A Bayesian semi-parametric mixture model for bivariate extreme value analysis with application to precipitation forecasting

Yuan Tian and Brian J. Reich

North Carolina State University

Supplementary Material

The supplementary file contains proofs of the theorems, computational details and the additional simulation results.

S1 Proof of Theorem 1.

Proof. Let $\bar{H}(y|\mu, \sigma_k, \xi_k) = \int_y^{+\infty} h(t|\mu, \sigma_k, \xi_k)dt$. $F(y|\Theta, \mathbf{p})$ is the cumulative distribution function of univariate MIXGP model in (2.2). Following Definition 1, we have

$$
\begin{aligned}
\alpha_+(F) &= \liminf_{y\to+\infty} \frac{-\log\left(1 - F(y|\Theta, \mathbf{p})\right)}{\log y} \\
&= \liminf_{y\to+\infty} \frac{-\log\left(\sum_{k=1}^K p_k[E\{\bar{H}(y|\mu, \sigma_k, \xi_k)I(\mu \in U_k) + I(\mu > y)\}]\right)}{\log y} \\
&= \liminf_{y\to+\infty} \frac{-\log\left(\sum_{k=1}^K p_k[E\{\bar{H}(y|\mu, \sigma_k, \xi_k)I(\mu \in U_k) + \bar{G}_k(y)\}]\right)}{\log y}
\end{aligned}
$$

where $U_k = [y + \frac{\sigma_k}{\xi_k}, y]$ for $\xi_k < 0$, and $U_k = (-\infty, y]$ for $\xi_k \geq 0$. Notice that

$$a_k = \liminf_{y \to +\infty} \frac{-\log\left[E\{\bar{H}(y|\mu, \sigma_k, \xi_k) I(\mu \in U_k)\} + \bar{G}_k(y)\right]}{\log y}.$$

We first show

$$a_k = \begin{cases} \alpha_+(G_k), & \xi_k \leq 0. \\ \min\{\alpha_+(G_k), \xi_k^{-1}\}, & \xi_k > 0. \end{cases}$$

Case 1: $\xi_k < 0$. By definition, we have

$$a_k = \liminf_{y \to +\infty} \frac{-\log\left[E\{(1 + \xi_k \frac{y-\mu}{\sigma_k})^{-1/\xi_k} I(y + \frac{\sigma_k}{\xi_k} \leq \mu \leq y)\} + \bar{G}_k(y)\right]}{\log y}.$$

It is immediate to see that

$$a_k \leq \liminf_{y \to +\infty} \frac{-\log \bar{G}_k(y)}{\log y} = \alpha_+(G_k). \tag{S1.1}$$

In addition, we have

$$a_k \geq \liminf_{y \to +\infty} \frac{-\log \bar{G}_k(y + \frac{\sigma_k}{\xi_k})}{\log(y + \frac{\sigma_k}{\xi_k}) + \log y - \log(y + \frac{\sigma_k}{\xi_k})} = \alpha_+(G_k).$$

This together with (S1.1) yields $a_k = \alpha_+(G_k)$.

Case 2: $\xi_k = 0$. By definition,

$$a_k = \liminf_{y \to +\infty} \frac{-\log\left[\mathrm{E}\left\{\exp(-\frac{y-\mu}{\sigma_k}) I(\mu \leq y)\right\} + \bar{G}_k(y)\right]}{\log y}.$$

Similar to (S1.1), we can show $a_k \leq \alpha_+(G_k)$. Consider, separately, the scenarios where $\{\alpha_+(G_k) = +\infty\}$, $\{0 < \alpha_+(G_k) < +\infty\}$ and $\{\alpha_+(G_k) = 0\}$. Below, we show $a_k \geq \alpha_+(G_k)$ in each scenario. This gives $a_k = \alpha_+(G_k)$.

(i) $\alpha_+(G_k) = +\infty$. By definition, for any $M > 0$, there exists a monotonically increasing sequence $\{y_j\}$ such that $\lim_j y_j = +\infty$ and for all $j \geq 1$,

$$\frac{-\log \bar{G}_k(y_j)}{\log y_j} \geq M,$$

and hence

$$\bar{G}_k(y_j) \leq y_j^{-M}. \tag{S1.2}$$

Notice that $\lim_{y\to+\infty} \exp\{-y/(2\sigma_k)\}y^M = 0$. There exists some positive integer J_0 such that

$$\exp\left(-\frac{y_j}{2\sigma_k}\right) \leq y_j^{-M}, \quad \forall j \geq J_0. \tag{S1.3}$$

Let $y_j^* = 2y_j$. With some calculations, we have

$$\begin{aligned}
& \mathrm{E}\left\{\exp\left(-\frac{y_j^* - \mu}{\sigma_k}\right) I(\mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*) \qquad\qquad (\text{S1.4}) \\
\leq\ & \mathrm{E}\left\{\exp\left(-\frac{y_j^* - \mu}{\sigma_k}\right) I(\mu \leq y_j) + I(y_j < \mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*) \\
\leq\ & \mathrm{E}\exp\left(-\frac{y_j^* - y_j}{\sigma_k}\right) + \bar{G}_k(y_j) = \mathrm{E}\exp\left(-\frac{y_j}{\sigma_k}\right) + \bar{G}_k(y_j).
\end{aligned}$$

Combining this together with (S1.2) and (S1.3), we obtain for all $j \geq J_0$,

$$\mathrm{E}\left\{\exp\left(-\frac{y_j^* - \mu}{\sigma_k}\right) I(\mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*) \leq 2y_j^{-M},$$

and hence

$$\frac{-\log\left[\mathrm{E}\left\{\exp\left(-\frac{y_j^*-\mu}{\sigma_k}\right) I(\mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*)\right]}{\log y_j^*} \geq \frac{-\log 2 + M \log y_j}{\log y_j + \log 2}.$$

Therefore,

$$\lim_{j\to+\infty} \frac{-\log\left[\mathrm{E}\left\{\exp\left(-\frac{y_j^*-\mu}{\sigma_k}\right) I(\mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*)\right]}{\log y_j^*} \geq M.$$

This implies $a_k \geq M$. Since M can be arbitrarily chose, we obtain $a_k = \alpha_+(G_k) = +\infty$.

(ii) $0 < \alpha_+(G_k) < +\infty$. By definition, for any sufficiently small $\varepsilon > 0$, there exists a monotonically increasing sequence $\{y_j\}$ such that $\lim_j y_j = +\infty$ and for all $j \geq 1$,

$$\frac{-\log \bar{G}_k(y_j)}{\log y_j} \geq \alpha_+(G_k) - \varepsilon.$$

In addition, we can find some positive integer J_0 such that $\exp\{-y_j/(2\sigma_k)\} \leq y_j^{\varepsilon-\alpha_+(G_k)}$, for all $j \geq J_0$. Set $y_j^* = 2y_j$. Using similar arguments in (i), we can show

$$\lim_{j\to+\infty} \frac{-\log\left[\mathrm{E}\left\{\exp\left(-\frac{y_j^*-\mu}{\sigma_k}\right) I(\mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*)\right]}{\log y_j^*} \geq \alpha_+(G_k) - \varepsilon.$$

Since ε can be arbitrarily chosen, this implies $a_k \geq \alpha_+(G_k)$.

(iii) $\alpha_+(G_k) = 0$. Thus, we have $a_k \leq 0$. By definition, a_k is nonnegative. Therefore, $a_k = \alpha_+(G_k) = 0$.

Case 3: $\xi_k > 0$. By definition, we have

$$a_k = \liminf_{y\to+\infty} \frac{-\log\left[\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(\mu \leq y)\right\} + \bar{G}_k(y)\right]}{\log y}.$$

Consider, separately, the scenarios where $\{\alpha_+(G_k) < \xi_k^{-1}\}$ and $\{\alpha_+(G_k) \geq \xi_k^{-1}\}$. In the first scenario, we show $a_k = \alpha_+(G_k)$. In the second scenario, we show $a_k = \xi_k^{-1}$.

(i) $\alpha_+(G_k) < \xi_k^{-1}$. When $\alpha_+(G_k) = 0$, using similar arguments in Case 2(iii), we can show $a_k = 0$. It suffices to consider the case where $\alpha_+(G_k) > 0$. Similar to (S1.1), we have $a_k \leq \alpha_+(G_k)$. It remains to show $a_k \geq \alpha_+(G_k)$. Since $\alpha_+(G_k) < \xi_k^{-1}$, for any sufficiently large y, we have

$$\left(1 + \xi_k \frac{y}{\sigma_k}\right)^{-\xi_k^{-1}} \leq y^{-\alpha_+(G_k)}. \tag{S1.5}$$

Therefore, using similar arguments in Case 2(ii), we can show there exists a diverging sequence $\{y_j^*\}$ such that

$$\lim_{j\to+\infty} \frac{-\log\left[\mathrm{E}\left\{\left(1+\frac{y_j^*-\mu}{\sigma_k}\right)^{-\xi_k^{-1}} I(\mu \leq y_j^*)\right\} + \bar{G}_k(y_j^*)\right]}{\log y_j^*} \geq \alpha_+(G_k) - \varepsilon,$$

for any $\varepsilon > 0$. Since ε can be arbitrarily chosen, this implies $a_k \geq \alpha_+(G_k)$.

(ii) $\alpha_+(G_k) \geq \xi_k^{-1}$. Notice that

$$\begin{aligned}
&\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(\mu \leq y)\right\} + \bar{G}_k(y) \geq \mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(\mu \leq y)\right\} \\
\geq\ &\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(-y \leq \mu \leq y)\right\} \geq \mathrm{E}\left\{\left(1+\xi_k\frac{2y}{\sigma_k}\right)^{-1/\xi_k} I(-y \leq \mu \leq y)\right\} \\
=\ &\left(1+\xi_k\frac{2y}{\sigma_k}\right)^{-1/\xi_k} \{G_k(y) - G_k(-y)\}.
\end{aligned}$$

Therefore,

$$a_k \leq \liminf_{y\to+\infty} \frac{-\log\left\{\left(1+\xi_k\frac{2y}{\sigma_k}\right)^{-1/\xi_k}\{G_k(y)-G_k(-y)\}\right\}}{\log y}. \tag{S1.6}$$

Since $G_k(y) - G_k(-y) \to 1$, the RHS of (S1.6) is equal to

$$\lim_{y\to+\infty} \frac{\xi_k^{-1}\log\left(1+\frac{2\xi_k y}{\sigma_k}\right)}{\log y} = \xi_k^{-1}.$$

Thus, we obtain $a_k \leq \xi_k^{-1}$.

