

Bayesian Mixture Model of $\mathcal{R}$ -Pareto Process for Extreme Precipitation

Student: Haitao Gao¹

Supervisors: Boris Béranger², Scott Sisson², Peng Zhong²

¹School of Computer Science and Engineering, UNSW

²School of Mathematics and Statistics, UNSW

November, 2025

Table of Contents

1 Recap

- Extreme Value Theory
- Motivation
- Previous Method

2 Bayesian Mixture Model

3 Simulation

- Two Distinguishable Subpopulations
- Two Indistinguishable Subpopulations
- Posterior Concentration
- Time Variant Weight Function Simulation

4 Application

5 Conclusion and Discussion

Table of Contents

1 Recap

- Extreme Value Theory
- Motivation
- Previous Method

2 Bayesian Mixture Model

3 Simulation

- Two Distinguishable Subpopulations
- Two Indistinguishable Subpopulations
- Posterior Concentration
- Time Variant Weight Function Simulation

4 Application

5 Conclusion and Discussion

What is Extreme Value (Temporal)

- 2008 **Global Financial Crisis**: severe global downturn, extreme volatility.
- Risky lending, housing bubble, complex derivatives; bailout turmoil.
- Stocks plunged, credit froze, millions lost jobs and homes.



Source: http://www.ritholtz.com/blog/wp-content/uploads/2008/09/crash_ws_j.png

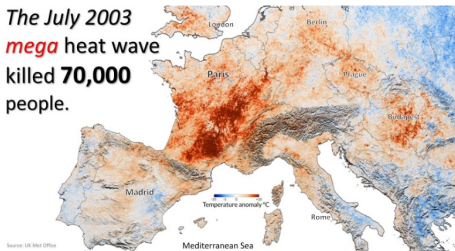


UNSW
SYDNEY

What is Extreme Value (Spatial)

- One of Europe's **hottest summers** on record
- Temperatures above 40°C , nearly 50°C in Portugal
- Drought, crop failure, and **70,000** heat-related deaths

The July 2003
mega heat wave
killed **70,000**
people.



Source: <https://fee.org/articles/lessons-from-the-deadly-european-h>



UNSW
SYDNEY

Extreme Value Theory (EVT)

Extreme events occur **rarely**, but they have a **significant impact** once they occur.

EVT

Extreme Value Theory (EVT) is an art of mathematical theory used to model these extreme values. EVT explores the asymptotic distribution of the extremely large or extremely small values, i.e. the tail distribution of a given probability function.

Motivation & Problem Definition

- Gap: few methods for **joint modelling of multiple extreme types**.
- Aim: **jointly model two types of extremes** in a unified framework.
- Approach: Bayesian mixture $\mathcal{R}$ -Pareto process with two subpopulations capturing different spatial dependence ranges.
- Novelty: to our knowledge, first work on this problem using $\mathcal{R}$ -Pareto processes.

- Try to fit the model by different subdata selected by **different risk functions L_p** .
- Issues: based on current theory, the likelihood function just can estimate L_1 unbiasedly. For other L_p normal, the likelihood is accurate only when the threshold is high enough asymptotically.
- Solution: just **use L_1 norm function** (accurate) but introduce two subpopulations with different range parameters to **build a mixture model**.

Table of Contents

- 1 Recap
 - Extreme Value Theory
 - Motivation
 - Previous Method
- 2 Bayesian Mixture Model
- 3 Simulation
 - Two Distinguishable Subpopulations
 - Two Indistinguishable Subpopulations
 - Posterior Concentration
 - Time Variant Weight Function Simulation
- 4 Application
- 5 Conclusion and Discussion

- Under climate change, extreme rainfall occurs across many scales: from ~ 1 km / 5 min to thousands of km^2 / several days.
- Characterising this behaviour is critical for monitoring and managing flood risk.
- Extreme rainfall arises from a **mixture of processes**:
 - **Convection**: local, short-lived, high-intensity storms (Schroeder *et al.*, 2018).
 - **Frontal**: large-scale, longer-duration, lower-intensity rainfall (Catto and Pfahl, 2013).

Types of Rainfall

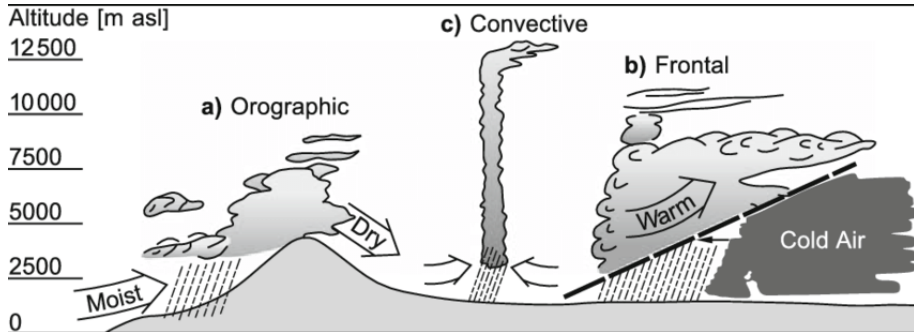


Figure: Types of Rainfall Formation: (a) Orographic, (b) Frontal, and (c) Convective. The primary drivers of extreme rainfall events are (b) Frontal lifting, which produces widespread, long-duration stratiform precipitation (affecting a large area), and (c) Convective instability, which generates highly intense, localized short-duration downpours. (a) Orographic rainfall is caused by air being forced up and over a mountain barrier.



