

Supplement to “Statistical Modeling of Multivariate Destructive Degradation Tests with Blocking”

Qiuzhuang Sun[†], Zhi-Sheng Ye[†], and Yili Hong[‡]

[†]Department of Industrial Systems Engineering and Management, National University of
Singapore

[‡]Department of Statistics, Virginia Tech

S.1 Expectation-Maximization (EM) Algorithm

S.1.1 Technical Details

M-step: By taking the first-order derivatives of the Q -function and letting them equal zero, the $\boldsymbol{\mu}$ parameter is updated as

$$\boldsymbol{\mu}^{(\tau+1)} = \frac{\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \left[u_i^{(\tau)} \mathbf{Y}_{ijk} - c_{ij}^{(\tau)} \mathbf{1}_d \right]}{K \sum_{i=1}^n \sum_{j=1}^m v_i^{(\tau)} t_j}. \quad (\text{S.1})$$

The covariance matrix $\boldsymbol{\Sigma}$ is updated as

$$\begin{aligned} \boldsymbol{\Sigma}^{(\tau+1)} = & \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^m \frac{1}{t_j} \sum_{k=1}^K \left[\mathbf{Y}_{ijk} \mathbf{Y}_{ijk}' - a_{ij}^{(\tau)} (\mathbf{Y}_{ijk} \mathbf{1}_d' + \mathbf{1}_d \mathbf{Y}_{ijk}') + b_{ij}^{(\tau)} \mathbf{1}_d \mathbf{1}_d' \right. \\ & \left. + c_{ij}^{(\tau)} t_j (\mathbf{1}_d \boldsymbol{\mu}'^{(\tau+1)} + \boldsymbol{\mu}^{(\tau+1)} \mathbf{1}_d') - u_i^{(\tau)} t_j (\mathbf{Y}_{ijk} \boldsymbol{\mu}'^{(\tau+1)} + \boldsymbol{\mu}^{(\tau+1)} \mathbf{Y}_{ijk}') + v_i^{(\tau)} t_j^2 \boldsymbol{\mu}^{(\tau+1)} \boldsymbol{\mu}'^{(\tau+1)} \right], \end{aligned} \quad (\text{S.2})$$

and the other parameters are updated as

$$\kappa^{(\tau+1)} = \left(\frac{\sum_{i=1}^n \sum_{j=1}^m b_{ij}^{(\tau)}}{nm} \right)^{\frac{1}{2}} \quad \text{and} \quad \omega^{(\tau+1)} = \left(\frac{\sum_{i=1}^n v_i^{(\tau)} - 2u_i^{(\tau)}}{n} + 1 \right)^{\frac{1}{2}},$$

where

$$u_i^{(\tau)} = \mathbb{E}[\zeta_i | \boldsymbol{\theta}^{(\tau)}, \mathbb{D}], \quad v_i^{(\tau)} = \mathbb{E}[\zeta_i^2 | \boldsymbol{\theta}^{(\tau)}, \mathbb{D}],$$

$$a_{ij}^{(\tau)} = \mathbb{E}[\epsilon_{ij} | \boldsymbol{\theta}^{(\tau)}, \mathbb{D}], \quad b_{ij}^{(\tau)} = \mathbb{E}[\epsilon_{ij}^2 | \boldsymbol{\theta}^{(\tau)}, \mathbb{D}], \quad \text{and} \quad c_{ij}^{(\tau)} = \mathbb{E}[\zeta_i \epsilon_{ij} | \boldsymbol{\theta}^{(\tau)}, \mathbb{D}].$$

A numerical issue is the positive-definiteness of the covariance matrix $\boldsymbol{\Sigma}$. Since our EM algorithm starts with an educated guess of model parameters as shown in Section S.1.2, the initial parameter estimation should be close to the true value so that $\boldsymbol{\Sigma}$ should be positive-definite for all iterations. Our simulation experiences show that the positive-definiteness of $\boldsymbol{\Sigma}$ is ensured for all iterations of the EM algorithm under different parameter settings. In practice, the positive-definiteness of $\boldsymbol{\Sigma}$ can be checked in each iteration of the EM algorithm. In case that the corresponding maximizer is not positive-definite in an iteration, we can re-parameterize $\boldsymbol{\Sigma}$ using its Cholesky decomposition. The maximization can be then numerically done.

E-step: The M-step needs values of $u_i^{(\tau)}$, $v_i^{(\tau)}$, $a_{ij}^{(\tau)}$, $b_{ij}^{(\tau)}$, and $c_{ij}^{(\tau)}$, $i \in [n]$, $j \in [m]$. To this end, the E-step derives the distribution of the missing data, which is capsulized as follows.

Theorem 1 *In the i th rig, the distribution of $[\epsilon_{i1}, \dots, \epsilon_{im}, \zeta_i]'$ conditional on $\mathbb{D}$ and $\boldsymbol{\theta}$ is multivariate normal with mean vector $\boldsymbol{\nu}_i$ and covariance matrix $\boldsymbol{\Xi}$, $i \in [n]$, where*

$$\boldsymbol{\nu}_i = \boldsymbol{\Xi} \begin{bmatrix} \mathbf{D}' \boldsymbol{\Sigma}_{\mathbf{x}}^{-1} \mathbf{y}_i \\ \mathbf{y}_i' \boldsymbol{\Sigma}_{\mathbf{x}}^{-1} \boldsymbol{\mu}_{\mathbf{y}} + \omega^{-2} \end{bmatrix} \quad (\text{S.3})$$

with

$$\mathbf{D} = \text{diag}(\mathbf{1}_{Kd}, \dots, \mathbf{1}_{Kd}) = \mathbf{I}_m \otimes \mathbf{1}_{Kd} \in \mathcal{R}^{M \times m} \quad (\text{S.4})$$

and

$$\mathbf{\Sigma}_{\mathcal{X}} = \text{diag}(\underbrace{\mathbf{\Sigma}t_1, \dots, \mathbf{\Sigma}t_1}_{K \text{ repetitions}}, \dots, \mathbf{\Sigma}t_m, \dots, \mathbf{\Sigma}t_m) = \text{diag}(\mathbf{t} \otimes \mathbf{1}_K) \otimes \mathbf{\Sigma} \in \mathcal{R}^{M \times M}. \quad (\text{S.5})$$

The expression of $\mathbf{\Xi}$ is given in the proof.

