MULTIVARIATE MODELING OF LOW, MODERATE, AND LARGE POSITIVE VALUES WITHOUT THRESHOLD SELECTION STEPS

Carlo Gaetan

Università Ca' Foscari di Venezia, Italy

gaetan@unive.it

SHARE kick-off meeting (Brest, France)

7 January 2026

BASED ON

Cornell University

arXiv > stat > arXiv:2510.02152

Statistics > Methodology

[Submitted on 2 Oct 2025]

Multivariate distributional modeling of low, moderate, and large intensities without threshold selection steps

Carlo Gaetan, Philippe Naveau

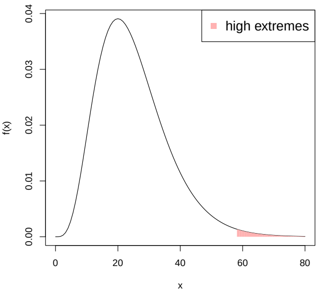
Warm up

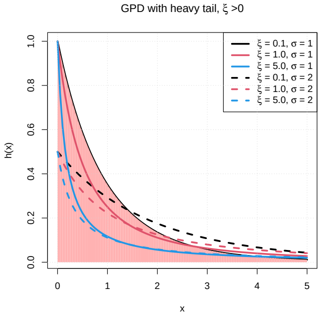
GENERALIZED PARETO DISTRIBUTION (GPD)

The distribution of X, when X exceeds a high threshold u, can be approximated by a GPD

$$H_\xi((x-u)/\sigma) = \begin{cases} 1-(1+\xi(x-u)/\sigma)_+^{-1/\xi} & \text{for } \xi \neq 0 \\ 1-\exp(-(x-u)/\sigma) & \text{for } \xi = 0 \end{cases}$$

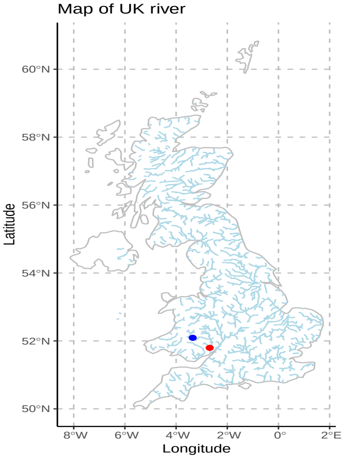
ξ **shape parameter**, $\sigma > 0$ **scale parameter** and $a_+ = \max(a, 0)$.

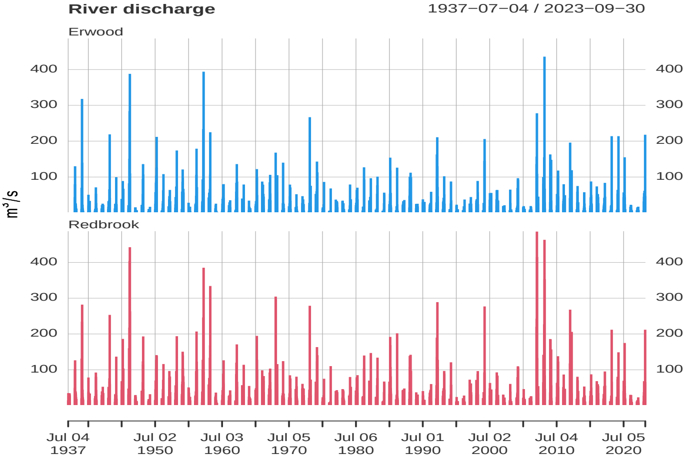




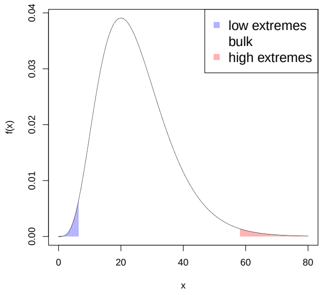
Example

EXAMPLE: WEEKLY MAXIMUM SUMMER RIVER DISCHARGES OF WYE RIVER



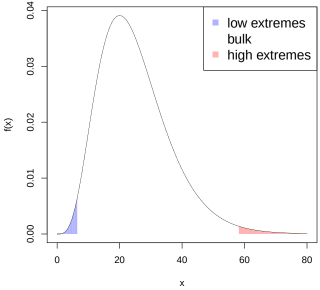


RIVER DISCHARGES



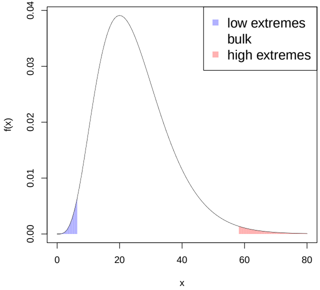
- **Flood risk managers** often focus on the analysis of high river flows
- **Farmers** may be interested in periods of low river runoffs to prevent food production shortages
- **Energy producers** in charge of electrical dams can be concerned by the full range of the variable of interest

RIVER DISCHARGES



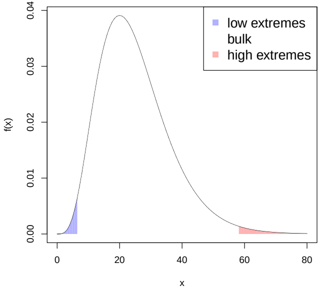
- **Flood risk managers** often focus on the analysis of high river flows
- **Farmers** may be interested in periods of low river runoffs to prevent food production shortages
- **Energy producers** in charge of electrical dams can be concerned by the full range of the variable of interest

RIVER DISCHARGES



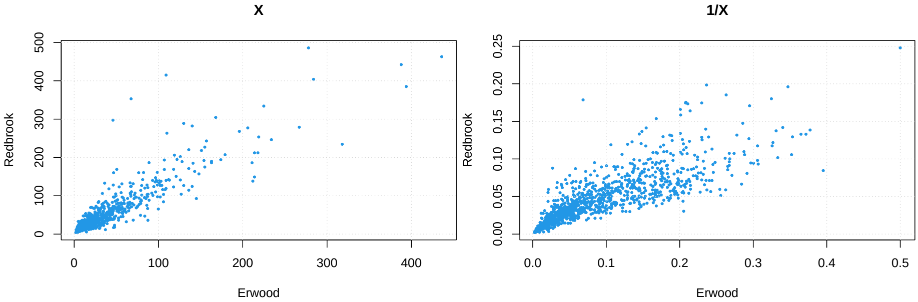
- **Flood risk managers** often focus on the analysis of high river flows
- **Farmers** may be interested in periods of low river runoffs to prevent food production shortages
- **Energy producers** in charge of electrical dams can be concerned by the full range of the variable of interest

RIVER DISCHARGES



- **Flood risk managers** often focus on the analysis of high river flows
- **Farmers** may be interested in periods of low river runoffs to prevent food production shortages
- **Energy producers** in charge of electrical dams can be concerned by the full range of the variable of interest

Sites along the same river basin as nearby measurements can be strongly dependent



What is an Extended Generalized Pareto Distribution ?

