Multidimensional EVT and max-stable processes

Laurens de Haan

Erasmus University Rotterdam, NL
University of Lisbon, PT

I start with two examples with which I have been concerned.

1) Heavy rainfall can be damaging for agriculture. In particular if the total rainfall in some area (i.e. the integral of the stochastic process) is extreme, there may be a problem. It seems that when one measures rainfall at two places, we have more or less independence if the places are at least 100 km apart.

At shorter range there is dependence. But the dependence may be different in the tail that is, if at least at one place there is a lot of rain. There may be less dependence.

Hence one should take into account dependence at higher levels. I come back to this example later.

2) There is a sizeable governmental research unit (Deltares) in The Netherlands working on models that can help improve the coastal defences against flooding. In particular they study wind storms on the North Sea that can lead to high still water levels as well as high waves near the coast. Since only really severe storms would be a threat and since these storms are not available in the observed storms over a period of say 50 years, there is a need to guess how these severe wind storms would look like.

Again I shall come back to this example.

This leads us to look for a framework that allows us to help solving these problems. The framework is the theory of max-stable processes that can be seen as the infinite-dimensional version of extreme value theory.

Let me start at the beginning, namely one-dimensional extreme value theory.

Let $X, X_1, X_2, \ldots$ be i.i.d. random variables and assume that the extreme value condition is fulfilled:

$$\lim_{n\to\infty} P\left\{\max_{1\le i\le n} \frac{X_i - b(n)}{a(n)} \le x\right\} = \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\}$$

for some $\gamma \in \mathbb{R}$ and suitable sequences $a(n) > 0$ and $b(n)$.

The condition is equivalent to the statement

$$\lim_{n\to\infty} P^n\left\{\frac{X - b(n)}{a(n)} \le x\right\} = \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\}$$

and (when taking logarithms)

$$\lim_{n\to\infty} -n \log P\left\{\frac{X-b(n)}{a(n)} \le x\right\} = (1+\gamma x)^{-1/\gamma},$$

In short: $-n\log P$ converges to a positive limit. This implies $P \to 1$ and hence $-\log P \sim 1-P$.

Consequently we get the equivalent statement

$$\lim_{n\to\infty} n \cdot P\left\{\frac{X-b(n)}{a(n)} > x\right\} = \lim_{n\to\infty} n(1-P) = (1+\gamma x)^{-1/\gamma}.$$

More generally:

$$\lim_{t\to\infty} t \cdot P\left\{ \frac{X - b(t)}{a(t)} > x \right\} = (1 + \gamma x)^{-1/\gamma}$$

Where t runs through the real numbers.

This is the well-known convergence to the generalized Pareto distribution from which the peaks-over-threshold method follows.

Alternatively we can write:

$$\lim_{t\to\infty} t\, P\left\{ \left(1 + \gamma \frac{X - b(t)}{a(t)} \right)^{1/\gamma} > x \right\} = \frac{1}{x}, \quad x > 1 .$$

Please keep in mind that the transformation

$$\left(1+\gamma\frac{X-b(t)}{a(t)}\right)^{1/\gamma}$$

(approximately) turns the tail of the distribution into a standard Pareto distribution, similar to what we do when defining a copula, but now only for the tail.

Now let us take i.i.d. random vectors $(X,Y),(X_1,Y_1),(X_2,Y_2),\dots$ and suppose that the extreme value condition holds:

$$\lim_{n\to\infty} P\left\{\max_{1\le i\le n}\frac{X_i-b_1(n)}{a_1(n)}\le x\ ,\ \max_{1\le i\le n}\frac{Y_i-b_2(n)}{a_2(n)}\le y\right\}=G(x,y)$$

for some distribution function G with non-degenerate marginals. Again we have a series of equivalent statements:

$$P^n\left\{\frac{X-b_1(n)}{a_1(n)}\le x\ ,\ \frac{Y-b_2(n)}{a_2(n)}\le y\right\}\to G(x,y)\ ,$$

$$-n\log P\left\{\frac{X-b_1(n)}{a_1(n)}\le x\ ,\ \frac{Y-b_2(n)}{a_2(n)}\le y\right\}\to -\log G(x,y)$$

and, since again $-\log P \sim 1-P$, this is equivalent to

$$nP\left\{\frac{X-b_1(n)}{a_1(n)}\le x\ ,\ \frac{Y-b_2(n)}{a_2(n)}\le y\right\}^c\to -\log G(x,y)$$

i.e.,

$$n\ P\left\{\frac{X-b_1(n)}{a_1(n)}> x\ \textbf{or}\ \frac{Y-b_2(n)}{a_2(n)}> y\right\}\to -\log G(x,y)\ .$$

We use the one-dimensional result to standardize each component:

$$n\,P\left\{\left(1+\gamma_1\frac{X-b_1(n)}{a_1(n)}\right)^{1/\gamma_1}>x\ \ \textbf{or}\ \ \left(1+\gamma_2\frac{Y-b_2(n)}{a_2(n)}\right)^{1/\gamma_2}>y\right\}$$

$$\to\log G\left(\frac{x^{\gamma_1}-1}{\gamma_1},\frac{x^{\gamma_2}-1}{\gamma_2}\right)=:-\log G_0(x,y)\ .$$

Then G_0 is standard i.e.,

$$-\log G_0(x,\infty)=-\log G_0(\infty,x)=1/x\ .$$

The defining property of G_0 is: for all $a,x,y>0$

$$-\log G_0(ax,ay)=-\frac{1}{a}\log G_0(x,y)$$

Comments:

1. Standardizing the marginal distributions leads to a limit that is a homogeneous function.
2. There exists a measure, υ say, such that

   $$-\log G_0(x,y) = \nu\{(s,t): s > x \ \textbf{or} \ t > y\} \text{ for } x, y > 0.$$

   Then for $a > 0$ and Borel sets B

   $$\nu(aB) = \frac{1}{a}\nu(B) .$$

3. Even non-extreme components are normalized.

Let us produce a random vector with distribution G_0.

Take a Poisson point process on $(0,\infty)$ with intensity measure $r^{-2}dr$.

Consider a realization $V_1, V_2, \ldots$ of this point process.

Take independently i.i.d. non-negative random vectors $\left(W^{(1)}, W^{(2)}\right), \left(W_1^{(1)}, W_1^{(2)}\right), \left(W_2^{(1)}, W_2^{(2)}\right), \ldots$ with $EW^{(i)} = 1$ for $i = 1, 2$.

Denote the distribution function of $\left(W^{(1)}, W^{(2)}\right)$ by H.

