# Supplementary material of "Remaining useful life prediction for two-phase degradation model based on reparameterized inverse Gaussian process"


Liangliang Zhuang[a], Ancha Xu[a,b,*], Yijun Wang[a], Yincai Tang[c]

[a]*School of Statistics and Mathematics, Zhejiang Gongshang University, Zhejiang 310018, China*
[b]*Collaborative Innovation Center of Statistical Data Engineering, Technology & Application Zhejiang Gongshang University, Zhejiang, China*
[c] *The KLATASDS-MOE, School of Statistics, East China Normal University, Shanghai 200241, China*


## S1 Proof of the additivity of rIG distribution

Firstly, we introduce the Bessel function and its related properties that will be utilized in this section.

**Definition 1.** *Let $K_\nu(\cdot)$ be the type I of Bessel function, with index $\nu$, which is defined as*

$$K_\nu(x) = \frac{1}{2}\int_0^\infty t^{\nu-1}\mathrm{e}^{-\frac{1}{2}x\left(t+t^{-1}\right)}\mathrm{d}t, \ \textit{where } x \in \mathbb{R}_{>0}. \tag{S1}$$

**Definition 2.** *Let $\mathrm{k}_\lambda(\cdot,\cdot)$ be the type II of Bessel function, with index $\lambda$, which is defined as*

$$\mathrm{k}_\lambda(\chi,\psi) = \int_0^\infty x^{\lambda-1}\exp\left[-\frac{1}{2}\left(\chi x^{-1}+\psi x\right)\right]\mathrm{d}x, \ \textit{where } \chi,\psi \in \mathbb{R}_{>0}. \tag{S2}$$

There are some nice properties of type I and type II of Bessel function, for example,

$$\begin{aligned} K_\nu(x) &= K_{-\nu}(x), \\ K_{1/2}(x) &= K_{-1/2}(x) = \sqrt{\frac{\pi}{2x}}\mathrm{e}^{-x}, \end{aligned} \tag{S3}$$

and

$$\begin{aligned} \mathrm{k}_\phi(\chi,\psi) &= 2\left(\frac{\chi}{\psi}\right)^{\phi/2}K_\phi(\sqrt{\chi\psi}), \\ \mathrm{k}_\phi(\chi,\psi) &= \mathrm{k}_{-\phi}(\psi,\chi), \end{aligned} \tag{S4}$$


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

where $v, \phi \in \mathbb{R}$ and $x, \chi, \psi \in \mathbb{R}_{>0}$. For more comprehensive information on the Bessel function, please refer to the relevant literature [1].

Assume that a random variable $Y$ follows the $rIG(\delta, \gamma)$. Using the properties of Bessel function, the MGF of $Y$ can be derived as follows:

$$\begin{aligned} M_Y(t) = E\left(\mathrm{e}^{ty}\right) &= \frac{\delta}{\sqrt{2\pi}} \mathrm{e}^{\delta\gamma} \int_0^{+\infty} y^{-\frac{3}{2}} \exp\left\{-\frac{\delta^2 y^{-1} + (\gamma^2 - 2t)\, y}{2}\right\} dy \\ &= \frac{\delta}{\sqrt{2\pi}} \mathrm{e}^{\delta\gamma} \mathrm{k}_{-\frac{1}{2}}\left(\delta^2, \gamma^2 - 2t\right) = \frac{\sqrt{2}\delta}{\sqrt{\pi}} \mathrm{e}^{\delta\gamma} \left(\frac{\delta^2}{\gamma^2 - 2t}\right)^{-\frac{1}{4}} K_{-\frac{1}{2}}\left(\sqrt{\delta^2\left(\gamma^2 - 2t\right)}\right) \\ &= \frac{\sqrt{2}\delta}{\sqrt{\pi}} \mathrm{e}^{\delta\gamma} \left(\frac{\delta^2}{\gamma^2 - 2t}\right)^{-\frac{1}{4}} \sqrt{\frac{\pi}{2\sqrt{\delta^2\left(\gamma^2 - 2t\right)}}} \mathrm{e}^{-\sqrt{\delta^2(\gamma^2 - 2t)}} \\ &= \mathrm{e}^{\delta\gamma\left(1 - \sqrt{1 - \frac{2t}{\gamma^2}}\right)}. \end{aligned} \tag{S5}$$

Assume that $Y_1 \sim rIG(\delta_1, \gamma)$, $Y_2 \sim rIG(\delta_2, \gamma)$, and $Y_1$ and $Y_2$ are independent of each other. Denote $Z = Y_1 + Y_2$, then the MGF of $Z$ is

$$\begin{aligned} M_z(t) = E\left[e^{t(y_1 + y_2)}\right] &= E\left(e^{ty_1}\right) E\left(e^{ty_2}\right) \\ &= e^{(\delta_1 + \delta_2)\left[\gamma\left(1 - \sqrt{1 - \frac{2t}{\gamma^2}}\right)\right]}. \end{aligned} \tag{S6}$$

Thus, $Z \sim IG\left(\delta_1 + \delta_2, \gamma\right)$, i.e., it has the convolution property: $rIG\left(\delta_1, \gamma\right) * rIG\left(\delta_2, \gamma\right) = rIG\left(\delta_1 + \delta_2, \gamma\right)$.