Water Resources Research

Research Article | Free Access

Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection

Philippe Naveau, Raphael Huser, Pierre Ribereau, Alexis Hannart

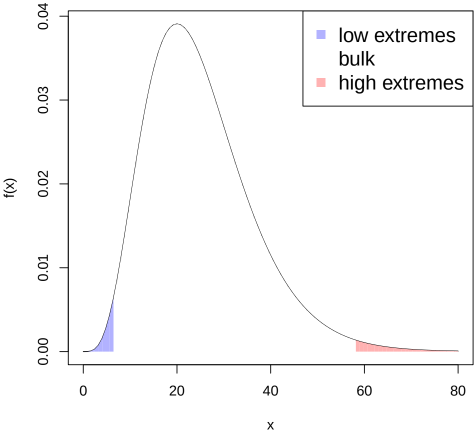
First published: 04 March 2016 | https://doi.org/10.1002/2015WR018552 | Citations: 110

Volume 52, Issue 4
April 2016
Pages 2753-2769

Figures | References | Related | Information

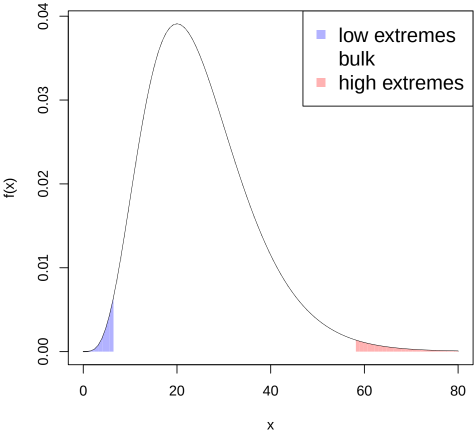
- Distribution for positive data !

TWO STRATEGIES TO MODEL JOINTLY EXTREMES AND BULK

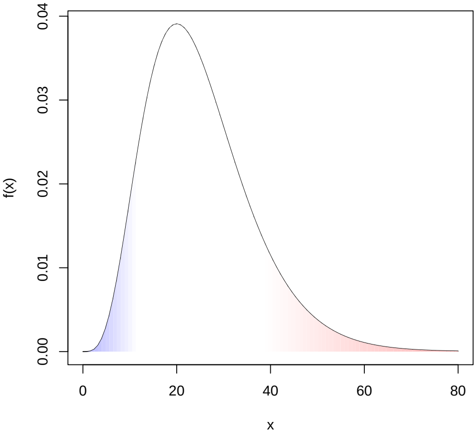


Classical approach

TWO STRATEGIES TO MODEL JOINTLY EXTREMES AND BULK



Classical approach



EGPD approach

THE THREE INGREDIENTS FOR A EGPD

High extremes

ξ the GPD parameter of X

THE THREE INGREDIENTS FOR A EGPD

Low extremes

$1/\kappa$ the GPD parameter of $1/X$

High extremes

ξ the GPD parameter of X

THE THREE INGREDIENTS FOR A EGPD

Low extremes

$1/\kappa$ the GPD parameter of $1/X$

Bulk

$B(u)$ a CDF on $[0, 1]$ with a pdf $b(u)$

High extremes

ξ the GPD parameter of X

Low extremes

$1/\kappa$ the GPD parameter of $1/X$

Bulk

$B(u)$ a CDF on $[0, 1]$ with a pdf $b(u)$

High extremes

ξ the GPD parameter of X

$$\Pr(X \leq x) = B(H_{\xi}(x)^{\kappa})$$

where the pdf $b(u)$ is such that

$$0 < b(0) < \infty \qquad \text{and} \qquad 0 < b(1) < \infty$$

Low extremes

$1/\kappa$ the GPD parameter of $1/X$

Bulk

$B(u)$ a CDF on $[0, 1]$ with a pdf $b(u)$

High extremes

ξ the GPD parameter of X

$$\Pr(X \leq x) = B(H_{\xi}(x)^{\kappa})$$

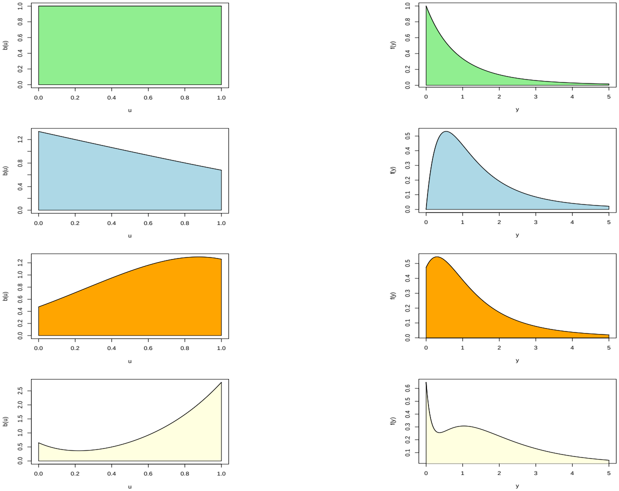
where the pdf $b(u)$ is such that

$$0 < b(0) < \infty \qquad \text{and} \qquad 0 < b(1) < \infty$$

- Scale parameter σ is absorbed in $B(u)$

EGPD EXAMPLES

$b(u)$ EGPD density



TAKE HOME MESSAGE

- $X \sim \text{EGPD}(\kappa, \xi, B)$
- $1/X \sim \text{EGPD}(1/\xi, 1/\kappa, \tilde{B})$
- $B(\cdot) \rightarrow \tilde{B}(\cdot)$

