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

Ana Ferreira
ISA, Univ Tecn Lisboa,
Tapada da Ajuda 1349-017 Lisboa, Portugal
and CEAUL, FCUL,
Bloco C6 - Piso 4 Campo Grande, 749-016 Lisboa, Portugal

Laurens de Haan
Erasmus University Rotterdam,
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
and CEAUL, FCUL,
Bloco C6 - Piso 4 Campo Grande, 749-016 Lisboa, Portugal

May 5, 2014

Abstract
In extreme value statistics the peaks-over-threshold method is widely used. The method is based on the Generalized Pareto distribution ([1], [13] in univariate theory and e.g. [4], [15] in multivariate theory) characterizing probabilities of exceedances over high thresholds. 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 any specification with the 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.

Keywords: domain of attraction, extreme value theory, generalized Pareto process, max-stable processes, functional regular variation

1 Introduction
We say that a stochastic process \( X \) in \( C(S) \) (the space of continuous real functions on \( S \), with \( S \) a compact subset of \( \mathbb{R}^d \)) is in the domain of attraction of a
max-stable process if, there are continuous functions \(a_n(s)\) positive and \(b_n(s)\) on \(S\) such that the processes

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

with \(X, X_1, \ldots, X_n\) independent and identically distributed, converge in distribution to a max-stable process \(Y\) in \(C(S)\). Necessary and sufficient conditions for this to happen are: uniform convergence of the marginal distributions and a convergence of measures (in fact a form of regular variation):

\[
\lim_{t \to \infty} tP(T_t X \in A) = \nu(A) \tag{1.1}
\]

where \(T_t X(s) := \left(1 + \gamma(s) \frac{X(s) - b_t(s)}{a_t(s)}\right)^{1/\gamma(s)}\) for all \(s \in S\) (with the notation \(x_+ = \max(0, x)\) for any real \(x\)); \(\nu\) is a homogeneous (of order -1) measure on \(C^+(S) := \{f \in C(S) : f \geq 0\}\) and \(A\) any Borel subset of \(C^+(S)\) verifying \(\nu(dA) = 0\) and \(\inf\{\sup_{s \in S} f(s) : f \in A\} > 0\) (de Haan and Lin (2001), cf. de Haan and Ferreira (2006) Section 9.5). The functions \(a_t(s)\) and \(b_t(s)\) are chosen in such a way that the marginal distributions are in standard form:

\[
\exp\left\{-(1 + \gamma(s)x)^{-1/\gamma(s)}\right\}, \quad 1 + \gamma(s)x > 0,
\]

for all \(x \in \mathbb{R}\) and \(s \in S\). Here \(\gamma\) 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 \(a_t(s) := \gamma(s)(b_{2t}(s) - b_t(s))/\left(2^{\gamma(s)} - 1\right)\).

From (1.1) it follows that

\[
P\left(\frac{1 + \gamma(\cdot)\frac{X(\cdot) - b_t(\cdot)}{a_t(\cdot)}}{a_t(\cdot)}^{1/\gamma(\cdot)} \in A\right) = P\left(\sup_{s \in S} \frac{X(s) - b_t(s)}{a_t(s)} > 0\right) \tag{\[1.2\]}
\]

converges as \(t \to \infty\) and so does

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

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

This reasoning is quite similar to how one gets the generalized Pareto distributions in \(\mathbb{R}\) (Balkema and de Haan, 1974) and in \(\mathbb{R}^d\) (Rootzén and Tajvidi, 2006; Falk, Hüsler and Reiss, 2010). 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 2.2. In the meantime, in Section 2.1.1.
is the discrete version of our approach to simple Pareto processes, leading to multivariate simple Pareto random vectors. 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 form independent and identically observations of storm fields.

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 e.g. we shall simply write \(\left(1 + \frac{X - b_t}{a_t}\right)^{1/\gamma}\) for \(\left\{\left(1 + \frac{X(s) - b_t(s)}{a_t(s)}\right)^{1/\gamma(s)}\right\}_{s \in S}\), with $X = \{X(s)\}_{s \in S}, a_t = \{a_t(s)\}_{s \in S}, b_t = \{b_t(s)\}_{s \in S}$ and $\gamma = \{\gamma(s)\}_{s \in S}$. In general, $C(S)$ is the space of continuous real functions on $S \subset \mathbb{R}^d$ equipped with the supremum norm.

2 Pareto processes

2.1 The simple Pareto process

Let $C^+(S)$ be the space of non-negative real continuous functions on $S$, with $S$ some compact subset of $\mathbb{R}^d$. We denote the Borel subsets of a metric space by $B(\cdot)$.

