Heavy tails and extreme events

A compact tour of extreme value distributions

Christian Y. ROBERT

Université Claude Bernard Lyon 1

Generative AI for Extreme events - Sorbonne University - March 19, 2026

Roadmap

- Why extremes are special?
- Univariate EV theory and distributions
- Multivariate EV theory and distributions
- Infinite-dimensional EV theory and distributions
- Point processes for EV

Introduction and motivation

Where EVT appears in risk

Risk types

- Finance/insurance: crashes, drawdowns, large losses, ...
- Climate: natural disaster like e.g. heatwaves, floods, windstorms, ...
- Engineering: loads, fatigue, reliability, safety margins, ...

Typical needed outputs

- High quantiles, high expectiles, expected shortfalls associated to high quantiles, ...
- Return levels taking potentially into account clustering in time, return periods, ...
- Joint exceedance probabilities, likelihood of co-occurences of extreme values, joint behavior of multiple variables in their tails, ...

EVT : Analysis of rare phenomena with (very) small probabilities... and prediction!

Three difficulties

- **Data scarcity**: the tail is where we have the least observations.
- **Extrapolation need**: we want to extrapolate beyond the observed range.
- **Intricate dependence**: extremes co-occur across space and/or cluster in time.

EVT viewpoint

Provides model *limits* of suitably normalized extremes, giving *universal* families of distributions to extrapolate tail probabilities.

Why this matters for generative models

A good generator must match: **marginal tail** (index/shape) + **(potentially high-dimensional) extremal dependence** (co-movements, clustering).

Simulation of extremes is not "just more samples"

- In heavy tails, **rare outcomes dominate** risk functionals.
- Heavy tailed distribution cannot natively be generated by neural networks.

EVT as a "specification layer"

EVT provides *constraints* and *diagnostics* for generators:

- *constraints* : affine normalization, max-stability, threshold stability, homogeneity properties ,...
- *diagnostics* : tail dependence functions, extremal coefficients, extremal index,...

A formal probabilistic setup for EVT

Probability space and state space

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let

$$X : \Omega \to \mathrm{E}$$

be a random element taking values in a **Polish space** E (complete, separable metric space).

Typical choices of E

- $\mathrm{E} = \mathbb{R}$ (univariate extremes),
- $\mathrm{E} = \mathbb{R}^d$ (multivariate extremes),
- $\mathrm{E} = C([0,1], \mathbb{R})$ (functional / infinite-dimensional extremes),
- $\mathrm{E} =$ point measures spaces.

A formal probabilistic setup for EVT

Two canonical viewpoints in Extreme Value Theory

- **Case 1 (statistics of extreme observations):** given $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} F$ (or a sub-vector of a strictly stationary sequence $(X_t)_{t\in\mathbb{Z}}$), study an *extreme-based* statistic $f_n(X_1, \ldots, X_n)$ (e.g., maxima, high order statistics, exceedance counts) and its weak limit as $n \to \infty$:

$$\mathcal{L}\big(f_n(X_1, \ldots, X_n)\big) \;\Rightarrow\; \mathcal{L}_\infty.$$

- **Case 2 (rare-event conditioning):** study the conditional law of X given a *rare event*, $A_u \subset E$ with $\mathbb{P}(X \in A_u) \to 0$ as $u \to \infty$ (e.g., threshold exceedances, generalized Pareto limits, and "typical shapes" of extreme episodes) and its weak limit:

$$\mathcal{L}\big(f_u(X) \mid X \in A_u\big) \;\Rightarrow\; \mathcal{L}_\infty.$$

Univariate extreme value theory and distributions

Let $X_1, \ldots, X_n$ be i.i.d. with cdf F on $\mathbb{R}$.

Two core "extreme" objects

- **Largest values:**
  - **Maxima:** $M_n = \max\{X_1, \ldots, X_n\}$.
  - **Record values**
  - **Upper order statistics:** $X_{(k)}$ is the k-th largest value among $X_1, \ldots, X_n$ (equivalently $X_{(1)} = M_n$).
  - **Exceedance count above** u**:** $N_n(u) = \sum_{i=1}^{n} \mathbf{1}\{X_i > u\}$.
- **High exceedances:** $X - u \mid X > u$ for a high threshold u.

Univariate max-stable distributions

$(M_n)_{n\geq 1}$ is an increasing sequence that converges almost surely to the extremal point of F: $x_F = \sup\{x : F(x) < 1\}$. A non-degenerate limit for M_n can therefore only arise after **normalization**.

Affine normalisation

Find sequences $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\frac{M_n - b_n}{a_n} \Rightarrow Z \quad (n \to \infty),$$

where Z has a non-degenerate distribution.

Contrast with the CLT

For maxima, the limit family is *not* Gaussian and the choice of (a_n, b_n) is not canonical: it depends on the tail of F.

Fisher–Tippett–Gnedenko theorem (FTG) - GEV distributions

Theorem (FTG)

If there exist $a_n > 0, b_n \in \mathbb{R}$ such that $\frac{M_n - b_n}{a_n} \Rightarrow Z$, then Z must follow a **Generalized Extreme Value (GEV)** distribution (up to location and scale parameters):

$$G_\xi(x) = \exp\left\{ -\left(1 + \xi x\right)^{-1/\xi} \right\}, \qquad 1 + \xi x > 0,$$

with the continuous limit $\xi \to 0$ giving the Gumbel form $G_0(x) = \exp\{-e^{-x}\}$.

Interpretation

EVT yields a **universal parametric limit family**: location/scale + **shape** ξ.

GEV shape parameter ξ

- **Fréchet** ($\xi > 0$): heavy tails, power-law decay (no finite upper endpoint).
- **Gumbel** ($\xi = 0$): light/medium tails (Gaussian, lognormal, exponential, ...).
- **Weibull** ($\xi < 0$): finite upper endpoint.

Risk takeaway

The sign and magnitude of ξ drives **extrapolation**: small changes in ξ can produce large changes in far-tail quantiles.

Domain of attraction: what does it mean?

Definition

We say F belongs to the **max-domain of attraction** of G_ξ, written $F \in \mathcal{D}(G_\xi)$, if there exist $a_n > 0, b_n$ such that

$$\mathbb{P}(M_n \leq a_n x + b_n) = F^n(a_n x + b_n) \longrightarrow G_\xi(x).$$

Key question

Given F, how can we:

- identify which ξ applies,
- choose (a_n, b_n),
- quantify the rate of convergence?

Max-stability

A distribution G is max-stable, if for every $m \geq 1$, there exist $a_m > 0, b_m$ such that

$$G^m(a_m x + b_m) = G(x).$$

Fact

GEV is the only max-stable family on $\mathbb{R}$ (up to affine changes). Therefore, any non-degenerate limit of normalized maxima must be GEV.

A central heavy-tail condition is **regular variation**:

$$\bar{F}(x) = 1 - F(x) = x^{-\alpha}L(x), \qquad x \to \infty,$$

where L is slowly varying and $\alpha > 0$ is the **tail index**.

Connection to GEV

Regular variation $\iff$ Fréchet domain with

$$\xi = \frac{1}{\alpha}.$$

Moment thresholds

If $\alpha < 1$ the mean is infinite; if $\alpha < 2$ the variance is infinite.

