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

Liangliang Zhuang, Yizhong Ma, Guanqi Fang, Ancha Xu

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.

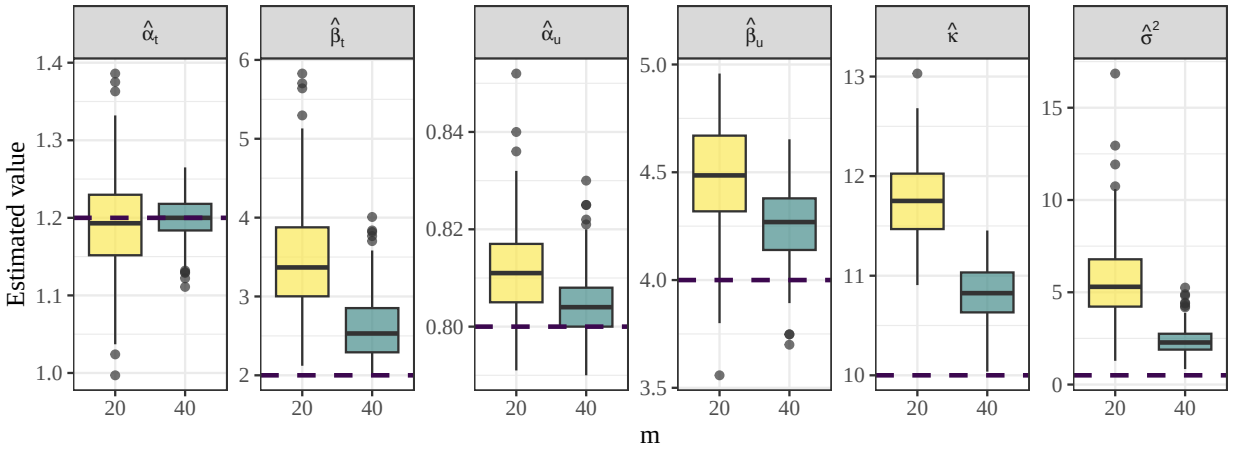


Figure S1: Performance of Initial Parameter Estimation (True Values Indicated by Purple Dashed Line).

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.

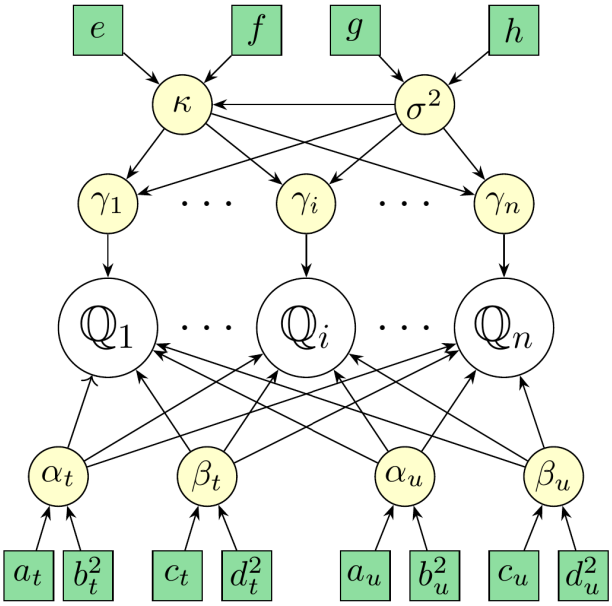


Figure S2: Dependency Diagram of Variables in Bayesian Framework: Arrows in the Directed graph Indicate Direct Influences. Green Rectangles Represent Hyperparameters, Yellow Circles Denote Variables, and White Circles Indicate Observed Data.

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, 2018). 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

