

The generalized Pareto process; with a view towards application and simulation

ANA FERREIRA^{1,3} and LAURENS DE HAAN^{2,3}

¹ISA, Univ Tecn Lisboa, Tapada da Ajuda 1349-017 Lisboa, Portugal. E-mail: anafh@isa.utl.pt

²Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

E-mail: ldehaan@ese.eur.nl

³CEAUL, FCUL, Bloco C6 – Piso 4 Campo Grande, 749-016 Lisboa, Portugal

In extreme value statistics, the peaks-over-threshold method is widely used. The method is based on the generalized Pareto distribution characterizing probabilities of exceedances over high thresholds in \mathbb{R}^d . We present a generalization of this concept in the space of continuous functions. We call this the generalized Pareto process. Differently from earlier papers, our definition is not based on a distribution function but on functional properties, and does not need a reference to a related max-stable process.

As an application, we use the theory to simulate wind fields connected to disastrous storms on the basis of observed extreme but not disastrous storms. We also establish the peaks-over-threshold approach in function space.

Keywords: domain of attraction; extreme value theory; functional regular variation; generalized Pareto process; max-stable processes; peaks-over-threshold

1. Introduction

Let $C(S)$ be the space of continuous real functions on S , equipped with the supremum norm, where S is a compact subset of \mathbb{R}^d . A stochastic process X in $C(S)$ is in the domain of attraction of some max-stable process Y if there are continuous functions $a_n(s)$ positive and $b_n(s)$ on S such that the processes

$$\left\{ \max_{1 \leq i \leq n} \frac{X_i(s) - b_n(s)}{a_n(s)} \right\}_{s \in S}$$

with X, X_1, \dots, X_n independent and identically distributed (i.i.d.), converge in distribution to Y in $C(S)$. Necessary and sufficient conditions are: uniform convergence of the marginal distributions and a convergence of measures (in fact a form of regular variation):

$$\lim_{t \rightarrow \infty} tP(T_t X(\cdot) \in A) = \nu(A), \tag{1.1}$$

where:

- $T_t X(s) = (1 + \gamma(s) \frac{X(s) - b_t(s)}{a_t(s)})_+^{1/\gamma(s)}$ for all $s \in S$ (with the notation $x_+ = \max(0, x)$ for any real x),
- ν is a homogeneous measure of order -1 on $C^+(S) := \{f \in C(S): f \geq 0\}$ and,
- A is any Borel subset of $C^+(S)$ satisfying $\inf\{\sup_{s \in S} f(s): f \in A\} > 0$ and $\nu(\partial A) = 0$.

Cf. de Haan and Lin [4] and de Haan and Ferreira [3], Section 9.5. Although in these references only the case $C = [0, 1]$ has been worked out, all arguments are valid for any compact subset of \mathbb{R}^d (as remarked in these references). The functions $a_t(s)$ and $b_t(s)$ are chosen in such a way that the marginal distributions are in standard form,

$$\lim_{t \rightarrow \infty} tP\left(\frac{X(s) - b_t(s)}{a_t(s)} > x\right) = (1 + \gamma(s)x)^{-1/\gamma(s)}, \quad 1 + \gamma(s)x > 0,$$

for all $x \in \mathbb{R}$ and $s \in S$. Here γ is a continuous function. In particular, one may take $b_t(s) := \inf\{x: P(X(s) \leq x) \geq 1 - 1/t\}$. This is how we choose $b_t(s)$ from now on. One possible choice of $a_t(s)$ is $a_t(s) := \gamma(s)(b_{2t}(s) - b_t(s))/(2^{\gamma(s)} - 1)$.

From (1.1), it follows that

$$\frac{P((1 + \gamma(\cdot)(X(\cdot) - b_t(\cdot))/a_t(\cdot))_+^{1/\gamma(\cdot)} \in A)}{P(\sup_{s \in S}((X(s) - b_t(s))/a_t(s)) > 0)}$$

converges, as $t \rightarrow \infty$, and so does

$$P\left(\left(1 + \gamma(\cdot)\frac{X(\cdot) - b_t(\cdot)}{a_t(\cdot)}\right)_+^{1/\gamma(\cdot)} \in A \mid \sup_{s \in S} \frac{X(s) - b_t(s)}{a_t(s)} > 0\right).$$

The limit constitutes a probability distribution on $C^+(S)$.

This reasoning is similar to how one obtains the generalized Pareto distributions in \mathbb{R} (Pickands [11]; Balkema and de Haan [1]) and in \mathbb{R}^d (Rootzén and Tajvidi [13]; Falk, Hüslér and Reiss [8]). It leads to what we call generalized Pareto processes.

The paper is organized as follows. The Pareto processes will be dealt with in Section 2. As in the finite-dimensional context, it is convenient to study first generalized Pareto processes in a standardized form. This is done in Section 2.1. The general process is discussed in Section 2.3. In Section 2.2, a discrete version of our approach is discussed leading to simple multivariate Pareto random vectors. The domain of attraction is discussed in Section 3. In Section 4, we show that by using the stability property of generalized Pareto processes one can create extreme storm fields starting from independent and identically observations of storm fields. We also establish the peaks-over-threshold approach in function space.

In the following, operations like $w_1 + w_2$ or $w_1 \wedge w_2$ with $w_1, w_2 \in C(S)$ mean, respectively, $\{w_1(s) + w_2(s)\}_{s \in S}$ and $\{w_1(s) \wedge w_2(s)\}_{s \in S}$. Then, with abuse of notation, operations like $w + x$ or $w \wedge x$ with $w \in C(S)$ and $x \in \mathbb{R}$ mean, respectively, $\{w(s) + x\}_{s \in S}$ and $\{w(s) \wedge x\}_{s \in S}$. Similarly for products and powers. Then, for example, we shall simply write $(1 + \gamma \frac{X - b_t}{a_t})^{1/\gamma}$ for $\{(1 + \gamma(s) \frac{X(s) - b_t(s)}{a_t(s)})^{1/\gamma(s)}\}_{s \in S}$, with $X = \{X(s)\}_{s \in S}$, $a_t = \{a_t(s)\}_{s \in S}$, $b_t = \{b_t(s)\}_{s \in S}$ and $\gamma = \{\gamma(s)\}_{s \in S}$.

Denote the class of Borel subsets of a metric space by $\mathcal{B}(\cdot)$.

2. Pareto processes

2.1. The simple Pareto process

Again, let $C^+(S)$ be the space of non-negative real continuous functions on S , with S some compact subset of \mathbb{R}^d .

Theorem 2.1. *Let W be a stochastic process in $C^+(S)$ and ω_0 a positive constant. The following three statements are equivalent:*

1. (Peaks-over-threshold):

- (a) The expectation $E(W(s)/\sup_{u \in S} W(u))$ is positive for all $s \in S$,
- (b) $P(\sup_{s \in S} W(s)/\omega_0 > x) = x^{-1}$, for $x > 1$ (standard Pareto distribution),
- (c)

$$P\left(\frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B \mid \sup_{s \in S} W(s) > r\right) = P\left(\frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) \tag{2.1}$$

for all $r > \omega_0$ and $B \in \mathcal{B}(\bar{C}_{\omega_0}^+(S))$ with

$$\bar{C}_{\omega_0}^+(S) := \left\{ f \in C^+(S) : \sup_{s \in S} f(s) = \omega_0 \right\}. \tag{2.2}$$

2. (Random functions):

- (a) $P(\sup_{s \in S} W(s) > \omega_0) = 1$,
- (b) $E(W(s)/\sup_{u \in S} W(u)) > 0$ for all $s \in S$,
- (c)

$$P(W \in rA) = r^{-1}P(W \in A) \tag{2.3}$$

for all $r > 1$ and $A \in \mathcal{B}(C_{\omega_0}^+(S))$, where rA means the set $\{rf, f \in A\}$, and

$$C_{\omega_0}^+(S) := \left\{ f \in C^+(S) : \sup_{s \in S} f(s) \geq \omega_0 \right\}. \tag{2.4}$$

3. (Constructive approach) $W(s) = YV(s)$, for all $s \in S$, for some Y and $V = \{V(s)\}_{s \in S}$ satisfying:

- (a) $V \in C^+(S)$ is a stochastic process satisfying $\sup_{s \in S} V(s) = \omega_0$ a.s., and $EV(s) > 0$ for all $s \in S$,
- (b) Y is a standard Pareto random variable, $P(Y \leq y) = 1 - 1/y$, $y > 1$,
- (c) Y and V are independent.

