obtained from the i-th replication. Smaller values indicate greater estimation accuracy.

4.1. Validation of model parameter estimators

For the ML method, point estimates are obtained using the EM algorithm described in Section 3.2. The algorithm iterates until the convergence criterion $|\boldsymbol{\Theta}^{(s+1)} - \boldsymbol{\Theta}^{(s)}| < 10^{-4}$ is satisfied, typically stabilizing within a reasonable number of iterations. For the Bayesian method, posterior samples of $\boldsymbol{\Theta}$ are generated via the HMC algorithm (details in supplementary Section S3). To reflect limited prior knowledge, we adopt weakly informative priors for the Bayesian inference: $(\kappa, \sigma^2) \sim \mathrm{NIGa}(10, 10^{-6}, 10^{-3}, 10^{-3})$, and independent normal priors $\mathcal{N}(1, 10^6)$ for $(\alpha_t, \alpha_u, \beta_t, \beta_u)$. Although the normal priors on β_t and β_u assign small probability mass to negative values, our sensitivity analysis using truncated-normal priors enforcing $\beta_t, \beta_u > 0$ (reported in supplementary Section S6.4) shows that the posterior estimates are very similar. After a burn-in of $\mathcal{L} = 1000$ iterations, an additional 4000 iterations are performed to obtain posterior samples, and parameter estimates are calculated as the posterior means of these samples.



(a) $m = 20$

(b) $m = 40$

Figure 3: RRMSE of estimates for both methods across various sample sizes and measurement frequencies.

Figure 3 illustrates the RRMSE of parameter estimates across various sample sizes (n) and measurement frequencies (m) for both inference methods. The results show that RRMSE decreases as n or m increases, highlighting that larger samples improve estimation accuracy. Both the Bayesian and EM methods exhibit comparable performance, with similar RRMSE values for most parameters. Notably, as n increases, more information about the random effects becomes available, significantly enhancing the estimation accuracy of random-effect parameters (e.g., κ and σ^2). On the other hand, when m and n are relatively small, the accuracy of other parameters is more impacted.

To further evaluate the method, we analyze the influence of degradation curvature (α) in the age scale t^α on estimation performance. One curvature is fixed at 1, while the other varies across 0.7, 1, and 1.3, representing concave, linear, and convex degradation trends, respectively. We also investigate the effects of different degradation scale ratios (smaller, equal, and larger) and sample heterogeneity. In addition, estimation accuracy is assessed

across the four model structures M_1–M_4. These analyses consistently demonstrate that the proposed inference procedures are both reliable and efficient. Furthermore, the interval estimation performance of the Bayesian and bootstrap approaches has also been evaluated through simulation. The Bayesian intervals remain close to the nominal level in small-sample settings, whereas the bootstrap intervals may exhibit slight undercoverage when n is small. Detailed results are provided in supplementary Section S6.5.

4.2. Validation of reliability estimation

To assess reliability estimation performance, we calculate the MTTF based on the estimated parameters and a failure threshold of 30 (i.e., $\mathcal{H} = 30$). The MTTF is computed for each simulated sample and compared to the true MTTF. Figure 4 shows box plots of the RRMSE for MTTF estimates across various sample sizes and measurement frequencies. Both inference methods produce MTTF estimates with low RRMSE values, which decrease with larger sample sizes, reducing outliers.

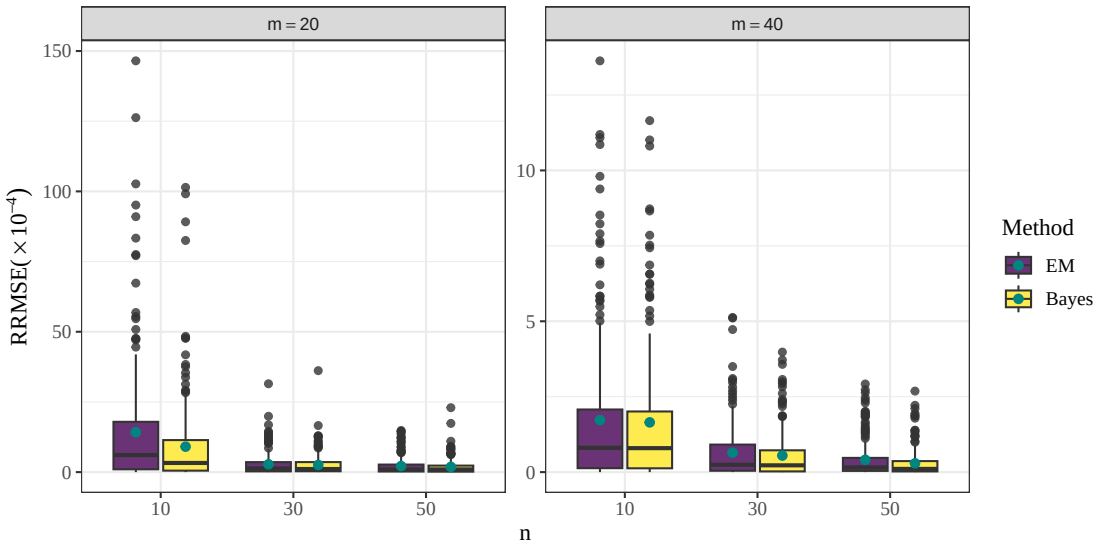
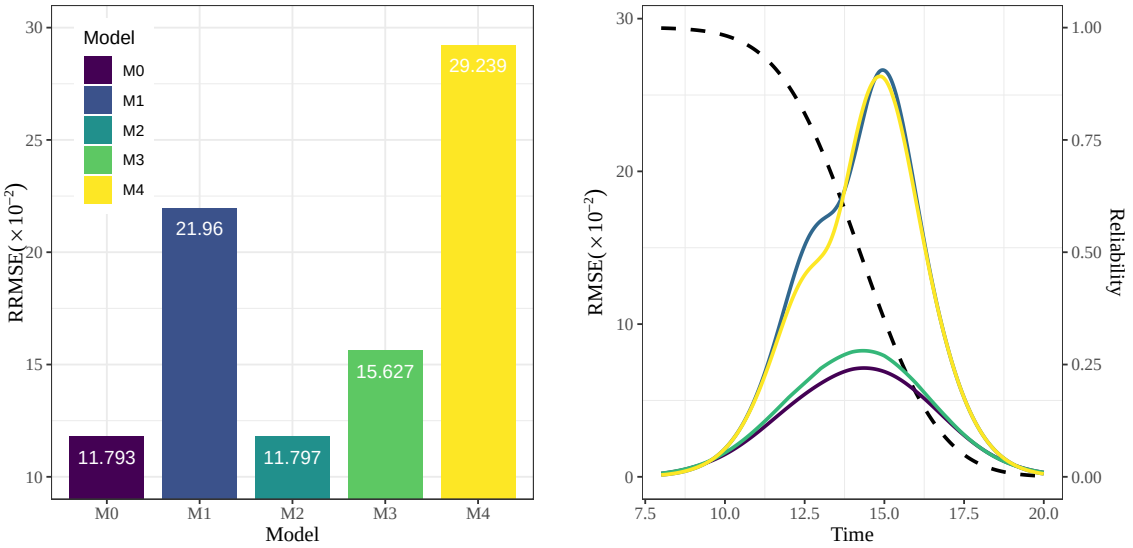


Figure 4: RRMSE of MTTF estimators for different sample sizes and measurement frequencies (green points indicate average RRMSEs).

Additionally, we evaluate the effect of model misspecification on reliability estimates. Using the parameter settings in Section 4, simulated data are generated, and different models are applied for parameter estimation. Reliability at various time points is calculated for each model and compared to the true reliability. Figure 5(a) presents the RRMSE of MTTF estimates across 200 simulations ($n = 10$ and $m = 30$), while Figure 5(b) illustrates



(a) MTTF (b) Predicted reliability curve

Figure 5: Reliability evaluation of different models, including (a) the comparison of MTTF prediction errors and (b) the predicted reliability curves over time.

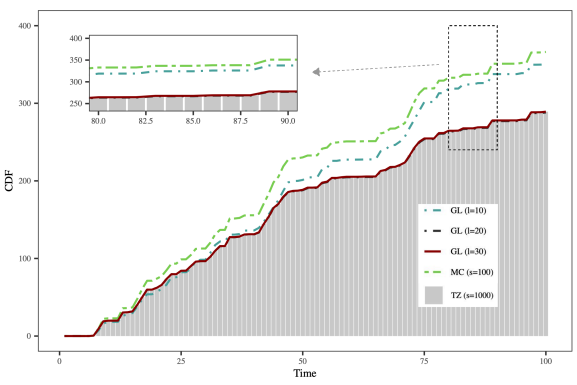
the corresponding reliability performance, with the black dashed line representing the true reliability curve. Since RRMSE may be undefined when values are zero, RMSE is used as an alternative metric here.

The figures demonstrate the importance of properly handling random effects and scales for accurate reliability estimation. Specifically, two-scale random effects models (e.g., M_0 and M_2) achieve lower RRMSE values and align more closely with the true reliability curve, indicating that incorporating random effects enhances estimation accuracy when such effects are present in the data. Conversely, fixed effects models (e.g., M_1 and M_4) yield higher estimation errors, underscoring the limitations of ignoring random effects. In terms of scale consideration, two-scale models (e.g., M_0 and M_1) outperform single-scale models (e.g., M_3 and M_4), as single-scale models lack the capacity to capture the complexity inherent in multi-scale degradation, resulting in greater bias in reliability estimates. Thus, accurately modeling two-scale random effects is essential for enhancing the precision of reliability predictions.

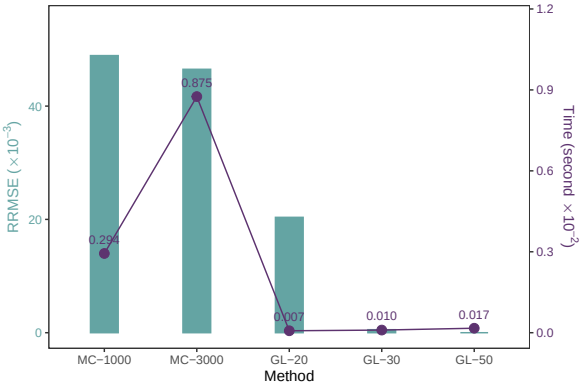
4.3. Performance of integral approximation methods

In this section, we evaluate the accuracy and efficiency of different approximation methods using (12) as an example, with a focus on $f(\Delta\varsigma_{ij} = 0.3 \mid \mathbb{Q}_1)$, where the denominator

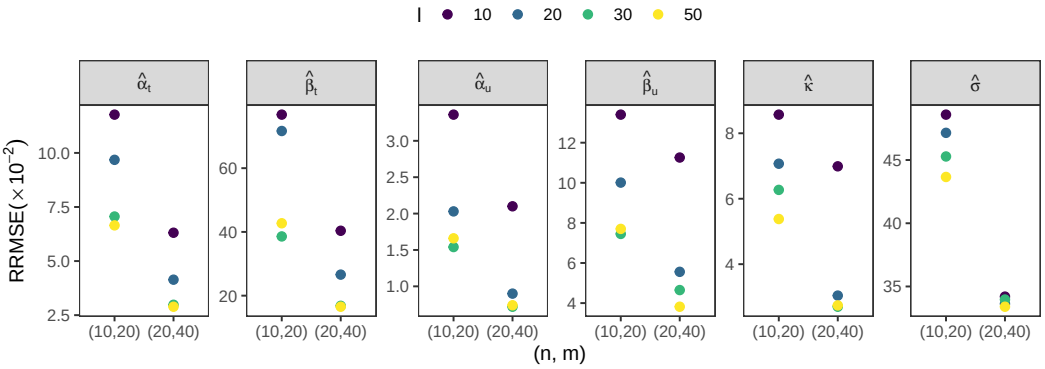
involves an expectation over $\Delta_{\varsigma_{ij}}$. Figure 6(a) shows the empirical PDF along with results from different approximation methods. The results show that the GL method with $l = 20$ and $l = 30$ closely matches the TZ method, which divides the interval $[0, 1]$ into 1000 subintervals and serves as a near-exact benchmark. However, the GL method at $l = 10$ performs poorly, and the MC method exhibits even lower accuracy. Using the TZ approximation as a reference, we compute the RRMSE and average computation time for each method (see Figure 6(b)). The GL method demonstrates higher efficiency compared to the MC method, achieving similar accuracy to the TZ method at $l = 30$ and $l = 50$.



(a) Empirical CDF of the conditional distribution



(b) computation and performance



(c) Performence of EM algorithm

Figure 6: Comparison of different approximation methods in terms of fitting accuracy and computational efficiency, including (a) empirical CDFs of the conditional distribution, (b) the trade-off between computation time and accuracy, and (c) EM algorithm performance for parameter estimation.

We further apply the GL method with varying values of l in the E-step of the EM algorithm, specifically to compute $\mathbb{E}_{\Delta_{\varsigma_{ij}}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[g(_{ij})]$. Through 100 simulations, we compare parameter estimates with true values, with RRMSE results shown in Figure 6(c) for different

sample sizes. The findings indicate that as l increases, the performance of the EM algorithm improves. For $(n, m) = (10, 20)$, the runtime for a single EM iteration with $l = 10, 20, 30, 50$ is 10.66, 53.93, 95.72, and 159.35 seconds respectively. Although the EM algorithm converges quickly with $l = 10$, it yields lower predictive accuracy. In contrast, $l = 30$ achieves nearly the same accuracy as $l = 50$ while reducing computation time by nearly half. Therefore, to balance accuracy and efficiency, we adopt the GL method with $l = 30$ for approximating the complex integral in this study.

5. Case study

In this section, we illustrate the implementation of the proposed methodology by analyzing the coating material degradation dataset as shown in Figure 1. We begin by fitting the dataset using model M_0 and its variants, where the two drift components follow the power-law forms $\Lambda^t(t) = \beta_t t^{\alpha_t}$ and $\Lambda^u(u) = \beta_u u^{\alpha_u}$. Statistical inference is performed using the corresponding EM or ML methods. Model convergence is evaluated through parameter estimate trajectories, as shown in Figure 7. This figure illustrates the iteration process of the EM algorithm for M_0, demonstrating that each parameter converges after a sufficient number of iterations. Similar convergence behavior for other models can be found in supplementary Section S7. Although Bayesian methods can also be applied for parameter estimation, they are not included here as they yield comparable results.

To further evaluate the practical value of the proposed two-scale rIG model, we also include several representative baseline models for comparison. (i) A single-scale reparameterized gamma process (RGa) (Zhou et al., 2023), which employs a linear drift function and introduces unit-specific random effects on the drift rate through a gamma distribution. In our analysis, this model is applied separately to the time and UV scales, denoted RGa(t) and RGa(u). (ii) A two-independent-scale gamma model (2IS-Ga), in which degradation is modeled as the sum of two independent gamma processes corresponding to the time and UV scales. Each process adopts the same power-law drift structure as in our formulation and both share a common shape parameter to ensure comparability across scales. All baseline models are estimated using the EM algorithm.

5.1. Model fitting and comparison

Tables 3-4 summarize the parameter estimates, log-likelihood values, and AIC results for all candidate models. For each specification, the log-likelihood is computed by substi-

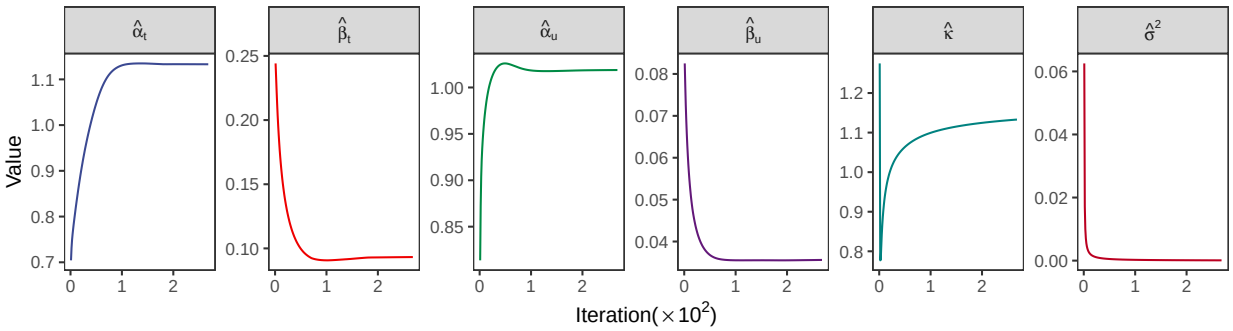


Figure 7: Iteration process of EM algorithm for parameters under model M_0.

Table 3: Parameter point estimates regarding the coating material data.

Type	Model	α_t	β_t	α_u	β_u	γ	κ	σ	Loglik	AIC
Two-scale	M_0	1.133	0.093	1.019	0.036	–	1.133	0.010	-262.799	537.598
	M_1	1.137	0.091	1.018	0.037	1.191	–	–	-262.573	**535.146**
	M_2	1.137	0.091	1.018	0.037	–	1.190	0.013	-262.589	537.178
Single-scale	M_3 (t)	1.010	0.057	–	–	–	1.193	0.016	-310.149	628.298
	M_3 (u)	–	–	1.111	0.247	–	0.917	0.037	-278.794	565.588
	M_4 (t)	1.086	0.277	–	–	0.912	–	–	-309.690	625.380
	M_4 (u)	–	–	0.980	0.070	1.171	–	–	-278.108	562.216

Table 4: AIC comparison of candidate models.

