arXiv:2601.03647v1 [stat.ME] 7 Jan 2026

相关极值指数的小区域估计

Koki Momoki[1]*和 Takuma Yoshida[2]†

[1,2]鹿儿岛大学理工学院，鹿儿岛县鹿儿岛市郡元*1-21-40*，邮编*890-8580*，日本

摘要

在极值分析中，重尾数据分布的尾部行为通过帕累托型分布建模，其中所谓的极值指数（EVI）控制着尾部行为。对于从多个总体子群或区域获得的重尾数据，本研究利用区域之间的信息，高效预测所有区域的极值指数。为此，我们提出一种混合效应模型，这是小区域估计中一种实用的方法。在该模型中，我们用称为随机效应的潜变量表示各区域极值指数的差异。通过引入区域间相关的随机效应，我们将区域之间的关系纳入模型。所得模型能够同时预测所有区域的极值指数。本文阐述该模型中的参数估计和随机效应预测，并明确估计量的理论性质。此外，我们通过数值实验展示所提方法的有效性。作为模型的一项应用，我们对日本的强降雨进行风险评估。

关键词：极值分析；混合效应模型；帕累托型分布；超阈值峰值法；风险评估；小区域估计

MSC 分类：*62E20*，*62G32*，*62P12*

1 引言

极值理论为分析各领域罕见事件发生的概率提供了优雅的统计方法。例如强降雨、极端温度、大额金融损失以及高额医疗费用。关于这一主题，已有大量教科书和综述论文问世，Gomes and Guillou（2015）对此作了介绍。极值分析中广为人知的方法包括区块极大值法和超阈值峰值法（POT）。区块极大值法使用广义极值分布对区块极大值建模

*电子邮箱：momoki@sci.kagoshima-u.ac.jp

†电子邮箱：yoshida@sci.kagoshima-u.ac.jp

例如年度最大值。超阈值法通过广义帕累托分布（GPD）拟合高阈值超越量，从而建立尾部模型。de Haan and Ferreira（2006，第一部分）讨论了这两种方法的数学细节及其联系。

在超阈值法的背景下，重尾分布这一类通常称为帕累托型分布，而非广义帕累托分布（Wang and Tsai 2009 的表1）。许多学者研究了针对帕累托型分布的超阈值法，包括 Gomes 等（2008a，2008b）、Beirlant 等（2009）和 Girard 等（2021）。该方法的主要任务是估计极值指数（EVI），这是一个决定帕累托型分布右尾行为的参数。Hill（1975）、Dekker 等（1989）和 Gomes 等（2008a）提出了多种极值指数估计量。其中，Hill（1975）提出的 Hill估计量是广泛使用的一种极值指数估计量。

本研究考察从多个总体子群中获得的重尾数据所适用的超阈值法。每个子群称为一个“区域”（Rao and Molina 2015；Molina 等，2022年），同一区域的数据在所关注的极端事件方面具有共同特征。预测极值指数最简单的策略是将 Hill估计量等经典方法直接应用于各个区域。然而，这种逐区域分析可能效率不高，因为超阈值法仅使用高阈值超越量，而且在许多应用中，每个区域的有效样本量往往较小。为克服这一难题，我们旨在利用区域之间的互信息，高效地预测所有区域的极值指数。在类似的背景下，Rohrbeck 和 Tawn（2019）提出了一种根据边际效应相似性对区域进行聚类的方法。然而，该方法会使某些区域的边际分布完全合并。此类方法或许便于解释，但可能无法捕捉区域之间的细微差异。与聚类方法不同，我们不假设区域之间具有同质性，也不合并不同区域的数据。作为利用区域间信息的另一种方法，我们考虑采用混合效应模型。

混合效应模型已在小区域估计框架下得到研究。Jiang and Nguyen（2007年）、Wu（2009）和Jiang（2017）介绍了小区域估计的标准方法。该模型通过借助协变量中的信息，尤其是在样本量较小的情形下，为各区域生成可靠的预测结果，因而十分有用。

然后，各区域之间模型参数的差异通过称为随机效应的潜变量来表示。通常假设各区域的随机效应彼此独立。在极值指数回归的背景下，Momoki和Yoshida（2025年）提出了采用混合效应模型的超阈值法。然而，他们的混合效应模型假设各区域的随机效应相互独立。因此，本研究开发了一种采用相关随机效应的方法，将区域间的关系直接纳入参数中。

本文的贡献与结构如下所述。第2 节介绍了我们针对帕累托型分布，采用混合效应模型的超阈值法。随后，我们通过相关随机效应反映区域之间的关系。针对这些相关随机效应，我们构建了用于估计参数的最大似然法，以及用于预测随机效应的条件众数法。结合这两种方法，我们可以预测极值指数。第 3节研究所提出估计量的数学性质。在采用独立随机效应的情况下，Nie（2007）和 Jiang 等人（2022）已确立广义混合效应模型中最大似然估计量的渐近正态性，而 Momoki和Yoshida（2025）则将他们的工作扩展至超阈值法。第 3 节中的渐近理论涵盖了相关随机效应。即使在广义混合效应模型和极值理论的背景下，这也可视为一项新颖成果。所得的渐近正态性表明，随着区域数量增加，所提出估计量的方差会减小。这一性质支持了我们纳入更多区域间信息的动机。对于所提出的方法，第 4 节通过模拟研究揭示了其数值表现特征。我们证实，通过恰当设定随机效应的相关性，我们的模型能够显著改善区域级估计。因此，第 5 节介绍了若干将区域间关系纳入随机效应相关性的示例。第 6 节将所提出的模型应用于日本极端降水的真实数据集。该数据集包含来自1138 个站点（即区域）的记录。对于这一大型数据集，与分区域估计方法相比，所提出的方法得出了更合理的结果。第 7 节总结了本文的要点。第 3 节中定理的证明见附录。

2 尾部概率的混合效应建模

我们考虑来自多个区域的观测数据。令 $J \in \mathbb{N}$ 表示区域数量，我们通过应用标签 $\mathrm{J} := \{1、2、\ldots\ldots、J\}$来标识这些区域。令

$$\{Y_{ij} \in \mathbb{R},\ i = 1, 2, \ldots, n_j,\ j \in \mathcal{J}\} \tag{1}$$

来自区域 $\mathcal{J}$的独立随机样本，其中 n_j是区域 $j \in \mathcal{J}$的样本量，且 Y_{ij}是区域 i 中的第 $j \in \mathcal{J}$个观测值。对于（1），我们假设来自同一区域的数据具有相同的区域特定的潜在分布。如第 1节所述，我们旨在利用区域之间关系的信息，预测所有区域 $\mathcal{J}$ 的潜在分布的右尾。为此，我们使用混合效应模型构建超阈值法。第 2.1 节介绍了针对（1）的最简单的超阈值法。随后，第 2.2 节和第 2.3节将阐述我们提出的超阈值法。

2.1 超阈值峰值法的经典方法

设 $F_j(y) := P(Y_{ij} \le y)$ 为从区域 $j \in \mathrm{J}$ 获取的数据 $\{Y_{ij}\}_{i=1}^{n_j}$ 的分布函数。极值分析通常关注准确评估潜在分布 $F_j(y)$、 $j \in \mathrm{J}$ 的右尾，其中这些目标可表示为

$$\begin{aligned} F_j(y \mid \omega_j) &:= P(Y_{ij} \le y \mid Y_{ij} > \omega_j) \\ &= \frac{F_j(y) - F_j(\omega_j)}{1 - F_j(\omega_j)}, \quad y > \omega_j,\ j \in \mathcal{J} \end{aligned} \tag{2}$$

对于某些高阈值$\{\omega_j\}_{j\in\mathcal{J}}$。为此，在极值分析中通常使用以下对（2）的近似（Smith 1987；Drees 等，2004）：

$$F_j(y \mid \omega_j) \approx 1 - \left(1 + \gamma_j \frac{y - \omega_j}{\psi_j}\right)_+^{-1/\gamma_j}, \quad j \in \mathcal{J}. \tag{3}$$

其中，式（3）右侧表示位置参数为 $\omega_j \in \mathbb{R}$、尺度参数为 $\psi_j > 0$、形状参数为 $\gamma_j \in \mathbb{R}$ 的广义帕累托分布的分布函数，且 $x_+ := \max\{0, x\}$（当 $x \in \mathbb{R}$ 时）。当 $\gamma_j = 0$ 时，式（3）右侧表示 $1 - \exp[-(y - \omega_j)/\psi_j]$，即指数分布。根据式（3），通过将高阈值超越量 $\{Y_{ij}\colon\ Y_{ij} > \omega_j,\ i = 1, 2, \ldots, n_j,\ j \in \mathrm{J}\}$ 拟合至广义帕累托分布并估计未知参数，可以预测目标 $F_j(y \mid \omega_j)$，$j \in \mathrm{J}$。

广义帕累托分布的未知参数。这种方法称为超阈值法，通常可依据极值理论在高阈值 $\{\omega_j\}_{j\in\mathcal{J}}$ 下得到论证（参见 de Haan和Ferreira（2006）第1章和第3章）。

本研究考察上述超阈值法的一种特殊类型。具体而言，我们假设底层分布 $F_j(y),\quad j\in \mathrm{J}$ 属于一类称为帕累托型分布的分布（Wang and Tsai 2009）。因此，$F_j(y),\quad j\in \mathrm{J}$ 可写为

$$F_j(y)=1-y^{-1/\gamma_j}\mathcal{L}_j(y),\quad y>0,\ j\in\mathcal{J}, \tag{4}$$

其中 $\gamma_j>0$ 是一个未知参数，称为极值指数（EVI）；$\mathcal{L}_j(y)$ 是一个缓变函数，也就是说，对于任意 $s>0$，$\mathcal{L}_j(ys)/\mathcal{L}_j(y)\to 1$ 当 $y\to\infty$ 时。如 Wang and Tsai（2009）的表1所示，帕累托型分布（4）涵盖许多重尾分布。在形式（4）下，右尾分布（2）可近似表示为（Wang and Tsai 2009）

$$F_j(y\mid\omega_j)\approx 1-\left(\frac{y}{\omega_j}\right)^{-1/\gamma_j},\quad y>\omega_j,\ j\in\mathcal{J}. \tag{5}$$

众所周知，近似式（5）与（3）密切相关，其中 $\gamma_j>0,\quad j\in \mathrm{J}$（de Haan and Ferreira 2006 的定理1.2.1）。在模型（5）中，若能估计参数 $\{\gamma_j\}_{j\in\mathcal{J}}$，则可检测目标 $F_j(y\mid\omega_j),\quad j\in \mathrm{J}$。

$\gamma_j>0$ 的一个广泛使用的估计量是Hill估计量（Hill 1975）：

$$\hat{\gamma}_j^{\mathrm{H}}:=\frac{\sum_{i=1}^{n_j}\log(Y_{ij}/\omega_j)I(Y_{ij}>\omega_j)}{\sum_{i=1}^{n_j}I(Y_{ij}>\omega_j)}, \tag{6}$$

其中，$I(\cdot)$ 是定义为如下形式的指示函数

$$I(Y_{ij}>\omega_j):=\begin{cases}0, & Y_{ij}\le\omega_j,\\ 1, & Y_{ij}>\omega_j.\end{cases}$$

经典估计量 $\{\hat{\gamma}_j^{\mathrm{H}}\}_{j\in\mathcal{J}}$，即分区域估计量，不利用区域之间的互信息。值得注意的是，每个 $\hat{\gamma}_j^{\mathrm{H}}$ 仅根据阈值超越量 $\{Y_{ij}:\ Y_{ij}>\omega_j,\quad i=1,\ 2,\ \ldots\ldots,\ n_j\}$ 构造。因此，在高阈值 ω_j 的设定下，其有效样本量通常较小。由此，极值分析中的区域级估计可能并不有效。为获得更有效的估计值，我们将区域 J 之间的关系纳入其中