## S2 Proof of theorems 1 and 2

We first prove the results of Theorem 1. Let $Y_1(t)$ and $Y_2(t)$ be the degradation processes before and after the change point $\tau$. Then we have

$$Y(t) = \begin{cases} Y_1(t), & t \leq \tau, \\ Y_1(\tau) + Y_2(t - \tau), & t > \tau. \end{cases}$$

When $0 \leq t \leq \tau$, conditioned on $\tau$, the reliability function of $T$, denoted as $\bar{F}_1(t|\tau)$, can be written as

$$\bar{F}_1(t \mid \tau) = P(T > t \mid \tau \geq t) = P(Y_1(t) < \mathcal{D} \mid \tau \geq t) = F_{r\mathcal{IG}}(\mathcal{D}|\delta_1 t, \gamma). \tag{S7}$$

When $t > \tau$, the reliability function of $T$ given $\tau$ can be formulated as

$$\begin{aligned}\bar{F}_2(t \mid \tau) =& P\left(Y(t) < \mathcal{D} \mid \tau < t\right) = P\left(Y_1(\tau) + Y_2(t-\tau) < \mathcal{D} \mid \tau < t\right) \\ =& \int_0^{\mathcal{D}} P\left(Y_2(t-\tau) < \mathcal{D} - y_\tau \mid \tau < t\right) f_1(y_\tau \mid \tau) \mathrm{d}y_\tau \\ =& \int_0^{\mathcal{D}} F_{r\mathcal{IG}}(\mathcal{D} - y_\tau | \delta_2(t-\tau), \gamma) f_1(y_\tau \mid \tau) \mathrm{d}y_\tau,\end{aligned} \tag{S8}$$

where $y_\tau$ represents the degradation value at the change point time $\tau$, and $f_1(y_\tau \mid \tau)$ is the PDF of $y_\tau$. According to the property of rIG process, we know that $f_1(y_\tau \mid \tau) = f_{r\mathcal{IG}}(y_\tau \mid \delta_1\tau, \gamma)$. From (S7) and (S8), the unconditional reliability function of $T$ is

$$\begin{aligned}R\left(t\right) &= P\left(Y(t) < \mathcal{D}, \tau \geq t\right) + P\left(Y(t) < \mathcal{D}, 0 < \tau < t\right) \\ &= \bar{F}_1\left(t \mid \tau\right) \bar{G}_\tau(t) + \int_0^t g_\tau(\tau|\mu_\tau, \sigma_\tau^2) \bar{F}_2\left(t \mid \tau\right) d\tau,\end{aligned} \tag{S9}$$

where $\bar{G}_\tau(t)$ is the survival function of random variable $\tau$. Given the reliability function, the MTTF can be computed as

$$\text{MTTF} = E(T) = \int_0^\infty R(t) dt. \tag{S10}$$

Next, we prove the results of Theorem 2. Let $y_t$ be the observed degradation value at time $t$. The RUL of the system $S_t$ at time $t$ is defined as $S_t = \inf\{x; Y\left(t+x\right) \geq \mathcal{D} \mid y_t < \mathcal{D}\}$, representing the minimum time the system continues normal operation under the condition $y_t < \mathcal{D}$. To calculate the probability that the system continues normal operation at time $t + x$, our objective is to calculate the reliability of the system's RUL. Considering different relationships among times $t$, $t+x$, and $\tau$, we initially derive three distinct reliabilities based on the condition of $\tau$.

(i) When $x + t \leq \tau$, the conditional reliability function of $S_t$ is

$$\begin{aligned}\bar{F}_{S_t,1}(x \mid \tau) &= P(Y(t+x) < \mathcal{D} \mid y_t < \mathcal{D}, x+t \leq \tau) \\ &= P(Y(t+x) - y_t < \mathcal{D} - y_t \mid y_t < \mathcal{D}, x+t \leq \tau) \\ &= F_{r\mathcal{IG}}(\mathcal{D} - y_t|\delta_1 x, \gamma).\end{aligned} \tag{S11}$$

(ii) When $t < \tau < x + t$, the conditional reliability function of $S_t$ is

$$\begin{aligned}\bar{F}_{S_t,2}(x \mid \tau) &= P(Y(t+x) < \mathcal{D} \mid y_t < \mathcal{D}, t < \tau < x+t) \\ &= P(Y_2(t+x-\tau) + Y_1(\tau) < \mathcal{D} \mid y_t < \mathcal{D}, t < \tau < x+t) \\ &= \int_0^{\mathcal{D}} F_{r\mathcal{IG}}(\mathcal{D} - y_\tau|\delta_2(t+x-\tau), \gamma) f_1(y_\tau \mid \tau) \mathrm{d}y_\tau.\end{aligned} \tag{S12}$$

(iii) When $\tau \le t$, the conditional reliability function of $S_t$ is

$$\bar{F}_{S_t,3}(x \mid \tau) = F_{r\mathcal{IG}}(\mathcal{D} - y_t | \delta_2 x, \gamma). \tag{S13}$$

Based on (S11) - (S13), the unconditional reliability function of RUL is

$$\begin{aligned} R_{S_t}(x) =& P(Y(t+x) < \mathcal{D}, t < x+t \le \tau) \\ &+ P(Y(t+x) < \mathcal{D}, t \le \tau < x+t) + P(Y(t+x) < \mathcal{D}, t > \tau) \\ =& \bar{F}_{S_t,1}(x \mid \tau)\, \bar{G}_\tau(x+t) + \int_t^{x+t} g_\tau(\tau|\mu_\tau, \sigma_\tau^2) \bar{F}_{S_t,2}(x \mid \tau)\, d\tau \\ &+ \int_0^t g_\tau(\tau) \bar{F}_{S_t,3}(x \mid \tau)\, d\tau. \end{aligned} \tag{S14}$$

The unconditional PDF of RUL function can be computed by

$$f_{S_t}(x) = -\frac{\partial R_{S_t}(x)}{\partial x}. \tag{S15}$$

The mean of RUL at time $t$ can be obtained by

$$\text{MRL} = E(S_t) = \int_0^\infty R_{S_t}(x) dx. \tag{S16}$$

## S3 Technical details of the EM algorithm

To elucidate the technical details of the EM algorithm, we first establish a set of notations. Note that the log-likelihood function in Eq. (13) of the manuscript can be divided into two parts with respect to $\tau_i$, i.e. $l_i(\boldsymbol{\theta_\tau}) = \boldsymbol{v}_i^\top(\tau_i)\boldsymbol{w}_i(\boldsymbol{\theta}_\tau)$ and $l_{i,j}(\boldsymbol{\eta}, \tau_i) = \sum_{k=1}^3 \lambda_{i,j}^{(k)} \boldsymbol{v}_{i,j}^{(k)\top}(\tau_i)\boldsymbol{w}_{i,j}^{(k)}(\boldsymbol{\eta})$, where

