Insight into spatial extremes

Marine Demangeot
IMAG – Université de Montpellier

Extreme values theory

deals with **extreme events**

Extreme event : ***rare*** *event which takes* ***very low or very large values***

Floods in Pakistan – *AFP*
(June - Oct. 2022)

$\approx$ 1 739 people killed
$\geq$ 2 million people homeless
\$14.9 billion of damage

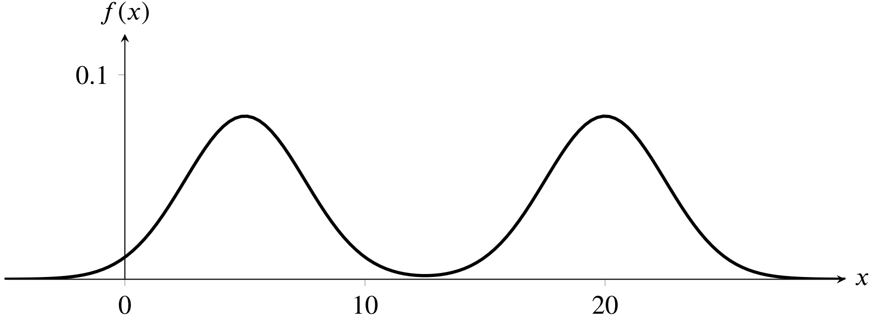
Wildfires in Gironde – AFP
(July 2022)

$\geq$ 20 800 hectares burnt
$\approx$ 36 000 people evacuated

Extreme values theory   deals with **extreme events**

Extreme event : ***rare*** *event which takes* ***very low or very large values***

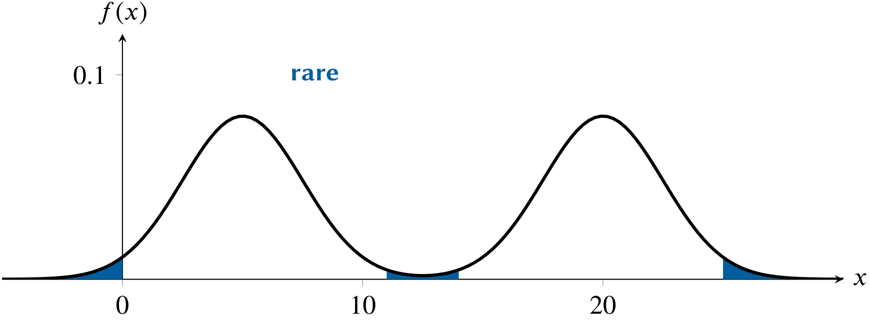
- phenomenon of interest $X \sim f$ (density)



Extreme values theory deals with **extreme events**

Extreme event *:* ***rare*** *event which takes* ***very low or very large values***

- phenomenon of interest $X \sim f$ (density)

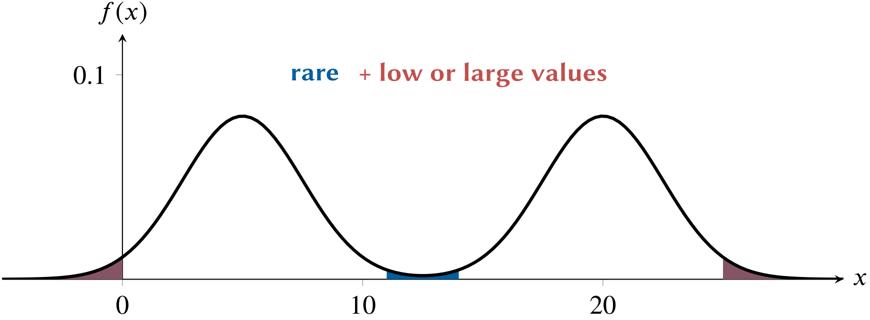


Extreme values theory

deals with **extreme events**

Extreme event** : **rare** event which takes **very low or very large values

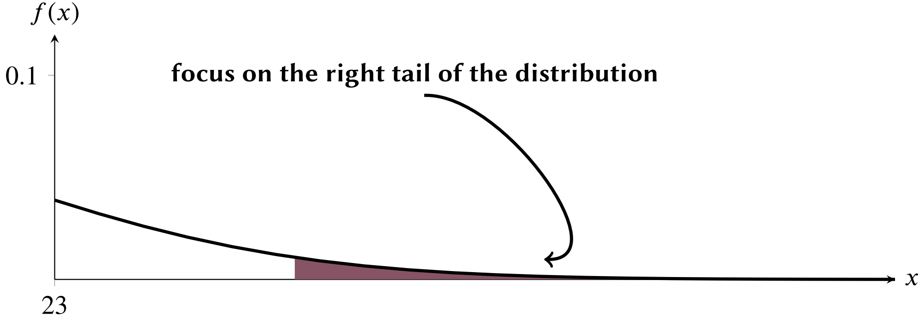
- phenomenon of interest $X \sim f$ (density)



Extreme values theory deals with **extreme events**

Extreme event : ***rare*** *event which takes* ***very low or very large values***

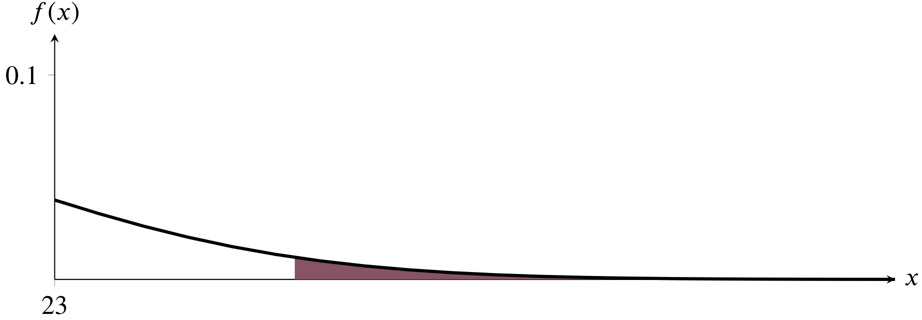
- phenomenon of interest $X \sim f$ (density)



Extreme values theory deals with **extreme events**

Extreme event *:* ***rare*** *event which takes* ***very low or very large values***

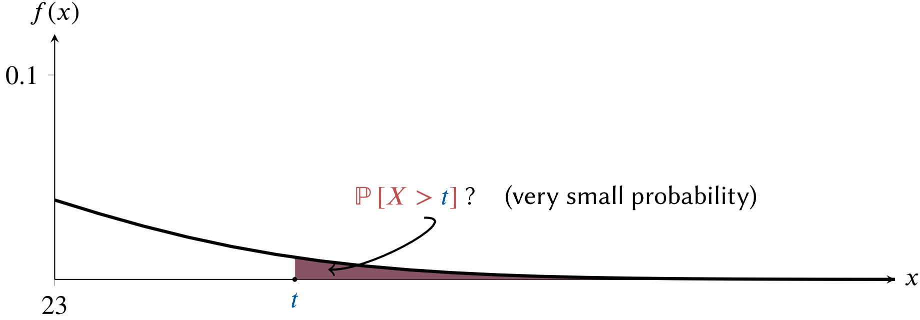
- X : **Daily amount of precipitation at a particular location**



Extreme values theory deals with **extreme events**

> ***Extreme event*** : ***rare*** *event which takes* ***very low or very large values***