Model	Two-scale				Single-scale	
	M_0	M_1	M_2	2IS-Ga	RGa(t)	RGa(u)
AIC	537.60	**535.15**	537.18	535.52	577.02	597.63

tuting the fitted parameters into (9), and the AIC is obtained as AIC $= 2\zeta - 2\ell$, where ℓ and ζ denote the log-likelihood and the number of parameters, respectively. Among all models, M_1 attains the lowest AIC, indicating the best overall fit for the coating dataset. The 2IS-Ga model produces an AIC value very close to that of M_1, largely because both models employ a two-scale power-law drift structure and therefore reproduce the global degradation trend similarly well. The key difference lies in their assumed increment distributions: 2IS-Ga imposes gamma-distributed increments with relatively light tails, whereas

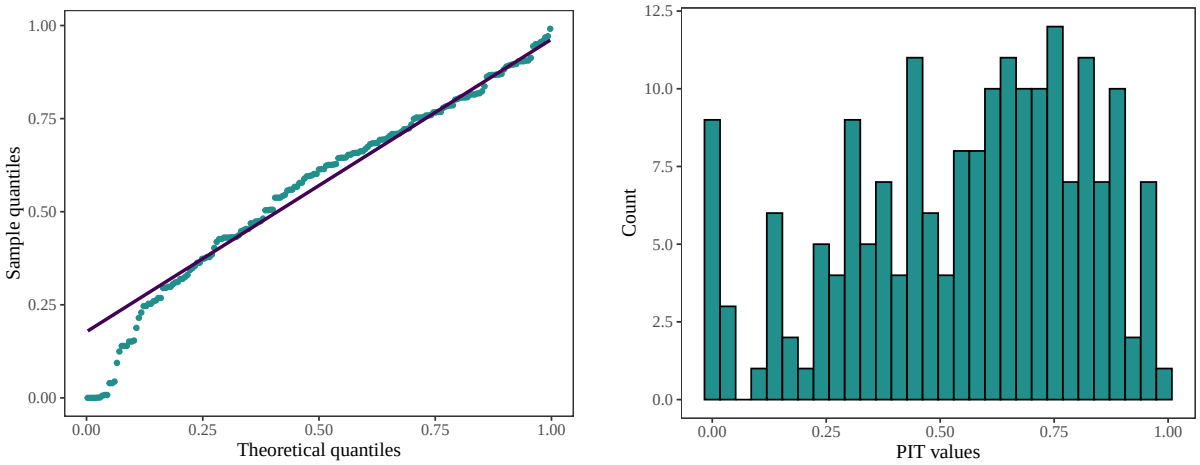
the proposed two-scale rIG process allows for heavier-tailed variation, which better reflects the local fluctuations observed in the coating data. For the random-effects models M_0 and M_2, the estimated values of σ are near zero, suggesting limited between-sample heterogeneity and supporting the simpler fixed-effect formulation adopted in M_1. In comparison, all single-scale models exhibit substantially larger AIC values, as modeling only the time or the UV scale is insufficient to capture their joint influence. The RGa(t) and RGa(u) models, which assume a linear drift form, perform even less favorably for this dataset.

The parameter estimates from the selected model M_1 provide interpretable insights into the coating degradation mechanism. The time-scale parameters (α_t, β_t) characterize calendar-time aging (e.g., oxidation and gradual chain scission), with $\alpha_t = 1.137$ indicating mildly accelerating degradation over time. Although the UV-scale parameters (α_u, β_u) are numerically smaller, this does not imply a weaker UV effect. In this experiment, the accumulated UV dose grows much faster than calendar time (e.g., when $t = 1$, the corresponding UV dose can already be on the order of several hundred units), so the UV-driven drift $\beta_u u^{\alpha_u}$ dominates $\beta_t t^{\alpha_t}$. This is consistent with established photo-oxidation mechanisms and is visually supported by the two-scale decomposition in Fig. 9. To quantify estimation uncertainty for M_1, we report percentile bootstrap 95% CIs based on 500 resamples: $\alpha_t \in [0.761, 1.263]$, $\beta_t \in [0.052, 0.220]$, $\alpha_u \in [0.945, 1.136]$, $\beta_u \in [0.017, 0.062]$, and $\gamma \in [1.066, 1.360]$. Under the two-scale decomposition, we can also examine the empirical dependence between the estimated time-driven and UV-driven components. Although both time and UV exposure contribute to degradation, the estimated cross-scale dependence is very small (near zero). This is plausible because calendar-time aging reflects slow, relatively uniform chemical oxidation and structural relaxation, whereas UV-induced degradation is driven primarily by the externally applied UV dose, which increases deterministically with the exposure schedule and is unrelated to the specimen's intrinsic aging state.

5.2. Model adequacy assessment

To assess whether the selected model provides an adequate fit to the coating degradation data, we conduct increment-level goodness-of-fit diagnostics based on probability integral transform (PIT) principles. Specifically, for each observed degradation increment, we evaluate its consistency with the model-implied conditional distribution under the selected model M_1. Figure 8(a) presents the Uniform Q–Q plot of the PIT values constructed from the fitted two-scale rIG model. Most points closely follow the reference line over the

central range, indicating that the model provides an adequate description of the distributional behavior of the degradation increments. Mild deviations are observed near the lower tail, suggesting slight local discrepancies for a small subset of increments, but no systematic lack-of-fit pattern is evident. Figure 8(b) further reports the histogram of the PIT values. The distribution is broadly consistent with a Uniform$(0, 1)$ shape, reinforcing the conclusion that the assumed power-law drift functions capture the primary degradation dynamics in the coating data.

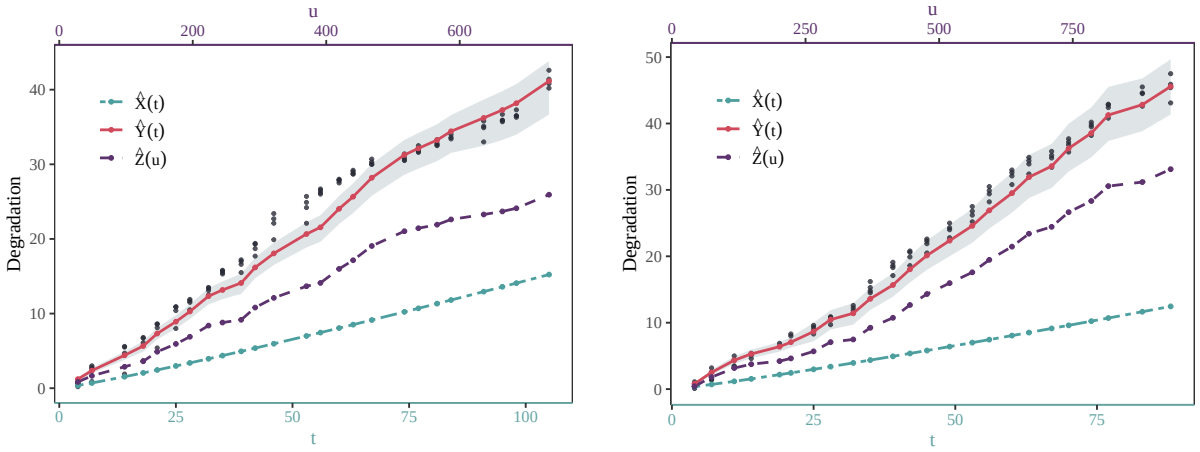


(a) Uniform Q–Q plot (b) PIT histogram

Figure 8: Increment-level goodness-of-fit diagnostics based on PIT under the selected model M_1: (a) Uniform Q–Q plot of PIT values; (b) histogram of PIT values.

Figure 9 presents the fitted degradation results for groups G9 and G15 under the proposed two-scale model. The black dots represent the observed degradation values, while the red solid curve shows the fitted overall degradation path $\hat{Y}(t)$ with its 90% confidence band constructed from 500 bootstrap replications (gray shading). The plot also decomposes the degradation into its two components: the time-driven part $\hat{X}(t)$ (teal dashed line) and the UV-driven part $\hat{Z}(u)$ (purple dotted line), illustrating how the two scales jointly contribute to the overall deterioration. As shown in the figure, the component $\hat{X}(t)$ grows slowly and approximately linearly, reflecting gradual calendar-time aging mechanisms such as thermal oxidation or structural relaxation in polymer coatings. In contrast, $\hat{Z}(u)$ increases much more rapidly, capturing UV-induced photochemical degradation (e.g., photo-oxidation or chain scission), which dominates the accelerated deterioration observed in the data. The sum of the two components yields the overall path $\hat{Y}(t)$, which closely matches the ob-

served trajectories. In addition, the 90% confidence band provides adequate coverage for most units, indicating that the proposed two-scale model effectively captures the physical degradation mechanisms of coating materials under outdoor exposure.



(a) G9 (b) G15

Figure 9: Two-scale decomposition of degradation for groups G9 and G15. Black points show observed values, the red curve is the fitted degradation path with its 90% confidence band, and the teal and purple curves represent the time- and UV-driven components.

5.3. Reliability analysis

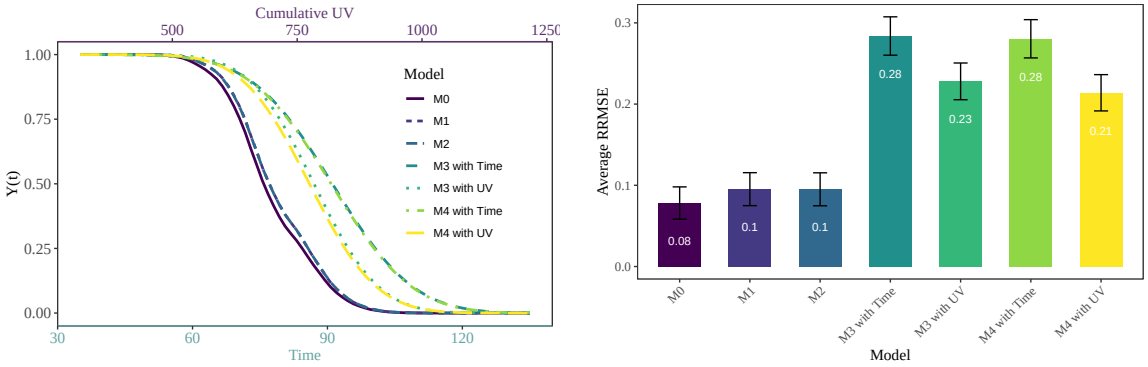
According to Hong et al. (2015); Zhai et al. (2023), the failure threshold for this coating material is $\mathcal{H} = 40$. As shown in Figure 1(b), all units fail within $t = 100$, as their degradation values exceed this threshold. Although the exact failure times are not known, we can identify the failure intervals and estimate the failure times using linear interpolation. Assuming the degradation value at time t_i is y_i and at time t_{i+1} is y_{i+1}, where $y_i < \mathcal{H}$ and $y_{i+1} > \mathcal{H}$, the interpolated failure time $t_{\mathcal{H}}$ can be calculated as:

$$t_{\mathcal{H}} = t_i + \frac{\mathcal{H} - y_i}{y_{i+1} - y_i} \times (t_{i+1} - t_i). \tag{15}$$

According to the top-to-bottom order in Figure 9, the estimated failure times for the four products in group G9 are 97.84, 97.31, 97.14, and 96.76, and in group G15 are 72.77, 69.76, 70.69, and 70.26.

Using the estimated failure times for each unit, we analyze the reliability predictions across different models. Since the results for G9 are similar to those for G15, we use G15 as

an example here. Figure 10(a) shows the reliability curves for the G15-11 coating material under various models, while Figure 10(b) compares the average RRMSE of four units with their estimated failure times calculated by (15), where error bars represent RRMSE deviations. The results indicate that models incorporating both time and UV exposure provide a more accurate depiction of the degradation process. Two-scale models (M_0, M_1, and M_2) capture the combined effects of time and UV exposure, leading to lower prediction errors and improved accuracy. In contrast, single-scale models, particularly calendar-time-based ones (M3/M4 with Time), tend to overestimate reliability, resulting in overly optimistic lifespan predictions. Thus, relying solely on a single degradation scale could misjudge material reliability and increase the risk of premature failure in high-UV environments.



(a) Reliability curves (b) Prediction performance evaluation

Figure 10: Reliability curves and prediction performance comparison with estimated failure times for the G15 coating material under different models.

6. Conclusion and discussion

In this study, we propose a novel two-scale rIG process model specifically designed for products with monotonic degradation paths. Based on physical mechanisms, the model integrates two independent degradation scales—age and usage—allowing it to more accurately describe degradation processes influenced by both time and operational demands. Additionally, we introduce heterogeneity by assigning unique model parameters to each unit, capturing differences between products. The model provides an explicit expression for failure time and describes its corresponding properties. We propose two statistical inference methods: an ML-based approach with bootstrap intervals and a Bayesian approach using

HMC sampling. These two strategies are complementary: the ML procedure is computationally efficient and provides reliable point estimates, whereas the Bayesian method offers uncertainty quantification and typically yields more stable interval estimates, especially in small-sample or higher-noise settings. This dual implementation allows practitioners to balance computational cost and uncertainty quantification according to the application context. Comprehensive simulations demonstrate the practical effectiveness of the proposed model. Key findings include: a) Increasing the sample size significantly improves the accuracy of point estimates, yielding more precise parameter values; b) The proposed estimation methods consistently produce reliable results across different settings, including changes in degradation curvature, degradation scale ratios, and sample heterogeneity; c) Ignoring either scale or random effects introduces substantial bias in MTTF estimates. Finally, we illustrate the practical applicability of the proposed approach through a case study. Note that this generalized model framework includes a range of specific variants. To improve accessibility and usability, we develop an R package incorporating the proposed model and its extensions.

While the proposed framework effectively captures two-scale monotonic degradation, several limitations and extensions deserve further investigation. First, the model assumes an additive structure over time and usage. In some systems, cross-scale interactions may be more complex, including nonlinear coupling or compensatory effects (e.g., intensive usage triggering more frequent maintenance and partially offsetting time-driven aging). Capturing such behaviors requires relaxing the additive assumption, for example via an interaction-augmented drift specification $\Lambda(t,u) = \Lambda^t(t) + \Lambda^u(u) + \beta_{tu} g(t,u)$ with $g(t,u) = t^{\alpha_t} u^{\alpha_u}$, or a link-based additive formulation $h\big(\mathcal{Y}_i(t,u)\big) = \mathcal{X}_i(t) + \mathcal{Z}_i(u)$, where $h(\cdot)$ is a monotone invertible link (e.g., log or Box–Cox). While these extensions increase structural flexibility, they fundamentally alter the stochastic representation of degradation by breaking additivity and conditional independence. As a result, the rIG closure is lost, likelihood evaluation becomes more involved, and inference typically relies on numerical integration or simulation-based methods rather than closed-form expressions. Second, the baseline model uses a shared random effect for the time- and usage-driven components. Allowing separate (possibly correlated) random effects provides a more granular representation of heterogeneity across degradation channels. However, the primary challenge here is not only computational complexity, but also statistical identifiability. Distinguishing multiple sources of heterogeneity requires sufficient non-collinear variation in the (t,u) design; otherwise, channel-specific

random effects may be weakly identified or confounded with scale-specific drift parameters. In such cases, inference becomes sensitive to prior specification and typically necessitates fully Bayesian sampling strategies or carefully designed experiments. Third, from a practical standpoint, future work may extend the framework to more heterogeneous datasets and additional reliability settings, such as competing risks, time-varying covariates, or reliability-driven decisions including maintenance optimization and warranty planning. The accompanying R package provides a modular basis for such extensions.

Data availability statement

The coating material degradation data that support the findings of this study are openly available in Gu et al. (2009).

Supplementary materials

Supplementary document: (S1) Proof of propositions; (S2) Technical details of the EM algorithm, covering conditional expectations, first-order partial derivatives of the Q-function, integral approximation methods, and initial parameter estimates; (S3) Statistical inference for model extensions; (S4) Technical details of the bootstrap method; (S5) Bayesian inference, including Bayesian formulation and Gibbs and HMC sampling algorithms; (S6) and (S7) Additional simulation experiments and case study. **Source Codes**: the R codes for this paper are provided.

References

Asgari, A., Si, W., Yuan, L., Krishnan, K., Wei, W., 2024. Multivariable degradation modeling and life prediction using multivariate fractional Brownian motion. Reliability Engineering & System Safety 248, 110146. doi:10.1016/j.ress.2024.110146.

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.

Fan, T.H., Dong, Y.S., Peng, C.Y., 2023. A complete Bayesian degradation analysis based on inverse gaussian processes. IEEE Transactions on Reliability 73, 536–548. doi:10.1109/TR.2023.3304673.

Fang, G., Pan, R., 2024. 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.

Gu, X., Stanley, D., Byrd, W.E., Dickens, B., Vaca-Trigo, I., Meeker, W.Q., Nguyen, T., Chin, J.W., Martin, J.W., 2009. Linking accelerated laboratory test with outdoor performance results for a model epoxy coating system, in: Service Life Prediction of Polymeric Materials: Global Perspectives, Springer. pp. 3–28.

Hong, Y., Duan, Y., Meeker, W.Q., Stanley, D.L., Gu, X., 2015. Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. Technometrics 57, 180–193. doi:10.1080/00401706.2014.915891.

Hu, J., Sun, Q., Ye, Z.S., Zhou, Q., 2020. Joint modeling of degradation and lifetime data for RUL prediction of deteriorating products. IEEE Transactions on Industrial Informatics 17, 4521–4531. doi:10.1109/TII.2020.3021054.

Li, Q., Li, H., Ma, Z., Liu, X., Guan, X., Zhang, X., 2023. Remaining useful life prediction of mechanical system based on improved adaptive fractional lévy stable motion with statistical dependence measurement error. Mechanical Systems and Signal Processing 200, 110646. doi:10.1016/j.ymssp.2023.110646.

Lin, C.P., Ling, M.H., Cabrera, J., Yang, F., Yu, D.Y.W., Tsui, K.L., 2021. Prognostics for lithium-ion batteries using a two-phase gamma degradation process model. Reliability Engineering & System Safety 214, 107797. doi:10.1016/j.ress.2021.107797.

Ling, M.H., Ng, H.K.T., Tsui, K.L., 2019. Bayesian and likelihood inferences on remaining useful life in two-phase degradation models under gamma process. Reliability Engineering & System Safety 184, 77–85. doi:10.1016/j.ress.2017.11.017.

Liu, H., Song, W., Zio, E., 2022. Fractional lévy stable motion with LRD for RUL and reliability analysis of li-ion battery. ISA transactions 125, 360–370. doi:10.1016/j.isatra.2021.07.002.

Lu, L., Wang, B., Hong, Y., Ye, Z., 2020. General path models for degradation data with multiple characteristics and covariates. Technometrics 63, 354–369. doi:10.1080/00401706.2020.1796814.