The inverse $1/X$ of a EGPD r.v. is still an EGPD

LEMMA

If

$$X \sim EGPD(\kappa, \xi, B), \quad \Leftrightarrow \quad \Pr(X \leq x) = B(H_{\xi}(x)^{\kappa})$$

then

$$1/X \sim \mathrm{EGPD}(1/\xi, 1/\kappa, \tilde{B}), \quad \Leftrightarrow \quad \Pr(1/X \leq \tilde{x}) = \tilde{B}(H_{1/\kappa}(\tilde{x})^{1/\xi})$$

with

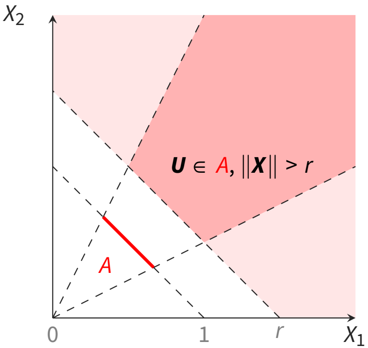
$$\tilde{b}(0) = \kappa\, \xi^{1/\xi}\, b(1) \quad \text{and} \quad \tilde{b}(1) = \xi\, b(0)$$

A multivariate EGPD

POLAR EXTREMES COORDINATES OF $\boldsymbol{X} = (X_1, X_2)$

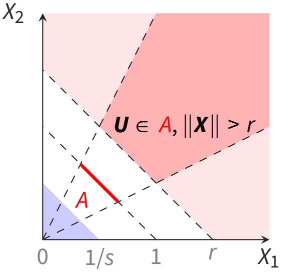
$$\boldsymbol{X} = \|\boldsymbol{X}\| \times \boldsymbol{U}$$

pseudo-radius $\|\boldsymbol{X}\| = X_1 + X_2$ and pseudo-angle $\boldsymbol{U} = (X_1/\|\boldsymbol{X}\|, X_2/\|\boldsymbol{X}\|)$



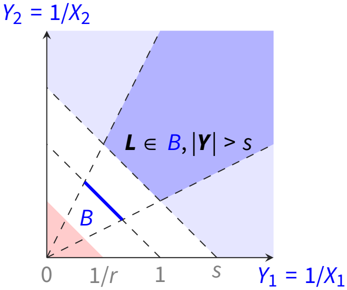
$$\boldsymbol{X} = \|\boldsymbol{X}\| \times \boldsymbol{U} \qquad \boldsymbol{Y} = |\boldsymbol{Y}| \times \boldsymbol{L}$$

Upper extreme's representation



Radius $\|\boldsymbol{X}\| = X_1 + X_2$ and $\boldsymbol{U} = \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|}$

Lower extreme's representation



$|\boldsymbol{Y}| = \frac{1}{\|\boldsymbol{1}/\boldsymbol{Y}\|}$ and $\boldsymbol{L} = \frac{\boldsymbol{Y}}{|\boldsymbol{Y}|}$

TAKE HOME MESSAGE

- $\boldsymbol{X} = (X_1, \ldots, X_d)$ with $X_i \sim \text{EGPD}(\kappa, \xi, B)$ for all $i = 1, \ldots, d$.
- Distribution of $\|\boldsymbol{X}\| = X_1 + \cdots + X_d$

The sum of dependent EGPDs is still an EGPD

- The upper tail parameter, ξ, remains unchanged from marginal behaviors to $\|\boldsymbol{X}\|$, while the lower tail behavior can be have a changing κ.
- $B(\cdot) \longrightarrow B_d(\cdot)$

PROPOSITION: DISTRIBUTION OF $\|\boldsymbol{X}\| = X_1 + \cdots + X_d$

Let be $\boldsymbol{X} = (X_1, \ldots, X_d)$ with $X_i \sim \text{EGPD}(\kappa, \xi, B)$ for all $i = 1, \ldots, d$.
If there exist three positive and finite constants, a, c_- and c_+ such that

$$\lim_{x \to \infty} \frac{\Pr(\|\mathbf{X}\| > x)}{\Pr(X_i > x)} = c_+ \quad (1)$$

and

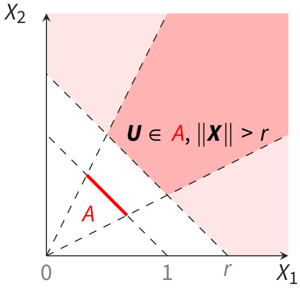
$$\lim_{x \to 0^+} \frac{\Pr(\|\mathbf{X}\| \leq x)}{[\Pr(X_i \leq x)]^a} = c_-, \quad (2)$$

then there exists a CDF B_d such that $\|\mathbf{X}\| \sim \text{EGPD}(a\kappa, \xi, B_d)$ with

$$b_d(0) = c_- \, b(0)^a \text{ and } b_d(1) = c_+ \, b(1)/a.$$

- Moreover $|\boldsymbol{Y}| = 1/\|\boldsymbol{X}\|$ is also $\text{EGPD}(1/\xi, 1/(a\kappa), \tilde{B}_d)$

MULTIVARIATE REGULAR VARIATION DISTRIBUTION



$$\mathbf{X} = \|\boldsymbol{X}\| \times \boldsymbol{U}$$

- $\|\mathbf{X}\|$ independent of $\boldsymbol{U}$ when $\|\mathbf{X}\|$ gets large
- $\Pr(\boldsymbol{U} \in A \mid \|\mathbf{X}\| > r)$ has a non-degenerate limit as $r \to \infty$, i.e.

$$\lim_{r \to \infty} \Pr(\boldsymbol{U} \in A \mid \|\mathbf{X}\| > r) = \Pr(\boldsymbol{U} \in A)$$

The main differences with classical EVT modelling are that

1. Our interest is not only on the upper extremal behaviour of $\boldsymbol{X}$, but also its lower extremal behaviour
2. The radial component is assumed to follow a EGPD, and consequently be in compliance with EVT for both small and large values of $||\mathbf{X}||$
3. In contrast to classical regular variation principles, the radial component **is not necessarily assumed independent** of the angular component
4. In particular, the degree of dependence will change according to the value $||\mathbf{X}||$

The main differences with classical EVT modelling are that

1. Our interest is not only on the upper extremal behaviour of $\boldsymbol{X}$, but also its lower extremal behaviour
2. The radial component is assumed to follow a EGPD, and consequently be in compliance with EVT for both small and large values of $||\mathbf{X}||$
3. In contrast to classical regular variation principles, the radial component **is not necessarily assumed independent** of the angular component
4. In particular, the degree of dependence will change according to the value $||\mathbf{X}||$