到 (4) 中，并将所有区域 $\mathcal{J}$ 的信息纳入单一模型。在第2.2节中，我们采用混合效应模型构建了针对 (4) 的复杂模型。混合效应模型广泛用于小区域估计（Jiang and Nguyen 2007年；Wu 2009）。

2.2 混合效应模型的应用

混合效应模型通过称为随机效应的潜变量，而非参数，来捕捉不同区域的数据分布差异。令 $V_1, V_2, \ldots, V_J$ 表示随机效应，其边际分布为正态分布$V_j \sim N(0, \sigma^2)$，$j \in$ J，均值为零，未知方差为 $\sigma^2 > 0$。在混合效应模型中，对随机效应采用高斯分布是标准做法（Jiang and Nguyen 2007年；Wu 2009）。

我们假设极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$表示为

$$\gamma_j = \exp(\mu + V_j) > 0, \quad j \in \mathcal{J}, \tag{7}$$

其中，$\mu \in \mathbb{R}$ 是所有区域 J 共有的未知参数（注释1）。Momoki和Yoshida（2025年）也采用指数连接函数，以确保（7）右侧为正。与第2.1节所述的经典方法不同，我们对极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$中各区域 J 之间的关系进行建模。为实现这一改进，我们的随机效应 $\{V_j\}_{j\in\mathcal{J}}$可以相关，其形式如下：

$$D_{j_1 j_2} \coloneqq Cor[V_{j_1}, V_{j_2}] \in [-1, 1], \quad j_1, j_2 \in \mathcal{J}$$

因此满足

$$\boldsymbol{V} \coloneqq (V_1, V_2, \ldots, V_J)^\top \sim N_J(\boldsymbol{0}, \sigma^2 \boldsymbol{D}) \tag{8}$$

对于已知的正定相关矩阵 $\boldsymbol{D} := [D_{j_1 j_2}]_{j_1,\ j_2\in\mathcal{J}}$(注2)。注意，$\boldsymbol{D}$ 的对角元素设为 1。粗略地说，如果两个区域$j_1 \in \mathcal{J}$和 $j_2 \in \mathcal{J}$的随机效应具有较强的正相关性 $D_{j_1 j_2}$，那么它们的极值指数$\gamma_{j_1} =\exp(\mu + V_{j_1})$和$\gamma_{j_2} =\exp(\mu + V_{j_2})$很可能具有相似的取值。从这个意义上说，具有相关随机效应7的极值指数（）8能够反映区域$\mathcal{J}$之间的关系。

区域的边缘分布 $j \in \mathcal{J}$表示为 $F_j(y \mid v_j) := P(Y_{ij} < y \mid V_j = v_j)$。根据（4），$F_j(y \mid v_j)$可建模为

$$F_j(y \mid v_j) = 1 - y^{-1/\exp(\mu+v_j)} \mathcal{L}_j(y; v_j), \quad y > 0,\ j \in \mathcal{J}, \tag{9}$$

其中，$\exp(\mu + v_j)$表示式（7）中所述的极值指数，且 $\mathcal{L}_j(y;\ v_j)$是关于 y的缓变函数。如式（8）所述，我们在该模型中使用相关随机效应 $\{V_j\}_{j\in\mathcal{J}}$，用于表示式（9）（备注 3）中区域 $\mathcal{J}$ 之间的关系。

备注1：所提出的用于极值指数（7）模型可以假设其服从对数正态分布。众所周知，对数正态分布是不对称的，其众数为 $\exp(\mu - \sigma^2)$，中位数为 $\exp(\mu)$，均值为 $\exp(\mu + \sigma^2/2)$。因此，模型（7）倾向于支持较大的极值指数值。由于较大的极值指数表明基础分布具有更厚的右尾，因此，这种极值指数结构意味着一种保守的风险预测方法。因此，我们的模型（7）对于避免低估极端事件是合理的。

备注2：随机效应的相关矩阵 $\boldsymbol{D}$ 由随机效应$\{V_j\}_{j\in\mathcal{J}}$的 $J(J-1)/2$个分量决定。因此，如果 J 较大且 $\boldsymbol{D}$ 的所有分量均完全未知，那么需要估计的参数数量将大幅增加。对于模型（9），如此众多的 $\boldsymbol{D}$未知参数可能会导致其他参数 μ 和 σ^2的估计值不确定。为避免这一困难， $\boldsymbol{D}$ 的构造与(μ, σ^2)的估计过程分开。构造 $\boldsymbol{D}$ 的示例将在第5节中讨论。

备注3。通常，对于样本量较小的区域，区域级估计的性能较差（Diallo和Rao 2018）。小区域估计旨在改善样本量较小区域的区域级估计值。混合效应模型有助于实现这一目标（Sugasawa 和 Kubokawa 2020）。通常假设混合效应模型的随机效应在各区域之间相互独立（Nie 2007, Jiang 等人 2022）。这意味着随机效应的相关矩阵 $\boldsymbol{D}$ 固定为单位矩阵 $\boldsymbol{I}_J$。在 EVI 回归的背景下，Momoki和Yoshida（2025）证明了独立随机效应的有效性。与他们的研究不同，本研究的创新之处在于采用非单位矩阵 $\boldsymbol{D} \neq \boldsymbol{I}_J$。通过这一修改，我们的混合效应模型（9）能够纳入区域$\mathcal{J}$之间极值指数$\{\gamma_j\}_{j\in\mathcal{J}}$的相似性。从这个意义上说，所提出的模型（9）实现了区域$\mathcal{J}$ 之间的“借力”（Dempster等 1981）。第5 节和第6 节给出了一些构建示例

$\boldsymbol{D} \neq \boldsymbol{I}_J$ 并展示了所提出模型在这些示例中的性能验证。

2.3 估计与预测方法

未知参数 μ 和 σ^2 采用最大似然法进行估计。此外，为了刻画各区域 $\mathcal{J}$ 的极值指数$\{\gamma_j\}_{j\in\mathcal{J}}$之间的差异，我们还必须预测随机效应 $\{V_j\}_{j\in\mathcal{J}}$。随机效应$\{V_j\}_{j\in\mathcal{J}}$采用条件众数法进行预测。这些估计与预测方法是广义混合效应模型中的标准方法（Jiang and Nguyen 2007年第2章第3.6.2节；Wu 2009年第11章）。估计量和预测变量的详细定义见以下各节 2.3.1 和 2.3.2。我们可以使用 R语言环境中的 Template Model Builder（TMB）软件包实现这些方法（Kristensen et al. 2016）。

2.3.1 近似最大似然估计

假设在给定条件 $\boldsymbol{V} = (V_1, V_2, \ldots\ldots, V_J)^\top$ 下，数据 $\{Y_{ij},\ i = 1, 2, \ldots\ldots, n_j,\ j \in \mathrm{J}\}$ 在同一区域内以及不同区域 J 之间均相互独立（Jiang 等人 2022年）。根据混合效应模型似然函数的标准定义（Wu 2009年第2章），阈值超越量 $\{Y_{ij}\colon Y_{ij} > \omega_j,\ i = 1, 2, \ldots\ldots, n_j,\ j \in \mathrm{J}\}$ 的似然可定义为

$$E_{\boldsymbol{V}}\left[\prod_{j\in\mathcal{J}}\prod_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j}\frac{\partial}{\partial y}F_j(y \mid \omega_j, V_j)\mid_{y=Y_{ij}}\right],$$

这涉及对随机效应分布取期望，因为随机效应 $\{V_j\}_{j\in\mathcal{J}}$是潜变量，其中

$$\begin{aligned} F_j(y \mid \omega_j, v_j) &:= P(Y_{ij} \le y \mid Y_{ij} > \omega_j, V_j = v_j) \\ &= \frac{F_j(y \mid v_j) - F_j(\omega_j \mid v_j)}{1 - F_j(\omega_j \mid v_j)}, \quad y > \omega_j,\ j \in \mathcal{J}. \end{aligned}$$

与式（5）类似，在模型（9）和第 3 节的假设（A1）下，概率密度函数 $(\partial/\partial y)F_j(y \mid \omega_j, v_j)$可近似表示为

$$\begin{aligned} \frac{\partial}{\partial y}F_j(y \mid \omega_j, v_j) &\approx \mathrm{pareto}_{\omega_j}(y; \exp(\mu + v_j)) \\ &:= \frac{1}{\omega_j \exp(\mu + v_j)}\left(\frac{y}{\omega_j}\right)^{-1/\exp(\mu+v_j)-1}, \quad y > \omega_j,\ j \in \mathcal{J}, \end{aligned} \tag{10}$$

其中 $\text{pareto}^{\omega_j}(y;\exp(\mu+v_j))$表示$Y_{ij}/^{\omega_j}$的帕累托分布密度，其参数为 $\exp(\mu+v_j)$。Wang and Tsai（2009，第2.2节）也给出了与式（10）相同的近似。因此，用于估计未知参数 μ 和 σ^2 的似然函数可近似写为

$$\begin{aligned} L(\mu,\sigma^2) &:= E_{\boldsymbol{V}}\left[\prod_{j\in\mathcal{J}}\prod_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j}\text{pareto}_{\omega_j}(Y_{ij};\exp(\mu+v_j))\right] \\ &= \int_{\mathbb{R}^J}\phi_J(\boldsymbol{v};\mathbf{0},\sigma^2\boldsymbol{D})\prod_{j\in\mathcal{J}}\prod_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j}\text{pareto}_{\omega_j}(Y_{ij};\exp(\mu+v_j))d\boldsymbol{v}, \end{aligned} \tag{11}$$

其中 $\boldsymbol{v}:=(v_1,\ v_2,\ \ldots\ldots,\ v_J)^\top\in\mathbb{R}^J$，而 $\phi_J(\boldsymbol{v},\ 0,\ \sigma^2\boldsymbol{D})$ 是 $N_J(0,\ \sigma^2\boldsymbol{D})$ 的联合概率密度函数。根据上文给出的似然函数 $L(\mu,\ \sigma^2)$，我们可以推导出 μ 和 σ^2 的最大似然估计量。将这些估计量记作 $\hat{}\mu\in\mathbb{R}$ 和 $\hat{}\sigma^2>0$。

2.3.2 用条件众数法预测随机效应

根据 Wu（2009），在给定 $\{Y_{ij}:\ Y_{ij}>\omega_j,\ \ i=1,\ 2,\ \ldots\ldots,\ n_j,\ \ j\in\mathrm{J}\}$ 的条件下，$\boldsymbol{V}=(V_1,\ V_2,\ \ldots\ldots,\ V_J)^\top$ 的条件联合密度函数与以下内容成正比：

$$\phi_J(\boldsymbol{v};\mathbf{0},\sigma^2\boldsymbol{D})\prod_{j\in\mathcal{J}}\prod_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j}\frac{\partial}{\partial y}F_j(y\mid\omega_j,v_j)\mid_{y=Y_{ij}}, \tag{12}$$

它是以下变量的函数：$\boldsymbol{v}=(v_1,\ v_2,\ \ldots\ldots,\ v_J)^\top$。与第 2.3.1 节类似，我们再次用帕累托密度 $\text{pareto}^{\omega_j}(y;\ \exp(\mu+v_j))$近似（12）的经验版本如下：

$$\phi_J(\boldsymbol{v};\mathbf{0},\hat{\sigma}^2\boldsymbol{D})\prod_{j\in\mathcal{J}}\prod_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j}\text{pareto}_{\omega_j}(Y_{ij};\exp(\hat{\mu}+v_j)), \tag{13}$$