Pei, H., Hu, C., Si, X., Zheng, J., Zhang, Q., Zhang, Z., Pang, Z., 2019. Remaining useful life prediction for nonlinear degraded equipment with bivariate time scales. IEEE Access 7, 165166–165180. doi:10.1109/ACCESS.2019.2951804.

Peng, C.Y., 2015. Inverse Gaussian processes with random effects and explanatory variables for degradation data. Technometrics 57, 100–111. doi:10.1080/00401706.2013.879077.

Peng, C.Y., Cheng, Y.S., 2020. Student-t processes for degradation analysis. Technometrics 62, 223–235. doi:10.1080/00401706.2019.1630008.

Peng, S., Jiang, W., Huang, W., Luo, Q., 2024. The impact of gamma usage processes on preventive maintenance policies under two-dimensional warranty. Reliability Engineering & System Safety 242, 109743. doi:10.1016/j.ress.2023.109743.

Peng, W., Li, Y.F., Yang, Y.J., Huang, H.Z., Zuo, M.J., 2014. Inverse Gaussian process models for degradation analysis: A Bayesian perspective. Reliability Engineering & System Safety 130, 175–189. doi:10.1016/j.ress.2014.06.005.

Si, X.S., Wang, W., Chen, M.Y., Hu, C.H., Zhou, D.H., 2013. A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution. European Journal of Operational Research 226, 53–66. doi:10.1016/j.ejor.2012.10.030.

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.

Wang, C., Liu, J., Yang, Q., Hu, Q., Yu, D., 2023a. Recoverability effects on reliability assessment for accelerated degradation testing. IISE Transactions 55, 698–710. doi:10.1080/24725854.2022.2089784.

Wang, P., Tang, Y., 2025. Remaining useful life prediction based on exponential dispersion process with random drifts. Statistical Theory and Related Fields doi:10.1080/24754269.2025.2555043.

Wang, X., Li, L., Xie, M., 2020. An unpunctual preventive maintenance policy under two-dimensional warranty. European Journal of Operational Research 282, 304–318. doi:10.1016/j.ejor.2019.09.025.

Wang, X., Wang, B.X., Hong, Y., Jiang, P.H., 2021. Degradation data analysis based on gamma process with random effects. European Journal of Operational Research 292, 1200–1208. doi:10.1016/j.ejor.2020.11.036.

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

Wang, X.L., 2023. Design and pricing of usage-driven customized two-dimensional extended warranty menus. IISE Transactions 55, 873–885.

Wang, Z., Zhai, Q., Shen, L., 2023b. Degradation modeling and RUL prediction in dynamic environments using a Wiener process with an autoregressive rate. IEEE Transactions on Reliability 73, 912–921. doi:10.1109/TR.2023.3319497.

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., Wang, W., 2025. Recursive Bayesian prediction of remaining useful life for gamma degradation process under conjugate priors. Scandinavian Journal of Statistics .

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

Zhai, Q., Li, Y., Chen, P., 2025. Modeling product degradation with heterogeneity: A general random-effects Wiener process approach. IISE Transactions 57, 1422–1435. doi:10.1080/24725854.2024.2434125.

Zhai, Q., Xu, A., Yang, J., Zhou, Y., 2023. Statistical modeling and reliability analysis for degradation processes indexed by two scales. IEEE Transactions on Industrial Informatics 20, 3675–3684. doi:10.1109/TII.2023.3313668.

Zhang, H., Chen, M., Shang, J., Yang, C., Sun, Y., 2021. Stochastic process-based degradation modeling and RUL prediction: from Brownian motion to fractional Brownian motion. Science China Information Sciences 64, 171201. doi:10.1007/s11432-020-3134-8.

Zhang, Z., Hu, C., Si, X., Zhang, W., 2017. Degradation modeling and remaining useful life prediction with bivariate time scale. Zidonghua Xuebao/Acta Automatica Sinica 43, 1789–98. doi:10.16383/j.aas.2017.c160509.

Zhang, Z., Si, X., Hu, C., 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. doi:10.1016/j.ejor.2018.02.033.

Zhao, X., Chen, P., Tang, L.C., 2025. Condition-based maintenance via Markov decision processes: A review. Frontiers of Engineering Management 12, 330–342.

Zhou, S., Xu, A., Tang, Y., Shen, L., 2023. Fast Bayesian inference of reparameterized gamma process with random effects. IEEE Transactions on Reliability 73, 399–412. doi:10.1109/TR.2023.3263940.

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.

Supplementary of "Modeling Two-Scale Degradation with Heterogeneity: A Unified Random-Effects Inverse Gaussian Framework"

The supplementary document is organized as follows: Section S1 contains proposition proofs. Section S2 details the EM algorithm, including conditional expectations, first-order partial derivatives of the Q-function, integral approximation, and initial parameter estimates. Section S3 discusses statistical inference for model extensions. Section S4 covers the bootstrap method. Bayesian inference, including formulation and Gibbs and HMC sampling, is in Section S5. Sections S6 and S7 provide additional simulation experiments and case study.

S1 Proof of propositions

S1.1 Proof of Proposition 1

To derive the marginal distribution of $\mathcal{Y}_i(t,u)$, we integrate out the random effect γ_i from its joint distribution with $\mathcal{Y}_i(t,u)$. Let $\Lambda_i(t,u) = \Lambda_i^t(t) + \Lambda_i^u(u)$, and recall that conditionally on γ_i, the response follows a reparameterized inverse Gaussian distribution:

$$\mathcal{Y}_i(t,u) \mid \gamma_i \sim r\mathcal{IG}(\Lambda_i(t,u), \gamma_i), \quad \gamma_i \sim \mathcal{N}(\kappa, \sigma^2).$$

The marginal density is given by:

$$\begin{aligned} f_{\mathcal{Y}_i(t,u)}(y) &= \int_{-\infty}^{\infty} f_{\mathcal{Y}_i(t,u)|\gamma_i}(y) \cdot f_{\gamma_i}(\gamma_i)\, d\gamma_i \\ &= \int_{-\infty}^{\infty} \frac{\Lambda_i(t,u)}{\sqrt{2\pi}} y^{-3/2} \exp\left[-\frac{1}{2}\left(\sqrt{y}\gamma_i - \frac{\Lambda_i(t,u)}{\sqrt{y}}\right)^2\right] \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\gamma_i-\kappa)^2}{2\sigma^2}\right] d\gamma_i. \end{aligned}$$

This integral is tractable since the integrand is the product of two Gaussian kernels. Completing the square and integrating yields:

$$f_{\mathcal{Y}_i(t,u)}(y) = y^{-3/2} \cdot \frac{\Lambda_i(t,u)}{\sqrt{2\pi(1+y\sigma^2)}} \cdot \exp\left[-\frac{\kappa^2 y - 2\kappa\Lambda_i(t,u) + \Lambda_i(t,u)^2/y}{2(y\sigma^2+1)}\right].$$

Alternatively, this result can also be obtained using a known result from Si and Zhou (2013), which provides a closed-form expression for expectations involving normal random variables within exponential-quadratic forms:

Lemma 1. *Let $\rho \sim \mathcal{N}(\kappa, \sigma^2)$, and let $\omega_1, \omega_2, A, B \in \mathbb{R}$, $C > 0$. Then:*

$$\mathbb{E}_\rho \left[(\omega_1 - A\rho) \exp\left(-\frac{(\omega_2 - B\rho)^2}{2C} \right) \right]$$
$$= \sqrt{\frac{C}{B^2\sigma^2 + C}} \left(\omega_1 - A\frac{B\sigma^2\omega_2 + \kappa C}{B^2\sigma^2 + C} \right) \exp\left(-\frac{(\omega_2 - B\kappa)^2}{2\,(B^2\sigma^2 + C)} \right).$$

Letting $\rho = \gamma_i$, $\omega_1 = y^{-3/2}\Lambda_i(t,u)/\sqrt{2\pi}$, $\omega_2 = -\Lambda_i(t,u)/y$, $A = 0$, $B = -1$, and $C = 1/y$, the expression from the lemma reduces to the same result as above.

Similarly, under model M_2, where $\mathcal{Y}_i(t,u) \mid \gamma_i \sim r\mathcal{IG}(\Lambda_i(t,u), \kappa\gamma_i)$ and $\gamma_i \sim \mathcal{N}(1, \sigma^2)$, the marginal distribution becomes:

$$f_{\mathcal{Y}_i(t,u)}(y) = y^{-3/2} \cdot \frac{\Lambda_i(t,u)}{\sqrt{2\pi(1 + y\kappa^2\sigma^2)}} \cdot \exp\left[-\frac{\kappa^2 y - 2\kappa\Lambda_i(t,u) + \Lambda_i(t,u)^2/y}{2(y\kappa^2\sigma^2 + 1)} \right].$$

We now derive several unconditional properties of the proposed two-scale rIG degradation process. By the law of total expectation,

$$\mathbb{E}[\mathcal{Y}_i(t,u)] = \mathbb{E}\big[\mathbb{E}(\mathcal{Y}_i(t,u) \mid \gamma_i)\big] = \Lambda(t,u)\,\mathbb{E}(\gamma_i^{-1}).$$

Similarly, the law of total variance yields

$$\begin{aligned}
\mathrm{Var}[\mathcal{Y}_i(t,u)] &= \mathbb{E}\big[\,\mathrm{Var}(\mathcal{Y}_i(t,u) \mid \gamma_i)\big] + \mathrm{Var}\big(\mathbb{E}[\mathcal{Y}_i(t,u) \mid \gamma_i]\big) \\
&= \mathbb{E}\Big[\frac{\Lambda(t,u)}{\gamma_i^3}\Big] + \mathrm{Var}\Big(\frac{\Lambda(t,u)}{\gamma_i}\Big) \\
&= \Lambda(t,u)\,\mathbb{E}(\gamma_i^{-3}) + \Lambda(t,u)^2\Big\{\mathbb{E}(\gamma_i^{-2}) - [\mathbb{E}(\gamma_i^{-1})]^2\Big\}.
\end{aligned}$$

For the covariance, applying the law of total covariance gives

$$\mathrm{Cov}(X, Z) = \mathbb{E}\big[\,\mathrm{Cov}(X, Z \mid \gamma_i)\big] + \mathrm{Cov}\big(\mathbb{E}[X \mid \gamma_i], \mathbb{E}[Z \mid \gamma_i]\big).$$

Since $\mathcal{X}_i$ and $\mathcal{Z}_i$ are conditionally independent given γ_i, the first term is zero. Hence,

$$\mathrm{Cov}\big(\mathcal{X}_i(t), \mathcal{Z}_i(u)\big) = \mathrm{Cov}\Big(\frac{\Lambda^t(t)}{\gamma_i}, \frac{\Lambda^u(u)}{\gamma_i}\Big) = \Lambda^t(t)\Lambda^u(u)\ \mathrm{Var}(1/\gamma_i).$$

The variances of $\mathcal{X}_i(t)$ and $\mathcal{Z}_i(u)$ follow from the same argument:

$$\mathrm{Var}(\mathcal{X}_i) = \Lambda^t(t)\,\mathbb{E}(\gamma_i^{-3}) + \big[\Lambda^t(t)\big]^2\,\mathrm{Var}(1/\gamma_i), \qquad \mathrm{Var}(\mathcal{Z}_i) = \Lambda^u(u)\,\mathbb{E}(\gamma_i^{-3}) + \big[\Lambda^u(u)\big]^2\,\mathrm{Var}(1/\gamma_i).$$

Combining these results, the correlation between the two components is

$$\mathrm{Corr}(\mathcal{X}_i, \mathcal{Z}_i) = \frac{\sqrt{\Lambda^t(t)\Lambda^u(u)}\ \mathrm{Var}(1/\gamma_i)}{\sqrt{\big[\mathbb{E}(\gamma_i^{-3}) + \Lambda^t(t)\,\mathrm{Var}(1/\gamma_i)\big]\big[\mathbb{E}(\gamma_i^{-3}) + \Lambda^u(u)\,\mathrm{Var}(1/\gamma_i)\big]}}.$$

This completes the proof.

S1.2 Proof of Proposition 2

Let $\mathcal{H}$ denote the pre-specified degradation threshold, and define the failure time $T_{\mathcal{H}}$ as the first-passage time at which the degradation path exceeds the threshold:

$$T_{\mathcal{H}} = \inf\{t > 0 : \mathcal{Y}_i(t, \varrho(t)) \geq \mathcal{H}\}.$$

Let $\Upsilon(t) = \Lambda^t(t) + \Lambda^u(\varrho(t))$ denote the cumulative scale over both age and usage dimensions. Given $\gamma_i \sim \mathcal{N}(\kappa, \sigma^2)$, the conditional distribution of $T_{\mathcal{H}}$ follows the rIG distribution. Thus, the conditional CDF is:

$$F_{T_{\mathcal{H}}}(t \mid \gamma_i) = 1 - F_{rIG}(\mathcal{H}; \Upsilon(t), \gamma_i),$$

where $F_{rIG}(\cdot)$ denotes the rIG CDF.

To obtain the unconditional distribution, we integrate out γ_i:

$$F_{T_{\mathcal{H}}}(t; \mathcal{H}, \Upsilon(t), \kappa, \sigma^2) = \int_{-\infty}^{\infty} (1 - F_{rIG}(\mathcal{H}; \Upsilon(t), \gamma)) \cdot \phi(\gamma; \kappa, \sigma^2)\, d\gamma,$$

where $\phi(\cdot)$ is the density of the normal distribution $\mathcal{N}(\kappa, \sigma^2)$. Using the analytical form of $F_{rIG}(\cdot)$, we obtain:

$$\begin{aligned} F_{T_{\mathcal{H}}}(t) = 1 &- \mathbb{E}_\gamma\left[\Phi\left(\sqrt{\mathcal{H}}\gamma - \frac{\Upsilon(t)}{\sqrt{\mathcal{H}}}\right)\right] \\ &- \mathbb{E}_\gamma\left[\exp(2\Upsilon(t)\gamma) \cdot \Phi\left(-\sqrt{\mathcal{H}}\gamma - \frac{\Upsilon(t)}{\sqrt{\mathcal{H}}}\right)\right]. \end{aligned}$$

These two expectations can be evaluated using the following lemma from Si and Zhou (2013):

Lemma 2. *Let $Z \sim \mathcal{N}(\kappa, \sigma^2)$ and $D, E, F \in \mathbb{R}$, then:*

$$\mathbb{E}\left[\exp(DZ) \cdot \Phi(E + FZ)\right] = \exp\left(D\kappa + \tfrac{1}{2}D^2\sigma^2\right) \cdot \Phi\left(\frac{E + F\kappa + DF\sigma^2}{\sqrt{1 + F^2\sigma^2}}\right).$$

Applying this lemma to each term. For the first expectation, take $D = 0$, $E = -\Upsilon(t)/\sqrt{\mathcal{H}}$, $F = \sqrt{\mathcal{H}}$:

$$\mathbb{E}_\gamma\left[\Phi\left(\sqrt{\mathcal{H}}\gamma - \frac{\Upsilon(t)}{\sqrt{\mathcal{H}}}\right)\right] = \Phi\left(\frac{-\Upsilon(t)/\sqrt{\mathcal{H}} + \kappa\sqrt{\mathcal{H}}}{\sqrt{1 + \mathcal{H}\sigma^2}}\right).$$

For the second expectation, take $D = 2\Upsilon(t)$, $E = -\Upsilon(t)/\sqrt{\mathcal{H}}$, $F = -\sqrt{\mathcal{H}}$:

$$\begin{aligned} &\mathbb{E}_\gamma\left[\exp(2\Upsilon(t)\gamma) \cdot \Phi\left(-\sqrt{\mathcal{H}}\gamma - \frac{\Upsilon(t)}{\sqrt{\mathcal{H}}}\right)\right] \\ &= \exp\left(2\Upsilon(t)\kappa + 2\Upsilon(t)^2\sigma^2\right) \cdot \Phi\left(\frac{-\Upsilon(t)/\sqrt{\mathcal{H}} - \kappa\sqrt{\mathcal{H}} - 2\Upsilon(t)\sqrt{\mathcal{H}}\sigma^2}{\sqrt{1 + \mathcal{H}\sigma^2}}\right). \end{aligned}$$

Substituting both expressions gives the final result:

$$F_{T_{\mathcal{H}}}(t;\mathcal{H},\Upsilon(t),\kappa,\sigma^2) = 1-\Phi\left(\frac{-\Upsilon(t)/\sqrt{\mathcal{H}}+\kappa\sqrt{\mathcal{H}}}{\sqrt{1+\mathcal{H}\sigma^2}}\right) - \exp\left(2\Upsilon(t)\kappa+2\Upsilon(t)^2\sigma^2\right)\cdot\Phi\left(\frac{-\Upsilon(t)/\sqrt{\mathcal{H}}-\kappa\sqrt{\mathcal{H}}-2\Upsilon(t)\sqrt{\mathcal{H}}\sigma^2}{\sqrt{1+\mathcal{H}\sigma^2}}\right),$$

which completes the proof.

S2 Technical details of the EM algorithm

This section provides the technical derivations underlying the proposed expectation-maximization (EM) algorithm. In Section S2.1, we present the complete data likelihood and the conditional distributions required in the E-step. Later sections cover additional theoretical results, integral approximation techniques, and initialization strategies.