Proof Let $\boldsymbol{\epsilon}_i = [\epsilon_{i1}, \dots, \epsilon_{im}]'$. Following Bayes' theorem, we have

$$\begin{aligned} p(\boldsymbol{\epsilon}_i, \zeta_i | \boldsymbol{\theta}, \mathbb{D}) &\propto p(\boldsymbol{\epsilon}_i, \zeta_i, \mathbb{D} | \boldsymbol{\theta}) \\ &\propto \exp \left\{ -\frac{1}{2} \left[(\mathbf{y}_i - \mathbf{D}\boldsymbol{\epsilon}_i - \zeta_i \boldsymbol{\mu}_{\mathbf{y}})' \mathbf{\Sigma}_{\mathcal{X}}^{-1} (\mathbf{y}_i - \mathbf{D}\boldsymbol{\epsilon}_i - \zeta_i \boldsymbol{\mu}_{\mathbf{y}}) + \frac{(\zeta_i - 1)^2}{\omega^2} + \boldsymbol{\epsilon}_i' \mathbf{\Sigma}_{\epsilon}^{-1} \boldsymbol{\epsilon}_i \right] \right\}, \end{aligned} \quad (\text{S.6})$$

where $\mathbf{\Sigma}_{\epsilon} = \kappa^2 \mathbf{I}_m \in \mathcal{R}^{m \times m}$. To find the quadratic form of (S.6), we should compare it with

$$p(\boldsymbol{\epsilon}_i, \zeta_i | \boldsymbol{\theta}, \mathbb{D}) \propto \exp \left[-\frac{1}{2} ([\boldsymbol{\epsilon}_i', \zeta_i]' - [\boldsymbol{\nu}_{i1}', \nu_{i2}']')' \mathbf{C} ([\boldsymbol{\epsilon}_i', \zeta_i]' - [\boldsymbol{\nu}_{i1}', \nu_{i2}']') \right], \quad (\text{S.7})$$

where

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & C_{22} \end{bmatrix}.$$

Here, $\boldsymbol{\nu}_{i1} \in \mathcal{R}^m$, $\nu_{i2} \in \mathcal{R}$, $\mathbf{C}_{11} \in \mathcal{R}^{m \times m}$, $\mathbf{C}_{12} \in \mathcal{R}^m$, and $C_{22} \in \mathcal{R}$. Solving (S.6) and (S.7) yields

$$\boldsymbol{\nu}_{i1} = \mathbf{C}^{-1} \mathbf{D}' \mathbf{\Sigma}_{\mathcal{X}}^{-1} \mathbf{y}_i, \quad \nu_{i2} = \mathbf{C}^{-1} (\mathbf{y}_i' \mathbf{\Sigma}_{\mathcal{X}}^{-1} \boldsymbol{\mu}_{\mathbf{y}} + \omega^{-2}),$$

$$\mathbf{C}_{11} = \mathbf{D}' \mathbf{\Sigma}_{\mathcal{X}}^{-1} \mathbf{D} + \mathbf{\Sigma}_{\epsilon}^{-1}, \quad \mathbf{C}_{12} = \mathbf{D}' \mathbf{\Sigma}_{\mathcal{X}}^{-1} \boldsymbol{\mu}_{\mathbf{y}}, \quad \text{and} \quad C_{22} = \boldsymbol{\mu}_{\mathbf{y}}' \mathbf{\Sigma}_{\mathcal{X}}^{-1} \boldsymbol{\mu}_{\mathbf{y}} + \omega^{-2}.$$

Then $\mathbf{\Xi}$ in the theorem is the inverse of $\mathbf{C}$. Let $S = C_{22} - \mathbf{C}_{12}' \mathbf{C}_{11} \mathbf{C}_{12}$ be the Schur complement of $\mathbf{C}_{11}$. We can calculate $\mathbf{\Xi} = \mathbf{C}^{-1}$ as

$$\mathbf{\Xi} = \begin{bmatrix} \mathbf{C}_{11}^{-1} + \mathbf{C}_{11}^{-1} \mathbf{C}_{12} \mathbf{C}_{12}' \mathbf{C}_{11}^{-1} / S & -\mathbf{C}_{11} \mathbf{C}_{12} / S \\ \text{symmetric} & 1/S \end{bmatrix}. \quad (\text{S.8})$$

Let $\mathbf{F} = \mathbf{C}_{11}^{-1} = (\mathbf{D}'\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{D} + \boldsymbol{\Sigma}_{\epsilon}^{-1})^{-1}$. Then we have the same form of $\boldsymbol{\nu}_i$ as shown in Theorem 1 and $\boldsymbol{\Xi}$ is given by

$$\boldsymbol{\Xi} = \begin{bmatrix} \mathbf{F} + \mathbf{F}(\mathbf{D}'\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\boldsymbol{\mu}_{\mathbf{y}}\boldsymbol{\mu}_{\mathbf{y}}'\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{D})\mathbf{F}/S & -\mathbf{F}\mathbf{D}'\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\boldsymbol{\mu}_{\mathbf{y}}/S \\ \text{symmetric} & 1/S \end{bmatrix}. \quad \blacksquare$$

Since the marginal distribution of a multivariate normal distribution is still normal, the values of $u_i^{(\tau)}$, $v_i^{(\tau)}$, $a_{ij}^{(\tau)}$, $b_{ij}^{(\tau)}$, and $c_{ij}^{(\tau)}$, for $i \in [n]$ and $j \in [m]$, can be readily obtained from Theorem 1 by replacing $\boldsymbol{\theta}$ with $\boldsymbol{\theta}^{(\tau)}$. Specifically, the values of $u_i^{(\tau)}$ and $a_{ij}^{(\tau)}$ are determined by $\boldsymbol{\nu}_i$. We can also derive the values of $v_i^{(\tau)}$, $b_{ij}^{(\tau)}$, and $c_{ij}^{(\tau)}$ from $\boldsymbol{\Xi}$, which completes the E-step.