**Theorem 2.1.** Let $W$ be a stochastic process in $C^+(S)$ and $\omega_0$ a positive constant. The following three statements are equivalent:

1. (POT - peaks-over-threshold stability)
   (a) $E(W(s)/\sup_{s \in S} W(s)) > 0$ 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)
   \[
   P\left(\frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B \mid \sup_{s \in S} W(s) > r\right) = P\left(\frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B\right),
   \]
   for all $r > \omega_0$ and $B \in B(\tilde{C}^+_{\omega_0}(S))$ with
   \[
   \tilde{C}^+_{\omega_0}(S) := \{f \in C^+(S) : \sup_{s \in S} f(s) = \omega_0\}. \tag{2.1}
   \]

2. (Random functions)
   (a) $P(\sup_{s \in S} W(s) > \omega_0) = 1$,
   (b) $E(W(s)/\sup_{s \in S} W(s)) > 0$ for all $s \in S$,
P(W \in rA) = r^{-1} P(W \in A),
\quad (2.3)
for all \( r > 1 \) and \( A \in \mathcal{B}(C^+_\omega(S)) \),
where \( rA \) means the set \( \{rf, f \in A\} \), and
\[
C^+_\omega(S) := \{f \in C^+(S) : \sup_{s \in S} f(s) \geq \omega_0\}. \quad (2.4)
\]

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

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

Definition 2.1. The process \( W \) characterized in Theorem 2.1, with threshold parameter \( \omega_0 \), is called simple Pareto process. The probability measure in \( (2.1) \) i.e.,
\[
\rho(B) = P\left( \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B \right), \quad \text{for } B \in \mathcal{B}(\bar{C}^+\omega_0(S)), \quad (2.5)
\]
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 \( s_1, s_2 \in S \), i.e. \( W(s_1) \) 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
\[
A_{r,B} = \left\{f \in C^+(S) : \sup_{s \in S} f(s)/\omega_0 > r, \frac{\omega_0 f}{\sup_{s \in S} f(s)} \in B\right\} = r \times A_{1,B},
\]
for all \( r > 1 \) and \( B \in \mathcal{B}(\bar{C}^+_\omega_0(S)) \). Then,
\begin{align*}
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{align*}
using in particular the independence of Y and V and $P \left( \sup_{s \in S} W(s)/\omega_0 > 1 \right) = 1$. Since $P(rA) = r^{-1}P(A)$ holds for any of the above sets, it holds for all Borel sets in the statement.

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

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

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

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

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.

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) $E (W(s)/\sup_{s \in S} W(s)) > 0$ for all $s \in S$,
   (b) $P \left( \sup_{s \in S} W(s)/\omega_0 > x \right) = x^{-1}$, for $x > 1$,
   (c)
   $$P \left( W \in rA \mid \sup_{s \in S} W(s) > r\omega_0 \right) = P \left( W \in A \right)$$
   \hspace{1cm} (2.6)
   for all $r > 1$ and $A \in \mathcal{B} \left( C^+_{\omega_0} (S) \right)$.

2. (a) $E (W(s)/\sup_{s \in S} W(s)) > 0$ for all $s \in S$,
   (b)
   $$P \left( \sup_{s \in S} W(s)/\omega_0 > r, \frac{\omega_0 W}{\sup_{s \in S} W(s)} \in B \right) = \frac{\rho(B)}{r}$$
   \hspace{1cm} (2.7)
   for all $r > 1$ and $B \in \mathcal{B} \left( \bar{C}^+_{\omega_0} (S) \right)$.

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 (2006), 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. and Definition 2.1.
Let \( w, W \in C^+(S) \). Conditions \( W \leq (>)w \) define the sets \( \{ f \in C^+(S) : f(s) \leq (>)w(s) \text{ for all } s \in S \} \) and \( (W \not\leq w) \) defines the set \( \{ f \in C^+(S) : f(s) > w(s) \text{ for at least one } s \in S \} \). Hence note that \( W > w \) is not the complement of \( W \leq w \).

Take for the conditional expectation,

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

defined in the usual sense and whenever \( \rho(B) = P(V \in B) > 0 \), with \( g \) a real functional (e.g. see Billingsley, 1995).

**Proposition 2.1.** Let \( w, W \in C^+(S) \), with \( W \) 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 \( B_0 = \{ f \in \bar{C}^+_w(S) : \inf_{s \in S_0} \frac{w(s)}{f(s)} \geq 1 \text{ and } f(s) = 0 \text{ for } s \in S_0 \} \). Then

\[
P(W \leq w) = \begin{cases} 
\rho(B_0) \left\{ 1 - E\left( \sup_{s \in \bar{S}_0} \frac{V(s)}{w(s)} \Big| V \in B_0 \right) \right\} & \text{if } \rho(B_0) > 0 \\
0 & \text{if } \rho(B_0) = 0.
\end{cases}
\]

(2.8)

**Proof.**

\[
P(W \leq w) = P(W(s) \leq w(s) \text{ for } s \in \bar{S}_0 \text{ and } W(s) \leq w(s) \text{ for } s \in S_0)
= P\left( Y \leq \inf_{s \in S_0} \frac{w(s)}{V(s)} \text{ and } V(s) = 0 \text{ for } s \in S_0 \right)
= P\left( Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \text{ and } V(s) = 0 \text{ for } s \in S_0 \text{ and } \inf_{s \in S_0} \frac{w(s)}{V(s)} \geq 1 \right)
+ P\left( Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \text{ and } V(s) = 0 \text{ for } s \in S_0 \text{ and } \inf_{s \in S_0} \frac{w(s)}{V(s)} < 1 \right)
= \int_{B_0} P\left( Y \leq \inf_{s \in \bar{S}_0} \frac{w(s)}{V(s)} \right) d\rho(v)
= \int_{B_0} 1 - \sup_{s \in \bar{S}_0} \frac{v(s)}{w(s)} d\rho(v) = \rho(B_0) - \int_{B_0} \sup_{s \in \bar{S}_0} \frac{v(s)}{w(s)} d\rho(v)
= \rho(B_0) \left\{ 1 - E\left( \sup_{s \in \bar{S}_0} \frac{V(s)}{w(s)} \Big| V \in B_0 \right) \right\}
\]