Definition 2.1. *The process W characterized in Theorem 2.1, with threshold parameter ω_0 , is called simple Pareto process. The probability measure in (2.1), that is,*

$$\rho(B) = P\left(\frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) \quad \text{for } B \in \mathcal{B}(\bar{C}_{\omega_0}^+(S)) \tag{2.5}$$

is called the spectral measure.

Some easy consequences of Theorem 2.1(3) are the following. The process W is stationary if and only if V is stationary. Independence at any two points $s_1, s_2 \in S$, that is, $W(s_1)$ and $W(s_2)$ being independent, is not possible. Complete dependence is equivalent to $V \equiv \omega_0$ a.s. We shall come back to some of these issues.

Proof of Theorem 2.1. We start by proving that 1 implies 3. By compactness and continuity, $\sup_{s \in S} W(s) < \infty$ a.s. Take:

$$Y = \frac{\sup_{s \in S} W(s)}{\omega_0} \quad \text{and} \quad V = \frac{\omega_0 W}{\sup_{s \in S} W(s)}.$$

Then (a), (b) and (c) are straightforward.

Next, we prove that 3 implies 2. Let

$$A_{r,B} = \left\{ f \in C^+(S): \sup_{s \in S} f(s)/\omega_0 > r, \frac{\omega_0 f}{\sup_{s \in S} f(s)} \in B \right\} = r A_{1,B}$$

for all $r > 1$ and $B \in \mathcal{B}(\bar{C}_{\omega_0}^+(S))$. Then,

$$\begin{aligned} P(W \in A_{r,B}) &= P\left(\sup_{s \in S} W(s)/\omega_0 > r, \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) \\ &= P(Y > r, V \in B) = P(Y > r)P(V \in B) \\ &= \frac{1}{r} P\left(\sup_{s \in S} W(s)/\omega_0 > 1, \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) \\ &= \frac{1}{r} P(W \in A_{1,B}) \end{aligned}$$

using in particular the independence of Y and V and $P(\sup_{s \in S} W(s)/\omega_0 > 1) = 1$. Since $P(rA) = r^{-1}P(A)$ holds for any of the above sets, it must also hold for all Borel sets in the statement.

Finally, check that 2 implies 1. For any $r > 1$, by (c) and (a),

$$P\left(\frac{\sup_{s \in S} W(s)}{\omega_0} > r\right) = \frac{1}{r} P\left(\frac{\sup_{s \in S} W(s)}{\omega_0} > 1\right) = \frac{1}{r}.$$

Also for any $B \in \mathcal{B}(\bar{C}_{\omega_0}^+(S))$,

$$\begin{aligned} P\left(\sup_{s \in S} W(s)/\omega_0 > r, \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) &= \frac{1}{r} P\left(\sup_{s \in S} W(s)/\omega_0 > 1, \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) \\ &= \frac{1}{r} P\left(\frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) \end{aligned}$$

since $\sup_{s \in S} W(s) > \omega_0$ holds a.s. That is, it follows that $\sup_{s \in S} W(s)/\omega_0$ is univariate Pareto distributed and, $\sup_{s \in S} W(s)$ and $W/\sup_{s \in S} W(s)$ are independent. \square

The following properties are direct consequences:

Corollary 2.1. For any simple Pareto process W , the random variable $\omega_0^{-1} \sup_{s \in S} W(s)$ has standard Pareto distribution.

Corollary 2.2. $W \in C^+(S)$ is a simple Pareto process if and only if any of the two equivalent statements hold:

1. (a) $E(W(s)/\sup_{u \in S} W(u)) > 0$ for all $s \in S$,
- (b) $P(\sup_{s \in S} W(s)/\omega_0 > x) = x^{-1}$, for $x > 1$,
- (c)

$$P\left(W \in rA \mid \sup_{s \in S} W(s) > r\omega_0\right) = P(W \in A) \tag{2.6}$$

for all $r > 1$ and $A \in \mathcal{B}(C_{\omega_0}^+(S))$.

2. (a) $E(W(s)/\sup_{u \in S} W(u)) > 0$ for all $s \in S$,
- (b)

$$P\left(\sup_{s \in S} \frac{W(s)}{\omega_0} > r, \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right) = \frac{\rho(B)}{r} \tag{2.7}$$

for all $r > 1$ and $B \in \mathcal{B}(\bar{C}_{\omega_0}^+(S))$.

From (2.6), we see that the probability distribution of W serves in fact as the exponent measure in max-stable processes (cf. de Haan and Ferreira [3], Section 9.3). Characterization 2 suggests ways for testing and modeling Pareto processes.

We proceed to express the distribution function of W in terms of the probability distribution of V from Theorem 2.1(3) and Definition 2.1.

Let $w, W \in C^+(S)$. The notation $W \leq w$ will mean $W(s) \leq w(s)$ for all $s \in S$. Similarly for $W > w$ and $W \not\leq w$. Clearly, the latter two are not the same.

Take for the conditional expectation,

$$E(g(V)|V \in B) = \frac{1}{\rho(B)} \int_B g(v) d\rho(v), \quad B \in \mathcal{B}(\bar{C}_{\omega_0}^+(S)),$$

defined in the usual sense and whenever $\rho(B) = P(V \in B) > 0$, with g a real functional (e.g., see Billingsley [2], Section 34).

Proposition 2.1. Let $w, W \in C^+(S)$, with W a simple Pareto process. Let $S_0 = \{s \in S: w(s) = 0\}$, $\bar{S}_0 = S \setminus S_0$ the complement of S_0 , and $B_0 = \{f \in \bar{C}_{\omega_0}^+(S): \inf_{s \in \bar{S}_0} \frac{w(s)}{f(s)} \geq 1 \text{ and } f(s) = 0 \text{ for } s \in S_0\}$. Then

$$P(W \leq w) = \begin{cases} \rho(B_0) \left\{ 1 - E\left(\sup_{s \in \bar{S}_0} \frac{V(s)}{w(s)} \mid V \in B_0\right) \right\}, & \text{if } \rho(B_0) > 0, \\ 0, & \text{if } \rho(B_0) = 0. \end{cases} \tag{2.8}$$

Proof.

$$\begin{aligned}
 P(W \leq w) &= P(W(s) \leq w(s) \text{ for } s \in \bar{S}_0 \text{ and } W(s) \leq w(s) \text{ for } s \in S_0) \\
 &= P\left(Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \text{ and } V(s) = 0 \text{ for } s \in S_0\right) \\
 &= P\left(Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \text{ and } V(s) = 0 \text{ for } s \in S_0 \text{ and } \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \geq 1\right) \\
 &\quad + P\left(Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \text{ and } V(s) = 0 \text{ for } s \in S_0 \text{ and } \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} < 1\right) \\
 &= \int_{B_0} P\left(Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{v(s)}\right) d\rho(v) \\
 &= \int_{B_0} 1 - \sup_{s \in \bar{S}_0} \frac{v(s)}{w(s)} d\rho(v) = \rho(B_0) - \int_{B_0} \sup_{s \in \bar{S}_0} \frac{v(s)}{w(s)} d\rho(v) \\
 &= \rho(B_0) \left\{ 1 - E\left(\sup_{s \in \bar{S}_0} \frac{V(s)}{w(s)} \mid V \in B_0\right) \right\},
 \end{aligned}$$

where the last but two equality follows by the fact that the second summand in the previous equality is zero and from the independence of Y and V . □

Corollary 2.3. *Under the conditions of Proposition 2.1,*

$$P(W \leq w) = 1 - E\left(\sup_{s \in \bar{S}_0} \frac{V(s)}{w(s)}\right) \quad \text{if } \rho(B_0) = 1. \tag{2.9}$$

The following is obtained in the particular case of w being strictly positive.