Mixture Model: Variables

Formally, $\{s_1, \dots, s_n\} \subset \mathbb{R}^2$ be fixed spatial sites. A Brown–Resnick (BR) max-stable process $\{Z(s) : s \in \mathbb{R}^2\}$ with semivariogram $\gamma : \mathbb{R}^2 \rightarrow [0, \infty)$ admits the spectral representation.

$$Z(s) = \max_{i \geq 1} R_i \exp\left(W_i(s) - \frac{1}{2} \gamma(s)\right), \quad (1)$$

where $\{R_i\}$ is a Poisson point process on $(0, \infty)$ with intensity $\xi^{-2} d\xi$, and $\{W_i\}$ are i.i.d. centred Gaussian processes with stationary increments such that

$$(W(s) - W(t)) = 2\gamma(s - t), \quad s, t \in \mathbb{R}^2. \quad (2)$$

On the finite grid, the induced $n \times n$ matrix is

$$\Sigma_{ij}(\gamma) = \gamma(s_i) + \gamma(s_j) - \gamma(s_i - s_j), \quad 1 \leq i, j \leq n. \quad (3)$$

The isotropic power semivariogram is utilised,

$$\gamma_{\theta, \vartheta}(h) = (\|h\|/\theta)^\vartheta, \quad h \in \mathbb{R}^2,$$

with $\vartheta = 1.5$ throughout, and write $\Sigma(\theta) \equiv \Sigma(\gamma_{\theta, 1.5})$.

Mixture Model: Likelihood Function

Zhong *et al.* (2024) proposed the intensity function for flexible $\mathcal{R}$ -Pareto process. In this thesis, we drop the skewness such that the intensity function is determined by $\Sigma(\theta)$,

$$\kappa(\mathbf{x}; \Sigma(\theta)) = \frac{|\Sigma|^{-1/2} (\mathbf{1}^\top \mathbf{q})^{-1/2}}{(2\pi)^{(D-1)/2} \prod_{k=1}^D w_k} \times \exp \left\{ -\frac{1}{2} \left[\log \mathbf{x}^\top \mathcal{M} \log \mathbf{x} + \log \mathbf{x}^\top \left(\frac{2}{\mathbf{1}^\top} \right. \right. \right. \\ \left. \left. \left. + \frac{\mathbf{q}^\top \omega^2 - 1}{\mathbf{1}^\top \mathbf{q}} + \frac{1}{4} \omega^{2,\top} \mathcal{M} \omega^2 \right) \right] \right\} \quad (5)$$

where $\mathbf{q} = \Sigma^{-1} \mathbf{1}$, $\omega = \sqrt{\text{diag}(\Sigma)}$ and $\mathcal{M} = \Sigma^{-1} - \mathbf{q} \mathbf{q}^\top / \mathbf{1}^\top \mathbf{q}$.

Mixture Model: Likelihood

The likelihood function, therefore, is formulated,

$$\mathcal{L}(x|\theta) = \prod_i^n \frac{\kappa(x; \theta)}{K(\mathcal{A}_r; \theta)} \quad (6)$$

As we apply L_1 risk function to define the extremes with the function, we, therefore, can drop the denominator in the likelihood function 6 as it is a constant such that the likelihood is simplified,

$$\mathcal{L}(x|\theta) = \prod_i^n \kappa(x; \theta) \quad (7)$$

Bayesian Mixture Model

The mixture r-Pareto Process is modelled in a Bayesian perspective. We assume that there are two distinct distributions ($M = 2$) in the observations: a Brown-Resnick Process with **long range** and a Brown-Resnick Process with **small range** corresponding to precipitation caused by frontal and convective respectively. Formally, for observation $y = \{y_1, y_2, \dots, y_n\}$, the M-components ($M = 2$) mixture distribution is

$$f(y_i | \theta, \lambda) = \sum_{i=1}^M \lambda_m f_m(y_i | \theta_m) \quad (8)$$

where,

- $f_m(y | \theta_m)$ depends on the parameter θ_m where f_m is the intensity function 5.
- λ_m is proportion of y_i from component m, $\sum_m \lambda_m = 1$

Bayesian Mixture Model: Indicator Variable

There is an unobservable indicator variable z_{im} :

$$z_{im} = \begin{cases} 1, & \text{if } y_i \text{ is drawn from the } m^{th} \text{ component} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

Given the λ , the distribution of $z_i = (z_{i1}, \dots, z_{iM})$ is,

$$z_i \mid \lambda \sim \text{Multinomial}(1 : \lambda_1, \dots, \lambda_M) \quad (10)$$

where

$$\pi(z_i \mid \lambda) = \frac{1!}{z_{i1}! \dots z_{iM}!} \prod_{m=1}^M \lambda_m^{z_{im}} = \prod_{m=1}^M \lambda_m^{z_{im}}, \quad (11)$$