It remains to show $a_k \geq \xi_k^{-1}$. Since $\alpha_+(G_k) \geq \xi_k^{-1}$, we have

$$\liminf_{y\to+\infty} \frac{-\log \bar{G}_k(y)}{\log y} \geq \xi_k^{-1}.$$

By definition, for any sufficiently small $0 < \varepsilon < \xi_k^{-1}$, there exists some $y_0 > 0$ such that

$$\frac{-\log \bar{G}_k(y)}{\log y} \geq \xi_k^{-1} - \varepsilon, \qquad \forall y \geq y_0,$$

and hence

$$\bar{G}_k(y) \leq y^{\varepsilon-\xi_k^{-1}}, \qquad \forall y \geq y_0. \tag{S1.7}$$

Similar to (S1.4), we have

$$\begin{aligned} \mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(\mu\leq y)\right\} + \bar{G}_k(y) &\leq \left(1+\xi_k\frac{y}{2\sigma_k}\right)^{-1/\xi_k} + \bar{G}_k\left(\frac{y}{2}\right) \\ \leq \left(1+\xi_k\frac{y}{2\sigma_k}\right)^{-1/\xi_k+\varepsilon} + \bar{G}_k\left(\frac{y}{2}\right) &\leq \{\xi_k/(2\sigma_k)\}^{\varepsilon-\xi_k^{-1}} y^{\varepsilon-\xi_k^{-1}} + \bar{G}_k\left(\frac{y}{2}\right). \end{aligned}$$

This together with (S1.7) yields

$$\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(\mu\le y)\right\}+\bar{G}_k(y)\le\left(1+\frac{\xi_k}{2\sigma_k}\right)^{\varepsilon-\xi_k^{-1}} y^{\varepsilon-\xi_k^{-1}},\quad \forall y\ge y_0.$$

Therefore, we have

$$\begin{aligned} a_k &= \liminf_{y\to+\infty}\frac{-\log\left[\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I(\mu\le y)\right\}+\bar{G}_k(y)\right]}{\log y} \\ &\ge \liminf_{y\to+\infty}\frac{-\log\left\{\left(1+\frac{\xi_k}{2\sigma_k}\right)^{\varepsilon-\xi_k^{-1}} y^{\varepsilon-\xi_k^{-1}}\right\}}{\log y}=\xi_k^{-1}-\varepsilon. \end{aligned}$$

Since ε can be chosen arbitrarily small, we've shown $a_k\ge\xi_k^{-1}$.

In the following, we prove $\alpha_+(F)=\min(a_1,\dots,a_K)$. For any $k\in\{1,\dots,K\}$, it follows from the definition of $\alpha_+(F)$ that

$$\begin{aligned} \alpha_+(F) &\le \liminf_{y\to+\infty}\frac{-\log\left(p_k[E\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu\in U_k)+\bar{G}_k(y)\}]\right)}{\log y} \\ &= \lim_{y\to+\infty}\frac{-\log p_k}{\log y}+\liminf_{y\to+\infty}\frac{-\log\left([E\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu\in U_k)+\bar{G}_k(y)\}]\right)}{\log y} \\ &= a_k. \end{aligned}$$

Hence, we obtain $\alpha_+(F)\le\min_{k\in\{1,\dots,K\}}a_k$.

It remains to show $\alpha_+(F)\ge\min_{k\in\{1,\dots,K\}}a_k$. Without loss of generality, let's assume $\min_{k\in\{2,\dots,K\}}a_k\ge a_1$. It suffices to prove $\alpha_+(F)\ge a_1$. Consider, separately, the following three scenarios where $\{0<\xi_1^{-1}\le\alpha_+(G_1)\}$, $\{\alpha_+(G_1)\le\xi_1^{-1}\}$, $\{\xi_1\le 0\}$. In Scenario 1, we show $\alpha_+(F)\ge\xi_1^{-1}$. In Scenarios 2 and 3, we show $\alpha_+(F)\ge\alpha_+(G_1)$. The proof is hence completed.

(i) $0 < \xi_1^{-1} \le \alpha_+(G_1)$. In this scenario, we have $a_1 = \xi_1^{-1}$. Since $\min(a_2, \dots, a_K) \ge \xi_1^{-1}$ and $a_k \le \alpha_+(G_k)$, we have $\min_{k\in\{2,\dots,K\}} \alpha_+(G_k) \ge \xi_1^{-1}$ and hence $\min_{k\in\{1,\dots,G\}} \alpha_+(G_k) \ge \xi_1^{-1}$. Similar to (S1.7), we can show there exists some $y_0 > 0$ such that

$$\bar{G}_k(y) \le y^{\varepsilon-\xi_1^{-1}}, \qquad \forall y \ge y_0,\ \forall 1 \le k \le K. \tag{S1.8}$$

Let $\mathbb{I}_{-1} = \{k : 2 \le k \le K, \xi_k < 0\}$, $\mathbb{I}_0 = \{k : 2 \le k \le K, \xi_k = 0\}$, $\mathbb{I}_1 = \{k : 1 \le k \le K, \xi_k > 0\}$. For $k \in \mathbb{I}_1$, we have $a_k \le \xi_k^{-1}$. Since $a_k \ge \xi_1^{-1}$, we have $\xi_k \le \xi_1$. For these k, we have

$$\left(1 + \xi_k \frac{y}{2\sigma_k}\right)^{-\xi_k^{-1}} \le \left(1 + \xi_k \frac{y}{2\sigma_k}\right)^{-\xi_1^{-1}} \le \left(\frac{\xi_k}{2\sigma_k}\right)^{-\xi_1^{-1}} y^{-\xi_1^{-1}},$$

for any $y > 0$. Combining this with (S1.8), we obtain

$$\begin{aligned}
& p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k) I(\mu \in U_k) + \bar{G}_k(y)\}] \qquad \text{(S1.9)}\\
= \;& p_k \left\{ \mathrm{E}\left(1 + \frac{\xi_k(y-\mu)}{\sigma_k}\right)^{-\xi_k^{-1}} I(\mu \le y) + \bar{G}_k(y) \right\} \\
\le \;& \left\{ \mathrm{E}\left(1 + \frac{\xi_k(y-\mu)}{\sigma_k}\right)^{-\xi_k^{-1}} I(\mu \le y) + \bar{G}_k(y) \right\} \le \left\{ \left(1 + \frac{\xi_k y}{2\sigma_k}\right)^{-\xi_k^{-1}} + \bar{G}_k(y/2) \right\} \\
\le \;& \left(\frac{\xi_k}{2\sigma_k}\right)^{-\xi_1^{-1}} y^{-\xi_1^{-1}} + 2^{\xi_1^{-1}-\varepsilon} y^{\varepsilon-\xi_1^{-1}} \le \omega_k y^{\varepsilon-\xi_1^{-1}},
\end{aligned}$$

for sufficiently large y, where $\omega_k = \{\xi_k/(2\sigma_k)\}^{-\xi_1^{-1}} + 2^{\xi_1^{-1}}$, $\forall k \in \mathbb{I}_1$.

For any $k \in \mathbb{I}_{-1}$, it follows from (S1.8) that

$$\begin{aligned} & p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu \in U_k) + \bar{G}_k(y)\}] \qquad \text{(S1.10)} \\ = \;& p_k\left[\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I\left(y+\frac{\sigma_k}{\xi_k} \le \mu \le y\right)\right\} + \bar{G}_k(y)\right] \\ \le \;& p_k \mathrm{E} I\left(y+\frac{\sigma_k}{\xi_k} \le \mu \le y\right) + \bar{G}_k(y) \le \mathrm{E} I\left(y+\frac{\sigma_k}{\xi_k} \le \mu \le y\right) + \bar{G}_k(y) \\ = \;& \bar{G}(y+\sigma_k/\xi_k) \le \bar{G}(y/2) \le 2^{\xi_1^{-1}-\varepsilon} y^{\varepsilon-\xi_1^{-1}} \le 2^{\xi_1^{-1}} y^{\varepsilon-\xi_1^{-1}} = \omega_k y^{\varepsilon-\xi_1^{-1}}, \end{aligned}$$

for sufficiently large y, where $\omega_k = 2^{\xi_1^{-1}}$, $\forall k \in \mathbb{I}_{-1}$.

For any $k \in \mathbb{I}_0$, similar to (S1.3), we can show

$$\exp\left(-\frac{y}{2\sigma_k}\right) \le y^{\varepsilon-\xi_1^{-1}},$$

for sufficiently large y. For these k, it follows from (S1.8) that

$$\begin{aligned} & p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu \in U_k) + \bar{G}_k(y)\}] \qquad \text{(S1.11)} \\ = \;& p_k\left\{\mathrm{E}\exp\left(-\frac{\xi_k(y-\mu)}{\sigma_k}\right) I(\mu \le y) + \bar{G}_k(y)\right\} \\ \le \;& \left\{\mathrm{E}\exp\left(-\frac{\xi_k(y-\mu)}{\sigma_k}\right) I(\mu \le y) + \bar{G}_k(y)\right\} \le \left\{\mathrm{E}\exp\left(-\frac{\xi_k y}{2\sigma_k}\right) + \bar{G}_k(y/2)\right\} \\ \le \;& y^{\varepsilon-\xi_1^{-1}} + 2^{\xi_1^{-1}-\varepsilon} y^{\varepsilon-\xi_1^{-1}} \le \omega_k y^{\varepsilon-\xi_1^{-1}}, \end{aligned}$$

for sufficiently large y, where $\omega_k = 1 + 2^{\xi_1^{-1}}$, $\forall k \in \mathbb{I}_0$.

Combining (S1.9)-(S1.11) yields

$$\sum_{k=1}^{K} p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu \in U_k) + \bar{G}_k(y)\}] \le \left(\sum_k \omega_k\right) y^{\varepsilon-\xi_1^{-1}},$$

for sufficiently large y. Hence, it follows from the definition of $\alpha_+(F)$ that

$$\alpha_+(F) \geq \liminf_{y\to\infty} \frac{-\log(\sum_k \omega_k) + (\xi_1^{-1} - \varepsilon)\log y}{\log y} = \xi_1^{-1} - \varepsilon. \qquad \text{(S1.12)}$$

Since ε can be chosen arbitrarily small, we have $\alpha_+(F) \geq \xi_1^{-1}$.

