UNIVERSITE CATHOLIQUE DE LOUVAIN
FACULTE DES SCIENCES
ECOLE DE STATISTIQUE, BIOSTATISTIQUE
ET SCIENCES ACTUARIELLES

Temporal analysis evolution of extreme values using climatological data

Promoteurs : Johan Segers
Co-promoteur : Prénom Nom
[] Lecteurs : Anna Kirilliouk
Michel Crucifix

Mémoire présenté en vue de l'obtention du
Master en statistiques, orientation générale
par : **Antoine Pissoort**

Juin 2017

Contents

List of Figures

Acknowledgements

I would first like to thank my thesis supervisor Johan Segers for all his help and his guidance during this year.

I also would like to thank the "Institut Royal de Météorologie" (IRM) of Belgium for his help, his guidance and his provided quality datasets.

Part I

Theoretical Framework

1 Introduction and Preliminaries

Unlike his counterparts (see for example credit risk analysis, financial applications,...), the extreme value analysis applied on the broad environnemental area like here for the meteorological data, has strong impacts on the people lives

The problem we are here facing in climate change evidence is that of le lack of past data to compare with her

Also, for such an analysis, the number of parameters to take into account is considerable (and tend to infinity)

Can make a parallelism with Chaos Theory and the well-known butterfly effect which have strong applications in weather models

We highly expect the climate change to affect the extreme weather

[extremes in climate change p.347]

The disadvantage of Q–Q plots is that the shape of the selected parametric distribution is no longer visible [book: stat. analysis of extreme values p.62]

"The first myth about climate extremes, which has been purported by researchers in climatology or hydrology, among them prominent names, is that “extremes are defined as rare events” or similar. This myth is debunked by a simple bimodal PDF (Fig. 6.12a). The events sitting in the tails of that distribution are not rare" [Mud14, pp.257]

Until now, studies on climate extremes that consider Europe have usually had a strong national signature , or have had to make use of either a dataset with daily series from a very sparse network of meteorological stations (e.g. eight stations in Moberg et al. (2000)) or standardized data analysis performed by different researchers in different countries along the lines of agreed methodologies (e.g. Brazdil et al., 1996; Heino et al., 1999) [KTWK+02]

Extrapolation !!!! See p154 [statistical analysis of extreme book]

Voir effet de l’ilot de chaleur –> urbanisation sur les tempés !

—> artificial warming on cities stations which were not(less) urbanized 100 years ago.

1.1 Statistical tools

1.1.1 Order statistics

First of all, we write the i-th order statistics $X_{(i)}$ which are the statistics ordered by increasing value

$$X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)} \tag{1.1}$$

We adopt this simpler notation by assuming that the number of our observations will be denoted by n.

One order statistics is of particular interest for our purpose, the maximum

$$X_{(n)} := \max_{1 \leq i \leq n} X_i \tag{1.2}$$

for the minimum, $X_{(1)} = \min_{1 \leq i \leq n} X_i$ that we will still use by converse (see section .)

$$X_{(1)} := \min_{1 \leq i \leq n} X_i = - \max_{1 \leq i \leq n} (-X_i) \tag{1.3}$$

We can retrieve the distribution of our statistic of interest $X_{(n)}$

$$\Pr\{X_{(n)} \leq x\} = \Pr\{X_1, ..., X_n \leq x\} \overset{(\perp\!\!\!\perp)}{=} \Pr\{X_1 \leq x\} ... \Pr\{X_n \leq x\} = F^n(x) \tag{1.4}$$

(write vertically)

where the ($\perp\!\!\!\perp$) follows from the iid assumption of the sequence $\{X_i\}$.

1.1.2 Miscellaneous

We also define the *survival* or *survivor* function $\bar{F} = 1 - F$ which is widely useful for this kind of (biostatistical) applications.

Finally, we decide to include in Appendix some concepts of convergence (A.), regularly and slowly varying functions (A.), as they will often appear in the text.

2 Extreme Value Theory : the Basics

There are two approaches, the block-maxima (section **2.1**) and the peaks-over-threshold approach (section **2.2**) yielding to different extreme value distribution. The former aims at while the latter models the...

Similar distribution functions We say that two distributions functions G and G^* are *similar* or are of the *same type* if, for constants $a > 0$ and b

$$G^*(az + b) = G(z) \qquad \forall x, \tag{2.1}$$

meaning that they differ only in location and scale. In the sequel, the concept of *similar* distributions will be useful to derive the three different families of extreme value distributions from other distributions of the *same type*. This is directly linked with max-stable process that we will define...

Max-stability [LLR83] or [Res87] We say that a distribution G is *max-stable* if, for each $n \in \mathbb{N}$

$$G^n(a_n z + b_n) = G(z), \qquad n = 2, 3, ... \tag{2.2}$$

for some constants $a_n > 0$ and b_n. In other words, taking powers of G results only in a change of location and scale. **??** This concept will be closely connected with the extremal limit laws in the following ().

Min-stability Anageously, from [RT07, pp.23], we say that a distribution function G is *min-stable* if

$$\Pr\{X_{(1)} > d_n + c_n z\} = \bar{G}^n(d_n + c_n z) = \bar{G}(z) \tag{2.3}$$

where $c_n = a_n$, $d_n = -b_n$ and for $X_{(1)}$ the minimum of the sample of size n (1.3).

Principles of stability Behind all the principles about Extreme Value Theory that will be covered during this thesis, will be influenced by the principle of *stability.*

2.1 Block-Maxima

Introduction and extremal types theorem ....

"Fisher-Tippett" extreme value theorem [extreme value and cluster ana. of euro... clustering 2011]

coherence from

This theorem, introduced by Fisher and Tippett [FT28], later revised by [Gne43] and finally streamlined by [**?**], is very important for its applications and states the following:

If the distribution of partial maxima of an independent and identically distributed sequence of random variables with common (unknown) distribution F, say, $X_{(n)}$, properly normalized, converges to a non-degenerate limiting distribution G, i.e.

$$\lim_{n\to\infty} \Pr\Big\{a_n^{-1}(X_{(n)} - b_n) \leq z\Big\} = F^n(a_n z + b_n) \longrightarrow G(z), \qquad n \to \infty \tag{2.4}$$

for some constants $a_n > 0$, $b_n \in \mathbb{R}$, meaning that F is said to be in the **domain of attraction**[1] of G, denoted by $F \in D(G)$, with G the *Generalized Extreme Value* (GEV) distribution :

$$G(z) = \exp\left\{ -\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]_{+}^{-\xi^{-1}} \right\} := G_\xi(z), \tag{2.5}$$

where we define $y_+ = \max(y, 0)$ which denotes in the above that $\{z : 1 + \xi\sigma^{-1}(z-\mu) > 0\}$ to ensure the term in the exponential is negative and the distribution function converge to 1 (...), $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$ with μ, σ and ξ being the three parameters of the model characterizing location, scale and shape respectively. We think important to point out that here, the location parameter μ does not represent the mean as in the classic statistical view, but does represent the "center" of the distribution, and the scale parameter σ is not the standard deviation, but does govern the "size" of the deviations around μ.

Max-stability and GEV distribution An important theorem says the following, for any distribution function F [Col01]

$$F \text{ is max-stable} \iff F \text{ is GEV} \tag{2.6}$$

To gain interesting insights of the implications of this theorem, we think useful to give a proof but only for the "$\Leftarrow$" as the converse requires too much mathematical backgrounds.