From maxima to exceedances: rare-event viewpoint

Fix a high level u and consider exceedances $X_i > u$.

Rare-event count: the law of small numbers

Let $N_n(u) = \sum_{i=1}^{n} \mathbf{1}\{X_i > u\}$. If $u = u_n$ is chosen so that $n\,\bar{F}(u_n) \to \tau > 0$, then

$$N_n(u_n) \Rightarrow \mathrm{Poisson}(\tau),$$

i.e. exceedances behave like **rare events**.

Interpretation

This is the gateway to the **point process** approach.

Key equivalence: order statistic $\leftrightarrow$ exceedance count

For any threshold u,

$$X_{(k)} \le u \iff \#\{i : X_i > u\} \le k-1 \iff N_n(u) := \sum_{i=1}^{n} \mathbf{1}\{X_i > u\} < k.$$

So the distribution of $X_{(k)}$ is governed by the distribution of the exceedance count $N_n(u)$.

GEV normalisation as a special case

If $u_n = a_n x + b_n$ is such that $[F(a_n x + b_n)]^n \to G(x)$ (max-stable distribution), then

$$\tau = -\log G(x), \qquad \mathbb{P}\left(\frac{X_{(k)} - b_n}{a_n} \le x\right) \to G(x) \sum_{i=0}^{k-1} \frac{\left(-\log G(x)\right)^i}{i!}.$$

How fast do maxima/order statistics converge?

Setup: GEV approximation and its error

Let G be a max-stable distribution. Assume $F \in \mathcal{D}(G)$ and choose $a_n > 0, b_n$ such that

$$F^n(a_n x + b_n) \;\to\; G(x).$$

What is the **rate** at which the approximation becomes accurate, e.g.

$$\Delta_n := \sup_{x \in \mathbb{R}} \left| F^n(a_n x + b_n) - G(x) \right| ?$$

Concrete intuition (heavy tails)

If $\bar{F}(x) = x^{-\alpha} L(x)$ is **regularly varying** (Fréchet case) and L satisfies a **second-order regular variation** with index $\rho < 0$, then there exists $a_n > 0$ such that

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left(\frac{M_n}{a_n} \le x \right) - \exp(-x^{-\alpha}) \right| \;=\; O(n^{\rho/\alpha}).$$

Record values distributions

Record values: definition and notation

Let F be a **continuous** cdf (continuity avoids ties).

Record event and record times

X_j is a **record value** if it exceeds all previous observations:

$$X_j > \max\{X_1, \ldots, X_{j-1}\}.$$

Define the **record times** $(L_n)_{n \geq 0}$ recursively by

$$L_0 = 1, \qquad L_n = \min\{j > L_{n-1} : X_j > X_{L_{n-1}}\} \quad (n \geq 1).$$

Record values sequence

The ***n*-th record value** is $R_n := X_{L_n}$, $n \geq 0$, so that $R_0 < X_{L_1} < \cdots$ almost surely.

Limit laws for record values (FTG analogue)

EVT question (record analogue of maxima)

Do there exist normalizing constants $a_n > 0$, b_n such that

$$\mathbb{P}\left(\frac{R_n - b_n}{a_n} \leq x\right) \longrightarrow H(x) \quad \text{with a non-degenerate limit law } H?$$

"FTG for records"

If there exist $a_n > 0$, b_n such that the convergence holds, then the limit must be of the form

$$H(x) = \Phi\Big(-\log\big(-\log G(x)\big)\Big),$$

where Φ is the standard normal cdf and G is an **max-stable distribution** (i.e. one of the FTG limits for maxima).

Three possible limit types (explicit)

Equivalently, up to location/scale changes, H must be of one of these forms:

$$\text{(i)} \quad H_1(x) = \Phi(x),$$

$$\text{(ii)} \quad H_{2,\alpha}(x) = \begin{cases} 0, & x \le 0, \\ \Phi(\alpha \log x), & x > 0, \end{cases} \qquad (\alpha > 0),$$

$$\text{(iii)} \quad H_{3,\alpha}(x) = \begin{cases} \Phi\big(\alpha \log(-x)^{-1}\big), & x < 0, \\ 1, & x \ge 0, \end{cases} \qquad (\alpha > 0).$$

Takeaway

Maxima: normalized $M_n \Rightarrow$ GEV G (FTG). **Records:** normalized $R_n \Rightarrow$ *Gaussian transform* of the same GEV class: $\Phi\big(-\log(-\log G)\big)$.

Generalized Pareto Distributions (GPD)

POT = Peaks Over Threshold

Conditional excesses and scaling - Generalized Pareto distributions (GPD)

Assume that $F \in \mathcal{D}(G_\xi)$, then there exists a scaling function $a(u) > 0$ such that, as $u \to \infty$,

$$\mathcal{L}\left(\frac{X-u}{a(u)} \,\middle|\, X > u\right) \Rightarrow \mathrm{GPD}_\xi.$$

Equivalently, for all $y \geq 0$ with $1 + \xi y > 0$,

$$\mathbb{P}\left(\frac{X-u}{a(u)} \leq y \,\middle|\, X > u\right) \longrightarrow G_\xi(y) := 1 - \left(1 + \xi y\right)^{-1/\xi}.$$

Define the conditional excess distribution above u:

$$F_u(y) = \mathbb{P}(X - u \le y \mid X > u), \qquad y \ge 0.$$

Theorem (PBdH)

If $F \in \mathcal{D}(G_\xi)$, then as $u \uparrow x_F$,

$$F_u(y) \approx \mathrm{GPD}_{\beta(u),\xi}(y) = 1 - \left(1 + \xi y/\beta(u)\right)^{-1/\xi}, \qquad 1 + \xi y/\beta(u) > 0.$$

Key message

GPD is the universal model for threshold exceedances.

Threshold stability

Let $Y \sim \mathrm{GPD}(\beta, \xi)$ with scale paremeter $\beta > 0$. For any $v \geq 0$ such that $1 + \xi v/\beta > 0$,

$$Y - v \,\big|\, Y > v \;\sim\; \mathrm{GPD}\big(\,\beta + \xi v, \xi\big).$$

Fact

GPD is the only threshold-stable family on $\mathbb{R}_+$ (up to scale changes). Therefore, any non-degenerate limit model for properly normalized threshold exceedances must be GPD, with a **shape parameter** ξ **that stays invariant** when the threshold increases.

Same shape parameter

The **same** ξ governs:

- maxima (GEV),
- upper order statistics,
- record values ($\Phi-$ GEV),
- exceedances (GPD),...

Multivariate extreme value theory and distributions

Why multivariate extremes?

- Catastrophes are **multi-dimensional**: wind + surge, rainfall + soil saturation, correlated defaults.
- Tail **dependence** controls joint risk (portfolio, spatial footprints).
- Classical correlation is **misleading** in tails.

Goal

Extend univariate statistics to multivariate ones (e.g. componentwise maxima) and derive limits taking into account the geometry of joint tails.

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be i.i.d. random vectors in $\mathbb{R}^d$ with joint cdf F.