Prior Distribution

Assuming that the prior for (θ, λ) is independent: $\pi(\theta, \lambda) = \pi(\theta)\pi(\lambda)$. We define the λ is from $\text{Dirichlet}(\alpha_1, \dots, \alpha_M)$:

$$\pi(\lambda) \propto \prod_{m=1}^M \lambda_m^{\alpha_m-1} \quad (12)$$

where

- We always give $\text{Dirichlet}(1, \dots, 1)$ that is a non-informative prior
- Dirichlet $\pi(\lambda)$ is conjugate with Multinomial $\pi(z \mid \lambda)$
- When $M = 2$, Dirichlet Distribution degenerates to **Beta Distribution** and Multinomial Distribution degenerates to **Binomial Distribution** in this thesis.



UNSW
SYDNEY

For range parameter θ , we give Beta distribution $\text{Beta}(1, 1)$. Since there is an implicit order $\theta_1 > \theta_2$, one mathematical trick we use here is that we give the prior distributions for θ_2 and the gap δ such that

$$\theta_2 \sim \text{Gamma}(1, 1) \quad \delta \sim \text{Gamma}(1, 1) \quad (13)$$

where $\theta_1 = \theta_2 + \delta$.

Posterior Distribution

Formally, the mixture posterior distribution is given by,

$$\pi(\theta, z, \lambda \mid y) \propto L(y \mid z, \theta) \pi(z \mid \lambda) \pi(\lambda) \pi(\theta) \quad (14)$$

where $\pi(\theta)$ denotes prior density for the parameters of the component models, commonly $\pi(\theta) = \prod_m \pi_m(\theta_m)$. So we have,

$$\pi(\theta, z, \lambda \mid y) \propto \prod_{i=1}^n \left[\prod_{m=1}^M (\lambda_m f(y_i \mid \theta_m))^{z_{im}} \right] \left[\prod_{m=1}^M \lambda_m^{\alpha_m - 1} \right] \pi(\theta) \quad (15)$$

We have three parameters needed update (θ, z, λ).

Markov Chain Monte Carlo (MCMC)

- Family of algorithms to sample from complex posteriors when analytic solutions are infeasible.
- **Gibbs sampling**: update each parameter from its full conditional.
- **Metropolis–Hastings**: propose a move and accept/reject to keep the target posterior as the stationary distribution.

Update θ_m The full conditional distribution is

$$\pi(\theta_m \mid y, z, \lambda, \theta_{-m}) \propto \prod_{i=1}^n f(y_i \mid \theta_m)^{z_{im}} \pi_m(\theta_m) \quad (16)$$

where **random walk proposal** is utilised here to get new θ_m .

Update λ The full conditional distribution is

$$\pi(\lambda \mid y, \theta, z) \propto \prod_{i=1}^n \prod_{m=1}^M \lambda_m^{z_{im}} \lambda_m^{\alpha_m - 1} = \prod_{m=1}^M \lambda_m^{\sum_i z_{im} + \alpha_m - 1}. \quad (17)$$

where

- the distribution is proportion to $\text{Dirichlet}(\alpha_1 + n_1, \dots, \alpha_M + n_M)$
- $n_m = \sum_i z_{im}$

Update z we denote z_i as the updated parameter. The full conditional distribution is

$$\pi(z_i \mid y, \theta, \lambda) \propto \prod_{m=1}^M [\lambda_m f(y_i \mid \theta_m)]^{z_{im}}. \quad (18)$$

where,

- Each z_i is conditionally independent given λ, θ .
- $\Pr(z_{im} = 1 \mid \lambda, \theta) \propto \lambda_m f(y_i \mid \theta_m) = p_{im}$.
- Normalised probabilities $\tilde{p}_{im} = p_{im} / \sum_m p_{im}$.
- Full conditional is then $z_i \sim \text{Multinomial}(1 : \tilde{p}_{i1}, \dots, \tilde{p}_{iM})$.

Table of Contents

1 Recap

- Extreme Value Theory
- Motivation
- Previous Method

2 Bayesian Mixture Model

3 Simulation

- Two Distinguishable Subpopulations
- Two Indistinguishable Subpopulations
- Posterior Concentration
- Time Variant Weight Function Simulation

4 Application

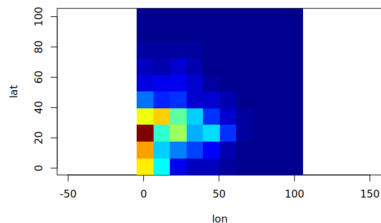
5 Conclusion and Discussion

Simulation Setup

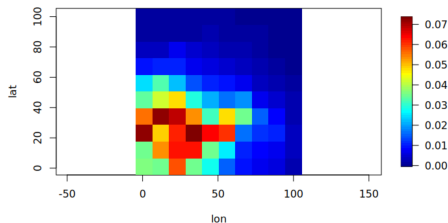
- Use a spatial grid to emulate a realistic precipitation field.
- Spatial dependence controlled by semivariogram parameter θ (distance-based).
- Simulated region: 10×10 square grid.
- Coordinates: latitude and longitude from 1 to 100 in steps of 10.