The main differences with classical EVT modelling are that

1. Our interest is not only on the upper extremal behaviour of $\boldsymbol{X}$, but also its lower extremal behaviour
2. The radial component is assumed to follow a EGPD, and consequently be in compliance with EVT for both small and large values of $\|\mathbf{X}\|$
3. In contrast to classical regular variation principles, the radial component **is not necessarily assumed independent** of the angular component
4. In particular, the degree of dependence will change according to the value $\|\mathbf{X}\|$

The main differences with classical EVT modelling are that

1. Our interest is not only on the upper extremal behaviour of $\boldsymbol{X}$, but also its lower extremal behaviour
2. The radial component is assumed to follow a EGPD, and consequently be in compliance with EVT for both small and large values of $||\mathbf{X}||$
3. In contrast to classical regular variation principles, the radial component **is not necessarily assumed independent** of the angular component
4. In particular, the degree of dependence will change according to the value $||\mathbf{X}||$

The main differences with classical EVT modelling are that

1. Our interest is not only on the upper extremal behaviour of $\boldsymbol{X}$, but also its lower extremal behaviour
2. The radial component is assumed to follow a EGPD, and consequently be in compliance with EVT for both small and large values of $||\mathbf{X}||$
3. In contrast to classical regular variation principles, the radial component **is not necessarily assumed independent** of the angular component
4. In particular, the degree of dependence will change according to the value $||\mathbf{X}||$

BIVARIATE EGPD $\boldsymbol{X} = (X_1, X_2)$, WITH FOUR INGREDIENTS

$$X_1 = \|\boldsymbol{X}\| U_1 \quad \text{and} \quad X_2 = \|\boldsymbol{X}\| U_2 = \|\boldsymbol{X}\|(1 - U_1)$$

$$\|\boldsymbol{X}\| \sim EGPD(\kappa, \xi, B_d)$$

Low extremes

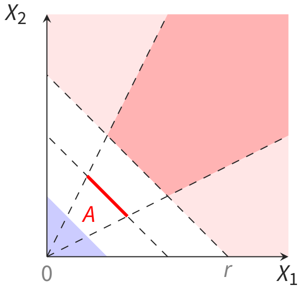
κ_d the GPD parameter of $1/\|\boldsymbol{X}\|$

Bulk

B_d a CDF function on $[0, 1]$ (with PDF b_d)

High extremes

ξ the GPD parameter of $\|\boldsymbol{X}\|$



$\boldsymbol{U} = \boldsymbol{X}/\|\boldsymbol{X}\|$

Bivariate conditional model

$$\left[\log\left(\frac{U_1}{1 - U_1}\right) \middle| \|\boldsymbol{X}\| = r\right] \stackrel{d}{=} \delta(r) Z$$

with Z standardized Gaussian $\perp \|\boldsymbol{X}\|$

Bivariate conditional model

$$\left[\log\left(\frac{U_1}{1-U_1}\right)\middle| \|\boldsymbol{X}\| = r\right] \stackrel{d}{=} \delta(r)Z$$

with Z standardized Gaussian $\perp \|\boldsymbol{X}\|$

- If $\delta(r)$ remains constant for large values of r, then we are in the multivariate regular variation framework
- We can specify other conditional distribution

$$U_1 \,|\, \|\boldsymbol{X}\| = r \sim f(\cdot; \delta(r))$$

  under the constraint $\mathbb{E}\left(U_1 \,|\, \|\boldsymbol{X}\| = r\right) = 1/2$
- Why Gaussian model? Flexible multivariate distribution that is easy to specify and estimate!

Bivariate conditional model

$$\left[\log\left(\frac{U_1}{1-U_1}\right)\middle|\|\boldsymbol{X}\| = r\right] \stackrel{d}{=} \delta(r)Z$$

with Z standardized Gaussian $\perp \|\boldsymbol{X}\|$

- If $\delta(r)$ remains constant for large values of r, then we are in the multivariate regular variation framework
- We can specify other conditional distribution

$$U_1 \mid \|\boldsymbol{X}\| = r \sim f(\cdot; \delta(r))$$

  under the constraint $\mathbb{E}\left(U_1 \mid \|\boldsymbol{X}\| = r\right) = 1/2$
- Why Gaussian model? Flexible multivariate distribution that is easy to specify and estimate!

Bivariate conditional model

$$\left[\log\left(\frac{U_1}{1-U_1}\right)\Bigg|\,\|\boldsymbol{X}\| = r\right] \overset{d}{=} \delta(r)Z$$

with Z standardized Gaussian $\perp \|\boldsymbol{X}\|$

- If $\delta(r)$ remains constant for large values of r, then we are in the multivariate regular variation framework
- We can specify other conditional distribution

$$U_1 | \|\boldsymbol{X}\| = r \sim f(\cdot; \delta(r))$$

under the constraint $\mathbb{E}\left(U_1 | \|\boldsymbol{X}\| = r\right) = 1/2$

- Why Gaussian model? Flexible multivariate distribution that is easy to specify and estimate!

Bivariate conditional model

$$\left[\log\left(\frac{U_1}{1-U_1}\right)\middle|\,\|\boldsymbol{X}\| = r\right] \stackrel{d}{=} \delta(r)Z$$

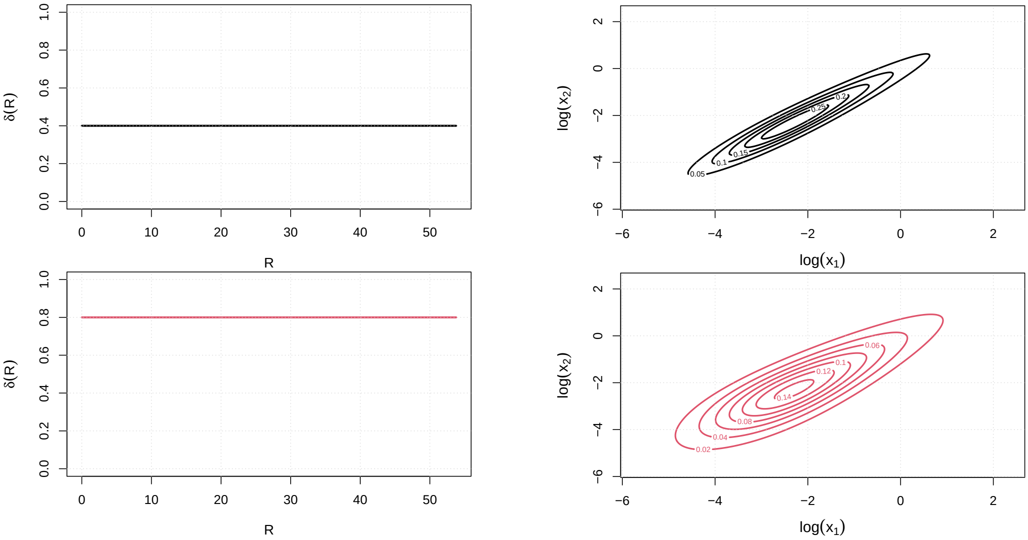
with Z standardized Gaussian $\perp \|\boldsymbol{X}\|$