Two core "extreme" objects

- **Largest values of a sample:**
  - **Componentwise maximum vector:** $\boldsymbol{M}_n = \big(\max_{1\le i\le n} X_{i1}, \ldots, \max_{1\le i\le n} X_{id}\big)$.
  - **Upper order statistics (radial):** choose a risk functional/norm $r(\boldsymbol{x})$ (e.g. $\|\boldsymbol{x}\|$, $x_1 + \cdots + x_d$, $\max_j x_j$), and study the largest $r(\boldsymbol{X}_i)$ and associated angles $\boldsymbol{X}_i / r(\boldsymbol{X}_i)$.
  - **Exceedance count of a tail region A:** $N_n(A) = \sum_{i=1}^n \mathbf{1}\{\boldsymbol{X}_i \in A\}$, for regions A capturing "large" multivariate events.
- **High exceedances:** conditional laws such as

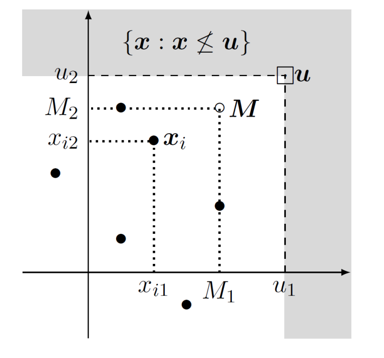
$$\mathcal{L}(\boldsymbol{X} - u\mathbf{1} \mid \boldsymbol{X} > u\mathbf{1}), \quad \mathcal{L}(\boldsymbol{X} \mid r(\boldsymbol{X}) > u), \quad \mathcal{L}(\boldsymbol{X} \mid \boldsymbol{X} \in A_u),$$

for high u (componentwise or radial/region-based).

Multivariate exceedances: "large" sets and competing notions

Several ways to define "extreme" in $\mathbb{R}^d$

- **Componentwise:** $\boldsymbol{X} > \boldsymbol{u}$ (all components exceed thresholds).
- **At least one large component:** $\max_j X_j > u$ (captures union-of-tails events).
- **Radial/region exceedance:** $r(\boldsymbol{X}) > u$ or $\boldsymbol{X} \in A_u$ (captures large magnitude under a chosen risk functional).



Example of tail regions in $\mathbb{R}^2$: componentwise thresholds (u_1, u_2), maxima (M_1, M_2), and a generic extreme set $\{\boldsymbol{x} : \boldsymbol{x} \not\leq \boldsymbol{u}\}$. Naveau and Segers (2026).

Asymptotic dependence vs independence

For two variables, the tail dependence coefficient

$$\chi = \lim_{u \uparrow 1} \mathbb{P}\big(F_1(X_1) > u \mid F_2(X_2) > u\big)$$

distinguishes:

- **Asymptotic extremal dependence**: $\chi > 0$ (joint extremes persist).
- **Asymptotic extremal independence**: $\chi = 0$ (joint extremes vanish faster).

Modeling consequence

Max-stable limits, Generalized Pareto limits naturally represent asymptotic dependence; asymptotic independence requires extensions (mutual asymptotic independence, k-wise asymptotic independence,...).

Multivariate Max-stable Distributions

Max-domain of attraction (MDA) in $\mathbb{R}^d$

Componentwise maxima and normalization

Define the componentwise maxima $\boldsymbol{M}_n := \max_{i\in[n]} \boldsymbol{X}_i$ and consider normalized maxima (componentwise operations)

$$\frac{\boldsymbol{M}_n - \boldsymbol{b}_n}{\boldsymbol{a}_n}, \qquad \boldsymbol{a}_n > 0,\ \boldsymbol{b}_n \in \mathbb{R}^d.$$

Definition (Max-domain of attraction)

We say that F is in the max-domain of attraction of a cdf G (write $F \in \mathcal{D}(G)$) if there exist $\boldsymbol{a}_n > 0, \boldsymbol{b}_n$ such that

$$F^n(\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n) \xrightarrow[n\to\infty]{} G(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in \mathbb{R}^d.$$

Max-stable distributions and marginal standardization

Definition (Max-stable cdf)

A cdf G on $\mathbb{R}^d$ is **max-stable** if for every $m \in \mathbb{N}$ there exist $\boldsymbol{a}_m > 0, \boldsymbol{b}_m$ such that

$$G^m(\boldsymbol{a}_m \boldsymbol{x} + \boldsymbol{b}_m) = G(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in \mathbb{R}^d.$$

Key fact

If $F \in \mathcal{D}(G)$, then the limit G is necessarily **max-stable**.

Max-stable distributions and marginal standardization

Margins and the “simple” max-stable version

Each marginal G_j is univariate GEV (Fréchet/Gumbel/Weibull). One can standardize margins to **unit Fréchet** and obtain an associated **simple max-stable** cdf $G^\star$ whose margins satisfy

$$G_j^\star(x) = \exp(-1/x), \qquad x > 0.$$

After this step, **all tail dependence is in $G^\star$**.

Characterizations of the dependence of $G^\star$

Five characterizations

- Exponent function, exponent measure and its polar decomposition
- Pickand's representation
- The stable tail dependence function
- The *D*-norm characterization
- The spectral representation

A first characterization: exponent function and exponent measure

Let $\boldsymbol{Z} = (Z_1, \ldots, Z_d)$ be **simple max-stable** with unit Fréchet margins $\mathbb{P}(Z_j \leq z) = \exp(-1/z)$ for $z > 0$.

Exponent function (a complete description of dependence)

The joint cdf can be written as

$$\mathbb{P}(Z_1 \leq z_1, \ldots, Z_d \leq z_d) = \exp\{-V(z_1, \ldots, z_d)\}, \qquad \boldsymbol{z} \in (0, \infty)^d,$$

where $V : (0, \infty)^d \to (0, \infty)$ is **homogeneous of order** -1:

$$V(t\boldsymbol{z}) = t^{-1} V(\boldsymbol{z}), \qquad t > 0,$$

and $V(\boldsymbol{e}_i) = 1$ with $(\boldsymbol{e}_i)_{i=1,\ldots,d}$ the standard basis of $\mathbb{R}^d$.

Extremal coefficient (quick dependence summary)

Define $\theta = V(1, \ldots, 1) \in [1, d]$ such that

$$\mathbb{P}(Z_1 \leq z, \ldots, Z_d \leq z) = \exp(-\theta/z).$$

Then $\theta = 1$ (complete dependence) and $\theta = d$ (independence).

Exponent measure viewpoint

Equivalently, there exists a Radon measure ν on $(0, \infty)^d$ such that

$$V(\boldsymbol{z}) = \nu\Big(\{\boldsymbol{x} \in (0, \infty)^d : \exists j,\ x_j > z_j\}\Big),$$

and ν is homogeneous: $\nu(tA) = t^{-1}\nu(A)$.