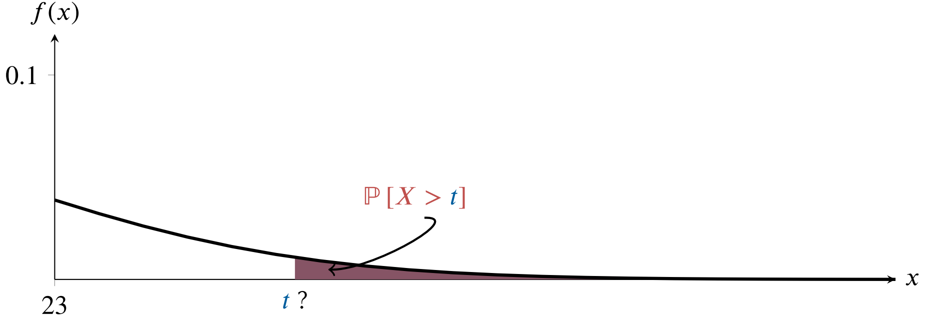
- X : **Daily amount of precipitation at a particular location**



Extreme values theory deals with **extreme events**

***Extreme event* : rare *event which takes* very low or very large values**

- X : **Daily amount of precipitation at a particular location**



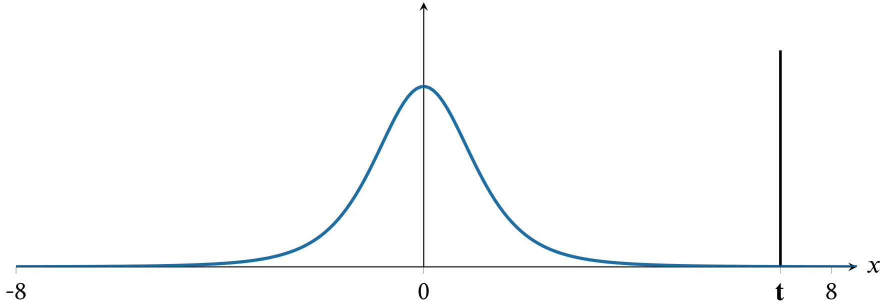
Extreme values theory deals with **extreme events**

Extreme event : ***rare*** *event which takes* ***very low or very large values***

Q Why do we need a specific theory to study extreme values ?

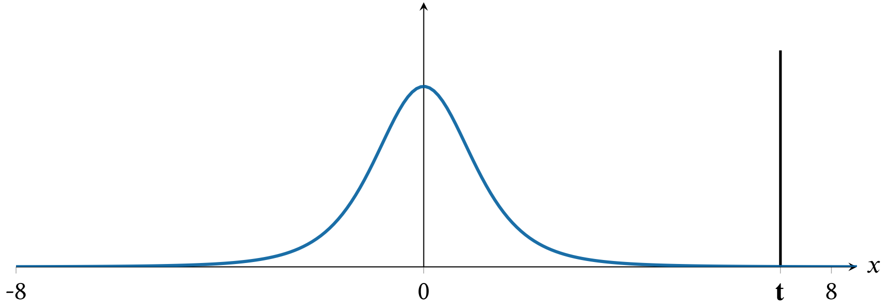
The issue of extreme events

$X \sim f$ (pdf) $\{X > \mathbf{t}\}$ $\mathbf{t}$ large



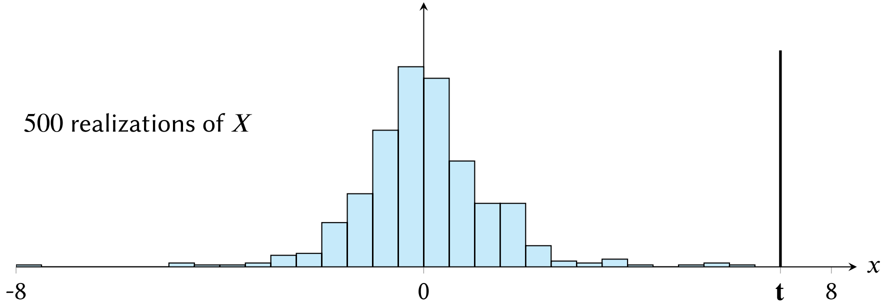
The issue of extreme events

$X \sim f$ (pdf) $\mathbf{P}[X > \mathbf{t}]$? $\mathbf{t}$ large



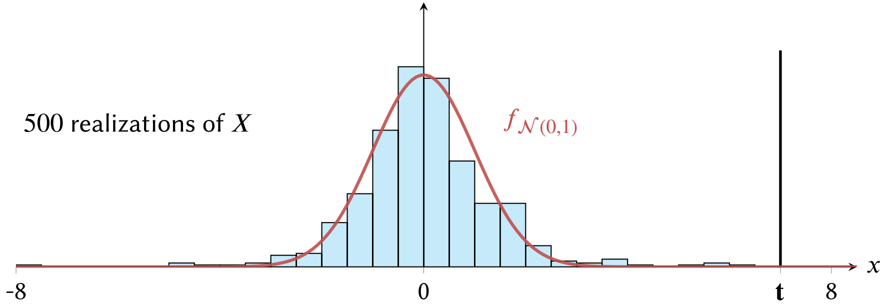
The issue of extreme events

$X \sim f$ (pdf) $\qquad$ $\mathbf{P}[X > \mathbf{t}]$? $\qquad$ $\mathbf{t}$ large



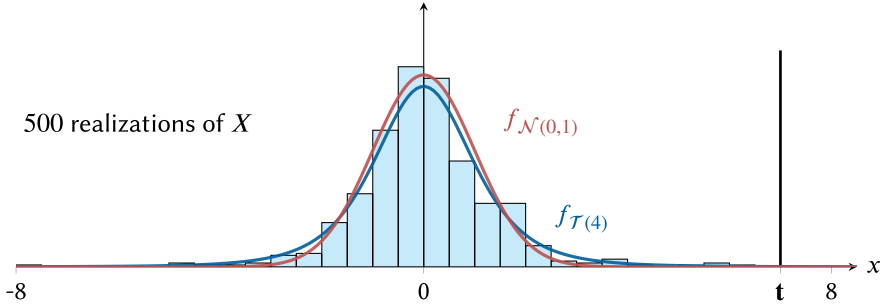
The issue of extreme events

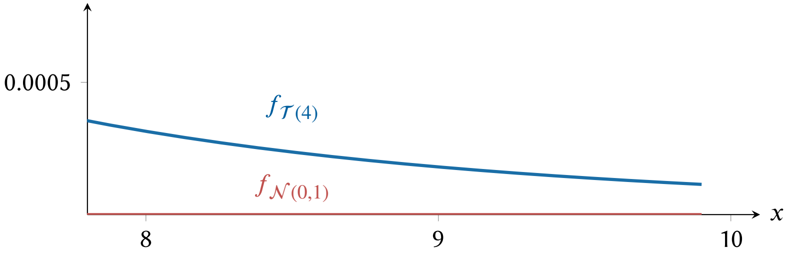
$X \sim f$ (pdf) $\mathbf{P}[X > \mathbf{t}]$? $\mathbf{t}$ large