其中 $\hat{}\mu$ 和 $\hat{}\sigma^2$ 分别表示 μ 和σ^2 的最大似然估计量。$\boldsymbol{V}=(V_1,V_2,\ldots,V_J)^\top$ 的预测变量13定义为式（

$$\begin{aligned} \tilde{\boldsymbol{v}} &= (\tilde{v}_1,\tilde{v}_2,\ldots,\tilde{v}_J)^\top \\ &:= \underset{\boldsymbol{v}\in\mathbb{R}^J}{\text{argmax}}\ \phi_J(\boldsymbol{v};\mathbf{0},\hat{\sigma}^2\boldsymbol{D})\prod_{j\in\mathcal{J}}\prod_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j}\text{pareto}_{\omega_j}(Y_{ij};\exp(\hat{\mu}+v_j)) \end{aligned}$$

（备注 4）。特别地，极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$预测为

$$\tilde{\gamma}_j := \exp\left(\hat{\mu}+\tilde{v}_j\right),\quad j\in\mathcal{J}. \tag{14}$$

备注4。(13) 的对数变换为

$$\log \phi_J(\boldsymbol{v}; 0, \hat{\sigma}^2 \boldsymbol{D}) + \sum_{j \in \mathcal{J}} \sum_{i=1,2,\ldots,n_j : Y_{ij} > \omega_j} \log \text{pareto}_{\omega_j}(Y_{ij}; \exp(\hat{\mu} + v_j)). \tag{15}$$

为简单起见，我们考察式（15）的结构，其中 $J = 2$, $\boldsymbol{V} = (V_1, V_2)^\top$,$\boldsymbol{v} = (v_1, v_2)^\top$, $\tilde{\boldsymbol{v}} = (\tilde{v}_1, \tilde{v}_2)^\top$, 且 $Cor[V_1, V_2] = \rho$。式（15）的第一项，即 $\log \phi_J(\boldsymbol{v}; 0, \hat{\sigma}^2 \boldsymbol{D})$，通过以下方式依赖于 $\boldsymbol{v}$

$$-\frac{\boldsymbol{v}^\top \boldsymbol{D}^{-1} \boldsymbol{v}}{2\hat{\sigma}^2} = -\frac{(v_1 - v_2)^2}{2\hat{\sigma}^2 (1 - \rho^2)} - \frac{v_1 v_2}{\hat{\sigma}^2 (1 + \rho)}. \tag{16}$$

于是，我们有 $\log \phi_J(\boldsymbol{v}; 0, \hat{\sigma}^2 \boldsymbol{D}) \to -\infty$ as $\rho \to 1$，而式（15）的第二项不会随 ρ 直接变化。因此，在最大化式（15）时，当$\rho \approx 1$，$\log \phi_J(\boldsymbol{v}; 0, \hat{\sigma}^2 \boldsymbol{D})$ 的作用远大于式（15）的第二项。于是，v_1 和 v_2 应彼此接近，以增大 $\log \phi_J(\boldsymbol{v}; 0, \hat{\sigma}^2 \boldsymbol{D})$。因此，使用 $\rho = Cor[V_1, V_2] \approx 1$ 意味着 $\tilde{v}_1$ 和 $\tilde{v}_2$的取值相近。若 $\rho = 0$，则 $\tilde{v}_1$ 和 $\tilde{v}_2$ 独立地得到，尽管它们的大小受到 $\hat{\sigma}^2$ 的限制。因此，由于 $\boldsymbol{D}$的相关结构，$(v_1, \ldots, v_J)$ 会以相互依赖的方式进行优化。

3 渐近理论

3.1 条件

我们在条件$n_j \to \infty$, $j \in \mathrm{J}$ 和 $J \to \infty$下，为所提出的模型（9）建立渐近理论。因此，区域数量 J以及所有样本量 $\{n_j\}_{j \in \mathcal{J}}$均假定足够大。Jiang 等人（2022）也讨论了相同的条件。

对于每个区域 $j \in \mathrm{J}$，假设ω_j是一个n_j序列，使得 $\omega_j \to \infty$当 $n_j \to \infty$ 时成立；这在极值理论中是标准做法（Smith 1987）。我们记$k_j := \sum_{i=1}^{n_j} I(Y_{ij} > \omega_j)$, $j \in \mathrm{J}$。那么，k_j表示区域 $j \in \mathrm{J}$ 的有效样本量。各区域有效样本量的平均值 $\{k_j\}_{j \in \mathcal{J}}$记为 $k := J^{-1} \sum_{j \in \mathcal{J}} k_j$。我们的渐近理论要求以下条件（A1）–（A5）对所有 $v_j \in \mathbb{R}$和所有 $j \in \mathrm{J}$ 均一致成立：

（A1）缓变函数 $\mathcal{L}_j(y; v_j)$ 在（9）中属于 Hall 类（Hall，1982）。换言之，它可以表示为

$$\mathcal{L}_j(y; v_j) = a_j(v_j) + b_j(v_j) y^{-\beta_j(v_j)} + o\left(y^{-\beta_j(v_j)}\right),$$

其中 $a_j(\cdot) > 0$、 $b_j(\cdot)$ 和 $\beta_j(\cdot) > 0$ 均为连续有界函数。

(A2) 在给定 $V_j = v_j$、 $k_j^{-1} \xrightarrow{P} 0$ 的条件下，当 $n_j \to \infty$, $j \in \mathrm{J}$ 和 $J \to \infty$ 时成立，其中“ $\xrightarrow{P}$ ”表示依概率收敛。

(A3) 存在一个函数 $\theta_j : \mathbb{R}^J \to \mathbb{R}$ 以及一些常数 $0 < \tau_1 < \tau_2 < \infty$，使得在给定 $\boldsymbol{V} = \boldsymbol{v}$ 的条件下，当 $n_j \to \infty$, $j \in \mathrm{J}$ 和 $J \to \infty$ 时， $k_j/k \xrightarrow{P} \theta_j(\boldsymbol{v}) \in (\tau_1, \tau_2)$。

(A4) 当 $n_j \to \infty$, $j \in \mathrm{J}$ 和 $J \to \infty$ 时， $k/J \xrightarrow{P} 0$ 。

(A5) 存在一些常数 $h_\mu \in \mathbb{R}$ 和 $h_{\sigma^2} \in \mathbb{R}$，使得

$$\left(\mathbf{1}_J^\top \boldsymbol{D}^{-1} \mathbf{1}_J\right)^{-1/2} \mathbf{1}_J^\top \boldsymbol{D}^{-1} E\left[\left[\varepsilon_j(V_j)\right]_{j \in \mathcal{J}}\right] \to h_\mu$$

and

$$2J^{-1/2} E\left[\boldsymbol{V}^\top \boldsymbol{D}^{-1}\left[\varepsilon_j(V_j)\right]_{j \in \mathcal{J}}\right] \to h_{\sigma^2}$$

当 $n_j \to \infty$, $j \in \mathrm{J}$ 和 $J \to \infty$ 时，其中 $\mathbf{1}_J := (1, 1, \ldots, 1)^\top \in \mathbb{R}^J$，并且

$$\varepsilon_j(V_j) := \frac{b_j(V_j) \exp(\mu + V_j) \beta_j(V_j)}{1 + \exp(\mu + V_j) \beta_j(V_j)} \omega_j^{-1/\exp(\mu + V_j) - \beta_j(V_j)}, \quad j \in \mathcal{J}.$$

首先，(A1)使 $F_j(y \mid \omega_j, \ v_j)$向帕累托分布的收敛更加规律（de Haan和Ferreira（2006）的第 2.3 节）。这一条件称为二阶条件。其次，(A2)–(A4)是关于有效样本量$\{k_j\}_{j \in \mathcal{J}}$的假设，用于控制阈值$\{\omega_j\}_{j \in \mathcal{J}}$的发散速率。Momoki和Yoshida（2025）第 3.1 节详细解释了(A2)–(A4)各自的含义。(A5)说明了估计量的偏差渐近消除的速率，这出于技术原因是渐近理论所必需的。

3.2 渐近性质

对于所提出的估计量$\hat{\mu}$和$\hat{\sigma}^2$，我们得到以下定理1。

T定理 1。假设*(A1)–(A5)*成立。那么，当 $n_j \to \infty$, $j \in \mathcal{J}$ 和 $J \to \infty$ ，

$$\begin{bmatrix} \sqrt{\mathbf{1}_J^\top \boldsymbol{D}^{-1} \mathbf{1}_J}\,(\hat{\mu} - \mu) \\ \sqrt{J}\left(\hat{\sigma}^2 - \sigma^2\right) \end{bmatrix} - \begin{bmatrix} h_\mu \\ h_{\sigma^2} \end{bmatrix} \xrightarrow{D} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix}\right).$$

下面给出关于上述定理1 的一些说明。

(R1) Jiang 等人（2022）研究了广义混合效应模型中最大似然估计量的渐近正态性。Momoki和Yoshida（2025）将Jiang 等人（2022）的结果扩展到了极值理论背景下的帕累托型分布。尽管Jiang 等人（2022）和Momoki和Yoshida（2025）假设区域之间的随机效应相互独立，即 $\boldsymbol{D} = \boldsymbol{I}_J$，但本研究中的随机效应 $\{V_j\}_{j\in\mathcal{J}}$可以像（8）那样在区域之间存在相关性。因此，我们的定理 1 的新颖之处在于，它为可能不相互独立的随机效应建立了更一般的理论。

(R2) 根据定理1，$\hat{}\mu$ 是 $\mathbf{1}_J^\top \boldsymbol{D}^{-1}\mathbf{1}_J$相合的。若 $\boldsymbol{D} = \boldsymbol{I}_J$，则其收敛速率为 $O(\sqrt{J})$，对应于最慢的情形。然而，如果$\boldsymbol{D}$的所有元素均非零且彼此高度相关，那么收敛速率将达到最快，即 $O(J)$。因此，预计将更多区域纳入模型能够显著改善估计量。当$\boldsymbol{D} \neq \boldsymbol{I}_J$时，这一性质可能尤为明显。

(R3) 在定理 1 中，h_μ 和 h_{σ^2} 源于采用近似式（10）进行超阈值峰值法操作。这些量由二阶条件（A1）刻画。若 h_μ 和 h_{σ^2} 较大，则估计量 $\hat{}\mu$ 和 $\hat{}\sigma^2$ 在 $n_j \to \infty,\ \ j \in \mathrm{J}$ 和 $J \to \infty$ 的情况下可能仍存在偏差。若要校正这些估计量的偏差，我们尤其需要估计（A1）中出现的二阶参数（函数）$\{\beta_j(v_j)\}_{j\in\mathcal{J}}$。尽管已有若干研究（Gomes 等人（2002）；de Wet 等人（2012））探讨了二阶参数的估计，但在混合效应模型背景下估计该参数的方法尚未得到开发。然而，当 J 较大时，$\hat{}\mu$ 和 $\hat{}\sigma^2$的标准差仍保持稳定。从这个意义上说，重要的是，$\hat{}\mu$ 和 $\hat{}\sigma^2$ 的渐近速率由 J 主导，而非由有效样本量 $\{k_j\}_{j\in\mathcal{J}}$ 主导。因此，使用混合效应模型的一个优点是，可以通过设置较高的阈值 $\{\omega_j\}_{j\in\mathcal{J}}$ 来减轻所提出估计量的偏差。

4 模拟研究

通过模拟研究，我们证实了所提出的混合效应模型的优势（9）。我们展示了所提出的方法（14）与使用 Hill估计量的区域级估计方法（6）之间的比较结果。

表 1：使用我们的方法（14）和使用 Hill估计量（6）获得的预测结果。

	均方误差($\{\tilde{\gamma}_j\}_{j\in\mathcal{J}}$)			均方误差($\{\hat{\gamma}_j^{\mathrm{H}}\}_{j\in\mathcal{J}}$)
	情况1：（18）	情况1：（19）	情况2：$\boldsymbol{D}=\boldsymbol{I}_J$	
$n=50$	$1。06\times10^{-4}$	$1。01\times10^{-4}$	$27。47\times10^{-4}$	$31。53\times10^{-4}$
$n=200$	$3。91\times10^{-5}$	3.73×10^{-5}	$75。03\times10^{-5}$	77.82×10^{-5}