(ii) $\alpha_+(G_1) \leq \xi_1^{-1}$. In this scenario, we have $a_1 = \alpha_+(G_1) < +\infty$. Since $\min(a_2, \dots, a_K) \geq \alpha_+(G_1)$ and $a_k \leq \alpha_+(G_k)$, we have $\min_{k\in\{2,\dots,K\}} \alpha_+(G_k) \geq \alpha_+(G_1)$ and hence $\min_{k\in\{1,\dots,G\}} \alpha_+(G_k) \geq \alpha_+(G_1)$. Consider separately the scenarios where $\{\alpha_+(G_1) = 0\}$ and $\{0 < \alpha_+(G_1) < +\infty\}$. When $\alpha_+(G_1) = 0$, we have $\alpha_+(F) \geq 0$. When $0 < \alpha_+(G_1) < +\infty$, similar to (S1.7), we can show there exists some $y_0 > 0$ such that

$$\bar{G}_k(y) \leq y^{\varepsilon - \alpha_+(G_1)}, \qquad \forall y \geq y_0,\ \forall 1 \leq k \leq K. \qquad \text{(S1.13)}$$

Let $\mathbb{I}_{-1} = \{k : 2 \leq k \leq K, \xi_k < 0\}$, $\mathbb{I}_0 = \{k : 2 \leq k \leq K, \xi_k = 0\}$, $\mathbb{I}_1 = \{k : 1 \leq k \leq K, \xi_k > 0\}$. For $k \in \mathbb{I}_1$, we have $a_k \leq \xi_k^{-1}$. Since $a_k \geq \alpha_+(G_1)$, we have $\alpha_+(G_1) \leq \xi_k^{-1}$. For these k, we have

$$\left(1 + \xi_k \frac{y}{2\sigma_k}\right)^{-\xi_k^{-1}} \leq y^{\varepsilon - \alpha_+(G_1)},$$

for any $\varepsilon > 0$ and any sufficiently large y. Combining this with (S1.13), we obtain

$$
\begin{aligned}
& p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu\in U_k)+\bar{G}_k(y)\}] \qquad\qquad \text{(S1.14)}\\
= \;& p_k\left\{\mathrm{E}\left(1+\frac{\xi_k(y-\mu)}{\sigma_k}\right)^{-\xi_k^{-1}} I(\mu\le y)+\bar{G}_k(y)\right\}\\
\le \;& \left\{\mathrm{E}\left(1+\frac{\xi_k(y-\mu)}{\sigma_k}\right)^{-\xi_k^{-1}} I(\mu\le y)+\bar{G}_k(y)\right\}\\
\le \;& \left\{\mathrm{E}\left(1+\frac{\xi_k(y-\mu)}{\sigma_k}\right)^{-\xi_k^{-1}} I(y/2\le\mu\le y)+\bar{G}_k(y/2)\right\}\\
\le \;& \left\{\left(1+\frac{\xi_k y}{2\sigma_k}\right)^{-\xi_k^{-1}} +\bar{G}_k(y/2)\right\}\\
\le \;& y^{\varepsilon-\alpha_+(G_1)}+2^{\alpha_+(G_1)-\varepsilon}y^{\varepsilon-\alpha_+(G_1)},
\end{aligned}
$$

for sufficiently large y, where $\omega_k = 1+2^{\alpha_+(G_1)-\varepsilon}$, $\forall k\in\mathbb{I}_1$.

For any $k\in\mathbb{I}_{-1}$, it follows from (S1.8) and (S1.13) that

$$
\begin{aligned}
& p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu\in U_k)+\bar{G}_k(y)\}] \qquad\qquad \text{(S1.15)}\\
= \;& p_k\left[\mathrm{E}\left\{\left(1+\xi_k\frac{y-\mu}{\sigma_k}\right)^{-1/\xi_k} I\left(y+\frac{\sigma_k}{\xi_k}\le\mu\le y\right)\right\}+\bar{G}_k(y)\right]\\
\le \;& p_k\mathrm{E}I\left(y+\frac{\sigma_k}{\xi_k}\le\mu\le y\right)+\bar{G}_k(y)\le \mathrm{E}I\left(y+\frac{\sigma_k}{\xi_k}\le\mu\le y\right)+\bar{G}_k(y)\\
= \;& \bar{G}(y+\sigma_k/\xi_k)\le\bar{G}(y/2)\le 2^{\alpha_+(G_1)-\varepsilon}y^{\varepsilon-\alpha_+(G_1)},
\end{aligned}
$$

for sufficiently large y, where $\omega_k = 2^{\alpha_+(G_1)-\varepsilon}$, $\forall k\in\mathbb{I}_{-1}$.

For any $k\in\mathbb{I}_0$, similar to (S1.3), we can show

$$
\exp\left(-\frac{y}{2\sigma_k}\right)\le y^{\varepsilon-\alpha_+(G_1)},
$$

for sufficiently large y. For these k, it follows from (S1.13) that

$$
\begin{aligned}
& p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu\in U_k)+\bar{G}_k(y)\}] \qquad \text{(S1.16)}\\
= \;& p_k\left\{\mathrm{E}\exp\left(-\frac{\xi_k(y-\mu)}{\sigma_k}\right)I(\mu\le y)+\bar{G}_k(y)\right\}\\
\le \;& \left\{\mathrm{E}\exp\left(-\frac{\xi_k(y-\mu)}{\sigma_k}\right)I(\mu\le y)+\bar{G}_k(y)\right\}\le\left\{\mathrm{E}\exp\left(-\frac{\xi_k y}{2\sigma_k}\right)+\bar{G}_k(y/2)\right\}\\
\le \;& y^{\varepsilon-\alpha_+(G_1)}+2^{\alpha_+(G_1)-\varepsilon}y^{\varepsilon-\alpha_+(G_1)}\le\omega_k y^{\varepsilon-\alpha_+(G_1)},
\end{aligned}
$$

for sufficiently large y, where $\omega_k = 1+2^{\alpha_+(G_1)-\varepsilon}$, $\forall k\in\mathbb{I}_0$.

Combining (S1.14)-(S1.16) yields

$$
\sum_{k=1}^{K}p_k[\mathrm{E}\{\bar{H}(y|\mu,\sigma_k,\xi_k)I(\mu\in U_k)+\bar{G}_k(y)\}]\le\left(\sum_k\omega_k\right)y^{\varepsilon-\alpha_+(G_1)},
$$

for sufficiently large y. Hence, it follows from the definition of $\alpha_+(F)$ that

$$
\alpha_+(F)\ge\liminf_{y\to\infty}\frac{-\log(\sum_k\omega_k)+\{\alpha_+(G_1)-\varepsilon\}\log y}{\log y}=\alpha_+(G_1)-\varepsilon.
$$

Since ε can be chosen arbitrarily small, we have $\alpha_+(F)\ge\alpha_+(G_1)$.

(iii) $\xi_1^{-1}\le 0$. In this scenario, we have $0\le a_1=\alpha_+(G_1)\le+\infty$. Since $\min(a_2,\ldots,a_K)\ge\alpha_+(G_1)$ and $a_k\le\alpha_+(G_k)$, we have $\min_{k\in\{2,\ldots,K\}}\alpha_+(G_k)\ge\alpha_+(G_1)$ and hence $\min_{k\in\{1,\ldots,G\}}\alpha_+(G_k)\ge\alpha_+(G_1)$. Consider separately the scenarios where $\{0\le\alpha_+(G_1)<+\infty\}$ and $\{\alpha_+(G_1)=+\infty\}$. When $\alpha_+(G_1)=+\infty$, we have $a_k=\alpha_+(G_k)=+\infty$ and $\xi_k\le 0$, $\forall k=1,\ldots,K$. Using similar arguments in *Case 2*(i), for any $M>0$, we can show there exists a monotonically increasing sequence $\{y_j\}$ such that $\lim_j y_j=+\infty$

and for all $j \geq 1$,

$$\bar{G}_k(y_j) \leq y_j^{-M}, \qquad \forall y \geq y_0,\ \forall 1 \leq k \leq K. \tag{S1.17}$$

Let $\mathbb{I}_{-1} = \{k : 2 \leq k \leq K, \xi_k < 0\}$, $\mathbb{I}_0 = \{k : 2 \leq k \leq K, \xi_k = 0\}$. For any $k \in \mathbb{I}_{-1}$, it follows from (S1.17) that

$$\begin{aligned}
& p_k[\mathrm{E}\{\bar{H}(y_j|\mu,\sigma_k,\xi_k)I(\mu \in U_k) + \bar{G}_k(y_j)\}] \qquad \text{(S1.18)}\\
= \;& p_k\left[\mathrm{E}\left\{\left(1+\xi_k\frac{y_j-\mu}{\sigma_k}\right)^{-1/\xi_k} I\left(y_j + \frac{\sigma_k}{\xi_k} \leq \mu \leq y_j\right)\right\} + \bar{G}_k(y_j)\right]\\
\leq \;& p_k \mathrm{E} I\left(y_j + \frac{\sigma_k}{\xi_k} \leq \mu \leq y_j\right) + \bar{G}_k(y_j) \leq \mathrm{E} I\left(y_j + \frac{\sigma_k}{\xi_k} \leq \mu \leq y_j\right) + \bar{G}_k(y_j)\\
= \;& \bar{G}(y_j + \sigma_k/\xi_k) \leq \bar{G}_k(y_j/2) \leq 2^M y_j^{-M}, \quad \forall j \geq J_0,
\end{aligned}$$

for sufficient large y_j. We set $\omega_k = 2^M$, $\forall k \in \mathbb{I}_{-1}$.

For any $k \in \mathbb{I}_0$, similar to (S1.3), we can show that there exists some positive integer J_0 such that

$$\exp\left(-\frac{y_j}{2\sigma_k}\right) \leq y_j^{-M}, \quad \forall j \geq J_0.$$

For these k, it follows from (S1.13) that

$$\begin{aligned}
& p_k[\mathrm{E}\{\bar{H}(y_j|\mu,\sigma_k,\xi_k)I(\mu \in U_k) + \bar{G}_k(y_j)\}] \qquad \text{(S1.19)}\\
= \;& p_k\left\{\mathrm{E}\exp\left(-\frac{\xi_k(y_j-\mu)}{\sigma_k}\right) I(\mu \leq y_j) + \bar{G}_k(y)\right\}\\
\leq \;& \left\{\mathrm{E}\exp\left(-\frac{\xi_k(y_j-\mu)}{\sigma_k}\right) I(\mu \leq y_j) + \bar{G}_k(y_j)\right\} \leq \left\{\mathrm{E}\exp\left(-\frac{\xi_k y_j}{2\sigma_k}\right) + \bar{G}_k(y/2)\right\}\\
\leq \;& y_j^{-M} + 2^M y_j^{-M} \leq \omega_k y_j^{-M}, \quad \forall j \geq J_0,
\end{aligned}$$

where $\omega_k = 1 + 2^M$, $\forall k \in \mathbb{I}_0$.

Combining (S1.18)-(S1.19) yields

$$\sum_{k=1}^{K} p_k[\mathrm{E}\{\bar{H}(y_j|\mu, \sigma_k, \xi_k) I(\mu \in U_k) + \bar{G}_k(y_j)\}] \leq \left(\sum_k \omega_k\right) y_j^{-M}, \forall j \geq J_0.$$

Hence, it follows from the definition of $\alpha_+(F)$ that

$$\alpha_+(F) \geq \lim_{j\to\infty} \frac{-\log(\sum_k \omega_k) + M \log y_j}{\log y_j} = M.$$

Since M can be chosen arbitrarily large, we have $\alpha_+(F) = \alpha_+(G_1) = +\infty$. When $0 \leq \alpha_+(G_1) < +\infty$, following the same argument in (ii) we can show $\alpha_+(F) \geq \alpha_+(G_1)$. The proof is hence completed.

□

S2 Proof of Theorem 2.