$$\boldsymbol{v}_i(\tau_i) = \left(1, \tau_i, \tau_i^2\right)^\top, \quad v_{i,j}^{(1)}(\tau_i) = 1, \quad \boldsymbol{v}_{i,j}^{(2)}(\tau_i) = \left(1, \log(\Delta A_{i,j} + \Delta B_i \tau_i), \tau_i, \tau_i^2\right)^\top, \quad v_{i,j}^{(3)}(\tau_i) = 1,$$

$$\boldsymbol{w}_i(\boldsymbol{\theta}_\tau) = \left(-\log\sqrt{2\pi}\sigma_\tau - \frac{\mu_\tau^2}{2\sigma_\tau^2}, \frac{\mu_\tau}{\sigma_\tau^2}, -\frac{1}{2\sigma_\tau^2}\right)^\top,$$

$$w_{i,j}^{(1)}(\boldsymbol{\eta}) = -\log\sqrt{2\pi} + \log\delta_{1,i}\Delta t_{i,j} + \gamma\delta_{1,i}\Delta t_{i,j} - \frac{3}{2}\log\Delta y_{i,j} - \frac{(\delta_{1,i}\Delta t_{i,j})^2}{2\Delta y_{i,j}} - \frac{\gamma^2\Delta y_{i,j}}{2},$$

$$\boldsymbol{w}_{i,j}^{(2)}(\boldsymbol{\eta}) = \left(-\log\sqrt{2\pi} - \frac{3}{2}\log\Delta y_{i,j} - \frac{\gamma^2\Delta y_{i,j}}{2} + \gamma\Delta A_{i,j} - \frac{\Delta A_{i,j}^2}{2\Delta y_{i,j}}, 1, \Delta B_i\gamma - \frac{\Delta A_{i,j}\Delta B_i}{\Delta y_{i,j}}, -\frac{\Delta B_i^2}{2\Delta y_{i,j}}\right)^\top,$$

$$w_{i,j}^{(3)}(\boldsymbol{\eta}) = -\log\sqrt{2\pi} + \log\delta_{2,i}\Delta t_{i,j} + \gamma\delta_{2,i}\Delta t_{i,j} - \frac{3}{2}\log\Delta y_{i,j} - \frac{[\delta_{2,i}\Delta t_{i,j}]^2}{2\Delta y_{i,j}} - \frac{\gamma^2\Delta y_{i,j}}{2},$$

where $\Delta A_{i,j} = \delta_{2,i}t_{i,j} - \delta_{1,i}t_{i,j-1}$ and $\Delta B_i = \delta_{1,i} - \delta_{2,i}$.

### S3.1 Derivation of the conditional expectations in the E-step

We need to calculate the expectations required in the EM algorithm with respect to $p(\tau_i \mid \boldsymbol{\Delta y}_i)$ in the E-step. In the following, we will suppress the dependence on $\boldsymbol{\vartheta}$ for simplicity. According to the independence of the degradation increments, the joint PDF of $\boldsymbol{\Delta Y}_i$ and $\tau_i$ is

$$f_{\boldsymbol{\Delta Y}_i,\tau_i}\left(\boldsymbol{\Delta y}_i, \tau_i\right) = \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_\tau\right). \tag{S17}$$

After integrating out $\tau_i$ from (S17), the marginal PDF of $\boldsymbol{\Delta Y}_i$ is

$$f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right) = \int_{-\infty}^{+\infty} \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_\tau\right) d\tau_i. \tag{S18}$$

Referring to Eq. (11) of the manuscript, we encounter three distinct scenarios involving $\tau$, $t_{i,j}$, and $t_{i,j-1}$. These scenarios give rise to three potential cases for $\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)$, leading to three corresponding forms of $f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right)$. Specifically, when $\tau_i < t_{i,0}$, the conditional PDF of $\Delta y_{i,j}$ is:

$$\begin{aligned} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) = {} & \frac{\Delta m_{i,j}^{(3)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)}{\sqrt{2\pi}} \exp\left\{\gamma \Delta m_{i,j}^{(3)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)\right\} \Delta y_{i,j}^{-3/2} \\ & \times \exp\left\{-\frac{\left[\Delta m_{i,j}^{(3)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)\right]^2 \Delta y_{i,j}^{-1} + \gamma^2 \Delta y_{i,j}}{2}\right\}, \end{aligned} \tag{S19}$$

For the sake of simplification, we denote the conditional PDF of $\Delta y_{i,j}$ in this case as $f_{i,j|(3)}\left(\Delta y_{i,j} \mid \delta_{2,i}, \gamma, \tau_i\right)$ where $\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right) = \Delta m_{i,j}^{(3)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)$. Similarly, for $t_{i,j-1} \leqslant \tau_i < t_{i,j}$ and $\tau_i \geqslant t_{i,n_i}$, their conditional PDF of $\Delta y_{i,j}$ are denoted as $f_{i,j|(2)}\left(\Delta y_{i,j} \mid \delta_{2,i}, \gamma, \tau_i\right)$, and $f_{i,j|(1)}\left(\Delta y_{i,j} \mid \delta_{2,i}, \gamma, \tau_i\right)$ with $\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right) = \Delta m_{i,j}^{(2)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)$, and $\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right) = \Delta m_{i,j}^{(1)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)$, respectively. Now, let's focus on decomposing $\prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right)$ in (S18). The three situations corresponding to the above are:

- For $\tau_i < t_{i,0}$,

$$\begin{aligned} \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) &= \prod_{j=1}^{n_i} f_{i,j|(3)}\left(\Delta y_{i,j} \mid \delta_{2,i}, \gamma, \tau_i\right), \\ &\triangleq L_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}, \tau_i\right). \end{aligned} \tag{S20}$$

- For $t_{i,j-1} \leqslant \tau_i < t_{i,j},\ j = 1, \ldots, n_i$,

$$\begin{aligned} \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) = & \left\{ \prod_{j'=1}^{j-1} f_{i,j'|(1)}\left(\Delta y_{i,j'} \mid \delta_{1,i}, \gamma, \tau_i\right) \right\} \\ & \times f_{i,j|(2)}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right), \\ & \times \left\{ \prod_{j'=j+1}^{n_i} f_{i,j'|(3)}\left(\Delta y_{i,j'} \mid \delta_{2,i}, \gamma, \tau_i\right) \right\} \\ \triangleq & \, M_{ij}\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}, \tau_i\right). \end{aligned} \tag{S21}$$