S2.1 Proof of Theorem 1

Given that $\gamma_i \sim \mathcal{N}(\kappa,\sigma^2)$, $\Delta x_{ij} \sim r\mathcal{IG}(\Delta\tau_{ij},\gamma_i)$, and $\Delta z_{ij} = \Delta y_{ij} - \Delta x_{ij} \sim r\mathcal{IG}(\Delta\nu_{ij},\gamma_i)$, the complete-data joint likelihood becomes:

$$\begin{aligned}
&f\left(\gamma_i,\boldsymbol{\Delta x}_i,\mathbb{Q}_i\right) = f\left(\mathbb{Q}_i \mid \gamma_i,\boldsymbol{\Delta x}_i\right) f\left(\boldsymbol{\Delta x}_i \mid \gamma_i\right) f\left(\gamma_i\right)\\
&= \left[\prod_{j=1}^{m_i} f\left(\Delta y_{ij} \mid \gamma_i,\Delta x_{ij}\right) f\left(\Delta x_{ij} \mid \gamma_i\right)\right] f\left(\gamma_i\right)\\
&= \left\{\prod_{j=1}^{m_i} \frac{\Delta\nu_{ij}}{\sqrt{2\pi}}(\Delta y_{ij}-\Delta x_{ij})^{-3/2}\exp\left[-\frac{1}{2}\left(\sqrt{\Delta y_{ij}-\Delta x_{ij}}\gamma_i - \frac{\Delta\nu_{ij}}{\sqrt{\Delta y_{ij}-\Delta x_{ij}}}\right)^2\right]\right.\\
&\quad \left.\times\frac{\Delta\tau_{ij}}{\sqrt{2\pi}}\Delta x_{ij}{}^{-3/2}\exp\left[-\frac{1}{2}\left(\sqrt{\Delta x_{ij}}\gamma_i - \frac{\Delta\tau_{ij}}{\sqrt{\Delta x_{ij}}}\right)^2\right]\right\} \times \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(\gamma_i-\kappa)^2}{2\sigma^2}\right] \qquad \text{(S1)}\\
&= \mathcal{C}_{1i}\times\prod_{j=1}^{m_i}\left\{\left[(\Delta y_{ij}-\Delta x_{ij})\Delta x_{ij}\right]^{-3/2}\exp\left[-\frac{1}{2}\left(\frac{\Delta\nu_{ij}^2}{\Delta y_{ij}-\Delta x_{ij}}+\frac{\Delta\tau_{ij}^2}{\Delta x_{ij}}\right)\right]\right\}\\
&\quad\times\exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2}+\sum_{j=1}^{m_i}\Delta y_{ij}\right)\left(\gamma_i-\frac{\kappa/\sigma^2+\sum_{j=1}^{m_i}(\Delta\nu_{ij}+\Delta\tau_{ij})}{1/\sigma^2+\sum_{j=1}^{m_i}\Delta y_{ij}}\right)^2\right],
\end{aligned}$$

where $\mathcal{C}_{1i} = (2\pi)^{-1/2-m_i}\sigma^{-1}\prod_{j=1}^{m_i}(\Delta\nu_{ij}\Delta\tau_{ij})\times\exp\left\{-\frac{1}{2}\left[\frac{\left(\sum_{j=1}^{m_i}(\Delta\nu_{ij}+\Delta\tau_{ij})+\kappa/\sigma^2\right)^2}{\sum_{j=1}^{m_i}\Delta y_{ij}+1/\sigma^2}+\frac{\kappa^2}{\sigma^2}\right]\right\}$ is a constant that does not depend on γ_i or Δx_{ij}. As we can see that $\gamma_i, \Delta x_{i1}, \Delta x_{i2}, \ldots, \Delta x_{im_i}$

are mutually independent. More importantly, the result implies that

$$\gamma_i \mid \mathbb{Q}_i \sim \mathcal{N}\left(\frac{\kappa/\sigma^2 + \sum_{j=1}^{m_i}(\Delta\nu_{ij} + \Delta\tau_{ij})}{1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij}}, \frac{1}{1/\sigma^2 + \sum_{j=1}^{m_i}\Delta y_{ij}}\right).$$

Next, we need to calculate the result of $f(\boldsymbol{\Delta x}_i \mid \mathbb{Q}_i)$. According to Bayes' theorem, we obtain:

$$f(\boldsymbol{\Delta x}_i \mid \mathbb{Q}_i) = \frac{f(\boldsymbol{\Delta x}_i, \mathbb{Q}_i)}{\int_0^{\Delta y_{ij}} \prod_{j=1}^{m_i} f(\Delta x_{ij}, \mathbb{Q}_i)\, \mathrm{d}\Delta x_{ij}}. \tag{S2}$$

First, we integrate over γ_i in equation (S1), yielding:

$$\begin{aligned}
f(\boldsymbol{\Delta x}_i, \mathbb{Q}_i) &= \int_{-\infty}^{\infty} f(\mathbb{Q}_i \mid \gamma_i, \Delta x_{ij})\, f(\Delta x_{ij} \mid \gamma_i)\, f(\gamma_i)\, d\gamma_i \\
&= \int_{-\infty}^{\infty} \left[\prod_{j=1}^{m_i} f(\Delta y_{ij} \mid \gamma_i, \Delta x_{ij})\, f(\Delta x_{ij} \mid \gamma_i)\right] f(\gamma_i)\, d\gamma_i \\
&= \mathcal{C}_{2i} \times \prod_{j=1}^{m_i} \left\{[(\Delta y_{ij} - \Delta x_{ij})\Delta x_{ij}]^{-3/2} \exp\left[-\frac{1}{2}\left(\frac{\Delta\nu_{ij}^2}{\Delta y_{ij} - \Delta x_{ij}} + \frac{\Delta\tau_{ij}^2}{\Delta x_{ij}}\right)\right]\right\},
\end{aligned}$$

where $\Delta x_{ij} \in [0, \Delta y_{ij}]$,

and $\mathcal{C}_{2i} = (2\pi)^{-m_i}\left(\sigma^2 \sum_{j=1}^{m_i}\Delta y_{ij} + 1\right)^{-1/2} \prod_{j=1}^{m_i}(\Delta\nu_{ij}\Delta\tau_{ij}) \exp\left\{-\frac{1}{2}\left[-\frac{\left[\sum_{j=1}^{m_i}(\Delta\nu_{ij}+\Delta\tau_{ij})+\kappa/\sigma^2\right]^2}{\sum_{j=1}^{m_i}\Delta y_{ij}+1/\sigma^2} + \frac{\kappa^2}{\sigma^2}\right]\right\}$.
Substitute the above result into (S2). At the same time, redefine the domain of the variable Δx_{ij} from $(0, \Delta y_{ij})$ to $(0, 1)$. This transformation is accomplished through a simple linear change of variable, where $\Delta\varsigma_{ij} = \Delta x_{ij}/\Delta y_{ij}$, leading to $\Delta x_{ij} = \Delta\varsigma_{ij}\Delta y_{ij}$. It is important to note that the Jacobian determinant $\mathrm{d}\Delta x_{ij}/\mathrm{d}\Delta\varsigma_{ij} = \Delta y_{ij}$ must be multiplied by the original density function. This eventually simplifies to give the density function of $\Delta\varsigma_{ij}$ as:

$$f(\Delta\varsigma_{ij} \mid \mathbb{Q}_i) = \frac{(1-\varsigma_{i,j})^{-3/2}\varsigma_{i,j}^{-3/2} \exp\left[-\frac{1}{2\Delta y_{i,j}}\left(\frac{\Delta\nu_{i,j}^2}{1-\varsigma_{i,j}} + \frac{\Delta\tau_{i,j}^2}{\varsigma_{i,j}}\right)\right]}{\int_0^1 (1-\varsigma_{i,j})^{-3/2}\varsigma_{i,j}^{-3/2} \exp\left[-\frac{1}{2\Delta y_{i,j}}\left(\frac{\Delta\nu_{i,j}^2}{1-\varsigma_{i,j}} + \frac{\Delta\tau_{i,j}^2}{\varsigma_{i,j}}\right)\right] \mathrm{d}\varsigma_{ij}}.$$

S2.2 Proof of Theorem 2

To derive the update rules in the M-step, we take the first-order partial derivatives of the Q-function with respect to each parameter in $\boldsymbol{\Theta}$ and equate them to zero. Let $\boldsymbol{\phi}_t = (\alpha_t, \beta_t)'$ and $\boldsymbol{\phi}_u = (\alpha_u, \beta_u)'$. The gradients are given by:

$$\frac{\partial Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right)}{\partial\kappa} = \frac{n\kappa}{\sigma^{2(s)}} - \frac{\sum_{i=1}^{n}\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]}{\sigma^{2(s)}} = 0, \tag{S3}$$

$$\frac{\partial Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right)}{\partial \sigma^2} = \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left[\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i^2] - 2\kappa^{(s)}\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]\right] - \frac{n}{2\sigma^2} + \frac{n\kappa^{2(s)}}{2\sigma^4} = 0, \tag{S4}$$

$$\frac{\partial Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right)}{\partial \boldsymbol{\phi_t}} = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\partial \Delta\tau_{ij}}{\partial \boldsymbol{\phi_t}}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i] - \frac{\Delta\tau_{ij}}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\Delta\varsigma_{ij}^{-1}] + \frac{1}{\Delta\tau_{ij}}\right\} = \mathbf{0}, \tag{S5}$$

$$\frac{\partial Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right)}{\partial \boldsymbol{\phi_u}} = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\partial \Delta\nu_{ij}}{\partial \boldsymbol{\phi_u}}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i] - \frac{\Delta\nu_{ij}}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[1-\Delta\varsigma_{ij}{}^{-1}] + \frac{1}{\Delta\nu_{ij}}\right\} = \mathbf{0}. \tag{S6}$$

Equations (S3) and (S4) yield closed-form updates for κ and σ^2:

$$\begin{aligned}\kappa^{(s+1)} &= \frac{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]}{n},\\ \sigma^{2(s+1)} &= \frac{\sum_{i=1}^{n}\left(\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i^2] - 2\kappa^{(s+1)}\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]\right)}{n} + \kappa^{2(s+1)}.\end{aligned} \tag{S7}$$

By contrast, Equations (S5) and (S6) involve nonlinear expressions without analytical solutions. We recommend solving these equations numerically via root-finding methods such as the `uniroot()` function in R, which is well-suited for univariate optimization within a bounded interval.

S2.3 Gain of the integral approximation method

Efficient and accurate integral approximation is a crucial aspect of implementing the EM algorithm for the proposed model. This section compares three commonly used techniques for numerical integration: the Trapezoidal Rule (TZ), Monte Carlo (MC) simulation, and Gauss–Legendre (GL) quadrature.

The TZ approximation estimates the value of a definite integral by dividing the interval $[a, b]$ into s equally spaced subintervals. The integral is then approximated by

$$\int_a^b h(x)\,\mathrm{d}x \approx \sum_{i=1}^{s}\frac{h(x_{i-1}) + h(x_i)}{2}\Delta x, \tag{S8}$$

where x_i denotes the ith evaluation point, and $\Delta x = (b-a)/s$ is the subinterval width.

The MC method estimates the integral via sampling, approximating the expected value by averaging s randomly drawn points:

$$\int_a^b h(x)\,\mathrm{d}x \approx \frac{b-a}{s}\sum_{i=1}^{s}h(\vartheta_i), \tag{S9}$$

where ϑ_i is a randomly sampled value from the interval $[a, b]$.

Although straightforward to implement, TZ and MC methods can suffer from slow convergence or large variance, particularly when the integrand exhibits high curvature or complexity (Swarztrauber, 2003). To overcome these limitations, we adopt the GL quadrature method, which is especially effective for smooth functions on bounded intervals. The GL quadrature approximates $\int_{-1}^{1} h(x)\,\mathrm{d}x$ as

$$\int_{-1}^{1} h(x)\,\mathrm{d}x \approx \sum_{q=1}^{l} w_q h(z_q), \tag{S10}$$

where $\{z_q\}$ are the l roots of the Legendre polynomial $P_l(x)$ and $\{w_q\}$ are the associated weights. These values are chosen such that the quadrature rule is exact for polynomials of degree up to $2l-1$ (Golub and Welsch, 1969). The weights are given by:

$$w_q = \frac{2}{\left(1-z_q^2\right)\left[P_l'(z_q)\right]^2}, \quad P_l(z_q) = \frac{1}{2^l l!}\frac{d^l}{dz_q^l}(z_q^2-1)^l.$$

In our implementation, we use GL quadrature to approximate conditional expectations of the form $\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i}[g(\Delta\varsigma_{ij})] = \int_0^1 g(\Delta\varsigma_{ij}) f(\Delta\varsigma_{ij} \mid \mathbb{Q}_i)\,\mathrm{d}\Delta\varsigma_{ij}$. To apply GL integration over $[-1, 1]$, we introduce the transformation $z_{ij} = 2\Delta\varsigma_{ij} - 1$, so that $\Delta\varsigma_{ij} = (z_{ij}+1)/2$ and $\mathrm{d}\Delta\varsigma_{ij} = 1/2\mathrm{d}z_{ij}$. The integral becomes:

$$\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i}[g(\Delta\varsigma_{ij})] = \frac{1}{2}\int_{-1}^{1} g\left(\frac{z_{ij}+1}{2}\right) f\left(\frac{z_{ij}+1}{2} \mid \mathbb{Q}_i\right)\mathrm{d}z_{ij}. \tag{S11}$$

Using GL quadrature, the expectation is approximated as:

$$\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i}[g(\Delta\varsigma_{ij})] \approx \frac{1}{2}\sum_{q=1}^{l} w_q g\left(\frac{z_q+1}{2}\right) f\left(\frac{z_q+1}{2} \mid \mathbb{Q}_i\right). \tag{S12}$$

This approach is particularly effective for capturing the non-uniform shape of $f(\Delta\varsigma_{ij} \mid \mathbb{Q}_i)$ and delivers accurate and stable results. In our implementation, we use the `gaussquad` package in R (Novomestky, 2022) to obtain roots and weights. The value $l = 30$ is selected to balance computational cost and numerical precision, as discussed in Section 4.3.

S2.4 Determine initial parameter estimation Values

Starting with well-chosen initial parameter estimates can facilitate the rapid convergence of the EM algorithm. These initial parameter estimates can be derived from an

approximate approach. First, a simple nonlinear rIG process, $r\mathcal{IG}(\Lambda^t(t) + \Lambda^u(u), \gamma_i)$, is fit individually to each degradation path, providing values for $\hat{\alpha}_{t,i}$, $\hat{\beta}_{t,i}$, $\hat{\alpha}_{u,i}$, $\hat{\beta}_{u,i}$, and $\hat{\gamma}_i$, $i = 1, \ldots, n$. Averaging over these individual fits gives the initial values: $\alpha_t^{(0)}$, $\beta_t^{(0)}$, $\alpha_u^{(0)}$, and $\beta_u^{(0)}$. To estimate $\kappa^{(0)}$ and $\sigma^{2(0)}$, an rIG distribution $\mathcal{N}(\kappa, \sigma^2)$ is fit to the collected $\hat{\gamma}_i$ values.

To demonstrate the effectiveness of the initial value estimation method, we follow the parameter settings from Section 4 of the main text for the simulation experiments. Under $n = 30$, Figure S1 compares the estimated results from the initial value estimation method with the true values (purple dashed line). As the sample size m increases, the estimated values of the parameters converge towards true values, and the distribution becomes more concentrated, indicating that larger sample sizes improve estimation accuracy. The initial value guesses perform well, with α and β showing relatively stable estimates even with smaller sample sizes, while the estimation accuracy for parameters like κ and σ^2 improves significantly as the sample size increases.

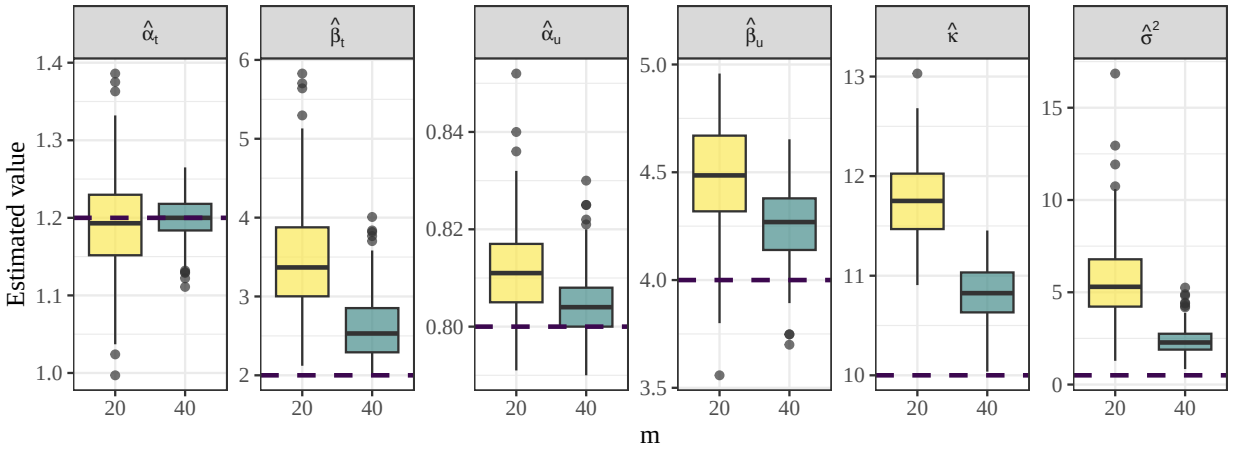


Figure S1: Performance of Initial Parameter Estimation (True Values Indicated by Purple Dashed Line).

S3 Statistical inference for model extensions

This section outlines the statistical inference procedures for the extended models under the unified framework. For Models M_1 through M_3, parameter estimation is performed using the EM algorithm. The derivations and implementation steps closely mirror those developed for Model M_0 and follow the same structure as described in Algorithm 1. In contrast, Model M_4, which does not incorporate random effects, permits direct inference via the classical maximum likelihood estimation (MLE) approach. For ease of use, all models (M_0–M_4) have

been integrated into a unified R function, `EM()`, where users can specify the desired model variant via the `model=` option.