Proposition 2.2. *Let $w, W \in C^+(S)$, with w positive and W a simple Pareto process. Then*

$$P(W \leq w) = E\left(\sup_{s \in S} \frac{V(s)}{w(s) \wedge \omega_0}\right) - E\left(\sup_{s \in S} \frac{V(s)}{w(s)}\right). \tag{2.10}$$

Proof. (i) First, consider the case $\inf_{s \in S} w(s) \geq \omega_0$. Use Theorem 2.1, part 3,

$$P(W \leq w) = P(YV \leq w) = P\left(Y \leq \inf_{s \in S} \frac{w(s)}{V(s)}\right) = 1 - E\left(\sup_{s \in S} \frac{V(s)}{w(s)}\right) \tag{2.11}$$

hence,

$$P(W \not\leq w) = E\left(\sup_{s \in S} \frac{V(s)}{w(s)}\right). \tag{2.12}$$

(ii) The probability measure of W on $C_{\omega_0}^+(S)$ can be extended to a measure ν on $C^+(S)$ while keeping the homogeneity relation (2.3) as follows: for any Borel set B such that

$$\sup_{f \in B} \sup_{s \in S} f(s) \leq \omega_0 \quad \text{and} \quad 0 < \varepsilon < \inf_{f \in B} \sup_{s \in S} f(s),$$

we define

$$\nu(B) := \frac{\omega_0}{\varepsilon} P\left(W \in \frac{\omega_0}{\varepsilon} B\right).$$

This measure (the same as in (1.1)) is homogeneous of order -1 :

$$\nu(rB) = r^{-1}\nu(B) \quad \text{for all } r > 0 \quad \text{and} \quad B \in \mathcal{B}(C^+(S)).$$

Then, the probability distribution of W is the restriction of ν to $C_{\omega_0}^+(S)$, that is, for $B \in \mathcal{B}(C^+(S))$,

$$P(W \in B) = \nu\left\{f \in B, \sup_{s \in S} f(s) > \omega_0\right\}. \tag{2.13}$$

Hence, by the homogeneity property of ν , (2.13) and (2.12) in that order:

$$\begin{aligned} \nu\{f \not\leq w\} &= \frac{\omega_0}{\inf_{s \in S} w(s)} \nu\left\{f \not\leq \frac{w\omega_0}{\inf_{s \in S} w(s)}\right\} \\ &= \frac{\omega_0}{\inf_{s \in S} w(s)} P\left(W \not\leq \frac{w\omega_0}{\inf_{s \in S} w(s)}\right) = E\left(\sup_{s \in S} \frac{V(s)}{w(s)}\right). \end{aligned} \tag{2.14}$$

By (2.13), elementary set-measure operations and (2.14) in that order:

$$\begin{aligned} P(W \not\leq w) &= \nu\{f \not\leq w, f \not\leq \omega_0\} \\ &= \nu\{f \not\leq w\} + \nu\{f \not\leq \omega_0\} - \nu\{f \not\leq w \text{ or } f \not\leq \omega_0\} \\ &= \nu\{f \not\leq w\} + \nu\{f \not\leq \omega_0\} - \nu\{f \not\leq w \wedge \omega_0\} \\ &= E\left(\sup_{s \in S} \frac{V(s)}{w(s)}\right) + 1 - E\left(\sup_{s \in S} \frac{V(s)}{w(s) \wedge \omega_0}\right). \end{aligned}$$

□

Note that $E(\sup_{s \in \bar{S}_0} \frac{V(s)}{w(s) \wedge \omega_0} | V \in B_0) = 1$, whenever $\rho(B_0) > 0$, which links the results of Propositions 2.1 and 2.2.

The following formulas might also be useful.

Corollary 2.4. *Let $w, W \in C^+(S)$, with W a simple Pareto process. Then:*

(a) With $B_1 = \{f \in \bar{C}_{\omega_0}^+(S) : \sup_{s \in S} \frac{w(s)}{f(s)} > 1 \text{ and } \inf_{s \in S} f(s) > 0\}$,

$$P(W > w) = \begin{cases} \rho(B_1)E\left(\inf_{s \in S} \frac{V(s)}{w(s)} \middle| V \in B_1\right), & \text{if } \rho(B_1) > 0, \\ 0, & \text{if } \rho(B_1) = 0. \end{cases} \tag{2.15}$$

In particular, if $P(V > 0) > 0$ and $\sup_{s \in S} w(s) > \omega_0$,

$$P(W > w) = P(V > 0)E\left(\inf_{s \in S} \frac{V(s)}{w(s)} \middle| V > 0\right). \tag{2.16}$$

(b) If $w > 0$ and $\sup_{s \in S} w(s) > \omega_0$,

$$P(W > w) = E\left(\inf_{s \in S} \frac{V(s)}{w(s)}\right). \tag{2.17}$$

(c) If $E(\inf_{s \in S} V(s)) > 0$, for $x \in \mathbb{R}$,

$$P(W > x | W > \omega_0) = \begin{cases} 1, & x \leq \omega_0, \\ \omega_0/x, & x > \omega_0. \end{cases} \tag{2.18}$$

(d) If $E(\inf_{s \in S} V(s)) > 0$, for $x \in \mathbb{R}$ and for each $s \in S$,

$$P(W(s) > x | W(s) > \omega_0) = \begin{cases} 1, & x \leq \omega_0, \\ \omega_0/x, & x > \omega_0. \end{cases} \tag{2.19}$$

Proof. For (2.15), similarly to the proof of Proposition 2.1,

$$\begin{aligned} P(W > w) &= P\left(Y \geq \sup_{s \in S} \frac{w(s)}{V(s)} \text{ and } \inf_{s \in S} V(s) > 0\right) \\ &= \int_{B_1} \inf_{s \in S} \frac{v(s)}{w(s)} d\rho(v) = \rho(B_1)E\left(\inf_{s \in S} \frac{V(s)}{w(s)} \middle| V \in B_1\right). \end{aligned}$$

For (2.17),

$$P(W > w) = P(YV > w) = P\left(Y > \sup_{s \in S} \frac{w(s)}{V(s)}\right) = E\left(\inf_{s \in S} \frac{V(s)}{w(s)}\right),$$

using Y standard Pareto and independent of V .

For (c) note that

$$\begin{aligned} P(W > w_0) &= P\left(Y \inf_{s \in S} V(s) > w_0\right) = E \min\left(1, \frac{\inf_{s \in S} V(s)}{\omega_0}\right) \\ &= \frac{1}{\omega_0} E \inf_{s \in S} V(s) > 0. \end{aligned}$$

Then (2.18) follows from (2.17).

For (d) note that, if $x > \omega_0$,

$$\begin{aligned}
 P(W(s) > x) &= P(YV(s) > x) = E \min\left(1, \frac{V(s)}{x}\right) \\
 &= x^{-1} E V(s) > 0.
 \end{aligned}
 \tag{2.20}$$

□

Relation (2.19) indicates that one-dimensional marginals, conditional on the process being larger than ω_0 , behave like Pareto; a similar observation has been done by Rootzén and Tajvidi [13] in the context of lower-dimensional distributions.