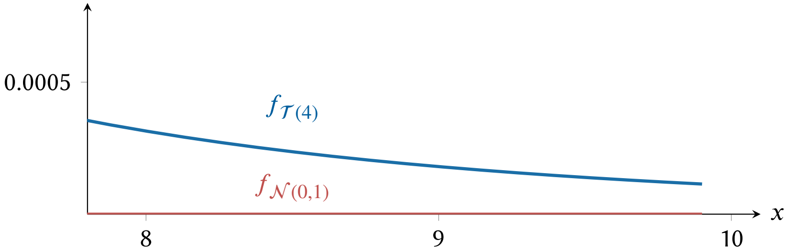


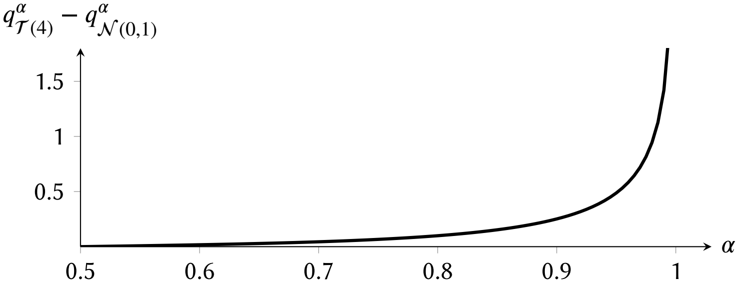
The issue of extreme events

$X \sim f$ (pdf) $\mathbf{P}[X > \mathbf{t}]$? $\mathbf{t}$ large



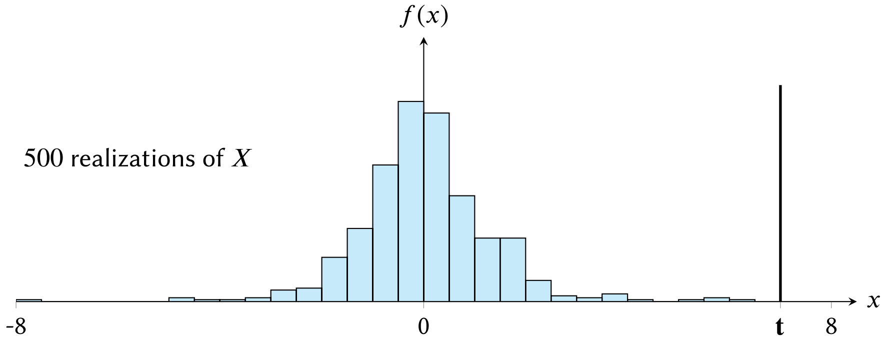

0.0005
$f_{\mathcal{T}(4)}$
$f_{\mathcal{N}(0,1)}$
x
8
9
10


0.0005
$f_{\mathcal{T}(4)}$
$f_{\mathcal{N}(0,1)}$
8
9
10
x


$q^{\alpha}_{\mathcal{T}(4)} - q^{\alpha}_{\mathcal{N}(0,1)}$
1.5
1
0.5
0.5
0.6
0.7
0.8
0.9
1
α

The issue of extreme events

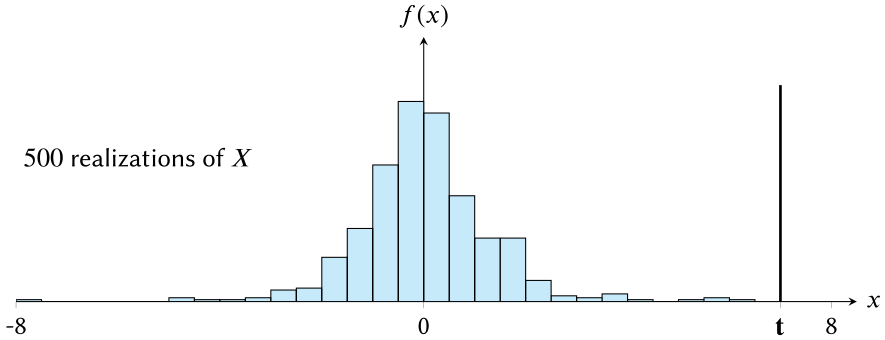
$X \sim f$ (pdf) $\quad \mathbf{P}[X > \mathbf{t}]$? $\quad \mathbf{t}$ large



Parametric estimation : risk of underestimating $\mathbf{P}[X > \mathbf{t}]$

The issue of extreme events

$X \sim f$ (pdf) $\qquad \mathbf{P}[X > \mathbf{t}]$? $\qquad \mathbf{t}$ large

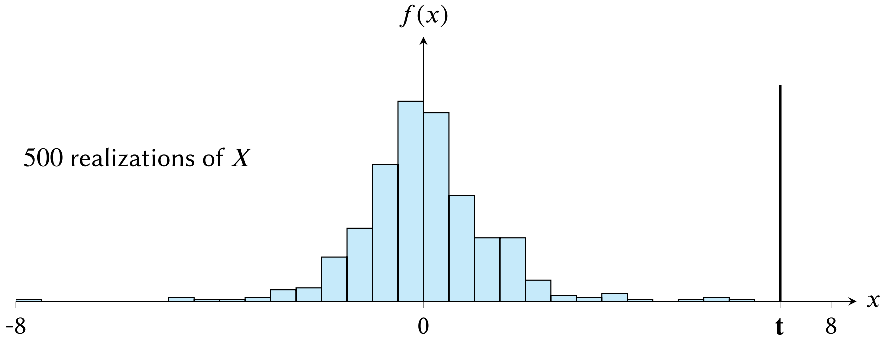


Parametric estimation : risk of underestimating $\mathbf{P}[X > \mathbf{t}]$

Non-parametric estimation : risk of underestimating $\mathbf{P}[X > \mathbf{t}]$

The issue of extreme events

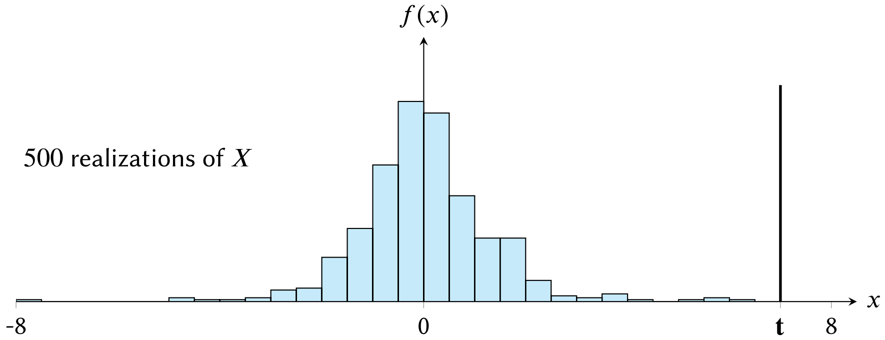
$X \sim f$ (pdf) $\quad\quad$ $\mathbf{P}[X > \mathbf{t}]$? $\quad\quad$ $\mathbf{t}$ large



Extrapolation of the tail behaviour beyond the range of the data

The issue of extreme events

$X \sim f$ (pdf) $\quad$ $\mathbf{P}[X > \mathbf{t}]$? $\quad$ $\mathbf{t}$ large

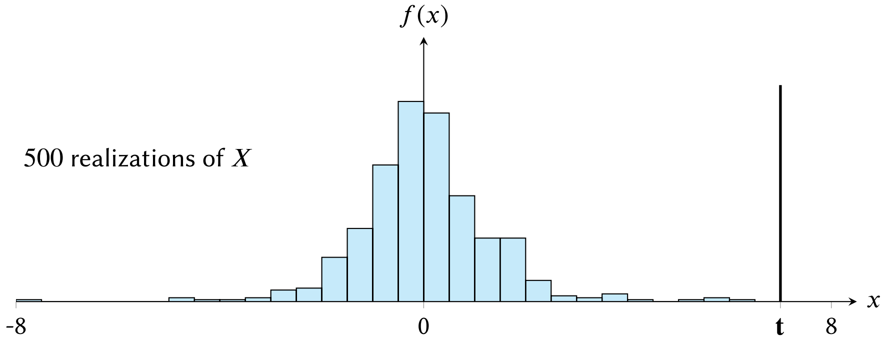


Extrapolation of the tail behaviour beyond the range of the data

Extreme value theory

The issue of extreme events

$X \sim f$ (pdf) $\mathbf{P}[X > \mathbf{t}]$? $\mathbf{t}$ large



Extrapolation of the tail behaviour beyond the range of the data

1 Study of the maximum **2 Study of threshold exceedances**

A Univariate extremes

Study of the maximum

$X \sim F$ random variable

F cumulative distribution function (c.d.f.)