Map Realisation

The subpopulations have two distinguishable range parameters $\theta_1 = 50$, $\theta_2 = 25$ **and** $\lambda = 0.7$



(a)



(b)

Figure: Simulated spatial-field realisations illustrating the effect of the range parameter θ on dependence. (a) $\theta = 25$ (short range): dependence decays rapidly and high values remain localised. (b) $\theta = 50$ (long range): dependence decays more slowly, producing broader, smoother regions of elevated values. Axes show grid longitude/latitude; the colour bar indicates precipitation magnitude.



UNSW
SYDNEY

Results: Traceplots

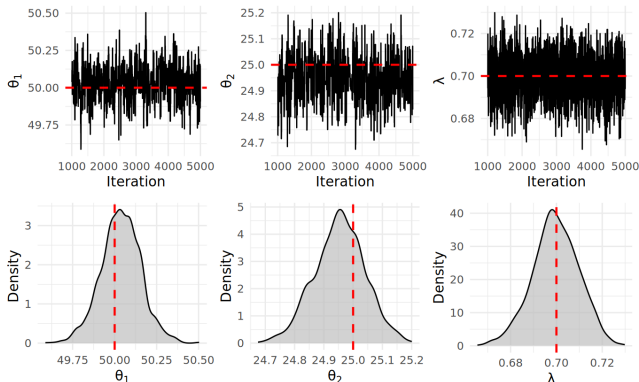


Figure: Trace (top) and marginal posterior density (bottom) for λ_1 , λ_2 , and π in the distinguishable-mixture simulation. The chains mix well and remain stable after burn-in. Red dashed lines mark the true parameter values; the posteriors concentrate near these targets, indicating accurate recovery by the proposed mixture model.



Posterior Mean of Indicator

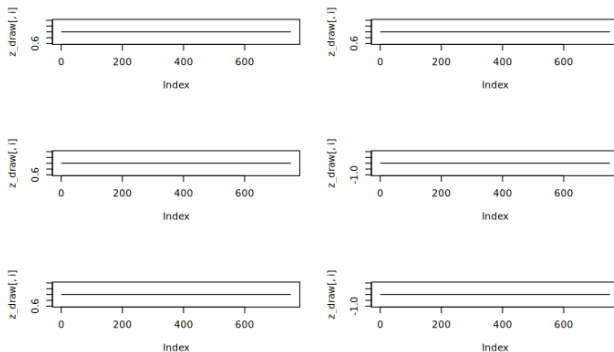


Figure: Posterior sample paths of the latent indicator z_i for six representative locations ($i \in \{1, 34, 56, 201, 344, 1002\}$). Each panel shows z_i across the kept MCMC iterations after burn-in. The traces are essentially flat at 0 or 1 with no label switching, implying posterior classification probabilities near 0 or 1 and strong separability between the frontal and convective components.



Spatial Dependency

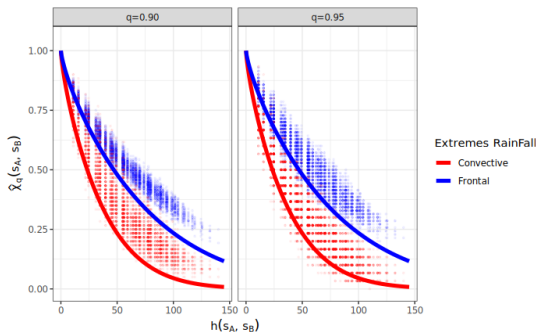


Figure: Empirical estimates of the upper tail dependence measure $\hat{\chi}_q(s_A, s_B)$ against spatial distance $h(s_A, s_B)$ for two rainfall components: frontal (blue) and convective (red) at quantiles $q = 0.90$ and $q = 0.95$. The solid lines are the fitted curves from the Gaussian Extremal Model. Frontal rainfall (blue) shows stronger and longer-range extremal dependence ($\hat{\chi}_q$ decays slower) compared to convective rainfall (red), which exhibits rapid decay.



Map Realisation

We set the range parameters to $\theta_1 = 50$, $\theta_2 = 40$ **and** $\lambda = 70$.