S.1.2 Determination of Starting Points for the EM Algorithm

The determination of starting point for the EM algorithm includes three steps. First note that the rig-layer block effect is an unknown but fixed value for each test rig. We can then obtain a rough estimate of degradation rate in the i th rig, denoted as $\hat{\boldsymbol{\mu}}_i$, $i \in [n]$. This can be done by resorting to the multivariate least square regression model (Renchner, 2003, Chapter 10), which minimizes the sum of squared error

$$\text{SSE} = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K (\mathbf{Y}_{ijk} - \boldsymbol{\mu}_i t_j)' (\mathbf{Y}_{ijk} - \boldsymbol{\mu}_i t_j).$$

According to the values of $\hat{\boldsymbol{\mu}}_i$, $i \in [n]$, we can determine the starting points for $\boldsymbol{\mu}$ and ω . For each test rig, we can estimate the covariance matrix by first assuming $\kappa = 0$, i.e., no gauge-layer block effect. Based on this covariance matrix, we can further determine the starting points for κ . The detailed procedure is as follows.

- For the i th test rig, get rough estimates of $\boldsymbol{\mu}_i$ using the multivariate least squared estimation as $\hat{\boldsymbol{\mu}}'_i = \sum_{k=1}^K (\mathbf{t}'\mathbf{t})^{-1} \mathbf{t}'\mathbf{Y}_{ik}/K$, where $\mathbf{Y}_{ik} = [\mathbf{Y}_{i1k}, \dots, \mathbf{Y}_{imk}]'$. The starting point of $\boldsymbol{\mu}$ for the EM algorithm is obtained as $\boldsymbol{\mu}^{(0)} = \sum_{i=1}^n \hat{\boldsymbol{\mu}}_i/n$.

- Get rough estimates of $\zeta_1, \dots, \zeta_n$ as $\hat{\zeta}_i = \frac{1}{d} \sum_{l=1}^d \hat{\mu}_{il} / \hat{\mu}_i$, where $\hat{\mu}_{il}$ and $\hat{\mu}_i$ are the l th coordinate of $\hat{\boldsymbol{\mu}}_i$ and the i th coordinate of $\hat{\boldsymbol{\mu}}$, respectively. Fit $\hat{\zeta}_1, \dots, \hat{\zeta}_n$ with a normal distribution $\mathcal{N}(1, \omega^2)$ to get the starting point $\omega^{(0)}$ for the EM algorithm.
- First get a rough estimate of the covariance matrix assuming that there is no gauge-layer block effect, i.e., $\kappa = 0$. It yields

$$\tilde{\boldsymbol{\Sigma}} = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \left(\frac{\mathbf{Y}_{ijk} - \hat{\boldsymbol{\mu}}_i t_j}{\sqrt{t_j}} \right) \left(\frac{\mathbf{Y}_{ijk} - \hat{\boldsymbol{\mu}}_i t_j}{\sqrt{t_j}} \right)'. \quad (\text{S.9})$$

Then we numerically maximize the likelihood function conditioning on the values of $\hat{\boldsymbol{\mu}}_i$, $i \in [n]$, to get the starting points $\boldsymbol{\Sigma}^{(0)}$ and $\kappa^{(0)}$. The above maximization can be efficiently done by many algorithms (e.g., the Nelder-Mead algorithm), where the starting points of covariance matrix for the numerical maximization is $\tilde{\boldsymbol{\Sigma}}$ in (S.9).

S.1.3 Inference with the Time Scale Transformation Function

This section develops the EM algorithm when there are unknown parameters in the time scale transformation function $\mathbf{H}(t) = \text{diag}(h_1(t), \dots, h_d(t))$. For the E-step, the distribution of missing data $[\epsilon_{i1}, \dots, \epsilon_{im}, \zeta_i]'$ conditional on the observed data $\mathbb{D}$ and model parameter $\boldsymbol{\theta}$ is still multivariate normal with mean vector $\boldsymbol{\nu}_i$ and covariance matrix $\boldsymbol{\Xi}$ as shown in Theorem 1. The only difference is that

$$\boldsymbol{\mu}_{\mathbf{y}} = [\underbrace{\boldsymbol{\mu}'\mathbf{H}(t_1), \dots, \boldsymbol{\mu}'\mathbf{H}(t_1)}_{K \text{ repetitions}}, \dots, \underbrace{\boldsymbol{\mu}'\mathbf{H}(t_m), \dots, \boldsymbol{\mu}'\mathbf{H}(t_m)}_{K \text{ repetitions}}]'$$

and

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \text{diag}(\underbrace{\boldsymbol{\Sigma}^{1/2}\mathbf{H}(t_1)\boldsymbol{\Sigma}^{1/2}, \dots, \boldsymbol{\Sigma}^{1/2}\mathbf{H}(t_1)\boldsymbol{\Sigma}^{1/2}}_{K \text{ repetitions}}, \dots, \boldsymbol{\Sigma}^{1/2}\mathbf{H}(t_m)\boldsymbol{\Sigma}^{1/2}, \dots, \boldsymbol{\Sigma}^{1/2}\mathbf{H}(t_m)\boldsymbol{\Sigma}^{1/2}).$$

For the M-step, the estimates $\kappa^{(\tau+1)}$ and $\omega^{(\tau+1)}$ are still updated as shown in Section S.1.1.