Polar (radial–angular) decomposition of the exponent measure

Polar transform: radius + angle

Choose a norm $\|\cdot\|$ on $\mathbb{R}^d$ and define, for $\boldsymbol{x} \neq \mathbf{0}$,

$$r = \|\boldsymbol{x}\| \in (0, \infty), \qquad \boldsymbol{w} = \boldsymbol{x}/\|\boldsymbol{x}\| \in \mathbb{S}_+^{d-1} := \{\boldsymbol{w} \geq 0 : \|\boldsymbol{w}\| = 1\}.$$

Decomposition into radial part and spectral measure

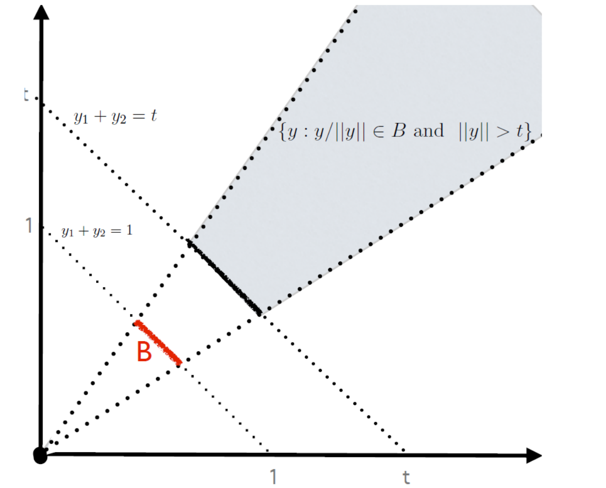
There exists a finite measure H on $\mathbb{S}_+^{d-1}$ (the **spectral measure**)

$$\nu(\mathrm{d}\boldsymbol{x}) = r^{-2}\,\mathrm{d}r\,H(\mathrm{d}\boldsymbol{w}), \qquad \boldsymbol{x} = r\boldsymbol{w},\ \ r > 0,\ \boldsymbol{w} \in \mathbb{S}_+^{d-1}.$$

Equivalently, for measurable $B \subset \mathbb{S}_+^{d-1}$ and $t > 0$,

$$\nu\big(\{\boldsymbol{x} : \|\boldsymbol{x}\| > t,\ \boldsymbol{x}/\|\boldsymbol{x}\| \in B\}\big) = t^{-1}H(B).$$

Polar (radial–angular) decomposition of the exponent measure



Naveau (2015)

Unit simplex and spectral measure

Consider the unit simplex ("L^1-sphere")

$$\Delta_{d-1} := \left\{ \boldsymbol{w} \in [0,1]^d : \sum_{j=1}^{d} w_j = 1 \right\}.$$

Pickands' characterization states that there exists a finite measure S on Δ_{d-1} (the **spectral measure**) such that the exponent function can be written as

$$V(\boldsymbol{z}) = \int_{\Delta_{d-1}} \max_{1 \le j \le d} \left(\frac{w_j}{z_j} \right) S(\mathrm{d}\boldsymbol{w}), \qquad \boldsymbol{z} \in (0, \infty)^d,$$

and $\int_{\Delta_{d-1}} w_j \, S(\mathrm{d}\boldsymbol{w}) = 1$, $j = 1, \ldots, d$.

How S encodes extremal dependence

- Mass near vertices e_j (one coordinate close to 1) $\Rightarrow$ extremes tend to occur **in one component at a time** (weak extremal dependence).
- Mass near the center $(1/d, \ldots, 1/d)$ $\Rightarrow$ extremes tend to be **balanced across components** (strong extremal dependence).
- Independence corresponds to S concentrated on the vertices (each with mass 1); complete dependence corresponds to S concentrated at $(1/d, \ldots, 1/d)$ (mass d).

Takeaway

Pickands' spectral measure S on Δ_{d-1} provides a concrete, geometric parameterization of all simple max-stable laws in $\mathbb{R}^d$.

Stable tail dependence function (stdf) and $G^\star$

stdf as a tail-dependence object

For a multivariate cdf F in a max-domain of attraction, the **stable tail dependence function** $\ell_F : (0,\infty)^d \to (0,\infty)$ is defined by the limit

$$\ell_F(\boldsymbol{x}) = \lim_{t\to\infty} t\,\mathbb{P}\Big(1 - F_1(X_1) \le x_1/t \text{ or } \cdots \text{ or } 1 - F_d(X_d) \le x_d/t\Big),$$

where F_j are the marginal cdfs.

Link between ℓ and the simple max-stable cdf

If G is max-stable and $G^\star$ is its associated simple max-stable version, then

$$\ell_G(\boldsymbol{x}) = -\log G^\star(1/\boldsymbol{x}) = V(1/\boldsymbol{x}), \qquad \boldsymbol{x} \in (0,\infty)^d.$$

Characterization via D-norms

Definition (D-norm)

Let $\mathbf{\Gamma} = (\Gamma_1, \ldots, \Gamma_d)$ be a nonnegative random vector such that $\mathbb{E}[\Gamma_j] = 1$ for all j. Then

$$\|\boldsymbol{x}\|_D := \mathbb{E}\left[\max_{j=1,\ldots,d}(|x_j|\Gamma_j)\right]$$

defines a norm on $\mathbb{R}^d$, called a **D-norm**, and $\mathbf{\Gamma}$ is a **generator**.

Theorem (simple max-stable $\Longleftrightarrow$ D-norm)

A cdf $G^\star$ on $\mathbb{R}^d$ is **simple max-stable** iff there exists a D-norm $\|\cdot\|_D$ such that for all $\boldsymbol{x} \in (0, \infty)^d$,

$$G^\star(\boldsymbol{x}) = \exp\left\{-\left\|\frac{1}{\boldsymbol{x}}\right\|_D\right\}.$$

In particular, for a max-stable G, $\ell_G(\boldsymbol{x}) = \|\boldsymbol{x}\|_D$.

A last caracterisation: the Spectral representations

For unit Fréchet max-stable $\boldsymbol{Z}$, there exists a spectral random vector $\boldsymbol{W}$ with $\mathbb{E}[W_j] = 1$ such that

$$\boldsymbol{Z} \stackrel{d}{=} \max_{k \geq 1} \zeta_k \boldsymbol{W}_k,$$

where (ζ_k) are points of a Poisson process on $(0, \infty)$ with intensity $\zeta^{-2} d\zeta$ and $(\boldsymbol{W}_k)$ are i.i.d. copies of $\boldsymbol{W}$.

Interpretation

Extremes arise from **random "storms"** (shapes $\boldsymbol{W}_k$) with random **magnitudes** (ζ_k^{-1}).

Finite-dimensional distributions based on the spectral representation)

Exponent function from the spectral random vector

For $\boldsymbol{z} = (z_1, \ldots, z_d) \in (0, \infty)^d$,

$$\mathbb{P}(Z_1 \leq z_1, \ldots, Z_d \leq z_d) = \exp\{-V(\boldsymbol{z})\},$$

with exponent function

$$V(\boldsymbol{z}) = \mathbb{E}\left[\max_{j=1,\ldots,d} \frac{W_j}{z_j}\right].$$

Setting $z_1 = \cdots = z_d = z$ gives

$$\mathbb{P}(Z_1 \leq z, \ldots, Z_d \leq z) = \exp\left(-\frac{\theta}{z}\right), \qquad \theta = \mathbb{E}\left[\max_{j=1,\ldots,d} W_j\right] \in [1, d].$$

$\theta = 1$ corresponds to complete dependence and $\theta = d$ to independence.