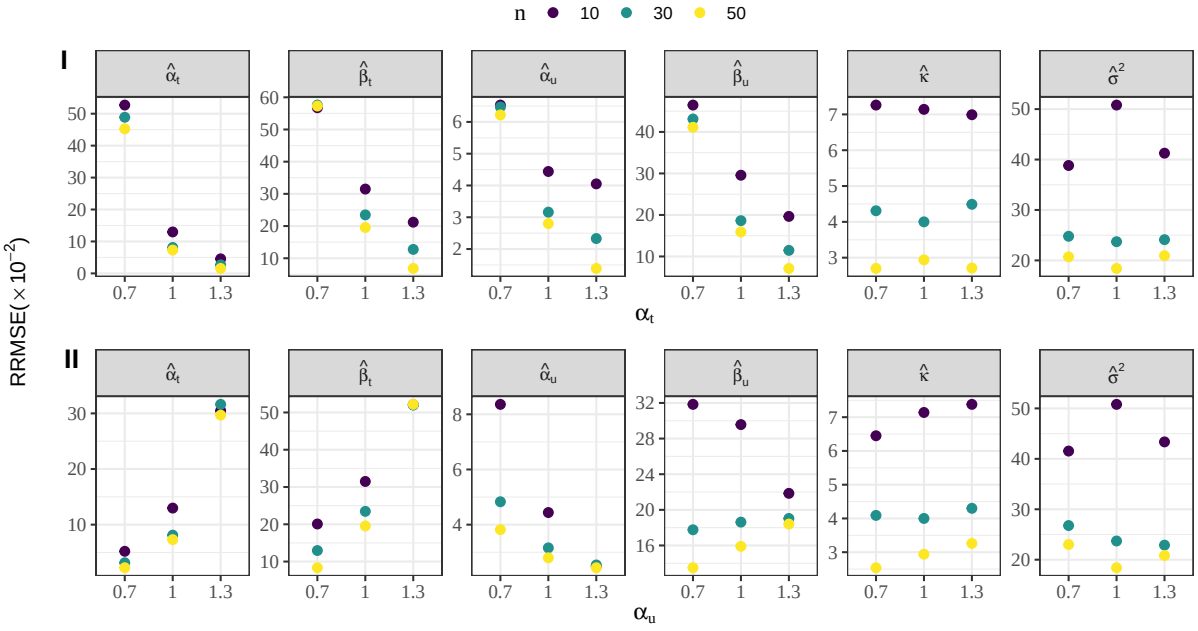
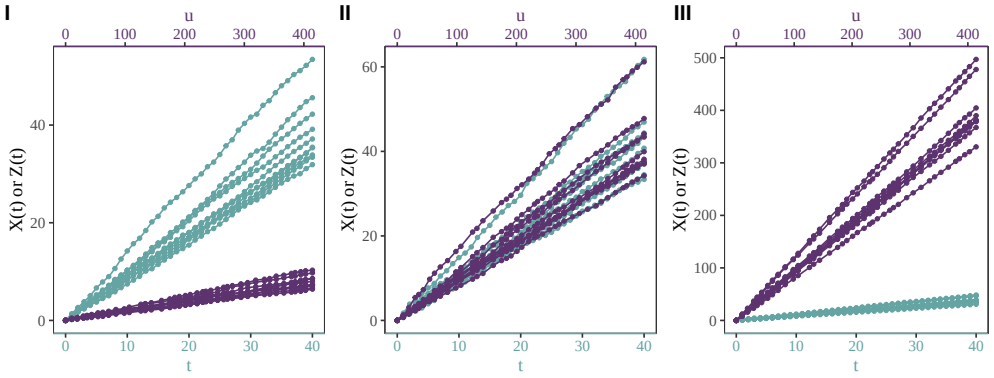


Figure S3: RRMSE of estimators for different degradation curvatures (scenario I: varying α_t, fixed $\alpha_u = 1$; scenario II: varying α_u, fixed $\alpha_t = 1$).

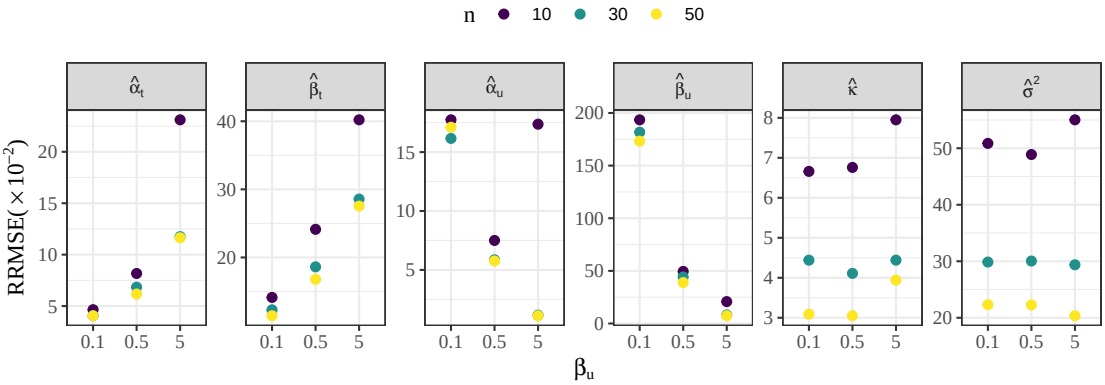
t-related parameters ($\hat{\alpha}_t$ and $\hat{\beta}_t$) with lower RRMSE. Conversely, the smaller degradation on the u-scale leads to poorer estimates for u-related parameters. As β_u increases, the u-scale degradation becomes more significant, improving the estimation for u-related parameters, while 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 $\beta_u = 5$.