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

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

(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 $W$ simple Pareto process. Then,

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

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

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

hence,

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

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

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

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:

$$\nu(rB) = r^{-1} \nu(B) \quad \text{for all } r > 0 \text{ and } B \in B \left( C^+ \omega_0(S) \right).$$

Then, the probability distribution of $W$ is the restriction of $\nu$ to $C^+ \omega_0(S)$ i.e., for $B \in B \left( C^+ \omega_0(S) \right)$,

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

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

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

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

$$P(W \not\leq w) = \nu \left\{ f \not\leq w, f \not\leq \omega_0 \right\} = \nu \left\{ f \not\leq w \right\} + \nu \left\{ f \not\leq \omega_0 \right\} - \nu \left\{ f \not\leq w \text{ or } f \not\leq \omega_0 \right\} = R \left( \sup_{s \in S} \frac{V(s)}{w(s)} \right) + 1 - E \left( \sup_{s \in S} \frac{V(s)}{w(s) \land \omega_0} \right). \quad \square$$
Note that \( E \left( \sup_{s \in S} \frac{V(s)}{w(s)} \mid V \in B_0 \right) = 1 \), 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 \( W \) simple Pareto process. Then:

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

\[
P(W > w) = \begin{cases} 
\rho(B_1) E \left( \inf_{s \in S} \frac{V(s)}{w(s)} \mid V \in B_1 \right) & \text{if } \rho(B_1) > 0 \\
0 & \text{if } \rho(B_1) = 0.
\end{cases}
\]

(2.15)

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

\[
P(W > w) = P(V > 0) E \left( \inf_{s \in S} \frac{V(s)}{w(s)} \mid V > 0 \right).
\]

(2.16)

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

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

(2.17)

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

\[
P(W > x \mid W > \omega_0) = \begin{cases} 
1, & x \leq \omega_0 \\
\omega_0/x, & x > \omega_0
\end{cases}.
\]

(2.18)

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

\[
P(W(s) > x \mid W(s) > \omega_0) = \begin{cases} 
1, & x \leq \omega_0 \\
\omega_0/x, & x > \omega_0
\end{cases}.
\]

(2.19)

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

\[
P(W > w) = P \left( Y \geq \sup_{s \in S} \frac{w(s)}{V(s)} \text{ and } \inf_{s \in S} V(s) > 0 \right)
\]

\[= \int_{B_1} \inf_{s \in S} \frac{v(s)}{w(s)} d\rho(v) = \rho(B_1) E \left( \inf_{s \in S} \frac{V(s)}{w(s)} \mid V \in B_1 \right).
\]

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

\[
P \left( W > w_0 \right) = P \left( Y \inf_{s \in S} V(s) > \omega_0 \right) = E \min \left( 1, \frac{\inf_{s \in S} V(s)}{\omega_0} \right)
\]

\[= \frac{1}{\omega_0} E \inf_{s \in S} V(s) > 0.
\]
Then (2.18) follows from (2.17).

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

\[
P(W(s) > x) = P(YV(s) > x) = E \min \left( 1, \frac{V(s)}{x} \right) = x^{-1}EV(s) > 0. \tag{2.20}
\]

Relation (2.19) indicates that one-dimensional marginals, conditional on the process being larger than \( \omega_0 \), behave like Pareto; similar observation is in Rootzén and Tajvidi (2006) 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 \( P(W(s_1) > c, W(s_2) > c) = P(W(s_1) > c)P(W(s_2) > c) \) for all \( c > \omega_0 \) is equivalent to verify \( E(V(s_1) \wedge V(s_2)) = c^{-1}E(V(s_1))E(V(s_2)) \) 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 well known properties useful for our purposes.

**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 \( \eta_1, \eta_2, \ldots \), independent and identically distributed (i.i.d.) copies of \( \eta \), 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,

\[
\max_{1 \leq i \leq n} \frac{\eta_i - d_n}{c_n} \overset{d}{=} \eta \quad \text{for all } n = 1, 2, \ldots.
\]

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

**Proposition 2.3** (Penrose(1992), 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 \( V_1, V_2, \ldots \) in \( C^+(\mathbb{R}) \) with \( EV_1(s) = 1 \) for all \( s \in \mathbb{R} \) and \( E \sup_{s \in \mathbb{R}} V(s) < \infty \). Then

\[
\bar{\eta} = \max_{i=1,2,\ldots} Z_i V_i. \tag{2.21}
\]

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 \( c > 0 \)).
The finite dimensional distributions of \( \bar{\eta} \) where computed in de Haan (1984).

With the above representation for simple max-stable processes, for \( s_1, \ldots, s_n \in \mathbb{R}, x_1, \ldots, x_n > 0, \) for all \( n \in \mathbb{N}, \)

\[
G(x_1, \ldots, x_n) = P(\bar{\eta}(s_1) \leq x_1, \ldots, \bar{\eta}(s_n) \leq x_n) = \exp\left(-E \max_{1 \leq i \leq n} \frac{V(s_i)}{x_i}\right).
\]

(2.22)

Note that \( \bar{\eta} \) is the maximum of infinitely many processes whereas \( W \) depends on just one of those processes (Theorem 2.1.3.).