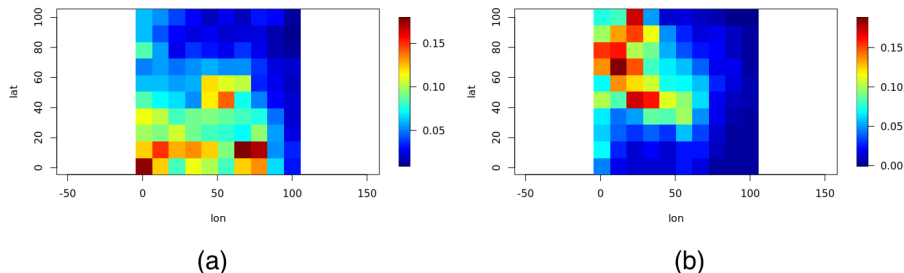


Figure: Simulated spatial-field realisations illustrating the effect of the range parameter θ on dependence, for two subpopulations with distinguishable ranges $\theta_1 = 50$ and $\theta_2 = 40$. (a) $\theta = 50$ (longer range): dependence decays more slowly, producing broader, smoother regions of elevated values. (b) $\theta = 40$ (shorter range): dependence decays more quickly, so high values remain more localised and spatial variability is more pronounced. Axes show grid longitude/latitude; colour bar indicates precipitation magnitude.



Results: Traceplot

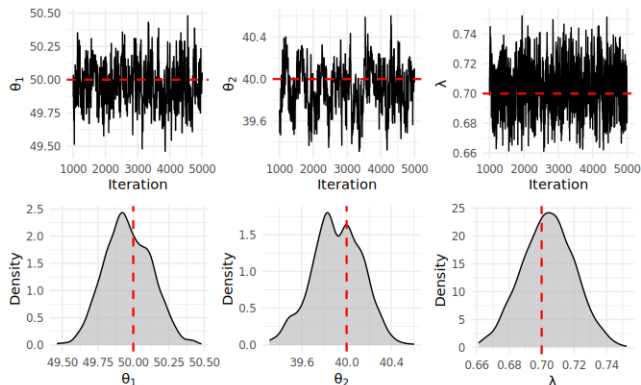


Figure: Trace (top) and marginal posterior density (bottom) for θ_1 , θ_2 , and λ in the similar-subpopulation setting. The two components have close parameter values, yielding broader and more overlapping posteriors. Red dashed lines mark the true values; the chains remain stable after burn-in and centre on these targets, indicating successful recovery despite weak separation.



Results: Posterior Mean of Indicator

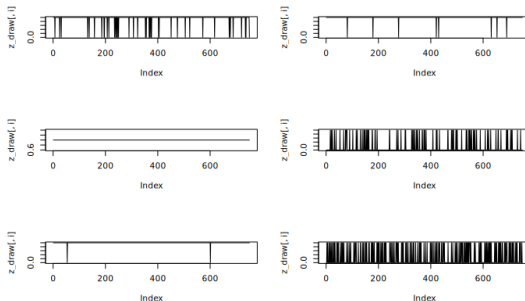


Figure: Posterior sample paths of the latent indicator z_i for six representative locations in the *similar-subpopulation* setting. Unlike the distinguishable case, several traces show frequent switches between 0 and 1, reflecting higher posterior uncertainty in class membership when component parameters are close. A few sites remain near a single state, indicating local identifiability, but overall the mixture is harder to separate under weak component separation.



Results: Spatial Dependency

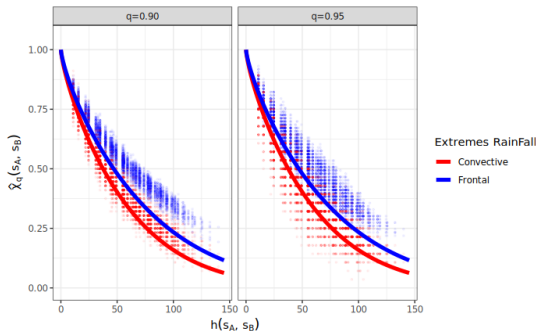


Figure: Empirical estimates of the upper tail dependence measure $\hat{\chi}_q(s_A, s_B)$ against spatial distance $h(s_A, s_B)$ for two rainfall components: frontal (blue) and convective (red) at quantiles $q = 0.90$ and $q = 0.95$. The solid lines are the fitted curves derived from the Gaussian Extremal Model. Frontal rainfall (blue) shows stronger and longer-range extremal dependence ($\hat{\chi}_q$ decays slower) compared to convective rainfall (red), which exhibits rapid decay.



Posterior Concentration

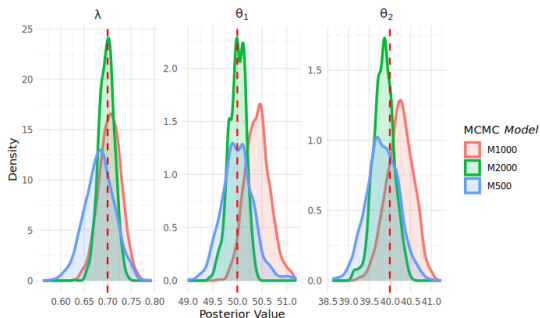


Figure: Posterior marginal densities for $(\lambda, \theta_1, \theta_2)$ under increasing sample size ($N \in \{500, 1000, 2000\}$) with true parameters $\theta_1 = 50$, $\theta_2 = 40$, and $\lambda = 0.7$. Curves correspond to fits M500 (blue), M1000 (red), and M2000 (green); vertical dashed lines mark the true values. As N grows, the posteriors contract and concentrate around the truth, illustrating posterior shrinkage and improved precision.