In comparison, $\boldsymbol{\mu}^{(\tau+1)}$ is updated as

$$\boldsymbol{\mu}^{(\tau+1)} = \left(K \sum_{i=1}^n \sum_{j=1}^m v_i^{(\tau)} \mathbf{H}^{(\tau+1)}(t_j) \right)^{-1} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \left[u_i^{(\tau)} \mathbf{Y}_{ijk} - c_{ij}^{(\tau)} \mathbf{1}_d \right], \quad (\text{S.10})$$

where $\mathbf{H}^{(\tau+1)}(\cdot)$ is the maximizer of $\mathbf{H}(\cdot)$ in the M-step. When all dimensions of degradation share a common transformed time scale but with unknown parameters, we can still have the closed-form maximizer of $\boldsymbol{\Sigma}$ similar to that as shown in the main text. Specifically, denote the common transformed time scale as $h(t) \equiv h_l(t)$, $l \in [d]$. The maximizers of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ at the $(\tau + 1)$ st iteration are given by replacing t_j in (S.1) and (S.2) with $h^{(\tau+1)}(t_j)$. Here, $h^{(\tau+1)}(\cdot)$ is the maximizer of $h(\cdot)$ at the $(\tau + 1)$ st iteration, which is given by solving the following equation (it is obtained by taking the first derivative of the likelihood function with respect to the unknown parameters in $h(\cdot)$ and letting it equal zero)

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \left\{ \left[\frac{d}{h(t_j)} - (\mathbf{Y}'_{ijk}(\boldsymbol{\Sigma}^{(\tau)})^{-1} \mathbf{Y}_{ijk} - 2a_{ij}^{(\tau)} \mathbf{Y}'_{ijk}(\boldsymbol{\Sigma}^{(\tau)})^{-1} \mathbf{1}_d \right. \right. \\ \left. \left. + b_{ij}^{(\tau)} \mathbf{1}'_d (\boldsymbol{\Sigma}^{(\tau)})^{-1} \mathbf{1}_d) h^{-2}(t_j) + v_i^{(\tau)} (\boldsymbol{\mu}^{(\tau)})' (\boldsymbol{\Sigma}^{(\tau)})^{-1} \boldsymbol{\mu}^{(\tau)} \right] \nabla h(t_j) \right\} = \mathbf{0},$$

where $\nabla h(t)$ is the gradient of $h(t)$ (assuming it exists) and $\mathbf{0}$ is the vector of zero with the same length of $\nabla h(t)$. If the time scale of each dimension of degradation is different, we cannot have closed-form maximizers for the parameters $\boldsymbol{\Sigma}^{(\tau+1)}$ and $\mathbf{H}^{(\tau+1)}(\cdot)$ in the M-step. Nevertheless, these maximizers can be numerically found by maximizing

$$\ell_c(\boldsymbol{\Sigma}, \mathbf{p}) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \left\{ \log[\det(\bar{\boldsymbol{\Sigma}}_j)] \right. \\ \left. + \mathbf{Y}'_{ijk} \bar{\boldsymbol{\Sigma}}_j^{-1} \mathbf{Y}_{ijk} + b_{ij}^{(\tau)} \mathbf{1}'_d \bar{\boldsymbol{\Sigma}}_j^{-1} \mathbf{1}_d + v_i^{(\tau)} \boldsymbol{\mu}(\mathbf{p})' \mathbf{H}(t_j) \bar{\boldsymbol{\Sigma}}_j^{-1} \mathbf{H}(t_j) \boldsymbol{\mu}(\mathbf{p}) \right. \\ \left. - 2a_{ij}^{(\tau)} \mathbf{Y}'_{ijk} \bar{\boldsymbol{\Sigma}}_j^{-1} \mathbf{1}_d - 2u_i^{(\tau)} \mathbf{Y}'_{ijk} \bar{\boldsymbol{\Sigma}}_j^{-1} \mathbf{H}(t_j) \boldsymbol{\mu}(\mathbf{p}) + 2c_{ij}^{(\tau)} \mathbf{1}'_d \bar{\boldsymbol{\Sigma}}_j^{-1} \mathbf{H}(t_j) \boldsymbol{\mu}(\mathbf{p}) \right\},$$

where $\bar{\boldsymbol{\Sigma}}_j = \boldsymbol{\Sigma}^{1/2} \mathbf{H}(t_j) \boldsymbol{\Sigma}^{1/2}$, $\mathbf{p}$ are the unknown parameters in $\mathbf{H}(\cdot)$, and $\boldsymbol{\mu}(\mathbf{p})$ is the function

following the form of (S.10). To ensure the positive-definiteness of the covariance matrix Σ in each iteration of the EM algorithm, we re-parameterize it as $\Sigma = \mathbf{L}\mathbf{L}'$, where $\mathbf{L}$ is a lower triangular matrix obtained from the corresponding Cholesky decomposition.

For the effectiveness of the EM algorithm, we also need to obtain some educated guesses for its starting points. The procedure provided in Section S.1.2 needs some slight changes. First, we fit the degradation data on each dimension using the univariate mixed-effect Wiener process (Wang, 2010) to obtain the initial estimates of the degradation rate $\boldsymbol{\mu}^{(0)}$ and the corresponding parameters $\mathbf{p}^{(0)}$ in $\mathbf{H}(\cdot)$. Then the starting point for $\omega^{(0)}$ can be obtained as the same as the linear case. Based on these initial estimates, the starting points $\Sigma^{(0)}$ and $\kappa^{(0)}$ can also be found similarly with the maximization of the corresponding likelihood function.

S.2 Interval Estimation

S.2.1 Derivation of Equation (5) in the Main Text

The matrix $\mathbf{A}_k$ is given by

$$\mathbf{A}_k \equiv (\Sigma_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \left[\frac{\partial \Sigma_{\mathbf{y}}}{\partial \theta_k} + 2\omega \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}' \frac{\partial \omega}{\partial \theta_k} + \omega^2 \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \boldsymbol{\mu}_{\mathbf{y}}' + \boldsymbol{\mu}_{\mathbf{y}} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}'}{\partial \theta_k} \right) \right].$$

We derive the above equation and Equation (5) in the main text as follows.