Multivariate Generalized Pareto Distributions (MGP)

Multivariate excesses over thresholds: "L" or "R" regions

For $d = 1$:

$$X \not\leq u \iff X > u.$$

For $d \geq 2$: different types of excesses over a threshold vector

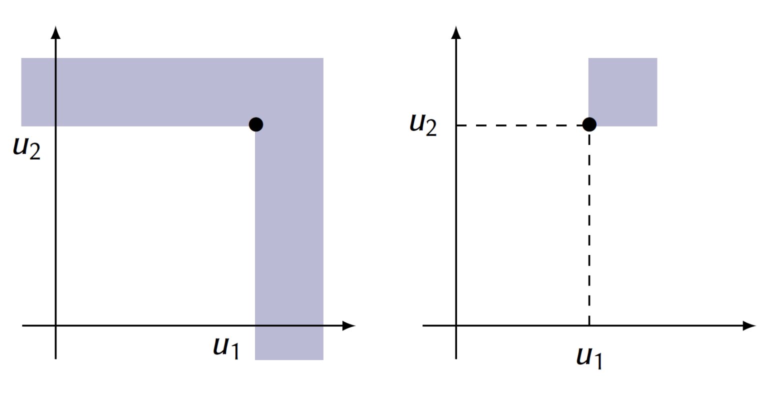
Let $\boldsymbol{u} = (u_1, \ldots, u_d)$. Two common notions are:

$$\boldsymbol{X} \not\leq \boldsymbol{u} \iff (X_j > u_j \text{ for some } j = 1, \ldots, d) \qquad (\text{"L"})$$

$$\boldsymbol{X} > \boldsymbol{u} \iff (X_j > u_j \text{ for all } j = 1, \ldots, d) \qquad (\text{"R"})$$

Events of the form $\ell(\boldsymbol{X}) > u$ for a loss functional $\ell(\boldsymbol{x}) = \max(\boldsymbol{x})$ or $\min(\boldsymbol{x})$.

Multivariate excesses over thresholds: "L" or "R" regions



Segers (2024)

MGP as a tail model (L-shaped exceedances)

Two-step GP approximation (multivariate POT)

Let $\boldsymbol{X} \sim F$ and let $\boldsymbol{u} = (u_1, \ldots, u_d)$ be a high threshold vector. For a failure set $B \subset \mathbb{R}^d$ contained in the **L-region** $L_{\boldsymbol{u}} := \{\boldsymbol{x} : \boldsymbol{x} \not\leq \boldsymbol{u}\}$,

$$\mathbb{P}(\boldsymbol{X} \in B) = \mathbb{P}(\boldsymbol{X} \not\leq \boldsymbol{u})\, \mathbb{P}(\boldsymbol{X} - \boldsymbol{u} \in B - \boldsymbol{u} \mid \boldsymbol{X} \not\leq \boldsymbol{u}).$$

Idea

Model the conditional excess distribution $\mathcal{L}(\boldsymbol{X} - \boldsymbol{u} \mid \boldsymbol{X} \not\leq \boldsymbol{u})$ by a **multivariate generalized Pareto** (MGP) distribution.

Standard MGP distribution

A random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_d) \in [-\infty, \infty)^d$ follows a **standard** multivariate generalized Pareto distribution if:

1. $E := \max(Z_1, \ldots, Z_d)$ is **unit-exponential**: $\mathbb{P}(E > t) = e^{-t}$, $t \geq 0$.
2. The non-positive vector $\boldsymbol{S} := \boldsymbol{Z} - E$ is **independent** of E, where

$$\boldsymbol{S} = \boldsymbol{Z} - E = (Z_1 - E, \ldots, Z_d - E) \in [-\infty, 0]^d,$$

   with $\mathbb{P}(S_j > -\infty) > 0$ for all $j = 1, \ldots, d$.

We write $\boldsymbol{Z} \sim \mathrm{MGP}(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$ and the representation is $\boldsymbol{Z} = E + \boldsymbol{S}$.

Interpretation

E is the **common shock** (overall severity), $\boldsymbol{S}$ describes **deviations from the maximum**.

Definition: *general* MGP and link to univariate GP margins

General MGP via marginal GP transforms

Let $\boldsymbol{Z} \sim \mathrm{MGP}(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$ be standard. For shape $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$ and scale $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_d) \in (0, \infty)^d$, define $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ by the componentwise transform

$$Y_j = \frac{\sigma_j}{\xi_j}\left(e^{\xi_j Z_j} - 1\right), \qquad \text{(with the limit } Y_j = \sigma_j Z_j \text{ if } \xi_j = 0\text{).}$$

Then $\boldsymbol{Y}$ is said to follow $\mathrm{MGP}(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{S})$.

Conditional margins are univariate GP

For each j, the conditional margin is GP:

$$(Y_j \mid Y_j > 0) \ \sim\ \mathrm{GP}(\sigma_j, \xi_j),$$

and Y_j and Z_j have the same sign.

Model

Let $\boldsymbol{X} = E + \boldsymbol{U} = (E + U_1, \ldots, E + U_d)$, where $E \sim \mathrm{Exp}(1)$ is independent of $\boldsymbol{U}$ and $0 < \mathbb{E}(e^{U_j}) < \infty$ for all j.

Proposition (high-threshold limit)

High-threshold excesses over the **L-event** $\{\max(\boldsymbol{X}) > u\}$ satisfy

$$\lim_{u \to \infty} \mathbb{P}(\boldsymbol{X} - u\boldsymbol{1} \leq \boldsymbol{x} \mid \max(\boldsymbol{X}) > u) = \mathbb{P}(\boldsymbol{Z} \leq \boldsymbol{x}),$$

where $\boldsymbol{Z} \sim \mathrm{MGP}(\boldsymbol{1}, \boldsymbol{0}, \boldsymbol{S})$ and $\boldsymbol{S} = \boldsymbol{T} - \max(\boldsymbol{T})$ with

$$\mathbb{P}(\boldsymbol{T} \in \mathrm{d}\boldsymbol{x}) = \frac{\mathbb{E}\left(e^{\max(\boldsymbol{U})}\boldsymbol{1}\{\boldsymbol{U} \in \mathrm{d}\boldsymbol{x}\}\right)}{\mathbb{E}\left(e^{\max(\boldsymbol{U})}\right)}.$$

A single exponential shock E plus “noise” $\boldsymbol{U}$ yields MGP limits for multivariate POT.

Why threshold stability?

If $\boldsymbol{X} - \boldsymbol{u} \mid \boldsymbol{X} \not\leq \boldsymbol{u}$ is MGP, then for higher thresholds $\boldsymbol{v} \geq \boldsymbol{u}$, the model should remain **internally consistent**: $\boldsymbol{X} - \boldsymbol{v} \mid \boldsymbol{X} \not\leq \boldsymbol{v}$ should still be MGP.

Proposition (standard case)

Let $\boldsymbol{Z} \sim \mathrm{MGP}(\boldsymbol{1}, \boldsymbol{0}, \boldsymbol{S})$ and $\boldsymbol{u} \in [0, \infty)^d$. Then

$$\boldsymbol{Z} - \boldsymbol{u} \mid \boldsymbol{Z} \not\leq \boldsymbol{u} \;\sim\; \mathrm{MGP}(\boldsymbol{1}, \boldsymbol{0}, \boldsymbol{S}_{\boldsymbol{u}}),$$

where $\boldsymbol{S}_{\boldsymbol{u}} = \boldsymbol{T}_{\boldsymbol{u}} - \max \boldsymbol{T}_{\boldsymbol{u}}$ and

$$\mathbb{P}(\boldsymbol{T}_{\boldsymbol{u}} \in \mathrm{d}\boldsymbol{x}) = \frac{\mathbb{E}\left(e^{\max(\boldsymbol{S}-\boldsymbol{u})} \mathbb{1}\{\boldsymbol{S} - \boldsymbol{u} \in \mathrm{d}\boldsymbol{x}\}\right)}{\mathbb{E}\left(e^{\max(\boldsymbol{S}-\boldsymbol{u})}\right)}.$$

Special case: if $u_1 = \cdots = u_d = u$, then $\boldsymbol{S}_{\boldsymbol{u}}$ has the same distribution as $\boldsymbol{S}$.