- For $\tau_i \geqslant t_{i,n_i}$,

$$\begin{aligned} \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) = & \prod_{j=1}^{n_i} f_{i,j|(1)}\left(\Delta y_{i,j} \mid \delta_{2,i}, \gamma, \tau_i\right), \\ \triangleq & \, R_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}, \tau_i\right). \end{aligned} \tag{S22}$$

Thus, the marginal PDF of $\boldsymbol{\Delta Y}_i$ in (S18) can be rewritten as

$$\begin{aligned} f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right) = & \int_{-\infty}^{+\infty} \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_\tau\right) d\tau_i \\ = & \int_{-\infty}^{t_{i,0}} L_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_\tau\right) d\tau_i \\ & + \sum_{j=1}^{n_i} \int_{t_{i,j-1}}^{t_{i,j}} M_{i,j}\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_\tau\right) d\tau_i \\ & + \int_{t_{i,n_i}}^{+\infty} R_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_\tau\right) d\tau_i, \end{aligned} \tag{S23}$$

Based on Bayes' theorem, we can calculate the conditional PDF $p(\tau_i \mid \boldsymbol{\Delta y}_i)$ as

$$p(\tau_i \mid \boldsymbol{\Delta y}_i) = \frac{f_{\boldsymbol{\Delta Y}_i, \tau_i}\left(\boldsymbol{\Delta y}_i, \tau_i\right)}{f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right)}. \tag{S24}$$

Then, the conditional expectations, $E_{\boldsymbol{\vartheta}_{(s)}}\left[\boldsymbol{v}_i \mid \boldsymbol{\Delta y}_i\right]$ and $E_{\boldsymbol{\vartheta}_{(s)}}\left[\lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \boldsymbol{\Delta y}_i\right]$ with respect to the conditional distribution can be derived. The conditional expectation of $\boldsymbol{v}_i$, $i = 1, \ldots, I$, is:

$$\begin{aligned}E_{\boldsymbol{\vartheta}_{(s)}}\left\{\boldsymbol{v}_i \mid \boldsymbol{\Delta y}_i\right\} =& \frac{1}{f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right)}\left(\int_{-\infty}^{t_{i,0}} \boldsymbol{v}_i(\tau_i) L_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right. \\ &+ \sum_{j=1}^{n_i}\int_{t_{i,j-1}}^{t_{i,j}} \boldsymbol{v}_i(\tau_i) M_{i,j}\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i \\ &\left.+ \int_{t_{i,n_i}}^{\infty} \boldsymbol{v}_i(\tau_i) R_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right).\end{aligned} \tag{S25}$$

The conditional expectations of $\lambda_{i,j}^{(k)} v_{i,j}^{(k)}$, $i = 1, \ldots, I$, $j = 1, \ldots, n_i$, and $k = 1, 2, 3$ are as follows.

$$\begin{aligned}E_{\boldsymbol{\vartheta}_{(s)}}\left\{\lambda_{i,j}^{(1)} v_{i,j}^{(1)} \mid \boldsymbol{\Delta y}_i\right\} = \frac{1}{f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right)} &\left(\sum_{j'=j+1}^{n_i}\int_{t_{i,j'-1}}^{t_{i,j'}} v_{i,j'}^{(1)}(\tau_i) M_{i,j'}\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right. \\ &\left.+ \int_{t_{i,n_i}}^{\infty} v_{i,j}^{(1)}(\tau_i) R_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right),\end{aligned} \tag{S26}$$

for $k = 1$.

$$\boldsymbol{E}_{\boldsymbol{\vartheta}_{(s)}}\left\{\lambda_{i,j}^{(2)} \boldsymbol{v}_{i,j}^{(2)} \mid \boldsymbol{\Delta y}_i\right\} = \frac{1}{f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right)}\left(\int_{t_{i,j-1}}^{t_{i,j}} \boldsymbol{v}_{i,j}^{(2)}(\tau_i) M_{i,j}\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right), \tag{S27}$$

for $k = 2$.

$$\begin{aligned}E_{\boldsymbol{\vartheta}_{(s)}}\left\{\lambda_{i,j}^{(3)} v_{i,j}^{(3)} \mid \boldsymbol{\Delta y}_i\right\} = \frac{1}{f_{\boldsymbol{\Delta Y}_i}\left(\boldsymbol{\Delta y}_i\right)} &\left(\int_{-\infty}^{t_{i,0}} v_{i,j}^{(3)}(\tau_i) L_i\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right. \\ &\left.+ \sum_{j'=1}^{j-1}\int_{t_{i,j'-1}}^{t_{i,j'}} v_{i,j'}^{(3)}(\tau_i) M_{i,j'}\left(\boldsymbol{\Delta y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i\right) g_\tau\left(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}\right) d\tau_i\right),\end{aligned} \tag{S28}$$

for $k = 3$. Thus, the Q-function within the E-step is:

$$\boldsymbol{Q}_{(s)}(\boldsymbol{\vartheta}) = \sum_{i=1}^{I} E_{\boldsymbol{\vartheta}_{(s)}}\left\{\boldsymbol{v}_i \mid \boldsymbol{\Delta} y\right\}^\top \boldsymbol{w}_i\left(\boldsymbol{\theta}_\tau\right) + \sum_{i=1}^{I}\sum_{j=1}^{n_i}\sum_{k=1}^{3} E_{\boldsymbol{\vartheta}_{(s)}}\left\{\lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \boldsymbol{\Delta y}\right\}^\top \boldsymbol{w}_{i,j}^{(k)}(\boldsymbol{\eta}). \tag{S29}$$

### S3.2 First-order partial derivative in M-step

Taking the first order partial derivatives of the Q-function in (S29) with respect to $\boldsymbol{\vartheta}$ and setting the derivations to be equal to zero to obtain $\boldsymbol{\vartheta}_{(s+1)}$.