Proof: If $a_n^{-1}(X_{(n)} - b_n)$ has limit distribution G for large n (2.4),

$$\Pr\Big\{a_n^{-1}(X_{(n)} - b_n) \le z\Big\} \approx G(z) \tag{2.7}$$

and so for any integer k, since nk is large,

$$\Pr\Big\{a_{nk}^{-1}(X_{(n)k} - b_{nk}) \le z\Big\} \approx G(z) \tag{2.8}$$

But, since $X_{(n)k}$ is the maximum of k variables having identical distribution as $X_{(n)}$,

$$\Pr\Big\{a_{nk}^{-1}(X_{(n)k} - b_{nk}) \le z\Big\} = \Big[\Pr\Big\{a_{nk}^{-1}(X_{(n)} - b_{nk}) \le z\Big\}\Big]^k \tag{2.9}$$

giving two expressions for the distribution of M_n, by (2.8) and (2.9) :

$$\Pr\{X_{(n)} \le z\} \approx G\Big(a_n^{-1}(z-b_n)\Big) \quad \text{and} \quad \Pr\{X_{(n)} \le z\} \approx G^{1/k}\Big(a_{nk}^{-1}(z-b_{nk})\Big), \tag{2.10}$$

so that G and $G^{1/k}$ are identical apart from location and scale coefficients. Hence, $G$is max-stable and therefore GEV. This gives proof of the **extremal types theorem** in 2.4-2.5.

The quantity $\xi \in \mathbb{R}$ in (2.5) is called the *extreme value index* (EVI) and is at the center of the analysis in extreme value theory. It determines, in some degree of accuracy, the type of the underlying distribution. Hence, from this general definition of the GEV distribution (2.5), we can directly retrieved three principal classes of EV distributions, from their *standard form*, in the *α-parametrization*, with $\alpha = \xi^{-1}$:

[1] We will more precisely define this concept in the next section.

$$\mathbf{I}: G_1(z) = \exp\{-e^{-z}\}, \qquad -\infty < z < \infty \tag{2.11}$$

$$\mathbf{II}: \; G_{2,\alpha}(z) = \begin{cases} 0, & z \leq 0 \\ \exp\{-(z)^{-\alpha}\}, & z > 0, \; \alpha > 0 \end{cases} \tag{2.12}$$

$$\mathbf{III}: \; G_{3,\alpha}(z) = \begin{cases} \exp\big\{ - (-z)^{\alpha}\big\}, & z > 0, \; \alpha > 0 \\ 1, & z \geq 0 \end{cases} \tag{2.13}$$

(mettre les indice aux fctions G + le shape parameter est "correct" ????)

where we replace z by $(z-\mu)/\sigma$ to add the location and scale parameters μ and σ in order to obtain the three *extreme value distributions* [RT07, pp.16].

By simply coming back in the ξ-parametrization by using $\xi = \alpha^{-1}$ in the above distribution functions, all these three classes of extreme distributions can be expressed in the same functional form as special cases of this single three-parameter[2] distribution (2.5). That is, when $\xi \to 0$ we retrieve the **type I** or *Gumbel* family (2.11) while $\xi > 0$ and $\xi < 0$ leads to the **type II** or *Fréchet* family and to the **type III** or *Weibull* family respectively ((2.12) and (2.13)). Both the Gumbel and Fréchet limiting distributions are unbounded[3].; that is, the upper endpoint tends to $+\infty$ while the Weibull distribution has a finite right endpoint. In the following, we will defined the left and the right endpoint $_*x$ and x_*, respectively by :

$$_*x = \inf\{x : F(x) > 0\} \quad \text{and} \quad x_* = \sup\{x : F(x) < 1\} \tag{2.14}$$

Density We give a representation of the density of these functions by considering the density of the GEV distribution (2.5), that is $g_\xi(z) = \frac{dG_\xi(z)}{dz}$. This is shown in in figure (2.1) for various shape parameters. We also give in appendix

$$g_\xi(z) = \sigma^{-1}\Big[1 + \xi\Big(\frac{z-\mu}{\sigma}\Big)\Big]^{-\xi^{-1}-1} \exp\Bigg\{ - \Big[1 + \xi\Big(\frac{z-\mu}{\sigma}\Big)\Big]^{-\xi^{-1}}\Bigg\} \tag{2.15}$$

?? dernier graphe weibull

In some ways, some people will feel this was unfortunate, because now it is common for people to model and fit the GEV without thinking very clearly about the specific form of their data and distributions [Extremes, distribution, etc..] That is the reason why we think it can be useful to explain in the following some examples of how we can construct such extreme distributions for the three classes in concrete cases (see next section **2.1.1**), playing with the appropriate choice of sequences a_n and b_n to retrieve the pertaining distribution family.

2.1.1 Concrete cases : examples

Type I or **Gumbel** distribution $G_1(x)$ can be retrieved by considering, for example, iid exponential distributed sequence $\{X_j\}$ of random variables, that is $X_j \overset{iid}{\sim} Exp(\lambda)$ and consider the

[2] Actually, there are just location and scale parameters in the type **I** extremal model (2.11) as $\xi \to 0$

[3] Actually, the Fréchet has a finite left endpoint, but this has no really interest here