Stability properties: threshold stability (general case)

Proposition (general case)

Let $\boldsymbol{Y} \sim \mathrm{MGP}(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{S})$ and let $\boldsymbol{v} \in [0, \infty)^d$ such that $\sigma_j + \xi_j v_j > 0$ for all j. Then

$$\boldsymbol{Y} - \boldsymbol{v} \mid \boldsymbol{Y} \not\leq \boldsymbol{v} \ \sim \ \mathrm{MGP}(\boldsymbol{\sigma} + \boldsymbol{\xi}\boldsymbol{v},\ \boldsymbol{\xi},\ \boldsymbol{S}_{\boldsymbol{u}}),$$

where $\boldsymbol{u}$ is the corresponding standard-scale threshold

$$u_j = \xi_j^{-1} \log\left(1 + \xi_j v_j / \sigma_j\right),$$

(with the limit $u_j = v_j/\sigma_j$ if $\xi_j = 0$), and $\boldsymbol{S}_{\boldsymbol{u}}$ is the updated deviation.

Univariate analogy (reminder)

$Y \sim \mathrm{GP}(\sigma, \xi) \ \Rightarrow \ Y - v \mid Y > v \sim \mathrm{GP}(\sigma + \xi v, \xi)$. Multivariate MGP extends this with an **updated dependence** through $\boldsymbol{S}_{\boldsymbol{u}}$.

Caveat: margins are not automatically MGP

For $d \geq 2$, lower-dimensional margins of an MGP distribution are **not** generally MGP.

Proposition (Sub-vectors)

Let $\boldsymbol{Y} \sim \mathrm{MGP}(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{S})$ on the L-shaped support $\{\boldsymbol{y} : \max_j y_j > 0\}$, and let $J \subset \{1, \ldots, d\}$ be nonempty. Then, conditioning on at least one positive component in the sub-vector,

$$\boldsymbol{Y}_J \,\big|\, \boldsymbol{Y}_J \not\leq \mathbf{0} \;\sim\; \mathrm{MGP}\big(\boldsymbol{\sigma}_J, \boldsymbol{\xi}_J, \boldsymbol{S}^{(J)}\big),$$

where $\boldsymbol{S}^{(J)}$ is an **updated dependence** object.

Practical message

When working with subsets of variables, either re-fit an MGP model on the subset, or propagate the dependence through the generator.

Proposition (Linear transformations; nonnegative mixing)

Assume $\boldsymbol{Y} \sim \mathrm{MGP}(\boldsymbol{\sigma}, \xi, \boldsymbol{S})$ and that the **shape is common across margins**, i.e. $\xi_1 = \cdots = \xi_d = \xi$. Let $A \in [0, \infty)^{m \times d}$ be such that, for each row i,

$$\mathbb{P}\left(\sum_{j=1}^{d} a_{ij} Y_j > 0\right) > 0.$$

Then the transformed vector is again MGP after conditioning on being in the L-region:

$$A\boldsymbol{Y} \,\big|\, A\boldsymbol{Y} \not\leq \boldsymbol{0} \;\sim\; \mathrm{MGP}\big(A\boldsymbol{\sigma},\ \xi\boldsymbol{1},\ \boldsymbol{S}_{\boldsymbol{\sigma},\xi,A}\big),$$

where $\boldsymbol{S}_{\boldsymbol{\sigma},\xi,A}$ is obtained by updating the dependence generator in a way that depends on $(\boldsymbol{S}, \boldsymbol{\sigma}, \xi, A)$.

Infinite dimensional extreme value theory and distributions

From vectors to functions

Many applications are linked to **spatial/temporal fields**:

$$\{X(s) : s \in \mathcal{S}\}, \qquad \mathcal{S} \subset \mathbb{R}^k,$$

with extremes occurring over *regions* (storm footprints) and *time windows*. We often care about functionals such as

$$\mathbb{P}\left(\sup_{s\in\mathcal{S}} X(s) > z\right) \quad \text{or} \quad \mathbb{P}\left(\int_{\mathcal{S}} X(s)\, ds > z_{\mathrm{crit}}\right),$$

which require both marginal tail behavior and **spatial extremal dependence**.

Questions

- What is the limit of pointwise maxima of random fields?
- What is the limit of (functional) exceedances of random fields?
- How to quantify extremal dependence over space and time?

Infinite-dimensional extremes: objects of interest

Let $X_1, \ldots, X_n$ be i.i.d. random elements in a function space E (e.g. $\mathrm{E} = C([0,1], \mathbb{R})$) with common law F.

Two core "extreme" objects (functional / spatial setting)

- **Largest values:**
  - **Pointwise maximum process:** $M_n(t) = \max_{1 \le i \le n} X_i(t)$, $t \in \mathcal{S}$.
  - **Extreme order statistics via a risk functional:** choose $r : \mathrm{E} \to \mathbb{R}$ (e.g. $r(x) = \|x\|_\infty = \sup_t x(t)$), then study the largest $r(X_i)$ together with the associated *shapes* $X_i / r(X_i)$ (normalized excursions/footprints).
  - **Exceedance counts of high-level sets:** $N_n(A) = \sum_{i=1}^n \mathbf{1}\{X_i \in A\}$, where $A \subset \mathrm{E}$ describes "large" functions (e.g. $\{\|x\|_\infty > u\}$).
- **High exceedances:** conditional laws such as $\mathcal{L}\big(f_u(X) \,\big|\, \|X\|_\infty > u\big)$, $\mathcal{L}\big(f_u(X) \,\big|\, X \in A_u\big)$,...

Max-stable processes

Definition (max-stable process)

Let $Z_1, Z_2, \dots$ be i.i.d. copies of a process Z on $\mathcal{X}$. If for every $n \geq 1$ there exist normalizing functions $a_n(x) > 0$, $b_n(x) \in \mathbb{R}$ such that

$$\max_{i=1,\dots,n} \frac{Z_i(x) - b_n(x)}{a_n(x)} \overset{d}{=} Z(x), \qquad x \in \mathcal{X},$$

then Z is **max-stable**.

Theorem (de Haan, 1984)

If $Y_1, Y_2, \dots$ are i.i.d. copies of Y and there exist $c_n(x) > 0, d_n(x)$ such that

$$\max_{i=1,\dots,n} \frac{Y_i(x) - d_n(x)}{c_n(x)} \Rightarrow Z(x) \quad \text{in } C(\mathcal{X}, \mathbb{R}),$$

with non-degenerate limit, then the limit process Z must be **max-stable**. Moreover, each margin $Z(x)$ is GEV-distributed.

Simple max-stable processes (unit Fréchet margins)

Standardization

It is convenient to work with **unit Fréchet** margins:

$$\mathbb{P}\{Z(x) \leq z\} = \exp(-1/z), \qquad z > 0,\ x \in \mathcal{X}.$$

A max-stable process with these margins is called **simple max-stable**.