$$\begin{aligned}\frac{\partial \boldsymbol{Q}_{(s)}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\theta}_\tau} &= \sum_{i=1}^{I}\left[\frac{\partial \boldsymbol{w}_i(\boldsymbol{\theta}_\tau)}{\partial \boldsymbol{\theta}_\tau}\right]^\top E_{\boldsymbol{\vartheta}_{(s)}}\left\{\boldsymbol{v}_i \mid \boldsymbol{\Delta y}\right\} = \boldsymbol{0},\\ \frac{\partial \boldsymbol{Q}_{(s)}(\boldsymbol{\vartheta})}{\partial (\delta_{1,i},\delta_{2,i})^\top} &= \sum_{j=1}^{n_i}\sum_{k=1}^{3}\left[\frac{\partial \boldsymbol{w}_{i,j}^{(k)}(\boldsymbol{\eta})}{\partial (\delta_{1,i},\delta_{2,i})^\top}\right]^\top E_{\boldsymbol{\vartheta}_{(s)}}\left\{\lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \boldsymbol{\Delta y}\right\} = \boldsymbol{0},\ i=1,\ldots,I,\\ \frac{\partial \boldsymbol{Q}_{(s)}(\boldsymbol{\vartheta})}{\partial \gamma} &= \sum_{i=1}^{I}\sum_{j=1}^{n_i}\sum_{k=1}^{3}\left[\frac{\partial \boldsymbol{w}_{i,j}^{(k)}(\boldsymbol{\eta})}{\partial \gamma}\right]^\top E_{\boldsymbol{\vartheta}_{(s)}}\left\{\lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \boldsymbol{\Delta y}\right\} = 0,\end{aligned} \tag{S30}$$

where

$$\begin{aligned}\frac{\partial \boldsymbol{w}_i(\boldsymbol{\theta}_\tau)}{\partial \boldsymbol{\theta}_\tau} &= \begin{pmatrix} -\frac{\mu_\tau}{\sigma_\tau^2}, & -\frac{1}{2\sigma_\tau^2}+\frac{\mu_\tau^2}{2\sigma_\tau^4}\\ \frac{1}{\sigma_\tau^2}, & \frac{-\mu_\tau}{\sigma_\tau^4}\\ 0, & \frac{1}{2\sigma_\tau^4}\end{pmatrix},\\ \frac{\partial w_{i,j}^{(1)}(\boldsymbol{\eta})}{\partial (\delta_{1,i},\delta_{2,i})^\top} &= \begin{pmatrix} \left(\frac{1}{\delta_{1,i}}+\gamma\right)\Delta t_{i,j}-\frac{\delta_{1,i}\Delta t_{i,j}^2}{\Delta y_{i,j}}, & 0\end{pmatrix},\\ \frac{\partial \boldsymbol{w}_{i,j}^{(2)}(\boldsymbol{\eta})}{\partial (\delta_{1,i},\delta_{2,i})^\top} &= \begin{pmatrix} -\gamma t_{i,j-1}+\frac{\Delta A_{i,j} t_{i,j-1}}{\Delta y_{i,j}}, & \gamma t_{i,j}-\frac{\Delta A_{i,j} t_{i,j}}{\Delta y_{i,j}}\\ 0, & 0\\ \gamma+\frac{\Delta B_i t_{i,j-1}-\Delta A_{i,j}}{\Delta y_{i,j}}, & -\gamma+\frac{\Delta A_{i,j}-\Delta B_i t_{i,j}}{\Delta y_{i,j}}\\ \frac{\delta_{2,i}-\delta_{1,i}}{\Delta y_{i,j}}, & \frac{\delta_{1,i}-\delta_{2,i}}{\Delta y_{i,j}}\end{pmatrix},\\ \frac{\partial w_{i,j}^{(3)}(\boldsymbol{\eta})}{\partial (\delta_{1,i},\delta_{2,i})^\top} &= \begin{pmatrix} 0, & \left(\frac{1}{\delta_{2,i}}+\gamma\right)\Delta t_{i,j}-\frac{\delta_{2,i}\Delta t_{i,j}^2}{\Delta y_{i,j}}\end{pmatrix},\\ \frac{\partial w_{i,j}^{(1)}(\boldsymbol{\eta})}{\partial \gamma} &= \delta_{1,i}\Delta t_{i,j}-\gamma\Delta y_{i,j},\\ \frac{\partial \boldsymbol{w}_{i,j}^{(2)}(\boldsymbol{\eta})}{\partial \gamma} &= \begin{pmatrix} -\gamma\Delta y_{i,j}+\Delta A_{i,j}, & 0, & \Delta B_i, & 0\end{pmatrix}^\top,\\ \frac{\partial w_{i,j}^{(3)}(\boldsymbol{\eta})}{\partial \gamma} &= \delta_{2,i}\Delta t_{i,j}-\gamma\Delta y_{i,j}.\end{aligned}$$

### S3.3 The procedure of EM algorithm

The EM algorithm in our study can be implemented using the following steps:

- **Step 1**. Initialize the parameters $\boldsymbol{\vartheta}$ to some random values $\boldsymbol{\vartheta}_{(\mathbf{0})}$, and setting the tolerance error $\epsilon$.

- **Step 2**. Calculate $E_{\boldsymbol{\vartheta}_{(s)}}\left[l_i\left(\boldsymbol{\theta}_\tau\right) \mid \boldsymbol{\Delta y}\right]$ and $E_{\boldsymbol{\vartheta}_{(s)}}\left[l_{i,j}(\boldsymbol{\eta},\boldsymbol{\tau}) \mid \boldsymbol{\Delta y}\right]$ based on the solution of the $s$-th iteration $\boldsymbol{\vartheta}_{(s)}$.

- **Step 3**. Calculate the solution of the $(s+1)$-th iteration $\boldsymbol{\vartheta}_{(s+1)}$ by Eq. (15) of the manuscript.

- **Step 4**. Repeat Steps 2 and 3 until $\left|\boldsymbol{\vartheta}_{(s+1)} - \boldsymbol{\vartheta}_{(s)}\right| < \epsilon$, where $|\cdot|$ is the $L_1$ distance, and $\epsilon$ is the error tolerance.