4.1 数据生成过程

在整个模拟研究中，我们将区域数量设为 $J=1000$。我们将极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$设为

$$\gamma_j = 2\left(\frac{j-1}{J-1}-\frac{1}{2}\right)^2+\frac{1}{5},\quad j\in\mathcal{J}. \tag{17}$$

使用这些极值指数，数据集 $\{Y_{ij},\ i=1,\ 2,\ \ldots\ldots,\ n_j,\ j\in\mathcal{J}\}$ 按照以下方式模拟为来自帕累托分布的独立同分布随机样本

$$P(Y_{ij}<y)=1-y^{-1/\gamma_j},\quad j\in\mathcal{J}.$$

为评估我们提出的方法和 Hill估计量的性能，我们生成了上述数据集的 $M=100$ 个重复样本，并对每个重复样本预测极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$。为简单起见，两种方法均采用固定阈值 $\omega_j=1,\ j\in\mathcal{J}$，即 $n_j=k_j,\ j\in\mathcal{J}$。在这一简单情形下，我们考虑了两种情况 $n:=n_1=n_2=\cdots=n_J=50$ 和 $n:=n_1=n_2=\cdots=n_J=200$。

4.2 结果

要使用我们提出的方法（14），我们必须首先确定$\boldsymbol{D}=[D_{j_1j_2}]_{j_1、j_2\in\mathcal{J}}$（第5节）。真实 EVI（17）在区域标签相似的区域中也相似。我们假设已知这些信息，然后采用

$$D_{j_1,j_2}=\exp\left(-\frac{|j_1-j_2|}{500}\right),\quad j_1,j_2\in\mathcal{J} \tag{18}$$

and

$$D_{j_1,j_2}=\exp\left(-\frac{|j_1-j_2|}{1000}\right),\quad j_1,j_2\in\mathcal{J}. \tag{19}$$

即使我们没有关于$\boldsymbol{D}$的信息，我们的方法对于 $\boldsymbol{D}=\boldsymbol{I}_J$仍然有用。因此，我们还考虑了 $\boldsymbol{D}=\boldsymbol{I}_J$。这三种设置（18）、（19）和 $\boldsymbol{D}=\boldsymbol{I}_J$ 分别被指定为“情况1”、“情况2”和“情形 3”。

表 1展示了使用我们的方法（14）和 Hill估计量（6）预测所得极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$的结果。在此表中，MSE($\{\tilde{\gamma}_j\}_{j\in\mathcal{J}}$)表示

我们方法（14）的均方误差（MSE），定义为

$$\mathrm{MSE}(\{\tilde{\gamma}_j\}_{j\in\mathcal{J}}) := \frac{1}{J\times M}\sum_{j=1}^{J}\sum_{m=1}^{M}\left(\tilde{\gamma}_j^{(m)}-\gamma_j\right)^2, \tag{20}$$

其中，$\{\tilde{\gamma}_j^{(m)}\}_{j\in\mathcal{J}}$ 是从第 m 个数据集中获得的预测变量。类似地，$\mathrm{MSE}(\{\hat{\gamma}_j^{\mathrm{H}}\}_{j\in\mathcal{J}})$ 表示 Hill估计量 (6) 的均方误差。如表1所示，对于 $\boldsymbol{D}$ 的所有设置，与使用 Hill估计量 (6) 得到的区域级估计的均方误差相比，我们的方法 (14) 的均方误差有所降低。尤其是，适当采用 $\boldsymbol{D}$，例如 (18) 或 (19)，能显著改善结果。因此，在下一节中，我们将深入讨论构造 $\boldsymbol{D}$ 的一些方法。即使在最简单的设置 $\boldsymbol{D}=\boldsymbol{I}_J$ 下，所提出的方法也略优于分区域估计方法。这一性能表明，$\boldsymbol{D}=\boldsymbol{I}_J$ 可作为 $\boldsymbol{D}$ 设置的候选方案。对于情况 $n=50$ 和 $n=200$，也得出了这些结论。

接下来，我们从偏差和标准误的角度考察表 1 中的结果。对于每种方法，我们都计算了各区域中预测的 EVI 的平均值以及第 5 和第 95 百分位数。图 1 展示了我们的方法在情况1（左上图）、情况2（右上图）和情形 3（左下图）下计算得到的平均值（实线）、百分位数（虚线）和真实 EVI（红线），以及区域级估计（右下图）；这些结果对应于 $n=50$。同样，图 2 展示了 $n=200$的结果。图 1 和 2 表明，除 $n=50$下的情形 3 外，预测变量的偏差都不大。在 $n=50$下的情形 3 中，预测变量在各区域均趋向于 $\hat{\mu}$，并呈现轻微偏差。不过，这一现象有助于提高预测的稳定性，因此其均方误差小于区域级估计的均方误差，如表 1所示。图 1 和 2 显示，不同方法的预测变量在标准误方面存在显著差异。特别是，显而易见，在情况1和情况2中纳入适当的 $\boldsymbol{D}$能够获得高度稳定的预测结果。

5 随机效应相关矩阵的设计

如第 2.3 节和 4.2节所述，合适的相关矩阵 $\boldsymbol{D}$ 随机效应的 $\{V_j\}_{j\in\mathcal{J}}$能显著提升我们方法的性能。

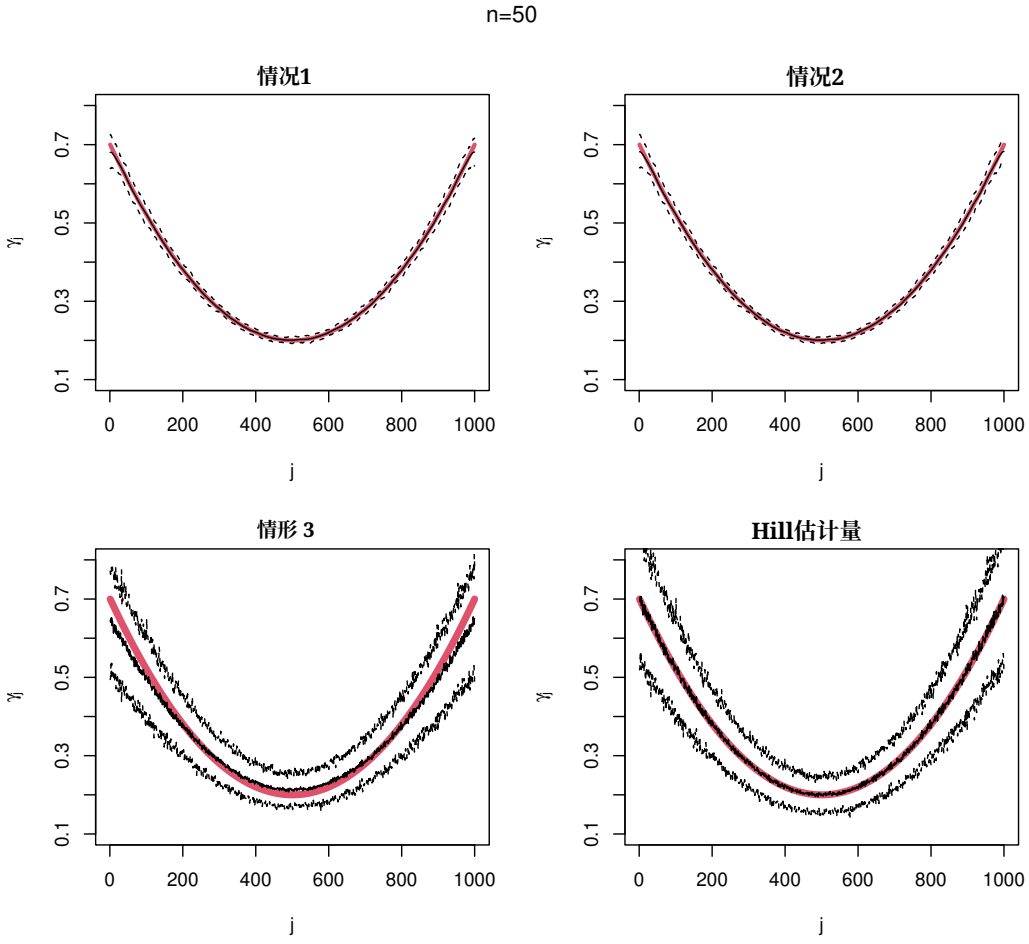


图 1：我们的方法在情况1（左上图）、情况2（右上图）和情形 3（左下图）下预测的 EVI 的平均值（实线）、第5和第95百分位数（虚线），以及真实 EVI（红线）；以及区域级估计（右下图），其中 $n = 50$

然而，随机效应 $\{V_j\}_{j\in\mathcal{J}}$是潜变量，并非可观测的数据。因此，我们无法直接根据定义（ $\boldsymbol{D}$ （8）进行评估。因此，我们必须从其他来源获得 $\boldsymbol{D}$，同时保留（8）的基本解释：较大的 $D_{j_1 j_2} = Cor[V_{j_1}, V_{j_2}]$， $j_1 \neq j_2 \in \mathcal{J}$表示极值指数$\gamma_{j_1} = \exp(\mu + V_{j_1})$和 $\gamma_{j_2} = \exp(\mu + V_{j_2})$的值较为接近。下面介绍三种构造 $\boldsymbol{D}$的方法。

(D1) 首先，我们考虑空间数据 $\{Y_{ij}$， $i = 1$、2、……、n_j， $j \in \mathrm{J}\}$。在这种情况下，区域 J 指的是地理位置，每个位置都具有纬度和经度等位置信息。对于此类空间数据，我们可以预期，彼此距离较近的区域，其极值指数 $\{\gamma_j\}_{j\in\mathcal{J}}$会较为相似。令 lon_j 和 lat_j 分别表示区域 $j \in \mathrm{J}$ 的经度和纬度。于是，我们可以设计出

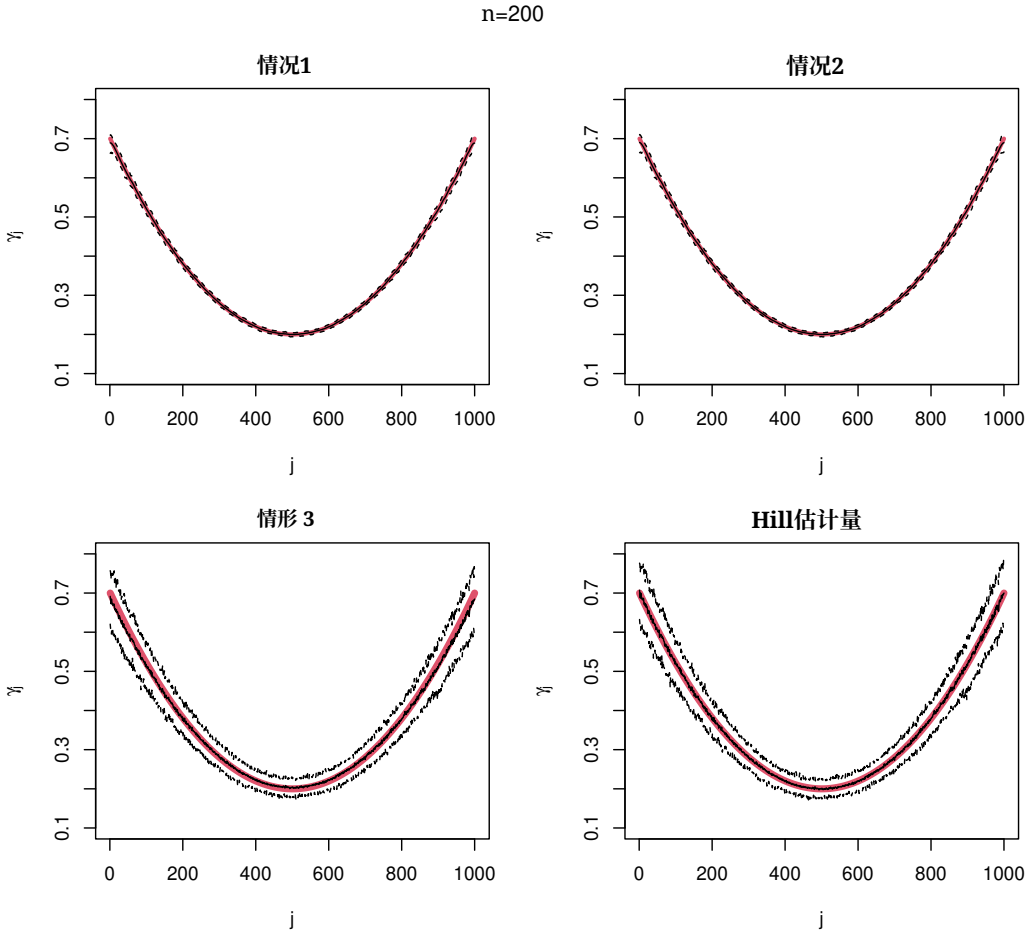


图 2：我们的方法在情况1（左上图）、情况2（右上图）和情形 3（左下图）下预测的 EVI 的平均值（实线）、第 5 和第 95 百分位数（虚线），以及真实 EVI（红线）；右下图为区域级估计，其中 $n = 200$