Why this helps

After marginal standardization, all modeling effort focuses on **extremal dependence across space**.

Spectral representation (Poisson storms)

Theorem (de Haan, Penrose): spectral representation

Any non-degenerate **simple** max-stable process **with continuous sample paths** admits a representation

$$Z(x) \overset{d}{=} \max_{i \geq 1} \zeta_i f_i(x), \qquad x \in \mathcal{X},$$

where $\{(\zeta_i, f_i)\}_{i \geq 1}$ are points of a Poisson process on $(0, \infty) \times C(\mathcal{X}, \mathbb{R}_+)$ with intensity $\zeta^{-2}\, d\zeta\, \nu(df)$ and

$$\int f(x)\, \nu(df) = 1, \qquad x \in \mathcal{X}.$$

Storm interpretation

ζ_i = **severity** (“radius”), $f_i(\cdot)$ = **spatial footprint** (“angle”). Max-stable processes are pointwise maxima over infinitely many storms.

Historically important models

- **Smith (1990):** mixed moving maxima with Gaussian kernel

$$Z(x) = \max_{i \geq 1} \zeta_i \, \varphi(x - U_i; 0, \Sigma).$$

- **Schlather (2002):** extremal Gaussian

$$Z(x) = \sqrt{2\pi} \max_{i \geq 1} \zeta_i \max\{0, W_i(x)\},$$

  with stationary Gaussian W (correlation ρ).

- **Brown–Resnick (Kabluchko et al. 2009):**

$$Z(x) = \max_{i \geq 1} \zeta_i \exp\{W_i(x) - \gamma(x)\},$$

  with fractional Gaussian process with semivariogram $\gamma(h)$.

Finite-dimensional distributions

Fix locations $x_1, \ldots, x_k \in \mathcal{X}$ and $\boldsymbol{z} \in (0, \infty)^k$. Let $Z(x) = \max_{i \geq 1} \zeta_i Y_i(x)$ with $\mathbb{E}[Y(x)] = 1$.

Exponent function from the spectral process

For a simple max-stable process,

$$\mathbb{P}\{Z(x_1) \leq z_1, \ldots, Z(x_k) \leq z_k\} = \exp\{-V_{\boldsymbol{x}}(\boldsymbol{z})\},$$

with

$$V_{\boldsymbol{x}}(z_1, \ldots, z_k) = \mathbb{E}\left[\max_{j=1,\ldots,k} \frac{Y(x_j)}{z_j}\right].$$

Extremal coefficient (summary of dependence)

Setting $z_1 = \cdots = z_k = z$ gives

$$V_{\boldsymbol{x}}(z, \ldots, z) = \frac{\theta(x_1, \ldots, x_k)}{z}, \qquad \theta(x_1, \ldots, x_k) = \mathbb{E}\left[\max_{j=1,\ldots,k} Y(x_j)\right] \in [1, k].$$

Generalized Pareto processes

POT for processes: exceedances via a cost functional ℓ

Let $X = \{X(t)\}_{t \in T}$ be a nonnegative heavy-tailed continuous process on a compact set T, viewed as a random element in $C := C(T, [0, \infty))$ with sup-norm $\|x\|_\infty = \sup_{t \in T} x(t)$.

Generalized notion of exceedance

Instead of $\sup_t X(t) > u$, define exceedances through a **cost functional**

$$\ell : C \to [0, \infty), \qquad \ell(\lambda x) = \lambda\, \ell(x) \;\; (\lambda > 0),$$

and choose thresholds $u \to \infty$ such that $\mathbb{P}\{\ell(X) > u\} \to 0$. Then the POT object is the conditional law of the rescaled process:

$$\mathcal{L}\big(u^{-1}X \;\big|\; \ell(X) > u\big)\,, \qquad u \to \infty.$$

Examples: $\ell(x) = \|x\|_\infty$ (supremum), $\ell(x) = \int_T x(t)\,dt$ (total mass), $\ell(x) = x(t_0)$ (site-specific), $\ell(x) = \inf_{t \in T} x(t)$ (all-large event).

The ℓ–Pareto (GPD) process: three equivalent characterizations

Theorem: equivalences

For a continuous process W, the following are equivalent:

(1) Constructive form: $W(t) = P\,Y(t)$, $t \in T$, where P is Pareto with tail index $\alpha > 0$, $\mathbb{P}(P > u) = u^{-\alpha}$ for $u \geq 1$, Y is continuous with $\ell(Y) \equiv 1$, and $P \perp Y$.

(2) Homogeneity (scaling) property: $\mathbb{P}\{\ell(W) > 1\} = 1$ and for all $u \geq 1$ and measurable $A \subset \{f : \ell(f) \geq 1\}$,

$$\mathbb{P}(W \in uA) = u^{-\alpha}\,\mathbb{P}(W \in A).$$

(3) Peaks-over-threshold stability: for all $u \geq 1$ with $\mathbb{P}\{\ell(W) > u\} > 0$,

$$\mathcal{L}\big(u^{-1}W \mid \ell(W) > u\big) = \mathcal{L}(W).$$

SPECTRAL MEASURE AND “RADIUS–SHAPE” DECOMPOSITION

Terminology

Such a W is called a **simple ℓ–Pareto process** (a functional GP process).

Spectral measure (angular law)

For a simple ℓ–Pareto process W, define the **spectral process**

$$Y = \frac{W}{\ell(W)} \qquad \Rightarrow \qquad \ell(Y) = 1.$$

Its distribution on the “ℓ-unit sphere”

$$S_\ell := \{f \in C : \ \ell(f) = 1\}$$

is the **spectral measure** σ_ℓ.

Spectral measure and "radius–shape" decomposition

Analogy with multivariate EVT

Radius: $P = \ell(W)$ (Pareto tail)
Angle/shape: $Y = W/\ell(W)$ (event profile), independent of P.

Why this matters for simulation

To generate extreme episodes: sample a severity P (Pareto) and a shape $Y \sim \sigma_\ell$, then set $W = PY$.

POT stability: an exact "threshold-stability" in function space

Peaks-over-threshold stability (re-stated)

If W is a simple ℓ–Pareto process, then for any $u \geq 1$,

$$\mathcal{L}\left(u^{-1}W \,\middle|\, \ell(W) > u\right) = \mathcal{L}(W).$$

So raising the threshold (in ℓ-space) and rescaling produces the **same law**.

Use as a diagnostic

Empirically, if the POT model is appropriate, the distribution of $u^{-1}X$ given $\ell(X) > u$ should be roughly invariant when u increases.

From regular variation to ℓ–Pareto limits

Assume X is **regularly varying** in C with index α and spectral measure σ on $S := \{f \in C : \|f\|_\infty = 1\}$.

Theorem: POT limit is ℓ–Pareto

If ℓ is continuous at the origin and does not vanish σ-a.e., then

$$\mathcal{L}\big(u^{-1}X \mid \ell(X) > u\big) \ \Rightarrow \ P^{\ell}_{\alpha,\sigma_\ell}, \qquad u \to \infty,$$

where $P^{\ell}_{\alpha,\sigma_\ell}$ is a simple ℓ–Pareto process.