Without loss of generality, we derive the Fisher information matrix for the case $d = 2$. The derivation can be readily generalized to any value of d . When $d = 2$, we have $\boldsymbol{\mu} = [\mu_1, \mu_2]'$ and denote the covariance matrix as

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Therefore, the model parameters are $\boldsymbol{\theta} = [\mu_1, \mu_2, \sigma_1, \sigma_2, \rho, \omega, \kappa]'$. Computing the first order

derivative of the likelihood function in Equation (4) of the main text, we have

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_k} = & -\frac{n}{2} \text{tr}(\mathbf{A}_k) + \frac{1}{2} \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}})' \mathbf{A}_k (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) \\ & + \sum_{i=1}^n \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}). \end{aligned}$$

Let

$$\eta_{ik} = \frac{1}{2} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}})' \mathbf{A}_k (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) + \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}). \quad (\text{S.11})$$

Note that $\mathbb{E} \left[\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_k} \right] = 0$ for all k . The (k, l) th entry of $\mathcal{I}(\boldsymbol{\theta})$ can be obtained as

$$\begin{aligned} [\mathcal{I}(\boldsymbol{\theta})]_{k,l} &= \mathbb{E} \left[\left(-\frac{n}{2} \text{tr}(\mathbf{A}_k) + \sum_{i=1}^n \eta_{ik} \right) \left(-\frac{n}{2} \text{tr}(\mathbf{A}_l) + \sum_{i=1}^n \eta_{il} \right) \right] \\ &= \frac{n^2}{4} \text{tr}(\mathbf{A}_k) \text{tr}(\mathbf{A}_l) - \frac{n}{2} \text{tr}(\mathbf{A}_l) \mathbb{E} \left[\sum_{i=1}^n \eta_{ik} \right] - \frac{n}{2} \text{tr}(\mathbf{A}_k) \mathbb{E} \left[\sum_{i=1}^n \eta_{il} \right] + \mathbb{E} \left[\sum_{i=1}^n \eta_{ik} \sum_{j=1}^n \eta_{jl} \right] \\ &= -\frac{n^2}{4} \text{tr}(\mathbf{A}_k) \text{tr}(\mathbf{A}_l) + \mathbb{E} \left[\sum_{i \neq j} \eta_{ik} \eta_{jl} + \sum_{i=1}^n \eta_{ik} \eta_{il} \right], \quad (\text{S.12}) \end{aligned}$$

where the last equality holds since $\mathbf{A}_k$ is symmetric, then $\mathbb{E}[\eta_{ik}] = \frac{1}{2} \text{tr}(\mathbf{A}_k)$ for all k (Seber and Lee, 2012, Theorem 1.5). When $i \neq j$, we have

$$\mathbb{E} \left[\sum_{i \neq j} \eta_{ik} \eta_{jl} \right] = \frac{n(n-1)}{4} \text{tr}(\mathbf{A}_k) \text{tr}(\mathbf{A}_l). \quad (\text{S.13})$$

Then it suffices to compute the value of $\mathbb{E} [\sum_{i=1}^n \eta_{ik} \eta_{il}]$. We further decompose $\eta_{ik} = \eta_{ik}^{(1)} + \eta_{ik}^{(2)}$, where $\eta_{ik}^{(1)}$ and $\eta_{ik}^{(2)}$ are the first and the second term of the right hand side of (S.11), respectively. Then, $\eta_{ik} \eta_{il} = \eta_{ik}^{(1)} \eta_{il}^{(1)} + \eta_{ik}^{(1)} \eta_{il}^{(2)} + \eta_{ik}^{(2)} \eta_{il}^{(1)} + \eta_{ik}^{(2)} \eta_{il}^{(2)}$. Among them,

$$\mathbb{E} [\eta_{ik}^{(1)} \eta_{il}^{(2)}] = \mathbb{E} [\eta_{ik}^{(2)} \eta_{il}^{(1)}] = 0 \quad (\text{S.14})$$

since they are the summation of the third order central moments of a multivariate normal random variable. Based on the properties of multivariate normal random variables, we can derive

$$\begin{aligned}
\mathbb{E} \left[\eta_{ik}^{(2)} \eta_{il}^{(2)} \right] &= \mathbb{E} \left[\left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_l} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) \right] \\
&= \mathbb{E} \left[\left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}})' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_l} \right) \right] \\
&= \mathbb{E} \left[\text{tr} \left((\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_l} \right) \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}})' \right) \right] \\
&= \text{tr} \left((\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_l} \right) \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \mathbb{E} [(\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}}) (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}})'] \right) \\
&= \text{tr} \left((\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_l} \right) \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}') \right) \\
&= \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_k} \right)' (\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1} \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \theta_l} \right).
\end{aligned} \tag{S.15}$$

Here, the second equality holds since $(\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-1}$ is symmetric. The third equality follows from the invariant property under cyclic permutations of the trace operator. The fourth equality holds since trace is a linear operator. To compute $\mathbb{E} \left[\eta_{ik}^{(1)} \eta_{il}^{(1)} \right]$, we note that $(\boldsymbol{\Sigma}_{\mathbf{y}} + \omega^2 \boldsymbol{\mu}_{\mathbf{y}} \boldsymbol{\mu}_{\mathbf{y}}')^{-\frac{1}{2}} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{y}})$ is multivariate normally distributed with zero mean vector and identity covariance matrix (Seber and Lee, 2012, Chapter 2). It then follows from Seber and Lee (2012, Theorem 1.6) that

$$\mathbb{E} \left[\eta_{ik}^{(1)} \eta_{il}^{(1)} \right] = \frac{\text{tr}(\mathbf{A}_k) \text{tr}(\mathbf{A}_l) + 2 \text{tr}(\mathbf{A}_k \mathbf{A}_l)}{4}. \tag{S.16}$$

Summing up (S.12)–(S.16), we can get the formula given in Equation (5) of the main text.

We can readily compute the analytical forms of the partial derivatives in (5) of the main text by the definition of the derivative of matrices. For example, we have

$$\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \sigma_1} = \text{diag} \left(\underbrace{\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_1} t_1, \dots, \frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_1} t_1}_{K \text{ repetitions}}, \dots, \underbrace{\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_1} t_m, \dots, \frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_1} t_m}_{K \text{ repetitions}} \right) \in \mathcal{R}^{M \times M},$$

where

$$\frac{\partial \Sigma}{\partial \sigma_1} = \begin{bmatrix} 2\sigma_1 & \rho\sigma_2 \\ \rho\sigma_2 & 0 \end{bmatrix}.$$

The rest of the partial derivatives in Equation (5) in the main text can be obtained similarly and are omitted here.

S.2.2 Bootstrap- t

The bootstrap- t can be conducted with the following steps.