矩阵 $\boldsymbol{D} = [D_{j_1 j_2}]_{j_1,\ j_2 \in \mathcal{J}}$ as

$$D_{j_1 j_2} = \text{欧几里得距离}_{j_1 j_2} := \exp\left[-\frac{\text{dis}_{j_1 j_2}}{c}\right] \in (0, 1], \qquad j_1, j_2 \in \mathcal{J},$$

其中，$\text{dis}_{j_1 j_2}$ 是区域 $j_1 \in \text{J}$ 和 $j_2 \in \text{J}$ 之间的欧几里得距离，即 $\text{dis}_{j_1 j_2} := \sqrt{(\text{lon}_{j_1} - \text{lon}_{j_2})^2 + (\text{lat}_{j_1} - \text{lat}_{j_2})^2}$，而 $c > 0$ 是一个常数（Dyrrdal 等，2015）。

（D2）关于用于极值分析的空间贝叶斯聚类，Rohrbeck 和 Tawn（2019）认为“同一聚类中的站点往往比不同聚类中的站点表现出更高程度的依赖性”，并采用了尾部依赖（Reiss 和 Thomas，2007年）。我们还认为，具有较强尾部依赖的区域具有相似的极值指数。我们基于尾部依赖构造 $\boldsymbol{D}$

如下所示。我们将数据 $\{Y_{ij},\ \ i=1、2、.....、\ n,\ \ j\in \mathrm{J}\}$ 视为具有相同联合分布的独立同分布 J 维随机向量 $\{(Y_{i1},\ Y_{i2},\ .....、\ Y_{iJ})^{\top},\ \ i=1、2、.....、\ n\}$。对于每个区域 $j\in \mathrm{J}$，$\{Y_{ij}\}_{i=1}^{n_j}$ 的边缘分布函数记为 $F_j(y):=P(Y_{ij}\le y)$。随后，区域 $j_1\in \mathrm{J}$ 与 $j_2\in \mathrm{J}$ 之间的尾部依赖定义为 $\mathrm{TD}_{j_1j_2}:=\lim_{p\uparrow 1}\mathrm{TD}_{j_1j_2}(p)$，其中

$$\mathrm{TD}_{j_1j_2}(p):=P(Y_{ij_1}>F_{j_1}^{-1}(p)\mid Y_{ij_2}>F_{j_2}^{-1}(p)),\quad j_1,j_2\in\mathcal{J},$$

且 $F_j^{-1}(\cdot)$ 是 $F_j(\cdot)$ 的反函数。根据 Reiss and Thomas (2007) 的第2.6节，上述 $\mathrm{TD}_{j_1j_2}$ 是对称的，即 $\mathrm{TD}_{j_1j_2}=\mathrm{TD}_{j_2j_1}$。此外，我们有 $0\le \mathrm{TD}_{j_1j_2}\le 1$，当 $\mathrm{TD}_{j_1j_2}=0$ 时表示尾部独立，当 $\mathrm{TD}_{j_1j_2}=1$ 时表示完全尾部依赖。粗略地说，较强的尾部依赖 $\mathrm{TD}_{j_1j_2}$ 表明，区域 $j_1\in \mathrm{J}$ 和 $j_2\in \mathrm{J}$ 中的极端事件更有可能同时发生。$\mathrm{TD}_{j_1j_2}$ 的样本版本定义为

$$\widehat{\mathrm{TD}}_{j_1j_2}(p):=\frac{1}{n(1-p)}\sum_{i=1}^{n}I_*(Y_{ij_1}>q_{j_1}(p),Y_{ij_2}>q_{j_2}(p))$$

对于固定的 $p\approx 1$，其中 $q_j(p)$ 是数据 $\{Y_{ij}\}_{i=1}^{n_j}$ 的 $(100\times p)\%$ 经验分位数，而 $I_*(\cdot,\ \cdot)$ 是满足下述条件的指示函数

$$I_*(Y_{ij_1}>q_{j_1}(p),Y_{ij_2}>q_{j_2}(p))$$
$$:=\begin{cases}1, & Y_{ij_1}>q_{j_1}(p)\text{ and }Y_{ij_2}>q_{j_2}(p),\\ 0, & \text{otherwise.}\end{cases}$$

因此，我们使用 $\boldsymbol{D}=[\mathrm{TD}_{j_1j_2}(p)]_{j_1,\ j_2\in\mathcal{J}}$ 作为 $[\mathrm{TD}_{j_1j_2}]_{j_1,\ j_2\in\mathcal{J}}$ 的替代方案。后者不包含位置信息，因此既适用于空间数据，也适用于非空间数据。

(D3) 我们假设给定的区域构成若干组。如果区域 $j_1\in \mathrm{J}$ 和 $j_2\in \mathrm{J}$ 属于不同的组，那么它们之间的相关性 $D_{j_1,\ j_2}$ 可以视为零。在这种情况下，相关矩阵 $\boldsymbol{D}$ 被设计为带有零子矩阵的分块矩阵。可以预期，这种更简单的构造方式能提高我们的方法中的优化过程效率。即使我们对给定区域的分组信息并不明确，也可以通过将 $[D_{j_1j_2}]_{j_1,\ j_2\in\mathcal{J}}$ 中的小绝对值设为零，对 $\boldsymbol{D}$ 施加类似于 (D1) 和 (D2) 的分组结构。$\boldsymbol{D}$ 最简单的分块结构是单位矩阵 $\boldsymbol{I}_J$，这表明

数据在不同区域之间互不相关。然而，$\boldsymbol{D} = \boldsymbol{I}_J$ 即使我们没有关于 $\boldsymbol{D}$ 的信息，仍然可以使用。正如第 4.2 节所示，这种简单的设置可能也比区域级估计更有效。

6 日本极端降水的应用

我们通过一个使用日本降水数据集的示例，展示了所提出的混合效应模型（9）的应用。该数据集可从日本气象厅网站获取（https://www.data.jma.go.jp/gmd/risk/obsdl/index.php），其中包含 $J = 1198$ 个气象站（即区域）1980—2022年 $n = 365 \times 43 = 15695$ 日降水量（毫米）记录。该时段约为 $n = 365 \times 43 = 15695$ 天。然而，根据各气象站的启用和关闭日期，许多站点缺少连续日期的记录。我们将该数据集记为 $\{(Y_{i1}, Y_{i2}, \ldots, Y_{iJ})^\top,\ i = 1, 2, \ldots, n\}$，其中 Y_{ij} 表示区域 iJ中的第 $j \in$ 个日降水量，索引 i 表示所有区域 $\mathcal{J}$ 共有的日期。在第 6.1节中，我们首先进行了探索性数据分析，以确定该数据集的模型设置。第 6.2 节利用我们的混合效应模型，对日本强降雨进行风险评估。

6.1 初步分析

与其他极端降水空间分析研究类似，我们没有考虑极端降水的时间变化，从而避免使用更复杂的模型和得出不确定的结果（Ragulinaand Reitan 2017；Rohrbeck和Tawn（2021））。

我们首先对各区域 $j \in$ J 的数据拟合广义帕累托分布，形式如下

$$
\begin{aligned}
&P(Y_{ij} \le y \mid Y_{ij} > \omega_j^{\mathrm{GPD}}) \\
&= 1 - \left(1 + \gamma_j^{\mathrm{GPD}} \frac{y - \omega_j^{\mathrm{GPD}}}{\psi_j^{\mathrm{GPD}}}\right)_+^{-1/\gamma_j^{\mathrm{GPD}}}, \quad y > \omega_j^{\mathrm{GPD}},\ j \in \mathcal{J},
\end{aligned} \tag{21}
$$

其中，ω_j^{GPD}是给定的高阈值，γ_j^{GPD}和 ψ_j^{GPD}是未知参数。如果$\gamma_j^{\mathrm{GPD}} \le 0$，则模型中的假设 $\gamma_j > 0$ 可能不成立（参见 de Haan and Ferreira 2006 的第 1 章）。因此，我们首先从数据集中剔除极值指数为负的区域$\{\gamma_j^{\mathrm{GPD}}\}$。为此，我们对每个区域 $j \in$ J进行如下假设检验：

$$
\mathrm{H}_{0j} : \gamma_j^{\mathrm{GPD}} > 0 \quad \text{vs.} \quad \mathrm{H}_{1j} : \gamma_j^{\mathrm{GPD}} \le 0, \tag{22}
$$

其中，H0j表示原假设，H1j表示备择假设。令$\hat{\gamma}_j^{\mathrm{GPD}}$为$\gamma_j^{\mathrm{GPD}}$的最大似然估计量（参见 de Haan and Ferreira 2006 的第 3.4 节），其中，式（21）中的阈值ω_j^{GPD}是通过采用 Wang and Tsai（2009）提出的差异度量的广义帕累托分布版本选取的。根据 Einmahl 等人（2019）第 3.1 节给出的检验统计量，以及 de Haan and Ferreira（2006）第 3.4 节所示的$\hat{\gamma}_j^{\mathrm{GPD}}$的渐近正态性，如果（$k_j^{\mathrm{GPD}}$）$^{1/2}\hat{\gamma}_j^{\mathrm{GPD}} \le z_\alpha$，则可以拒绝原假设 H0$j$，其中$k_j^{\mathrm{GPD}} := \sum_{i=1}^{n_j} I$（$Y_{ij} > \omega_j^{\mathrm{GPD}}$），$z_\alpha$是$N$（0, 1）的第$100 \times \alpha$百分位点，$\alpha$是给定的显著性水平。我们对每个区域$j \in \mathcal{J}$采用$\alpha = 0.05$进行上述假设检验（22）。根据检验结果，60个区域的原假设 H0j被拒绝。因此，在后续使用我们提出的模型进行的分析中，省略了这60个根据上述假设检验推断为极值指数为负的区域，其中J由$J = 1198$改为$J = 1138$。图3展示了这些区域的位置分布图$J = 1138$。

为实现我们的模型（9），接下来我们构造了随机效应的相关矩阵$\boldsymbol{D} \in \mathbb{R}^{J \times J}$。现在，我们可以使用第 5 节中描述的两种方法（D1）和（D2）。图 4 展示了所有$j_1 > j_2 \in \mathcal{J}$对的欧几里得距离 ED$_{j_1 j_2}$与尾部依赖 TD$_{j_1 j_2}$(0.95) 的散点图。如图 c4所示，$\boldsymbol{D} = [\mathrm{ED}_{j_1 j_2}]_{j_1, j_2 \in \mathcal{J}}$和$\boldsymbol{D} = [\mathrm{TD}_{j_1 j_2}(0.95)]_{j_1, j_2 \in \mathcal{J}}$的构成相似，尤其是接近 1 的元素。因此，我们仅报告 c $\boldsymbol{D} = [$TD$_{j_1 j_2}$(0.95)$]_{j_1, j_2 \in \mathcal{J}}$的结果，即（D2）。实际上，在本应用中，使用（D1）得到的分析结果与使用（D2）相比并无显著差异。