Proof. Under the assumption that there exists a nonempty cluster $o \in \{1, \ldots, K\}, s.t. \gamma_o \in [0, 1), \xi_{o,1} = \max_{k=1,\ldots,K}\{\xi_{k,1}\}, \xi_{o,2} = \max_{k=1,\ldots,K}\{\xi_{k,2}\}$, and $\xi_{o,1} > 0, \xi_{o,2} > 0$, we have

$$\begin{aligned}
\chi(t) &= \Pr\{Y > F_Y^{-1}(t)|X > F_X^{-1}(t)\} \\
&= \frac{\sum_{k=1}^{K} p_k E\{\Pr(X > F_X^{-1}(t), Y > F_Y^{-1}(t)|\boldsymbol{\mu}, \boldsymbol{\sigma}_k, \boldsymbol{\xi}_k, \gamma_k)\}}{1-t} \\
&= \frac{\sum_{k=1}^{K} p_k E\{\bar{F}_{X|\mu_1,\sigma_{k,1},\xi_{k,1}}(F_X^{-1}(t)) + \bar{F}_{Y|\mu_2,\sigma_{k,2},\xi_{k,2}}(F_Y^{-1}(t)) - 1 + F_{X,Y|\boldsymbol{\mu}_k,\boldsymbol{\sigma}_k,\boldsymbol{\xi}_k,\gamma_k}(F_X^{-1}(t), F_Y^{-1}(t))\}}{1-t} \\
&\geq \frac{p_o E\{\bar{F}_{X|\mu_1,\sigma_{o,1},\xi_{o,1}}(F_X^{-1}(t)) + \bar{F}_{Y|\mu_2,\sigma_{o,2},\xi_{o,2}}(F_Y^{-1}(t)) - 1 + F_{X,Y|\boldsymbol{\mu}_o,\boldsymbol{\sigma}_o,\boldsymbol{\xi}_o,\gamma_o}(F_X^{-1}(t), F_Y^{-1}(t))\}}{1-t}.
\end{aligned}$$

where

$$F_X(x) = 1 - \sum_{k=1}^{K} p_k[E\{\bar{H}(x|\mu_1, \sigma_{k,1}, \xi_{k,1})I(\mu_1 \in U_{k,1}) + I(\mu_1 > x)\}],$$
$$F_Y(y) = 1 - \sum_{k=1}^{K} p_k[E\{\bar{H}(y|\mu_2, \sigma_{k,2}, \xi_{k,2})I(\mu_2 \in U_{k,2}) + I(\mu_2 > y)\}],$$

$U_{k,1} = [x + \frac{\sigma_{k,1}}{\xi_{k,1}}, x]$ for $\xi_{k,1} < 0$, and $U_{k,1} = (-\infty, x]$ for $\xi_{k,1} \geq 0$, $U_{k,2} = [y + \frac{\sigma_{k,2}}{\xi_{k,2}}, y]$ for $\xi_{k,2} < 0$, and $U_{k,2} = (-\infty, y]$ for $\xi_{k,2} \geq 0$.

Denote $t_x = \bar{F}_{X|\mu_1,\sigma_{o,1},\xi_{o,1}}(F_X^{-1}(t)), t_y = \bar{F}_{Y|\mu_2,\sigma_{o,2},\xi_{o,2}}(F_Y^{-1}(t))$, then

$$\chi(t) \geq \frac{p_o E\left\{t_x + t_y - 1 + \exp[-\{(-\log(1-t_x))^{1/\gamma} + (-\log(1-t_y))^{1/\gamma}\}^{\gamma}]\right\}}{1-t}$$

Denote $z_1 = -\log(1 - t_x), z_2 = -\log(1 - t_y)$, we have $t_x = 1 - \exp(-z_1), t_y = 1 - \exp(-z_2)$, and

$$\chi(t) \geq \frac{p_o E[1 - \exp(-z_1) - \exp(-z_2) + \exp\{-(z_1^{1/\gamma} + z_2^{1/\gamma})^{\gamma}\}]}{1-t}.$$

Let $l(z_1, z_2) = 1 - \exp(-z_1) - \exp(-z_2) + \exp\{-(z_1^{1/\gamma} + z_2^{1/\gamma})^{\gamma}\}$. Notice that

$$\frac{\partial l(z_1, z_2)}{\partial z_1} = \exp(-z_1) - \exp\{-(z_1^{1/\gamma} + z_2^{1/\gamma})^{\gamma}\}(z_1^{1/\gamma} + z_2^{1/\gamma})^{\gamma-1} z_1^{1/\gamma-1} \geq 0,$$
$$\frac{\partial l(z_1, z_2)}{\partial z_2} = \exp(-z_2) - \exp\{-(z_1^{1/\gamma} + z_2^{1/\gamma})^{\gamma}\}(z_1^{1/\gamma} + z_2^{1/\gamma})^{\gamma-1} z_2^{1/\gamma-1} \geq 0.$$

For any fixed z_1 (or z_2), $l(z_1, z_2)$ is a monotonously increasing function of z_2 (or z_1).

Let $G_{k,1}$ and $G_{k,2}$ be the marginal distribution of $G_k(\cdot, \cdot), k = 1, \ldots, K$. Under the assumption that $G_k(\cdot, \cdot) \sim N(u_k, \Sigma)$, we have $\alpha_+(G_{k,1})) =$

$\alpha_+(G_{k,2}) = +\infty, k = 1, \ldots, K$. Since $\xi_{o,1} = \max_{k=1,\ldots,K}\{\xi_{k,1}\}$ and $\xi_{o,2} = \max_{k=1,\ldots,K}\{\xi_{k,2}\}$, we have $0 < \xi_{o,1}^{-1} \leq \min_{k=1,\ldots,K}\{\alpha_+(G_{k,1})\}$ and $0 < \xi_{o,2}^{-1} \leq \min_{k=1,\ldots,K}\{\alpha_+(G_{k,2})\}$.

Similarly, we can show there exists some $x_0 > 0, y_0 > 0$ such that

$$\begin{aligned} \bar{G}_{k,1}(x) &\leq x^{\varepsilon - \xi_{o,1}^{-1}}, & \forall x \geq x_0,\ \forall 1 \leq k \leq K, \\ \bar{G}_{k,2}(y) &\leq y^{\varepsilon - \xi_{o,2}^{-1}}, & \forall y \geq y_0,\ \forall 1 \leq k \leq K, \end{aligned} \tag{S2.1}$$

for any sufficient small ϵ.

Let $\mathbb{I}_{-1}^{(x)} = \{k : 1 \leq k \leq K, \xi_{k,1} < 0\}$, $\mathbb{I}_{0}^{(x)} = \{k : 1 \leq k \leq K, \xi_{k,1} = 0\}$, $\mathbb{I}_{1}^{(x)} = \{k : 1 \leq k \leq K, \xi_{k,1} > 0\}$, $\mathbb{I}_{-1}^{(y)} = \{k : 1 \leq k \leq K, \xi_{k,2} < 0\}$, $\mathbb{I}_{0}^{(y)} = \{k : 1 \leq k \leq K, \xi_{k,2} = 0\}$, $\mathbb{I}_{1}^{(y)} = \{k : 1 \leq k \leq K, \xi_{k,2} > 0\}$.

For $k \in \mathbb{I}_1^{(x)}$, we have

$$p_k[\mathrm{E}\{\bar{H}(x|\mu_1, \sigma_{k,1}, \xi_{k,1})I(\mu_1 \in U_{k,1}) + \bar{G}_{k,1}(x)\}] \leq \omega_{k,1} x^{\varepsilon - \xi_{o,1}^{-1}}, \tag{S2.2}$$

for sufficiently large x, where $\omega_{k,1} = \{\xi_{k,1}/(2\sigma_{k,1})\}^{-\xi_{o,1}^{-1}} + 2^{\xi_{o,1}^{-1}}$, $\forall k \in \mathbb{I}_1^{(x)}$.

For any $k \in \mathbb{I}_{-1}^{(x)}$, we have

$$p_k[\mathrm{E}\{\bar{H}(x|\mu_1, \sigma_{k,1}, \xi_{k,1})I(\mu_1 \in U_{k,1}) + \bar{G}_{k,1}(x)\}] \leq \omega_k x^{\varepsilon - \xi_1^{-1}}, \tag{S2.3}$$

for sufficiently large x, where $\omega_{k,1} = 2^{\xi_{o,1}^{-1}}$, $\forall k \in \mathbb{I}_{-1}^{(x)}$.

For any $k \in \mathbb{I}_0^{(x)}$, we have

$$p_k[\mathrm{E}\{\bar{H}(x|\mu_1, \sigma_{k,1}, \xi_{k,1})I(\mu_1 \in U_{k,1}) + \bar{G}_{k,1}(x)\}] \leq \omega_{k,1} x^{\varepsilon - \xi_{o,1}^{-1}}, \tag{S2.4}$$

for sufficiently large x, where $\omega_{k,1} = 1 + 2^{\xi_{o,1}^{-1}}$, $\forall k \in \mathbb{I}_0^{(x)}$.

Combining (S2.2)-(S2.4) yields

$$\begin{aligned}\bar{F}_X(x) &= \sum_{k=1}^{K} p_k[\mathrm{E}\{\bar{H}(x|\mu_1, \sigma_{k,1}, \xi_{k,1})I(\mu_1 \in U_{k,1}) + \bar{G}_{k,1}(x)\}] \\ &\leq \left(\sum_k \omega_{k,1}\right) x^{\varepsilon - \xi_{k,1}^{-1}},\end{aligned}$$

for sufficiently large x. Since ε can be chosen arbitrarily small, we have $\bar{F}_X(x) \leq (\sum_k \omega_{k,1})\, x^{-\xi_{o,1}^{-1}}$. Using similar arguments, we have $\bar{F}_Y(y) \leq (\sum_k \omega_{k,2})\, y^{-\xi_{o,2}^{-1}}$. Therefore,

$$F_X^{-1}(t) \leq \left(\frac{1-t}{\sum_k \omega_{k,1}}\right)^{-\xi_{o,1}} \quad \text{and} \quad F_Y^{-1}(t) \leq \left(\frac{1-t}{\sum_k \omega_{k,2}}\right)^{-\xi_{o,2}}. \tag{S2.5}$$

Notice that

$$\begin{aligned}t_x &= \bar{F}_{X|\mu_1,\sigma_{o,1},\xi_{o,1}}(F_X^{-1}(t)) \\ &= \bar{H}(F_X^{-1}(t)|\mu_1, \sigma_{o,1}, \xi_{o,1})I(\mu_1 \in U_o) + I(\mu_1 > F_X^{-1}(t)) \qquad \text{(S2.6)} \\ &\geq (1 + \xi_{o,1}\frac{F_X^{-1}(t) - \mu_1}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}} I(\mu_1 \leq F_X^{-1}(t))\end{aligned}$$

Similarly, we can show $t_y \geq (1 + \xi_{o,2}\frac{F_Y^{-1}(t)-\mu_2}{\sigma_{o,2}})^{-\xi_{o,2}^{-1}} I(\mu_2 \leq F_Y^{-1}(t))$.