Since the $\left(W_i^{(1)}, W_i^{(2)}\right)$ are random vectors, we can alternatively consider a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^2$ with intensity measure

$$Q\left(dr, du, dv\right) = r^{-2} dr\ H\left(du, dv\right).$$

We then get the realizations:

$$\left(V_1, W_1^{(1)}, W_1^{(2)}\right), \left(V_2, W_2^{(1)}, W_2^{(2)}\right), \ldots$$

Consider the vector

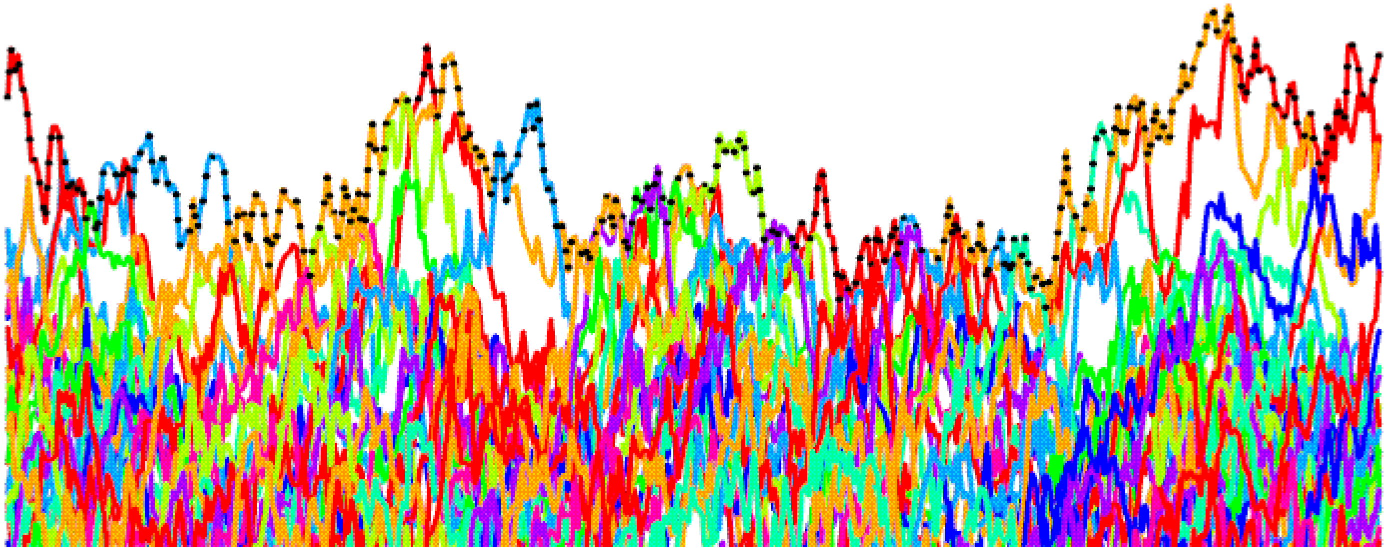
$$Z := \left(\bigvee_{i=1}^{\infty} V_i W_i^{(1)}, \bigvee_{i=1}^{\infty} V_i W_i^{(2)} \right)$$

Let us calculate the distribution function of Z.

Example 1: $\sigma^2(t) = |t|$ (Brown, Resnick, 1977)

Let W be a standard Brownian motion. The stationarity of $\{V_i(\cdot), i \in \mathbb{N}\}$ follows e.g. from the following remark:

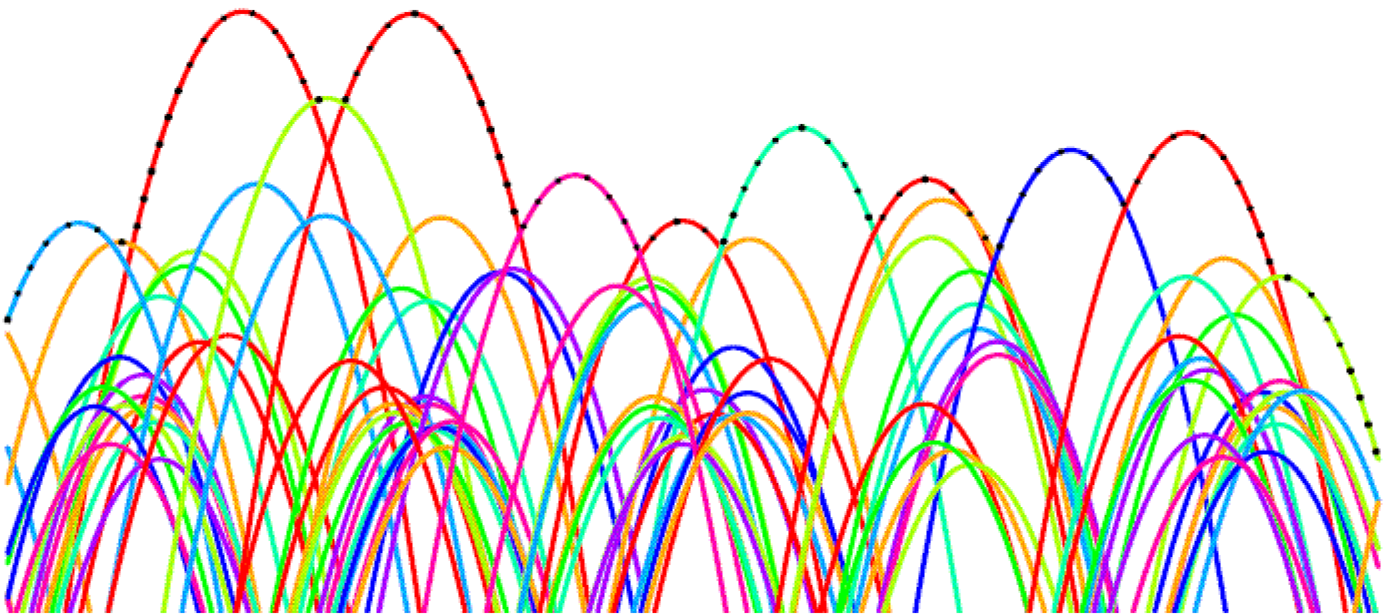
> *The measure $e^{-u}du$ is an invariant measure for the diffusion $W(t) - t/2$.*



Example 2: $\sigma^2(t) = t^2$ (Eddie, Gale, 1981, Hüsler, Hooghiemstra, 1996

In this case we have

> *The process* $\{W(t), t \in \mathbb{R}\}$ *has the same law as* $\{tN, t \in \mathbb{R}\}$, *where* $N \sim \mathcal{N}(0,1)$.



$$G_0^*(x,y) := P\left\{ \bigvee_{i=1}^{\infty} V_i \, W_i^{(1)} \le x \, , \, \bigvee_{i=1}^{\infty} V_i \, W_i^{(2)} \le y \right\}$$

$$= P\left\{ V_i \le \frac{x}{W_i^{(1)}} \text{ and } V_i \le \frac{y}{W_i^{(2)}} \text{ for } i = 1, 2, \ldots \right\}$$

$$= P\left\{ V_i \le \frac{x}{W_i^{(1)}} \wedge \frac{y}{W_i^{(2)}} \text{ for } i = 1, 2, \ldots \right\}$$

$$= \exp\left[-Q\left\{ (r,u,v) : r > \frac{x}{u} \wedge \frac{y}{v} \right\} \right]$$

$$= \exp\left\{ -\iint_{z > \frac{x}{u} \wedge \frac{y}{v}} \frac{dz}{z^2} \, dH(u,v) \right\}$$

$$= \exp\left\{ -\int_0^{\infty} \int_0^{\infty} \frac{u}{x} \vee \frac{v}{y} \, dH(u,v) \right\}.$$

Then clearly for $a, x, y > 0$

$$-\log G_0^*(ax, ay) = -\frac{1}{a} \log G_0^*(x, y).$$

Conversely we can prove: for each extreme value distribution G_0 there exists an H such that

$$G_0(x, y) = \exp\left\{ -\int_0^\infty \int_0^\infty \frac{u}{x} \vee \frac{v}{y} \, dH(u, v) \right\}.$$

Comments:

1. H is called the "spectral measure" of G_0.
2. If $\left(Z_1^{(1)}, Z_1^{(2)}\right), \left(Z_2^{(1)}, Z_2^{(2)}\right), \ldots$ are i.i.d. copies of the random vector $\left(Z^{(1)}, Z^{(2)}\right)$ then for all k

$$\left(\frac{1}{k}\bigvee_{i=1}^{k} Z_i^{(1)}, \frac{1}{k}\bigvee_{i=1}^{k} Z_i^{(2)}\right) \stackrel{d}{=} \left(Z^{(1)}, Z^{(2)}\right)$$

and $Z^{(1)}$ and $Z^{(2)}$ are standard Fréchet,

hence the vector Z is "simple max-stable".