- Based on the maximum likelihood (ML) estimate $\hat{\boldsymbol{\theta}}$ in Section 3.1 of the main text, generate B pseudo-datasets by Monte Carlo simulations, where B is a sufficiently large number.
- Obtain the ML estimates of the desired indexes $g^{(b)}(\hat{\boldsymbol{\theta}})$ for the b th dataset, $b \in [B]$.
- Compute $z^{*(b)} = [g^{(b)}(\hat{\boldsymbol{\theta}}) - g(\hat{\boldsymbol{\theta}})] / \sqrt{\text{Avar}(g^{(b)}(\hat{\boldsymbol{\theta}}))}$, where the asymptotic variance can be calculated by Equation (6) in the main text.
- Let z_q^* be the q th empirical quantile based on the $z^{*(b)}$ for $b \in [B]$. The two-sided equal-tailed $(1 - \alpha)$ bootstrap- t confidence interval for $g(\boldsymbol{\theta})$ is given by

$$\left[g(\hat{\boldsymbol{\theta}}) - z_{1-\alpha/2}^* \sqrt{\text{Avar}(g(\hat{\boldsymbol{\theta}}))}, g(\hat{\boldsymbol{\theta}}) - z_{\alpha/2}^* \sqrt{\text{Avar}(g(\hat{\boldsymbol{\theta}}))} \right].$$

S.3 Likelihood Ratio Test and Goodness-of-Fit

The likelihood ratio (LR) test is established as follows. Consider two hypothesis tests $H_0 : \kappa = 0$ versus $H_1 : \kappa > 0$, and $H_0 : \omega = 0$ versus $H_1 : \omega > 0$. We can fit the data using Equation (2) in the main text without and with each of the restriction above to compute the LR statistic. For each test, parameters of interest are on the boundary of the parameter space under H_0 . [Self and Liang \(1987\)](#) suggested that the LR statistic then converges in

Table S.1: Biases ($\times 10$, averaged from the 1,000 replications) and RMSEs of the ML estimates for μ_l and σ_l , $l = 1, 2, 3$, by using the proposed two-layer block-effects model to fit the data in Section 5.2 of the main text.

n	m	K	μ_1		μ_2		μ_3		σ_1		σ_2		σ_3	
			bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE
5	5	1	0.21	0.73	0.33	0.96	0.42	1.15	1.14	3.03	1.68	3.08	1.69	3.01
		5	0.12	0.45	0.11	0.72	0.085	0.90	0.044	0.073	0.029	0.083	0.021	0.096
	10	1	0.095	0.46	0.16	0.73	0.21	0.91	-0.62	0.19	-0.38	0.17	-0.35	0.18
		5	-0.036	0.44	-0.047	0.71	-0.064	0.89	0.018	0.050	0.0085	0.060	0.0024	0.067
	5	1	-0.036	0.35	-0.049	0.55	-0.068	0.68	-0.28	1.11	0.27	1.16	0.29	1.10
		5	-0.028	0.32	-0.024	0.51	-0.062	0.63	0.0095	0.054	0.010	0.059	0.012	0.072
	10	1	-0.029	0.32	-0.045	0.50	-0.047	0.63	-0.071	0.12	-0.25	0.13	-0.20	0.13
		5	0.0026	0.31	0.0034	0.50	-0.0014	0.62	0.0014	0.035	-0.0012	0.042	-0.0023	0.048

distribution to a 50 : 50 mixture of χ_0^2 and χ_1^2 . The null hypothesis is rejected if the LR statistic is large.

The quantile-quantile (Q-Q) plot is made as follows. First, we vectorize the degradation data at the j th measurement from the i th rig as $\mathbf{y}_{ij} = [\mathbf{Y}'_{ij1}, \dots, \mathbf{Y}'_{ijK}]'$, $i \in [n]$ and $j \in [m]$. Conditioning on ζ_i , $\mathbf{y}_{ij}$ is multivariate normal with mean vector $\zeta_i \boldsymbol{\mu}_{\mathbf{y}_{ij}} \equiv [\boldsymbol{\mu}' \zeta_i t_j, \dots, \boldsymbol{\mu}' \zeta_i t_j] \in \mathcal{R}^{dK}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}_{ij}} \equiv \text{diag}(\boldsymbol{\Sigma} t_j, \dots, \boldsymbol{\Sigma} t_j) + \kappa^2 \mathbf{1}_{dK} \mathbf{1}_{dK}' \in \mathcal{R}^{dK \times dK}$. Since the ML estimator $\hat{\boldsymbol{\theta}}$ is consistent, the statistics $(\mathbf{y}_{ij} - \hat{\zeta}_i \hat{\boldsymbol{\mu}}_{\mathbf{y}_{ij}})' \hat{\boldsymbol{\Sigma}}_{\mathbf{y}_{ij}}^{-1} (\mathbf{y}_{ij} - \hat{\zeta}_i \hat{\boldsymbol{\mu}}_{\mathbf{y}_{ij}})$ approximately follow a chi-square distribution with dK degrees of freedom. Here, $\hat{\boldsymbol{\mu}}_{\mathbf{y}_{ij}}$ and $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}_{ij}}$ are the ML estimates of the corresponding model parameters. The estimate $\hat{\zeta}_i \equiv \mathbb{E}[\zeta_i | \hat{\boldsymbol{\theta}}, \mathbf{y}_i]$ can be calculated from Theorem 1.