Footnotes goodbye

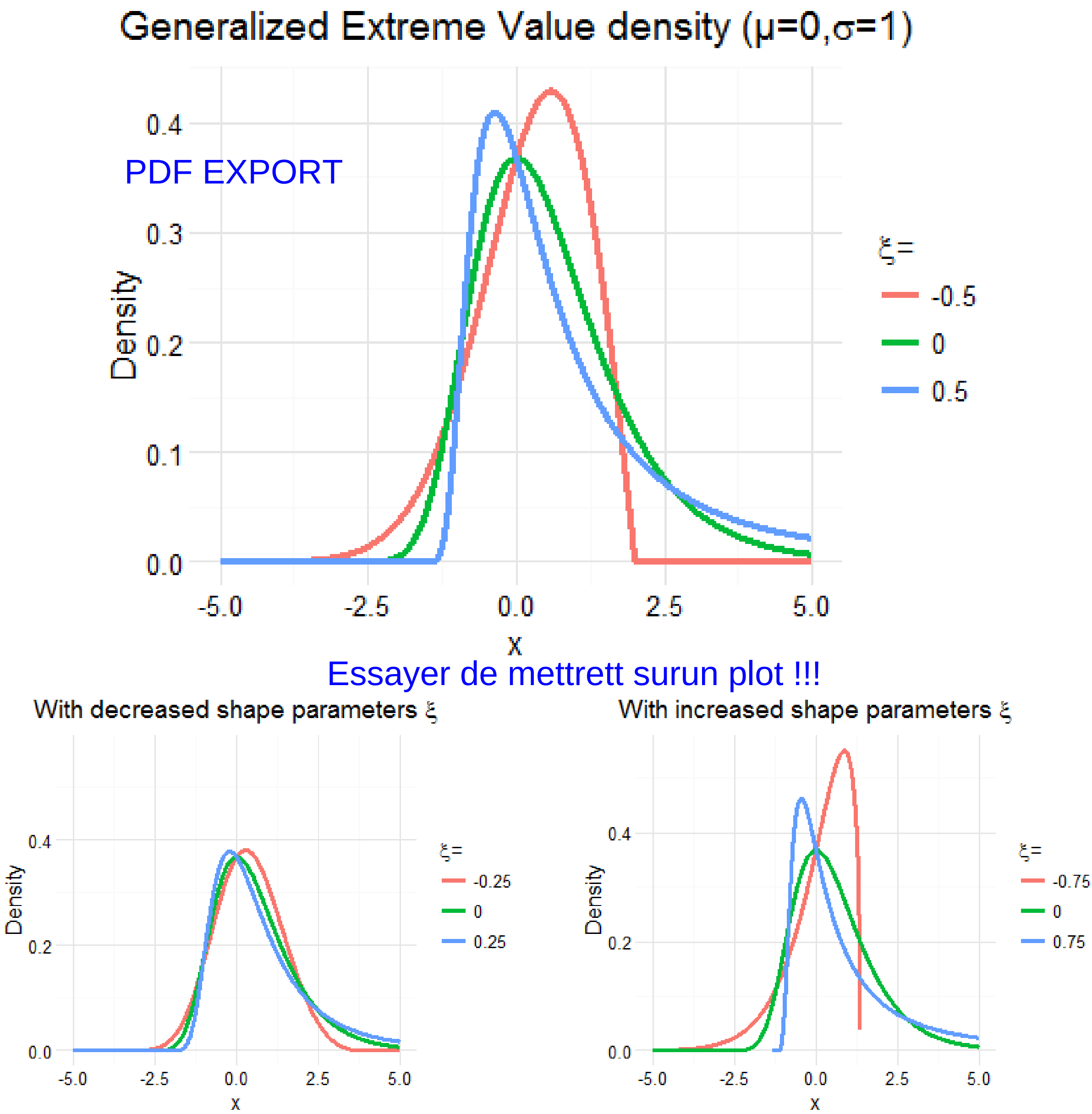

Figure 2.1: Representation of the GEV densities for various values of the shape parameter ξ : $|\xi| = 0.5$ (1) $|\xi| = 0.25$, (2) and $|\xi| = 0.75$ (3) for the Fréchet density [red], Weibull density [blue] while the Gumbel density [green] is kept fixed with $\xi = 0$.

largest of these values $X_{(n)}$ as defined earlier. By definition, we know $F(x) = 1 - \exp^{-x}$. Our goal is to find non-random sequences $\{b_n\}$, $\{a_n > 0\}$ such that

$$\lim_{n\to\infty} \Pr\{a_n^{-1}(X_{(n)} - b_n) \leq z\} = G_1(z). \tag{2.16}$$

We can easily find that

$$\begin{aligned} \Pr\{a_n^{-1}(X_{(n)} - b_n) \leq z\} &= \Pr\{X_{(n)} \leq b_n + a_n z\} \\ &= \Big[\Pr\{X_1 \leq b_n + a_n z\}\Big]^n \\ &= \Big[1 - e^{-\lambda(b_n + a_n z)}\Big]^n \end{aligned} \tag{2.17}$$

from the iid assumption of the random variables and their exponential distribution. Hence, by choosing the sequences $a_n = \lambda^{-1} \log n$ and $b_n = \lambda^{-1}$,

$$\begin{aligned} \Big[1 - e^{-\lambda(b_n + a_n z)}\Big]^n &= \Big[1 - \frac{1}{n}e^{-z}\Big]^n \\ &\to \exp(-e^{-z}) := G_1(z), \end{aligned} \tag{2.18}$$

[mettre les exponentielles en exp() et en e ??]

we find the the so-called standard *Gumbel* distribution in the limit[4].

We can show the same with iid standard normal random variables, $X_j \overset{iid}{\sim} N(0,1)$, with sequences $a_n = -\Phi^{-1}(1/n)$ and $b_n = 1/a_n$. (see appendix [extremes, distributions pdf])

Typically, unbounded distributions like the Exponential and Normal (as well as the Gamma, Lognormal, Weibull, etc.) whose tails fall off exponentially or faster will have this same Gumbel limiting distribution for the maxima, and will have medians (and other quantiles) that grow as $n \to \infty$ at the rate of (some power of) $\log n$. This is typical example of light-tailed distribution (i.e., decays exponentially, as defined in section **1.1**).

<u>**Type II**</u> or **Fréchet-Pareto type** distrbution $G_2(x)$

When starting with a sequence $\{X_j\}$ iid random variables (block-maxima?) following a *basic*[5] Pareto distribution with shape parameter $\alpha \in (0, \infty)$, $X_j \sim Pa(\alpha)$, we have that

$$F(x) = 1 - x^{-\alpha}, \quad \text{for } x \in [1, \infty) \tag{2.19}$$

so that we can write, by choosing appropriately $b_n = 0$

$$\begin{aligned} -n\bar{F}(a_n z + b_n) &= -n(a_n z + b_n)^{-\alpha} \\ &= \Big[Q(1 - \frac{1}{n})\Big]^{\alpha} (a_n)^{-\alpha}(-z^{-\alpha}). \end{aligned} \tag{2.20}$$

where $Q(1 - \frac{1}{n})$ is the quantile function (see). Hence, it is easy to see that by setting the constant $a_n = Q(1 - 1/n)$ and keeping $b_n = 0$, we have that

$$\Pr\{a_n^{-1} X_{(n)} \leq z\} \to \exp(-z^{-\alpha}) \tag{2.21}$$

showing that for this particular values of the normalizing constants, we retrieve the Fréchet distribution in the limit from a strict Pareto distribution. The fact that b_n is set to zero can be understood intuitively since for heavy-tailed distribution (see) such as the Pareto distribution, a correction for location is not necessary to obtain non-degenerate limit distribution. [BTV96, pp.51]

[see p.28 memoire other si ft autre exemple]

More generally, we can state the more general following theorem :

[4] We remind the following property : $\lim_{n \to \infty} (1 + \frac{x}{n})^n = \exp(x)$.

[5] Or *generalized* with scale parameter set to 1 so that it vanishes.

Theorem x : If for the same choice of normalizing constants as above, that is $a_n = Q(1 - \frac{1}{n})$ and $b_n = 0$ and for any $x \in \mathbb{R}$

$$n[1 - F(a_n x)] = \frac{1 - F(a_n x)}{1 - F(a_n)} \to x^{-\alpha} \qquad \text{as n} \to \infty \tag{2.22}$$

then we can obtain the Fréchet distribution in the limit, or written formally "$\bar{F}$ *is of Pareto-type*" or, more technically, "$\bar{F}$ *is regularly varying with index -α*". We let the concepts of **regularly varying functions**, together with **slowly varying functions** be defined in appendix A.1 with some useful theorems and properties, according to [BTV96, p.51-54] and supported by [BGST06, p.49, 77-82].

[BGST06, pp.75] !!!!

Type III or Weibull family (?) of distributions $G_3(x)$ are, for example, in the limit of n iid uniform random variables $X_j \sim U[L, R]$ where $L \in \mathbb{R}$ and $R \in \mathbb{R}$ denote respectively the left and the right endpoint of the domain of definition, $R > L$. We have then

$$F(x) = \frac{x}{R - L}, \qquad \text{for } x \in [L, R] \tag{2.23}$$

$= 0$ for $x < L$, $= 1$ for $x > R$. Assuming the general case ($[L, R]$ can be $\neq [0, 1]$), we have for the maximum $X_{(n)}$:

$$\begin{aligned} \Pr\{a_n^{-1}(X_{(n)} - b_n) \leq z\} &= \Pr\{X_{(n)} \leq b_n + a_n z\} \\ &= \Big[1 - \frac{R - b_n - a_n z}{R - L}\Big]^n \qquad if L \leq b_n + a_n z \leq R \\ &= (1 + \frac{z}{n})^n \to e^z \qquad if z \leq 0 \text{ and } n > |z| \end{aligned} \tag{2.24}$$

when choosing $a_n = R$ and $b_n = (R - L)/n$, we find the unit Reversed Weibull distribution $We(1, 1)$ in the limit as expected.