Max-stable processes form the infinite-dimensional extension of this:

Let $X, X_1, X_2, \ldots$ be i.i.d. stochastic processes in $C(S)$ with S a compact subset of $\mathbb{R}^d$.

Assume that for suitable continuous norming functions $a_s(n) > 0$ and $b_s(n)$ the sequence of processes

$$\left\{ \max_{1 \le i \le n} \frac{X_i(s) - b_s(n)}{a_s(n)} \right\}_{s \in S}$$

converges in distribution, in $C(S)$, to a stochastic process

$$\{Y(s)\}_{s\in S}.$$

Then, obviously (using the standardizing transformation),

$$\left\{\max_{1\le i\le n}\left(1+\gamma(s)\frac{X_i(s)-b_s(n)}{a_s(n)}\right)^{1/\gamma(s)}\right\}_{s\in S}$$

converges to

$$\{Z(s)\} := \left\{\left(1+\gamma(s)Y(s)\right)^{1/\gamma(s)}\right\}_{s\in S}.$$

Here $\gamma(s)$ is the extreme value index at location s.

Then the limit process Z is simple max-stable i.e. if $Z_1, Z_2, \ldots$ are i.i.d. copies of Z,

$$\left\{ \frac{1}{k} \max_{1 \le i \le k} Z_i(s) \right\}_{s \in S} \stackrel{d}{=} \left\{ Z_i(s) \right\}_{s \in S} .$$

Correspondingly if $Y_1, Y_2, \ldots$ are i.i.d. copies of Y_1 then for some continuous $A_s(n) > 0$ and $B_s(n) > 0$

$$\left\{ \max_{1 \le i \le n} \left(\frac{Y_i(s) - B_s(n)}{A_s(n)} \right) \right\}_{s \in S} \stackrel{d}{=} \left\{ Y_i(s) \right\}_{s \in S} .$$

i.e. the process Y is max-stable.

The analysis of a simple max-stable process is similar to the finite-dimensional case:

$$\left\{ \max_{1\le i\le n} \left(1+\gamma(s)\frac{X_i(s)-b_s(n)}{a_s(n)} \right)^{1/\gamma(s)} \right\}_{s\in S} \xrightarrow{d} \{Z(s)\}_{s\in S}$$

in space $C(S)$, is equivalent to: for Borel sets $B\in C(S)$

$$\lim_{t\to\infty} t\, P\left\{ \left(1+\gamma(\cdot)\frac{X(\cdot)-b_\cdot(t)}{a_\cdot(t)} \right)^{1/\gamma(\cdot)} \in B \right\} = \nu(B)$$

provided $\nu(\partial B)=0$. Then for the measure ν :

$$\nu(aB)=\frac{1}{a}\nu(B) \text{ for } a>0,\ B \text{ Borel in } C(S).$$

There is a representation as before:

Take a Poisson process on $(0,\infty)$ with intensity measure $r^{-2}dr$. Consider a realization $V_1, V_2, \ldots$ of this point process.

Take independently i.i.d. continuous stochastic processes $W, W_1, W_2, \ldots$ on S with $W(s) \geq 0$ and $EW(s) = 1$ for $s \in S$ and $E \sup_{s \in S} W(s) < \infty$.

Consider for $s \in S$

$$Z(s) := \bigvee_{i=1}^{\infty} V_i \; W_i(s)$$

Simple auxiliary result:

Statement Consider a Poisson point process on $(0,\infty)$ with intensity measaure $r^{-2}dr$.

The following two operations lead to the same point process (in distribution):

1) multiply all points by an integer k

2) change the intensity measure from $r^{-2}dr$ to $kr^{-2}dr$.

Proof

The distribution of any Poisson process is determined by probability (for any $0 < a < b < \infty$) that the process has no points in the interval $[a,b]$.

For the original Poisson points process the probability is $\exp\left\{-\int_a^b r^{-2} dr\right\}$.

Now the number of points in $[a,b]$ for the transformed process (transformed by multiplication) equals the number of points for the original process in $\left[\frac{a}{k},\frac{b}{k}\right]$.

Hence the probability of no point of the new process in $[a,b]$ is

$$\exp\left\{-\int_{a/k}^{b/k} r^{-2}dr\right\} = \exp\left\{-\int_a^b kr^{-2}dr\right\}. \qquad \square$$

Remark

This holds also for our point process with "points"

$$\left\{V_i W_i\left(s\right)\right\}.$$

I now argue that Z **is** a simple max-stable process i.e., that

$$\left\{ \frac{1}{k} \overset{k}{\underset{i=1}{\vee}} Z_j(s) \right\}_{s \in S} \overset{d}{=} \{ Z(s) \}_{s \in S}$$

Check:

On one side we have

$$k\ Z(s) = k \overset{\infty}{\underset{i=1}{\vee}} V_i\ W_i(s)$$

(all points multiplied by k).

On the other side we have

$$\bigvee_{i=1}^{k} Z_i(s) = \bigvee_{i=1}^{\infty} \bigvee_{j=1}^{k} V_{i,j} \; W_{i,j}(s).$$

When forming $\bigvee_{j=1}^{k} V_{i,j} \; W_{i,j}(s)$ we have effectively joined all the points of k independent processes into one process.

This is the same (in distribution) as multiplying the intensity measure by k. $\square$

Conversely every simple max-stable process can be represented in this way.

Hence the probability distribution of simple max-stable processes is completely determined by the distribution of the underlying stochastic process W:

$$\{Z(s)\} = \left\{ \bigvee_{i=1}^{\infty} V_i \, W_i(s) \right\}.$$

The probability measure of the process W is called the **spectral measure** of the process Z .

Remark

Note that the process W *may be zero somewhere but (Giné, Hahn, Vatan, 1991)*

$$P\{Z(s)>0 \quad \text{for} \quad s\in S\}=1.$$

It is not always easy to detect the structure that we found in concrete examples.

Examples

1. Take $Z_1, Z_2, \ldots$ i.i.d and $P\{Z_1 \le x\} = \exp(-1/x), x > 0$.

   For $0 \le s \le 1$ define

   $$Z(s) := \max\left(sZ_1 (1-s) Z_2\right)$$

   It is easy to verify directly that this process is simple max-stable.