S.4 Supplementary Results

S.4.1 Additional Simulation Results

Tables S.1 and S.2 provide the biases and the root mean squared errors (RMSEs) for the ML estimation in Section 5.2 of the main text. Meanwhile, we also provide the simulation result to see the performance of the EM algorithm when the dimension of degradation d increases slightly. Without loss of generality, we let $d = 10$, fix $N = 5$, $m = 10$, and $K = 5$, and set an identical degradation rate $\mu = 5$, degradation volatility $\sigma = 1$, and correlation coefficient $\rho = 0.5$ for all dimensions. By keeping the values of ω and κ as the previous

Table S.2: Biases ($\times 10$, averaged from the 1,000 replications) and RMSEs of the ML estimates for ρ_{kl} , $k, l = 1, 2, 3$, $k \neq l$, κ , and ω , by using the proposed two-layer block-effects model to fit the data in Section 5.2 of the main text.

n	m	K	ρ_{12}		ρ_{13}		ρ_{23}		ω		κ	
			bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE
5	5	1	-1.51	0.47	-0.89	0.35	-0.59	0.25	-0.32	0.22	4.83	5.90
		5	0.062	0.074	0.041	0.063	0.022	0.047	-0.28	0.072	-1.01	0.30
	10	1	-0.57	0.23	-0.34	0.17	-0.14	0.11	-0.30	0.078	-0.61	0.54
		5	0.034	0.050	0.015	0.043	0.0096	0.034	-0.27	0.071	-0.39	0.25
10	5	1	-1.20	0.35	-0.69	0.25	-0.36	0.15	-0.17	0.089	1.63	2.04
		5	0.040	0.052	0.021	0.044	0.021	0.033	-0.14	0.048	-0.49	0.22
	10	1	-0.30	0.13	-0.17	0.096	-0.12	0.073	-0.14	0.048	-0.35	0.42
		5	0.013	0.036	0.0047	0.030	0.0086	0.024	-0.12	0.048	-0.30	0.17

Table S.3: Biases ($\times 10$) and RMSEs of the ML estimates for μ , σ , ρ , ω , and κ , based on 1,000 simulation replications when the dimension of data is $d = 10$. Here, we fix $(n, m, K) = (10, 10, 5)$ and set the measurement time as $t_j = j$, $j = 1, \dots, 10$.

μ		σ		ρ		ω		κ	
bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE
0.96	0.31	-0.0092	0.033	0.0019	0.034	-0.13	0.050	-0.065	0.10

case, the computational time for convergence of the EM algorithm is around eight seconds. Since the parameters of each dimension of degradation are set as the same, we aggregate the results for the parameters μ , σ , and ρ without loss of generality. The biases and the RMSEs based on 1,000 replications are summarized in Table S.3, where it illustrates the satisfactory performance of the proposed model even when d increases slightly.

To illustrate the serious results when the block effects are incorrectly overlooked, Ta-

Table S.4: Biases ($\times 10$, averaged from the 1,000 replications) and RMSEs of the point estimates for μ_l and σ_l , $l = 1, 2, 3$, by using the pure random-effects model to fit the data in Section 5.2 of the main text.

n	m	K	μ_1		μ_2		μ_3		σ_1		σ_2		σ_3	
			bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE
5	5	1	-0.28	1.07	-0.60	1.59	-0.83	1.90	2.4	3.26	3.64	3.37	4.28	3.35
		5	0.22	0.45	0.35	0.72	0.44	0.90	-0.52	0.27	-0.48	0.40	-0.57	0.51
	10	1	0.063	0.48	0.12	0.76	0.16	0.95	-0.39	0.37	-0.020	0.52	0.075	0.66
		5	-0.031	0.45	-0.041	0.71	-0.056	0.89	-0.36	0.22	-0.49	0.34	-0.58	0.42
10	5	1	-0.20	0.50	-0.31	0.79	-0.40	0.95	1.54	2.62	2.38	2.76	2.76	2.79
		5	-0.021	0.32	-0.041	0.51	-0.051	0.64	-0.35	0.23	-0.42	0.35	-0.51	0.44
	10	1	-0.036	0.32	-0.057	0.50	-0.061	0.63	-0.38	0.29	-0.36	0.43	-0.40	0.54
		5	0.0066	0.31	0.0080	0.50	0.0038	0.62	-0.26	0.18	-0.37	0.28	-0.47	0.35

Table S.5: Biases ($\times 10$, averaged from the 1,000 replications) and RMSEs of the point estimates for ρ_{kl} , $k, l = 1, 2, 3$, $k \neq l$, κ , and ω , by using the pure random-effects model to fit the data in Section 5.2 of the main text.

n	m	K	ρ_{12}		ρ_{13}		ρ_{23}		ω		κ	
			bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE
5	5	1	-1.03	0.51	-0.73	0.45	-0.92	0.42	-0.46	0.11	5.22	5.95
		5	-1.62	0.41	-1.18	0.34	-1.41	0.35	-0.35	0.073	-0.95	0.48
	10	1	-1.41	0.47	-1.03	0.41	-1.06	0.38	-0.33	0.072	-0.64	0.60
		5	1.19	0.33	-0.96	0.28	-1.07	0.28	-0.30	0.071	-0.63	0.42
10	5	1	-1.30	0.46	-0.92	0.41	-0.89	0.35	-0.18	0.082	3.17	4.49
		5	-1.27	0.35	-0.98	0.28	-1.11	0.29	-0.16	0.049	-0.52	0.40
	10	1	-1.39	0.40	-1.10	0.35	-1.21	0.34	-0.19	0.049	-0.45	0.48
		5	-0.84	0.25	-0.69	0.22	-0.73	0.21	-0.14	0.047	-0.48	0.34

Table S.6: Coverage probability (in %) of the 95% interval estimators for the model parameters based on 1,000 simulation replications. In the simulations, $m = 10$ and $K = 5$.

n		μ_1	μ_2	μ_3	σ_1	σ_2	σ_3	ρ_{12}	ρ_{13}	ρ_{23}	ω	κ
5	Proposed model	94.2	94.1	94.3	95.1	95.5	93.6	95.4	95.2	95.4	95.2	92.6
	Random-effects model	22.7	22.2	21.3	87.4	83.6	81.1	81.8	80.3	78.1	27.0	87.6
10	Proposed model	95.0	95.1	95.4	95.5	95.5	95.6	96.0	95.2	95.6	95.3	95.6
	Random-effects model	24.1	23.8	23.4	92.5	90.5	89.5	87.7	86.1	84.5	31.3	84.8

bles S.4 and S.5 show the biases and RMSEs of the point estimates for the model parameters θ based on the pure random-effects model in Section 5.3 of the main text. The results are obtained by using this model to fit the degradation data in Section 5.2 of the main text. Finally, Table S.6 shows the coverage probability of the 95% confidence intervals for the model parameters based on 1,000 simulation replications from two candidate models as demonstrated in Section 5.3 of the main text.