We think it is important to summarize all these extreme value distributions in the same plot

no diff

2.1.2 Maximum(?) domain of attraction

limoits aresum stable

The preceding results can be more easily summarized and obtained when considering *maximum domain of attraction* (MDA) theory.

max stable

max--> dsitci,tion avec sum

Domain of attraction We say that a distribution F is in the (*maximum*) *domain of attraction* of an extreme value type distribution G_k ($k = 1, 2, 3$ denoting the Gumbel, Fréchet and Weibull family respectively), denoted by $F \in D(G_k)$, if there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that the distribution of $a_n^{-1}(X_{(n)} - b_n)$ converges weakly (see **??**) to G_k where $X_{(n)}$ is as defined earlier with distribution F.
Before going further with the characterization of the three domains of attraction of our purpose, we think important to introduce a new theorem from Gnedenko [**?**]

Convergence to Types Theorem Let F_n be a sequence of random variables converging weakly (see appendix A.1.1) to F. Let $a_n > 0$ and $b_n \in \mathbb{R}$ such that $a_n F_n + b_n \Rightarrow F'$, where both F and F' are non-degenerate (see **??**). Then,

$$a_n \to a \quad \text{and} \quad b_n \to b, \qquad a > 0 \ \text{ and } \ b \in \mathbb{R} \tag{2.25}$$

Equivalently, if G_n, G, G' are distribution functions with G, G' being non-degenerate, and there exists $a_n, a'_n > 0$ and $b_n, b'_n \in \mathbb{R}$ such that

$$G_n(a_n x + b_n) \xrightarrow{d} G(x) \quad \text{ and } \quad G_n(a'_n x + b'_n) \xrightarrow{d} G'(x) \tag{2.26}$$

at all continuity points of F, respectively F', then there exists constants $A > 0$ and $B \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \frac{a_n}{a'_n} \to A, \qquad \lim_{n\to\infty} \frac{(b_n - b'_n)}{a'_n} \to B \tag{2.27}$$

and $G'(Ax + B) = G(x)\ \forall x \in \mathbb{R}$.

We have now all the necessary tools to the pertaining domains of attractions. But, before proceeding, we would like to point out that the fact that the characterization of the first domain of attraction (Gumbel class) is much more complex than the two following (Fréchet and Weibull class) and requires much more technicalities going beyond the scope of this thesis. Moreover, despite this class is important in theory, it is less relevant for our purpose of modelling extremes. It often requires other generalizations, for instance with additional parameters to surpass the issues of fitting empirical data. [PF15]

In each of the characterization of the domains of attractions, we will present some of their most useful, necessary and sufficient conditions ... together with their *von Mises conditions*, initially from [] but revisited in [**?**]. The latter are very important in practice and sometimes more intuitive because they make use of the *hazard function*, defined by :

$$r(x) = \frac{f(x)}{\bar{F}(x)} \tag{2.28}$$

involving the density function $f(x) = \frac{dF(x)}{d(x)}$ in the numerator.[**?**]

I. Domain of attraction for Gumbel distribution G_1 We derive here two ways of formulating necessary and sufficient condition for a distribution function F to be in the domain of attraction of G_1, namely $F \in D(G_1)$.

- From [HF06, pp.20], for finite or infinite right endpoint x_* with $\int_x^{x_*} \int_t^{x_*} (1 - F(s)) ds dt < \infty$, the function

  bar

$$h(x) := \frac{(1 - F(x)) \int_x^{x_*} \int_t^{x_*} (1 - F(s))\ ds\ dt}{\left(\int_x^{x_*} (1 - F(s)) ds \right)^2} \tag{2.29}$$

  ????

  must satisfy $\lim_{t \uparrow x_*} h(t) = 1$.[6]

[6]We remind the (intuitive here) notation $\lim_{t\uparrow y}(.)$ simply means that t is approaching y from below, i.e. from values smaller than y in a increasing manner. Alternatively, one may write $\lim_{t\to y^-}(.)$. The converse holds for $\lim_{t\downarrow y}(.)$.

- From [**?** , pp.72], for some auxiliary function b, for every $v > 0$, the condition

$$\frac{\bar{F}(x + b(x)v)}{\bar{F}(x)} \to e^{-v} \quad \text{gpd} \tag{2.30}$$

  must hold as $x \to x_*$. Then,

$$\frac{b(x + vb(x))}{b(x)} \to 1. \tag{2.31}$$

A lot of more precise characterizations and conditions together with proofs can be found, for example in [HF06, pp.20-33]. We can also mention a condition that is based on the von Mises function[].

However, we present his ***von Mises criterion*** as in [**?** , pp.73]:

If the *hazard function* $r(x)$ (2.28) is ultimately positive in the neighbourhood of x_*, is differentiable there and satisfies

$$\lim_{x \uparrow x_*} \frac{dr(x)}{dx} = 0 \tag{2.32}$$

then $F \in D(G_1)$.

More intuitively, we can remark that all distributions which are exponentially decaying will have this propensity to be in the Gumbel domain of attraction. For instance, the *Exponential*, the *Gamma*, the *Weibull*, the *logistic*, etc. To see that, by a Taylor expansion, we have that

$$\bar{G}_1(x) = 1 - \exp(-e^{-x}) \sim e^{-x}, \qquad \text{as } x \to \infty \tag{2.33}$$

and hence, the Gumbel domain of attractions G_1 decays exponentially (as tend their pertaining distributions).

II. Domain of attraction for Fréchet distribution $G_{2,\alpha}$ Let $\alpha := \xi^{-1} > 0$ be the *index* of the Fréchet distribution $G_{2,\alpha}$ (see (2.12)). Then, a distribution F is in the *domain of attraction of $G_{2,\alpha}$* if and only if

PAS F !!!!!!!!

$$\bar{F}(x) = x^{-\alpha} L(x) \tag{2.34}$$

for some slowly varying function L. In this case and with $b_n = 0$,

$$F^n(a_n x) \to G_2(x), \qquad \forall x \in \mathbb{R} \tag{2.35}$$

with

$$a_n := F^{\leftarrow}(1 - \frac{1}{n}) = \left(\frac{1}{\bar{F}}\right)^{\leftarrow}(n), \tag{2.36}$$

where we define the quantity $F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}$ for $t < 0 < 1$ as the *generalized inverse of F* with which we can retrieve $x_t = F^{\leftarrow}(t)$, the t-quantile of F.

This previous theorem informs us that all distribution functions $F \in D(G_{2,\alpha})$ have necessarily an infinite right endpoint, that is $x_* = \sup\{x : F(x) < 1\} = \infty$. These distributions are all with regularly varying right-tail with index $-\alpha$. In short,

$$F \in D(G_{2,\alpha}) \Longleftrightarrow \bar{F} \in R_{-\alpha}. \tag{2.37}$$

Finally, we must also present the (revisited) ***Von Mises condition*** for this domain of attraction which state the following [FM93] : if F is absolutely continuous with density f and right endpoint $x_* = \infty$, such that

$$\lim_{x \uparrow \infty} x \ r(x) = \alpha > 0, \tag{2.38}$$

where $r(x)$ is the *hazard function* (2.28), then $F \in D(G_{2,\alpha})$. We illustrate this with the standard Pareto distribution case (as previsouly in), that is

$$F(x) = \left(1 - (\frac{x_m}{x})^\alpha\right) 1_{x \geq x_m}, \quad \alpha > 0 \quad \text{and} \quad x_m > 0 \tag{2.39}$$

Clearly, we can see that by setting $K = x_m^\alpha$, we have

$$\bar{F}(x) = K x^{-\alpha}. \tag{2.40}$$

Therefore, we have that $a_n = (Kn)^{\alpha^{-1}}$ and $b_n = 0$.