- **Step 5**. The MLE of $\boldsymbol{\vartheta}$ can be obtained as $\hat{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta}_{(s+1)}$.

## S4 Technical details of the Bayesian analysis

The full conditional posterior distribution for each parameter can be computed as follows.

1. Given $\boldsymbol{\theta}_{\backslash(\mu_\tau,\sigma_\tau^2)}$ and $\boldsymbol{\Delta Y}$, the full conditional posterior distribution of $(\mu_\tau,\sigma_\tau^2)$ is

$$\left(\mu_\tau,\sigma_\tau^2\right) \mid \boldsymbol{\theta}_{\backslash(\mu_\tau,\sigma_\tau^2)}, \boldsymbol{\Delta Y} \sim NIGa\left(\beta_\tau',\eta_\tau',v_\tau',\xi_\tau'\right),$$

   where $\beta_\tau' = \beta_\tau + I, \eta_\tau' = \left(\beta_\tau\eta_\tau + \sum_{i=1}^{I}\tau_i\right)/\left(\beta_\tau + I\right), v_\tau' = I/2 + v_\tau$, and $\xi_\tau' = \xi_\tau + \beta_\tau\eta_\tau^2/2 + \sum_{i=1}^{I}\tau_i^2/2 - \left(\beta_\tau\eta_\tau + \sum_{i=1}^{I}\tau_i\right)^2/\left(2\left(\beta_\tau + I\right)\right).$

2. Given $\boldsymbol{\theta}_{\backslash\left(\mu_1,\sigma_1^2\right)}$ and $\boldsymbol{\Delta Y}$, the full conditional posterior distribution of $(\mu_1,\sigma_1^2)$ is

$$\left(\mu_1,\sigma_1^2\right) \mid \boldsymbol{\theta}_{\backslash\left(\mu_1,\sigma_1^2\right)}, \boldsymbol{\Delta Y} \sim NIGa\left(\beta_1',\eta_1',v_1',\xi_1'\right),$$

   where $\beta_1' = \beta_1 + I,\quad \eta_1' = \left(\beta_1\eta_1 + \sum_{i=1}^{I}\delta_i^1\right)/\left(\beta_1 + I\right),\ v_1' = I/2 + v_1$, and $\xi_1' = \xi_1 + \beta_1\eta_1^2/2 + \sum_{i=1}^{I}\delta_{1,i}^2/2 - \left(\beta_1\eta_1 + \sum_{i=1}^{I}\delta_i^1\right)^2/\left(2\left(\beta_1 + I\right)\right)$.

3. Given $\boldsymbol{\theta}_{\backslash\left(\mu_2,\sigma_2^2\right)}$ and $\boldsymbol{\Delta Y}$, the full conditional posterior distribution of $(\mu_2,\sigma_2^2)$ is

$$\left(\mu_2,\sigma_2^2\right) \mid \boldsymbol{\theta}_{\backslash\left(\mu_2,\sigma_2^2\right)}, \boldsymbol{\Delta Y} \sim NIGa\left(\beta_2',\eta_2',v_2',\xi_2'\right),$$

   where $\beta_2' = \beta_2 + I,\ \eta_2' = \left(\beta_2\eta_2 + \sum_{i=1}^{I}\delta_{2,i}\right)/\left(\beta_2 + I\right),\ v_2' = I/2 + v_2$, and $\xi_2' = \xi_2 + \beta_2\eta_2^2/2 + \sum_{i=1}^{I}\delta_{2,i}^2/2 - \left(\beta_2\eta_2 + \sum_{i=1}^{I}\delta_{2,i}\right)^2/\left(2\left(\beta_2 + I\right)\right)$.

4. Given $\boldsymbol{\theta}_{\backslash\gamma}$ and $\boldsymbol{\Delta Y}$, the full conditional posterior distribution of $\gamma$ is

$$\gamma \mid \boldsymbol{\theta}_{\backslash\gamma}, \boldsymbol{\Delta Y} \sim N\left(\omega', \kappa'\right).$$

where $\omega' = \left(\omega + \kappa N\right) / \left(1 + \kappa N\right), \kappa' = \kappa^2 / \left(1 + \kappa^2 N\right)$ and $N = \sum_{i=1}^{I}\sum_{j=1}^{n_i}\Delta y_{i,j}$.

5. For $\delta_{1,i}, i = 1, \ldots, I$, the full conditional posterior distribution has the following form

$$\pi\left(\delta_{1,i} \mid \boldsymbol{\theta}_{\backslash\delta_{1,i}}, \boldsymbol{\Delta Y}\right) \propto \exp\left\{\frac{2\mu_1\delta_{1,i} - \delta_{1,i}^2}{2\sigma_1^2} + \gamma\sum_{j=1}^{n_i}\Delta\mathcal{M}_{1,i,j} - \sum_{i=1}^{I}\sum_{j=1}^{n_i}\frac{\Delta\mathcal{M}_{1,i,j}^2}{2\Delta y_{i,j}}\right\}\prod_{j=1}^{n_i}\Delta\mathcal{M}_{1,i,j},$$

where $\Delta\mathcal{M}_{1,i,j} = \delta_{1,i}t_{i,j}\lambda_{i,j}^{(1)} + \left[\left(\delta_{1,i} - \delta_{2,i}\right)\tau_i + \delta_{2,i}t_{i,j+1} - \delta_{1,i}t_{i,j}\right]\lambda_{i,j}^{(2)}$.

6. For $\delta_{2,i}, i = 1, \ldots, I$, the full conditional posterior distribution is

$$\pi\left(\delta_{2,i} \mid \boldsymbol{\theta}_{\backslash\delta_{2,i}}, \boldsymbol{\Delta Y}\right) \propto \exp\left\{\frac{2\mu_2\delta_{2,i} - \delta_{2,i}^2}{2\sigma_2^2} + \gamma\sum_{j=1}^{n_i}\Delta\mathcal{M}_{2,i,j} - \sum_{i=1}^{I}\sum_{j=1}^{n_i}\frac{\Delta\mathcal{M}_{2,i,j}^2}{2\Delta y_{i,j}}\right\}\prod_{j=1}^{n_i}\Delta\mathcal{M}_{2,i,j},$$

where $\Delta\mathcal{M}_{2,i,j} = \left[\left(\delta_{1,i} - \delta_{2,i}\right)\tau_i + \delta_{2,i}t_{i,j+1} - \delta_{1,i}t_{i,j}\right]\lambda_{i,j}^{(2)} + \delta_{2,i}t_{i,j}\lambda_{i,j}^{(3)}$.