- If $\delta(r)$ remains constant for large values of r, then we are in the multivariate regular variation framework
- We can specify other conditional distribution

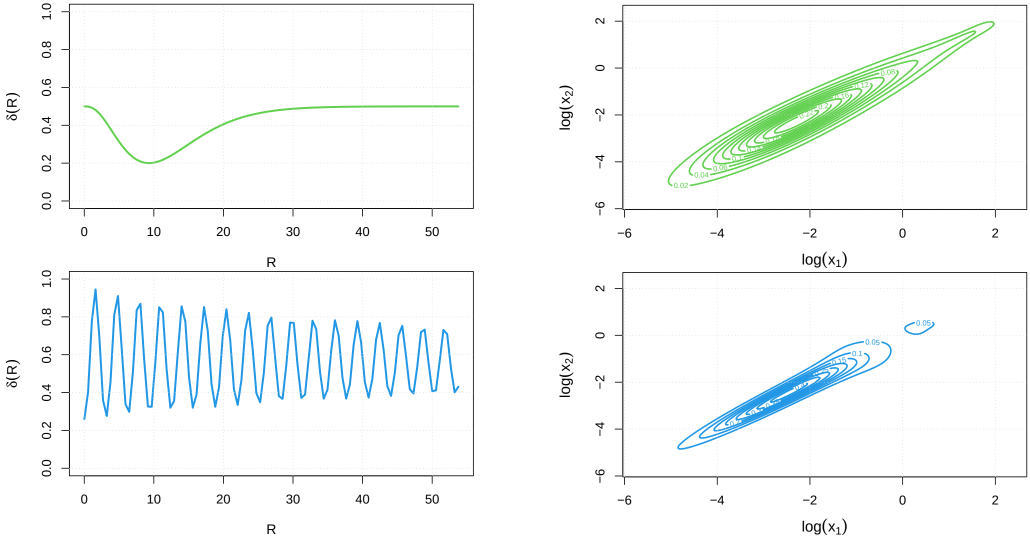
$$U_1 | \|\boldsymbol{X}\| = r \sim f(\cdot; \delta(r))$$

  under the constraint $\mathbb{E}\left(U_1 | \|\boldsymbol{X}\| = r\right) = 1/2$
- Why Gaussian model? Flexible multivariate distribution that is easy to specify and estimate!

FLEXIBILITY (I)



FLEXIBILITY (II)



Does this work in practice?

ESTIMATION IN TWO STEPS: FIRST STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $r_i = x_{i,1} + x_{i,2}$

1. Maximize the EGPD log-likelihood

$$l_R(\kappa, \xi) = \sum_{i=1}^{n} \left\{ \log \kappa + (\kappa - 1) \log H_\xi(r_i) + \log h_\xi(r_i) + \log \widehat{b}(H_\xi(r_i)^\kappa) \right\}.$$

- Density $b(u)$ is approximated with **Bernstein polynomials**

$$\widehat{b}(u) = m \sum_{k=1}^{m} \omega_{k,m} \beta_{k,m-k+1}(u)$$

  with $\beta_{i,j}(u) = \binom{j}{i} u^i (1-u)^{j-i}$ and $\omega_{k,m} = \mathbb{B}_n(k/m) - \mathbb{B}_n((k-1)/m)$ [$\mathbb{B}_n$ ECDF of $b_i = H_\xi(\|\boldsymbol{x}_i\|)^\kappa$]

- **Ad-hoc R program but simple to write!**

```
evd::dgpd(x = x,loc = 0,scale = 1,shape = xi)
evd::pgpd(x = x,loc = 0,scale = 1,shape = xi)
ecdf(b)
dbeta(u,shape1 = i, shape2 = j-i+1)
```

ESTIMATION IN TWO STEPS: FIRST STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $r_i = x_{i,1} + x_{i,2}$

1. Maximize the EGPD log-likelihood

$$l_R(\kappa, \xi) = \sum_{i=1}^{n} \left\{ \log \kappa + (\kappa - 1) \log H_\xi(r_i) + \log h_\xi(r_i) + \log \widehat{b}(H_\xi(r_i)^\kappa) \right\}.$$

- Density $b(u)$ is approximated with **Bernstein polynomials**

$$\widehat{b}(u) = m \sum_{k=1}^{m} \omega_{k,m} \beta_{k,m-k+1}(u)$$

with $\beta_{i,j}(u) = \binom{j}{i} u^i (1-u)^{j-i}$ and $\omega_{k,m} = \mathbb{B}_n(k/m) - \mathbb{B}_n((k-1)/m)$ [$\mathbb{B}_n$ ECDF of $b_i = H_\xi(\|\boldsymbol{x}_i\|)^\kappa$]

- **Ad-hoc R program but simple to write!**

```
evd::dgpd(x = x,loc = 0,scale = 1,shape = xi)
evd::pgpd(x = x,loc = 0,scale = 1,shape = xi)
ecdf(b)
dbeta(u,shape1 = i, shape2 = j-i+1)
```

ESTIMATION IN TWO STEPS: FIRST STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $r_i = x_{i,1} + x_{i,2}$

1. Maximize the EGPD log-likelihood

$$l_R(\kappa, \xi) = \sum_{i=1}^{n} \left\{ \log \kappa + (\kappa - 1) \log H_\xi(r_i) + \log h_\xi(r_i) + \log \widehat{b}(H_\xi(r_i)^\kappa) \right\}.$$