These distributions are typically very-fat tailed (and hence, heavy-tailed, see) distributions, such that $E(X_+)^\delta = \infty$ for $\delta > \alpha$. This class of distributions is appropriate for phenomena with extremely large maxima (like...). [**?**] Common distributions include Pareto, Cauchy, Burr, stable distributions with $\alpha < 2$, etc. An example to see that, is again by Taylor expansion at the tail of $G_{2,\alpha}$ with $\alpha > 0$

$$\bar{G}_{2,\alpha}(x) = 1 - \exp(-x^{-\alpha}) \sim x^{-\alpha}, \quad \text{as } x \to \infty \tag{2.41}$$

showing that $G_{2,\alpha}$ tends to decrease as a *power law.*

III. Domain of attraction for Weibull distribution $G_{3,\alpha}$ We say that a distribution function F is in the *domain of attraction* of $G_{3,\alpha}$ (2.13) with index $\alpha > 0$ if and only if there exists finite right endpoint $x_* \in \mathbb{R}$ such that

$$\bar{F}(x_* - \frac{1}{x}) = x^{-\alpha} L(x) \tag{2.42}$$

where L is a slowly varying function.

For $F \in D(G_{3,\alpha})$, we have

$$a_n = x_* - F^{\leftarrow}(1 - \frac{1}{n}), \quad \text{and} \quad b_n = x_* \tag{2.43}$$

and hence

$$a_n^{-1}(X_{(n)} - b_n) \xrightarrow{d} G_{3,\alpha}. \tag{2.44}$$

[see references [domain of attraction course]]

Finally, we still present the ***Von Mises condition*** [FM93] related to the $G_{3,\alpha}$) domain of attraction. It states that for F having positive derivative on some $[x_0, x_*)$, with finite right endpoint $x_* < \infty$, then $F \in D(G_{3,\alpha})$ if

$$\lim_{x \uparrow x_*} (x_* - x) r(x) = \alpha > 0, \quad \text{with} \quad \int_{-\infty}^{x_*} \bar{F}(u) du < \infty \tag{2.45}$$

where $r(x)$ is again the *hazard function* defined in (2.28).

compare hazar conver gence rates of the three types

qq mots d'expl sur les formules !!!!!!

We can observe that the Weibull domain of attraction thus includes all the distribution functions that are bounded to the right ($x_* < \infty$). As most phenomena are typically bounded, we will think as the Weibull for the most attractive and flexible class for modelling extremes. But, in practice, the Fréchet one is often more preferablen extreme analysis context because allowing for arbitrarily large values.

[put general case pp.73-75 beirlant] ?

An interesting property of all these three types of domain of attraction $D(G_{k,\alpha})$ is that those are *closed* under tail-equivalence **??**. It this sense,

- For the **Gumbel** domain of attraction, let $F \in D(G_{1,\alpha})$. If H is another distribution function such that, for some $b > 0$,

$$\lim_{x \uparrow x_*} \frac{\bar{F}(x)}{\bar{H}(x)} = e^b, \tag{2.46}$$

  then H is also in the domain of attraction of the Gumbel distribution, $H \in D(G_{1,\alpha})$.

- For the **Fréchet** domain of attraction, let $F \in D(G_{2,\alpha})$. If H is another distribution function such that, for some $c > 0$,

$$\lim_{x \to \infty} \frac{\bar{F}(x)}{\bar{H}(x)} = c^\alpha, \tag{2.47}$$

  then H is also in the domain of attraction of the Fréchet distribution, $H \in D(G_{2,\alpha})$.

- For the **Weibull** domain of attraction, let $F \in D(G_{3,\alpha})$. If H is another distribution function such that, for some $c > 0$,

$$\lim_{x \uparrow x_*} \frac{\bar{F}(x)}{\bar{H}(x)} = c^{-\alpha}, \tag{2.48}$$

  then H is also in the domain of attraction of the Weibull distribution, $H \in D(G_{3,\alpha})$

2.2 Method for Minima

transposer ! enoncé modeles (~1pg) pas tout

2.3 Peaks-Over-Threshold Method

Seuils meteo : 0C (gel permanent), 25C et 30C pour les Tx 0C (gel) et 20C pour les Tn

2.3.1 Preliminaries: intuitive ideas

The threshold models relying on the *Peaks-Over-Threshold* (POT) method are useful to propose a better (?) alternative than the blocking method in **2.1**. With this new method, we consider a more natural way of determining whether an observation is extreme or not, by focusing only on all observations that are greater than a pre-specified *threshold*. It avoids the problem that can arise by considering the maximum of blocks only (), but this method also brings its own problems ().

Let's consider a sequence $\{X_j\}$ of n iid random variables having marginal distribution function F. We are then regarding for observations that exceed a well-chosen (see) threshold u, which must obviously be smaller than the right endpoint $x_* = \sup\{x : F(x) < 1\}$ of F. The aim here is to here is to find a "child" probability distribution function (fig.? -video youtube), say H, from the underlying (parent) distribution F, that will allow us to model the exceedance $Y = X - u$, and with H then expressed as $\boxed{H(y) = \Pr\{X - u \leq y | X > u\}}$. Typically, threshold models can therefore be regarded as the conditional survival function of the exceedances Y, knowing that the threshold u is exceeded [BGST06, pp.147] :

$$\Pr\{Y > y \mid Y > 0\} = \Pr\{X - u > y \mid X > u\} = \frac{\bar{F}(u + y)}{\bar{F}(u)}. \tag{2.49}$$

or in terms of the exceedance distribution function $F^{[u]}(x) = \Pr\{X \leq u + x | X > u\}$ [RT07, pp.12], [**?**] and [Ros15] :

$$F^{[u]}(x) = \frac{\Pr\{X - u \leq x, X > u\}}{\Pr\{X > u\}} = \frac{F(x + u) - F(u)}{\bar{F}(u)} \tag{2.50}$$

making use of the well-known conditional probability law. One can remark that (2.49) is actually the survivor of the exceedance distribution function, that is $\bar{F}^{[u]}$.

These intuitive characterizations we have given above about the modelling of the threshold exceedances in term of probability distribution function can be useful to understand the following.

However, if the parent distribution F were known, we would be able to compute the distribution of the threshold exceedances in (2.49). [Col01, pp.74] But as for the GEV in the method of block-maxima (section **2.1**), the distribution F is not known in practice, as we will see also in (..). Hence, and as usual in statistics[7], we must again rely on approximations. This time, we will try to approximate (2.50)

2.3.2 Characterization of the Generalized Pareto Distribution

Anageously to the *Fisher-Tippett* theorem in section **2.1** which applies for the block maxima, we have now to define a new theorem which applies for values above a predefined threshold. From this result 2.5(?), these two theorems form together the basis of Extreme Value Theory.