Without loss of generality, we assume $t_x = \min(t_x, t_y)$. By the monotonicity of $l(\cdot,\cdot)$, we have

$$\chi(t) \geq \frac{p_o E\{2t_x - 1 + (1-t_x)^{2\gamma_o}\}}{1-t}.$$

According to (S2.6) we have

$$
\begin{aligned}
\chi(t) &\geq \frac{p_o E[\{2(1+\xi_{o,1}\frac{F_X^{-1}(t)-\mu_1}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}} - 1 + (1-(1+\xi_{o,1}\frac{F_X^{-1}(t)-\mu_1}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}})^{2^{\gamma_o}}\} I(\mu_1 \leq F_X^{-1}(t))]}{1-t} \\
&\geq \frac{p_o E[\{2(1+\xi_{o,1}\frac{F_X^{-1}(t)-\mu_1}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}} - 1 + (1-(1+\xi_{o,1}\frac{F_X^{-1}(t)-\mu_1}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}})^{2^{\gamma_o}}\} I(0 \leq \mu_1 \leq F_X^{-1}(t))]}{1-t} \\
&\geq \frac{p_o[2(1+\xi_{o,1}\frac{F_X^{-1}(t)}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}} - 1 + (1-(1+\xi_{o,1}\frac{F_X^{-1}(t)}{\sigma_{o,1}})^{-\xi_{o,1}^{-1}})^{2^{\gamma_o}}]}{1-t} \\
&\geq \frac{p_o[2(\frac{2\xi_{o,1}}{\sigma_{o,1}}F_X^{-1}(t))^{-\xi_{o,1}^{-1}} - 1 + (1-(\frac{2\xi_{o,1}}{\sigma_{o,1}}F_X^{-1}(t))^{-\xi_{o,1}^{-1}})^{2^{\gamma_o}}]}{1-t}
\end{aligned}
$$

for sufficiently large t. It follows from (S2.5) that

$$
\chi(t) \geq \frac{p_o[\{1-\frac{1-t}{\sum_k \omega_{k,1}}(\frac{2\xi_{o,1}}{\sigma_{o,1}})^{\xi_{o,1}^{-1}}\}^{2^{\gamma_o}} + 2\frac{1-t}{\sum_k \omega_{k,1}}(\frac{2\xi_{o,1}}{\sigma_{o,1}})^{\xi_{o,1}^{-1}} - 1]}{1-t}.
$$

According to the L'Hospital's rule,

$$
\begin{aligned}
\chi = \lim_{t\to 1}\chi(t) \geq \lim_{t\to 1} &\frac{p_o[\{1-\frac{1-t}{\sum_k \omega_{k,1}}(\frac{2\xi_{o,1}}{\sigma_{o,1}})^{\xi_{o,1}^{-1}}\}^{2^{\gamma_o}} + 2\frac{1-t}{\sum_k \omega_{k,1}}(\frac{2\xi_{o,1}}{\sigma_{o,1}})^{\xi_{o,1}^{-1}} - 1]}{1-t} \\
&\geq \frac{2p_o}{\sum_k \omega_{k,1}}(\frac{2\xi_{o,1}}{\sigma_{o,1}})^{\xi_{o,1}^{-1}}(1-2^{\gamma_o - 1}).
\end{aligned}
$$

If $\gamma_o \in [0, 1)$, then $\chi > 0$. □

S3 Computational details

We use Metropolis-within-Gibbs MCMC to implement the model. The hierarchical form of the bivariate model is

$$\begin{pmatrix} Y_i \\ X_i \end{pmatrix} \Bigg| \boldsymbol{\mu}^{(i)}, \boldsymbol{\Theta}, \boldsymbol{\gamma}, Z_i = k \sim \text{BGPD}(\boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}_k, \boldsymbol{\xi}_k; \gamma_k)$$

$$\boldsymbol{\mu}^{(i)} | \boldsymbol{u}_1, \ldots, \boldsymbol{u}_K, \Sigma, Z_i = k \sim N(\boldsymbol{u}_k, \Sigma)$$

$$Z_i | \mathbf{p} \sim \text{Cat}(\mathbf{p})$$

where $\boldsymbol{\Theta} = \{\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_K, \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K\}, \boldsymbol{\gamma} = \{\gamma_1, \ldots, \gamma_K\}$, $\mathbf{p} = (p_1, \ldots, p_K)$. The priors are set to be

$$\log(\boldsymbol{\sigma}_k) | \boldsymbol{\nu}_0, \Delta_0 \sim N(\boldsymbol{\nu}_0, \Delta_0)$$

$$\boldsymbol{\xi}_k | \boldsymbol{\iota}_0, \Omega_0 \sim N(\boldsymbol{\iota}_0, \Omega_0)$$

$$\gamma_k | \alpha_0, \beta_0 \sim N(\alpha_0, \beta_0)$$

$$\boldsymbol{u}_k | \boldsymbol{u}_0, \Sigma_0 \sim N(\boldsymbol{u}_0, \Sigma_0)$$

$$\Sigma | m_0, \Psi_0 \sim IW(\Psi_0, m_0)$$

$$\mathbf{p} | \boldsymbol{\rho} \sim Dir(\boldsymbol{\rho})$$

For $\{\boldsymbol{u}_k, k = 1, \ldots, K\}, \Sigma$ and $\mathbf{p}$, we can present the full conditional posterior densities,

$$f(\boldsymbol{u}_k|\mathbf{X},\mathbf{Y},\mathbf{Z},\cdot) \propto N(\tilde{\mu},\widetilde{\Sigma})$$

$$f(\Sigma|\mathbf{X},\mathbf{Y},\mathbf{Z},\cdot) \propto IW(\sum_{k=1}^{K}\sum_{i:Z_i=k}(\boldsymbol{\mu}^{(i)}-\boldsymbol{u}_k)(\boldsymbol{\mu}^{(i)}-\boldsymbol{u}_k)^T+\Psi_0, n+m_0)$$

$$f(\{p_1,\ldots,p_K\}|\mathbf{X},\mathbf{Y},\mathbf{Z},\cdot) \propto \mathrm{Dir}(\rho_1+n_1,\ldots,\rho_K+n_K)$$

$$f(Z_i=k|\mathbf{X},\mathbf{Y},\mathbf{Z},\cdot) \propto \frac{p_k g(\mu_1^{(i)},\mu_2^{(i)}|\boldsymbol{u}_k,\Sigma)h(X_i,Y_i|\boldsymbol{\sigma}_k,\boldsymbol{\xi}_k,\boldsymbol{\mu}^{(i)};\gamma_k)}{\sum_{k=1}^{K} p_k g(\mu_1^{(i)},\mu_2^{(i)}|\boldsymbol{u}_k,\Sigma)h(X_i,Y_i|\boldsymbol{\sigma}_k,\boldsymbol{\xi}_k,\boldsymbol{\mu}^{(i)};\gamma_k)}$$

where

$$\widetilde{\Sigma}^{-1} = \Sigma_0^{-1} + n\Sigma^{-1}$$

$$\tilde{\mu} = \widetilde{\Sigma}\{\Sigma_0^{-1}\boldsymbol{u}_0 + \Sigma^{-1}(\sum_{i:Z_i=k}\boldsymbol{\mu}^{(i)})\}$$

$$n_k = \sum_{i=1}^{n} I(Z_i=k), k=1,\ldots,K$$

The remains parameters and the latent variable $\boldsymbol{\mu}^{(i)}, i=1,\ldots,n$ are updated using Metropolis sampling. For the GPD related parameter $\{(\log(\boldsymbol{\sigma}_k), \boldsymbol{\xi}_k, \gamma_k),\ k=1,\ldots,K\}$, we generate the candidate of $(\log(\boldsymbol{\sigma}_k), \boldsymbol{\xi}_k, \gamma_k)$ from the normal distribution

$$\begin{pmatrix} \log(\boldsymbol{\sigma}_k^{(c)}) \\ \boldsymbol{\xi}_k^{(c)} \\ \Phi^{-1}(\gamma_k^{(c)}) \end{pmatrix} \sim N\left(\begin{pmatrix} \log(\boldsymbol{\sigma}_k^{(r-1)}) \\ \boldsymbol{\xi}_k^{(r-1)} \\ \Phi_{mean=0.05}^{-1}(\gamma_k^{(r-1)}) \end{pmatrix}, s_1^2 I\right).$$

The acceptance ratios are

$$R_{\boldsymbol{\sigma}_k} = \min\left\{1, \frac{\prod_{i:Z_i=k} l(X_i, Y_i|\boldsymbol{\sigma}_k^{(c)}, rest)}{\prod_{i:Z_i=k} l(X_i, Y_i|\boldsymbol{\sigma}_k^{(r-1)}, rest)} \times \frac{q(\boldsymbol{\sigma}_k^{(c)}|\boldsymbol{\sigma}_k^{(r-1)})}{q(\boldsymbol{\sigma}_k^{(r-1)}|\boldsymbol{\sigma}_k^{(c)})}\right\}$$

$$R_{\boldsymbol{\xi}_k} = \min\left\{1, \frac{\prod_{i:Z_i=k} l(X_i, Y_i|\boldsymbol{\xi}_k^{(c)}, rest)}{\prod_{i:Z_i=k} l(X_i, Y_i|\boldsymbol{\xi}_k^{(r-1)}, rest)} \times \frac{q(\boldsymbol{\xi}_k^{(c)}|\boldsymbol{\xi}_k^{(r-1)})}{q(\boldsymbol{\xi}_k^{(r-1)}|\boldsymbol{\xi}_k^{(c)})}\right\}$$

$$R_{\gamma_k} = \min\left\{1, \frac{\prod_{i:Z_i=k} l(X_i, Y_i|\gamma_k^{(c)}, rest)}{\prod_{i:Z_i=k} l(X_i, Y_i|\gamma_k^{(r-1)}, rest)} \times \frac{q(\gamma_k^{(c)}|\gamma_k^{(r-1)})}{q(\gamma_k^{(r-1)}|\gamma_k^{(c)})}\right\}$$

where $l(\cdot)$ is the likelihood function. For the latent parameter $\boldsymbol{\mu}^{(i)}, i = 1, \ldots, n$, we generate $\boldsymbol{\mu}^{(i,c)} \sim N(\boldsymbol{\mu}^{(i,r-1)}, \Sigma^{(r-1)})$, the acceptance ratio for latent parameter $\boldsymbol{\mu}^{(i)}, i = 1, \ldots, n$ is

$$R_{\boldsymbol{\mu}^{(i)}} = \min\left\{1, \frac{l(X_i, Y_i|\boldsymbol{\mu}^{(i,c)}, rest)}{l(X_i, Y_i|\boldsymbol{\mu}^{(i,r-1)}, rest)} \times \frac{q(\boldsymbol{\mu}^{(i,c)}|\boldsymbol{\mu}^{(i,r-1)})}{q(\boldsymbol{\mu}^{(i,r-1)}|\boldsymbol{\mu}^{(i,c)})}\right\}.$$

The present procedure of fitting bivariate MIXGP model can be modified to fit univariate MIXGP model directly.