7. For $\tau_i, i = 1, \ldots, I$, the full conditional posterior distribution is given by

$$\pi\left(\tau_i \mid \boldsymbol{\theta}_{\backslash\tau_i}, \boldsymbol{\Delta Y}\right) \propto \exp\left\{\frac{2\mu_\tau\tau_i - \tau_i^2}{2\sigma_\tau^2} + \gamma\sum_{j=1}^{n_i}\Delta\mathcal{M}_{3,i,j} - \sum_{i=1}^{I}\sum_{j=1}^{n_i}\frac{\Delta\mathcal{M}_{3,i,j}^2}{2\Delta y_{i,j}}\right\}\prod_{j=1}^{n_i}\Delta\mathcal{M}_{3,i,j},$$

where $\Delta\mathcal{M}_{3,i,j} = \left[\left(\delta_{1,i} - \delta_{2,i}\right)\tau_i + \delta_{2,i}t_{i,j+1} - \delta_{1,i}t_{i,j}\right]\lambda_{i,j}^{(2)}$.

## S5 Additional results of the case studies

For the Bayesian method, we monitor the convergence of the ARMS-Gibbs algorithm through trace plots, and ergodic mean plots, which can be found in figures S1, and S2, respectively. These plots confirm the convergence of the Markov chains. Figure S3 illustrates the iterations of model parameter estimation based on the EM algorithm. Clearly, after 100 iterations, the estimated values converge to a relatively stable state.

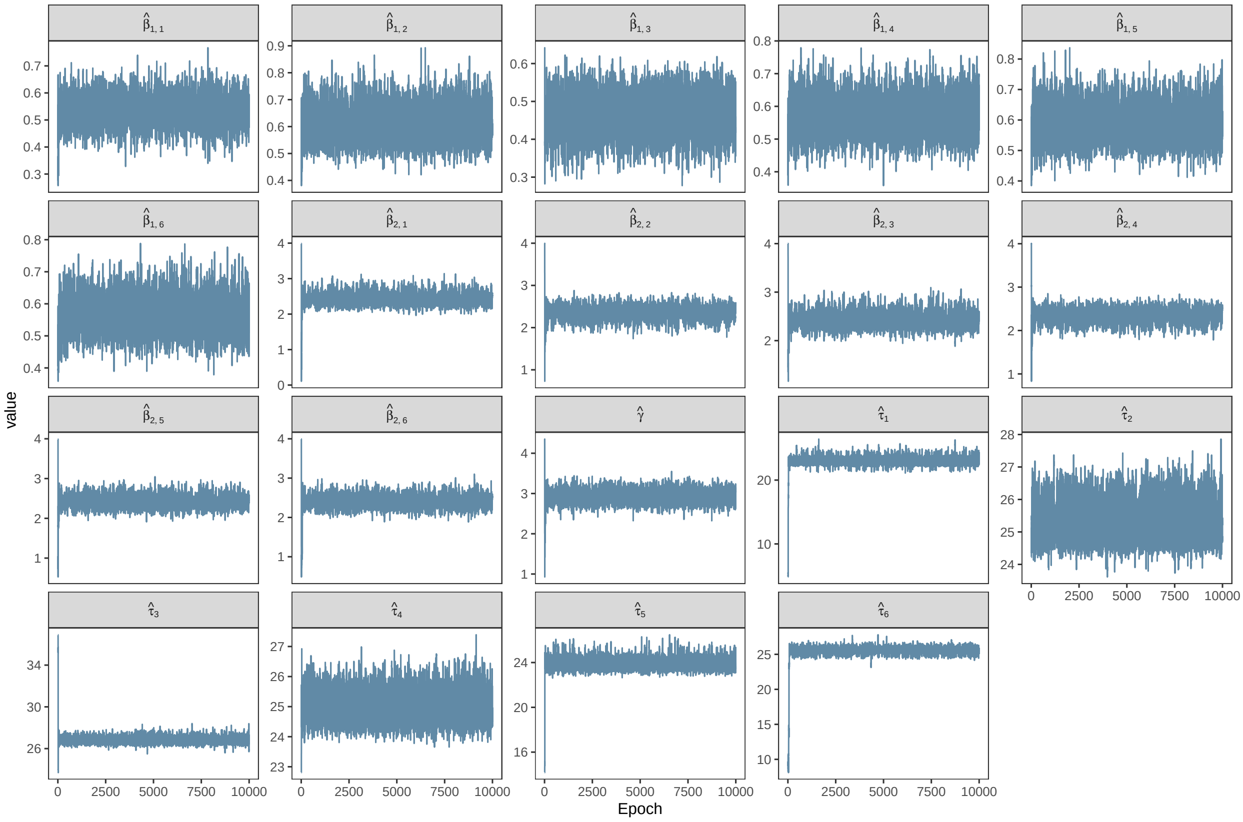

Figure S1: Trace plots of posterior samples of the model parameters.

## References

[1] Paolella, M.S., 2007. Intermediate probability: A computational approach. John Wiley & Sons.