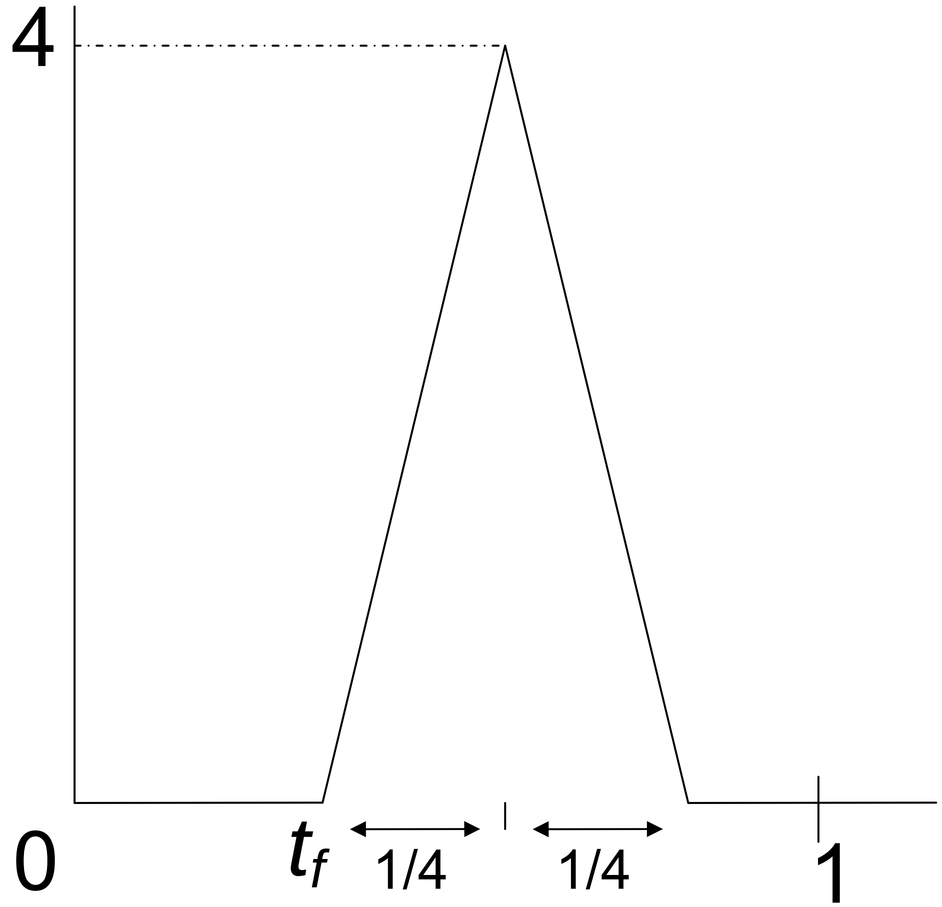
   In standard form:

   $$Z(s) := \bigvee_{i=1}^{k} V_i . \left(s(1-Q_i) + (1-s) Q_i\right)$$

   where $P\{Q_i = 0\} = P\{Q_i = 1\} = 1/2$.

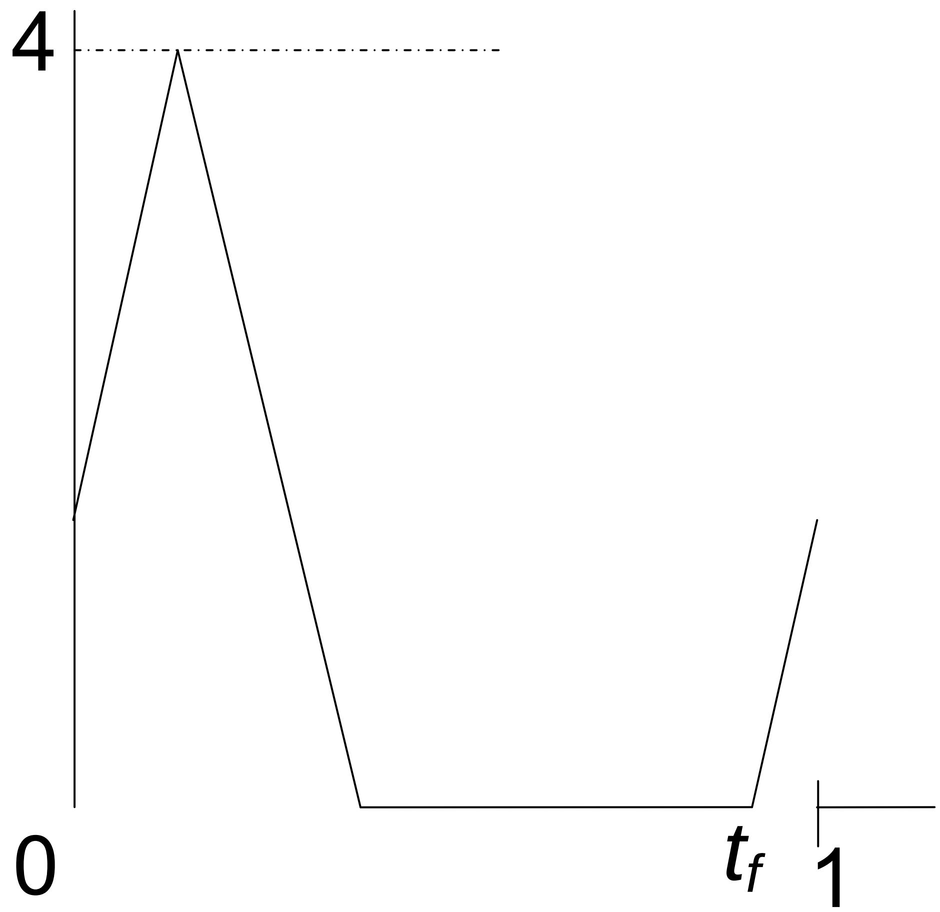
   Note that $Z(0)$ and $Z(1)$ are independent.

2. Define the class of functions f_t by



for $t =: t_f \leq 1/2$

and by



for $t > 1/2$.

Introduce a probability distribution on this class of functions by

$$\rho\left\{ f : a \le t_f \le b \right\} = b - a \quad \text{for} \quad a, b \in (0,1), a < b$$

(uniform distribution).

Then (side condition)

$$\int f(s)\, d\rho(f) = \frac{1}{4} \cdot 4 = 1 \quad \text{for} \quad s \in [0,1] \, .$$

This is the stochastic process W.

Note that $Z(a)$ and $Z(b)$ are independent for $b - a > 1/2$.

Generalization of this example:

We can take quite general probability density functions and move them horizontally over the entire real line. Such processes are called *max-moving average* processes.

3. If $P\{W(s)=0 \text{ for some } s\}=0$, then $Z(s_0)$ and $Z(s_1)$ are dependent for all s_0 and s_1.

Example (S is not compact here : $S=\mathbb{R}$):

Recall the representation

$$Z(s) := \bigvee_{i=1}^{\infty} V_i \, W_i(s).$$

Take $W(s) := \exp\{B(s) - |s|/2\}$ with B Brownian motion (in two directions from zero).

In that case Z is stationary but W is clearly not (Brown and Resnick 1977)!.

4. If in contrast

$$P\{W(s) > 0 \text{ for all } s\} = 0$$

then we say that the process is "asymptotically independent" generalizing a concept from finite-dimensional extremes (Ledford and Tawn 1998).

Applications

We go back to our two examples/applications in the beginning.

1. Rainfall

   Daily rainfall is monitored at about 30 stations throughout North Holland. The question is:

   what is the probability that the total rainfall in the area on one day exceeds 60 mm in some given year?

The meteorologists agree that the shape parameter $\gamma(s)$ is constant in the area of interest i.e. $\gamma(s) \equiv \gamma \approx 0.1$.