2.1.1 The finite-dimensional setting

Simple Pareto random vectors can be constructed similarly as in Theorem 2.1 – Definition 2.1.

**Theorem 2.2.** Let \( (W_1, \ldots, W_d) \) be a random vector in \( \mathbb{R}_+^d = [0, \infty)^d \) and \( \omega_0 \) a positive constant. The following three statements are equivalent:

1. (a) \( E (W_i / \max_{1 \leq i \leq d} W_i) > 0 \) for all \( i = 1, \ldots, d, \)
   (b) \( P (\max_{1 \leq i \leq d} W_i / \omega_0 > x) = x^{-1}, \) for \( x > 1 \) (standard Pareto distribution),
   (c) \[
P \left( \begin{array}{c}
\omega_0 (W_1, \ldots, W_d) \\
\max_{1 \leq i \leq d} W_i
\end{array} \right) \in B \left| \begin{array}{c}
\max_{1 \leq i \leq d} W_i > r\omega_0
\end{array} \right) = P \left( \begin{array}{c}
\omega_0 (W_1, \ldots, W_d) \\
\max_{1 \leq i \leq d} W_i
\end{array} \right) \in B,
\] \( \quad \) (2.23)

for all \( r > 1 \) and \( B \in \mathcal{B} (\hat{D}_{\omega_0}^+) \) with
\[
\hat{D}_{\omega_0}^+ := \{(w_1, \ldots, w_d) \in \mathbb{R}_+^d : \max_{1 \leq i \leq d} w_i = \omega_0\}. \quad (2.24)
\]

2. (a) \( P (\max_{1 \leq i \leq d} W_i > \omega_0) = 1, \)
   (b) \( E (W_i / \max_{1 \leq i \leq d} W_i) > 0 \) for all \( i = 1, \ldots, d, \)
   (c) \[
P ((W_1, \ldots, W_d) \in r A) = r^{-1} P ((W_1, \ldots, W_d) \in A),
\] \( \quad \) (2.25)

for all \( r > 1 \) and \( A \in \mathcal{B} (D_{\omega_0}^+) \) with
\[
D_{\omega_0}^+ := \{(w_1, \ldots, w_d) \in \mathbb{R}_+^d : \max_{1 \leq i \leq d} w_i \geq \omega_0\}. \quad (2.26)
\]

3. \( (W_1, \ldots, W_d) = Y (V_1, \ldots, V_d), \) for some \( Y \) and random vector \( (V_1, \ldots, V_d) \in \mathbb{R}_+^d \) verifying:
   (a) \( \max_{1 \leq i \leq d} V_i = \omega_0 \) a.s., and \( EV_i > 0 \) for all \( i = 1, \ldots, d, \)
   (b) \( Y \) is a standard Pareto random variable, \( F_Y (y) = 1 - 1/y, \) \( y > 1, \)
   (c) \( Y \) and \( V \) are independent.
Proof. Similar to the proof of Theorem 2.1.

Definition 2.3. The random vector \((W_1, \ldots, W_d) \in \mathbb{R}^d_+\) characterized in Theorem 2.1 with threshold parameter \(\omega_0\), is simple Pareto. The probability measure in (2.23), i.e.,

\[
\rho(B) = P\left(\frac{\omega_0 (W_1, \ldots, W_d)}{\max_{i=1, \ldots, d} W_i} \in B\right), \quad \text{for } B \in \mathcal{B}(\bar{D}_{\omega_0}^+),
\]

is called the spectral measure.

From the finite dimensional representation it follows again that for having all marginals Pareto, one would have to have \(\max(V(s_1), \ldots, V(s_n)) = \omega_0\), for all \(s_1, \ldots, s_n \in S\) and all integer \(n\), which corresponds to \(V \equiv \omega_0\) a.s., i.e. the complete dependence case.

Nonetheless, we see that it is still possible that some finite dimensional marginals of a Pareto process are still Pareto. For example, consider a situation where the maximum of the process may only occur at some fixed locations in \(S\) and the related finite dimensional random vector includes these locations.

One can give formulas for distribution functions, following similar reasoning as before. The correspondent to Proposition 2.1 is given by,

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

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

The correspondent to Proposition 2.2 with \(w_i > 0\) for all \(i = 1, \ldots, d\), is given by,

\[
P(W_1 \leq w_1, \ldots, W_d \leq w_d) = E \left(\max_{1 \leq i \leq d} \frac{V_i}{w_i} \wedge \omega_0\right) - E \left(\max_{1 \leq i \leq d} \frac{V_i}{w_i}\right),
\]

which corresponds to the one given in Rootzén and Tajvidi, in the correspondent region.