- Density $b(u)$ is approximated with **Bernstein polynomials**

$$\widehat{b}(u) = m \sum_{k=1}^{m} \omega_{k,m} \beta_{k,m-k+1}(u)$$

with $\beta_{i,j}(u) = \binom{j}{i} u^i (1-u)^{j-i}$ and $\omega_{k,m} = \mathbb{B}_n(k/m) - \mathbb{B}_n((k-1)/m)$ $[\mathbb{B}_n$ ECDF of $b_i = H_\xi(\|\boldsymbol{x}_i\|)^\kappa]$

- **Ad-hoc R program but simple to write!**

```
evd::dgpd(x = x,loc = 0,scale = 1,shape = xi)
evd::pgpd(x = x,loc = 0,scale = 1,shape = xi)
ecdf(b)
dbeta(u,shape1 = i, shape2 = j-i+1)
```

ESTIMATION IN TWO STEPS: FIRST STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $r_i = x_{i,1} + x_{i,2}$

1. Maximize the EGPD log-likelihood

$$l_R(\kappa, \xi) = \sum_{i=1}^{n} \left\{ \log \kappa + (\kappa - 1) \log H_\xi (r_i) + \log h_\xi (r_i) + \log \widehat{b}(H_\xi(r_i)^\kappa) \right\}.$$

- Density $b(u)$ is approximated with **Bernstein polynomials**

$$\widehat{b}(u) = m \sum_{k=1}^{m} \omega_{k,m} \beta_{k,m-k+1}(u)$$

with $\beta_{i,j}(u) = \binom{j}{i} u^i (1-u)^{j-i}$ and $\omega_{k,m} = \mathbb{B}_n(k/m) - \mathbb{B}_n((k-1)/m)$ [$\mathbb{B}_n$ ECDF of $b_i = H_\xi(\|\boldsymbol{x}_i\|)^\kappa$]

- **Ad-hoc R program but simple to write!**

```
evd::dgpd(x = x,loc = 0,scale = 1,shape = xi)
evd::pgpd(x = x,loc = 0,scale = 1,shape = xi)
ecdf(b)
dbeta(u,shape1 = i, shape2 = j-i+1)
```

ESTIMATION IN TWO STEPS: SECOND STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $v_i = \log(x_{i,1}) - \log(x_{i,2})$

2. Maximize the penalized Gaussian log-likelihood

$$PL_V(\gamma) = -\sum_{i=1}^{n} \left\{ \log(\delta(r_i)) + 0.5 \left(\frac{v_i}{\delta(r_i)} \right)^2 \right\} + \lambda \boldsymbol{\gamma}^\top \boldsymbol{P} \boldsymbol{\gamma}.$$

   * Linear combination of K basis functions $S_j(r)$ (cubic splines)

$$\log \delta(r) = \gamma_0 + \sum_{j=1}^{K} \gamma_j S_j(r), \qquad \boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_K)^\top$$

   * $\lambda > 0$ smoothing parameter and $\boldsymbol{P}$ is a positive semi-definite matrix
   * **R code**

```
mgcv::gam(list(v~1,~s(r),method = "REML",family=gaulss())
```

ESTIMATION IN TWO STEPS: SECOND STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $v_i = \log(x_{i,1}) - \log(x_{i,2})$

2. Maximize the penalized Gaussian log-likelihood

$$PL_V(\gamma) = -\sum_{i=1}^{n} \left\{ \log(\delta(r_i)) + 0.5 \left(\frac{v_i}{\delta(r_i)} \right)^2 \right\} + \lambda \boldsymbol{\gamma}^\top \boldsymbol{P} \boldsymbol{\gamma}.$$

* Linear combination of K basis functions $S_j(r)$ (cubic splines)

$$\log \delta(r) = \gamma_0 + \sum_{j=1}^{K} \gamma_j S_j(r), \qquad \gamma = (\gamma_0, \ldots, \gamma_K)^\top$$

* $\lambda > 0$ smoothing parameter and $\boldsymbol{P}$ is a positive semi-definite matrix
* **R code**

```
mgcv::gam(list(v~1,~s(r),method = "REML",family=gaulss())
```

ESTIMATION IN TWO STEPS: SECOND STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $v_i = \log(x_{i,1}) - \log(x_{i,2})$

2. Maximize the penalized Gaussian log-likelihood

$$PL_V(\boldsymbol{\gamma}) = -\sum_{i=1}^{n} \left\{ \log(\delta(r_i)) + 0.5 \left(\frac{v_i}{\delta(r_i)} \right)^2 \right\} + \lambda \boldsymbol{\gamma}^\top \boldsymbol{P} \boldsymbol{\gamma}.$$

* Linear combination of K basis functions $S_j(r)$ (cubic splines)

$$\log \delta(r) = \gamma_0 + \sum_{j=1}^{K} \gamma_j S_j(r), \qquad \boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_K)^\top$$

* $\lambda > 0$ smoothing parameter and $\boldsymbol{P}$ is a positive semi-definite matrix
* **R code**

```
mgcv::gam(list(v~1,~s(r),method = "REML",family=gaulss())
```

ESTIMATION IN TWO STEPS: SECOND STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $v_i = \log(x_{i,1}) - \log(x_{i,2})$

2. Maximize the penalized Gaussian log-likelihood

$$PL_V(\gamma) = -\sum_{i=1}^{n} \left\{ \log(\delta(r_i)) + 0.5 \left(\frac{v_i}{\delta(r_i)} \right)^2 \right\} + \lambda \boldsymbol{\gamma}^\top \boldsymbol{P} \boldsymbol{\gamma}.$$

* Linear combination of K basis functions $S_j(r)$ (cubic splines)

$$\log \delta(r) = \gamma_0 + \sum_{j=1}^{K} \gamma_j S_j(r), \qquad \boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_K)^\top$$

* $\lambda > 0$ smoothing parameter and $\boldsymbol{P}$ is a positive semi-definite matrix
* **R code**

```
mgcv::gam(list(v~1,~s(r),method = "REML",family=gaulss())
```

ESTIMATION IN TWO STEPS: SECOND STEP

Transform $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})^T$ into $v_i = \log(x_{i,1}) - \log(x_{i,2})$

2. Maximize the penalized Gaussian log-likelihood

$$PL_V(\boldsymbol{\gamma}) = -\sum_{i=1}^{n} \left\{ \log(\delta(r_i)) + 0.5 \left(\frac{v_i}{\delta(r_i)} \right)^2 \right\} + \lambda \boldsymbol{\gamma}^\top \boldsymbol{P} \boldsymbol{\gamma}.$$