$X_1, \ldots, X_n$ copies i.i.d. of X

Objective Study of $\max_{i=1,\ldots,n} X_i$

$$\forall x \in \mathbb{R} \quad \mathbf{P}\left[\max_{i=1,\ldots,n} X_i \leq x\right] = \mathbf{P}\left[X_1 \leq x\right] \times \cdots \times \mathbf{P}\left[X_n \leq x\right] = (F(x))^n$$

Objective Approximation of $\mathcal{L}\left(\max_{i=1,\ldots,n} X_i\right)$ as $n \to +\infty$

Limit law for the maximum

Theorem *Fisher & Tippett (1928) Gnedenko (1943)*

*Let $\left(X_i\right)_{i\in\mathbb{N}^*}$ be i.i.d. random variables . **If** for any $n \in \mathbb{N}^*$, it exists $a_n > 0$ and $b_n \in \mathbb{R}$ s.t.*

$$\max_{i=1,\dots,n} \frac{X_i - b_n}{a_n} \underset{n\to\infty}{\overset{d}{\to}} Z \qquad \textit{(non degenerate)}$$

then** Z is **max-stable

$Z, Z_1, \dots, Z_n$ i.i.d. random variables (non degenerated)

Definition Z is **max-stable** for any $n \in \mathbb{N}^*$, it exists $\alpha_n > 0$ and $\beta_n \in \mathbb{R}$ s.t.

$$\max_{i=1,\dots,n} \frac{Z_i - \beta_n}{\alpha_n} \underset{n\to\infty}{\overset{d}{\to}} Z$$

Limit law for the maximum

Theorem

Fisher & Tippett (1928) Gnedenko (1943)

*Let $(X_i)_{i \in \mathbb{N}^*}$ be i.i.d. random variables . **If** for any $n \in \mathbb{N}^*$, it exists $a_n > 0$ and $b_n \in \mathbb{R}$ s.t.*

$$\max_{i=1,\ldots,n} \frac{X_i - b_n}{a_n} \underset{n\to\infty}{\overset{d}{\rightarrow}} Z \qquad \textit{(non degenerate)}$$

then** Z is **max-stable

Characterization

$Z \sim G$ (c.d.f.) is max-stable if and only if

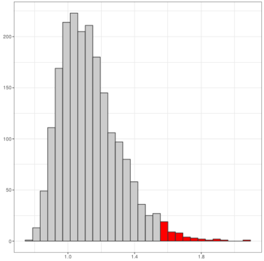
$$G(x) = G_\xi(x) := \exp\left\{-\left(1 + \xi \frac{x-\beta}{\alpha}\right)^{-1/\xi}\right\} \qquad x \in \mathbb{R}^d$$

$\xi \in \mathbb{R} \quad \alpha > 0 \quad \beta \in \mathbb{R} \quad 1 + \xi x > 0$ by convention $G_0(x) = \exp\{-e^{-x}\}$

ξ **extreme value index**

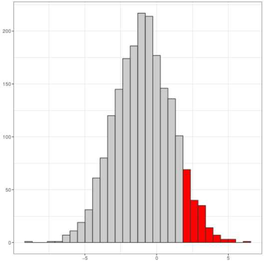
Remark **The tail index determines the tail decay rate**

Heavy tail
(polynomial decay)
$\xi > 0$



➤ Pareto, Student's t distributions

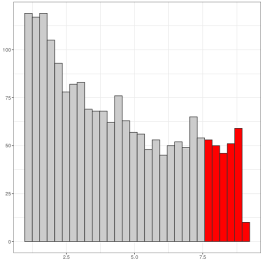
Light tail
(exponential decay)
$\xi = 0$



T. Opitz - ATHENS courses 2023

➤ Normal, exponential, gamma distributions

Bounded tail
$\xi < 0$



➤ Uniform, beta distributions

Estimation

For n large enough,

$$\max_{i=1,\ldots,n} \frac{X_i - b_n}{a_n} \overset{distribution}{\approx} Z \sim G_\xi$$

with G_ξ a max-stable distribution

$$\Rightarrow \quad \mathbb{P}\left[X > x\right] \approx -\frac{1}{n} \log\left\{G_\xi\left(\frac{x - b_n}{a_n}\right)\right\} \qquad \text{for } n \text{ large}$$

Objective Estimation of ξ, a_n, b_n

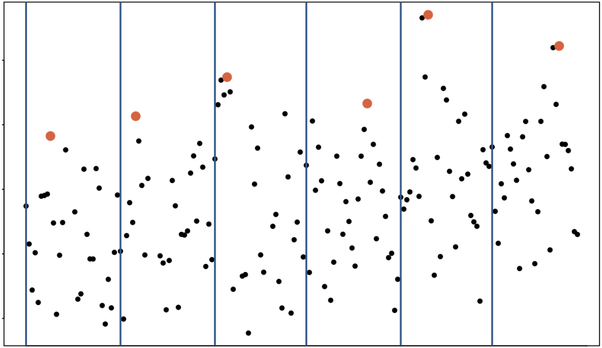
G_ξ has a parametric form : **maximum likelihood estimation** of ξ, a_n and b_n
ex : *ismev* package (R)

How to get i.i.d. replications of $\max_{i=1,\ldots,n} X_i$?

Block maxima method

Dividing the original sample of size n in k blocks of size m to have (approximately) k i.i.d. maxima

➤ e.g. annual maxima



E. Chautru - ATHENS courses 2023

Bais-variance trade-off between the choice of k and m

$m = m(n) \to +\infty$ and $m/n \to 0$ as $n \to +\infty$

Return levels

From the limit distribution of the maximum, we can compute $\mathbb{P}\left[X > x\right]$ but we can also compute **return levels**

- k blocks of size m : k i.i.d. replications of $\max\limits_{i=1,\dots,m} X_i$

Definition Let $K \in \mathbb{N}^*$. The K $m-$block return level $\boldsymbol{R_{m,K}}$ is the quantile associated to the probability $1/K$:

$$\mathbb{P}\left[\max_{i=1,\dots,m} X_i > \boldsymbol{R_{m,K}}\right] = \frac{1}{K}$$

i.e. it is the level which is exceeded in one out of every m $n-$blocks, on average.

$$R_{m,K} \approx G_\xi^{-1}\left(\left(1 - \frac{1}{K} - b_m\right)/a_m\right)$$

Exemple We observe daily rainfall in Brest for many years. We compute the annual maximum for each year ($m = 365$). If $K = 1000$ then the return level $R_{365,1000}$ represent the amount of rainfall that will be observed once every 1000 years, on average.

Remark

Loss of information when considering only maxima.