POT-stability [RT07, pp.25] The max-stability theorem in **??** can be applied and are formulated here by the fact that the GP distribution functions H are the only continuous one such that, for certain choice of constants a_u and b_u,

$$F^{[u]}(a_u x + b_u) = F(x). \tag{2.51}$$

This will be useful for modelling the exceedances in the following theorem (?). And for the examples (see ex. p.25)

[7] We would like to quote here the well-known phrase in statistics "*All models are wrong, but some are useful*" from George Box & Draper (1987), *Empirical model-building and response surfaces*, Wiley, p.424

***Pickands–Balkema–de Haan* theorem** discovered by [BH74] and [Iii75] which showed that the distribution of a threshold u of normalized excesses $F^{[u]}(x)(b_u + a_u x)$, as the threshold approaches the endpoint u_* of F, is the Generalized Pareto Distribution (**GPD**) $H_{\xi,\sigma_u}(y)$. That is, if X is a random variable for which (2.4) holds, and for the approximating GP distribution function possessing the same left endpoint u as the exceedance distribution function $F^{[u]}$, we have [RT07, pp.27]:

$$|F^{[u]}(x) - H_{\xi,\sigma_u}(x)| \longrightarrow 0, \qquad u \to u_*, \tag{2.52}$$

or, in an other, maybe more intuitive formulation [Col01] :

$$\Pr\{X \leq y \mid X > u\} \longrightarrow H_{\xi,\sigma_u}(y), \qquad u \to u_* \tag{2.53}$$

where the **GPD** is defined as :

$$H_{\xi,\sigma_u}(y) = 1 - \Big(1 + \frac{\xi y}{\sigma_u}\Big)_+^{-\xi^{-1}}, \qquad y > 0 \ \text{ and } \ (1 + \xi y \sigma_u^{-1}) > 0 \tag{2.54}$$

where the scale parameter is denoted σ_u to emphasize its dependency with the chosen threshold u :

$$\sigma_u = \sigma + \xi(u - \mu), \tag{2.55}$$

where one can also remark that the location parameter μ does not appear anymore in (2.54) as it does appear in 2.63.

Outline proof of the GPD and justification from GEV As we did for block-maxima approach in section **2.1.1** (2.7-2.10), we think it is interesting to have a formal and comprehensive, and still not too technical, intuitive view of where are the GPD from. Hence, we aim here at retrieving the GPD $H_{\xi,\sigma_u}(y)$ (2.53-2.54) from probability distributions as expressed in (2.49-2.50).

- We start with X having distribution function F. From the GEV theorem in section **2.1.** (see 2.4-2.5), we have for the largest order statistic, for large enough n,

$$F_{X_{(n)}}(z) = F^n(z) \approx \exp\bigg\{ - \bigg[1 + \xi\Big(\frac{z-\mu}{\sigma}\Big)\bigg]^{-\xi^{-1}}\bigg\}, \tag{2.56}$$

  with $\mu, \sigma > 0$ and ξ the GEV parameters. hence, by simply taking logarithm on both sides, we have

$$n \ln F(z) \approx -\bigg[1 + \xi\Big(\frac{z-\mu}{\sigma}\Big)\bigg]^{-\xi^{-1}} \tag{2.57}$$

- We also have that, for large enough z, by a Taylor expansion :

$$\ln F(z) \approx -[1 - F(z)]. \tag{2.58}$$

  as both sides go to zero as $x \to \infty$. Therefore, substituting into (2.57), we get the following for large u :

$$1 - F(u) \approx n^{-1}\Big[1 + \xi\Big(\frac{u-\mu}{\sigma}\Big)\Big]^{-\xi^{-1}}. \tag{2.59}$$

Or, specially expressed for our purpose of retrieving something in the form of (2.49-2.50), with $y > 0$,

$$1 - F(u+y) \approx n^{-1}\Big[1 + \xi\Big(\frac{u+y-\mu}{\sigma}\Big)\Big]^{-\xi^{-1}}. \tag{2.60}$$

- Whence we get for 2.49, with some mathematical manipulations, as $u \to u_* = \sup\{x : F(x) < 1\}$:

$$\begin{aligned}
\Pr\{X > u+y \mid X > u\} = \frac{\bar{F}(u+y)}{\bar{F}(u)} &\approx \frac{n^{-1}\big[1 + \xi\sigma^{-1}(u+y-\mu)\big]^{-\xi^{-1}}}{n^{-1}\big[1 + \xi\sigma^{-1}(u-\mu)\big]^{-\xi^{-1}}} \\
&= \Big[1 + \frac{\xi\sigma^{-1}(u+y-\mu)}{1+\xi\sigma^{-1}(u-\mu)}\Big]^{-\xi^{-1}} \\
&= \Big[1 + \frac{\xi y}{\sigma_u}\Big]^{-\xi^{-1}}
\end{aligned} \tag{2.61}$$

where σ_u is still linear in the threshold u (2.63), that is $\sigma_u = \sigma + \xi(u - \mu)$. By simply reverting the probability as in (2.50), we have then

$$\begin{aligned}
\Pr\{X - u \leq y \mid X > u\} &= 1 - \Pr\{X > u+y \mid X > u\} \\
&= 1 - \Big(1 + \frac{\xi y}{\sigma_u}\Big)^{-\xi^{-1}}
\end{aligned} \tag{2.62}$$

which is $GPD(\xi, \tilde{\sigma})$ as required and σ_u is as defined in (2.63)

More comprehension can come from [RT07, pp.-27-28] or if one wants to compute rates of convergence.

Dependence of the scale parameter σ We chose to express the scale parameter as σ_u to emphasize its dependency with the threshold u. If we increase the threshold, say to $u' > u$, then the scale parameter will be adjusted following :

$$\sigma_{u'} = \sigma_u + \xi(u' - u), \tag{2.63}$$

and in particular, this adjusted parameter $\sigma_{u'}$ will increase if $\xi > 0$ and decrease if $\xi < 0$. If $\xi = 0$, there would be no change in the scale parameter[8]. We think important to point out the fact that, similarly as mentioned for the GEV models in (2.5), the scale parameter σ_u for GPD models is not the usual standard deviation, but does govern the "size" of the excesses. [AEH$^+$13a, pp.20]

[8]This is consistent with the *memoryless property* of the exponential distribution H_{0,σ_u} (**??**), for which we give more details in

Three different types of GPD and duality with GEV One will remark the similarity with the GEV distributions as the parameters of the GPD of the threshold excesses are uniquely determined by the corresponding GEV distribution parameters of block-maxima (see outline proof in the above to convince yourself). Hence, the shape parameter ξ of the GPD is equal to that of the corresponding GEV and, most of all, it is invariant[9] while the computation of σ_u will not be affected by changes of the corresponding μ or σ in the GEV, from the self-compensation arising in (2.63). [Col01, pp.76]

Hence, as for the block-maxima approach, there are also three possible families of the GPD depending on the value of the shape parameter ξ which determines the qualitative behaviour of the GPD. [**?**], [**?**]

- The **first** type, call it $H_{0,\sigma_u}(y)$, comes by letting the shape parameter $\xi \to 0$ in 2.54, giving :

$$H_{0,\sigma_u}(y) = 1 - \exp\Big(-\frac{y}{\sigma_u}\Big), \qquad y > 0 \tag{2.64}$$

  where we recognize that it corresponds to an **exponential distribution** with parameter $1/\sigma_u$, that is $Y \sim \exp(\sigma_u^{-1})$.

- The **second** and the **third** types, that is when $\xi < 0$ and $\xi > 0$ (resp.), differ only by their support :

$$H_{\xi,\sigma_u}(y) = 1 - \left(1 + \frac{\xi y}{\sigma_u}\right)^{-\xi^{-1}} \quad \text{for} \quad \begin{cases} y > 0 & \text{if} \quad \xi > 0; \\ 0 < y < -\sigma_u/|\xi| & \text{if} \quad \xi < 0. \end{cases} \tag{2.65}$$

  Therefore, if $\xi > 0$ the corresponding GPD has no upper limit and when $\xi < 0$, the associated GPD has an upper bound in $y_* = -\sigma_u/|\xi|$. ??

Some plots ?