Let $s_1, s_2 \in S$ and $x > \omega_0$. From (2.20),

$$P(W(s_i) > x) = \frac{E(V(s_i))}{x} > 0, \quad i = 1, 2,$$

and, similarly

$$P(W(s_1) > x, W(s_2) > x) = \frac{E(V(s_1) \wedge V(s_2))}{x}.$$

Hence the statement $P(W(s_1) > c, W(s_2) > c) = P(W(s_1) > c)P(W(s_2) > c)$ for all $c > \omega_0$ is equivalent to the statement $E(V(s_1) \wedge V(s_2)) = c^{-1}E(V(s_1))E(V(s_2))$ for all $c > \omega_0$, which is impossible. That is, independence in the Pareto process between any two points is impossible.

For later use, we define next max-stable processes and give a well-known property.

Definition 2.2. A process $\eta = \{\eta(s)\}_{s \in \mathbb{R}} \in C(\mathbb{R})$ with non-degenerate marginals is called max-stable if, for η_1, η_2, \dots , i.i.d. copies of η , there are real continuous functions $c_n = \{c_n(s)\}_{s \in \mathbb{R}} > 0$ and $d_n = \{d_n(s)\}_{s \in \mathbb{R}}$ such that,

$$\max_{1 \leq i \leq n} \frac{\eta_i - d_n}{c_n} \stackrel{d}{=} \eta \quad \text{for all } n = 1, 2, \dots$$

The process is called simple if its marginal distributions are standard Fréchet, and then it will be denoted by $\bar{\eta}$.

Proposition 2.3 (Penrose [10], Theorem 5). All simple max-stable processes can be generated in the following way. Consider a Poisson point process on $(0, \infty]$ with mean measure $r^{-2} dr$. Let $\{Z_i\}_{i=1}^\infty$ be a realization of this point process. Further consider i.i.d. stochastic processes V_1, V_2, \dots in $C^+(\mathbb{R})$ with $E V_1(s) = 1$ for all $s \in \mathbb{R}$ and $E \sup_{s \in \mathbb{R}} V(s) < \infty$. Then

$$\bar{\eta} \stackrel{d}{=} \max_{i=1,2,\dots} Z_i V_i.
 \tag{2.21}$$

Conversely, each process with this representation is simple max-stable (and one can take V such that $\sup_{s \in \mathbb{R}} V(s) = c$ a.s. with $c > 0$).

Note that $\bar{\eta}$ depends on infinitely many processes V_i whereas W depends on just one of those processes (Theorem 2.1(3)).

2.2. The finite-dimensional setting

The theory of simple Pareto random *vectors* (r.v.) can be obtained from the results of Theorem 2.1, by taking a discrete set for S , $S = \{s_1, \dots, s_d\}$ say. Consequently, consider for this section the r.v. $(W_1, \dots, W_d) = (W(s_1), \dots, W(s_d))$.

Definition 2.3. *The r.v. $(W_1, \dots, W_d) \in \mathbb{R}_+^d$ with threshold parameter ω_0 is called simple Pareto. The probability measure*

$$\rho(B) = P\left(\frac{\omega_0(W_1, \dots, W_d)}{\max_{i=1, \dots, d} W_i} \in B\right) \tag{2.22}$$

for $B \in \mathcal{B}(\{(w_1, \dots, w_d) \in \mathbb{R}_+^d: \max(w_1, \dots, w_d) = \omega_0\})$ is again called the spectral measure.

It follows again that for having all marginals Pareto, one would need $\max(V(s_1), \dots, V(s_n)) = \omega_0$, for all $s_1, \dots, s_n \in S$ and all $n = 1, \dots, d$, which corresponds to $V \equiv \omega_0$ a.s., that is, the complete dependence case.

Nonetheless, we see that it is possible that some finite-dimensional marginals of a Pareto process have a Pareto distribution. For example, consider a situation where the maximum of the process occurs a.s. at some fixed locations in S , s_1, \dots, s_d say. Then $(W(s_1), \dots, W(s_d))$ is a d -dimensional simple Pareto r.v. with threshold parameter ω_0 . Moreover, any $(W(s'_1), \dots, W(s'_D))$ where $\{s_1, \dots, s_d\} \subset \{s'_1, \dots, s'_D\}$ is a D -dimensional simple Pareto r.v. with the same threshold parameter ω_0 .

One can give formulas for distribution functions, following similar reasoning as before. The statement corresponding to Proposition 2.1 is

$$P(W_1 \leq w_1, \dots, W_d \leq w_d) = \begin{cases} \rho(B_0) \left\{ 1 - E\left(\max_{i \in \bar{I}_0} \frac{V_i}{w_i} \mid B_0\right) \right\}, & \text{if } \rho(B_0) > 0, \\ 0, & \text{if } \rho(B_0) = 0, \end{cases} \tag{2.23}$$

where $I_0 = \{1 \leq i \leq d: w_i = 0\}$, $\bar{I}_0 = \{1 \leq i \leq d: w_i \neq 0\}$ and $B_0 = \{(V_1, \dots, V_d): V_i = 0 \text{ for } i \in I_0 \text{ and } \min_{i \in \bar{I}_0} \frac{w_i}{V_i} \geq 1\}$.

The statement corresponding to Proposition 2.2, with $w_i > 0$ for all $i = 1, \dots, d$, is

$$P(W_1 \leq w_1, \dots, W_d \leq w_d) = E\left(\max_{1 \leq i \leq d} \frac{V_i}{w_i \wedge \omega_0}\right) - E\left(\max_{1 \leq i \leq d} \frac{V_i}{w_i}\right), \tag{2.24}$$

which corresponds to formula (2) in Definition 2.1 from Rootzén and Tajvidi [13].

Note that Rootzén and Tajvidi’s formula, (2) in Definition 2.1, only holds for a vector (x, y) – for simplicity, we take bivariate vectors in the following discussion – larger than the vector of the lower endpoints of the marginal distributions. The following example illustrates this fact.

Apply Rootzén and Tajvidi’s formula with $G(x, y) = e^{-((x+1)^{-1} + (y+1)^{-1})}$, for $x > -1$ and $y > -1$, that is, the r.v. constructed from two independent unit Fréchet random variables shifted

by -1 . Then, their formula for the Pareto r.v. shifted by $(1, 1)$ is,

$$H(x, y) = \begin{cases} \frac{1}{2} \left(\frac{1}{x \wedge 0} - \frac{1}{y \wedge 0} - \frac{1}{x} - \frac{1}{y} \right) & (x > 0 \text{ and } y \geq 1) \text{ or } (y > 0 \text{ and } x \geq 1), \\ 0, & 0 < x \leq 1, 0 < y \leq 1. \end{cases} \quad (2.25)$$

Now note that this does not properly accommodate the positive mass that exists on the axis.

Our alternative approach leads, in this case, to the following distribution function. Consider the bivariate Pareto r.v. $(YB, Y(1 - B))$, with B Bernoulli $(1/2)$. Then, by direct calculations or by applying (2.23) one obtains the distribution function

$$P(YB \leq x, Y(1 - B) \leq y) = \begin{cases} \frac{1}{2} \left(2 - \frac{1}{x} - \frac{1}{y} \right), & \text{if } x \geq 1, y \geq 1, \\ \frac{1}{2} \left(1 - \frac{1}{x} \right), & \text{if } x \geq 1, 0 \leq y < 1, \\ \frac{1}{2} \left(1 - \frac{1}{y} \right), & \text{if } y \geq 1, 0 \leq x < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.26)$$

Regard that (2.25) and (2.26) are the same except when $x = 0$ or $y = 0$.