S4 Additional results

Table 1: S1: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S1	0.50	0.50	**17.18**	-96.09	-103.72	-105.16	53.67	-165.76	53.67
			(2.60)	(0.84)	(1.38)	(2.01)	(1.33)	(8.43)	(1.33)
		0.95	**-18.57**	-95.97	-99.32	-100.06	-38.39	-130.56	-38.39
			(1.02)	(0.36)	(0.60)	(0.87)	(0.57)	(8.49)	(0.57)
		0.99	**-23.49**	-36.33	-37.53	-37.81	-43.25	-122.24	-42.15
			(0.62)	(0.13)	(0.22)	(0.31)	(0.43)	(7.74)	(0.63)
		0.995	**-22.38**	-22.62	-185.23	-23.54	-41.52	-117.23	-33.19
			(0.58)	(0.08)	(0.68)	(0.19)	(0.40)	(7.54)	(0.95)
		0.999	-19.20	-7.05	-178.60	**-7.33**	-37.08	-108.08	-17.22
			(0.53)	(0.02)	(0.67)	(0.06)	(0.35)	(7.24)	(0.64)
	0.9	0.50	**-5.75**	-87.08	-109.57	-90.38	43.59	-391.27	43.59
			(2.76)	(1.18)	(1.18)	(1.71)	(2.13)	(4.25)	(2.13)
		0.95	6.14	-77.10	-85.87	-78.44	**1.12**	-182.16	4.97
			(1.44)	(0.45)	(0.48)	(0.66)	(0.61)	(15.47)	(0.98)
		0.99	-8.88	-28.75	-31.75	-29.21	-15.47	-163.11	**1.32**
			(1.05)	(0.15)	(0.16)	(0.23)	(0.38)	(14.23)	(0.31)
		0.995	-10.52	-18.07	-144.47	-18.35	-16.54	-156.20	**-0.70**
			(0.95)	(0.09)	(0.67)	(0.14)	(0.34)	(13.91)	(0.22)
		0.999	-10.21	-5.43	-138.11	-5.51	-15.68	-144.05	**-2.07**
			(0.83)	(0.03)	(0.64)	(0.04)	(0.28)	(13.43)	(0.14)

Table 2: S2: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S2	0.50	0.50	**34.32**	86.92	127.72	36.68	55.01	-111.27	55.01
			(8.63)	(1.17)	(5.30)	(4.00)	(1.28)	(3.21)	(1.28)
		0.95	**-0.76**	-22.62	-2.00	-9.04	-15.14	-10.91	-15.14
			(0.97)	(0.63)	(1.33)	(1.48)	(0.54)	(0.51)	(0.54)
		0.99	**-2.58**	-18.55	-8.00	-9.51	-12.19	-14.75	-5.16
			(0.51)	(0.35)	(0.64)	(0.84)	(0.27)	(0.34)	(0.31)
		0.995	**-2.26**	-15.80	-7.57	-8.48	-10.34	-13.87	-3.90
			(0.35)	(0.28)	(0.50)	(0.67)	(0.21)	(0.29)	(0.23)
		0.999	**-1.28**	-10.83	-5.65	-6.05	-6.92	-11.06	-2.49
			(0.18)	(0.18)	(0.33)	(0.43)	(0.13)	(0.23)	(0.13)
	0.9	0.50	**2.39**	133.54	78.57	115.00	109.33	-56.74	109.33
			(5.06)	(1.36)	(4.37)	(4.04)	(2.16)	(3.40)	(2.16)
		0.95	**-12.98**	-90.05	-99.52	-65.05	-49.30	-89.92	-49.30
			(1.98)	(0.96)	(2.12)	(2.60)	(0.91)	(1.36)	(0.91)
		0.99	**-2.32**	-79.25	-80.98	-54.61	-39.59	-72.08	-38.08
			(0.94)	(0.78)	(2.04)	(2.22)	(0.58)	(1.03)	(0.75)
		0.995	**-2.13**	-71.65	-71.96	-48.24	-34.03	-64.40	-27.83
			(0.75)	(0.72)	(1.96)	(2.06)	(0.49)	(0.94)	(0.70)
		0.999	**-1.63**	-56.43	-55.12	-36.00	-23.65	-50.26	-13.97
			(0.61)	(0.60)	(1.78)	(1.75)	(0.36)	(0.80)	(0.41)

Table 3: S3: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S3	0.50	0.50	**12.08**	-96.60	-107.63	-108.04	54.60	-133.39	54.60
			(2.80)	(0.69)	(1.17)	(1.71)	(1.35)	(11.04)	(1.35)
		0.95	**-14.95**	-127.66	-133.31	-133.62	-27.28	-140.86	-27.28
			(1.18)	(0.35)	(0.60)	(0.86)	(0.61)	(18.32)	(0.61)
		0.99	**-18.69**	-58.10	-60.57	-60.71	-31.73	-133.62	-21.25
			(0.58)	(0.15)	(0.26)	(0.37)	(0.42)	(17.48)	(0.84)
		0.995	**-17.47**	-40.06	-200.36	-41.84	-30.19	-128.79	-14.78
			(0.56)	(0.10)	(0.57)	(0.26)	(0.38)	(17.21)	(0.60)
		0.999	-13.60	-13.20	-190.63	-13.80	-24.87	-117.55	**-8.01**
			(0.46)	(0.03)	(0.55)	(0.09)	(0.31)	(16.72)	(0.35)
	0.9	0.50	50.71	-60.61	-62.47	-53.39	**48.44**	-109.93	48.44
			(2.86)	(1.41)	(1.88)	(4.16)	(2.31)	(10.27)	(2.31)
		0.95	**-56.12**	-106.65	-107.57	-104.18	-87.14	-188.52	-87.14
			(1.61)	(0.64)	(0.86)	(1.92)	(1.16)	(15.79)	(1.16)
		0.99	-50.88	-48.20	-48.59	**-47.26**	-81.63	-180.42	-81.63
			(1.34)	(0.27)	(0.36)	(0.81)	(0.89)	(16.01)	(0.89)
		0.995	-46.57	-32.12	-238.21	**-31.52**	-76.56	-174.40	-76.56
			(1.27)	(0.18)	(0.73)	(0.53)	(0.82)	(16.01)	(0.82)
		0.999	-39.04	-12.57	-232.69	**-12.35**	-67.53	-163.76	-67.53
			(1.20)	(0.07)	(0.72)	(0.20)	(0.73)	(15.99)	(0.73)

Table 4: S4: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S4	0.50	0.50	**7.91**	96.41	106.16	104.47	55.65	-95.18	55.65
			(2.26)	(0.71)	(1.19)	(2.01)	(1.27)	(3.00)	(1.27)
		0.95	**-9.27**	-107.27	-100.24	-101.70	-34.12	-58.29	-34.12
			(1.09)	(0.53)	(0.86)	(1.48)	(0.57)	(3.38)	(0.57)
		0.99	**-4.92**	-92.46	-86.99	-88.20	-25.66	-43.83	-16.53
			(0.54)	(0.41)	(0.67)	(1.16)	(0.34)	(2.75)	(0.57)
		0.995	**-3.71**	-82.28	-77.35	-78.47	-20.61	-269.70	-9.76
			(0.47)	(0.37)	(0.60)	(1.05)	(0.27)	(34.85)	(0.43)
		0.999	**-2.38**	-63.63	-59.61	-60.56	-12.61	-256.02	-4.28
			(0.33)	(0.31)	(0.49)	(0.86)	(0.18)	(35.04)	(0.25)
	0.9	0.50	**28.68**	86.95	108.54	87.98	51.69	-96.96	51.69
			(7.54)	(1.14)	(1.17)	(1.64)	(2.05)	(3.67)	(2.05)
		0.95	**10.76**	-117.37	-101.34	-116.72	-26.04	-58.10	-26.04
			(1.23)	(0.86)	(0.84)	(1.25)	(0.78)	(3.47)	(0.78)
		0.99	**-4.15**	-118.79	-104.94	-118.25	-32.25	-59.02	-25.85
			(0.66)	(0.75)	(0.72)	(1.09)	(0.52)	(3.16)	(0.87)
		0.995	**-4.96**	-113.45	-100.27	-112.94	-29.86	-267.28	-19.52
			(0.52)	(0.72)	(0.68)	(1.03)	(0.45)	(33.58)	(0.67)
		0.999	**-3.35**	-94.91	-83.37	-94.48	-20.30	-250.45	-8.81
			(0.33)	(0.63)	(0.59)	(0.91)	(0.32)	(33.34)	(0.38)

Table 5: S1: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S1	0.50	0.50	**0.98**	9.30	10.95	11.47	3.06	34.66	3.06
			(0.21)	(0.16)	(0.29)	(0.42)	(0.15)	(4.98)	(0.15)
		0.95	**0.45**	9.22	9.90	10.09	1.51	24.33	1.51
			(0.04)	(0.07)	(0.12)	(0.17)	(0.05)	(4.77)	(0.05)
		0.99	**0.59**	1.32	1.41	1.44	1.89	20.99	1.82
			(0.04)	(0.01)	(0.02)	(0.02)	(0.04)	(4.03)	(0.05)
		0.995	0.54	**0.51**	34.36	0.56	1.74	19.48	1.19
			(0.03)	(<0.01)	(0.26)	(0.01)	(0.03)	(3.80)	(0.07)
		0.999	0.40	**0.05**	31.94	0.05	1.39	16.98	0.34
			(0.03)	(<0.01)	(0.24)	(<0.01)	(0.03)	(3.44)	(0.03)
	0.9	0.50	**0.80**	7.72	12.15	8.47	2.36	154.91	2.36
			(0.13)	(0.21)	(0.27)	(0.33)	(0.19)	(3.57)	(0.19)
		0.95	0.25	5.97	7.40	6.20	**0.04**	57.34	0.12
			(0.07)	(0.07)	(0.08)	(0.11)	(0.01)	(13.90)	(0.02)
		0.99	0.19	0.83	1.01	0.86	0.25	47.05	**0.01**
			(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	(11.67)	(<0.01)
		0.995	0.20	0.33	20.92	0.34	0.29	43.93	**0.01**
			(0.06)	(<0.01)	(0.20)	(0.01)	(0.01)	(11.07)	(<0.01)
		0.999	0.18	0.03	19.11	0.03	0.25	38.96	**0.01**
			(0.05)	(<0.01)	(0.18)	(<0.01)	(0.01)	(10.14)	(<0.01)