After looking at the behaviour of the density of these functions, we will procure a more comprehensive view by defining some examples of how to retrieve these different types of Generalized Pareto Distributions.

Density functions of the GPD

Examples of the GPD as limiting distribution for exceedances We have seen in the previous paragraph that if we can have an approximate distribution G for block-maxima, then threshold excess will have a corresponding distribution given by a member of the Generalized Pareto family. Whence the shape parameter ξ, as for GEV distributions, is determinant for controlling the behaviour of the GPD, and thus leads to the three different types in ((2.64)-(2.65)).

1. The first type

The choice of a threshold will be discussed in section **3.5.1**.

From [BGST06, p.147-],

[9] For instance, choosing different block size in the GEV modelling would shift its (estimated) parameters while GPD (estimated) parameters are *stable*.

2.4 Stationary Series

From now, we considered $X_{(n)} = \max_{1\leq i\leq n} X_i$ where we have assumed $X_1, ..., X_n$ are independent random variables. For sake of simplicity, we abandon this notation. In the sequel, this will be denoted by $\tilde{X}_{(n)} = \max_{1\leq i\leq n} \tilde{X}_i$ where $\tilde{X}_1, ..., \tilde{X}_n$ will typically denote a sequence of independent random variables, so that the maximum $\tilde{X}_{(n)}$ is composed of (plays with) independent random variables only. We are now interested by modelling $X_{(n)} = \max_{1\leq i\leq n} X_i$ where the $\{X_i\}$ will now denote a *stationary* sequence of random variables sharing the same marginal distribution as $\{\tilde{X}_i\}$, F.

Stationarity We say that the sequence $\{X_i\}$ is stationary if

More generally, for $h \geq 0$ and $n \geq 1$, the distribution of the lagged random vector $(X_{1+h}, ..., X_{n+h})$ does not depend on h when the sequence is said to be (strongly) stationary.

Condition $\boldsymbol{D(u_n)}$ for stationary sequences says, following [BGST06, Col01, pp.373-374, pp.93]

This condition ensures that, when the sets of variables are separated by a relatively short distance, typically $s_n = o(n)$, the long-range dependence between such events is limited, in a sense that is sufficiently close to zero to have no effect on the limit extremal laws.

From this result, we can retrieve the *extreme-value theorem*

Result is remarkable in the sense that, provided a series has limited long-range dependence at extreme levels (D(u_n)) condition makes precise), maxima of stationary series follow the same distributional limit laws as those of independent series. [S.Coles 2001 p.94]

2.4.1 The extremal index

.. However, the maximum has a tendency to decrease as .... [Col01, pp.96]

2.5 Nonstationary Series

2.6 Return Levels and Return Periods

Assuming for this introductory example our time unit reference is in year as usually assumed in climatological analysis, which is very common for climatic data, let us consider the *m-year return level* r_m and define it as the high quantile for which the probability that the annual maximum exceeds this quantile is 1/m, which is called the *return period*.

That is, we can directly retrieve it from our estimation of the three GEV parameters

$$F(r_m) = \Pr\{X_{(n),y} \leq r_m\} = 1 - 1/m \tag{2.66}$$

$$\Big[1 + \xi\sigma^{-1}(r_m - \mu)\Big]^{-\xi^{-1}} = \frac{1}{m} \tag{2.67}$$

leading directly to (after some calculations)

$$r_m = \begin{cases} \mu + \sigma\xi^{-1}(m^{\xi} - 1) & \text{if } \xi \neq 0 \\ \mu + \sigma \log(n) & \text{if } \xi = 0 \end{cases}$$

(See for eq. index ())

However, we recall that the definition of return period is easily misinterpreted and the given above is thus not universally accepted. To evaporate (vanish) this issue, it is important to distinguish stationary from non-stationary sequences.

2.6.1 Stationarity

Under an assumption of a stationary sequence, the return level is the same for all years, and this gives rise to the notion of the return period (or m-year event). Hence, the return period of a particular event is the inverse of the probability that the event will be exceeded in any given year. The m-year return level is associated with a return period of m years. However, there are two main interpretations in this context for return periods.

[AEH+13b, pp.100]

Denoting $X_{(n),y}$ the annual maximum for year y. Omitting the notational dependence on block size n, we assume $\{X_{(n),y}\} \overset{iid}{\sim} F$.

1. The first interpretation of the m-year event is **the expected waiting time until an exceedance occurs**. To see that, letting T be the year of the first exceedance, we we can write

ok pas depassé

$$\begin{aligned}
\Pr\{T = t\} &= \Pr\{X_{(n),1} \leq r_m, ..., X_{(n),t-1} \leq r_m, X_{(n),t} > r_m\} \\
&= \Pr\{X_{(n),1} \leq r_m\}...\Pr\{X_{(n),t-1} \leq r_m\}\Pr\{X_{(n),t} > r_m\} && \text{[from iid assumption]} \\
&= \Pr\{X_{(n),1} \leq r_m\}^{t-1} \Pr\{X_{(n),1} > r_m\} && \text{[from stationarity]} \\
&= F^{t-1}(r_m)(1 - F(r_m)) \\
&= (1 - 1/m)^{t-1}(1/m)
\end{aligned} \tag{2.68}$$

   We easily recognize that T has geometric density with parameter $1/m$. From simple properties of geometric distributions, we found its expected values is $1/(1/m)$, that is the expected waiting time for an m-year event is m years.

2. The second interpretation of an m-year event is that **the expected number of events in a period of m years is exactly 1**. To see that, define $N = \sum_{y=1}^{m} I(X_{(n),y} > r_m)$ as the random variable representing the number of exceedances in m years (where I is indicator function). We can view each year as a "trial", and from the fact that we have assumed $\{X_{(n),y}\}$ are iid, we can compute the probability that the number of exceedances in m-years is k

$$\Pr\{N = k\} = \binom{m}{k}(1/m)^k(1 - 1/m)^{m-k} \tag{2.69}$$

   from which we recognize a well-know distribution, that is $N \sim$ Binomial$(m, 1/m)$. Again from properties of this distribution, we easily find that N has an expected value of 1.

2.6.2 Non-stationarity

From the definition on non-stationary process, the modelling of return period will change over time. Hence, we introduce the notation of the distribution function F_y of a particular $X_{(n),y}$. We must study

$p(y) = \Pr(X_{(n),y} > r) = 1 - F_y(r)$. If we estimate F_y, we can retrieve easily p(y). $F_y(r_p(y)) = 1 - p$ with the exceedance level $r_p(y)$ changing with year. It shows the changing nature of "risk".

Return period as expected waiting time

Return period as expected number of events

2.7 Point Process Approach

Following [Col01],

Furthermore, we will see that the parametrization of the point process model is invariant to threshold choice so that this variation would only affect the well-known (already mentioned) bias-variance trade-off in the inference. Interesting if seasonal modelling.

3 Methods of Inference

We decide to present in this section the two mains methods of inference for GEV distributions. First, the likelihood-based methods for their wide applicability, and easy interpretability. Then, we will broadly present the bayesian methods for their increasing supports in this domain, easily adjustable. Finally, we will present some other well-known methods that are also widely used to estimate GEV parameters like the Hill or the moment estimator.

3.1 Likelihood-based Methods

(return level) [extremesin climate change p.106] "Approximate confidence intervals for the return level can be obtained by the delta method (Casella and Berger, 2002, Sect. 5.5.4) which relies on the asymptotic normality of maximum-likelihood estimators and produces a symmetric confidence interval. Alternatively, profile likelihood methods (Coles 2001, Sect. 2.6.6) provide asymmetric confidence intervals, which better capture the skewness generally associated with return level estimates." stationary case

where the density function of the GEV distribution $g_\xi(x)$ in the previous section in 2.15.