S3.1 Two-scale with fixed-effects model (M_1)

Let $\boldsymbol{\Theta} = (\alpha_t, \beta_t, \alpha_u, \beta_u, \gamma)'$ represent the vector of unknown parameters, and denote the complete data by $\{\mathbb{Q}, \boldsymbol{\Delta x}\}$. The log-likelihood function can then be expressed as

$$\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\Delta x}) = \ell\left(\alpha_u, \beta_u; \mathbb{Q} \mid \boldsymbol{\Delta x}\right) + \ell\left(\alpha_t, \beta_t; \boldsymbol{\Delta x}\right),$$

where

$$\begin{aligned}
\ell\left(\alpha_u, \beta_u; \mathbb{Q} \mid \boldsymbol{\Delta x}\right) = &\sum_{i=1}^{n}\sum_{j=1}^{m_i}\Bigg[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\left[\log(1-\Delta x_{ij}/\Delta y_{ij}) + \log\Delta y_{ij}\right] + \log\Delta\nu_{ij} \\
&-\frac{1}{2}\left(\gamma\sqrt{\Delta y_{ij}-\Delta x_{ij}} - \Delta\nu_{ij}/\sqrt{\Delta y_{ij}-\Delta x_{ij}}\right)^2\Bigg], \\
\ell\left(\alpha_t, \beta_t; \boldsymbol{\Delta x}\right) = &\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\log\Delta x_{ij} + \log\Delta\tau_{ij} - \frac{1}{2}\left(\gamma\sqrt{\Delta x_{ij}} - \Delta\tau_{ij}/\sqrt{\Delta x_{ij}}\right)^2\right].
\end{aligned}$$

Given the parameter estimate $\boldsymbol{\Theta}^{(s)}$ obtained at the s-th iteration, the algorithm continues with the following steps:

- **E-step**: The Q-function, based on the current parameter estimate $\boldsymbol{\Theta}^{(s)}$, is computed as

$$\begin{aligned}
&Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right) = \mathbb{E}\left[\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\Delta x}) \mid \mathbb{Q}, \boldsymbol{\Theta}^{(s)}\right] \\
&= \sum_{i=1}^{n}\Bigg\{-\frac{1}{2}\gamma^2\sum_{j=1}^{m_i}\Delta y_{ij} + \gamma\sum_{j=1}^{m_i}(\Delta\tau_{ij}+\Delta\nu_{ij}) \\
&\quad -\frac{3}{2}\sum_{j=1}^{m_i}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\log(1-\Delta\varsigma_{ij})] - \frac{1}{2}\sum_{j=1}^{m_i}\frac{\Delta\nu_{ij}^2}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}\left[(1-\Delta\varsigma_{ij})^{-1}\right] \\
&\quad -\frac{3}{2}\sum_{j=1}^{m_i}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}\left[\log(\Delta\varsigma_{ij})\right] - \frac{1}{2}\sum_{j=1}^{m_i}\frac{\Delta\tau_{ij}^2}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}\left[{\Delta\varsigma_{ij}}^{-1}\right]\Bigg\} + \mathcal{C}_1,
\end{aligned} \tag{S13}$$

  where $\mathcal{C}_1 = -M\log(2\pi) + \sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[\log\Delta\nu_{ij}\Delta\tau_{ij} - 3\log\Delta y_{ij}\right]$. The expectations concerning $\Delta\varsigma_{ij}$ within the Q-function can be derived using Theorem 1, and then incorporated into (S13).

- **M-step**: Maximize the Q-function to obtain the updated parameter estimate $\boldsymbol{\Theta}^{(s+1)}$. The parameter $\gamma^{(s+1)}$ is updated as: $\gamma^{(s+1)} = \sum_{i=1}^{n}\sum_{j=1}^{m_i}(\Delta\tau_{ij}+\Delta\nu_{ij}) / \sum_{i=1}^{n}\sum_{j=1}^{m_i}\Delta y_{ij}$.

The values of $\boldsymbol{\phi}_t^{(s+1)} = (\alpha_t^{(s+1)}, \beta_t^{(s+1)})'$, and $\boldsymbol{\phi}_u^{(s+1)} = (\alpha_u^{(s+1)}, \beta_u^{(s+1)})'$ are then determined by solving the following equations:

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} \frac{\partial \Delta\tau_{ij}}{\partial \boldsymbol{\phi_t}} \left\{ \gamma^{(s)} - \frac{\Delta\tau_{ij}}{\Delta y_{ij}} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\Delta\varsigma_{ij}^{-1}] + \frac{1}{\Delta\tau_{ij}} \right\} = \mathbf{0},$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} \frac{\partial \Delta\nu_{ij}}{\partial \boldsymbol{\phi_u}} \left\{ \gamma^{(s)} - \frac{\Delta\nu_{ij}}{\Delta y_{ij}} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[1 - \Delta\varsigma_{ij}{}^{-1}] + \frac{1}{\Delta\nu_{ij}} \right\} = \mathbf{0}.$$

S3.2 Two-scale with identity-trick model (M_2)

Let $\boldsymbol{\Theta} = (\alpha_t, \beta_t, \alpha_u, \beta_u, \kappa, \sigma^2)'$ represent the vector of unknown parameters, and denote the complete data by $\{\mathbb{Q}, \boldsymbol{\Delta x}, \boldsymbol{\gamma}\}$. The log-likelihood function can then be expressed as

$$\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\Delta x}, \boldsymbol{\gamma}) = \ell\left(\alpha_u, \beta_u, \kappa; \mathbb{Q} \mid \boldsymbol{\Delta x}, \boldsymbol{\gamma}\right) + \ell\left(\alpha_t, \beta_t, \kappa; \boldsymbol{\Delta x} \mid \boldsymbol{\gamma}\right) + \ell(\sigma^2; \boldsymbol{\gamma})$$

where

$$\begin{aligned}
\ell\left(\alpha_u, \beta_u, \kappa; \mathbb{Q} \mid \boldsymbol{\Delta x}, \boldsymbol{\gamma}\right) =& \sum_{i=1}^{n}\sum_{j=1}^{m_i} \Bigg[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\left[\log(1 - \Delta x_{ij}/\Delta y_{ij}) + \log \Delta y_{ij}\right] + \log \Delta\nu_{ij} \\
& -\frac{1}{2}\left(\kappa\gamma_i\sqrt{\Delta y_{ij} - \Delta x_{ij}} - \Delta\nu_{ij}/\sqrt{\Delta y_{ij} - \Delta x_{ij}}\right)^2 \Bigg], \\
\ell\left(\alpha_t, \beta_t, \kappa; \boldsymbol{\Delta x} \mid \boldsymbol{\gamma}\right) =& \sum_{i=1}^{n}\sum_{j=1}^{m_i} \Bigg[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\log \Delta x_{ij} + \log \Delta\tau_{ij} \\
& -\frac{1}{2}\left(\kappa\gamma_i\sqrt{\Delta x_{ij}} - \Delta\tau_{ij}/\sqrt{\Delta x_{ij}}\right)^2 \Bigg], \\
\ell(\sigma^2; \boldsymbol{\gamma}) =& \sum_{i=1}^{n} -\frac{1}{2}\left[\log(2\pi) + \log\sigma^2 + (\gamma_i - 1)^2/\sigma^2\right].
\end{aligned}$$

Given the parameter estimate $\boldsymbol{\Theta}^{(s)}$ obtained at the s-th iteration, the algorithm continues with the following steps:

- **E-step**: The Q-function, based on the current parameter estimate $\boldsymbol{\Theta}^{(s)}$, is computed

as

$$
\begin{aligned}
&Q\left(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(s)}\right) = \mathbb{E}\left[\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\Delta x}, \boldsymbol{\gamma}) \mid \mathbb{Q}, \boldsymbol{\Theta}^{(s)}\right] \\
&= \sum_{i=1}^{n} \left\{ -\frac{1}{2}\left(\kappa^2 \sum_{j=1}^{m_i} \Delta y_{ij} + \frac{1}{\sigma^2}\right) \mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i^2] + \left[\kappa \sum_{j=1}^{m_i} (\Delta\tau_{ij} + \Delta\nu_{ij}) + \frac{1}{\sigma^2}\right] \mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i] \right. \\
&\quad -\frac{3}{2}\sum_{j=1}^{m_i} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\log(1-\Delta\varsigma_{ij})] - \frac{1}{2}\sum_{j=1}^{m_i} \frac{\Delta\nu_{ij}^2}{\Delta y_{ij}} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[(1-\Delta\varsigma_{ij})^{-1}\right] \\
&\quad \left. -\frac{3}{2}\sum_{j=1}^{m_i} \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[\log(\Delta\varsigma_{ij})\right] - \frac{1}{2}\sum_{j=1}^{m_i} \Delta\tau_{ij}^2 \mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}\left[{\Delta\varsigma_{ij}}^{-1}\right] \right\} + \mathcal{C}_2,
\end{aligned}
\tag{S14}
$$

where $\mathcal{C}_2 = -(M + n/2)\log(2\pi) + \sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[\log\left(\Delta\nu_{ij}\Delta\tau_{ij}\right) - 3\log\Delta y_{ij}\right] - n\log\sigma - n/(2\sigma^2)$. The expectations related to $\Delta\varsigma_{ij}$ can be calculated according to Theorem 1. Additionally, since $\gamma_i \mid \mathbb{Q}_i$ follows a normal distribution with mean $\frac{\kappa\sum_{j=1}^{m_i}(\Delta\nu_{ij}+\Delta\tau_{ij})+1/\sigma^2}{\kappa^2\sum_{j=1}^{m_i}\Delta y_{ij}+1/\sigma^2}$, and variance $\frac{1}{\kappa^2\sum_{j=1}^{m_i}\Delta y_{ij}+1/\sigma^2}$, corresponding expectations can be derived.

- **M-step**: Maximize the Q-function to obtain the updated parameter estimate $\boldsymbol{\Theta}^{(s+1)}$. The parameters $\kappa^{(s+1)}$ and $\sigma^{2(s+1)}$ are updated as:

$$
\begin{aligned}
\kappa^{(s+1)} &= \frac{\sum_{i=1}^{n}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i]\sum_{j=1}^{m_i}(\Delta\tau_{ij}+\Delta\nu_{ij})\right\}}{\sum_{i=1}^{n}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i^2]\sum_{j=1}^{m_i}\Delta y_{ij}\right\}} \\
\sigma^{2(s+1)} &= \frac{\sum_{i=1}^{n}\left(\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i^2] - 2\mathbb{E}_{\gamma_i^2|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i]\right)}{n} + 1.
\end{aligned}
$$

The update of $\boldsymbol{\phi}_t^{(s+1)}$, and $\boldsymbol{\phi}_u^{(s+1)}$ are implemented by solving the following functions:

$$
\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\partial\Delta\tau_{ij}}{\partial\boldsymbol{\phi_t}}\left\{\kappa^{(s+1)}\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i] - \frac{\Delta\tau_{ij}}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\Delta\varsigma_{ij}^{-1}] + \frac{1}{\Delta\tau_{ij}}\right\} = \mathbf{0}, \\
&\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\partial\Delta\nu_{ij}}{\partial\boldsymbol{\phi_u}}\left\{\kappa^{(s+1)}\mathbb{E}_{\gamma_i|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[\gamma_i] - \frac{\Delta\nu_{ij}}{\Delta y_{ij}}\mathbb{E}_{\Delta\varsigma_{ij}|\mathbb{Q}_i, \boldsymbol{\Theta}^{(s)}}[1-{\Delta\varsigma_{ij}}^{-1}] + \frac{1}{\Delta\nu_{ij}}\right\} = \mathbf{0}.
\end{aligned}
$$

S3.3 Single-scale with random-effects model (M_3)

Let $\boldsymbol{\Theta} = (\alpha, \beta, \kappa, \sigma^2)'$ represent the vector of unknown parameters, and denote the complete data by $\mathbb{Q} = \{\mathbb{Q}_1, \ldots, \mathbb{Q}_n\}$, where $\mathbb{Q}_i = \{\boldsymbol{\Delta y}_i, \boldsymbol{t}_i\}$. The log-likelihood function can then be expressed as

$$
\ell(\boldsymbol{\Theta}; \mathbb{Q}, \boldsymbol{\gamma}) = \ell(\alpha, \beta; \mathbb{Q} \mid \boldsymbol{\gamma}) + \ell(\kappa, \sigma^2; \boldsymbol{\gamma})
$$

where

$$\ell(\alpha,\beta;\mathbb{Q}\mid\boldsymbol{\gamma})=\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[-\frac{1}{2}\log(2\pi)-\frac{3}{2}\log\Delta y_{ij}+\log\Delta\tau_{ij}-\frac{1}{2}\left(\gamma_i\sqrt{\Delta y_{ij}}-\Delta\tau_{ij}/\sqrt{\Delta y_{ij}}\right)^2\right],$$

$$\ell(\kappa,\sigma^2;\boldsymbol{\gamma})=\sum_{i=1}^{n}-\frac{1}{2}\left[\log(2\pi)+\log\sigma^2+(\kappa-\gamma_i)^2/\sigma^2\right].$$

Given the parameter estimate $\boldsymbol{\Theta}^{(s)}$ obtained at the s-th iteration, the algorithm continues with the following steps:

- **E-step**: The Q-function, based on the current parameter estimate $\boldsymbol{\Theta}^{(s)}$, is computed as

$$\begin{aligned}Q\left(\boldsymbol{\Theta}\mid\boldsymbol{\Theta}^{(s)}\right)&=\mathbb{E}\left[\ell(\boldsymbol{\Theta};\mathbb{Q},\boldsymbol{\gamma})\mid\mathbb{Q},\boldsymbol{\Theta}^{(s)}\right]\\&=\sum_{i=1}^{n}\left\{-\frac{1}{2}\left(\sum_{j=1}^{m_i}\Delta y_{ij}+\frac{1}{\sigma^2}\right)\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i^2]+\left(\sum_{j=1}^{m_i}\Delta\tau_{ij}+\frac{\kappa}{\sigma^2}\right)\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]\right\}+\mathcal{C}_3,\end{aligned}\tag{S15}$$

  where $\mathcal{C}_3=-(M+n)/2\log(2\pi)+\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[\log\Delta\tau_{ij}-3/2\log\Delta y_{ij}\right]-n\log\sigma-(n\kappa^2)/(2\sigma^2)$. The expectations are calculated concerning the conditional distribution $f\left(\gamma_i,\mid\mathbb{Q},\boldsymbol{\Theta}^{(s)}\right)$. Note that $\gamma_i\mid\mathbb{Q}_i\sim\mathcal{N}\left(\frac{\kappa/\sigma^2+\sum_{j=1}^{m_i}\Delta\tau_{ij}}{1/\sigma^2+\sum_{j=1}^{m_i}\Delta y_{ij}},\frac{1}{1/\sigma^2+\sum_{j=1}^{m_i}\Delta y_{ij}}\right)$. The two conditional expectations involved can be obtained as

$$\mathbb{E}_{\gamma_i|\mathbb{Q}_i}[\gamma_i]=\frac{\kappa/\sigma^2+\sum_{j=1}^{m_i}\Delta\tau_{ij}}{1/\sigma^2+\sum_{j=1}^{m_i}\Delta y_{ij}},\text{ and }\mathbb{E}_{\gamma_i|\mathbb{Q}_i}[\gamma_i^2]=\frac{1/\sigma^2+\sum_{j=1}^{m_i}\Delta y_{ij}+\left(\kappa/\sigma^2+\sum_{j=1}^{m_i}\Delta\tau_{ij}\right)^2}{\left(1/\sigma^2+\sum_{j=1}^{m_i}\Delta y_{ij}\right)^2},\tag{S16}$$

  then put them into (S15).

- **M-step**: Maximize the Q-function to obtain the updated parameter estimate $\boldsymbol{\Theta}^{(s+1)}$. The parameters $\kappa^{(s+1)}$ and $\sigma^{2(s+1)}$ are updated as:

$$\kappa^{(s+1)}=\frac{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]}{n},\text{ and }\sigma^{2(s+1)}=\frac{\sum_{i=1}^{n}\left(\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i^2]-2\kappa^{(s+1)}\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]\right)}{n}+\kappa^{2(s+1)}.$$

  The update of $\boldsymbol{\phi}^{(s+1)}=(\alpha^{(s+1)},\beta^{(s+1)})'$ can be implemented by solving the following functions:

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\partial\Delta\tau_{ij}}{\partial\boldsymbol{\phi}}\left\{\mathbb{E}_{\gamma_i|\mathbb{Q}_i,\boldsymbol{\Theta}^{(s)}}[\gamma_i]-\frac{\Delta\tau_{ij}}{\Delta y_{ij}}+\frac{1}{\Delta\tau_{ij}}\right\}=\mathbf{0}.$$

S3.4 Single-scale with fixed-effects model (M_4)

Let $\boldsymbol{\Theta} = (\alpha, \beta, \gamma)'$ represent the vector of unknown parameters, and define the observed data by $\mathbb{Q} = \{\mathbb{Q}_1, \mathbb{Q}_2, \ldots, \mathbb{Q}_n\}$, where $\mathbb{Q}_i = \{\boldsymbol{\Delta y}_i, \boldsymbol{t}_i\}$. Here, $\boldsymbol{t}_i = (t_{i0}, \ldots, t_{im_i})'$ and $\boldsymbol{\Delta y}_i = (\Delta y_{i1}, \ldots, \Delta y_{im_i})'$. Then, the log-likelihood function is given by

$$\ell(\boldsymbol{\Theta}; \mathbb{Q}) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \left[-\frac{1}{2}\log(2\pi) - \frac{3}{2}\log \Delta y_{ij} + \log \Delta \tau_{ij} - \frac{1}{2}\left(\gamma\sqrt{\Delta y_{ij}} - \Delta \tau_{ij}/\sqrt{\Delta y_{ij}}\right)^2 \right].$$

We apply the ML estimation method, determining parameter estimates by equating the first-order partial derivatives to zero. The resulting parameter estimates are as follows: $\hat{\gamma} = \sum_{i=1}^{n}\sum_{j=1}^{m_i} \Delta \tau_{ij} / \sum_{i=1}^{n}\sum_{j=1}^{m_i} \Delta y_{ij}$, and $\sum_{i=1}^{n}\sum_{j=1}^{m_i} \frac{\partial \Delta \tau_{ij}}{\partial \phi} \left\{\gamma - \frac{\Delta \tau_{ij}}{\Delta y_{ij}} + \frac{1}{\Delta \tau_{ij}}\right\} = \mathbf{0}$.