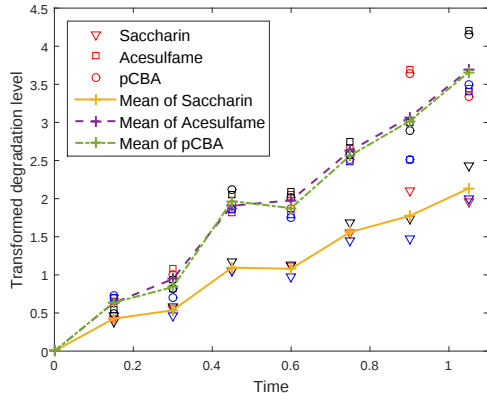
S.4.2 Additional Results for Case Study

The linearized data of the six test rigs are shown in Figure S.1. We can see that different rigs exhibit different (transformed) degradation rate, which implies the rig-layer block effect due to the variation of the hydroxyl radical concentration. This variation can be caused by the preparation of the test solution in each replication. The uncontrollable environmental conditions within the test stand/rig may also lead to the variation of the (transformed)

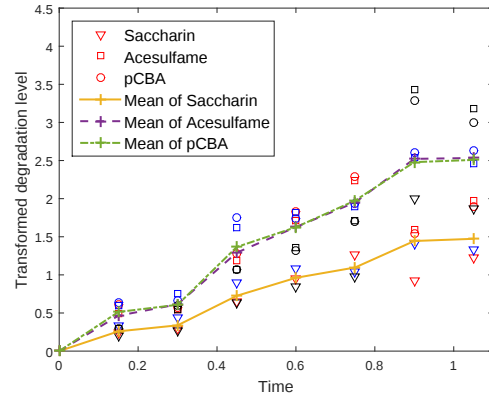
degradation rate in each replication. In addition, we can also see some simultaneous humps and dips in the mean degradation path. For example, there is an obvious dip for the fifth measurement of Rig 4. The dip may attribute to the impurity contained in the *tertiary*-butanol that was added to quench the degradation or the unknown problems for the liquid chromatography tandem mass spectrometry (LC-MS/MS) measurement system. The transformed data based on Equation (9) in the main text are also included as an Excel file in the supplementary material. For the model selection in the main text, we use the proposed model with time scale transformation $\mathbf{H}(t) = \text{diag}(h_1(t), \dots, h_d(t))$ to fit the data. We can also let the d dimensions of degradation share the same (transformed) time scale $h(t)$. The Akaike information criterion (AIC) value corresponding to $h(t) = t^q$ is -627.34 . AIC still prefers the proposed model without the time scale transformation. We also use the Bayesian information criterion (BIC) for model selection. The results are consistent with AIC, i.e., the proposed two-layer block-effects model without time scale transformation has the smallest BIC value.

References

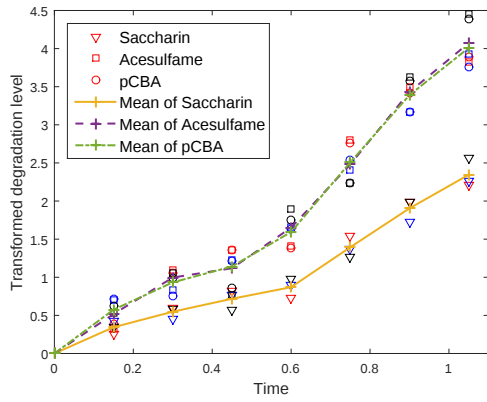
- Rencher, A. C. (2003), *Methods of Multivariate Analysis*, John Wiley & Sons.
- Seber, G. A. and Lee, A. J. (2012), *Linear Regression Analysis*, John Wiley & Sons.
- Self, S. G. and Liang, K.-Y. (1987), “Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions,” *Journal of the American Statistical Association*, 82(398), 605–610.
- Wang, X. (2010), “Wiener processes with random effects for degradation data,” *Journal of Multivariate Analysis*, 101(2), 340–351.



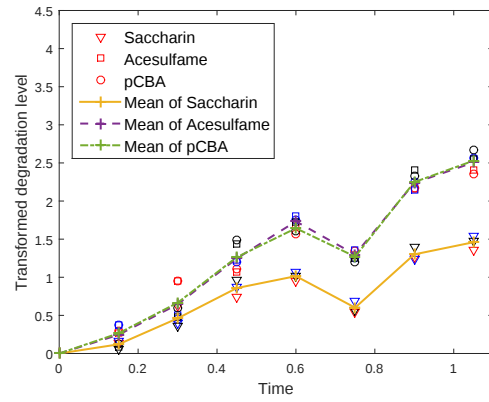
(a) Rig 1



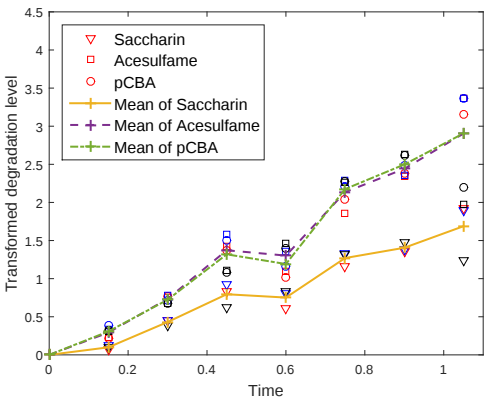
(b) Rig 2



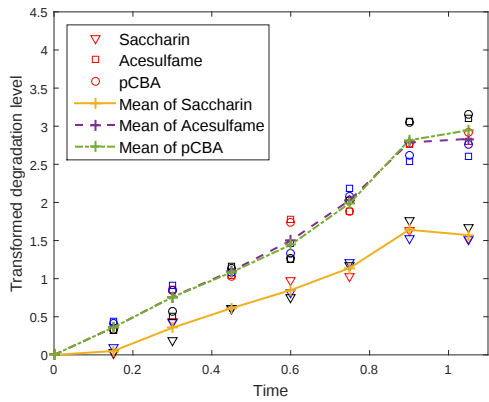
(c) Rig 3



(d) Rig 4



(e) Rig 5



(f) Rig 6

Figure S.1: Linearized degradation data from the six test rigs after the transformation based on Equation (9) in the main text.