Another remark on Rootzén and Tajvidi [13]: their Theorem 2.2(ii) is not completely correct. It is not sufficient to require condition (6) of the same paper for $x, y > 0$. A counter example is given by

$$P(X > x \text{ or } Y > y) = \left(\frac{1}{2}e^{-2(x \vee 0)} + \frac{1}{2}e^{-2(y \vee 0)} \right)^{1/2}, \quad x \vee y \geq 0,$$

and zero elsewhere. This distribution satisfies (6) for $x, y > 0$ but not for all (x, y) and it is not a generalized Pareto distribution.

2.3. The generalized Pareto process

The more general processes with continuous extreme value index function $\gamma = \{\gamma(s)\}_{s \in S}$, location and scale functions $\mu = \{\mu(s)\}_{s \in S}$ and $\sigma = \{\sigma(s)\}_{s \in S}$ is defined as follows.

Definition 2.4. Let W be a simple Pareto process, $\mu, \sigma, \gamma \in C(S)$ with $\sigma > 0$. The generalized Pareto process $W_{\mu, \sigma, \gamma} \in C(S)$ is defined by,

$$W_{\mu, \sigma, \gamma} = \mu + \sigma \frac{W^\gamma - 1}{\gamma} \quad (2.27)$$

with all operations taken componentwise (recall the convention explained in the end of Section 1).

The result corresponding to Corollary 2.1 is the following.

Corollary 2.5. *The random variable $\sup_{s \in S} \{ (1 + \gamma(s) \frac{W_{\mu, \sigma, \gamma}(s) - \mu(s)}{\sigma(s)})^{1/\gamma(s)} \} \omega_0^{-1}$ has standard Pareto distribution.*

The process satisfies the following stability property.

Proposition 2.4. *For any generalized Pareto process $W_{\mu, \sigma, \gamma}$,*

$$P\left(\left(1 + \gamma \frac{W_{\mu, \sigma, \gamma} - \mu}{\sigma}\right)^{1/\gamma} \in rA\right) = r^{-1} P\left(\left(1 + \gamma \frac{W_{\mu, \sigma, \gamma} - \mu}{\sigma}\right)^{1/\gamma} \in A\right) \tag{2.28}$$

for all $r > 1$ and $A \in \mathcal{B}(C_{\omega_0}^+(S))$. Moreover, there exist normalizing functions $u(r)$ and $s(r)$ such that

$$\begin{aligned} P\left(\left(1 + \gamma \frac{W_{\mu, \sigma, \gamma} - u(r)}{s(r)}\right)^{1/\gamma} \in A \mid \sup_{s \in S} \left(1 + \gamma \frac{W_{\mu, \sigma, \gamma} - u(r)}{s(r)}\right)^{1/\gamma} > \omega_0\right) \\ = P\left(\left(1 + \gamma \frac{W_{\mu, \sigma, \gamma} - \mu}{\sigma}\right)^{1/\gamma} \in A\right) \end{aligned} \tag{2.29}$$

for all $r > 1$ and $A \in \mathcal{B}(C_{\omega_0}^+(S))$.

Conversely, if (2.29) holds and $\sup_{s \in S} \{ (1 + \gamma(s) \frac{W_{\mu, \sigma, \gamma}(s) - \mu(s)}{\sigma(s)})^{1/\gamma(s)} \} \omega_0^{-1}$ has a standard Pareto distribution, then (2.28) holds.

Proof. Relation (2.28) is direct from Definition 2.4 and (2.3). Then, with $u(r) = \mu + \sigma(r^\gamma - 1)/\gamma$ and $s(r) = \sigma r^\gamma$, relation (2.29) is easily shown to be true by (2.28) and Corollary 2.5.

Conversely, for all $r > 1$ and $A \in \mathcal{B}(C_{\omega_0}^+(S))$,

$$\begin{aligned} & \frac{P((1 + \gamma(W_{\mu, \sigma, \gamma} - u(r))/s(r))^{1/\gamma} \in A)}{P(\sup_{s \in S} (1 + \gamma(W_{\mu, \sigma, \gamma} - u(r))/s(r))^{1/\gamma} \omega_0^{-1} > 1)} \\ &= \frac{P((1 + \gamma(W_{\mu, \sigma, \gamma} - \mu)/\sigma)^{1/\gamma} \in rA)}{r^{-1}} \\ &= P\left(\left(1 + \gamma \frac{W_{\mu, \sigma, \gamma} - \mu}{\sigma}\right)^{1/\gamma} \in A\right) \end{aligned}$$

by (2.29) and $\sup_{s \in S} \{ (1 + \gamma(s) \frac{W_{\mu, \sigma, \gamma}(s) - \mu(s)}{\sigma(s)})^{1/\gamma(s)} \} \omega_0^{-1}$ being standard Pareto distributed. \square

The result corresponding to Proposition 2.2 on distribution functions is now, for $w > 0$:

$$\begin{aligned} P(W_{\mu, \sigma, \gamma} \leq w) &= E \left\{ \sup_{s \in S} V(s) \left(\left(1 + \gamma(s) \frac{w(s) - \mu(s)}{\sigma(s)} \right)^{1/\gamma(s)} \wedge \omega_0 \right)^{-1} \right\} \\ &\quad - E \left\{ \sup_{s \in S} V(s) \left(1 + \gamma(s) \frac{w(s) - \mu(s)}{\sigma(s)} \right)^{-1/\gamma(s)} \right\} \end{aligned}$$

for $1 + \gamma(w - \mu)/\sigma \in C^+(S)$.

3. Domain of attraction

Let us start with the characterization of the domain of attraction of a max-stable process. This result will lead directly to a characterization of the domain of attraction of a generalized Pareto process. The following is a slight variation and extension of Theorem 9.5.1 of de Haan and Ferreira [3].

Denote by $\bar{\eta} = \{\bar{\eta}(s)\}_{s \in S}$ any simple max-stable process in $C^+(S)$ (cf. Definition 2.2). Any max-stable process $\eta = \{\eta(s)\}_{s \in S}$ in $C(S)$ can be represented by $\eta = (\bar{\eta}^\gamma - 1)/\gamma$, for some $\bar{\eta}$ and continuous function $\gamma = \{\gamma(s)\}_{s \in S}$. For simplicity, we always take here

$$C_1^+(S) = \left\{ f \in C^+(S) : \sup_{s \in S} f(s) \geq 1 \right\}$$

that is, w.l.g. consider the constant ω_0 introduced in Section 2 equal to 1. For X a random element of $C(S)$, suppose the marginal distribution functions $F_s(x) = P(X(s) \leq x)$ are continuous in x , for all $s \in S$.

Theorem 3.1. *Suppose X, X_1, X_2, \dots are i.i.d. random elements of $C(S)$. The following statements are equivalent.*

1. *There exists a max-stable stochastic process $\eta \in C(S)$ with continuous index function γ , and $a_n > 0$ and b_n in $C(S)$ such that*

$$\left\{ \max_{1 \leq i \leq n} \frac{X_i(s) - b_n(s)}{a_n(s)} \right\}_{s \in S} \rightarrow^d \{\eta(s)\}_{s \in S} \tag{3.1}$$

in $C(S)$ (\rightarrow^d denotes weak convergence or convergence in distribution). The normalizing functions are w.l.g. chosen in such a way that $-\log P(\eta(s) \leq x) = (1 + \gamma(s)x)^{-1/\gamma(s)}$ for all x with $1 + \gamma(s)x > 0$, $s \in S$.