6.2 混合效应模型分析

为明确模型（9）的未知结构，我们首先获得了估计值$\hat{\mu}$和$\hat{\sigma}^2$，以及预测变量$\tilde{\boldsymbol{v}} = (\tilde{v}_1, \tilde{v}_2, \ldots\ldots, \tilde{v}_J)^\top \in \mathbb{R}^J$对于阈值$\{\omega_j\}_{j \in \mathcal{J}}$，这些阈值是采用与第 6.1节所述方法类似的方法选取的。估计值为

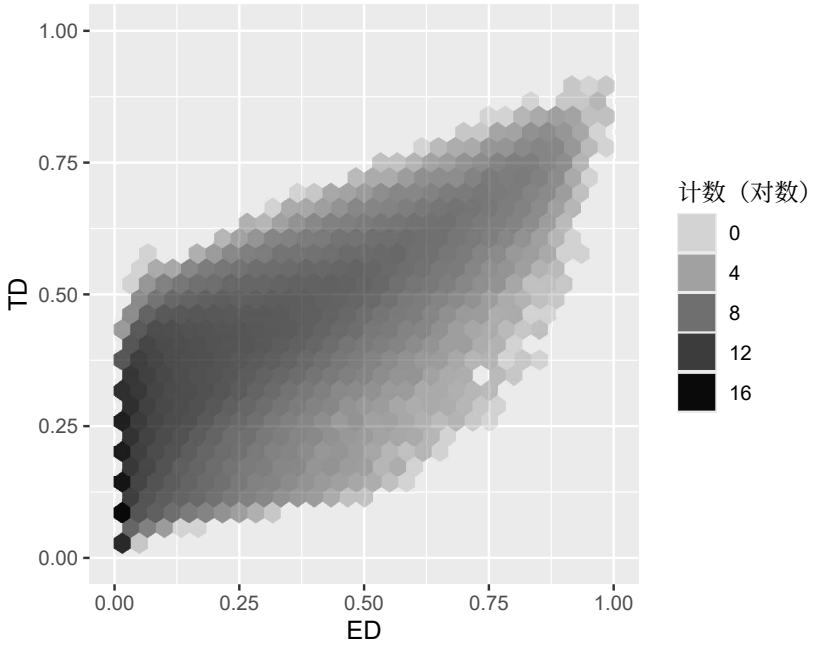
$$\hat{\mu} = -1.1642\ (\pm 0.1023) \quad \text{and} \quad \hat{\sigma}^2 = 0.0284\ (\pm 0.0023),$$

括号中的数值表示根据定理 1中的渐近正态性推导出的 95%置信区间的宽度。图 5 展示了所有区域的预测 EVI 热力图，即$\tilde{\gamma}_j = \exp(\hat{\mu} + \tilde{v}_j)$，$j \in \mathcal{J}$。在这张图中，区域边界是使用 deldir(https://CRAN.R-project.



图3：日本 $J = 1138$ 气象站（区域）的位置

org/package=deldir) 包在 R语言中构建的。白色区域表示原假设 $H0_j$ 在第 6.1 节中被拒绝。同样，图 6 展示了使用 Hill估计量进行区域级估计（第 2.1 节）得到的预测 EVI 热力图 $\{\hat{\gamma}_j^H\}$。图 5 和6 表明，与使用 Hill估计量进行区域级估计相比，我们提出的方法能够更清晰地捕捉极值指数的空间变化。提出的方法之所以能取得这一自然的结果，是因为它通过尾部依赖纳入了区域之间关系的信息。如图 5所示，部分区域的预测 EVI 存在局部变化。图 7 突出显示了图 5 中的区域 A 和 B 及其海拔，其中各点表示区域的位置及其预测 EVI。区域 A 包含名为日本阿尔卑斯山脉的山脉，而区域 B 则包含日本最大的湖泊琵琶湖。从图 7可以明显看出，在区域 A，日本阿尔卑斯山脉东侧的极值指数较大值多于西侧。在区域 B，周边的极值指数较低



c 图4：所有 $j_1 > j_2 \in \mathrm{J}$ 的成对组合中，欧几里得距离 ED $_{j_1 j_2}$与尾部依赖 TD$_{j_1 j_2}(0.95)$ 的散点图。每个六边形的颜色表示其中的数据点数量

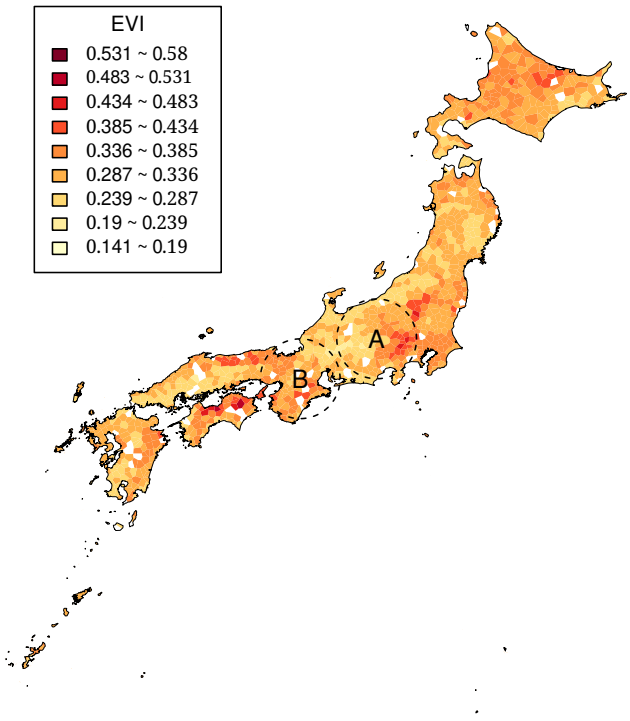


图5：使用所提出的方法得到的预测的 EVI热力图 $\{\tilde{\gamma}_j\}_{j \in \mathcal{J}}$

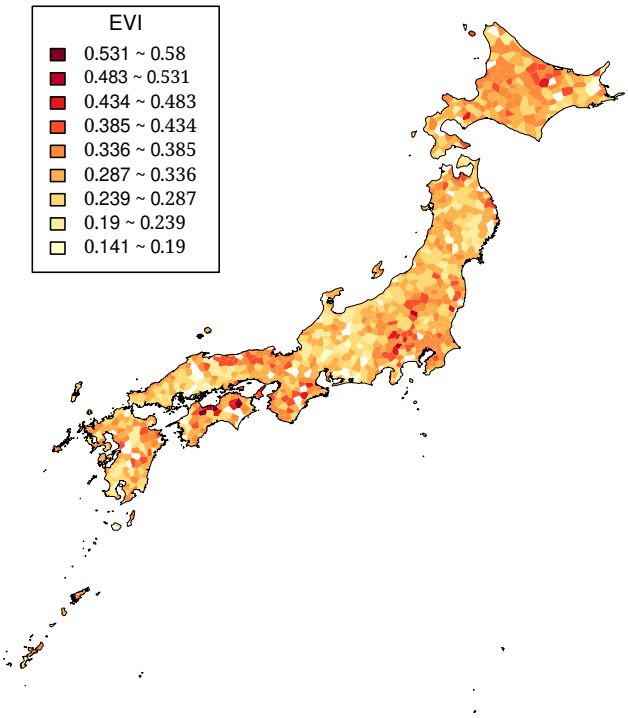


图6：使用 Hill估计量得到的预测的 EVI $\{\hat{\gamma}_j^{\text{H}}\}_{j\in\mathcal{J}}$

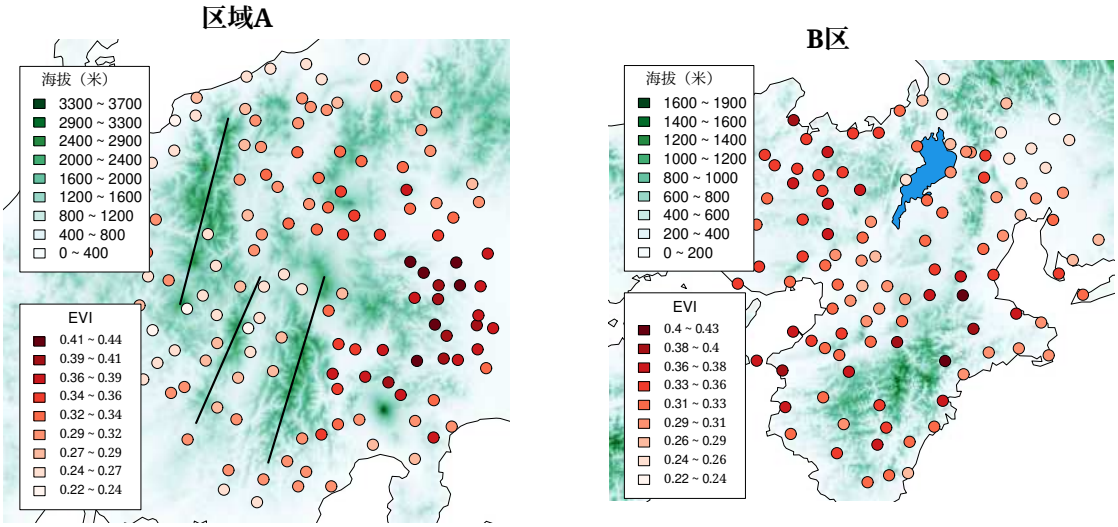


图7：图5中区域 A 和 B 的地形图：点表示区域的位置和预测的 EVI，三条直线表示日本阿尔卑斯山脉；蓝色表示琵琶湖

琵琶湖以及南部山脉的高处。

其次，我们预测了每个区域的极端降雨高分位数。根据（9）和（10），可以评估每个区域的极端降雨频率，表示为

$$\begin{aligned} & P(Y_{ij} > y \mid V_j = \tilde{v}_j) \\ & \approx P(Y_{ij} > \omega_j \mid V_j = \tilde{v}_j) \left(\frac{y}{\omega_j}\right)^{-1/\tilde{\gamma}_j} \\ & \approx \frac{k_j}{n_j} \left(\frac{y}{\omega_j}\right)^{-1/\tilde{\gamma}_j}, \quad j \in \mathcal{J}. \end{aligned} \tag{23}$$

对于每个区域 $j \in \mathcal{J}$，令 $z_j^{(R)}$表示第 R个区域的 j年重现水平，其定义为在特定年份以概率 R^{-1} 超过的高分位数（Cooley 等（2007））。通过求解方程P（$Y_{ij} > z_j^{(R)} \mid V_j = \tilde{v}_j$） $=$（$365R$）$^{-1}$，我们可以得到 $z_j^{(R)}\}_{j\in\mathcal{J}}$。对于每个 $j \in \mathcal{J}$，利用近似式（23），$\{z_j^{(R)}\}_{j\in\mathcal{J}}$ 可预测为

$$\tilde{z}_j^{(R)} := \omega_j \left(\frac{365Rk_j}{n_j}\right)^{\tilde{\gamma}_j}, \quad j \in \mathcal{J}$$

（de Haan 和 Ferreira（2006）第4章）。图 8 展示了所有区域预测的50年重现水平$\{\tilde{z}_j^{(50)}\}_{j\in\mathcal{J}}$的热力图。从图 8可以明显看出，太平洋一侧（即地图的西南部）的重现水平往往较高，这是因为台风经常影响这些区域（Troˇselj 和 Lee（2021））。图 9展示了图 8 中C区的地形图及其预测的50年重现水平，具体信息见图注。在C区，由于四国山脉地势陡峭，来自南方的温暖湿润空气形成上升气流，使太平洋一侧的降水增多，正如日本气象厅网站所述。然而，在这些山脉以北的地区，强降雨发生的可能性较低。图 9 中重现水平的巨大差异反映了这种天气状况。从图 9可知，所有区域中最高的50年重现水平为1271毫米。根据该数据集，具有这一最高重现水平的区域在1980—2022年间经历的最大日总降水量为764毫米。因此，在此类区域，我们应警惕前所未有的更强降雨发生。