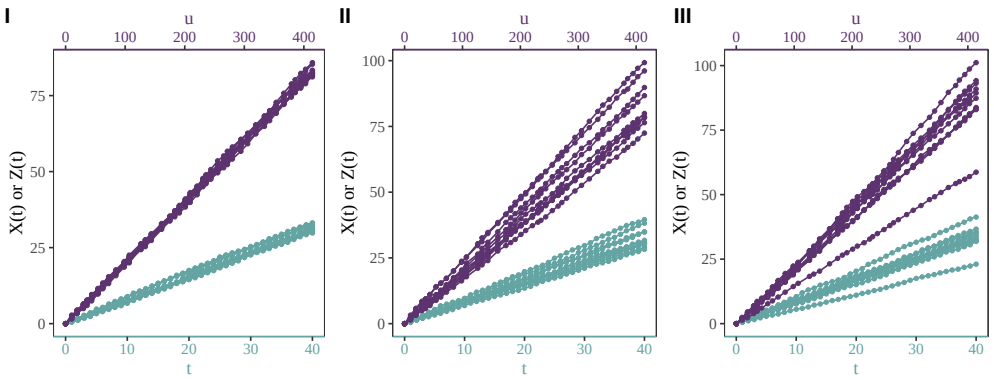
especially as the sample size increases, with RRMSE remaining largely consistent.

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

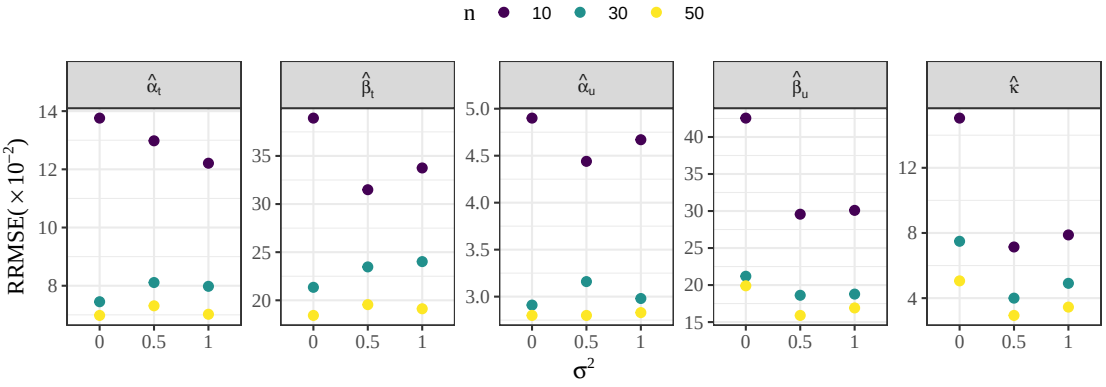
The parameter estimation performance for different model extensions is presented. Models M_1 to M_3 employ the EM algorithm, while M_4 uses the traditional MLE method. The corresponding results are shown in Tables S2 to S5.

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.

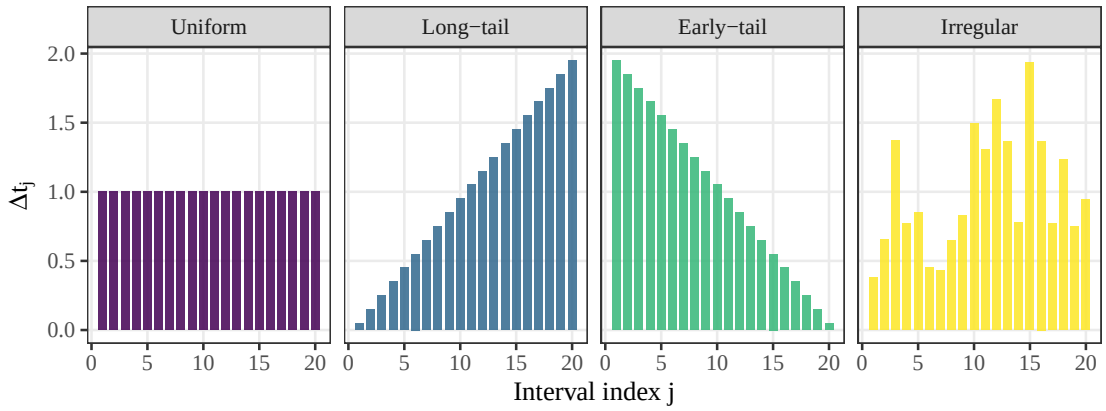


Figure S6: Illustration of four inspection–time designs considered in the simulation study: (i) uniform spacing, (ii) long-tail spacing, (iii) early-tail spacing, and (iv) irregular spacing generated from random gaps.