Note that Rootzén and Tajvidi’s formula (i.e. (2) in Definition 2.1) only holds for the vector \(x\) larger than the vector of the lower endpoints of the marginal distributions. The following example illustrates this. Consider the model related to two independent unit Fréchet. Then Rootzén and Tajvidi’s formula corresponds to

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

but, it is undetermined otherwise. Then note that their formula is not able to capture the positive mass on the axis.
By direct calculations, or alternatively applying (2.28) with the bivariate Pareto random vector \((YB, Y(1 - B))\), with \(B\) Bernoulli \((1/2)\), the correct distribution function is,

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

(2.30)

Another remark on Rootzéns and Tajvidi (2006): their Theorem 2.2(ii) is not completely correct. It is not sufficient to require condition (6) of the same paper for \(x > 0\). 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 \(x > 0\) but not for all \(x\) and it is not a generalized Pareto distribution.

### 2.2 The generalized Pareto process

Let \(C(S)\) be the space of real continuous functions on \(S\) with \(S \subset \mathbb{R}^d\) compact. 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:

**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,

\[
W_{\mu, \sigma, \gamma} = \mu + \sigma \frac{W^\gamma - 1}{\gamma} \tag{2.31}
\]

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

The result corresponding to Corollary 2.1 is,

**Corollary 2.5.** The random variable \(\sup_{s \in S} \left( 1 + \gamma(s) \frac{W_{\mu, \sigma, \gamma}(s) - \mu(s)}{\sigma(s)} \right)^{1/\gamma(s)} \omega_{0}^{-1}\)

has standard Pareto distribution.

Related with stability or homogeneity properties we have:

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

\[
P \left( \left( 1 + \gamma \frac{W_{\mu, \sigma, \gamma} - \mu}{\sigma} \right)^{1/\gamma} \in rA \right) = r^{-1} P \left( \left( 1 + \gamma \frac{W_{\mu, \sigma, \gamma} - \mu}{\sigma} \right)^{1/\gamma} \in A \right), \tag{2.32}
\]
for all $r > 1$ and $A \in \mathcal{B}(C^+_\infty(S))$. Moreover, there exist normalizing functions $u(r)$ and $s(r)$ such that

\[
P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} \right| \sup_{s \in S} \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} > \omega_0 \right) = P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \in A \right| A \right),
\]

(2.33)

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

Conversely, (2.33) and $\sup_{s \in S} \left\{ \left( 1 + \gamma(s) \frac{W_{\mu,\sigma,\gamma}(s) - \mu(s)}{\sigma(s)} \right)^{1/\gamma(s)} \right\} \omega_0^{-1}$ being standard Pareto distributed, imply (2.32).

**Proof.** Relation (2.32) is direct from Definition 2.4 and (2.3). Then, with $u(r) = \mu + \sigma (r^{\gamma - 1})/\gamma$ and $s(r) = \sigma r$, and, for all $r > 1$ and $A \in \mathcal{B}(C^+_\infty(S))$,

\[
P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} \in A \right| \sup_{s \in S} \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} \not\leq \omega_0 \right)
= P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \in rA \right| \sup_{s \in S} \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} \not\leq \omega_0 \right)
= \frac{P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \in rA \right| rA \right) \omega_0^{-1}}{P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \not\leq \omega_0 \right) \right)} = P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \in A \right| A \right)
\]

by (2.32) and Corollary 2.5.

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

\[
P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} \in A \right| \sup_{s \in S} \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - u(r)}{s(r)} \right)^{1/\gamma} \omega_0^{-1} > 1 \right)
= \frac{P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \in rA \right| rA \right) \omega_0^{-1}}{P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \not\leq \omega_0 \right) \right)} = P \left( \left. \left( 1 + \gamma \frac{W_{\mu,\sigma,\gamma} - \mu}{\sigma} \right)^{1/\gamma} \in A \right| A \right)
\]

by (2.33) and $\sup_{s \in S} \left\{ \left( 1 + \gamma(s) \frac{W_{\mu,\sigma,\gamma}(s) - \mu(s)}{\sigma(s)} \right)^{1/\gamma(s)} \right\} \omega_0^{-1}$ being standard Pareto distributed.

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

\[
P(W_{\mu, \sigma, \gamma} \leq w) = E \left\{ \sup_{s \in S} V(s) \left( \left( 1 + \gamma(s) \frac{w(s) - \mu(s)}{\sigma(s)} \right)^{\frac{1}{\gamma(s)}} \wedge \omega_0 \right)^{-1} \right\}
- E \left\{ \sup_{s \in S} V(s) \left( 1 + \gamma(s) \frac{w(s) - \mu(s)}{\sigma(s)} \right)^{-\frac{1}{\gamma(s)}} \right\},
\]

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

### 3 Domain of attraction

The maximum domain of attraction of extreme value distributions in infinite-dimensional space has been characterized in de Haan and Lin (2001). This result leads directly to a characterization of the domain of attraction of a generalized Pareto process.

Let \( C(S) \) be the space of real continuous functions in \( S \), with \( S \subset \mathbb{R}^d \) some compact subset, equipped with the supremum norm. The convergences below \( \rightarrow^d \) denote weak convergence or convergence in distribution. 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

\[
\eta = (\bar{\eta} \gamma - 1)^{\frac{1}{\gamma}}, \quad \text{for some } \bar{\eta} \text{ and continuous function } \gamma = \{ \gamma(s) \}_s \in S.
\]

For simplicity we always take here

\[
C^+_1(S) = \{ f \in C^+(S) : \sup_{s \in S} f(s) \geq 1 \},
\]

i.e., w.l.g. consider the constant \( \omega_0 \) introduced in Section 2 equal to 1.