We are interested in the total rainfall in the area i.e. $\int_s X(s)\,ds$, Hence we need the extreme value condition for this random variable. We have the following theorem (stated by Coles and Tawn 1996).

Theorem *Suppose that process* X *is in the domain of attraction of the max-stable process* Z *i.e.,*

$$\left\{ \max_{1\le i\le n} \left(1+\gamma(s)\frac{X_i(s)-b_s(n)}{a_s(n)} \right)^{1/\gamma(s)} \right\}_{s\in S} \xrightarrow{d} \{Z(s)\}_{s\in S}$$

with spectral measure ρ*.*

Under certain mild conditions,

$$\lim_{t\to\infty} tP\left\{ \frac{\int_S X(s)\,ds - \int_S b_s(t)\,ds}{\int_S a_s(t)\,du} > x \right\} = \theta_\gamma (1+\gamma x)^{-1/\gamma}.$$

where

$$\theta_\gamma = \int_{\overline{C}_1^+(s)} \left(\int_S A(s) g^\gamma (s/ds) \right)^{1/\gamma} d\rho(g), \qquad A(s) := \frac{a_s(t)}{\int_S a_s(t)\,du}$$

and ρ *is the spectral measure acting on*

$$\overline{C}_1^+(s) := \{ f \in C(s) : f \geq 0, \|f\|_\infty = 1 \}. \qquad \square$$

Hence $\int_S X(s)\,ds$ is in a one-dimensional domain of attraction.

We can estimate all elements of this formula.

We find that the return period of the event is about 250 years.

2. Wind storms

In the Netherlands much effort goes into the protection of the sea coast against flooding. Since only severe storms are a threat to the coast, extreme value theory comes in into the analysis.

Models are available for how the atmospheric conditions influence the general water level near the coast as well as the wave activity.

These in turn are crucial for decisions on how to build the sea defences.

At present the input for the atmospheric conditions and the water model consists of wind observation at Schiphol airport as well as wave observations at a few locations not far from the coast on the North Sea.

In the future the entire wind field over a certain period will be available for the analysis. These are artificial data.

These data will contain some rather big storms but not really devastating storms (since that would require unreasonably long data sets).

The question is to guess how devastating storms look like (i.e. the wind fields) if we only observe big but not really severe storms.

This can be done by using the homogeneity property of the limit measure ν. Remember

$$\nu(aB) = \frac{1}{a}\nu(B) \qquad \text{for} \quad a > 0, B \in C(S) \; .$$

The measure ν is closely related to the probability distribution of the peaks-over-threshold observations i.e. the storms for which the windspeed at some location exceeds a certain threshold.

Hence we can standardize the marginal distribution of the selected storms, then **multiply** the standardized wind speed everywhere by a certain factor $a > 1$ and then transform back.

This procedure results in peak-over-threshold observations at a much higher threshold and it can be considered as a kind of extreme value bootstrap.

Let us have a closer look.

For some time people have attempted to build a big scale model of what happens:

- start from atmospheric conditions: wind fields in a quite big area relevant to the Dutch coast.
- a model is available on how such wind field produces high still water levels and waves off shore
- another model is available on how this translates into high still water levels and waves near the coast.

In this setup – that will be used in the future – we need wind fields. They will be created by a climate model.

This leads to the following question.

The climate model will produce a number of wind storms. Some of them may be quite severe but not severe enough to create a danger to the coast. Can we use these windstorms to produce other wind storms (from the same model) that are really severe and could create problems?

(consultant firm Deltares – Sofia Caires) .

Use homogeneity property:

- We have n observations that are spatial stochastic processes: $\{X_i(s)\}_{s,i}$.
- We want to produce some extreme storms that is, stochastic processes from the same distribution and such that $X(s_0) > c$ at some main point of interest s_0 where c is large or chosen such that $P\{X(s_0) > c\} = p$ with p a given small number.

For the explanation it is useful to start in $\mathbb{R}^1$, go then to $\mathbb{R}^2$ and finally to the function space.

In $\mathbb{R}^1$: domain of attraction condition $\Rightarrow$

$$\lim_{t\to\infty} P\left\{ \frac{X-b(t)}{a(t)} > x \,\middle|\, X > b(t) \right\} = (1+\gamma x)^{-1/\gamma}$$

where $b(t) := \inf\{s : 1-F(s) \geq 1/t\}$.

Out of i.i.d. observations $X_1, X_2, \ldots X_n$ select those that are bigger than $b(t)$.

Let us denote the selected observations as

$$X_1^{(1)}, X_2^{(1)}, \ldots, X_k^{(1)}$$

These are again i.i.d. and satisfy (with $X^{(1)}$ one of the $X_1^{(1)}$'s)

$$P\left\{\frac{X^{(1)}-b(t)}{a(t)} > x\right\} \approx (1+\gamma x)^{-1/\gamma}$$

This is the peaks-over-threshold method.

Next we may replace the variable x at both sides in this relation by $(x^{\gamma}-1)/\gamma$ where $x>1$.

This results in

$$P\left\{\frac{X^{(1)}-b(t)}{a(t)}>\frac{x^{\gamma}-1}{\gamma}\right\}\approx\left(1+\gamma\frac{x^{\gamma}-1}{\gamma}\right)^{-1/\gamma}=\frac{1}{x}$$

or equivalently

$$P\left\{\left(1+\gamma\frac{X^{(1)}-b(t)}{a(t)}\right)^{1/\gamma}>x\right\}\approx\frac{1}{x}$$

for $x>1$ (standardized peak observation).

Now consider for $t_0 > 1$

$$X_i^{(2)} := t_0 \left(1 + \gamma \frac{X_i^{(1)} - b(t)}{a(t)} \right)^{1/\gamma}, i = 1, 2, 3 \ldots, k$$

Then

$$P\left\{ X_i^{(2)} > x \right\} \approx 1 \wedge \frac{t_0}{x}$$

Finally write (take the standardization away)

$$X_i^{(3)} := a(t) \frac{\left(X_i^{(2)} \right)^{\gamma} - 1}{\gamma} + b(t) \ , \ i = 1, 2, \ldots, k$$

We are going to evaluate the probability distribution of $X^{(3)}$.

We need the following result:

Lemma : For $c > 0$

$$\lim_{t\to\infty} \frac{b(tc)-b(t)}{a(t)} = \frac{c^{\gamma}-1}{\gamma}$$

$$\lim_{t\to\infty} \frac{a(tc)}{a(t)} = c^{\gamma}$$

and find

$$P\left\{\frac{X^{(3)}-b(tt_0)}{a(tt_0)}>x\right\}=P\left\{X^{(3)}>b(tt_0)+x\,a(tt_0)\right\}$$

$$=P\left\{a(t)\frac{\left(X^{(2)}\right)^{\gamma}-1}{\gamma}+b(t)>b(tt_0)+x\,a(tt_0)\right\}$$

$$=P\left\{X^{(2)}>\left(1+\gamma\frac{x\,a(tt_0)+b(tt_0)-b(t)}{a(t)}\right)^{1/\gamma}\right\}$$

$$=P\left\{\left(1+\gamma\frac{X^{(1)}-b(t)}{a(t)}\right)^{1/\gamma}>t_0^{-1}\left(1+\gamma\frac{x\,a(tt_0)+b(tt_0)-b(t)}{a(t)}\right)^{1/\gamma}\right\}$$

$$=P\left\{\left(1+\gamma\frac{X^{(1)}-b(t)}{a(t)}\right)^{1/\gamma}>\left(t_0^{-\gamma}+\gamma\frac{x\,a(tt_0)+b(tt_0)-b(t)}{t_0^{\gamma}a(t)}\right)^{1/\gamma}\right\}$$

$$=P\left\{1+\gamma\frac{X^{(1)}-b(t)}{a(t)}>t_0^{-\gamma}+\gamma\left[x\left(1+o(1)\right)+t_0^{-\gamma}\cdot\frac{t_0^{\gamma}-1}{\gamma}\left(1+o(1)\right)\right]\right\}$$

$$= P\left\{1+\gamma\frac{X^{(1)}-b(t)}{a(t)} > \left(1+o(1)\right)+\gamma\left[x\left(1+o(1)\right)+\frac{t_0^{-\gamma}}{\gamma}o(1)\right]\right\}$$

$$= P\left\{\frac{X^{(1)}-b(t)}{a(t)} > x\left(1+o(1)\right)+o(1)\right\} \approx \left(1+\gamma x\right)^{-1/\gamma}.$$