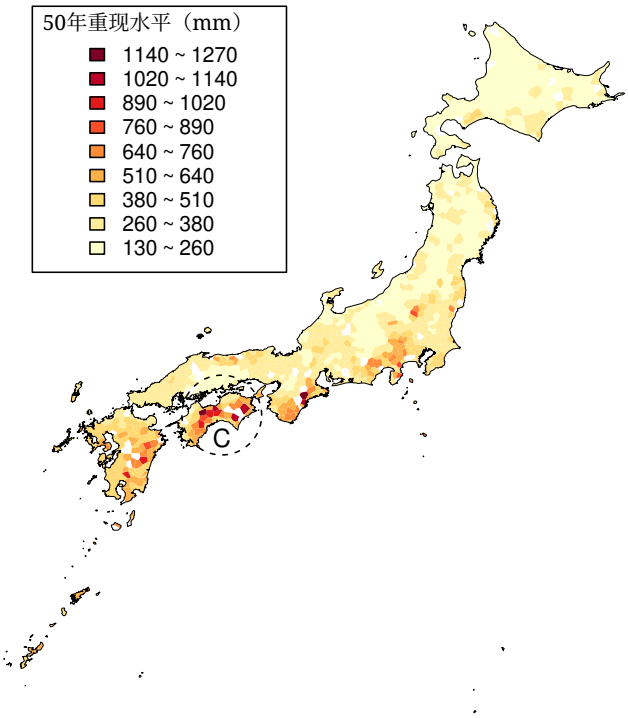


图8：预测的50年重现水平热力图 $\{\tilde{z}_j^{(50)}\}_{j\in\mathcal{J}}$

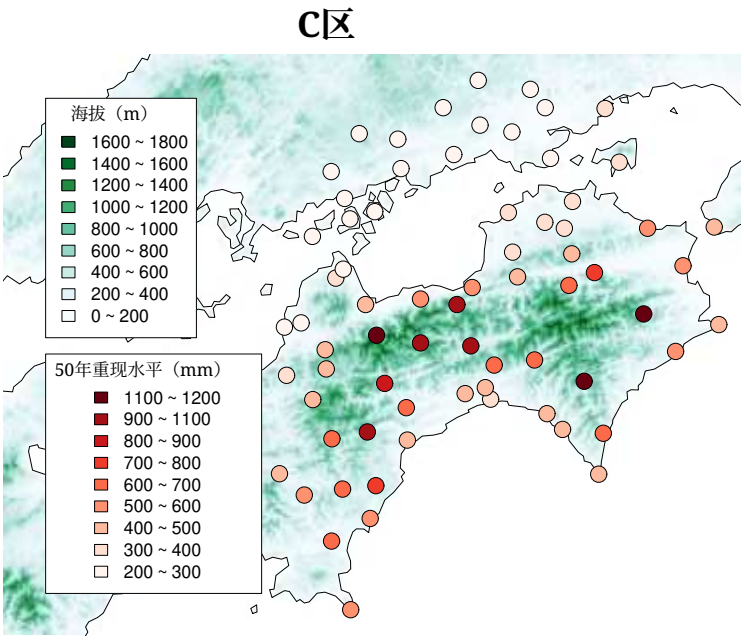


图9：图8中C区的地形图：每个点显示区域的位置及其预测的50年重现水平

7 讨论

对于来自多个区域的数据进行极值分析，我们研究了基于混合效应模型的新超阈值法。由于极值分析中各区域的有效样本量通常较小，Hill估计量等经典估计方法往往会产生不确定的结果。因此，本研究利用了区域之间关系的信息。我们采用带有相关随机效应的混合效应模型来实现这一模型。根据第3节中的渐近理论，混合效应模型支持我们利用更丰富的区域信息这一设想。事实上，随着区域数量 J的增加，我们模型中估计量的性能也会提升（第3节中的R2）。因此，正如第6节的应用所示，我们的模型适用于分析包含众多区域的大型数据。此外，第4.2 节所述的数值实验表明，在极值分析中，混合效应模型中适当相关的随机效应能够显著改善区域级估计。第6节的应用中，与分区域估计方法相比，我们的方法给出了更自然的分析结果（图5 和6）。第5节讨论了构建随机效应相关矩阵的一些示例。

本研究针对帕累托型分布提出了采用混合效应模型的超阈值法。更一般地，该方法也可以针对2.1节所述的广义帕累托分布进行拓展。然而，这种拓展面临以下挑战。首先，由于超阈值法的特性，广义帕累托分布中的尺度参数通常与其极值指数和阈值相关（de Haan and Ferreira（2006）的定理1.2.5）。例如，若假设极值指数服从混合效应模型，那么尺度参数也会受到随机效应的影响。因此，针对广义帕累托分布建立混合效应模型可能比针对帕累托型模型更加复杂。其次，在广义帕累托分布的混合效应模型中，极值指数可以是任意实数，但必须大于 $-1/2$，以确保稳定性。另一方面，尺度参数的行为取决于极值指数的符号（de Haan and Ferreira（2006）的定理1.2.5）。因此，当用广义帕累托分布的密度替换帕累托型密度，代入定义为（11）的似然函数时，随机效应的积分在极值指数等于零的点处会出现不连续。这种不连续可能会使最大化对数似然的优化过程变得复杂

。如果所有区域的极值指数符号相同（均为正或均为负），则或许可以避免这些困难。本文给出的结果对应于所有极值指数均为正的情况。若所有区域的极值指数均为负，则我们的方法可以拓展到广义帕累托分布的情形。例如，可以将极值指数建模为 $\gamma_j = -2^{-1}(1+\exp[\mu+V_j])^{-1}$， $j \in \mathrm{J}$，使得 $-1/2 < \gamma_j < 0$。这将留待未来作为一项重要研究课题。

附录

本附录给出了证明定理 1 所需的关键结果。令

$$l(\boldsymbol{v};\mu) := (Jk)^{-1} \sum_{j\in\mathcal{J}} \sum_{i=1,2,\ldots,n_j:Y_{ij}>\omega_j} \left[\mu + v_j + \left\{\frac{1}{\exp(\mu+v_j)} + 1\right\} \log \frac{Y_{ij}}{\omega_j}\right].$$

我们将 ¤$\boldsymbol{v} = ($ ¤$_1$、¤$_2$、……、¤$_J)^\top := \mathrm{argmin}_{\boldsymbol{v}\in\mathbb{R}^J} l(\boldsymbol{v};\ \mu)$ 记为。在给定 $\boldsymbol{V} = \boldsymbol{v}$ 的条件下，类似于 Momoki 和 Yoshida （2025）中的（35），我们有

$$\dot{v}_j = v_j + J\frac{k}{k_j}\frac{\partial}{\partial v_j} l(\boldsymbol{v};\mu)\,(1+o_P(1)), \quad j \in \mathcal{J}. \tag{24}$$

似然函数 $L(\mu、\ \sigma^2)$ 满足

$$\begin{aligned}&\frac{\partial}{\partial\mu}\log L(\mu,\sigma^2)\\ &= \frac{\int_{\mathbb{R}^J} \frac{\partial}{\partial\mu}\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D})\exp\left[-(Jk)\,l(\boldsymbol{v}-\mu\mathbf{1}_J;\mu)\right]d\boldsymbol{v}}{\int_{\mathbb{R}^J}\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D})\exp\left[-(Jk)\,l(\boldsymbol{v}-\mu\mathbf{1}_J;\mu)\right]d\boldsymbol{v}}\end{aligned} \tag{25}$$

and

$$\begin{aligned}&\frac{\partial}{\partial\sigma^2}\log L(\mu,\sigma^2)\\ &= \frac{\int_{\mathbb{R}^J} \frac{\partial}{\partial(\sigma^2)}\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D})\exp\left[-(Jk)\,l(\boldsymbol{v}-\mu\mathbf{1}_J;\mu)\right]d\boldsymbol{v}}{\int_{\mathbb{R}^J}\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D})\exp\left[-(Jk)\,l(\boldsymbol{v}-\mu\mathbf{1}_J;\mu)\right]d\boldsymbol{v}},\end{aligned} \tag{26}$$

其中我们有

$$\frac{\partial}{\partial\mu}\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D}) = \mathbf{1}_J^\top\left(\sigma^2\boldsymbol{D}\right)^{-1}(\boldsymbol{v}-\mu\mathbf{1}_J)\,\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D}) \tag{27}$$

and

$$\begin{aligned}&\frac{\partial}{\partial\left(\sigma^2\right)}\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D})\\ &= \left(\sigma^2\right)^{-1}\left[\frac{(\boldsymbol{v}-\mu\mathbf{1}_J)^\top\left(\sigma^2\boldsymbol{D}\right)^{-1}(\boldsymbol{v}-\mu\mathbf{1}_J)-1}{2}\right]\phi_J(\boldsymbol{v};\mu\mathbf{1}_J,\sigma^2\boldsymbol{D}).\end{aligned} \tag{28}$$

我们将（24）作为

$$\begin{aligned}
&\left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \{(\dot{\boldsymbol{v}} + \mu 1) - \mu 1_J\} \\
&= \left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \boldsymbol{V} \\
&\quad + \left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \left[J \frac{k}{k_j} \frac{\partial}{\partial v_j} l(\boldsymbol{V}; \mu)\right]_{j \in \mathcal{J}} (1 + o_P(1)) .
\end{aligned} \tag{29}$$

由（8），（29）右侧的第一项依分布收敛于 N（0，（σ^2）$^{-1}$），当 $J \to \infty$时。（29）右侧的第二项满足

$$\begin{aligned}
&\left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \left[J \frac{k}{k_j} \frac{\partial}{\partial v_j} l(\boldsymbol{V}; \mu)\right]_{j \in \mathcal{J}} \\
&\quad - \left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \left[E\left[\varepsilon_j(V_j)\right]\right]_{j \in \mathcal{J}} \\
&= k^{-1/2} \left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \\
&\quad \times \left[\left(\frac{k_j}{k}\right)^{-1/2} \left\{J k k_j^{-1/2} \frac{\partial}{\partial v_j} l(\boldsymbol{V}; \mu) - k_j^{1/2} \varepsilon_j(V_j)\right\}\right]_{j \in \mathcal{J}} \\
&\quad + \left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \left[\varepsilon_j(V_j) - E\left[\varepsilon_j(V_j)\right]\right]_{j \in \mathcal{J}} ,
\end{aligned} \tag{30}$$

其中，（30）右侧当 $n_j \to \infty$ 时依概率收敛于 0，$j \in \mathcal{J}$和 $J \to \infty$（Momoki 和 Yoshida 2025 的引理 2）。因此，我们得到

$$\begin{aligned}
&\left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \{(\dot{\boldsymbol{v}} + \mu 1) - \mu 1_J\} \\
&\quad - \left(1_J^\top \boldsymbol{D}^{-1} 1_J\right)^{-1/2} 1_J^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \left[E\left[\varepsilon_j(V_j)\right]\right]_{j \in \mathcal{J}} \\
&\quad \xrightarrow{D} N\left(0, \left(\sigma^2\right)^{-1}\right)
\end{aligned} \tag{31}$$