   * Linear combination of K basis functions $S_j(r)$ (cubic splines)

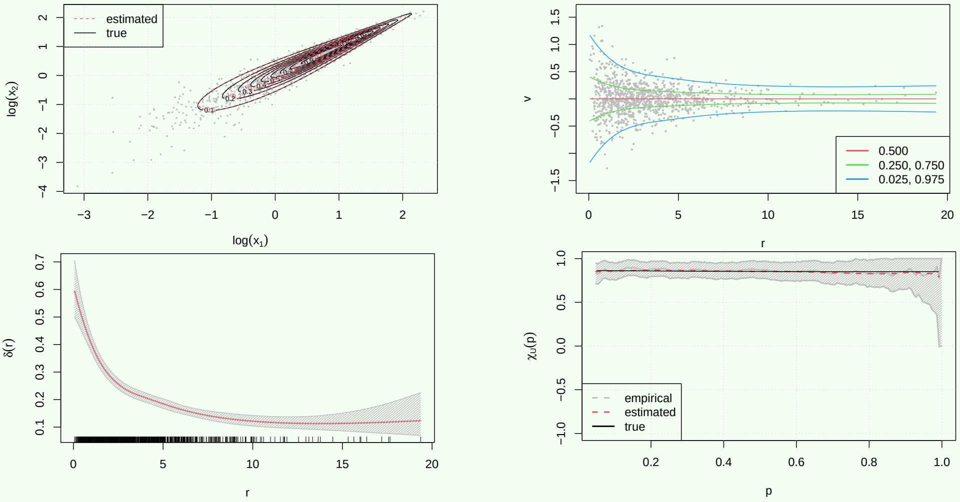
$$\log \delta(r) = \gamma_0 + \sum_{j=1}^{K} \gamma_j S_j(r), \qquad \boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_K)^\top$$

   * $\lambda > 0$ smoothing parameter and $\boldsymbol{P}$ is a positive semi-definite matrix
   * **R code**

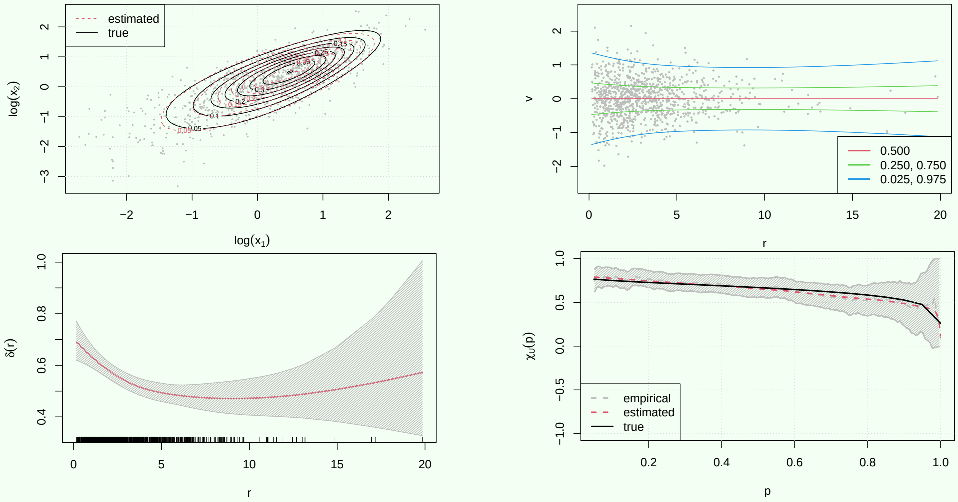
```
mgcv::gam(list(v~1,~s(r),method = "REML",family=gaulss())
```

Can we approximate common (bivariate) copula models ?

MRV: SYMMETRIC LOGISTIC COPULA

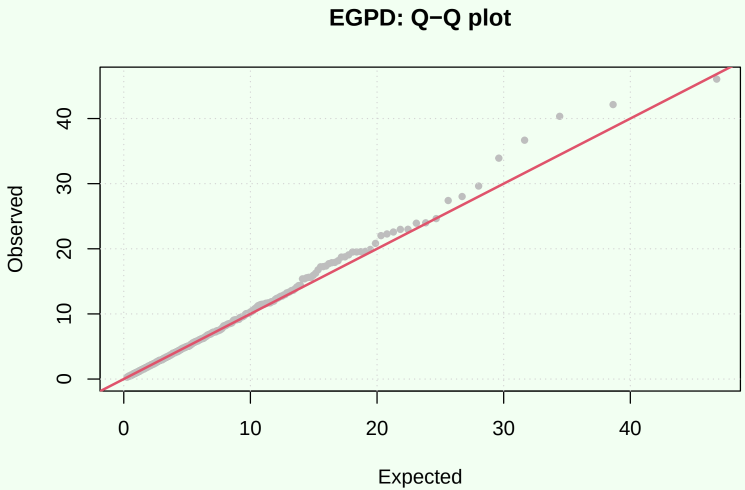


NO MRV: GAUSSIAN COPULA



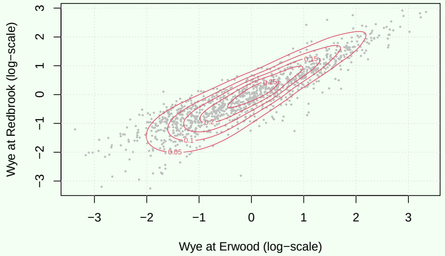
Coming back to the real data example ...