S4 Technical details of bootstrap

The bootstrap method is a popular resampling technique for estimating the distribution of a parameter, especially useful when the sample size is too small to rely on asymptotic properties. In small or moderate samples, the bootstrap approach provides an effective way to construct confidence intervals. This technique generates multiple resamples from the original dataset to approximate the sampling distribution of the parameter, allowing for the construction of approximate $100(1-\alpha)\%$ confidence intervals for functions of the model parameters. Algorithm S1 outlines the bootstrap procedure for model M_0. For extended models, the process is similar, with only minor adjustments needed on lines 3, 5, and 6 of the algorithm.

S5 Bayesian inference

In this section, we introduce the Bayesian inference procedure. Section S5.1 provides the Bayesian formulation for the proposed model. Sections S5.2 and S5.3 detail the posterior sampling methods, specifically Gibbs and Hamiltonian Monte Carlo (HMC) sampling, respectively.

S5.1 Bayesian formulation

Bayesian analysis plays a crucial role in reliability modeling and decision-making due to its ability to incorporate prior knowledge and quantify uncertainty (Taylor et al., 2024; Puli et al., 2023). One of its primary advantages is the ability to incorporate existing

Algorithm S1: Bootstrap Algorithm Procedure.

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

Output: B resamples of the estimate $\left\{\hat{\boldsymbol{\Theta}}_1^*, \ldots, \hat{\boldsymbol{\Theta}}_B^*\right\}$.

```
for b = 1 to B do
    for i = 1 to n do
        Generate γ̆_i from N(κ̂, σ̂²);
        for k = 1 to m_i do
            Given α̂_t, γ̂_t, α̂_u, γ̂_u, generate τ̆_ij, and ν̆_ij.
            Given γ̆_i, generate y̆_ij from rIG(τ̆_ij + ν̆_ij, γ̆_i);
        end
    end
    Obtain the bootstrapped degradation data Q̆;
    Obtain Θ̂_b^* based on Q̆ using the proposed EM algorithm.
end
```

information, often in the form of prior distributions, which encapsulate our prior beliefs about the parameters of interest. By integrating this prior knowledge with observed data, Bayesian analysis offers a powerful means to refine parameter estimates. The Bayesian framework in our model M_0 is set as follows:

$$\mathcal{Y}_i(t,u) = \mathcal{X}_i(t) + \mathcal{Z}_i(u),\ \gamma_i \sim \mathcal{N}(\kappa, \sigma^2),\ i = 1, \ldots, n, \tag{S17}$$

$$\mathcal{X}_i(t) \mid \gamma_i \sim r\mathcal{IG}\left(\Lambda^t(t; \alpha_t, \beta_t), \gamma_i\right),\ \mathcal{Z}_i(u) \mid \gamma_i \sim r\mathcal{IG}\left(\Lambda^u(u; \alpha_u, \beta_u), \gamma_i\right), \tag{S18}$$

$$\kappa \mid \sigma^2 \sim \mathcal{N}\left(e, \sigma^2/f\right),\ \sigma^2 \sim \mathcal{IG}a(g,h), \tag{S19}$$

$$\alpha_t \sim \mathcal{N}\left(a_t, b_t^2\right), \alpha_u \sim \mathcal{N}\left(a_u, b_u^2\right), \tag{S20}$$

$$\beta_t \sim \mathcal{N}\left(c_t, d_t^2\right), \beta_u \sim \mathcal{N}\left(c_u, d_u^2\right), \tag{S21}$$

where $\mathcal{IG}a(\cdot)$ denotes the inverse gamma distribution. The dependence of the variables are shown in Figure S2. Combining $\pi(\kappa \mid \sigma^2)$ and $\pi(\sigma^2)$ in (S19), then we have $\pi(\kappa, \sigma^2) \sim \mathcal{NIG}a\,(e, f, g, h)$, i.e., normal-inverse gamma distribution as below.

$$\pi\left(\kappa, \sigma^2\right) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{f}}{\sigma} \exp\left(-\frac{1}{2} \frac{f(\kappa - e)^2}{\sigma^2}\right) \times \frac{h^g \left(\sigma^2\right)^{-g-1}}{\Gamma(g)} \exp\left(-\frac{h}{\sigma^2}\right). \tag{S22}$$

Note that we use the normal-inverse gamma distribution as the conjugate prior for the mean and variance parameters of the normal distribution to ensure analytical tractability of our

formulas (Bernardo and Smith, 2009). In (S19), we establish priors for the shared parameters within the model. This enables us to improve the accuracy of γ_i estimation by incorporating information from other samples. Next, we provide a prior for the drift parameters in (S20) and (S21) with normal priors. While the parameters might occasionally be negative, this becomes highly improbable when the prior distribution's mean to standard deviation ratio is small. In Bayesian modeling, these rare extremes are typically considered negligible if the prior is well-designed.

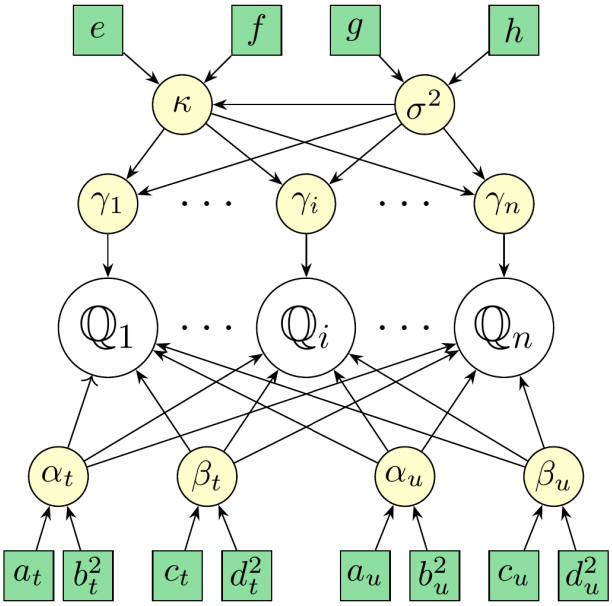


Figure S2: Dependency Diagram of Variables in Bayesian Framework: Arrows in the Directed graph Indicate Direct Influences. Green Rectangles Represent Hyperparameters, Yellow Circles Denote Variables, and White Circles Indicate Observed Data.

Recall $\boldsymbol{\Theta} = (\alpha_t, \beta_t, \alpha_u, \beta_u, \kappa, \sigma^2)'$, and according to Bayes' theorem, the joint posterior distribution of $\boldsymbol{\Theta}$ can be derived as

$$\begin{aligned} \pi\left(\boldsymbol{\Theta} \mid \mathbb{Q}\right) \propto & f_{\mathcal{Y}(t,u)}\left(\boldsymbol{\Delta y} \mid \boldsymbol{\Theta}\right) \pi(\boldsymbol{\gamma} \mid \kappa, \sigma^2) \pi\left(\kappa, \sigma^2 \mid e, f, g, h\right) \\ & \times \pi\left(\alpha_t \mid a_t, b_t^2\right) \pi\left(\alpha_u \mid a_u, b_u^2\right) \pi\left(\beta_t \mid c_t, d_t^2\right) \pi\left(\beta_u \mid c_u, d_u^2\right), \end{aligned} \tag{S23}$$

where

$$f_{Y(t,u)}\left(\boldsymbol{\Delta y} \mid \boldsymbol{\Theta}\right) = \prod_{i=1}^{n}\prod_{j=1}^{m_i} f\left(\boldsymbol{\Delta y} - \boldsymbol{\Delta x} \mid \boldsymbol{\Theta}\right) f\left(\boldsymbol{\Delta x} \mid \boldsymbol{\Theta}\right)$$
$$= \prod_{i=1}^{n}\prod_{j=1}^{m_i} \left\{ \frac{\Delta\nu_{ij}}{\sqrt{2\pi}} (\Delta y_{ij} - \Delta x_{ij})^{-3/2} \exp\left[-\frac{1}{2}\left(\sqrt{\Delta y_{ij} - \Delta x_{ij}}\gamma_i - \frac{\Delta\nu_{ij}}{\sqrt{\Delta y_{ij} - \Delta x_{ij}}} \right)^2 \right] \right. \quad \text{(S24)}$$
$$\left. \times \frac{\Delta\tau_{ij}}{\sqrt{2\pi}} \Delta x_{ij}^{-3/2} \exp\left[-\frac{1}{2}\left(\sqrt{\Delta x_{ij}}\gamma_i - \frac{\Delta\tau_{ij}}{\sqrt{\Delta x_{ij}}} \right)^2 \right] \right\}.$$

Given the complexity of $\pi(\boldsymbol{\Theta} \mid \mathbb{Q})$, a direct derivation of Bayesian estimates appears to be unfeasible. In this paper, we introduce two methods for posterior sampling. The specific details are provided in the following sections.

S5.2 Gibbs sampling algorithm

The Gibbs sampling algorithm iteratively samples each parameter based on its full conditional distribution, given the current values of the other parameters (Gelman et al., 1995). This approach simplifies sampling from complex posterior distributions by breaking it into manageable steps. A key aspect of implementing Gibbs sampling is calculating these conditional distributions for each parameter. Let $\boldsymbol{\Theta}_{\setminus\eta}$ represent the elements remaining in $\boldsymbol{\Theta}$ after removing η. By combining (S24) and $\pi(\alpha_t) \sim \mathcal{N}(a_t, b_t^2)$, we have

$$\pi(\alpha_t \mid \boldsymbol{\Theta}_{\setminus\alpha_t}, \mathbb{Q}) \propto \prod_{i=1}^{n}\prod_{j=1}^{m_i} \Delta\tau_{ij} \times \exp\left\{ \sum_{i=1}^{n}\sum_{j=1}^{m_i} \left(\gamma_i \Delta\tau_{ij} - \frac{\Delta\tau_{ij}^2}{2\Delta x_{ij}} \right) - \frac{(\alpha_t - a_t)^2}{2b_t^2} \right\}. \quad \text{(S25)}$$

Similarly, by combining (S24) and $\pi(\alpha_u) \sim \mathcal{N}(a_u, b_u^2)$, we have

$$\pi(\alpha_u \mid \boldsymbol{\Theta}_{\setminus\alpha_u}, \mathbb{Q}) \propto \prod_{i=1}^{n}\prod_{j=1}^{m_i} \Delta\nu_{ij} \times \exp\left\{ \sum_{i=1}^{n}\sum_{j=1}^{m_i} \left(\gamma_i \Delta\nu_{ij} - \frac{\Delta\nu_{ij}^2}{2(\Delta y_{ij} - \Delta x_{ij})} \right) - \frac{(\alpha_u - a_u)^2}{2b_u^2} \right\}. \quad \text{(S26)}$$

By combining (S24) and $\pi(\beta_t) \sim \mathcal{N}(c_t, d_t^2)$, we have

$$\pi(\beta_t \mid \boldsymbol{\Theta}_{\setminus\beta_t}, \mathbb{Q}) \propto \prod_{i=1}^{n}\prod_{j=1}^{m_i} \Delta\tau_{ij} \times \exp\left\{ \sum_{i=1}^{n}\sum_{j=1}^{m_i} \left(\gamma_i \Delta\tau_{ij} - \frac{\Delta\tau_{ij}^2}{2\Delta x_{ij}} \right) - \frac{(\beta_t - c_t)^2}{2d_t^2} \right\}. \quad \text{(S27)}$$

Similarly, by combining (S24) and $\pi(\beta_u) \sim \mathcal{N}(c_u, d_u^2)$, we have

$$\pi(\beta_u \mid \boldsymbol{\Theta}_{\setminus\beta_u}, \mathbb{Q}) \propto \prod_{i=1}^{n}\prod_{j=1}^{m_i} \Delta\nu_{ij} \times \exp\left\{ \sum_{i=1}^{n}\sum_{j=1}^{m_i} \left(\gamma_i \Delta\nu_{ij} - \frac{\Delta\nu_{ij}^2}{2(\Delta y_{ij} - \Delta x_{ij})} \right) - \frac{(\beta_u - c_u)^2}{2d_u^2} \right\}. \quad \text{(S28)}$$

By combining (S24) and (S22), we have

$$\pi(\kappa \mid \boldsymbol{\Theta}_{\setminus\kappa}, \mathbb{Q}) \propto \exp\left[-\frac{n+f}{2\sigma^2}\left(\kappa - \frac{\sum_{i=1}^{n}\gamma_i + ef}{n+f}\right)^2\right]. \tag{S29}$$

Hence, $\pi(\kappa \mid \boldsymbol{\Theta}_{\setminus\kappa}, \mathbb{Q})$ follows a normal distribution:

$$\mathcal{N}\left(\frac{\sum_{i=1}^{n}\gamma_i + ef}{n+f}, \frac{\sigma^2}{n+f}\right).$$

Similarly, by combining (S24) and (S22), we can obtain:

$$\pi(\sigma^2 \mid \boldsymbol{\Theta}_{\setminus\sigma^2}, \mathbb{Q}) \propto (\sigma^2)^{-(\frac{1}{2}+\frac{n}{2}+g+1)} \times \exp\left(-\frac{\frac{1}{2}\sum_{i=1}^{n}(\gamma_i-\kappa)^2 + h + \frac{1}{2}f(\kappa-e)^2}{\sigma^2}\right), \tag{S30}$$

hence, $\pi(\sigma^2 \mid \boldsymbol{\Theta}_{\setminus\sigma^2}, \mathbb{Q})$ follows an inverse gamma distribution:

$$IGa\left(\frac{1}{2}+\frac{n}{2}+g, \frac{1}{2}\sum_{i=1}^{n}(\gamma_i-\kappa)^2 + h + \frac{1}{2}f(\kappa-e)^2\right).$$

Overall, $\pi((\kappa,\sigma^2) \mid \boldsymbol{\Theta}_{\setminus(\kappa,\sigma^2)}, \mathbb{Q})$ follows a normal-inverse gamma distribution:

$$\mathcal{NIG}a\left(e', f', g', h'\right), \tag{S31}$$

where $e' = \left(ef + \sum_{i=1}^{n}\gamma_i\right)/(n+f)$, $f' = n+f$, $g' = 1/2+n/2+g$, $h' = 1/2\sum_{i=1}^{n}(\gamma_i-\kappa)^2 + h + 1/2f(\kappa-e)^2$.

Note that the full conditional posterior distributions for κ and σ^2 are known, allowing their samples to be generated directly using statistical software. On the other hand, for α_t, β_t, α_u, and β_u, we can employ the Adaptive Rejection Metropolis Sampling (ARMS) algorithm (Gilks et al., 2022). Overall, the ARMS-Gibbs sampling algorithm, outlined in Algorithm S2, generates posterior samples of the parameters for subsequent Bayesian inference.

S5.3 HMC algorithm

HMC is a sampling method particularly effective for high-dimensional models, with its core advantage being the use of Hamiltonian dynamics to explore parameter space efficiently (Betancourt, 2017). Unlike Gibbs sampling, which relies on step-by-step conditional sampling and can suffer from slow convergence in complex parameter spaces, HMC leverages gradient information to enable larger, more directed jumps in the parameter space. By simulating the trajectory of a particle under the influence of an energy function, HMC

Algorithm S2: ARMS-Gibbs Sampling Algorithm.

Input: $\mathbb{Q}$, initial values $\boldsymbol{\Theta}^{(0)}$.

Output: Posterior samples of $\boldsymbol{\Theta}$.

for $s = 1$ **to** $\mathcal{S}$ **do**

  Generate $\left(\kappa^{(s)}, \sigma^{2(s)}\right)$ from $NIGa\left(e'^{(s)}, f'^{(s)}, g'^{(s)}, h'^{(s)}\right)$, detailed in (S31);

  Use the ARMS algorithm to sample $\alpha_t^{(s)}$, $\beta_t^{(s)}$, $\alpha_u^{(s)}$, and $\beta_u^{(s)}$ from $\pi(\alpha_t \mid \boldsymbol{\Theta}^{(s-1)} \backslash \alpha_t, \mathbb{Q})$, $\pi(\beta_t \mid \boldsymbol{\Theta}^{(s-1)} \backslash \beta_t, \mathbb{Q})$, $\pi(\alpha_u \mid \boldsymbol{\Theta}^{(s-1)} \backslash \alpha_u, \mathbb{Q})$, and $\pi(\beta_u \mid \boldsymbol{\Theta}^{(s-1)} \backslash \beta_u, \mathbb{Q})$, respectively (see (S25)–(S28)).

end

After discarding the initial $\mathcal{L}$ burn-in samples, the remaining $\mathcal{S} - \mathcal{L}$ samples are used to construct point estimates and credible intervals for each parameter.

samples along smooth, continuous paths, significantly reducing sample autocorrelation and improving sampling efficiency. This mechanism avoids the common issue of Gibbs sampling becoming stuck in high-dimensional spaces and allows HMC to handle strong correlations between parameters more effectively (Fang et al., 2024).

In HMC, the parameter vector $\boldsymbol{\Theta}$ is treated as a position variable, while an auxiliary momentum variable $\boldsymbol{p}$ is introduced, transforming the sampling problem into a physical motion simulation. The auxiliary distribution for $\boldsymbol{p}$ is a multivariate normal distribution, independent of $\boldsymbol{\Theta}$, and defined as $\boldsymbol{p} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{M})$, where $\boldsymbol{M}$ is a symmetric, positive-definite mass matrix that scales the momentum. The resulting Hamiltonian function, representing the total energy of the system, is defined as $H(\boldsymbol{\Theta}, \boldsymbol{p}) = U(\boldsymbol{\Theta}) + K(\boldsymbol{p})$, where the potential energy $U(\boldsymbol{\Theta})$ is the negative log-posterior density, i.e., $U(\boldsymbol{\Theta}) = -\ln \pi(\boldsymbol{\Theta} \mid \mathbb{Q})$, and the kinetic energy $K(\boldsymbol{p})$ is defined as $K(\boldsymbol{p}) = \boldsymbol{p}^T \boldsymbol{M}^{-1} \boldsymbol{p}/2$. Assuming the s-th step with parameters $\boldsymbol{\Theta}^{(s)}$, the procedure to obtain the new posterior samples $\boldsymbol{\Theta}^{(s+1)}$ is as follows:

(i) **Momentum sampling**: Sample a new momentum $\boldsymbol{p}^{(s)}$ from $\mathcal{N}(\mathbf{0}, \boldsymbol{M})$.