当 $n_j \to \infty$ 时，$j \in \mathcal{J}$ 和 $J \to \infty$。此外，由（28），可写为

$$\begin{aligned}
&J^{-1/2} \left(\sigma^2\right)^{-1} \left[\frac{\{(\dot{\boldsymbol{v}} + \mu 1) - \mu 1_J\}^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \{(\dot{\boldsymbol{v}} + \mu 1) - \mu 1_J\} - 1}{2}\right] \\
&= 2^{-1/2} \left(\sigma^2\right)^{-1} \left[\frac{\boldsymbol{V}^\top \left(\sigma^2 \boldsymbol{D}\right)^{-1} \boldsymbol{V} - 1}{(2J)^{1/2}}\right] \\
&\quad + J^{-1/2} \left(\sigma^2\right)^{-2} \boldsymbol{V}^\top \boldsymbol{D}^{-1} \left[J \frac{k}{k_j} \frac{\partial}{\partial v_j} l(\boldsymbol{V}; \mu)\right]_{j \in \mathcal{J}} (1 + o_P(1)) .
\end{aligned} \tag{32}$$

由于随机变量 $\boldsymbol{V}^\top (\sigma^2 \boldsymbol{D})^{-1} \boldsymbol{V}$服从卡方分布 χ_J，其自由度为 J，因此（32）右侧的第一项收敛

至 $N(0, 2^{-1}(\sigma^2)^{-2})$的分布，当 $J \to \infty$时。与（29）右侧的第二项类似，（32）右侧的第二项渐近等价于 $J^{-1/2}(\sigma^2)^{-2}E[\boldsymbol{V}^\top \boldsymbol{D}^{-1}[\varepsilon_j(V_j)]_{j\in\mathcal{J}}]$当 $n_j \to \infty$, $j \in$J 和 $J \to \infty$ 时。因此，我们得到

$$\begin{aligned} & J^{-1/2}\left(\sigma^2\right)^{-1}\left[\frac{\{(\dot{\boldsymbol{v}}+\mu\mathbf{1})-\mu\mathbf{1}_J\}^\top\left(\sigma^2\boldsymbol{D}\right)^{-1}\{(\dot{\boldsymbol{v}}+\mu\mathbf{1})-\mu\mathbf{1}_J\}-1}{2}\right] \\ & -J^{-1/2}\left(\sigma^2\right)^{-2}E\left[\boldsymbol{V}^\top\boldsymbol{D}^{-1}\left[\varepsilon_j(V_j)\right]_{j\in\mathcal{J}}\right] \\ & \xrightarrow{D} N\left(0, 2^{-1}\left(\sigma^2\right)^{-2}\right) \end{aligned} \tag{33}$$

当 $n_j \to \infty$、 $j \in \mathcal{J}$ 和 $J \to \infty$。我们对（25）和（26）应用 Miyata（2004）附录A中的拉普拉斯近似。利用（31）和（33），进而得到

$$\begin{aligned} & \left(\mathbf{1}_J^\top\boldsymbol{D}^{-1}\mathbf{1}_J\right)^{-1/2}\frac{\partial}{\partial\mu}\log L(\mu,\sigma^2) \\ & -\left(\mathbf{1}_J^\top\boldsymbol{D}^{-1}\mathbf{1}_J\right)^{-1/2}\mathbf{1}_J^\top\left(\sigma^2\boldsymbol{D}\right)^{-1}\left[E\left[\varepsilon_j(V_j)\right]\right]_{j\in\mathcal{J}} \\ & \xrightarrow{D} N\left(0,\left(\sigma^2\right)^{-1}\right) \end{aligned} \tag{34}$$

and

$$\begin{aligned} & J^{-1/2}\frac{\partial}{\partial\left(\sigma^2\right)}\log L(\mu,\sigma^2)-J^{-1/2}\left(\sigma^2\right)^{-2}E\left[\boldsymbol{V}^\top\boldsymbol{D}^{-1}\left[\varepsilon_j(V_j)\right]_{j\in\mathcal{J}}\right] \\ & \xrightarrow{D} N\left(0,2^{-1}\left(\sigma^2\right)^{-2}\right) \end{aligned} \tag{35}$$

当 $n_j \to \infty$、 $j \in \mathcal{J}$ 和 $J \to \infty$。定理1 的其余证明与 Momoki和Yoshida（2025）给出的定理1证明类似。

声明

竞争利益

作者没有需要披露的相关财务或非财务利益。

资金支持

本研究得到 JSPS KAKENHI（项目编号：22K11935 和 23K28043）的资助。

作者贡献

所有作者均参与了研究构思与设计。研究方法、调查和数据分析主要由 Koki Momoki 完成。稿件

主要由 Koki Momoki 撰写。所有作者均审阅了稿件，并批准了稿件的最终版本。

致谢

我们谨此感谢 FASTEK（https://www.fastekjapan.com/）为本文提供英文润色。本研究获得 JSPS KAKENHI（项目编号：22K11935 和 23K28043）的经费支持。

数据可用性

本研究分析的数据集由日本气象厅网站（https://www.data.jma.go.jp/gmd/risk/obsdl/index.php）获取的数据处理而成。经合理请求，可向通讯作者获取该数据集。模拟研究所用的 R语言 （https://www.r-project.org/）代码也可经合理请求向通讯作者获取。

遵守伦理标准

潜在利益冲突披露 不适用。**涉及人类参与者和/或动物的研究** 不适用。**知情同意** 不适用。

参考文献

[1] Beirlant, J.、Joossens, E. 和 Segers, J.：重尾分布的二阶精细化超阈值峰值建模统计规划与推断杂志 **139**（8），2800-2815（2009）。doi:10.1016/j.jspi.2009.01.006

[2] Cooley, D.、Nychka, D. 和 Naveau, P.：极端降水重现水平的贝叶斯空间建模。美国统计学会会刊**102**（479），824-840（2007年）。http://www.jstor.org/stable/27639928

[3] de Haan, L. 和 Ferreira, A.：极值理论：导论。Springer纽约，（2006）。doi:10.1007/0-387-34471-3

[4] Dekkers, A. L. M., Einmahl, J. H. J. 和 de Haan, L.：极值分布指数的矩估计量。统计学年刊 **17**（4），1833-1855（1989）。10.1214/aos/1176347397

[5] Dempster, A. P.、Rubin, D. B. 和 Tsutakawa, R. K.：协方差分量模型中的估计。美国统计学会会刊**76**（374），341-353（1981）。
doi:10.1080/01621459.1981.10477653

[6] de Wet, T.、Goegebeur, Y. 和 Guillou, A.：二阶尺度参数的加权矩估计量。 应用概率的方法与计算**14**，753-783（2012）。
doi:10.1007/s11009-011-9263-6

[7] Diallo, M. 和 Rao, J. N. K.：单位层级模型下带偏斜正态误差的复杂参数小区域估计。 斯堪的纳维亚统计学杂志**45**（4），1092-1116（2018）。
doi:10.1111/sjos.12336

[8] Drees, H.、Ferreira, A. 和 de Haan, L.：极值指数的最大似然估计。应用概率年刊**14**（3），1179-1201（2004）。
doi:10.1214/105051604000000279

[9] Einmahl, J. J.、Einmahl, J. H. J. 和 de Haan, L.：通过极值理论研究人类寿命的极限。 美国统计学会会刊**114**（527），1075-1080（2019）。
doi:10.1080/01621459.2018.1537912

[10] Girard, S., Stupfler, G. 和 Usseglio-Carleve, A.：重尾异方差回归模型中的极端条件期望分位数估计。统计学年刊**49**(6), 3358-3382 (2021)。
doi:10.1214/21-AOS2087

[11] Gomes, M. I., Alves, M. I. F. 和 Santos, P.A.：重尾模型的 PORT Hill 估计量和矩估计量。统计学通讯——模拟与计算**37**(7), 1281-1306 (2008a)。
doi:10.1080/03610910802050910

[12] Gomes, M. I., de Haan, L. 和 Peng, L.：极值统计中二阶参数的半参数估计。极值**5**, 387-414 (2002)。doi:10.1023/A:1025128326588

[13] Gomes, M. I., de Haan, L. 和 Rodrigues, L. H.：重尾模型的尾指数估计：加权对数超额中的偏差调整。英国皇家统计学会会刊*B*辑：统计方法 **70**(1), 31-52 (2008b)。doi:10.1111/j.1467-9868.2007.00620.x

[14] Gomes, M. I. 和 Guillou, A.：单变量极值的极值理论与统计：综述。*International StatisticalReview* **83**，263-292（2015年）。
doi:10.1111/insr.12058

[15] Hall, P.：正则变异指数的若干简单估计。皇家统计学会会刊：B辑（方法论）44(1)，37-42（1982年）。
doi:10.1111/j.2517-6161.1982.tb01183.x[16] Hill, B. M.：关于分布尾部推断的一种简单通用方法。《统计年鉴》3(5)，1163-1174（1975）。
doi:10.1214/aos/1176343247[17] Jiang, J.：混合效应模型的渐近分析：理论、应用与未决问题。CRC出版社（2017年）。
doi:10.1201/9781315119281[18]Jiang, J. 和 Nguyen, T.：线性与广义线性混合模型及其应用。Springer纽约（2007年）。
doi:10.1007/978-0-387-47946-0[19] Jiang, J.、Wand, M. P. 和 Bhaskaran, A.：广义线性混合模型分析与设计的实用且精确的渐近结果。皇家统计学会会刊：B辑（统计方法论）84(1)，55-82（2022年）。
doi:10.1111/rssb.12473[20] Kristensen, K.、Nielsen, A.、Berg, C. W.、Skaug, H. 和 Bell, B. M.：TMB：自动微分与拉普拉斯近似。统计软件杂志70(5)，1-21（2016年）。doi:10.18637/jss.v070.i05[21] Miyata, Y.：使用渐近众数的全指数拉普拉斯近似。美国统计学会会刊99(468)，1037-1049（2004年）。doi:10.1198/016214504000001673[22] Molina, I.、Corral, P. 和 Nguyen, M.：小区域贫困与不平等估计：综述与讨论。TEST 31(4)，1143-1166（2022年）。doi:10.1007/s11749-022-00822-1[23] Momoki, K. 和 Yoshida, T.：极值指数回归的混合效应模型。统计学通讯：理论与方法54(19)，6154-6171（2025年）。
doi:10.1080/03610926.2025.2450769[24] Nie, L.：广义线性和非线性混合效应模型中最大似然估计的收敛速度：理论与应用。统计规划与推断杂志137(6)，1787-1804（2007年）。doi:10.1016/j.jspi.2005.06.010[25] Ragulina, G. 和 Reitan, T.：利用长时间序列和贝叶斯方法研究极端降水的广义极值分布形状参数及其性质

方法。《水文科学杂志》**62**(6), 863-879 (2017)。
doi:10.1080/02626667.2016.1260134

[26] Rao, J. N. K. 和 Molina, I.：小区域估计。John Wiley & Sons（2015）。
doi:10.1002/9781118735855

[27] Reiss, R. D. 和 Thomas, M.：极值统计分析：在保险、金融、水文学及其他领域中的应用。Birkh¨auser Basel（2007年）。
doi:10.1007/978-3-7643-7399-3

[28]Rohrbeck, C. 和 Tawn, J. A.：水文变量极端行为的贝叶斯空间聚类。*arXiv*（2019）。doi:10.48550/arXiv.1906.08522

[29] Smith, R. L.：概率分布尾部的估计。《统计年鉴》**15**(3), 1174-1207（1987）。https://www.jstor.org/stable/2241822

[30] Sugasawa, S. 和 Kubokawa, T.：《混合模型的小区域估计：综述》。《日本统计与数据科学杂志》**3**(2), 693-720 (2020年).
doi:10.1007/s42081-020-00076-x

[31] Tro˘selj, J. 和 Lee, H. S.：台风引发的河流极端流量建模：日本台风海贝思案例研究。《水文学杂志：区域研究》**34**, 100776(2021).
doi:10.1016/j.ejrh.2021.100776

[32] Wang, H. 和 Tsai, C. L.：尾指数回归。《美国统计学会杂志》**104**(487), 1233-1240 (2009). doi:10.1198/jasa.2009.tm08458

[33] Wu, L.：复杂数据混合效应模型。CRC出版社 (2009)。
doi:10.1201/9781420074086