We want to take into account the largest order statistics when analysing the tail of the distribution : study of threshold exceedances

Limit law for exceedances

- X a random variable with $x^* := \sup\{x \in \mathbb{R} : F(x) < 1\} \in \mathbb{R} \cup \{+\infty\}$ the upper endpoint of its distribution

Theorem *Pickands (1975) Balkema & de Hann (1974)*

Let $\left(X_i\right)_{i \in \mathbb{N}^*}$ be independent copies of X. The two following propositions are equivalent

- $\forall n \in \mathbb{N}^*$, *it exists $a_n > 0$ and $b_n \in \mathbb{R}$ s.t.*

$$\max_{i=1,\ldots,n} \frac{X_i - b_n}{a_n} \xrightarrow[n\to\infty]{d} Z \qquad \textit{max-stable}$$

- *There exists a scaling function $m : \mathbb{R} \to (0, +\infty)$ such that*

$$\frac{X - t}{m(t)} \mid X > t \xrightarrow[t \to x^*]{} Y$$

where Y follow a ***generalized Pareto distribution***

Generalized Pareto distribution

Definition $Y \sim H$ (c.d.f.) has a generalized Pareto distribution if and only if

$$H(x) = H_\xi(x) := 1 - \left(1 + \xi\frac{x}{\sigma}\right)^{-1/\xi} \qquad x \in (0, +\infty) \text{ s.t. } 1 + \xi x > 0$$

$\xi \in \mathbb{R}$ extreme value index

$\sigma \in (0, +\infty)$ (scale parameter)

by convention $H_0 = \exp\{-x/\sigma\}$

Threshold stability Let y^* the upper endpoint of H_ξ then

$Y - t' | Y > t'$ has a generalized Pareto distribution too

Estimation

For t large enough,

$$\frac{X-t}{m(t)} \mid X > t \overset{distribution}{\approx} Y \sim H_\xi$$

with H_ξ a generalized Pareto distribution

$$\Rightarrow \quad \forall x > t \quad \mathbb{P}[X > x] \approx \mathbb{P}[X > t] \, H_\xi\left(\frac{x-t}{m(t)}\right)$$

Objective Estimation of ξ, $m(t)$, $\mathbb{P}[X > t]$

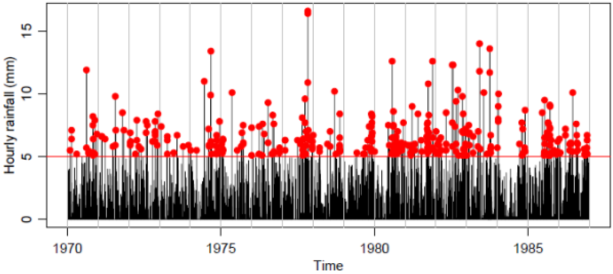
G_ξ has a parametric form : **maximum likelihood estimation** of ξ, $m(t)$
(and empirical estimation of $\mathbb{P}[X > t]$)
ex : *ismev* package (R)

How to get i.i.d. replications of $X - t | X > t$?

Peak over threshold method

Set a high enough threshold t and select the observations exceeding t

Remark threshold exceedances that occur in groups $\longrightarrow$ declustering to obtain a set of threshold excesses that are approximately independent

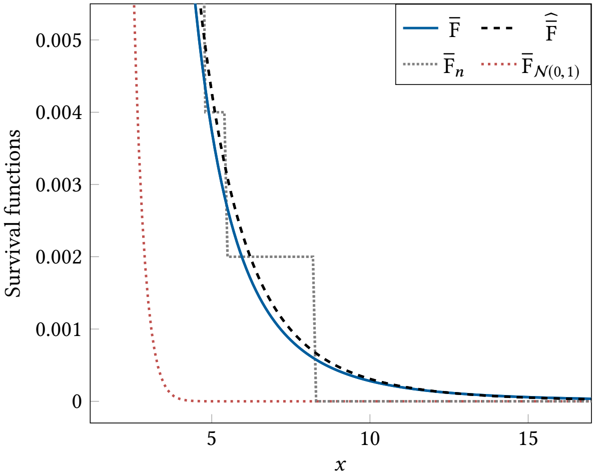


J. Koh - ATHENS courses 2023

Bais-variance trade-off for the choice of the threshold t (there exist methods to help select t)

Illustration

- $x_1, \ldots, x_{500}$, 500 realizations of $X \sim \mathcal{T}(4)$



B

Spatial extremes

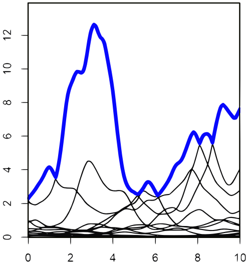
Spatial maxima

Z **random field** (RF) on $\mathbb{R}^d$ with non-degenerate margins

$Z_1, \ldots, Z_n$ i.i.d. copies of Z

Definition Z is **max-stable** if for any $n \in \mathbb{N}^*$, $\exists\, \alpha_n(.) > 0$ and $\beta_n(.) \in \mathbb{R}$ s.t.

$$\left(\max_{i=1,\ldots,n} \frac{Z_i(x) - \beta_n(x)}{\alpha_n(x)} \right)_{x \in \mathbb{R}^d} \overset{f.d.d.}{=} (Z(x))_{x \in \mathbb{R}^d}$$



M. Oesting - Workshop EPFL 2017

Spatial maxima

Z **random field** (RF) on $\mathbb{R}^d$
with non-degenerate margins

$Z_1, \dots, Z_n$ i.i.d. copies of Z

Definition Z is **max-stable** if for any $n \in \mathbb{N}^*$, $\exists\, \alpha_n(.) > 0$ and $\beta_n(.) \in \mathbb{R}$ s.t.

$$\left(\max_{i=1,\dots,n} \frac{Z_i(x) - \beta_n(x)}{\alpha_n(x)} \right)_{x \in \mathbb{R}^d} \overset{f.d.d.}{=} (Z(x))_{x \in \mathbb{R}^d}$$

Remark The margins of Z are max-stable

Spatial maxima

Z **random field** (RF) on $\mathbb{R}^d$
with non-degenerate margins

$Z_1, \dots, Z_n$ i.i.d. copies of Z

Definition Z is **max-stable** if for any $n \in \mathbb{N}^*$, $\exists\, \alpha_n(.) > 0$ and $\beta_n(.) \in \mathbb{R}$ s.t.

$$\left(\max_{i=1,\dots,n} \frac{Z_i(x) - \beta_n(x)}{\alpha_n(x)} \right)_{x \in \mathbb{R}^d} \overset{f.d.d.}{=} (Z(x))_{x \in \mathbb{R}^d}$$

Remark The margins of Z are max-stable

Theorem

Let $\left(X_i\right)_{i \in \mathbb{N}^*}$ be i.i.d. RFs. ***If*** *for any $n \in \mathbb{N}^*$, $\exists\, a_n(.) > 0$ and $b_n(.) \in \mathbb{R}$ s.t.*

$$\left(\max_{i=1,\dots,n} \frac{X_i(x) - b_n(x)}{a_n(x)} \right)_{x \in \mathbb{R}^d} \xrightarrow[n \to \infty]{f.d.d.} (Z(x))_{x \in \mathbb{R}^d}$$

then *Z is a* ***max-stable RF***

Spectral representation

Max-stable margins: $\forall x \in \mathbb{R}^d \quad Z(x) \sim G_{\xi_x} \qquad \xi_x \in \mathbb{R}$

Spectral representation

Unit-Fréchet margins: $\forall x \in \mathbb{R}^d \quad \mathbf{P}\left[Z(x) \leq z\right] = e^{-1/z}\, \mathbf{1}_{\{z>0\}} \quad z \in (0, +\infty)$

$\gamma = 1$: heavy-tailed

Unit-Fréchet margins: $\forall x \in \mathbb{R}^d \quad \mathbf{P}[Z(x) \leq z] = e^{-1/z}\, \mathbf{1}_{\{z>0\}} \quad z \in (0, +\infty)$

$\gamma = 1$: heavy-tailed

Spectral representation

de Haan (1984) Giné et al. (1990)

If *Z is continuous in probability* ***then***

$$(Z(x))_{x\in\mathbb{R}^d} \overset{f.d.d.}{=} \left(\max_{T\,\in\,\Pi} \frac{Y_T(x)}{T} \right)_{x\in\mathbb{R}^d}$$

Π *homogeneous Poisson point process on* $(0, +\infty)$ *with unit rate*

Y_T *i.i.d copies of a continuous in probability and nonnegative RF Y with* $E[Y(x)] = 1$ *for any* $x \in \mathbb{R}^d$

Finite-dimensional distributions

- $x_1, \dots, x_k \in \mathbb{R}^d, \quad z_1, \dots, z_k \in (0, +\infty)$ $\qquad k \in \mathbb{N}^*$

From the spectral representation, we get an expression for the finite-dimensional distributions of Z :

$$\begin{aligned} \mathbf{P}\left[Z(x_1) \leq z_1, \dots, Z(x_k) \leq z_k\right] &= \exp\left\{ \boldsymbol{E}\left[\max_{i \in \{1,\dots,k\}} \frac{Y(x_i)}{z_i}\right]\right\} \\ &= \exp\left\{-\boldsymbol{V}_{\boldsymbol{x}}(\boldsymbol{z})\right\} \end{aligned}$$

$$x = \{x_1, \dots, x_k\} \quad z = \{z_1, \dots, z_k\}$$

The **exponent function $\boldsymbol{V}_{\boldsymbol{x}}(\boldsymbol{z})$** is a measure of the dependence structure of Z

No parametric characterization for the exponent function and more generally for the dependence structure of a max-stable RF. To estimate $\boldsymbol{V}_{\boldsymbol{x}}(\boldsymbol{z})$ with, e.g, a maximum likelihood method, we need to assume a parametric model for $\boldsymbol{V}_{\boldsymbol{x}}(\boldsymbol{z})$.