The maximum domain of attraction condition in \( C(S) \) can be stated as:

**Condition 3.1.** For \( X, X_1, X_2, \ldots \) i.i.d. random elements of \( C(S) \), there exists a max-stable stochastic process \( \eta \in C(S) \) with continuous index function \( \gamma \), and \( a_s(n) > 0 \) and \( b_s(n) \) in \( C(S) \) such that

\[
\left\{ \max_{1 \leq i \leq n} \frac{X_i(s) - b_s(n)}{a_s(n)} \right\}_{s \in S} \rightarrow^d \{ \eta(s) \}_{s \in S}
\]

(3.1)

on \( C(S) \). The normalizing functions are w.l.g. chosen in such a way that

\[
- \log P(\eta(s) \leq x) = (1 + \gamma(s)x)^{-1/\gamma(s)} \text{ for all } x \text{ with } 1 + \gamma(s)x > 0, s \in S.
\]

Next is an equivalent characterization of the domain of attraction condition, in terms of ‘exceedances’.

For \( X \) a random elements of \( C(S) \), suppose the marginal distribution functions \( F_s(x) = P(X(s) \leq x) \) are continuous in \( x \), for all \( s \in S \). To simplify notation write the normalized process,

\[
T_t X = \left( 1 + \gamma \frac{X - b_t}{a_t} \right)^{1/\gamma},
\]

(3.2)
for \( \gamma = \{\gamma(s)\}_{s \in S} \) and normalizing functions \( a_t = \{a_t(s)\}_{s \in S} > 0 \) and \( b_t = \{b_t(s)\}_{s \in S} \), all in \( C(S) \).

**Condition 3.2.** For \( X \in C(S) \) suppose, for some \( a_t > 0 \) and \( b_t \) in \( C(S) \),
\[
\lim_{t \to \infty} t P \left( \frac{X(s) - b_t(s)}{a_t(s)} > x \right) = (1 + \gamma(s)x)^{-1/\gamma(s)}, \quad 1 + \gamma(s)x > 0, \tag{3.3}
\]
uniformly in \( s \),
\[
\lim_{t \to \infty} \frac{P(\sup_{s \in S} T_t X(s) > x)}{P(\sup_{s \in S} T_t X(s) > 1)} = \frac{1}{x}, \quad \text{for all } x > 1, \tag{3.4}
\]
and
\[
\lim_{t \to \infty} P \left( \frac{T_t X}{\sup_{s \in S} T_t X(s)} \in B \mid \sup_{s \in S} T_t X(s) > 1 \right) = \rho(B), \tag{3.5}
\]
for each \( B \in B \left( \bar{C}_1^+ (S) \right) \) with \( \rho(\partial B) = 0 \), with \( \rho \) some probability measure on \( \bar{C}_1^+ (S) \).

Note that this is the same as for max-stable processes; cf. Theorem 9.5.1 in de Haan and Ferreira (2006). Note also that (3.4)–(3.5) specify a simple Pareto process in the limit.

**Theorem 3.1.** Conditions 3.1 and 3.2 are equivalent. Moreover, the limiting probability measure in (3.5) is the probability measure of \( V \) in representation (2.21).

**Proof.** Cf. Theorem 9.5.1 in de Haan and Ferreira (2006). The normalization there is
\[
\left\{ \frac{1}{t(1 - F_s(X(s)))} \right\}_{s \in S},
\]
instead of (3.2) but the results are the same.

The following is a direct consequence:

**Corollary 3.1.** Relations (3.4)–(3.5) imply
\[
\lim_{t \to \infty} P \left( T_t X \in A \mid \sup_{s \in S} T_t X(s) > 1 \right) = P(W \in A),
\]
with \( A \in B \left( C_1^+ (S) \right) \), \( P(\partial A) = 0 \) and \( W \) some simple Pareto process.

The converse statement of Corollary 3.1 is as follows:

**Theorem 3.2.** 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 \( P(X(s) > \tilde{b}_u(s)) \) for some \( s \in S \) \( \to 0 \) as \( u \to \infty \), 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 \( B(C(S)) \),
\[
\lim_{u \to \infty} P \left( \frac{X - \tilde{b}_u}{\tilde{a}_u} \in A \mid X(s) - \tilde{b}_u(s) > 0 \text{ for some } s \in S \right) = \tilde{P}(A),
\]
for all \( A \in B(C(S)) \) and \( \tilde{P}(\partial A) = 0 \). Then Conditions 3.1 and 3.2 are fulfilled.
Proof. By the conditions on $\tilde{b}_u$, we can determine $q = q(t)$ such that $P(X(s) > \tilde{b}_{q(t)}(s)$ for some $s \in S) = 1/t$. Then with $b_t(s) = \tilde{b}_{q(t)}(s)$ and $a_t(s) = \tilde{a}_{q(t)}(s)$,

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

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

$$\lim_{t \to \infty} tP \left( \frac{X - b_t}{a_t} \in C \right) = \hat{P}(C). \tag{3.6}$$

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

$$\lim_{t \to \infty} tP(X(s) > b_t(s) + xa_t(s)) = \hat{P}\{f : f(s) > x\}.$$

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

$$\lim_{t \to \infty} \frac{U_tx(s) - b_t(s)}{a_t(s)} = V_x(s), \quad \text{for } x > 0.$$

It follows (Lemma 10.4.2, p.340, in de Haan and Ferreira (2006)) that for some real $\gamma(s)$ and all $x > 0$

$$\lim_{t \to \infty} \frac{b_t(s) - b_t(s)}{a_t(s)} = \frac{x^{\gamma(s)} - 1}{\gamma(s)} \quad \text{and} \quad \lim_{t \to \infty} \frac{a_t(s)}{a_t(s)} = x^{\gamma(s)}. \tag{3.7}$$

Since the limit process has continuous paths, the function $\gamma$ must be continuous on $S$.