Time Variant Weight

we define the time-varying mixing weight as

$$\lambda(t) = 0.5 - 0.5 \sin t \quad (19)$$

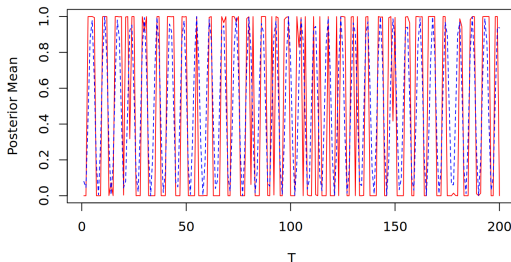


Figure: Time-varying classification: posterior mean of the indicator z_t (blue) versus the ground-truth mixing weight $\lambda(t)$ (red). The close alignment shows that the model recovers the temporal pattern of frontal-convective prevalence; small discrepancies appear mainly near rapid transitions where uncertainty is highest.

Time Variant Weight

we investigate the generalisation of the proposed model to a **different** time-varying weight function to show the model can really capture the correct probability of frontal extremes.

$$\lambda(t) = 0.5 \sin t + 0.5 \quad (20)$$

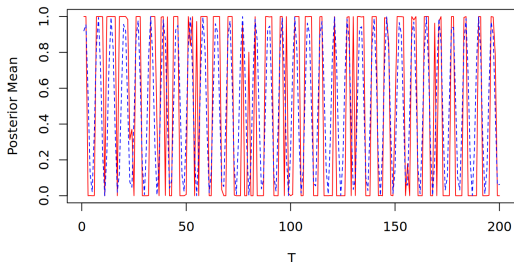


Figure: Time-varying classification with an *inverted* mixing function: posterior mean of the indicator z_t (blue) versus the inverted ground-truth curve $1 - \lambda(t)$



Based on the four simulation study, our proposed model:

- Can estimate the correct parameters in both distinguishable and Indistinguishable subpopulations.
- The posterior mean of indicator z works properly (i) confident to assign membership when they are distinguishable (ii) many label switching when the subpopulations are close
- The posterior distribution is concentrated when the sample size is getting larger
- Can capture the true weight parameters

Table of Contents

- 1 Recap
 - Extreme Value Theory
 - Motivation
 - Previous Method
- 2 Bayesian Mixture Model
- 3 Simulation
 - Two Distinguishable Subpopulations
 - Two Indistinguishable Subpopulations
 - Posterior Concentration
 - Time Variant Weight Function Simulation
- 4 Application
- 5 Conclusion and Discussion

Dataset

Precipitation over the Tampa Bay region using weather radar measurements sampled every 15 minutes from 1995–2019, yielding 139,881 images with 4,449 spatial cells per image. Figure 13 shows a representative radar snapshot. We restrict attention to the full calendar year 1995.

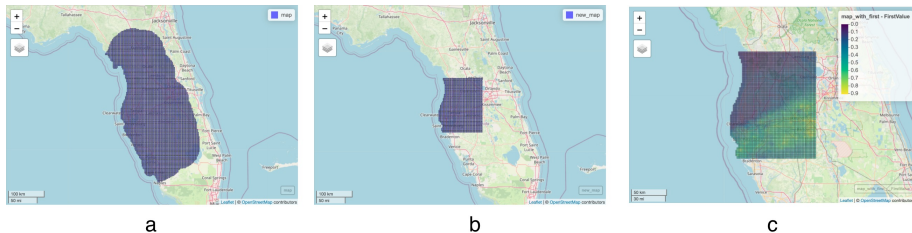


Figure: Spatial maps of the study area over Florida. (a) The full Tampa Bay coverage area, showing the complete spatial grid. (b) The restricted radar range of the Bay, focusing on the central analysis region. (c) One replicate of the daily rainfall field, mapped onto the radar grid with one replicate observed rainfall values visualised using a continuous colour scale.



- Upsampling rainfall data to hourly resolution.
- Use EPSG:3577; coordinates normalised to grid-cell units.
- Define a 14×14 grid: latitude $[16,357,547, 16,388,530]$, longitude $[-7,910,861, -7,886,519]$.

Framed Map

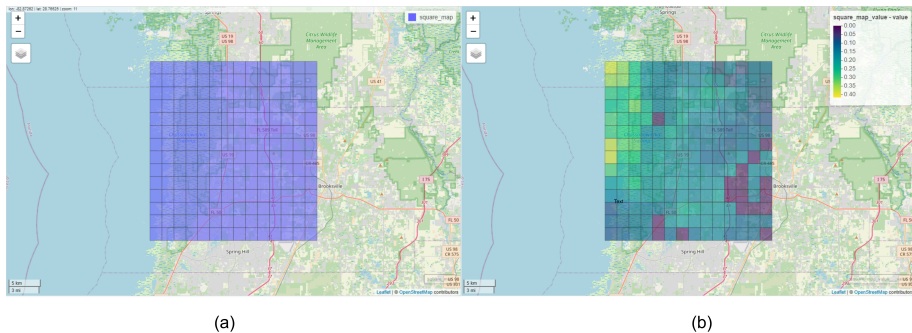


Figure: Framed map over spatial map. (a) The map is framed by 14×14 squared grid. (b) One realisation of hourly rainfall over the squared grid region.