2. *There exist continuous functions $\gamma, a_t > 0$ and b_t such that*

$$\lim_{t \rightarrow \infty} tP\left(\frac{X(s) - b_t(s)}{a_t(s)} > x\right) = (1 + \gamma(s)x)^{-1/\gamma(s)}, \quad 1 + \gamma(s)x > 0, \tag{3.2}$$

uniformly for $s \in S$ and, for the normalized process

$$T_t X = \left(1 + \gamma \frac{X - b_t}{a_t}\right)_+^{1/\gamma}$$

we have

$$\lim_{t \rightarrow \infty} \frac{P(\sup_{s \in S} T_t X(s) > x)}{P(\sup_{s \in S} T_t X(s) > 1)} = \frac{1}{x} \quad \text{for all } x > 1 \tag{3.3}$$

and

$$\lim_{t \rightarrow \infty} P\left(\frac{T_t X}{\sup_{s \in S} T_t X(s)} \in B \mid \sup_{s \in S} T_t X(s) > 1\right) = \rho(B) \tag{3.4}$$

for each $B \in \mathcal{B}(\bar{C}_1^+(S))$ with $\rho(\partial B) = 0$, with ρ some probability measure on $\bar{C}_1^+(S)$.

The following shows that the same conditions are valid for the domain of attraction of a generalized Pareto process.

Theorem 3.2. 1. *The conditions of Theorem 3.1 imply*

$$\lim_{t \rightarrow \infty} P\left(T_t X \in A \mid \sup_{s \in S} T_t X(s) > 1\right) = P(W \in A)$$

with $A \in \mathcal{B}(C_1^+(S))$, $P(\partial A) = 0$ and W some simple Pareto process.

2. *Conversely suppose that there exists a function $\tilde{b}_u = \{\tilde{b}_u(s)\}_{s \in S}$, that is continuous in s for each u and increasing in u , and with the property that $P(X(s) > \tilde{b}_u(s) \text{ for some } s \in S) \rightarrow 0$ as $u \rightarrow \infty$, and a continuous function (in s), $\tilde{a}_u = \{\tilde{a}_u(s)\}_{s \in S} > 0$ such that, for some probability measure \tilde{P} on $\mathcal{B}(C(S))$,*

$$\lim_{u \rightarrow \infty} P\left(\frac{X - \tilde{b}_u}{\tilde{a}_u} \in A \mid X(s) - \tilde{b}_u(s) > 0 \text{ for some } s \in S\right) = \tilde{P}(A)$$

for all $A \in \mathcal{B}(C(S))$ and $\tilde{P}(\partial A) = 0$. Then the results of Theorem 3.1 hold.

Proof. Statement 1 follows directly from Theorem 3.1.

We prove statement 2:

By the conditions on \tilde{b}_u , we can determine $q = q(t)$ such that $P(X(s) > \tilde{b}_{q(t)}(s) \text{ for some } s \in S) = 1/t$. Then with $b_t(s) = \tilde{b}_{q(t)}(s)$ and $a_t(s) = \tilde{a}_{q(t)}(s)$,

$$\lim_{t \rightarrow \infty} t P\left(\frac{X - b_t}{a_t} \in C \text{ and } X(s) > b_t(s) \text{ for some } s \in S\right) = \tilde{P}(C)$$

for all $C \in \mathcal{B}(C(S))$ and $\tilde{P}(\partial C) = 0$. In particular, if $\inf\{\sup_{s \in S} f(s) : f \in C\} > 0$ we have

$$\lim_{t \rightarrow \infty} t P\left(\frac{X - b_t}{a_t} \in C\right) = \tilde{P}(C). \tag{3.5}$$

We proceed as usual in extreme value theory. Fix for the moment $s \in S$. It follows that for $x > 0$

$$\lim_{t \rightarrow \infty} t P(X(s) > b_t(s) + x a_t(s)) = \tilde{P}\{f : f(s) > x\}.$$

Let U_s be the inverse function of $1/P(X(s) > x)$ and $V(s)$ be the inverse function of $1/\tilde{P}\{f : f(s) > x\}$. Then

$$\lim_{t \rightarrow \infty} \frac{U_{tx}(s) - b_t(s)}{a_t(s)} = V_x(s) \quad \text{for } x > 0.$$

It follows (Lemma 10.4.2, p. 340, in de Haan and Ferreira [3]) that for some real $\gamma(s)$ and all $x > 0$

$$\lim_{t \rightarrow \infty} \frac{b_{tx}(s) - b_t(s)}{a_t(s)} = \frac{x^{\gamma(s)} - 1}{\gamma(s)} \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{a_{tx}(s)}{a_t(s)} = x^{\gamma(s)}. \quad (3.6)$$

Since the limit process has continuous paths, the function γ must be continuous on S .

Now replace t in (3.5) by ct where $c > 0$. Then

$$\lim_{t \rightarrow \infty} tP \left(\frac{b_t(s) - b_{tc}(s)}{a_{tc}(s)} + \frac{a_t(s)}{a_{tc}(s)} \frac{X - b_t}{a_t} \in C \right) = \frac{1}{c} \tilde{P}(C)$$

hence, by (3.6)

$$\lim_{t \rightarrow \infty} tP \left(\left(1 + \gamma \frac{X - b_t}{a_t} \right)^{1/\gamma} \in c(1 + \gamma C)^{1/\gamma} \right) = \frac{1}{c} \tilde{P}(C)$$

and by (3.5)

$$\lim_{t \rightarrow \infty} tP \left(\left(1 + \gamma \frac{X - b_t}{a_t} \right)^{1/\gamma} \in (1 + \gamma C)^{1/\gamma} \right) = \tilde{P}(C).$$

Write $P(A) = \tilde{P}((A^\gamma - 1)/\gamma)$. Then

$$\lim_{t \rightarrow \infty} tP(T_t X \in A) = P(A)$$

with $P(cA) = c^{-1}P(A)$, for all $c > 0$ and $A \in \mathcal{B}(C(S))$ such that $\inf\{\sup_{s \in S} f(s) : f \in A\} > 1$ and $P(\partial A) = 0$. The rest is like the proof of the equivalence between (2b) and (2c) of Theorem 9.5.1 in de Haan and Ferreira [3]. \square

Example 3.1. Any max-stable process is in the domain of attraction of a generalized Pareto process, with ρ given by the probability measure of V from (2.21).

Example 3.2. Any Pareto process with spectral measure ρ is in the domain of attraction of a max-stable process where the underlying process V (cf. representation (2.21)) has probability measure ρ .

Example 3.3. The finite-dimensional distributions of the moving maximum processes obtained in de Haan and Pereira [6] can be applied to obtain the finite-dimensional distributions of the corresponding Pareto process.

Example 3.4 (Regular variation (de Haan and Lin [4], Hult and Lindskog [9])). A stochastic process X in $C(S)$ is regularly varying if and only if there exists an $\alpha > 0$ and a probability measure ρ such that,

$$\frac{P(\sup_{s \in S} X(s) > tx, X / \sup_{s \in S} X(s) \in \cdot)}{P(\sup_{s \in S} X(s) > t)} \rightarrow^d x^{-\alpha} \rho(\cdot), \quad x > 0, t \rightarrow \infty, \quad (3.7)$$

on $\{f \in C(S) : \sup_{s \in S} f(s) = 1\}$. Hence, a regularly varying process such that (3.2) holds for the marginals, satisfies the conditions of Theorem 3.1, in particular with $\gamma = 1/\alpha$, $b_t = t$ and $a_t = t/\alpha$; note that the index function is constant.

On the other hand, the normalized process $T_t X$, with $T_t X$ satisfying (3.3)–(3.4), is regularly varying with $\alpha = 1$ and spectral measure ρ on $\bar{C}_1^+(S)$.

Remark 3.1. As seen in Section 2.2, our analysis is also valid in the finite-dimensional set-up. The main difference from Rootzén and Tajvidi [13] is that their analysis is entirely based on distribution functions whereas ours is more structural. Here are some remarks on their domain of attraction results.