Now replace $t$ in (3.6) by $ct$ where $c > 0$. Then

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

hence, by (3.7)

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

and by (3.6)

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

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

$$\lim_{t \to \infty} tP(T_tX \in A) = P(A),$$

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 (2006).
**Example 3.1.** Any max-stable process verifies Condition (3.2) with $\rho$ given by the probability measure of $V$ from (2.21).

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

**Example 3.3.** The finite dimensional distributions of the moving maximum processes obtained in de Haan and Pereira (2006) can be applied to obtain the correspondent finite dimensional distributions of the correspondent Pareto process, in the appropriate region.

**Example 3.4** (Regular variation (de Haan and Lin 2001, Hult and Lindskog 2005)). A stochastic process $X$ in $C(S)$ is regularly varying if and only if there exists an $\alpha > 0$ and a probability measure $\rho$ such that

\[
P(\sup_{s \in S} X(s) > tx, X/\sup_{s \in S} X(s) \in \cdot) \rightarrow x^{-\alpha} \rho(\cdot), \quad x > 0, \ t \to \infty,
\]

on $\{ f \in C(S) : \sup_{s \in S} f(s) = 1 \}$. Hence, a regularly varying process such that (3.3) holds for the marginals, verifies Condition (3.2) with $\gamma = 1/\alpha$, $b_t = t$ and $a_t = t/\alpha$; note that the index function is constant in this case.

On the other hand, the normalized process $tT_t X$ with $T_t X$ verifying (3.4)–(3.5), verifies regular variation with $\alpha = 1$ and spectral measure $\rho$ on $C^+_1(S)$.

**Remark 3.1.** As seen in Section 2.1.1, our analysis is also valid in the finite-dimensional set-up. The main difference from Rootzén and Tajvidi (2006) 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 $F = 1 - F$ with $F$ some $d$-variate distribution function, $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, and $u(\cdot) = (u_1(\cdot), u_2(\cdot), \ldots, u_d(\cdot))$ and $\sigma$ the normalizing functions considered in Rootzén and Tajvidi (2006) (see e.g. their definition of $X_u$). By using $\sigma(x)/\sigma(t) = (x_1^{\gamma_1}, x_2^{\gamma_2}, \ldots, x_d^{\gamma_d})$ and

\[
(u(x) - u(t))/\sigma(t) \to \left(\frac{x_1^{\gamma_1} - 1}{\gamma_1}, \frac{x_2^{\gamma_2} - 1}{\gamma_2}, \ldots, \frac{x_d^{\gamma_d} - 1}{\gamma_d}\right), \ t \to \infty,
\]

for some reals $\gamma_1, \gamma_2, \ldots, \gamma_d$ (cf. proof of Theorem 2.1(ii) in Rootzén and Tajvidi (2006)) and by

\[
F^*(x) := F(u_1(x_1), u_2(x_2), \ldots, u_d(x_d)),
\]

one simplifies their relation (19) to

\[
tF^*(tx) \to - \log G \left(\frac{x_1^{\gamma_1} - 1}{\gamma_1}, \frac{x_2^{\gamma_2} - 1}{\gamma_2}, \ldots, \frac{x_d^{\gamma_d} - 1}{\gamma_d}\right),
\]

and one simplifies their relation (6) to

\[
P(X^* \leq tx|X^* \leq t) = P(X^* \leq x)
\]

for $t \geq 1$. Hence one can take $u(t) := \left(\frac{t^{\gamma_1} - 1}{\gamma_1}, \frac{t^{\gamma_2} - 1}{\gamma_2}, \ldots, \frac{t^{\gamma_d} - 1}{\gamma_d}\right)$ in Theorem 2.2 of that paper.
4 View towards application and simulation

4.1 Towards application

The following remark may be useful for application. Suppose the domain of attraction condition (1.1) holds. Define \( B = \{ f \in C^+(S) : f > 1 \} \). Let \( A \) be a Borel set in \( C^+(S) \). Then applying (1.1) twice we get

\[
\lim_{t \to \infty} P(T_t X \in A | T_t X \in B) = \frac{\nu(A \cap B)}{\nu(B)}.
\]

This is the content of Corollary 3.1. 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 \( B' = \{ f \in C^+(S) : T_t f(s_i) \geq 1 \text{ for } i = 1, \ldots, p \} \). Then

\[
\lim_{t \to \infty} P(T_t X \in A | T_t X \in B') = \frac{\nu(A \cap B')}{\nu(B')}
\]

which again is a probability distribution.