Full likelihood

Assumption $V_x(z) = V_x(z|\,\beta)$ (β : vector of parameters) and all derivatives of $V_x(z|\,\beta)$ exists

Density function of $Z(x_1), \ldots, Z(x_k)$

$$f_x(z) = \exp\left\{-V_x(z|\,\beta)\right\} \sum_{P \in \mathcal{P}_z} \prod_{S \in P} \left(-\frac{\partial}{\partial z_S} V_x(z|\,\beta)\right)$$

Main issues regarding the computation of $f_x(z)$:

- Closed-form expressions for $V_x(z|\,\beta)$ and $\dfrac{\partial}{\partial z_S} V_x(z|\,\beta)$ are not always available
- $2^n - 1$ (complex) partial derivatives must be computed
- $\mathrm{card}(\mathcal{P}_z)$ = Bell number which grows more than exponentially with n : storage of very large data structures is required. ex: $n = 10 \to 115\,000$ partitions

The maximum likelihood estimator cannot be used

Composite likelihood

Use a log-composite likelihood, *i.e.* a linear combination of (smaller) log-likelihood entities

- $Z_1,\dots,Z_n$, i.i.d. simple max-stable RFs with
- $\forall i \in \{1,\dots,n\} \quad z_i(x) = (z_i(x_1),\dots,z_i(x_k))$ realization of Z_i observed in locations $x_1,\dots,x_k$

Pairwise log-likelihood
Padoan et al (2010)

For max-stable processes, most of the time, only the bivariate densities can "easily" be derived analytically

$$\ell_p(z_1(x),\dots,z_n(x) \mid \beta) = \sum_{\ell=1}^{n} \sum_{1\le i<j\le k} w_{i,j} \log\left\{ f\left(z_\ell(x_i), z_\ell(x_j) \right) \right\}$$

where $\{w_{i,j}, 1 \le i < j \le k\}$ are suitable weights that improve efficiency or reduce the computational cost

Remark Under mild conditions the maximum pairwise likelihood estimator is consistent and asymptotically normal

Composite likelihood

Use a log-composite likelihood, *i.e.* a linear combination of (smaller) log-likelihood entities

- $Z_1, \ldots, Z_n$, i.i.d. simple max-stable RFs with
- $\forall i \in \{1, \ldots, n\} \quad z_i(x) = (z_i(x_1), \ldots, z_i(x_k))$ realization of Z_i observed in locations $x_1, \ldots, x_k$

Pairwise log-likelihood
Padoan et al (2010)

For max-stable processes, most of the time, only the bivariate densities can "easily" be derived analytically

$$\ell_p(z_1(x), \ldots, z_n(x) \mid \beta) = \sum_{\ell=1}^{n} \sum_{1 \leq i < j \leq k} w_{i,j} \log \left\{ f\left(z_\ell(x_i), z_\ell(x_j) \right) \right\}$$

where $\{w_{i,j}, 1 \leq i < j \leq k\}$ are suitable weights that improve efficiency or reduce the computational cost

Remark Pairwise log-likelihood can handle the estimation of both the margins and the dependence structure : better understanding of overall uncertainty even if it increases the computational complexity

Example of multivariate models

- k locations $x_1, \dots, x_k$ $\quad x = (x_1, \dots, x_k)$

Symmetric logistic model

for all $z = (z_1, \dots, z_k) \in (0, +\infty)^k$

$$V_x(z) = \left(\sum_{i=1}^{k} z_i^{-1/\alpha} \right)^{\alpha}$$

Only one parameter to estimate, α, but lack of flexibility

Asymmetric logistic model

for all $z = (z_1, \dots, z_k) \in (0, +\infty)^k$

$$V_x(z) = -\sum_{c \in C} \left(\sum_{i \in c} \left(\frac{w_{i,c}}{z_i} \right)^{\alpha_c} \right)^{1/\alpha_c}$$

where C is the set of all non-empty subsets of $\{1, \dots, k\}$ and with constraints on α_c and $w_{i,c}$

More flexible model but the number of parameters to be estimated is of order 2^k

Example of multivariate models

The multivariate models do not take into account the spatial features of the phenomenon under study

We can follow ideas from Geostatistics

Geostatistical context

Geostatistics regionalized phenomenon modelled by a random field (RF) Z

> ***In many cases, we partially observe only one realization of Z***

➤ mining resources estimation, soil contamination evaluation

Assumption Z is second-order stationary

$\forall x,\ h \in \mathbb{R}^d$

$$\begin{aligned} \mu &:= \mathbf{E}\left[Z(x)\right] < +\infty \\ \sigma^2 &:= \mathbf{Var}\left[Z(x)\right] < +\infty \\ \rho(h) &:= \mathbf{Corr}\left[Z(x), Z(x+h)\right] \end{aligned}$$

Geostatistical context

Geostatistics regionalized phenomenon modelled by a random field (RF) Z

In many cases, we partially observe only* one *realization of Z

- mining resources estimation, soil contamination evaluation

Assumption Z is second-order stationary

Dependence measure

$$\gamma : \mathbb{R}^2 \to \mathbb{R}_+ \qquad \textbf{variogram}$$

$$\gamma(h) := \frac{1}{2}\mathbf{E}\left[\left(Z(h) - Z(0)\right)^2\right]$$

Stationary max-stable RF

Let Z be a simple max-stable RF

⚠ Z is not second-order stationary

Assumption Z simple **stationary** max-stable RF

$$\mathbf{P}\left[Z(x_1) \leq z_1, \ldots, Z(x_n) \leq z_n\right] = \mathbf{P}\left[Z(x_1 + h) \leq z_1, \ldots, Z(x_n + h) \leq z_n\right]$$

$$h, x_1, \ldots, x_n \in \mathbb{R}^d \qquad z_1, \ldots, z_n \in \mathbb{R}$$

Remark In the stationary case, bivariate extremal dependence is parameterized in terms of separating vector h (and, in the isotropic case, in terms of distance), which drastically reduces the number of required parameters

Extremal coefficient function

Extremal coefficient function (ECF)

$$\theta : h \in \mathbb{R}^d \longrightarrow \boldsymbol{E}\Big[\max\big(Y(0), Y(h)\big)\Big] \in [1,2]$$

$\forall h \in \mathbb{R}^d \quad z \in (0, +\infty)$

$$\mathbf{P}\left[Z(0) \leq z, Z(h) \leq z\right] = \exp\left\{-\frac{\theta(h)}{z}\right\}$$

- $\theta(h) = 1$ **full dependence**
- $\theta(h) = 2$ **independence**

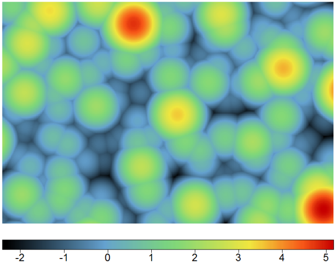
Example - Smith process

$$Z(x) := \max_{(S,T)\in\Pi} \left(\frac{f(x-S)}{T} \right) \qquad x \in \mathbb{R}^d$$

Π homogeneous Poisson point process on $\mathbb{R}^d \times (0, +\infty)$ with unit rate
$f : \mathbb{R}^d \to \mathbb{R}_+$ centred d-variate **Gaussian** p.d.f. with covariance matrix Σ

$$\theta(h) = 2\Phi\left(\frac{\|h\|_{\Sigma^{-1}}}{2} \right)$$

Φ standard Gaussian c.d.f.