(ii) **Hamiltonian Integration**: Integrate the equations of motion (e.g., with the leapfrog algorithm (Birdsall and Langdon, 2018)) to simulate the particle' s trajectory in parameter space.

(iii) **Metropolis-Hastings Correction**: Apply a Metropolis-Hastings step to obtain the updated sample $\boldsymbol{\Theta}^{(s+1)}$.

As with Gibbs sampling, the HMC procedure generates $\mathcal{S}$ posterior samples, with the initial $\mathcal{L}$ iterations discarded as burn-in. The remaining $\mathcal{S} - \mathcal{L}$ samples are then used to compute posterior summaries for statistical inference. In this study, the HMC algorithm is implemented using the probabilistic programming language Stan, accessed via its R interface *rstan* (Carpenter et al., 2017).

S6 Additional simulation experiments

S6.1 Parameter estimation under different settings for model M_0

S6.1.1 Parameter estimation v.s. sample size

The estimation performance of the estimator of model M_0 under different sample sizes and measurement frequencies is shown in Table S1.

S6.1.2 Parameter estimation v.s. degradation curvature

We evaluate the model's estimation performance under different combinations of degradation curvatures. The model parameters are the same as in Section 4 of the main paper, except that α_t and α_u are varied, which can be set to 0.7, 1, and 1.3. Sample sizes are $n = 10$, 30, and 50, with $m = 20$. To simplify, we fix one curvature at 1 and vary the other. Figure S3 shows the average RRMSE for all parameters. Key findings include: i) Larger sample sizes improve estimation accuracy by reducing uncertainty. ii) A convex degradation trend ($\alpha_t > 1$) generally provides the best performance. However, in Scenario II, as α_u increases, the performance of α_t and β_t worsens due to the large difference between the two scales' degradation paths, making it harder to capture subtle changes in time-scale parameters. We will discuss this phenomenon in more detail in the next section. iii) $\hat{\kappa}$ and $\hat{\sigma^2}$ estimates remain stable across different curvatures, influenced more by sample size.

S6.1.3 Parameter estimation v.s. degradation value ratio

We further compare the degradation value ratio of the two scales. We conduct simulations with $\alpha_t = \alpha_u = 1$ and $\beta_t = 5$, keeping other parameters the same as in Section 4. By varying β_u (0.1, 0.5, and 5), we illustrate different ratios of degradation between the t and u scales (see Figure 4(a)), corresponding to scenarios where the t-scale degradation is much greater than, approximately equal to, or much less than the u-scale degradation, respectively. Figure 4(b) presents the RRMSE of parameter estimates for the three β_u values. When β_u is small, the degradation on the t-scale dominates, resulting in better estimation accuracy for

Table S1: RRMSE($\times 10^{-}2$) of estimators across various sample sizes and measurement frequencies for model M_0.

(n, m)	Method	Matrix	α_t	β_t	α_u	β_u	κ	σ^2
$(10, 20)$	EM	RB	-3.75	24.52	0.51	-0.54	2.39	-2.59
		RRMSE	7.10	41.04	1.61	8.29	6.38	48.68
	Bayes	RB	-3.75	35.72	1.27	-3.23	1.31	21.59
		RRMSE	10.15	87.85	3.45	12.17	5.79	65.14
$(30, 20)$	EM	RB	-4.02	22.12	0.49	-2.42	0.66	1.94
		RRMSE	5.21	29.33	0.99	4.99	3.05	30.56
	Bayes	RB	-0.87	5.06	0.13	-1.21	-0.23	8.66
		RRMSE	3.65	19.18	0.83	4.42	2.87	33.49
$(50, 20)$	EM	RB	-3.91	20.28	0.43	-2.37	0.33	-1.25
		RRMSE	4.68	24.69	0.85	4.23	2.49	20.17
	Bayes	RB	-0.36	0.72	-0.09	-0.80	-0.61	2.79
		RRMSE	2.77	12.24	0.65	3.38	2.46	20.97
$(10, 40)$	EM	RB	-1.52	10.56	0.30	-1.19	0.61	-11.18
		RRMSE	3.99	24.59	1.06	6.26	4.15	45.38
	Bayes	RB	-1.00	6.69	0.16	-1.52	-0.15	12.75
		RRMSE	4.05	25.91	1.19	6.49	4.02	58.12
$(30, 40)$	EM	RB	-1.62	10.49	0.19	-1.00	0.58	0.36
		RRMSE	2.67	15.83	0.69	4.09	2.48	28.42
	Bayes	RB	-0.38	1.53	-0.09	-0.50	-0.18	8.17
		RRMSE	2.01	10.13	0.62	3.66	2.36	31.46
$(50, 40)$	EM	RB	-1.79	10.40	0.25	-1.50	0.16	-1.26
		RRMSE	2.34	13.75	0.52	3.07	1.86	20.29
	Bayes	RB	-0.51	0.59	-0.09	-1.00	-0.67	3.29
		RRMSE	1.55	7.71	0.44	2.67	1.93	21.26

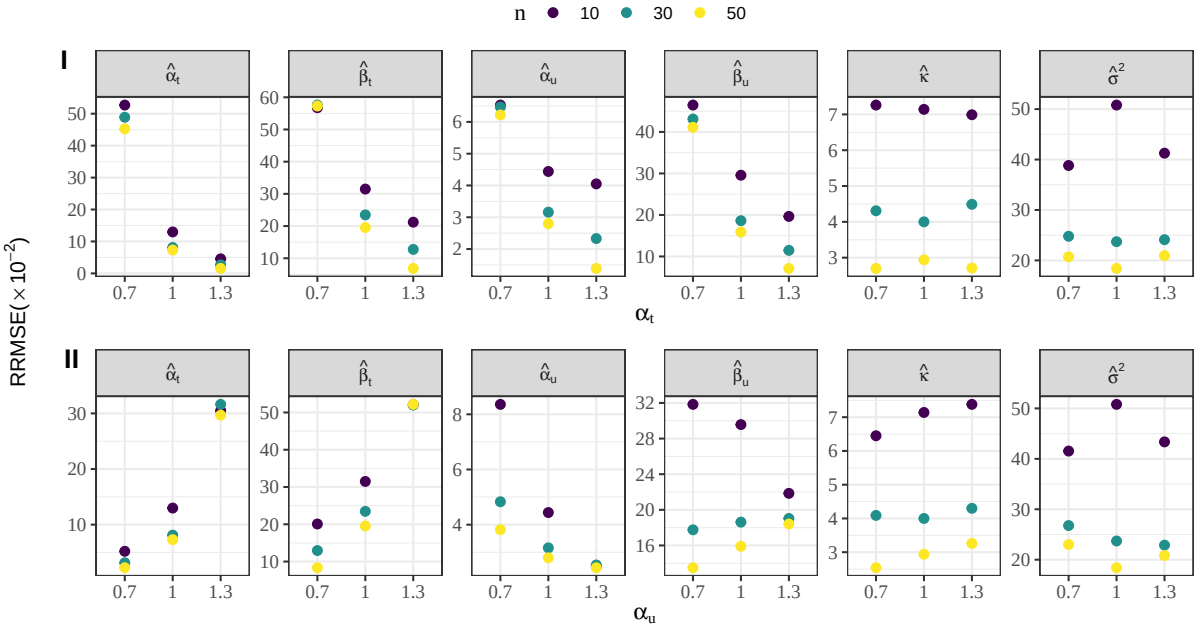
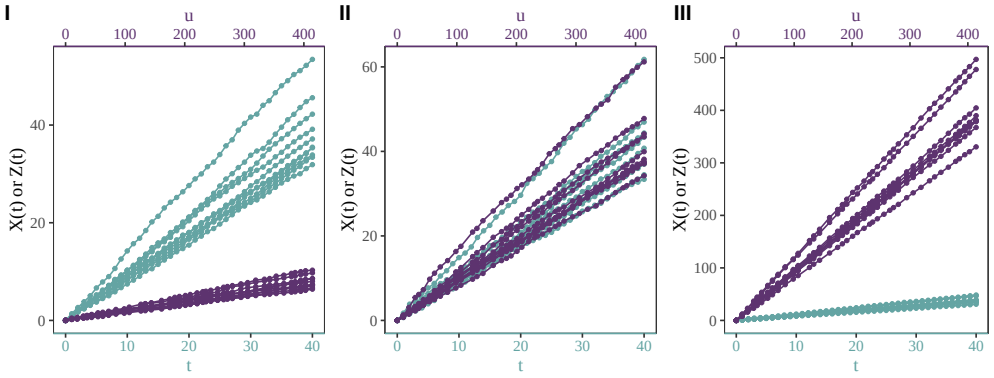


Figure S3: RRMSE of estimators for different degradation curvatures (scenario I: varying α_t, fixed $\alpha_u = 1$; scenario II: varying α_u, fixed $\alpha_t = 1$).

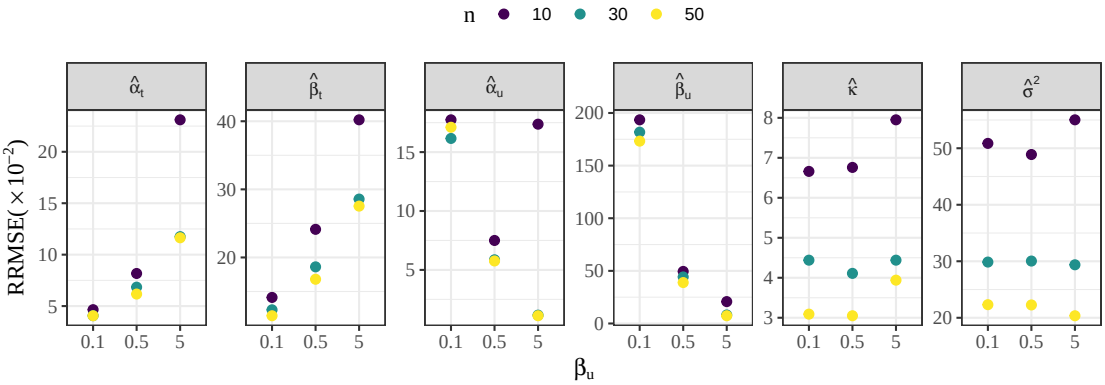
t-related parameters ($\hat{\alpha}_t$ and $\hat{\beta}_t$) with lower RRMSE. Conversely, the smaller degradation on the u-scale leads to poorer estimates for u-related parameters. As β_u increases, the u-scale degradation becomes more significant, improving the estimation for u-related parameters, while the estimation accuracy for t-related parameters declines. The estimates for $\hat{\kappa}$ and $\hat{\sigma}^2$ remain relatively unaffected by changes in β_u, showing minimal variation in RRMSE.

S6.1.4 Parameter estimation v.s. sample heterogeneity

We perform a simulation analysis for different levels of random effects. We set $\alpha_t = \alpha_u = 1$ and conduct the simulation under the same conditions as described in Section 4. By varying σ^2, we assess the impact of sample heterogeneity on estimation performance. The values of σ^2 are set to 0, 0.5, and 1, with the corresponding degradation paths shown in Figure 5(a). As σ^2 increases, the degradation paths become increasingly dispersed. Figure 5(b) presents the RRMSE results for the parameters under these three scenarios. We observe that the RRMSE for each parameter shows no significant variation across different σ^2 values. Whether $\sigma^2 = 0$, 0.5, or 1, the estimation errors remain relatively stable. This indicates that sample heterogeneity has little impact on the estimation performance for these parameters,



(a) Two-scale degradation paths



(b) RRMSE of estimators

Figure S4: Performance of two-scale degradation paths for $\beta_t = 5$ where I for $\beta_u = 0.1$, II for $\beta_u = 0.5$, and III for $\beta_u = 5$.

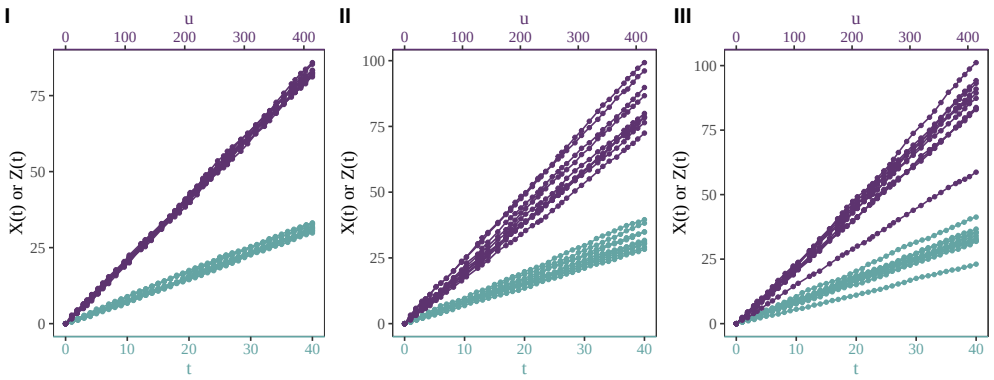
especially as the sample size increases, with RRMSE remaining largely consistent.

S6.2 Parameter estimation for models $M_1 - M_4$ v.s. sample size

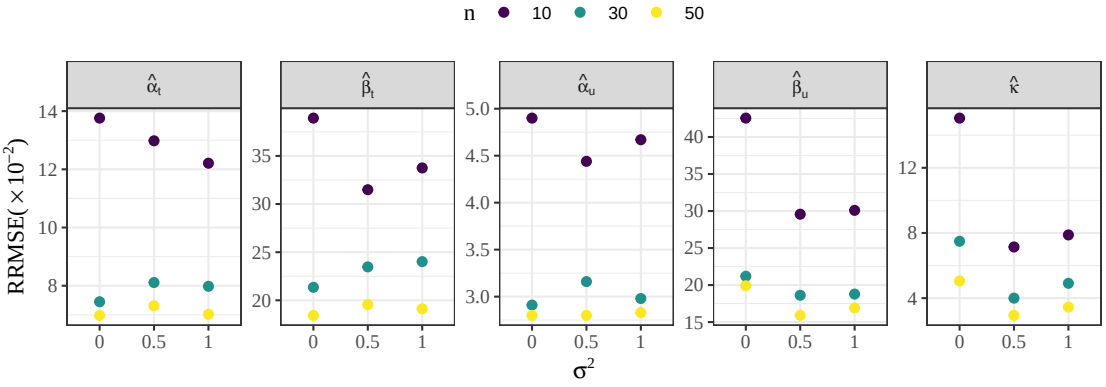
The parameter estimation performance for different model extensions is presented. Models M_1 to M_3 employ the EM algorithm, while M_4 uses the traditional MLE method. The corresponding results are shown in Tables S2 to S5.

S6.3 Validation under different inspection–time designs

To assess the robustness of the proposed method with respect to non-uniform inspection schedules, we consider three representative alternatives to the uniform inspection grid used in the baseline simulations:



(a) Two-scale degradation paths



(b) RRMSE of estimators

Figure S5: Performance of different sample heterogeneity, where I for $\sigma^2 = 0$, II for $\sigma^2 = 0.5$, and III for $\sigma^2 = 1$.

Table S2: Performance ($\times 10^{-2}$) of estimators across Various sample sizes and measurement frequencies for model M_1.

(n, m)	Matrix	α_t	β_t	α_u	β_u	γ
$(10, 20)$	RB	-1.98	17.77	0.52	-0.93	1.36
	RRMSE	8.92	51.81	2.26	10.33	5.37
$(30, 20)$	RB	-3.08	16.37	0.45	-2.95	-0.44
	RRMSE	5.94	29.8	1.40	6.73	3.22
$(50, 20)$	RB	-2.53	14.8	0.18	-0.81	0.61
	RRMSE	4.86	24.47	1.02	4.93	2.29
$(10, 40)$	RB	-0.98	9.14	0.29	-0.52	0.77
	RRMSE	5.33	33.5	1.45	7.30	3.53
$(30, 40)$	RB	-1.18	7.67	0.22	-1.33	0.08
	RRMSE	3.23	18.38	0.87	4.16	2.01
$(50, 40)$	RB	-0.70	4.7	0.08	-0.47	0.11
	RRMSE	2.36	13.31	0.69	3.61	1.67

Table S3: Performance ($\times 10^{-2}$) of estimators across Various sample sizes and measurement frequencies for model M_2.

(n, m)	Matrix	α_t	β_t	α_u	β_u	κ	σ^2
$(10, 20)$	RB	-7.16	-5.60	-6.70	-8.25	1.24	-3.76
	RRMSE	26.47	27.13	26.51	27.99	12.83	55.12
$(30, 20)$	RB	-1.98	-0.92	-2.21	1.93	0.62	0.98
	RRMSE	14.27	19.58	14.44	23.82	6.56	29.55
$(50, 20)$	RB	-0.99	1.46	-1.21	-0.52	0.45	-2.33
	RRMSE	10.50	16.98	11.36	19.74	5.92	20.66
$(10, 40)$	RB	-2.04	1.95	-1.90	2.39	1.21	-4.42
	RRMSE	14.16	15.04	14.25	17.41	10.72	42.70
$(30, 40)$	RB	-1.11	-1.52	-0.91	-1.53	-0.45	-1.09
	RRMSE	10.04	11.42	10.26	16.29	6.33	30.86
$(50, 40)$	RB	-0.07	0.27	-0.47	-0.94	-0.10	-0.08
	RRMSE	7.74	10.04	5.97	14.41	4.63	21.91

Table S4: Performance ($\times 10^{-2}$) of estimators across Various sample sizes and measurement frequencies for model M_3.

(n, m)	Matrix	α	β	κ	σ^2
$(10, 20)$	RB	0.07	0.17	-0.44	-10.54
	RRMSE	1.52	7.03	6.27	49.81
$(30, 20)$	RB	0.04	0.61	0.70	-0.33
	RRMSE	0.96	5.05	5.10	32.05
$(50, 20)$	RB	-0.14	0.57	0.03	-5.20
	RRMSE	0.75	3.08	3.12	25.92
$(10, 40)$	RB	-0.07	0.65	0.29	-17.74
	RRMSE	1.18	6.13	5.77	44.98
$(30, 40)$	RB	-0.04	0.62	0.65	-3.74
	RRMSE	0.53	2.77	3.11	26.79
$(50, 40)$	RB	-0.05	0.42	0.07	-5.11
	RRMSE	0.41	2.36	2.61	21.05

Table S5: Performance ($\times 10^{-2}$) of estimators across various sample sizes and measurement frequencies for model M_4.