Let $\bar{F} = 1 - F$ with F some d -variate distribution function, $\mathbf{x} = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$, and $\mathbf{u}(\cdot) = (u_1(\cdot), u_2(\cdot), \dots, u_d(\cdot))$ and σ the normalizing functions considered in Rootzén and Tajvidi [13] (see, e.g., their definition of $\mathbf{X}_\mathbf{u}$). By using $\sigma(\mathbf{x}t)/\sigma(t) \rightarrow (x_1^{\gamma_1}, x_2^{\gamma_2}, \dots, x_d^{\gamma_d})$ and $(\mathbf{u}(\mathbf{x}t) - \mathbf{u}(t))/\sigma(t) \rightarrow (\frac{x_1^{\gamma_1}-1}{\gamma_1}, \frac{x_2^{\gamma_2}-1}{\gamma_2}, \dots, \frac{x_d^{\gamma_d}-1}{\gamma_d})$, $t \rightarrow \infty$, for some reals $\gamma_1, \gamma_2, \dots, \gamma_d$ (cf. proof of Theorem 2.1(ii) in Rootzén and Tajvidi [13]) and by

$$\bar{F}^*(\mathbf{x}) := \bar{F}(u_1(x_1), u_2(x_2), \dots, u_d(x_d)),$$

one simplifies their relation (19) to

$$t \bar{F}^*(t\mathbf{x}) \rightarrow -\log G\left(\frac{x_1^{\gamma_1} - 1}{\gamma_1}, \frac{x_2^{\gamma_2} - 1}{\gamma_2}, \dots, \frac{x_d^{\gamma_d} - 1}{\gamma_d}\right),$$

and one simplifies their relation (6) to

$$P(\mathbf{X}^* \leq t\mathbf{x} | \mathbf{X}^* \not\leq t\mathbf{1}) = P(\mathbf{X}^* \leq \mathbf{x})$$

for $t \geq 1$. Hence, one can take $\mathbf{u}(t) := (\frac{t^{\gamma_1}-1}{\gamma_1}, \frac{t^{\gamma_2}-1}{\gamma_2}, \dots, \frac{t^{\gamma_d}-1}{\gamma_d})$ in Theorem 2.2 of that paper.

4. View towards application and simulation

4.1. Towards application

Suppose the domain of attraction condition (1.1) holds. Define $B = \{f \in C^+(S) : \sup_{s \in S} f(s) > 1\}$. Let A be a Borel set in $C^+(S)$. Then applying (1.1) twice we get

$$\lim_{t \rightarrow \infty} P(T_t X \in A | T_t X \in B) = \frac{\nu(A \cap B)}{\nu(B)} = P(W \in A)$$

with W a simple Pareto process. This is the content of Theorem 3.2(1) and it gives the basis to the peaks-over-threshold method in function space, as it gives a limit probability distribution on B .

A similar reasoning holds with B replaced by a different set B' as long as $\inf\{\sup_{s \in S} f(s) : f \in B'\} > 0$. Consider in particular $B' = \{f \in C^+(S) : \max_{i=1, \dots, p} T_i f(s_i) \geq 1\}$, for some integer p . Then

$$\lim_{t \rightarrow \infty} P(T_t X \in A | T_t X \in B') = \frac{\nu(A \cap B')}{\nu(B')} = \frac{P(W \in A \cap B')}{P(W \in B')}, \tag{4.1}$$

which is a generalized Pareto distribution.

Now, we proceed as in the peaks-over-threshold method for scalar observations: let $k = k(n)$ be a sequence of integers with $\lim_{n \rightarrow \infty} k(n) = \infty$ and $\lim_{n \rightarrow \infty} k(n)/n = 0$, as $n \rightarrow \infty$. Suppose that we have n independent observations of the process X in the domain of attraction. Select those observations satisfying $X(s_i) > b_{n/k}(s_i)$, for some $i = 1, 2, \dots, p$. The probability distribution of those selected observations is approximately the right-hand side of (4.1), that is, generalized Pareto. This seems a useful applicable form of the peaks-over-threshold method in this framework as it suggests estimating the spectral measure using observations that exceed a threshold at some discrete points in the space only.

4.2. Towards simulation

‘Deltares’ is an advisory organization of the Dutch government concerning (among others) the safety of the coastal defenses against severe wind storms. One studies the impact of severe storms on the coast, storms that are so severe that they have never been observed. In order to see how these storms look like, it is planned to simulate wind fields on and around the North Sea using certain climate models. These climate models simulate independent and identically distributed (i.i.d.) wind fields similar to the ones that could be observed (but that are only partially observed). Since the model runs during a limited time, some of the wind fields will be connected with storms of a certain severity but we do not expect to see really disastrous storms that could endanger the coastal defenses. The question put forward by Deltares is: can we get an idea how the really disastrous wind fields look like on the basis of the ‘observed’ wind fields? We want to show that this can be done using the generalized Pareto process.

Consider a continuous stochastic processes $\{X(s)\}_{s \in S}$ where S is a compact subset of \mathbb{R}^d . Suppose that the probability distribution of the process is in the domain of attraction of some max-stable process, that is, there exist functions $a_n > 0$ and b_n such that the sequence of i.i.d. processes

$$\left\{ \max_{1 \leq i \leq n} \frac{X_i(s) - b_n(s)}{a_n(s)} \right\}_{s \in S}$$

converges to a continuous process, say η , in distribution in $C(S)$. Then η is a max-stable process.

Define

$$T_t X(s) := \left(1 + \gamma(s) \frac{X(s) - b_t(s)}{a_t(s)} \right)_+^{1/\gamma(s)},$$

$$R_{T_t X} := \sup_{s \in S} T_t X(s).$$

Then

$$T_t^{\leftarrow} f(s) := a_t(s) \frac{(f(s))^{\gamma(s)} - 1}{\gamma(s)} + b_t(s) \quad \text{for } f \in C^+(S).$$

As before, suppress the s from now on. Then, with t_0 some large constant,

$$\begin{aligned} &P\left(\frac{T_t^{\leftarrow} t_0 T_t X - b_{t t_0}}{a_{t t_0}} \in A \mid R_{T_t X} > 1\right) \\ &= P\left\{\frac{a_t}{a_{t t_0}} \frac{[t_0(1 + \gamma(X - b_t)/a_t)^{1/\gamma}]^\gamma - 1}{\gamma} - \frac{b_{t t_0} - b_t}{a_{t t_0}} \in A \mid R_{T_t X} > 1\right\} \\ &= P\left\{\frac{a_t t_0^\gamma}{a_{t t_0}} \frac{1 + \gamma(X - b_t)/a_t - t_0^{-\gamma}}{\gamma} - \frac{b_{t t_0} - b_t}{a_{t t_0}} \in A \mid R_{T_t X} > 1\right\} \tag{4.2} \\ &= P\left\{\frac{a_t t_0^\gamma}{a_{t t_0}} \left(\frac{X - b_t}{a_t} - t_0^{-\gamma} \left[\frac{b_{t t_0} - b_t}{a_t} - \frac{t_0^\gamma - 1}{\gamma}\right]\right) \in A \mid R_{T_t X} > 1\right\} \\ &= P\left\{\frac{X - b_t}{a_t} \in \frac{a_{t t_0} t_0^{-\gamma}}{a_t} A + t_0^{-\gamma} \left(\frac{b_{t t_0} - b_t}{a_t} - \frac{t_0^\gamma - 1}{\gamma}\right) \mid R_{T_t X} > 1\right\}. \end{aligned}$$

Since,

$$\frac{a_{t t_0}(s) t_0^{-\gamma(s)}}{a_t(s)} \rightarrow 1 \quad \text{and} \quad \frac{b_{t t_0}(s) - b_t(s)}{a_t(s)} - \frac{t_0^{\gamma(s)} - 1}{\gamma(s)} \rightarrow 0$$

uniformly for $s \in S$, the limit of this probability, as $t \rightarrow \infty$, is the same as the limit of

$$P\left(\frac{X - b_t}{a_t} \in A \mid R_{T_t X} > 1\right), \tag{4.3}$$

which is generalized Pareto by Theorem 3.2(1).