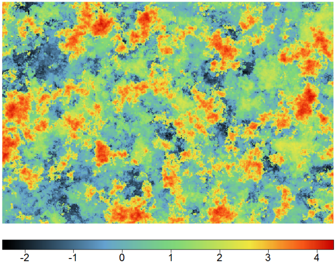
Example - Extremal Gaussian process

$$Z(x) := \sqrt{2\pi} \max_{T \in \Pi} \left(\frac{W_T(x)}{T} \right) \qquad x \in \mathbb{R}^d$$

Π homogeneous Poisson point process on $(0, +\infty)$

W_T i.i.d. copies of a standard Gaussian process with correlation function ρ

$$\theta(h) = 1 + \sqrt{\frac{1 - \rho(h)}{2}}$$

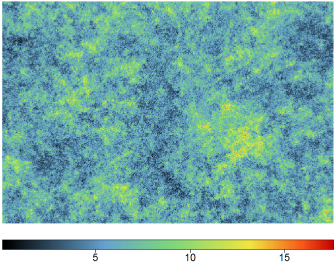


Example - Brown-Resnick process

$$Z(x) := \max_{T \in \Pi} \left(\exp\left\{ W_T(x) - \frac{\sigma_W^2(x)}{2} \right\} \right) \qquad x \in \mathbb{R}^d$$

Π homogeneous Poisson point process on $(0, +\infty)$

W_T i.i.d. copies of a intrinsically Gaussian process with $\mathrm{E}\left[W_T(x)\right] = 0$, $\sigma_W^2(x) = \mathrm{Var}\left[W_T(x)\right]$ and variogram γ



Example - Brown-Resnick process

$$Z(x) := \max_{T \in \Pi} \left(\exp \left\{ W_T(x) - \frac{\sigma_W^2(x)}{2} \right\} \right) \qquad x \in \mathbb{R}^d$$

Π homogeneous Poisson point process on $(0, +\infty)$

W_T i.i.d. copies of a intrinsically Gaussian process with $\mathrm{E}\left[W_T(x)\right] = 0$, $\sigma_W^2(x) = \mathrm{Var}\left[W_T(x)\right]$ and variogram γ

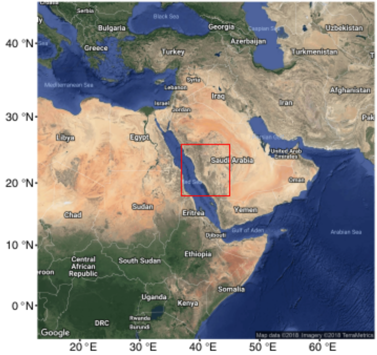
$$\theta(h) = 2\Phi\left(\sqrt{\frac{\gamma(h)}{2}}\right)$$

Φ standard Gaussian c.d.f.

Illustration - Saudi Arabian rainfall

Davison et al (2019) *SpatialExtremes* package (R)

- Jeddah liable to intense (but rare) strong convective rainstorms, leading to flash floods, extensive damage and deaths
- three-hourly rainfall satellite measurements (in mm/hr) available at 750 grid cells over 17 years (1998-2014)



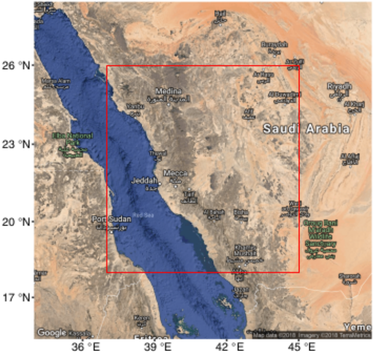
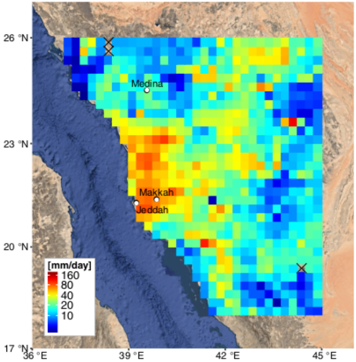


Illustration - Saudi Arabian rainfall

- Computation of the 17 annual daily rainfall maxima for each cell.
- 17 maps of annual daily rainfall maxima : 17 realizations of a max-stable process observed at 750 locations (space-rich but time-poor dataset)



Annual maximum for 2009

Illustration - Saudi Arabian rainfall

- Computation of the 17 annual daily rainfall maxima for each cell.
- 17 maps of annual daily rainfall maxima : 17 realizations of a max-stable process observed at 750 locations (space-rich but time-poor dataset)

The region is arid : some grid cells experienced no rain in certain years. In total, six annual maxima, in different grid cells, are equal to zero

$\longrightarrow$ censored likelihood : for the observed maxima that do not exceed the threshold $u = 3$mm/day, we only give to the likelihood the information that these observations are below u

Estimation of the margins

Remark Estimations of the margins parameters and of the dependence structure are done separately

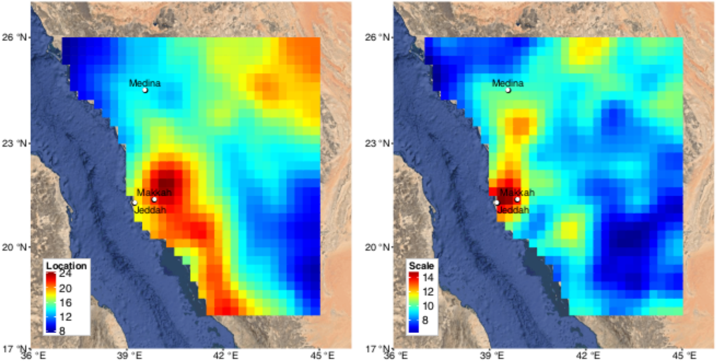
For each cell $s \in \{1, \dots, 750\}$, we want to estimate

- the location parameter – $b_n(s)$
- the scale parameter – $a_n(s)$
- the shape parameter or extreme value index – $\xi(s)$

Remark Only 17 temporal replicates : it is difficult to find simple relationships that well capture the spatial variation of the marginal parameters.
$\longrightarrow$ neglecting spatial dependence : local censored likelihood approach to estimate the 3 parameters at each location

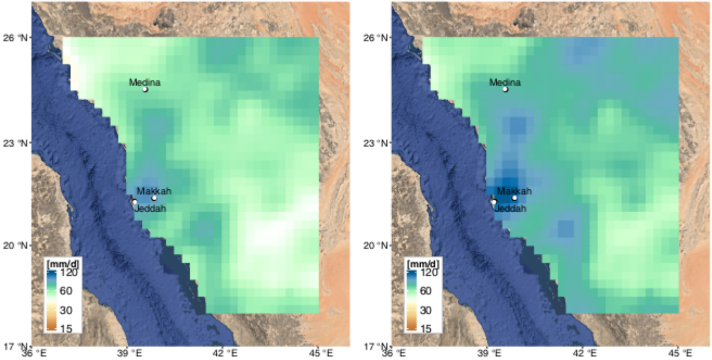
Result Good estimation of the location and scale parameter but large variation of $\xi(s)$ (from 0.28 to 0.49) that imply very different tail behavior
It is assumed that $\xi(s) = \xi$. The estimation gives $\widehat{\xi} = 0.14$: the rainfall distribution is slightly heavy-tailed