Hence for all x

$$P\left\{\frac{X^{(3)}-b(tt_0)}{a(tt_0)} > x\right\} \approx P\left\{\frac{X^{(1)}-b(t)}{a(t)} > x\right\}.$$

As before, in this relation we can take

$$b(tt_0) := \inf\{s : 1 - F(s) \le 1/(tt_0)\}$$

i.e., the probability that X exceeds $b(tt_0)$ is exactly $1/(tt_0)$.

Now $b(tt_0)$ is also a quantile of the probability distribution of X, but a much higher quantile than $b(t)$ since it is connected with the exceedance probability $1/(tt_0)$ instead of $1/t$.

This means that we have transformed peaks-over-threshold observations $X_1^{(1)},\ldots,X_k^{(1)}$ with threshold $b(t)$ to peaks-over-threshold observations $X_1^{(3)},\ldots,X_k^{(3)}$ with threshold $b(tt_0)$.

Hence we have taken observations that are somewhat exceptional but not really extreme and "lifted" them to observations that are really extreme $(\text{since } t_0 > 1)$.

In $\mathbb{R}^2$: domain of attraction condition $\Rightarrow$

$$P\left\{X > b_1(t) + x\,a_1(t) \;\; \underline{\underline{\text{or}}} \;\; Y > b_2(t) + y\,a_2(t) \,\middle|\, X > b_1(t) \;\; \underline{\underline{\text{or}}} \;\; Y > b_2(t)\right\}$$

$$\to \frac{-\log G(x,y)}{-\log G(0,0)}$$

for $x > 0$ or $y > 0$.

We have a sample (i.e. i.i.d. observations) $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ from this distribution.

Select those observations (X_i,Y_i) for which $X_i > b_1(t)$ **or** $X_i > b_2(t)$.

Let us denote those vectors as

$$\left(X_1^{(1)},Y_1^{(1)}\right),\left(X_2^{(1)},Y_2^{(1)}\right),\ldots,\left(X_k^{(1)},Y_k^{(1)}\right) .$$

These are again i.i.d. random vectors and they satisfy the approximate relation

$$P\left\{\frac{X^{(1)}-b_1(t)}{a_1(t)}>x \quad \underline{\underline{\text{or}}} \quad \frac{Y^{(1)}-b_2(t)}{a_2(t)}>y\right\}\approx\frac{-\log G(x,y)}{-\log G(0,0)} \quad .$$

Let us remark that as before b_1 and b_2 are essentially quantiles:

$$b_1(t)=\inf\left\{s:P\left\{X>s\right\}\geq 1/t\right\}$$

$$b_2(t)=\inf\left\{s:P\left\{Y>s\right\}\geq 1/t\right\} .$$

Next we proceed as before. Consider the random vectors

$$\left(X_i^{(2)},Y_i^{(2)}\right):=\left(t_0\left(1+\gamma_1\frac{X_i^{(1)}-b_1(t)}{a_1(t)}\right)^{1/\gamma_1},t_0\left(1+\gamma_2\frac{Y_i^{(1)}-b_2(t)}{a_2(t)}\right)^{1/\gamma_2}\right)$$

and finally introduce

$$\left(X_i^{(3)},Y_i^{(3)}\right):=\left(a_1(t)\frac{\left(X_i^{(2)}\right)^{\gamma_1}-1}{\gamma_1}+b_1(t),a_2(t)\frac{\left(X_i^{(2)}\right)^{\gamma_2}-1}{\gamma_2}+b_2(t)\right)$$

for $i=1,2,\ldots,k$.

Then

$$P\left\{\frac{X^{(3)}-b_1(tt_0)}{a_1(tt_0)}>x \quad \underline{\underline{\text{or}}} \quad \frac{Y^{(3)}-b_2(tt_0)}{a_2(tt_0)}>y\right\}$$

$$=P\left\{X^{(3)}>b_1(tt_0)+x\,a_1(tt_0) \quad \underline{\underline{\text{or}}} \quad Y^{(3)}>b_2(tt_0)+y\,a_2(tt_0)\right\}$$

$$=P\left\{a_1(t)\ \frac{\left(X^{(2)}\right)^{\gamma_1}-1}{\gamma_1}-b_1(t)>b_1(tt_0)+x\,a_1(tt_0)\ \ \underline{\underline{\text{or}}}\ a_2(t)\ \frac{\left(Y^{(2)}\right)^{\gamma_2}-1}{\gamma_2}-b_2(t)>b_2(tt_0)+y\,a_2(tt_0)\right\}$$

$$=P\left\{X^{(2)}>\left(1+\gamma_1\frac{x\,a_1(tt_0)+b_1(tt_0)-b_1(t)}{a_1(t)}\right)^{1/\gamma_1} \quad \underline{\underline{\text{or}}}\ \ Y^{(2)}>\left(1+\gamma_2\frac{y\,a_2(tt_0)+b_2(tt_0)-b_2(t)}{a_2(t)}\right)^{1/\gamma_2}\right\}$$

$$=P\left\{\left(1+\gamma_1\frac{X^{(1)}-b_1(t)}{a_1(t)}\right)^{1/\gamma_1}>t_0^{-1}\left(1+\gamma_1\frac{x\,a_1(tt_0)+b_1(tt_0)-b_1(t)}{a_1(t)}\right)^{1/\gamma_1} \quad \underline{\underline{\text{or}}}\right.$$

$$\left.\left(1+\gamma_2\frac{Y^{(1)}-b_2(t)}{a_2(t)}\right)^{1/\gamma_2}>t_0^{-1}\left(1+\gamma_2\frac{y\,a_2(tt_0)+b_2(tt_0)-b_1(t)}{a_1(t)}\right)^{1/\gamma_2}\right\}$$

From the *Lemma* we know that

$$t_0^{-1}\left(1+\gamma_1\frac{x\,a_1(tt_0)+b_1(tt_0)-b_1(t)}{a_1(t)}\right)^{1/\gamma_1}=\left(t_0^{-\gamma_1}+\gamma_1\frac{x\,a_1(tt_0)+b_1(tt_0)-b_1(t)}{t_0^{\gamma_1}a_1(t)}\right)^{1/\gamma_1}$$

converges to $(1+\gamma_1 x)^{1/\gamma_1}$ at $t\to\infty$ and that

$$t_0^{-1}\left(1+\gamma_2\frac{ya_2(tt_0)+b_2(tt_0)-b_2(t)}{a_2(t)}\right)^{1/\gamma_2}=\left(t_0^{-\gamma_2}+\gamma_2\frac{ya_2(tt_0)+b_2(tt_0)-b_2(t)}{t_0^{\gamma_2}a_2(t)}\right)^{1/\gamma_2}$$

converges to $(1+\gamma_2 y)^{1/\gamma_2}$.