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.

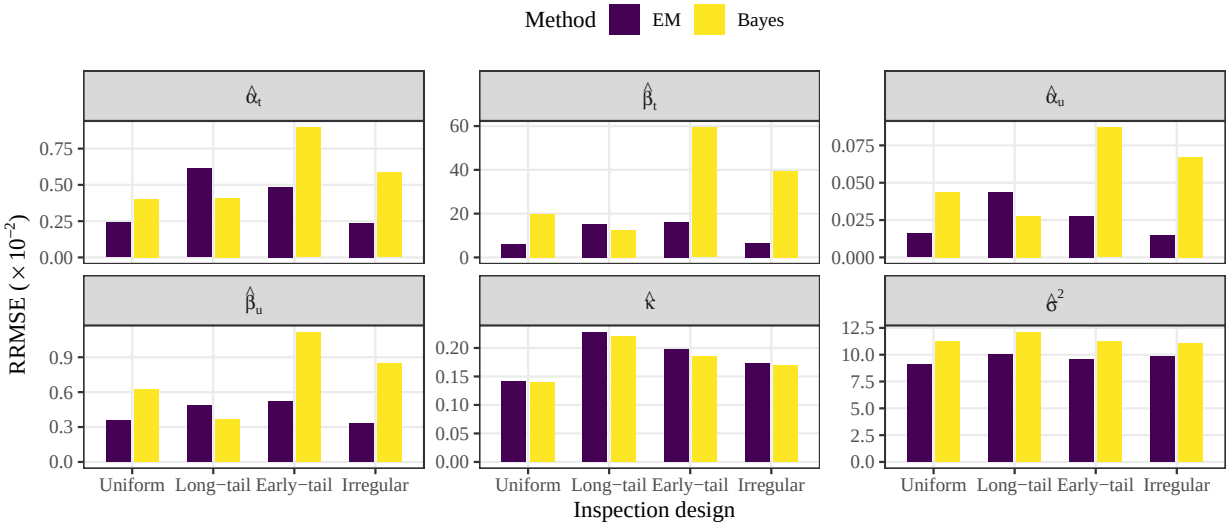


Figure S7: RRMSE ($\times 10^{-2}$) comparison of EM and Bayesian estimators under different inspection-interval designs ($n = 30$, $m = 20$).

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

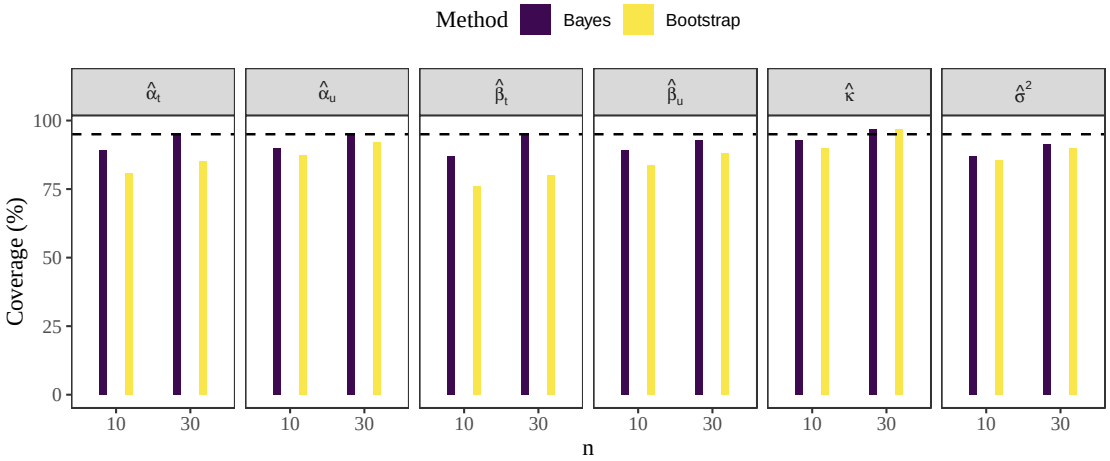


Figure S8: Coverage probabilities of the 95% confidence intervals for both Bayesian and bootstrap methods under $(n, m) = (10, 20)$ and $(30, 20)$.