RIVER DISCHARGES: DISTRIBUTION OF $\|\boldsymbol{X}\| = X_1 + X_2$

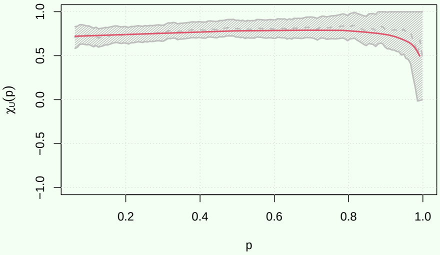


RIVER DISCHARGES: DISTRIBUTION OF $\boldsymbol{X} = (X_1, X_2)^T$

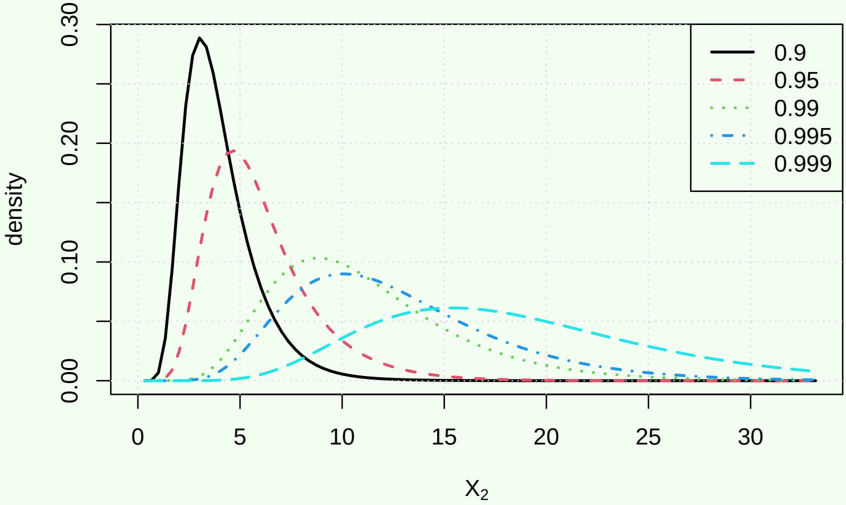
Bivariate density



Goodness of fit



RIVER DISCHARGES: DISTRIBUTION OF $X_2|X_1$



TAKE HOME MESSAGE

A multivariate EGPD exists

RELATED WORKS I

- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve
- Rainfall measurement are discrete ...
- ...and most of the time, it is not raining, i.e. zero inflation !

RELATED WORKS I

- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve
- Rainfall measurement are discrete ...
- ...and most of the time, it is not raining, i.e. zero inflation !

RELATED WORKS I

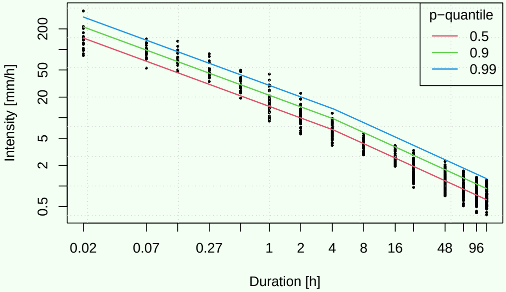
- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve
- Rainfall measurement are discrete ...
- ...and most of the time, it is not raining, i.e. zero inflation !

RELATED WORKS I

- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve
- Rainfall measurement are discrete ...
- ...and most of the time, it is not raining, i.e. zero inflation !

RELATED WORKS I

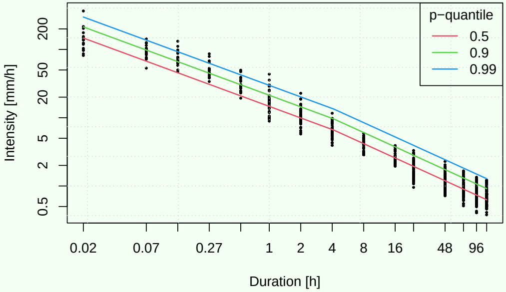
- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve



- Rainfall measurement are discrete ...

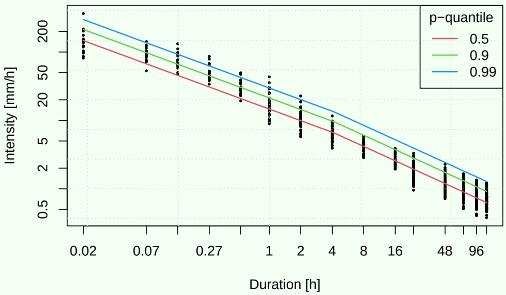
RELATED WORKS I

- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve



- Rainfall measurement are discrete ...

- Ailliot, P., **Gaetan C. and Naveau, P.**, (2025+) A parsimonious tail compliant multiscale statistical model for aggregated rainfall, *submitted*.
- $X_1 + \ldots + X_d \sim EGPD(\kappa, \xi, B_d)$
- Duration d
- Intensity Duration Frequency (IDF) curve



- Rainfall measurement are discrete ...

RELATED WORKS II

Cornell University

arXiv > stat > arXiv:2509.05982

Statistics > Methodology

[Submitted on 7 Sep 2025]

Joint modeling of low and high extremes using a multivariate extended generalized Pareto distribution

Noura Alotaibi, Matthew Sainsbury-Dale, Philippe Naveau, Carlo Gaetan, Raphaël Huser

- Weighted sum of latent variables
- Amortized neural inference approach

RELATED WORKS II

Cornell University

arXiv > stat > arXiv:2509.05982

Statistics > Methodology

[Submitted on 7 Sep 2025]

Joint modeling of low and high extremes using a multivariate extended generalized Pareto distribution

Noura Alotaibi, Matthew Sainsbury-Dale, Philippe Naveau, Carlo Gaetan, Raphaël Huser

- Weighted sum of latent variables
- Amortized neural inference approach

RELATED WORKS II

Cornell University

arXiv > stat > arXiv:2509.05982

Statistics > Methodology

[Submitted on 7 Sep 2025]

Joint modeling of low and high extremes using a multivariate extended generalized Pareto distribution

Noura Alotaibi, Matthew Sainsbury-Dale, Philippe Naveau, Carlo Gaetan, Raphaël Huser

- Weighted sum of latent variables
- Amortized neural inference approach

Thanks !!!

Merci !!!

Let $\boldsymbol{X} = \|\mathbf{X}\| \times \boldsymbol{U}$ be a random vector that satisfies (1) and (2).
We say that $\mathbf{X}$ follows a **multivariate logistic-heteroscedastic EGPD** if the log-ratio of its angular component $\boldsymbol{U}$ can be expressed, given the radius $\|\mathbf{X}\| = r$, as

$$V_i := \log(U_i/U_d) = \delta(r)\, Z_i, \ \text{ for } i = 1, \ldots, (d-1), \tag{3}$$

where the (d-1) dimensional vector $\boldsymbol{Z} = (Z_1, \ldots, Z_{d-1})^\top$ is a **zero-mean exchangeable random vector** independent of $\|\mathbf{X}\|$ and $\delta(\cdot)$ is a positive measurable function such that, uniformly on any compact of the real line,

$$\lim_{r \to 0^+} \delta(r) = \delta_- \text{ and } \lim_{r \to \infty} \delta(r) = \delta_+, \tag{4}$$

for some finite positive constants δ_- and δ_+.