Table 6: S2: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S2	0.50	0.50	8.73	7.69	19.15	2.96	3.19	13.42	3.19
			(3.50)	(0.19)	(0.60)	(0.31)	(0.14)	(0.76)	(0.14)
		0.95	**0.10**	0.55	0.18	0.30	0.26	0.15	0.26
			(0.02)	(0.03)	(0.07)	(0.04)	(0.02)	(0.01)	(0.02)
		0.99	**0.03**	0.36	0.11	0.16	0.16	0.23	0.04
			(0.01)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(<0.01)
		0.995	**0.02**	0.26	0.08	0.12	0.11	0.20	0.02
			(<0.01)	(0.01)	(0.02)	(0.02)	(<0.01)	(0.01)	(<0.01)
		0.999	**0.01**	0.12	0.04	0.06	0.05	0.13	0.01
			(<0.01)	(<0.01)	(0.01)	(0.01)	(<0.01)	(0.01)	(<0.01)
	0.9	0.50	**2.60**	18.02	8.10	14.88	12.42	4.39	12.42
			(0.75)	(0.36)	(0.61)	(0.91)	(0.48)	(0.43)	(0.48)
		0.95	**0.57**	8.20	10.36	4.91	2.51	8.27	2.51
			(0.08)	(0.18)	(0.42)	(0.37)	(0.09)	(0.25)	(0.09)
		0.99	**0.09**	6.34	6.98	3.48	1.60	5.30	1.51
			(0.03)	(0.12)	(0.33)	(0.27)	(0.05)	(0.15)	(0.06)
		0.995	**0.06**	5.19	5.57	2.75	1.18	4.24	0.82
			(0.03)	(0.10)	(0.28)	(0.23)	(0.03)	(0.13)	(0.04)
		0.999	**0.04**	3.22	3.36	1.60	0.57	2.59	0.21
			(0.03)	(0.07)	(0.20)	(0.15)	(0.02)	(0.08)	(0.01)

Table 7: S3: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S3	0.50	0.50	**0.94**	9.38	11.72	11.97	3.16	30.10	3.16
			(0.35)	(0.14)	(0.26)	(0.34)	(0.15)	(6.56)	(0.15)
		0.95	**0.36**	16.31	17.81	17.93	0.78	53.74	0.78
			(0.03)	(0.09)	(0.16)	(0.23)	(0.03)	(15.41)	(0.03)
		0.99	**0.38**	3.38	3.68	3.70	1.02	48.71	0.52
			(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(14.00)	(0.04)
		0.995	0.34	1.61	40.18	1.76	0.93	46.51	**0.25**
			(0.02)	(0.01)	(0.23)	(0.02)	(0.02)	(13.49)	(0.02)
		0.999	0.21	0.17	36.37	0.19	0.63	42.06	**0.08**
			(0.02)	(<0.01)	(0.21)	(0.00)	(0.02)	(12.48)	(0.01)
	0.9	0.50	3.40	3.88	4.26	4.60	**2.88**	22.74	2.88
			(0.38)	(0.18)	(0.25)	(0.59)	(0.24)	(5.10)	(0.24)
		0.95	**3.41**	11.42	11.65	11.23	7.73	60.72	7.73
			(0.22)	(0.14)	(0.19)	(0.44)	(0.21)	(13.48)	(0.21)
		0.99	2.77	2.33	2.37	**2.30**	6.74	58.43	6.74
			(0.17)	(0.03)	(0.04)	(0.08)	(0.15)	(13.52)	(0.15)
		0.995	2.33	1.03	56.80	**1.02**	5.93	56.31	5.93
			(0.15)	(0.01)	(0.35)	(0.04)	(0.13)	(13.33)	(0.13)
		0.999	1.67	0.16	54.19	**0.16**	4.61	52.63	4.61
			(0.12)	(< 0.01)	(0.34)	(0.01)	(0.10)	(12.96)	(0.10)

Table 8: S4: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S4	0.50	0.50	**0.58**	9.35	11.41	11.32	3.26	9.97	3.26
			(0.13)	(0.14)	(0.26)	(0.40)	(0.14)	(0.81)	(0.14)
		0.95	**0.21**	11.54	10.12	10.56	1.20	4.55	1.20
			(0.03)	(0.11)	(0.17)	(0.32)	(0.04)	(1.02)	(0.04)
		0.99	0.05	8.57	7.61	7.92	0.67	2.69	**0.31**
			(0.01)	(0.08)	(0.12)	(0.22)	(0.02)	(0.66)	(0.02)
		0.995	**0.04**	6.78	6.02	6.27	0.43	195.42	0.11
			(0.01)	(0.06)	(0.09)	(0.17)	(0.01)	(34.95)	(0.01)
		0.999	**0.02**	4.06	3.58	3.74	0.16	189.60	0.02
			(<0.01)	(0.04)	(0.06)	(0.11)	(<0.01)	(35.16)	(<0.01)
	0.9	0.50	6.59	7.69	11.92	8.01	**3.10**	10.76	3.10
			(2.74)	(0.20)	(0.26)	(0.29)	(0.23)	(0.74)	(0.23)
		0.95	**0.27**	13.85	10.34	13.78	0.74	4.59	0.74
			(0.04)	(0.20)	(0.17)	(0.30)	(0.04)	(1.01)	(0.04)
		0.99	**0.06**	14.17	11.07	14.10	1.07	4.49	0.74
			(0.01)	(0.18)	(0.15)	(0.26)	(0.03)	(0.89)	(0.05)
		0.995	**0.05**	12.92	10.10	12.86	0.91	185.33	0.43
			(0.01)	(0.16)	(0.14)	(0.24)	(0.03)	(34.82)	(0.03)
		0.999	**0.02**	9.05	6.99	9.01	0.42	174.98	0.09
			(0.01)	(0.12)	(0.10)	(0.18)	(0.01)	(34.16)	(0.01)

Table 9: S5: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S5	0.50	0.50	7.07	64.42	71.89	77.00	335.90	**-0.10**	335.90
			(2.92)	(0.86)	(1.17)	(1.36)	(1.48)	(28.42)	(1.48)
		0.95	**-13.06**	-81.31	-76.42	-73.14	24.43	-251.26	21.29
			(1.38)	(0.58)	(0.77)	(0.89)	(0.42)	(31.10)	(0.54)
		0.99	-5.15	-43.45	-40.74	-38.94	2.98	-232.18	**-1.37**
			(0.68)	(0.32)	(0.43)	(0.49)	(0.16)	(32.41)	(0.39)
		0.995	-3.29	-28.74	-26.80	-25.53	**1.32**	-82.61	-2.51
			(0.49)	(0.23)	(0.31)	(0.35)	(0.09)	(6.01)	(0.32)
		0.999	-1.38	-8.00	-7.32	-6.88	**0.40**	-46.71	-1.67
			(0.23)	(0.08)	(0.11)	(0.12)	(0.02)	(3.37)	(0.17)
	0.9	0.50	**-18.96**	61.87	48.70	60.45	260.65	-151.94	260.65
			(3.96)	(1.71)	(3.06)	(3.18)	(3.38)	(31.12)	(3.38)
		0.95	**-7.62**	-49.30	-58.03	-50.96	18.59	-343.41	15.10
			(1.71)	(0.99)	(1.89)	(1.94)	(0.81)	(31.40)	(0.79)
		0.99	**-2.64**	-22.06	-26.52	-23.13	2.73	-300.21	-3.50
			(0.66)	(0.46)	(0.93)	(0.96)	(0.24)	(31.60)	(0.59)
		0.995	-1.78	-12.97	-16.00	-13.76	**1.51**	-153.21	-3.96
			(0.44)	(0.30)	(0.63)	(0.64)	(0.13)	(11.72)	(0.48)
		0.999	-0.90	-2.42	-3.33	-2.71	**0.56**	-93.32	-2.30
			(0.21)	(0.08)	(0.18)	(0.19)	(0.02)	(7.19)	(0.26)

Table 10: S6: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses.. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S6	0.50	0.50	**6.18**	84.27	38.63	95.79	198.11	44.16	198.11
			(4.78)	(0.83)	(3.09)	(4.03)	(1.82)	(20.93)	(1.82)
		0.95	-11.06	-92.37	-127.24	-85.93	**2.91**	-122.11	3.91
			(2.24)	(0.58)	(2.23)	(2.83)	(0.57)	(28.58)	(0.70)
		0.99	-2.42	-54.42	-77.12	-51.13	-1.28	-120.53	**-0.59**
			(1.76)	(0.36)	(1.41)	(1.75)	(0.20)	(29.31)	(0.40)
		0.995	-3.12	-37.95	-55.57	-35.71	**-3.94**	-123.33	-5.60
			(0.51)	(0.27)	(1.08)	(1.32)	(0.17)	(6.54)	(0.42)
		0.999	-1.46	-12.42	-20.33	-11.77	**-0.81**	-71.85	-2.47
			(0.26)	(0.11)	(0.47)	(0.56)	(0.05)	(3.92)	(0.23)
	0.9	0.50	**19.05**	97.05	69.59	82.00	33.18	-89.36	33.18
			(5.26)	(1.28)	(7.58)	(3.84)	(3.58)	(17.63)	3.58
		0.95	**-4.60**	-94.27	-119.57	-106.32	-42.21	-168.24	-42.21
			(2.14)	(0.90)	(5.41)	(2.76)	(1.69)	(28.09)	(1.69)
		0.99	**-2.66**	-57.77	-76.36	-66.00	-13.70	-144.61	-11.48
			(1.04)	(0.57)	(3.49)	(1.77)	(0.64)	(29.18)	(0.81)
		0.995	-2.73	-41.08	-56.37	-47.63	**-4.28**	-191.11	-6.44
			(0.71)	(0.44)	(2.71)	(1.38)	(0.29)	(10.31)	(0.60)
		0.999	-1.46	-14.24	-22.22	-17.40	**-0.70**	-122.88	-2.86
			(0.32)	(0.19)	(1.25)	(0.62)	(0.07)	(6.82)	(0.32)

Table 11: S7: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S7	0.50	0.50	**4.76**	117.17	11.86	115.84	199.29	45.70	199.29
			(5.83)	(0.70)	(4.63)	(5.71)	(1.82)	(20.98)	(1.82)
		0.95	**-13.98**	-124.21	-216.62	-127.81	-35.92	-161.95	-35.92
			(1.34)	(0.54)	(3.79)	(4.69)	(0.85)	(27.72)	(0.85)
		0.99	**-5.05**	-88.57	-161.58	-92.51	-23.49	-147.51	-21.13
			(0.67)	(0.40)	(2.89)	(3.56)	(0.45)	(28.85)	(0.63)
		0.995	**-3.78**	-68.75	-131.98	-72.59	-15.59	-46.12	-13.66
			(0.53)	(0.33)	(2.47)	(3.02)	(0.31)	(6.65)	(0.52)
		0.999	**-1.99**	-31.10	-70.67	-34.17	-4.03	-23.61	-4.39
			(0.36)	(0.18)	(1.51)	(1.80)	(0.11)	(4.28)	(0.27)
	0.9	0.50	**20.98**	101.44	99.96	89.32	158.95	56.28	158.95
			(4.49)	(1.09)	(11.81)	(4.91)	(3.37)	(22.13)	(3.37)
		0.95	**-4.16**	-167.65	-178.79	-179.73	-56.22	-180.11	-56.22
			(2.06)	(0.93)	(9.28)	(4.24)	(1.86)	(27.84)	(1.86)
		0.99	**-4.59**	-130.99	-144.65	-141.70	-32.73	-164.21	-31.17
			(1.09)	(0.75)	(7.32)	(3.47)	(0.99)	(28.90)	(1.20)
		0.995	**-4.13**	-108.05	-122.02	-117.90	-21.50	-65.09	-19.14
			(0.89)	(0.67)	(6.41)	(3.07)	(0.69)	(9.12)	(0.87)
		0.999	**-3.03**	-59.01	-71.77	-66.34	-5.49	-37.21	-5.89
			(0.61)	(0.44)	(4.23)	(2.08)	(0.23)	(6.48)	(0.42)