3.2 Bayesian Methods

see evdbayes pdf package r

3.3 Other Methods

[BGST06, pp.140]

3.3.1 The Hill estimator

3.3.2 Pickands estimator

3.3.3 The moment estimator

3.3.4 Estimators based on generalized quantile

3.4 Bootstrap Methods

3.5 Model Selection, Diagnostics and Graphical Tools

All QQ-plots

Gumbel plots

Z and W statistic plots

3.5.1 Threshold choice for excess of a threshold models

Mean residual life function or *mean excess function* , following again [BGST06, pp.14-19], [Col01, pp.78-80],

$$\begin{aligned} mrl(u_0) &:= E(X - u_0 \mid X > u_0) \\ &= \frac{\int_{u_0}^{x_+}(1 - F(u))du}{\bar{F}(u_0)} \end{aligned} \tag{3.1}$$

for X having survival function $\bar{F}(u_0)$ computed at u_0, with $x_* = sup\{x : F(x) < 1\}$ denoting the right endpoint of the support of F. It denotes, in an actuarial context, the expected remaining quantity or amount to be paid out when a level u_0 has been chosen. However, even if it is mainly applied in an actuarial context or in survival analysis in the literature (see [**?**] for a weel-known example), there are also interesting and reliable applications in our more environmental purposes as we will see in the following.

This can be particularly interesting for our purpose when considering threshold models. For this case, we can suppose the excesses of a threshold generated by the sequence $\{X_i\}$ follow a generalized Pareto distribution (see 2.2). Knowing the theoretical mean of this distribution, we retrieve, provided the shape parameter $\xi < 1$ and denoting σ_u the scale parameter corresponding to excess of a threshold $u > u_0$,

$$\begin{aligned} mrl(u) := E(X - u \mid X > u) &= \frac{\sigma_u}{1 - \xi} \\ &= \frac{\sigma_{u_0} + \xi u}{1 - \xi} \end{aligned} \tag{3.2}$$

from the threshold u dependence with the scale parameter σ (see 2.63). Hence, we remark that $mrl(u)$ is linearly increasing in u. Furthermore, we can estimate empirically this function intuitively by

$$\widehat{mrl}(u) = \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{[i]} - u) \tag{3.3}$$

where the $x_{[i]}$'s denote the n_u observations that exceed u. This leads lead to an interesting tool for our purpose, the *mean residual life plot*. It comes from combining the linearity detected between $mrl(u)$ and u in (3.2) with (3.3). Therefore, a reliable information can be retrieved from the point of the points

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{[i]} - u) \right) : u < x_{max} \right\}. \tag{3.4}$$

Even if its interpretation is not easy, this graphical procedure will give insights for the choice of a suitable threshold u_0 to model extremes via general Pareto distribution, that is the threshold u_0 above which we can detect linearity in the plot. Relying on this well-chosen threshold u_0, the generalized Pareto distribution should be a good approximation. Remind however that its interpretation is often subjective. Furthermore, information in the far right-hand-side of this plot is unreliable. Variability is high due to the limited amount of data (exceedances) above very high thresholds. This can be seen for example on larger confidence intervals.

From [Col01, pp.83]

Fromm [Col01, pp.84]

3.5.2 Dispersion index plot

4 Dealing with Non-Stationary Sequences

As we are dealing with time varying sequences, we can

5 Conclusion

"Another approach would be to use something other than time as the covariate in the model. For instance, one could imagine linking temperature data directly to CO2 level rather than time. However, linking to a climatological covariate makes extrapolation into the future more difficult, as one would need to extrapolate the covariate as well. No obvious climatological covariate comes to mind for the Red River application. "

Timescale-uncertainty effects on extreme value analyses seem not to have been studied yet. For stationary models (Sect. 6.2), we anticipate sizable effects on block extremes–GEV estimates only when the uncertainties distort strongly the blocking procedure. For nonstationary models (Sect. 6.3), one may augment confidence band construction by inserting a timescale simulation step (after Step 4 in Algorithm 6.1) [Mud14, pp.262]

(!!!!! delete not useful equation numbering!!!!)

A Appendices

A.1 Statistical concepts

A.1.1 Convergence concepts

Weakly convergence We say that a sequence of random variables X_n *converges weakly* to

A.1.2 Varying functions !!!!

A.2 Figures

Beirlant: 2004 !!

Bibliography

[AEH+13a] Amir AghaKouchak, David Easterling, Kuolin Hsu, Siegfried Schubert, and Soroosh Sorooshian, editors. *Extremes in a Changing Climate*, volume 65 of *Water Science and Technology Library*. Springer Netherlands, Dordrecht, 2013.

[AEH+13b] Amir AghaKouchak, David Easterling, Kuolin Hsu, Siegfried Schubert, and Soroosh Sorooshian, editors. *Extremes in a Changing Climate*, volume 65 of *Water Science and Technology Library*. Springer Netherlands, Dordrecht, 2013.

[BGST06] Jan Beirlant, Yuri Goegebeur, Johan Segers, and Jozef Teugels. *Statistics of Extremes: Theory and Applications*. John Wiley & Sons, March 2006. Google-Books-ID: jqmRwfG6aloC.

[BH74] A. A. Balkema and L. de Haan. Residual Life Time at Great Age. *The Annals of Probability*, 2(5):792–804, October 1974.

[BTV96] Jan Beirlant, Jozef L. Teugels, and Petra Vynckier. *Practical Analysis of Extreme Values*. Leuven University Press, 1996. Google-Books-ID: ylR3QgAACAAJ.

[Col01] Stuart Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer London, London, 2001.

[FM93] Michael Falk and Frank Marohn. Von Mises Conditions Revisited. *The Annals of Probability*, 21(3):1310–1328, July 1993.

[FT28] R. A. Fisher and Leonard H. C. Tippett. Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. *ResearchGate*, 24(02):180–190, January 1928.

[Gne43] B. Gnedenko. Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire. *Annals of Mathematics*, 44(3):423–453, 1943.

[HF06] L. de Haan and Ana Ferreira. *Extreme value theory: an introduction*. Springer series in operations research. Springer, New York ; London, 2006. OCLC: ocm70173287.

[Iii75] James Pickands Iii. Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics*, 3(1):119–131, January 1975.

[KTWK+02] A. M. G. Klein Tank, J. B. Wijngaard, G. P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A. F. V. van Engelen, E. Forland, M. Mietus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar, J. Antonio López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L. V. Alexander, and P. Petrovic. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology*, 22(12):1441–1453, October 2002.

[LLR83] M. R. Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer New York, New York, NY, 1983.

[Mud14] Manfred Mudelsee. *Climate Time Series Analysis*, volume 51 of *Atmospheric and Oceanographic Sciences Library*. Springer International Publishing, Cham, 2014.

[PF15] E. C. Pinheiro and S. L. P. Ferrari. A comparative review of generalizations of the Gumbel extreme value distribution with an application to wind speed data. *arXiv:1502.02708 [stat]*, February 2015. arXiv: 1502.02708.

[Res87] Sidney I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer New York, New York, NY, 1987.

[Ros15] Gianluca Rosso. Extreme Value Theory for Time Series using Peak-Over-Threshold method. *arXiv preprint arXiv:1509.01051*, 2015.

[RT07] Rolf-Dieter Reiss and Michael Thomas. *Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields ; [includes CD-ROM]*. Birkhäuser, Basel, 3. ed edition, 2007. OCLC: 180885018.