Hence

$$P\left\{\frac{X^{(3)}-b_1(tt_0)}{a_1(tt_0)}>x \quad \underline{\underline{\text{or}}} \quad \frac{Y^{(3)}-b_2(tt_0)}{a_2(tt_0)}>y\right\}$$

$$\approx\left\{\frac{X^{(1)}-b_1(t)}{a_1(t)}>x \quad \underline{\underline{\text{or}}} \quad \frac{Y^{(1)}-b_2(t)}{a_2(t)}>y\right\} \quad .$$

Once again we have transformed peaks-over-threshold observations $\left(X_1^{(1)},Y_1^{(1)}\right),\left(X_2^{(1)},Y_2^{(1)}\right),\ldots,\left(X_k^{(1)},Y_k^{(1)}\right)$ with threshold $b(t)$ to peaks-over-threshold observations $\left(X_1^{(3)},Y_1^{(3)}\right),\left(X_2^{(3)},Y_2^{(3)}\right),\ldots,\left(X_k^{(3)},Y_k^{(3)}\right)$ with threshold $b(tt_0)$.

Note:

1. Selection criterion: at least one component is big.
2. Also the non-extreme components are normalized.

Spatial observations

I start again with a brief introduction to part of the theory.

Consider n independent stochastic processes

$$\left\{X_i(s)\right\}_{s\in C}, i=1,2,\ldots,n$$

with the same probability distribution where C is a compact set (for example C could be the area around the North Sea coast that we consider and $X_i(s)$ the wind speed at location s at hour i).

We assume that $X_i(s)$ is continuous in s.

We suppose as before that the maximum of those processes converges in distribution that is, the sequence of stochastic processes

$$\left\{ \max_{1 \le i \le n} \frac{X_i(s) - b_s(n)}{a_s(n)} \right\}_{s \in C}$$

converges to some continuous process $\{Y(s)\}_{s \in C}$ in distribution (in C-space).

If we have such convergence, the limit process $Y(s)$ is **stable** with respect to taking the maximum i.e., there are continuous functions $A_s(n) > 0$ and $B_s(n)$ such that, if $Y_1, Y_2, \ldots, Y_n$ are i.i.d. copies of Y, the process

$$\left\{ \max_{1 \le i \le n} \frac{Y_i(s) - B_s(n)}{A_s(n)} \right\}_{s \in C}$$

has the same probability distribution as $\{Y(s)\}_{s \in C}$ itself.

There is also a peaks-over-threshold method in this context.

By way of introduction I formulate the relation in $\mathbb{R}^2$ in a different way.

It says that the random vector

$$\left(\frac{X - b_1(n)}{a_1(n)}, \frac{Y - b_2(n)}{a_2(n)} \right),$$

under the condition that

$$\max\left(\frac{X - b_1(n)}{a_1(n)}, \frac{Y - b_2(n)}{a_2(n)} \right) > 0,$$

has approximately a fixed probability distribution not depending on n (if n is large).

The direct analogue of this statement holds for stochastic processes (but it is much more difficult to prove, see e.g. section 9.3 of our book).

The process

$$\left\{\frac{X(s)-b_s(t)}{a_s(t)}\right\}_{s\in C},$$

has, under the condition that

$$\max_{s\in C}\frac{X(s)-b_s(t)}{a_s(t)}>0,$$

approximately (for large t) a fixed probability distribution, say the probability distribution of some stochastic process $\{V(s)\}_{s\in C}$ (the generalized Pareto process).

Note that

$$\max_{s \in C} \frac{X(s) - b_s(t)}{a_s(t)} > 0$$

is equivalent to:

$$X(s) > U_s(t)$$

for at least one s .

As a result we get the peaks-over-threshold method in this context: out of the n i.i.d. stochastic processes $\{X_i(s)\}$ select those for which

$$\max_{s\in C} \frac{X_i(s)-b_s(t)}{a_s(t)} > 0 .$$

Denote the selected processes as

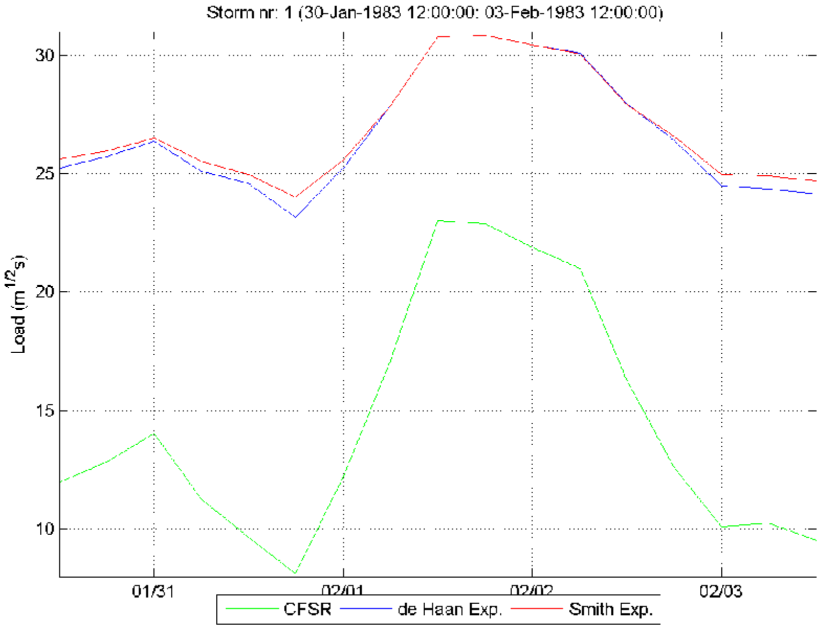
$$\left\{X_1^{(1)}(s)\right\}_{s\in C}, \left\{X_2^{(1)}(s)\right\}_{s\in C}, \ldots, \left\{X_k^{(1)}(s)\right\}_{s\in C} .$$

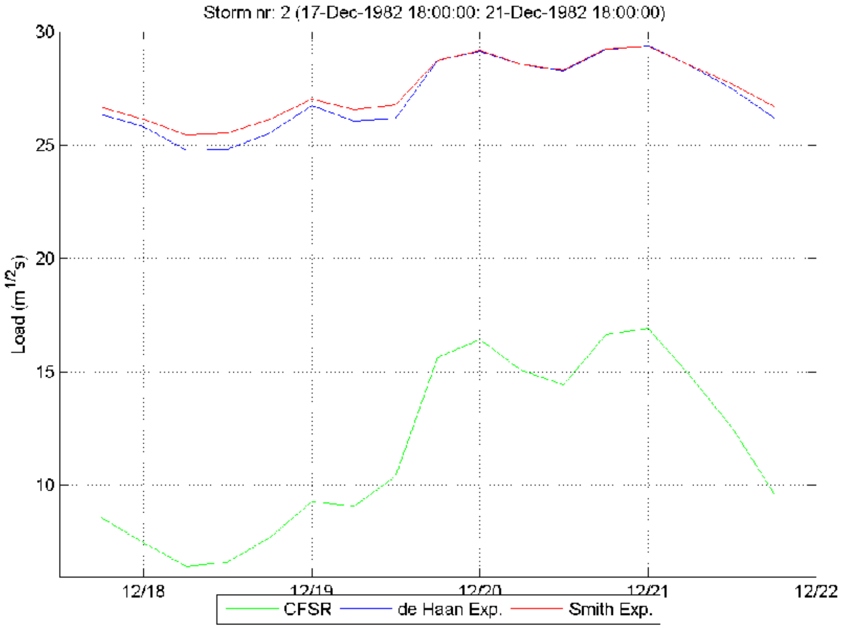
These processes are i.i.d. and their probability distribution is approximately (for large t) the same as the distribution of the process $\{V(s)\}_{s \in C}$.