Table 12: S8: $(E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x})] - t_y) \times 1000$(Bias) for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S8	0.50	0.50	7.53	**-2.07**	-40.77	-18.51	-49.67	-209.61	-49.67
			(3.74)	(1.25)	(2.15)	(3.90)	(1.38)	(5.10)	(1.38)
		0.95	2.41	**-0.00**	-15.17	-4.39	-0.70	-35.04	-1.19
			(2.04)	(0.58)	(1.00)	(1.30)	(0.85)	(13.02)	(0.80)
		0.99	**-0.02**	-0.08	-5.02	-1.30	3.88	-34.00	-1.11
			(0.95)	(0.25)	(0.45)	(0.50)	(0.19)	(13.59)	(0.32)
		0.995	-1.01	**-0.10**	-3.08	-0.86	2.81	-21.95	-0.83
			(0.69)	(0.16)	(0.31)	(0.33)	(0.08)	(9.73)	(0.21)
		0.999	-1.09	**-0.09**	-1.02	-0.41	0.85	-17.13	-0.32
			(0.36)	(0.06)	(0.13)	(0.13)	(0.01)	(9.82)	(0.08)
	0.9	0.50	-42.90	**-3.70**	-33.67	32.56	-37.50	-124.17	-37.50
			(6.08)	(2.28)	(3.91)	(4.60)	(2.58)	(6.45)	(2.58)
		0.95	-6.51	**-0.89**	-18.98	7.59	29.39	-23.67	21.68
			(2.63)	(0.90)	(1.67)	(1.64)	(0.59)	(13.18)	(0.60)
		0.99	-1.24	**-0.40**	-7.62	1.43	9.01	-29.60	4.39
			(0.92)	(0.32)	(0.72)	(0.54)	(0.05)	(13.66)	(0.24)
		0.995	-1.29	**-0.30**	-5.06	0.53	4.77	-18.80	2.17
			(0.61)	(0.20)	(0.51)	(0.34)	(0.01)	(9.77)	(0.15)
		0.999	-1.18	**-0.15**	-2.00	-0.11	1.00	-15.60	0.38
			(0.34)	(0.07)	(0.24)	(0.13)	(<0.01)	(9.84)	(0.05)

Table 13: S5: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S5	0.50	0.50	**0.91**	4.23	5.31	6.12	113.05	81.59	113.05
			(0.29)	(0.12)	(0.17)	(0.20)	(1.01)	(8.38)	(1.01)
		0.95	**0.36**	6.64	5.90	5.43	0.61	160.80	0.48
			(0.05)	(0.09)	(0.12)	(0.14)	(0.02)	(31.41)	(0.02)
		0.99	0.07	1.90	1.68	1.54	**0.01**	160.01	0.02
			(0.02)	(0.03)	(0.04)	(0.04)	(<0.01)	(34.65)	(<0.01)
		0.995	0.04	0.83	0.73	0.66	**<0.01**	10.47	0.02
			(0.01)	(0.01)	(0.02)	(0.02)	(<0.01)	(0.90)	(<0.01)
		0.999	0.01	0.06	0.05	0.05	**<0.01**	3.33	0.01
			(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(0.29)	(<0.01)
	0.9	0.50	**1.94**	4.12	3.32	4.68	69.09	120.93	69.09
			(0.41)	(0.20)	(0.31)	(0.35)	(1.78)	(7.31)	(1.78)
		0.95	0.35	2.53	3.73	2.98	0.41	217.52	**0.29**
			(0.08)	(0.11)	(0.25)	(0.26)	(0.03)	(29.99)	(0.03)
		0.99	0.05	0.51	0.79	0.63	**0.01**	191.02	0.05
			(0.01)	(0.02)	(0.06)	(0.06)	(<0.01)	(33.58)	(0.01)
		0.995	0.02	0.18	0.30	0.23	**<0.01**	37.34	0.04
			(<0.01)	(0.01)	(0.02)	(0.03)	(<0.01)	(3.18)	(0.01)
		0.999	0.01	0.01	0.01	0.01	**<0.01**	13.93	0.01
			(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(1.20)	(<0.01)

Table 14: S6: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S6	0.50	0.50	**2.35**	7.17	2.45	10.81	39.58	46.20	39.58
			(0.86)	(0.14)	(0.50)	(0.69)	(0.73)	(6.93)	(0.73)
		0.95	0.63	8.57	16.69	8.19	**0.04**	97.39	0.06
			(0.20)	(0.11)	(0.48)	(0.55)	(0.01)	(27.07)	(0.01)
		0.99	0.32	2.97	6.15	2.92	**0.01**	101.30	0.02
			(0.24)	(0.04)	(0.19)	(0.20)	(<0.01)	(29.46)	(<0.01)
		0.995	0.04	1.45	3.21	1.45	**0.02**	19.53	0.05
			(0.01)	(0.02)	(0.11)	(0.11)	(<0.01)	(1.19)	(0.01)
		0.999	0.01	0.16	0.44	0.17	<**0.01**	6.71	0.01
			(<0.01)	(<0.01)	(0.02)	(0.02)	(<0.01)	(0.42)	(<0.01)
	0.90	0.50	3.16	9.58	10.64	8.21	2.40	39.40	**2.40**
			(0.72)	(0.25)	(1.51)	(0.73)	(0.30)	(8.23)	(0.30)
		0.95	**0.48**	8.97	17.25	12.07	2.07	108.03	2.07
			(0.09)	(0.17)	(1.18)	(0.59)	(0.15)	(27.07)	(0.15)
		0.99	0.12	3.37	7.06	4.67	0.23	106.93	**0.20**
			(0.03)	(0.07)	(0.50)	(0.24)	(0.02)	(29.37)	(0.02)
		0.995	0.06	1.71	3.92	2.46	**0.03**	47.25	0.08
			(0.02)	(0.04)	(0.29)	(0.14)	(<0.01)	(2.87)	(0.01)
		0.999	0.01	0.21	0.65	0.34	<**0.01**	19.80	0.02
			(<0.01)	(0.01)	(0.06)	(0.02)	(<0.01)	(1.21)	(<0.01)

Table 15: S7: $E[\hat{P}(Y \leq Y^C_{t_y}|X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S7	0.50	0.50	**3.46**	13.78	2.30	16.72	40.05	46.54	40.05
			(2.39)	(0.16)	(0.96)	(1.07)	(0.74)	(6.92)	(0.74)
		0.95	**0.38**	15.46	48.37	18.56	1.36	103.86	1.36
			(0.06)	(0.14)	(1.24)	(1.43)	(0.06)	(27.03)	(0.06)
		0.99	**0.07**	7.86	26.95	9.84	0.57	105.85	0.49
			(0.01)	(0.07)	(0.73)	(0.80)	(0.02)	(29.38)	(0.02)
		0.995	**0.04**	4.74	18.04	6.19	0.25	6.60	0.21
			(0.01)	(0.05)	(0.51)	(0.54)	(0.01)	(1.68)	(0.01)
		0.999	**0.02**	0.97	5.22	1.49	0.02	2.41	0.03
			(0.01)	(0.01)	(0.17)	(0.16)	(<0.01)	(0.68)	(<0.01)
	0.9	0.50	**2.48**	10.41	24.09	10.41	26.41	52.62	26.41
			(0.60)	(0.23)	(4.07)	(1.01)	(1.11)	(6.78)	(1.11)
		0.95	**0.45**	28.19	40.67	34.12	3.51	110.75	3.51
			(0.10)	(0.31)	(2.97)	(1.52)	(0.21)	(27.08)	(0.21)
		0.99	**0.14**	17.22	26.34	21.29	1.17	111.33	1.12
			(0.04)	(0.20)	(1.95)	(0.99)	(0.07)	(29.34)	(0.07)
		0.995	**0.10**	11.72	19.04	14.86	0.51	12.64	0.44
			(0.03)	(0.15)	(1.44)	(0.74)	(0.03)	(3.05)	(0.03)
		0.999	0.05	3.50	6.96	4.84	**0.04**	5.63	0.05
			(0.01)	(0.05)	(0.57)	(0.29)	(<0.01)	(1.47)	(0.01)

Table 16: S8: $E[\hat{P}(Y \leq Y_{t_y}^C | X_{t_x}) - t_y]^2$ for covariate quantile levels $t_x = 0.5, 0.9$ and conditional response quantile level $t_y = 0.5, 0.95, 0.99, 0.995, 0.999$. Standard errors are given in parentheses. Bold represents the smallest MSE among four models under the same setting. "MIXGP" stands for our proposed MIXGP model, "GEPD" stands for the GEPD model, "C3" stands for the CGPD model with 3 clusters, "C5" stands for the CGPD model with 5 clusters, "GAM" stands for the gamma model, "GEV" stands for the GEV model and "MIX-t" stands for the tail mixture model.

Setting	t_x	t_y	MIXGP	GEPD	C3	C5	GAM	GEV	MIX-t
S8	0.50	0.50	1.47	**0.16**	2.13	1.88	2.66	46.57	2.66
			(0.24)	(0.02)	(0.17)	(0.19)	(0.13)	(3.16)	(0.13)
		0.95	0.43	**0.03**	0.33	0.19	0.07	18.35	0.07
			(0.13)	(<0.01)	(0.03)	(0.02)	(0.01)	(12.69)	(0.01)
		0.99	0.09	**0.01**	0.05	0.03	0.02	19.82	0.01
			(0.04)	(<0.01)	(0.01)	(<0.01)	(<0.01)	(13.79)	(<0.01)
		0.995	0.05	<**0.01**	0.02	0.01	0.01	10.05	0.01
			(0.03)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(9.90)	(<0.01)
		0.999	0.01	<**0.01**	<0.01	<0.01	<0.01	10.03	<0.01
			(0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(9.98)	(<0.01)
	0.9	0.50	5.58	**0.54**	2.68	3.19	2.08	19.62	2.08
			(0.85)	(0.06)	(0.35)	(0.37)	(0.21)	(3.42)	(0.21)
		0.95	0.74	**0.08**	0.64	0.33	0.90	18.12	0.51
			(0.16)	(0.01)	(0.08)	(0.05)	(0.03)	(12.70)	(0.03)
		0.99	0.09	**0.01**	0.11	0.03	0.08	19.71	0.02
			(0.02)	(<0.01)	(0.01)	(<0.01)	(<0.01)	(13.79)	(<0.01)
		0.995	0.04	<**0.01**	0.05	0.01	0.02	9.99	0.01
			(0.01)	(<0.01)	(0.01)	(<0.01)	(<0.01)	(9.90)	(<0.01)
		0.999	0.01	<**0.01**	0.01	<0.01	<0.01	10.01	<0.01
			(<0.01)	(<0.01)	(<0.01)	(<0.01))	(0.00)	(9.98)	(<0.01)