Now we proceed as in the peaks-over-threshold method for scalar observations: let \( k = k(n) \) be a sequence of integers with \( \lim_{n \to \infty} k(n) = \infty \) and \( \lim_{n \to \infty} k(n)/n = 0 \), as \( n \to \infty \). 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 \( i = 1, \ldots, p \). The probability distribution of those selected observations is approximately the right-hand side of (1.1). 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 i.e., there exist functions $a_n(s) > 0$ and $b_n(s)$ ($s \in S$) such that the sequence of i.i.d. processes

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

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

Define

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

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

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

$$P \left( \frac{T_t - t_0 T_t X - b_{t_0}}{a_{t_0}} \in A \middle| R_{T_t} > 1 \right) = \begin{array}{l}
P \left( a_t \left( t_0 \left( 1 + \gamma(s) \frac{X(s) - b_{t_0}}{a_t(s)} \right)^{1/\gamma(s)} \right)^{\gamma} - b_{t_0} - b_t \in A \middle| R_{T_t} > 1 \right) \\
+ P \left( a_t \left( t_0 \left( 1 + \gamma(s) \frac{X(s) - b_{t_0}}{a_t(s)} \right)^{1/\gamma(s)} \right)^{\gamma} - b_{t_0} - b_t \in A \middle| R_{T_t} > 1 \right)
\end{array}$$

Since,

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

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

$$P \left( \frac{X - b_t}{a_t} \in A \middle| R_{T_t} > 1 \right)$$

which is $P(W \in A)$ by Corollary 3.1.

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 i.e. ones that satisfy $X \leq b_n$. These are the moderately heavy storms. However we want
to know how the storm field of a really heavy storm (i.e. \( X \leq b_N \) with \( N > n \)) looks like. That is exactly what relation (4.2) does. Take a moderately heavy storm \( X \) and transform it to \( T_{n/k}^{-1} T_n X \). 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 \( \gamma, a \) 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 \( X_1, X_2, \ldots, X_n \) 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 \( \gamma, a \) and \( b \) (de Haan and Lin (2003), Einmahl and Lin (2006)); 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

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

those that satisfy \( X_i(s) > \hat{b}_{n/k}(s) \) for some \( s \in S \) i.e. 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 i.e., in the end we obtain the processes

\[
\tilde{T}_{n/k}^{-1} t_0 \hat{T}_{n/k} X_i \text{ for those } X_i \text{ for which } \hat{R}_{T_{n/k}} X_i > 1.
\]

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

**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 GP from there. But estimation of spectral measure is more difficult (de Haan and Lin (2003)) although this procedure is less restrictive on the number of observations than 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 R-package due to Ribatet (2011). In Figures 1-2 are represented the 11 out of 20 of these processes, normalized and for which $\hat{R}_{T_n} > 1$, and the corresponding lifted processes $\tilde{T}_{n/k} t_0 T_{n/k} X_i$ 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.

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

References

[1] Balkema, A.A. and de Haan, L. (1974) Residual life time at great age. *Ann. Probab.* 2, 792–804

[2] Billingsley, P. (1995) Probability and Measure (3rd. edition). John Wiley & Sons Inc.

[3] Einmahl, J.H.J., and Lin, T. (2006). Asymptotic normality of extreme value estimators on C[0,1]. *Ann. Statist.* 34, 469-492.
Figure 2: All processes: $\hat{T}_{n/k}X_i$ - slashed lines - and $\hat{T}_{n/k}^\leftarrow t_0\hat{T}_{n/k}X_i$ - continuous lines for which $\hat{R}_{n/k}X_i > 1$

[4] Falk, M., Hüsler, J. and Reiss, R.-D. (2010) Laws of Small Numbers: Extremes and Rare Events, Birkhäuser, Springer.

[5] de Haan, L. (1984) A spectral representation for max-stable processes. *Ann. Prob.* 12, 1194–1204

[6] de Haan, L. and Ferreira, A. (2006) Extreme Value Theory: An Introduction. Springer, Boston.

[7] de Haan, L. and Lin, T. (2001) On convergence toward an extreme value distribution in $C[0,1]$. *Ann. Prob.* 29, 467–483

[8] de Haan, L. and Lin, T. (2003). Weak consistency of extreme value estimators in $C[0,1]$. *Ann. Statist.* 31, 1996–2012.

[9] de Haan, L. and Pereira, T.T. (2006) Spatial Extremes: the stationary case. *Ann. Statist.* 34, 146–168

[10] de Haan, L. and Pickands, J. (1986) Stationary min-stable stochastic processes. *Probab. Th. Rel. Fields* 72, 477–492

[11] Hult, H. and Lindskog, F. (2005) Extremal behaviour of regularly varying stochastic processes. *Stoch. Proc. Appl.* 1, 249–274

[12] Penrose, M.D. (1992) Semi-min-stable processes. *Ann. Probab.* 20, 1450–1463
[13] Pickands, J. III (1975) Statistical inference using extreme order statistics. *Ann. Statist.* 3, 119-131.

[14] Ribatet, M. (2011) SpatialExtremes: Modelling Spatial Extremes. R package

[15] H. Rootzén and N. Tajvidi (2006) Multivariate generalized Pareto distributions. *Bernoulli* 12, 917–930