UNSW
SYDNEY

Results: Traceplot

The posterior means are $\hat{\theta}_1 = 6.856$, $\hat{\theta}_2 = 2.094$, and $\hat{\lambda} = 0.419$.

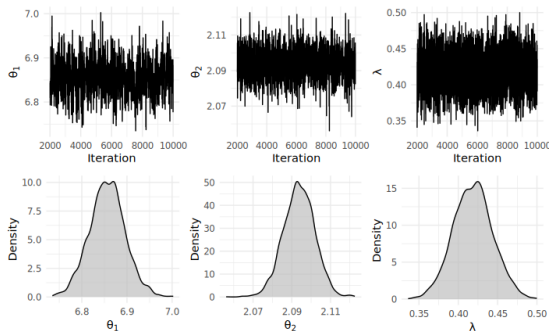


Figure: Tampa Bay application: trace (top) and marginal posterior density (bottom) for θ_1 , θ_2 , and mixing weight λ . The chains show stable mixing after burn-in with no apparent drifts, and the marginal posteriors are unimodal and reasonably concentrated, indicating well-identified spatial ranges and mixture proportion for the Tampa Bay rainfall field.



Results: Posterior Mean of Indicator

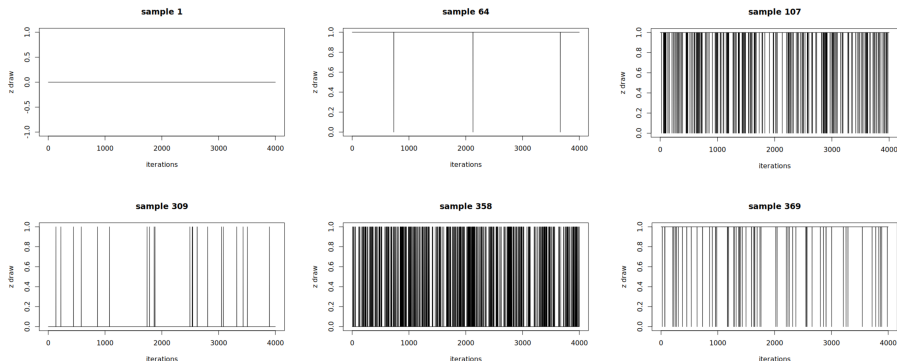


Figure: Posterior sample paths of the latent indicator z_i for six representative locations ($i \in \{1, 64, 107, 309, 385, 369\}$). Each panel shows z_i across the kept MCMC iterations after burn-in. The plots show the different confidence to assign the membership to different samples

Results: Spatial Dependency

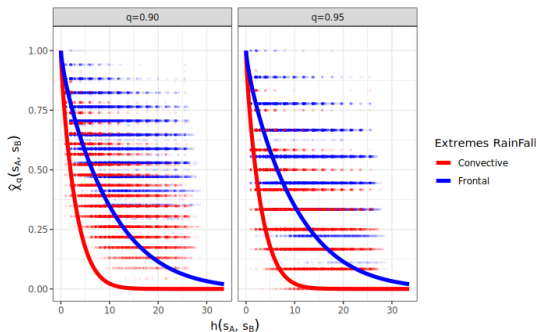


Figure: Empirical extremal dependence by distance for Tampa Bay. Each point is a site pair (s_A, s_B) with vertical coordinate $\hat{\chi}_q(s_A, s_B)$ and horizontal axis the intersite distance $h(s_A, s_B)$, shown for $q = 0.90$ (left) and $q = 0.95$ (right). Colours denote posterior regime: Comp1 (frontal) and Comp2 (convective). The convective pairs display a steep decline of $\hat{\chi}_q$ toward zero with distance, while the frontal pairs retain non-negligible extremal dependence at larger separations, corroborating the fitted short vs. long range structure.

Results: Comparison with Existing Literature

Given the working grid resolution (one grid unit ≈ 1.38 km), these correspond to an effective long-range dependence of approximately **9.46 km** for the frontal regime and a localised dependence of about **2.89 km** for the convective regime.