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

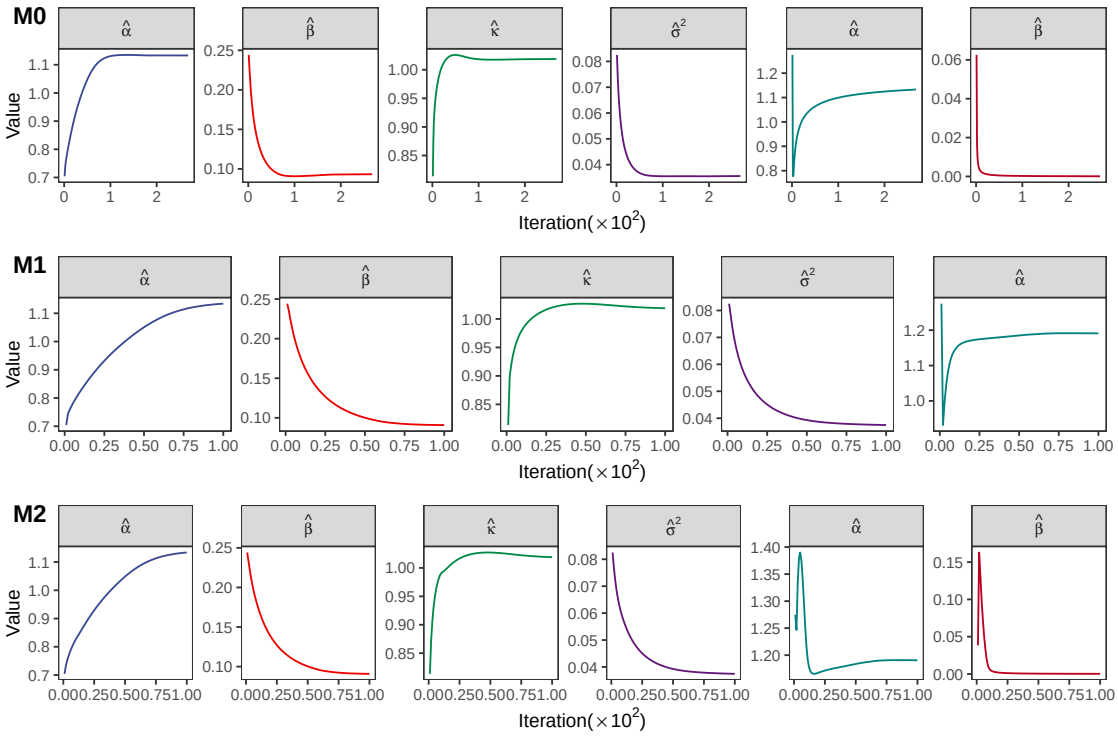


Figure S9: Iteration process of EM algorithm for parameters under two-scale models.

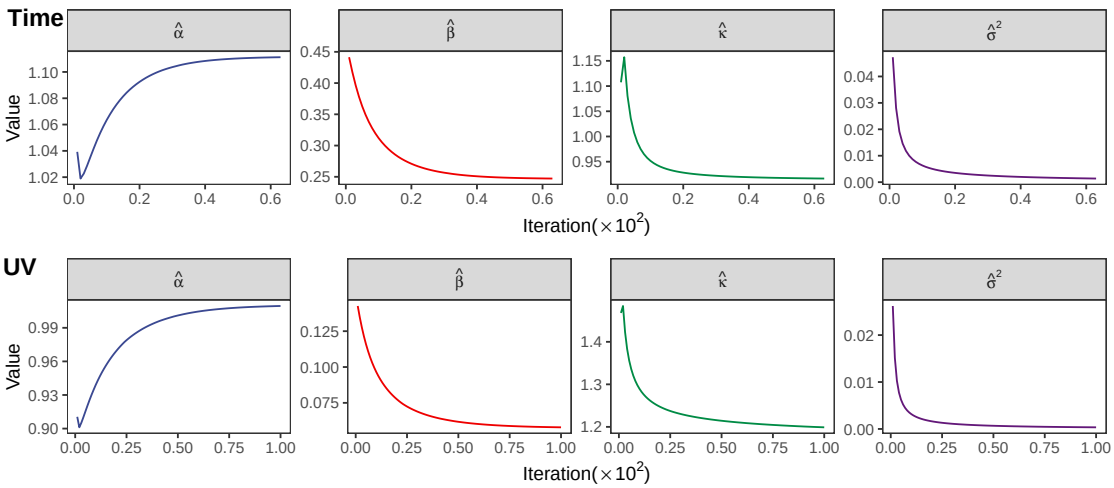


Figure S10: Iteration process of EM algorithm for parameters under M_3 with different scales.

References

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

Betancourt, M., 2018. A conceptual introduction to Hamiltonian Monte Carlo. arXiv:1701.02434.

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.