In this subsection, we are not so much interested in estimating the joint limit distribution (which is the peaks-over-threshold method) but in the fact that the two conditional distributions (4.2) and (4.3) are approximately the same.

Suppose, for example, that we have observed wind fields over a certain area during some time. Then we are likely to find some rather heavy storms, that is, ones that satisfy $X \not\leq b_n$. These are the moderately heavy storms. However, we want to know how the storm field of a really heavy storm (i.e., $X \not\leq b_N$ with $N > n$) looks like. That is exactly what relation (4.2) does. Take a moderately heavy storm X and transform it to $T_n^{\leftarrow} \frac{N}{n} T_n X$. This results in the storm field of a really heavy storm by relation (4.2).

Notice then what we do here is similar to prediction or kriging, not estimating a distribution function.

The reasoning above also holds with estimated functions of γ , a and b , on the basis of k th upper order statistics and taking $t = n/k$.

Under the above framework, we propose the following simulation method:

- (1) Let X_1, X_2, \dots, X_n be i.i.d. and let the underlying distribution satisfy the conditions above, namely that the probability distribution is in the domain of attraction of some max-stable process.
- (2) Estimate the functions γ , a and b (de Haan and Lin [5], Einmahl and Lin [7]); denote the estimators by $\hat{\gamma}$, \hat{a} and \hat{b} . Note that this procedure provides us with a number k that reflects the threshold for estimating the parameters.
- (3) Select from the normalized processes

$$\hat{T}_{n/k} X_i := \left(1 + \hat{\gamma} \frac{X_i - \hat{b}_{n/k}}{\hat{a}_{n/k}} \right)_+^{1/\hat{\gamma}}, \quad i = 1, \dots, n,$$

those that satisfy $X_i(s) > \hat{b}_{n/k}(s)$ for some $s \in S$, that is, for which $\hat{R}_{T_{n/k} X_i} > 1$.

- (4) Multiply these processes by a (large) factor t_0 ; this brings the processes to a higher level without changing the distribution essentially.
- (5) Finally undo the normalization, that is, in the end we obtain the processes

$$\hat{T}_{n/k}^{\leftarrow} t_0 \hat{T}_{n/k} X_i \text{ for those } X_i \text{ for which } \hat{R}_{T_{n/k} X_i} > 1.$$

These processes are peaks-over-threshold processes with respect to a much higher threshold (namely b_{t_0}) than the processes X_i for which $\hat{R}_{T_{n/k} X_i} > 1$ (with threshold b_t).

Remark 4.1. Note that an alternative procedure under the maximum domain of attraction condition would be, first to estimate the spectral measure and then to simulate a generalized Pareto process from there. But the estimation of the spectral measure is more difficult (de Haan and Lin [4]) although this procedure is less restrictive on the number of observations that can be simulated.

4.3. Simulations

We exemplify the lifting procedure with the process $X(s) = Z(s)^{\gamma(s)}$, with $\gamma(s) = 1 - s(1 - s)^2$, $s \in [0, 1]$, and Z is the moving maximum process with standard Gaussian density. The Z process can be easily simulated in the R-package due to Ribatet [12]. Figure 1 is represented by the 11 out of 20 realizations, normalized for which $\hat{R}_{T_{n/k} X_i} > 1$, and lifted ones $\hat{T}_{n/k}^{\leftarrow} t_0 \hat{T}_{n/k} X_i$ with $t_0 = 10$.

Acknowledgements

An anonymous referee pointed out an error in the original version of Proposition 2.2; we are grateful for this.

The presented work in Section 4.2 is part of the SBW (Strength and Load on Water Defences) project, commissioned by Rijkswaterstaat Centre for Water Management in the Netherlands and carried out by Deltares, The Netherlands.

Research partially supported by FCT Project PTDC/MAT/112770/2009 and FCT-PEst-OE/MAT/UI0006/2011.

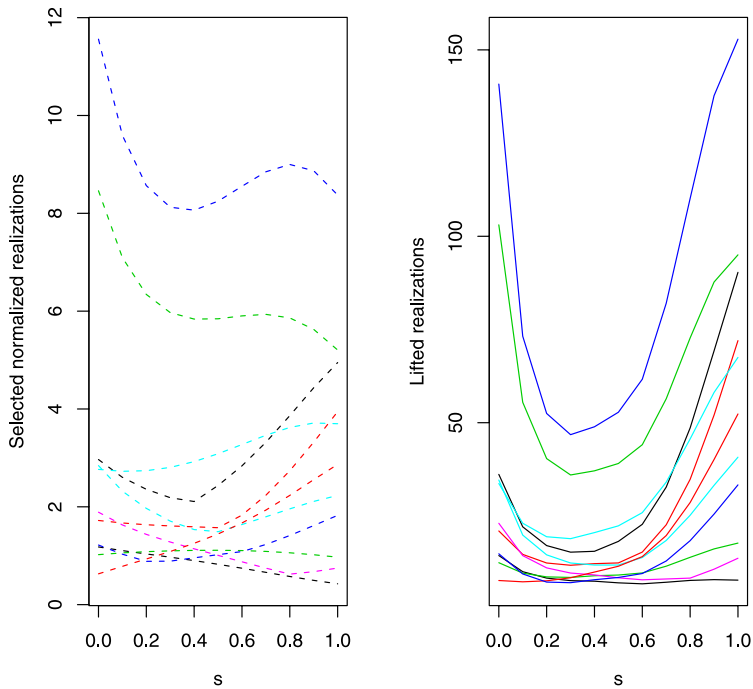


Figure 1. (Left): the realizations $\hat{T}_{n/k}X_i$, for which $\hat{R}_{T_{n/k}X_i} > 1$, obtained from the moving maximum process with standard Gaussian density; (right) the lifted realizations $\hat{T}_{n/k}^{\leftarrow t_0} \hat{T}_{n/k}X_i$ with $t_0 = 10$.

References

- [1] Balkema, A.A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probab.* **2** 792–804. [MR0359049](#)
- [2] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. [MR1324786](#)
- [3] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. [MR2234156](#)
- [4] de Haan, L. and Lin, T. (2001). On convergence toward an extreme value distribution in $C[0, 1]$. *Ann. Probab.* **29** 467–483. [MR1825160](#)
- [5] de Haan, L. and Lin, T. (2003). Weak consistency of extreme value estimators in $C[0, 1]$. *Ann. Statist.* **31** 1996–2012.
- [6] de Haan, L. and Pereira, T.T. (2006). Spatial extremes: The stationary case. *Ann. Statist.* **34** 146–168.
- [7] Einmahl, J.H.J. and Lin, T. (2006). Asymptotic normality of extreme value estimators on $C[0, 1]$. *Ann. Statist.* **34** 469–492. [MR2275250](#)
- [8] Falk, M., Hüsler, J. and Reiss, R.D. (2010). *Laws of Small Numbers: Extremes and Rare Events*. Basel: Birkhäuser.
- [9] Hult, H. and Lindskog, F. (2005). Extremal behavior of regularly varying stochastic processes. *Stochastic Process. Appl.* **115** 249–274. [MR2111194](#)

- [10] Penrose, M.D. (1992). Semi-min-stable processes. *Ann. Probab.* **20** 1450–1463. [MR1175271](#)
- [11] Pickands, J. III (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119–131. [MR0423667](#)
- [12] Ribatet, M. (2011). SpatialExtremes: Modelling spatial extremes. R package.
- [13] Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli* **12** 917–930. [MR2265668](#)

Received November 2012 and revised April 2013