Table: Model Comparison and Proposed Mixture Results

Model	Risk Function	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}$
BR	L_1			5.91 (1.71)
sBR	L_1			5.81 (1.74)
Mixture Model	L_1	9.46 (0.054)	2.89 (0.011)	6.17 (0.028)

Example Realisations

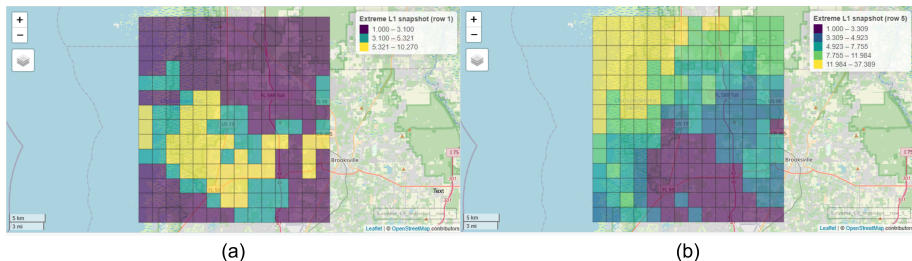


Figure: Tampa Bay (hourly resolution): two representative extreme snapshots used for regime assignment. Each grid cell side is ≈ 1.38 km; colours map intensity with yellow = higher values. (a) Localised, cellular yellow patches indicative of a short-range convective pattern (estimated range ≈ 2.76 km from $\hat{\theta}_2 = 2.093$). (b) Broad, coherent swaths of elevated intensity consistent with a long-range frontal pattern (estimated range ≈ 9.46 km from $\hat{\theta}_1 = 6.856$). These examples align with the posterior mixing weight $\hat{\lambda} = 0.420$.



Surprise

The seasonal pattern found by our model is matched with U.S. Geological Survey (Yates *et al.*, 2011).

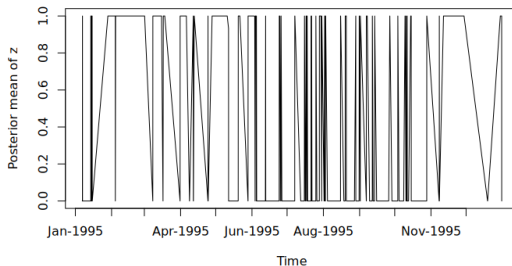


Figure: The posterior mean of indicator z : In the wet season (June - September) Tempa Bay contains more localised convective rainfall and it changes rapidly. In dry season (October - June) Tempa Bay includes more frontal rainfall and it is more stable than wet season.



UNSW
SYDNEY

Climate and Weather

The region has a humid subtropical climate with an average annual temperature of about 72 °F and average annual rainfall that ranges from 50 to 55 in. in different parts of the watershed (Lewis and Estevez, 1988; Wolfe and Drew, 1990). About 60 percent of the annual rainfall usually occurs during the summer (mid-June through September) rainy season, in the form of localized convective thunderstorms and occasional tropical storms and hurricanes. During the dry season, which generally extends from October through early June, the rainfall that occurs is usually associated with the passage of large-scale frontal systems. Rain events associated with frontal passages are most common during the January-through-March period, producing a period of somewhat elevated rainfall during an otherwise dry season (Flannery, 1989). The months of lowest rainfall are usually November, April, and May. Mean daily rainfall values for the period of record (1900s–2007) at four locations in the Tampa Bay watershed are shown in fig. 2–3.

Figure: A screenshot of the paragraph of the U.S Geological Survey



Table of Contents

- 1 Recap
 - Extreme Value Theory
 - Motivation
 - Previous Method
- 2 Bayesian Mixture Model
- 3 Simulation
 - Two Distinguishable Subpopulations
 - Two Indistinguishable Subpopulations
 - Posterior Concentration
 - Time Variant Weight Function Simulation
- 4 Application
- 5 Conclusion and Discussion

- **Fixed Parameters:** Smoothness (ϑ) and isotropy were fixed; future work can model these with increased computational resources.
- **Spatial Extent:** The model was not fitted over the full spatial lattice due to computational constraints.
- **Likelihood Choice:** Employed an L_1 -based angular likelihood; exploring other L_p likelihoods may improve estimation accuracy in non-asymptotic regimes.
- **Fixed Components:** Assumed two subpopulations ($M = 2$); allowing an unknown number of components requires hierarchical priors and identification safeguards.

- **Richer Spatial Kernels:** Jointly infer θ , ϑ (smoothness), and anisotropy within each component for enhanced spatial modeling.
- **General Risk Functions:** Directly integrate L_∞ or score matching likelihoods and compare to current L_1 -based approximations.
- **Adaptive Model Complexity:** Place a prior on the number of components (M) to automatically discover multiple extremal regimes.
- **Location Mixtures:** Allow site-level mixture memberships, enabling different locations to exhibit distinct extreme event types simultaneously via spatially-varying weights or latent labels.

Thanks for your attention.

Q&A

References I

- Catto, J. L. and Pfahl, S. (2013) The importance of fronts for extreme precipitation. *Journal of Geophysical Research: Atmospheres* **118**, 10–791.
- Schroeder, K., Kirchengast, G. and O, S. (2018) Strong dependence of extreme convective precipitation intensities on gauge network density. *Geophysical Research Letters* **45**, 8253–8263.
- Yates, K. K., Greening, H. and Morrison, G. (eds) (2011) *Integrating Science and Resource Management in Tampa Bay, Florida*. Number 1348 in U.S. Geological Survey Circular. Reston, Virginia: U.S. Geological Survey.
- Zhong, P., Sisson, S. A. and Beranger, B. (2024) Flexible max-stable processes for fast and efficient inference. *arXiv preprint arXiv:2407.13958*.