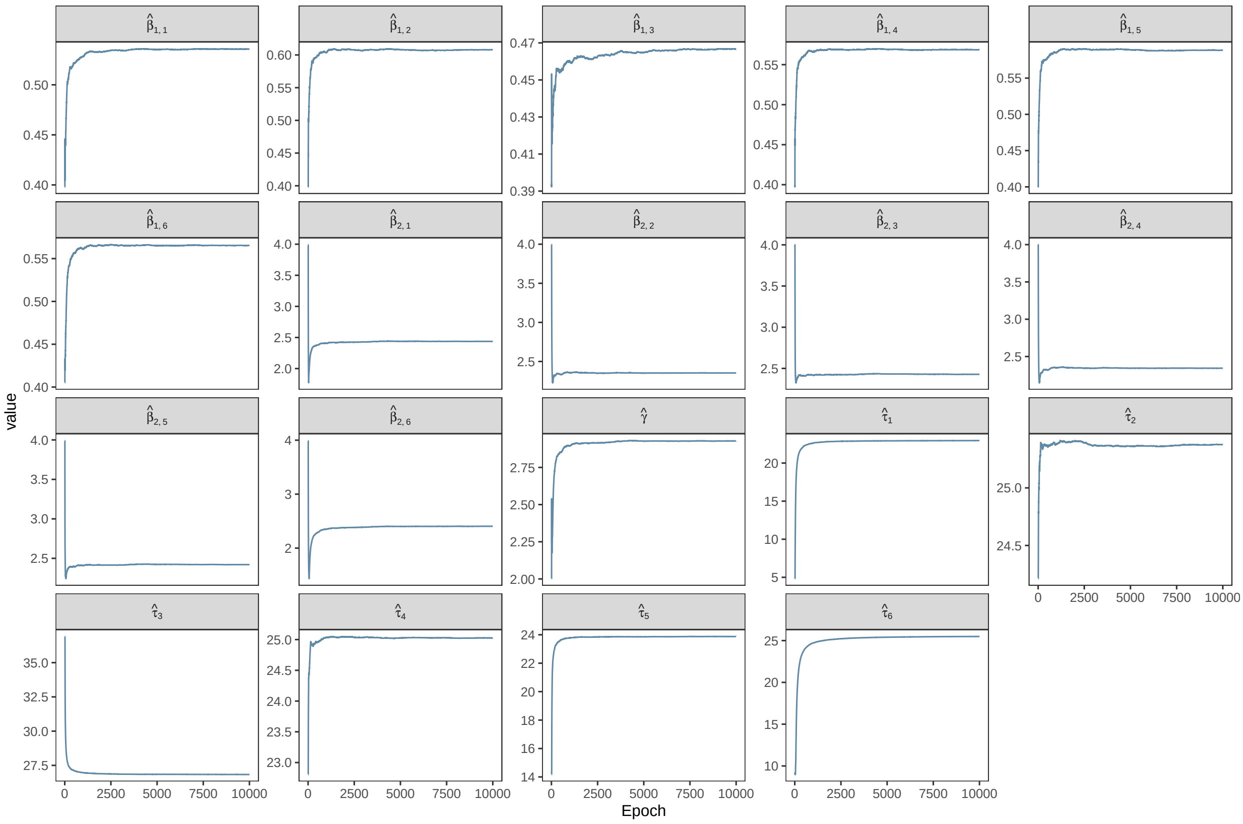


Figure S2: Ergodic mean plots of posterior samples of the model parameters.

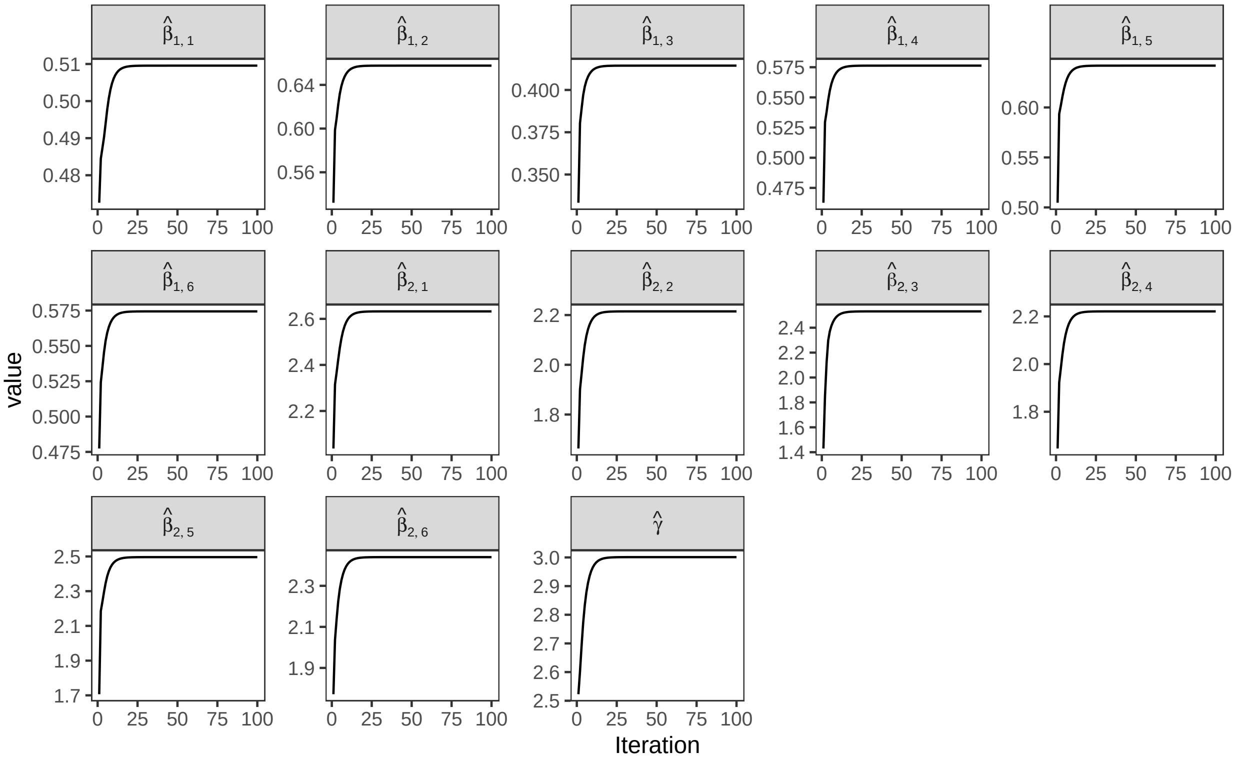


Figure S3: The estimates for the model parameters versus the iteration in the EM algorithm.