(n, m)	Matrix	α	β	γ
$(10, 20)$	RB	-0.47	14.34	15.99
	RRMSE	1.56	29.48	32.06
$(30, 20)$	RB	-0.48	3.84	3.84
	RRMSE	1.06	16.73	17.28
$(50, 20)$	RB	-0.36	2.48	2.7
	RRMSE	0.81	11.14	11.67
$(10, 40)$	RB	-0.1	11.45	11.85
	RRMSE	0.97	28.27	30.83
$(30, 40)$	RB	-0.07	3.48	3.28
	RRMSE	0.66	15.37	15.33
$(50, 40)$	RB	-0.07	1.09	0.97
	RRMSE	0.45	9.65	10.33

1. Long-tail: inspection gaps gradually increase with j, constructed by $t_j = T(j/m)^2$. This reflects practical settings where inspections are more frequent in early stages but become sparse later (e.g., periodic maintenance of long-running equipment).

2. Early-tail: the opposite pattern of Long-tail, yielding dense early inspections that become sparser over time, generated by $t_j = T[1-(1-j/m)^2]$. This design mimics strategies where assets receive intensive monitoring shortly after deployment, with reduced frequency once they stabilize.

3. Irregular: total duration is fixed but adjacent intervals fluctuate randomly, $\Delta t_j = \frac{g_j}{\sum_{k=1}^{m} g_k} T$, where $g_j \sim \text{Gamma}(\text{shape} = 5, \text{rate} = 5)$. This emulates inspection schedules influenced by operational interruptions or resource availability.

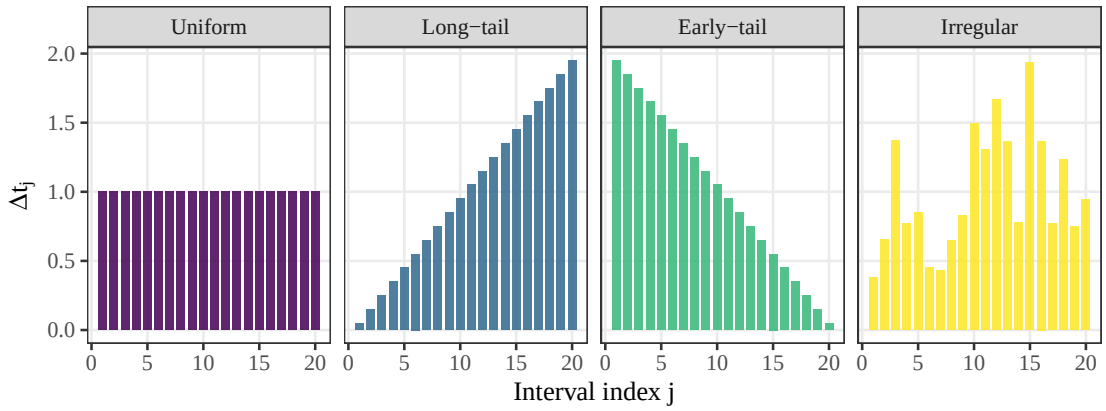


Figure S6: Illustration of four inspection–time designs considered in the simulation study: (i) uniform spacing, (ii) long-tail spacing, (iii) early-tail spacing, and (iv) irregular spacing generated from random gaps.

Figure S6 visualizes the four inspection–time designs. Under the setting $(n = 30, m = 20)$, each design is simulated independently with $B = 100$ replications, and all other model settings follow those in Section 3. Figure S7 reports the resulting RRMSE (scaled by 10^{-2}) for both EM and Bayesian estimators. Across all designs, the RRMSE levels remain similar. The uniform, long-tail, and early-tail schedules exhibit nearly identical accuracy, indicating that front-loaded or back-loaded inspection patterns have only minor influence on estimation. The Irregular design shows slightly increased error for some time-scale parameters (e.g., α_t, β_t) under the Bayesian estimator, but the magnitude remains within the same

order. Effects on κ and σ^2 are negligible. Overall, the four inspection designs yield comparable accuracy, demonstrating that the proposed method is robust to the inspection-interval pattern.

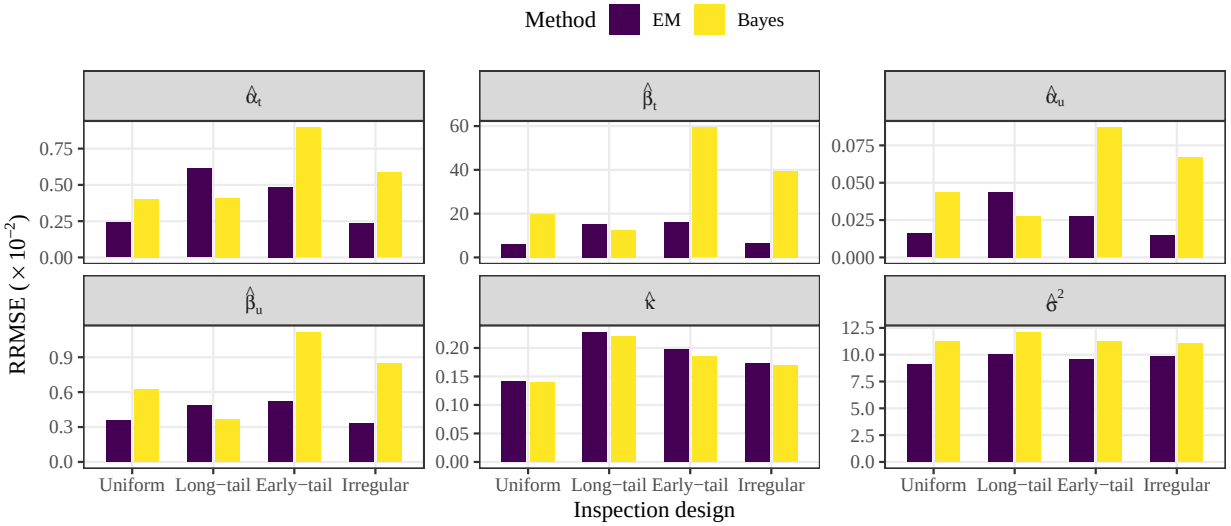


Figure S7: RRMSE ($\times 10^{-2}$) comparison of EM and Bayesian estimators under different inspection-interval designs ($n = 30$, $m = 20$).

Finally, the model does not impose additional assumptions on the number of inspection points m. The method only requires enough observations to form valid increments $\Delta Y_{ij} = Y_i(t_{j+1}, u_{j+1}) - Y_i(t_j, u_j)$. In practice, a moderate number of measurements (e.g., $m \geq 5$) is typically sufficient to ensure parameter identifiability.

S6.4 Validation under different prior specifications

In the Bayesian inference of the main paper, weakly informative normal priors $\mathcal{N}(1, 10^6)$ are assigned to $(\alpha_t, \alpha_u, \beta_t, \beta_u)$ to reflect minimal prior knowledge. Since the drift functions $\Lambda^t(t) = \beta_t t^{\alpha_t}$ and $\Lambda^u(u) = \beta_u u^{\alpha_u}$ require $\beta_t, \beta_u > 0$ to ensure the positivity of the $rIG(\theta, \gamma)$ parameter, we conduct a prior sensitivity analysis to examine whether enforcing positivity through truncated priors leads to materially different results.

Specifically, we replace the priors for (β_t, β_u) with truncated normal distributions $\beta_t, \beta_u \sim \mathcal{TN}(1, 10^6; 0, \infty)$, while keeping all other priors unchanged. Under the same simulation setting as in the main text ($n = 10$ units and $m = 20$ inspections per path), we perform 100 replications using both the normal and truncated-normal priors, and compare the RMSE of point estimates, 95% coverage probabilities, and the average lengths of the 95% credible intervals. The results are summarized in Table S6.

Across all parameters, the differences between the two prior choices are minimal: the RMSEs vary only slightly, and both the coverage rates and interval lengths remain nearly identical. This indicates that, for the data size and noise level considered here, the inference is not sensitive to whether the positivity of (β_t, β_u) is enforced through truncation. Thus, although truncated priors impose the positivity constraint more strictly, the resulting estimates are very close to those obtained under weakly informative normal priors.

Table S6: Comparison of Bayesian estimation performance under normal vs. truncated-normal priors (100 replications, $n = 10$, $m = 20$).

Para.	RMSE		95% Coverage		Avg. CI Length	
	Normal	Trunc. Normal	Normal	Trunc. Normal	Normal	Trunc. Normal
α_t	0.189	0.176	0.900	0.920	0.541	0.548
β_t	2.842	2.795	0.880	0.910	7.013	6.823
α_u	0.044	0.045	0.920	0.910	0.118	0.118
β_u	0.763	0.717	0.890	0.910	2.131	2.173
κ	0.368	0.336	0.920	0.950	1.379	1.399
σ^2	0.276	0.322	0.940	0.920	1.215	1.284

S6.5 Validation of interval estimation

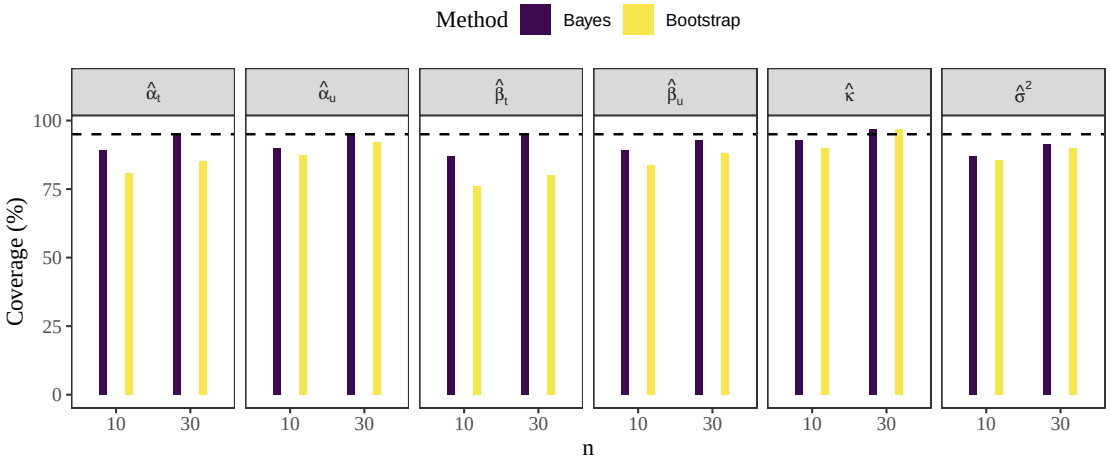


Figure S8: Coverage probabilities of the 95% confidence intervals for both Bayesian and bootstrap methods under $(n, m) = (10, 20)$ and $(30, 20)$.

To evaluate the interval estimation performance of the proposed methods, we conducted 200 replications under the settings $n = 10, 30$ and $m = 20$, and constructed the 95% confidence intervals using both the Bayesian approach and the bootstrap procedure. Figure S8 presents the coverage probabilities of model parameters. Overall, both methods yield reasonable coverage, while the Bayesian intervals exhibit noticeably more stable behavior and remain closer to the nominal 95% level across sample sizes. In contrast, the bootstrap method shows undercoverage for some parameters (e.g., β_t, β_u, and σ^2) when $n = 10$, but the performance improves markedly when n increases to 30, at which point the coverage results of the two methods become more comparable. This indicates that the Bayesian intervals are more reliable in small-sample settings, whereas the difference between the two methods diminishes as the sample size grows.

S7 Additional results of case studies

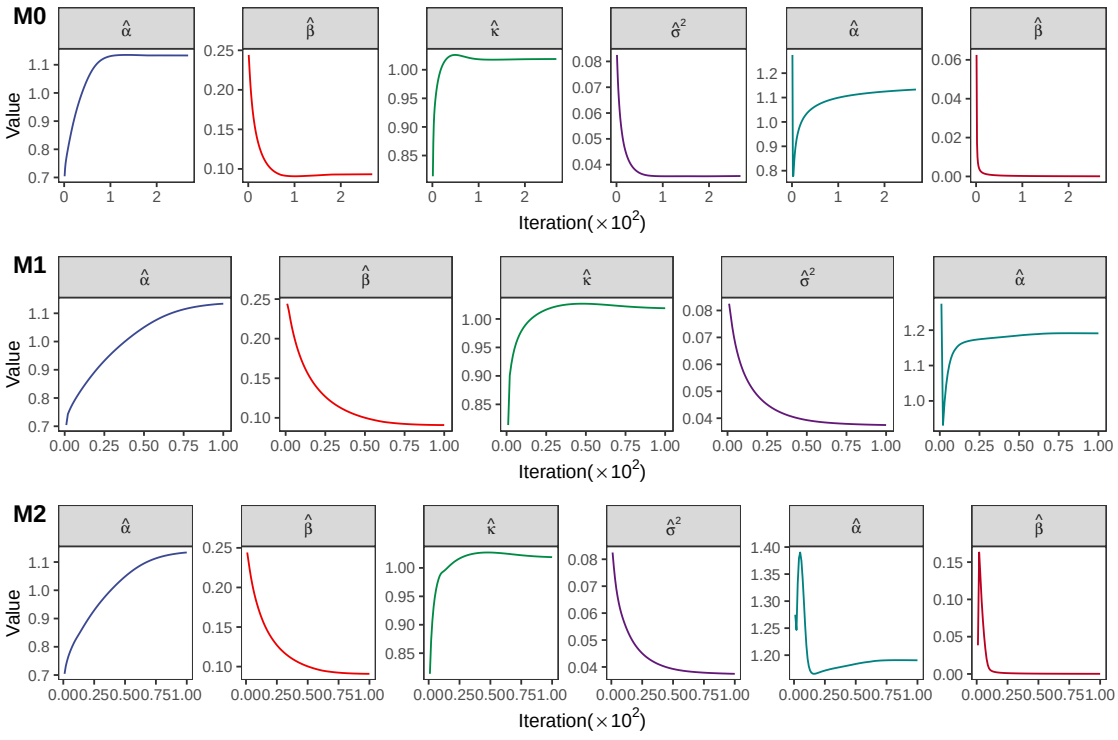


Figure S9: Iteration process of EM algorithm for parameters under two-scale models.

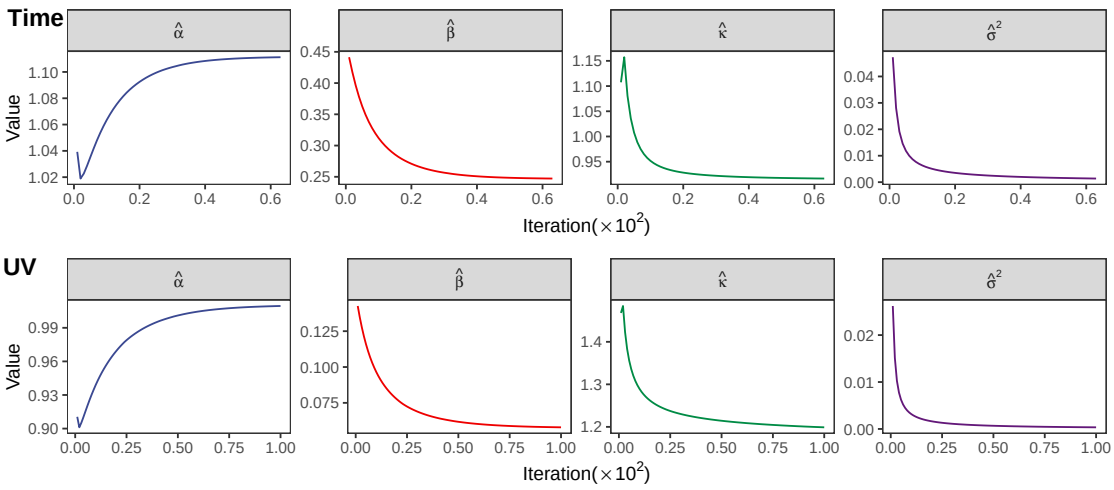


Figure S10: Iteration process of EM algorithm for parameters under M_3 with different scales.

References

Bernardo, J.M., Smith, A.F., 2009. Bayesian Theory. John Wiley & Sons.

Betancourt, M., 2017. A conceptual introduction to Hamiltonian Monte Carlo. arXiv .

Birdsall, C.K., Langdon, A.B., 2018. Plasma Physics Via Computer Simulation. CRC press.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M.A., Guo, J., Li, P., Riddell, A., 2017. Stan: A probabilistic programming language. Journal of statistical software 76. doi:10.18637/jss.v076.i01.

Fang, G., Pan, R., Wang, Y., He, H., Yu, W., 2024. Exploring initiation-growth correlation in accelerated degradation testing data: Both classical and Bayesian approaches. IEEE Transactions on Reliability doi:10.1109/TR.2024.3471796.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman and Hall/CRC.

Gilks, W., Best, N., Tan, K., 2022. Adaptive rejection Metropolis sampling within gibbs sampling. Applied Statistics 44, 455–472. doi:10.2307/2986138.

Golub, G.H., Welsch, J.H., 1969. Calculation of Gauss quadrature rules. Mathematics of Computation 23, 221–230.

Novomestky, F., 2022. gaussquad: collection of functions for Gaussian quadrature. doi:10.32614/CRAN.package.gaussquad.

Puli, V.K., Chiplunkar, R., Huang, B., 2023. Sparse robust dynamic feature extraction using Bayesian inference. IEEE Transactions on Industrial Electronics 71, 6201–6209. doi:10.1109/TIE.2023.3290235.

Si, X.S., Zhou, D., 2013. A generalized result for degradation model-based reliability estimation. IEEE Transactions on Automation Science and Engineering 11, 632–637. doi:10.1109/TASE.2013.2260740.

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.

Taylor, D., Rigdon, S.E., Pan, R., Montgomery, D.C., 2024. Bayesian D-optimal design for life testing with censoring. Quality and Reliability Engineering International 40, 71–90. doi:10.1002/qre.322.