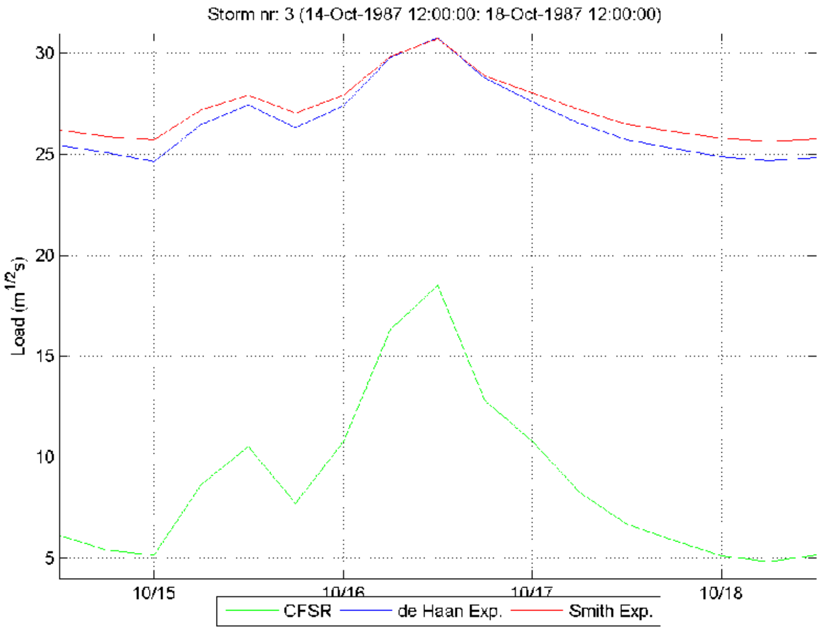
The rest of the reasoning is the same as in the finite-dimensional cases.

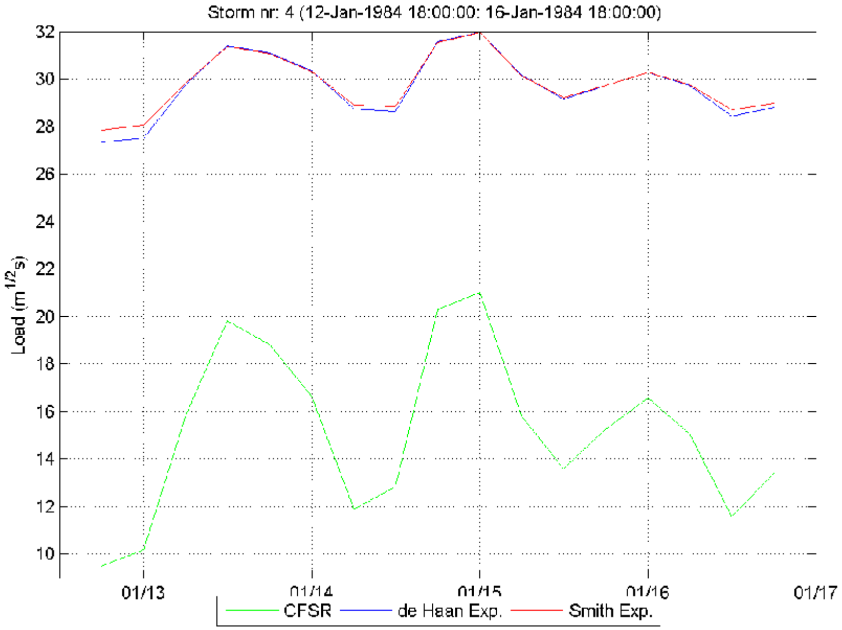
Result:

Out of k relatively extreme stochastic processes (wind storms) we have obtained k really extreme stochastic processes with the same – up scaled – (Pareto) distribution.


Storm nr: 1 (30-Jan-1983 12:00:00: 03-Feb-1983 12:00:00)
Load (m^1/2 s)
30
25
20
15
10
01/31
02/01
02/02
02/03
CFSR
de Haan Exp.
Smith Exp.


Storm nr: 2 (17-Dec-1982 18:00:00: 21-Dec-1982 18:00:00)
Load (m^1/2 s)
30
25
20
15
10
12/18
12/19
12/20
12/21
12/22
CFSR
de Haan Exp.
Smith Exp.


Storm nr: 3 (14-Oct-1987 12:00:00: 18-Oct-1987 12:00:00)
Load (m^1/2 s)
30
25
20
15
10
5
10/15
10/16
10/17
10/18
CFSR
de Haan Exp.
Smith Exp.


Storm nr: 4 (12-Jan-1984 18:00:00: 16-Jan-1984 18:00:00)
Load (m^1/2 s)
32
30
28
26
24
22
20
18
16
14
12
10
01/13
01/14
01/15
01/16
01/17
CFSR
de Haan Exp.
Smith Exp.

Some references

B. M. Brown and S.I. Resnick (1977). J. Appl. Prob. 14, 732-739.

L. de Haan (1984). Ann. Prob. 12, 1194-1204.

E. Giné *et al* (1990). Prob. Th. Related Fields 87, 139-165.

M. D. Penrose (1992). Ann. Prob. 20, 1450-1463.

S.G. Coles and J.A. Tawn (1996). J. Royal Statist. Soc. B 58, 329-347.

T. Lin *et al* (2001). Ann. Prob. 29, 467-482;
(2003) Ann. Stat. 31, 1996-2012; (2006) Ann. Stat. 34, 469-492.

M. Schlather (2002). Extremes 5, 33-44.

L. de Haan and A. Ferreira (2006). *Extreme value theory*, Springer.

The 8th Conference on Extreme Value Analysis
July 8-12, 2013
Fudan University, Shanghai, China

We are pleased to announce that the 8th Conference on Extreme Value Analysis will take place from July 8 to 12, 2013 at Fudan University, Shanghai, China.

Organizers: Deyuan Li, Liang Peng,
Zhengjun Zhang, Ming Zheng

Email: eva2013sh@yahoo.com
Website: http://eva.fudan.edu.cn

Topics:

- Univariate, multivariate, infinite dimensional extreme value theory
- Order statistics and records
- Rare events and risk analysis
- Spatial/spatio-temporal extremes
- Heavy tails in actuarial sciences
- Other related applications

History: Previous EVA conferences have been held in Leuven, Belgium (2001), Lyon, France (2011), Vimeiro, Portugal (1983), Aveiro, Portugal (2004), twice in Gothenburg, Sweden (1998 and 2005), Bern, Switzerland (2007), Fort Collins, USA (2009).