Estimation of the margins



Estimation, in each grid point, of the location parameter (left) and the scale parameter (right)

Estimation of the margins



Estimation, in each grid point, of the 50-year (left) and 100-year (right) return levels (in mm/day)

Estimation of the dependence structure

Remark the data are transformed to a common unit Fréchet scale

4 stationary isotropic max-stable processes to assess the spatial dependence of extreme rainfall events :

- **Brown-Resnick process** with variogram $\gamma(s_1, s_2) = (\|s_1 - s_2\|/\lambda)^\kappa$
  $\lambda > 0 \quad \kappa \in (0, 2]$
- **Extremal Gaussian process** (Schalther process) with correlation function $\rho(s_1, s_2) = \exp\{-\gamma(s_1, s_2)\}$
- **Extremal-t process** with correlation function $\rho(s_1, s_2) = \exp\{-\gamma(s_1, s_2)\}$ and degrees of freedom $\alpha > 0$
- **Smith process** with covariance matrix $\Sigma = \lambda^2 I_2$

In addition : the same models are considered in a geometrically anisotropic version. They are obtained by replacing the Euclidean distance $\|s_1 - s_2\|$ by the Mahalanobis distance in the models

Estimation of the dependence structure

Estimation Pairwise (conditional) likelihood + 95% confidence intervals obtained using a non-parametric bootstrap

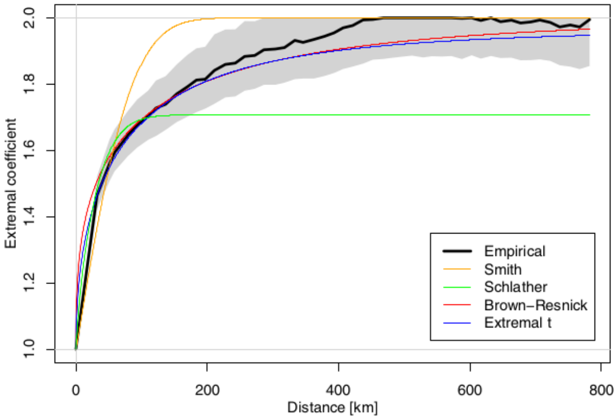
Model	λ [km]	κ	α	a	θ	CLIC
Smith	$34_{[26,39]}$					124
Schl.	$44_{[34,53]}$	$1.46_{[1.19,1.84]}$				362
B.–R.	$13_{[8,16]}$	$0.71_{[0.52,0.94]}$				23
Ext.-t	$333_{[165,1357]}$	$0.90_{[0.63,1.13]}$	$5.9_{[3.9,13.1]}$			0

Anisotropic max-stable models

Model	λ [km]	κ	α	a	θ	CLIC
Smith	$31_{[28,37]}$			$0.82_{[0.72,1.33]}$	$0.19_{[-1.48,2.02]}$	119
Schl.	$42_{[31,54]}$	$1.47_{[1.20,1.82]}$		$0.89_{[0.70,1.28]}$	$0.23_{[-0.32,1.26]}$	362
B.–R.	$12_{[7,21]}$	$0.72_{[0.53,0.95]}$		$0.71_{[0.52,1.81]}$	$-0.12_{[-0.25,1.41]}$	21
Ext.-t	$424_{[176,1352]}$	$0.90_{[0.64,1.10]}$	$6.2_{[4.1,14.8]}$	$1.37_{[0.55,1.66]}$	$1.37_{[-0.30,1.42]}$	41

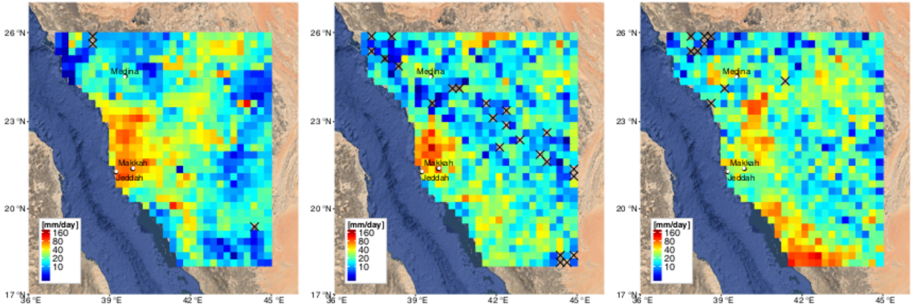
CLIC : model selection criterion (to minimize)

- No strong evidence against isotropy
- Extremal-t model looks best but $\widehat{\alpha}$ and $\widehat{\lambda}$ are very unstable
- Isotropic Brown–Resnick model : good balance of fit and parsimony



For each model and each distance h the estimation of $\theta(h)$ may be compared to the empirical estimation of $\theta(h)$

- Brown–Resnick and extremal-t processes : reasonable fit
- Smith model : too rigid to capture the decay of dependence
- Schlather model : unable to capture the long-range independence.



Annual rainfall maxima for 2009 (left), and two simulated maps (middle and right) based on the fitted isotropic Brown–Resnick max-stable model

- The spatial patterns and forms of dependence observed in the simulations tend to agree with the data (although the 2009 annual maxima seem to be slightly smoother)

Spatial aggregated risk

- 10000 simulations from the fitted isotropic Brown-Resnick model Z
- Focus on 14 grid cells s that are less than 50km from Jeddah and Makkah (2nd and 3rd largest cities) – ($s \in S^*$)

We use the simulation to estimate the the probability $p(v)$ that the annual maximum averaged over the 14 grid cells be larger than a select threshold v :

$$p(v) = \mathbf{P}\left[\frac{1}{|S^*|}\sum_{s \in S^*} Z(s) > v\right]$$

- p(50mm/day) = 0.072 (return period = 13.9 years)
- p(**71.1mm/day**) = 0.019 (return period = 1/0.019 = **53.6 years**)
  71.1: daily rainfall level observed on November 25, 2009, during the Jeddah floods, which caused 122 fatalities
- p(100mm/day) = 0.0041 (return period = 245.1 years)

Recap

- Estimating parametrically the dependence structure of multivariate extremes is difficult in high dimension (lack of flexibility or a rapid increase in the number of parameters)

- Same problem (even worse) for non-parametric estimation (not presented here)

- Stationary max-stable processes are more parsimonious models but

  1. They rely on the strong assumption of stationarity
  2. They require prior domain knowledge on the spatial locations of gauging stations

Need for sparse structure for multivariate extremes

So far, 3/4 main approaches :

- Adaptation of unsupervised learning methods : clustering approaches and principal component analysis
- Concomitant extremes
- Graphical models

See a review in *Engelke & Ivanovs (2021)*

References

A.A. Balkema and L. de Haan, Residual Life Time at Great Age, *The Annals of Probability*, 2(5), 1974

A. Davison, R. Huser and E. Thibaud, Spatial Extremes, *CRC Press*, 2019

L. de Haan, A Spectral Representation for Max-stable Processes, *Ann. Probab.*, 12(4):1194-1204, 1984

s. Engelke and I. Jevgenijs, Sparse Structures for Multivariate Extremes, *Annual Review of Statistics and Its Application*, 8(1):241-270, 2021

R.A. Fisher and L.H.C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):180-190, 1928

B. Gnedenko, Sur la distribution limite du terme maximum d'une serie aleatoire, *Annals of Mathematics*, 44(3):423-453, 1943