Explicit spectral measure update

The limiting spectral measure σ_ℓ on S_ℓ is

$$\sigma_\ell(B) = \frac{1}{c_\ell}\int_S \ell(f)^\alpha \,\mathbf{1}\left\{\frac{f}{\ell(f)} \in B\right\}\sigma(df), \qquad c_\ell = \int_S \ell(f)^\alpha \sigma(df).$$

Point processes theory for extreme values

Point processes: the object behind exceedances

Counting measures and point processes

Given a sequence of points $(x_i)_{i\geq 1}$ in a measurable space $(E,\mathcal{E})$, the map $m(A) = \sum_{i\geq 1} \mathbf{1}\{x_i \in A\}$, $A \in \mathcal{E}$, defines a **counting measure** (finite on compacts). Let $M_p(E)$ be the space of point measures on E. A **point process** on E is a measurable map

$$N : (\Omega, \mathcal{A}, \mathbb{P}) \to \big(M_p(E), \mathcal{M}_p(E)\big).$$

Examples of two exceedance point processes

- **Times of exceedances:** $N_n^{(\tau)}(B) = \sum_{i=1}^n \mathbf{1}\{i/n \in B,\ X_i > u_n(\tau)\}$ on $E = (0,1]$.
- **Times + normalized magnitudes:** $N_n^{(\tau)}(B) = \sum_{i=1}^n \mathbf{1}\{(i/n, (X_i - b_n)/a_n) \in B\}$ on $E = (0,1] \times \mathbb{R}$.

Characterizing a point process via its Laplace functional

Finite-dimensional distributions

The law of N is determined by the joint laws of $\big(N(A_1), \dots, N(A_m)\big)$, $A_1, \dots, A_m \in \mathcal{E}$.

Laplace functional

For $g \geq 0$ measurable with suitable support,

$$L_N(g) := \mathbb{E}\Big[\exp\Big\{-\int_E g\,dN\Big\}\Big].$$

Equivalently, for $m \in M_p(E)$, $\int_E g\,dm$ is a finite sum over the points of m.

Weak convergence of point processes is conveniently checked through convergence of Laplace functionals for nonnegative test functions with compact support.

Poisson point processes (PPP)

Definition

A point process N on $(E, \mathcal{E})$ is a **Poisson random measure** with mean measure μ if:

- for any $A \in \mathcal{E}$, $N(A) \sim \mathrm{Poisson}(\mu(A))$;
- for disjoint $A_1, \ldots, A_m$, the counts $N(A_1), \ldots, N(A_m)$ are independent.

Laplace functional of a PPP

$$L_N(g) = \exp\left\{-\int_E \left(1 - e^{-g(x)}\right) \mu(dx)\right\}.$$

Rare-event heuristic: Exceedances above high thresholds behave like points of a PPP: *rare* and *approximately independent*.

Compound Poisson processes and clustering of extremes

Compound Poisson point process

A **compound Poisson** point process has Poisson-distributed cluster centers, but each center generates a random **cluster size** with distribution π.

Laplace functional (cluster size π)

Let $L_\pi(\cdot)$ be the Laplace transform of the cluster size. Then

$$L_N(g) = \exp\left\{ - \int_E \left(1 - L_\pi(g(x))\right) \mu(dx) \right\}.$$

EVT message: For dependent sequences (time series), extremes tend to appear in **clusters**. The limiting exceedance process is typically **compound Poisson**, not Poisson.

Weak convergence of point processes and the i.i.d. exceedance limit

Weak convergence in $M_p(E)$

We say $N_n \Rightarrow N$ in $M_p(E)$ if, for all $A_1, \ldots, A_m$ with $\mathbb{P}(N(\partial A_i) = 0) = 1$,

$$\big(N_n(A_1), \ldots, N_n(A_m)\big) \Rightarrow \big(N(A_1), \ldots, N(A_m)\big).$$

A key criterion is: $N_n \Rightarrow N \iff L_{N_n}(g) \to L_N(g)$ for all $g \geq 0$ with compact support.

The i.i.d. exceedance process converges to a PPP

If $u_n(\tau)$ is such that $n\bar{F}(u_n(\tau)) \to \tau$ (with $\bar{F} = 1 - F$), then

$$N_n^{(\tau)} \Rightarrow N,$$

where N is a homogeneous PPP on $(0, 1]$ with intensity measure $\tau \, |\cdot|$ (Lebesgue measure).

Dependent case

Mixing-type condition $\Delta(u_n(\tau))$

Let $\mathcal{F}_{p,q}(\tau)$ be the sigma-field generated by $\{X_i > u_n(\tau)\}$, $p \le i \le q$. Define

$$\alpha_{n,\ell}(\tau) = \sup\left\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{1,t}(\tau),\ B \in \mathcal{F}_{t+\ell,\infty}(\tau),\ t \ge 1\right\}.$$

Condition $\Delta(u_n(\tau))$ holds if there exists $\ell_n = o(n)$ with $\ell_n \to \infty$ and $\alpha_{n,\ell_n}(\tau) \to 0$.

Limit theorem

Assume $\Delta(u_n(\tau))$ and $n\bar{F}(u_n(\tau)) \to \tau$. If $N_n^{(\tau)}$ converges, the limit is a **homogeneous compound Poisson** process with intensity $\tau\theta\,|\cdot|$ and cluster size distribution π, where $\theta \in (0, 1]$ is the **extremal index**. Moreover, if $\Delta(u_n(\tau))$ holds for all $\tau > 0$, then θ and π do not depend on τ and

$$\theta^{-1} = \sum_{k \ge 1} k\,\pi(k) = \mathbb{E}[\text{cluster size}].$$

EVT deliverables that a generator should reproduce

- **Univariate**: tail index / GEV-GPD parameters; high-quantile calibration.
- **Multivariate**: tail dependence function, extremal coefficients, asymptotic (in)dependence.
- **Time/space**: extremal index, tail/spectral tail process, cluster size distribution.

A useful workflow

Fit univariate EVT structure *first*, then train a generator constrained/regularized to respect these tail features.

Thank you

References (a short list of books)

- J. Beirlant et al. (2004). *Statistics of Extremes*. Wiley.
- A. C. de Carvalho, R. Huser, P. Naveau and B. J. Reich (2026). *Handbook on Statistics of Extremes*. Chapman & Hall/CRC.
- S. Coles (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.
- P. Embrechts, C. Klüppelberg, T. Mikosch (1997). *Modelling Extremal Events*. Springer.
- L. de Haan and A. Ferreira (2006). *Extreme Value Theory: An Introduction*. Springer.
- R. Kulik and P. Soulier (2020). *Heavy-tailed Time Series*. Springer.
- T. Mikosch and O. Wintergberger (2025). *Extreme Value Theory for Time Series: Models with Power-law Tails*. Springer.
- S. Resnick (2007). *Heavy-Tail Phenomena*. Springer.

Backup: a minimal glossary

- **GEV** (Generalized Extreme Value): limit law for normalized maxima.
- **GPD** (Generalized Pareto Distributions): limit law for threshold exceedances.
- **(Multivariate) Regular variation**: power-law tails; tail index α.
- **Max-stable**: stable under maxima; multivariate and process analogues.
- **Threshold-stable**: stable under thresholding; multivariate and process analogues.
- **Spectral measures / functions**: measures / functions for angle components of an EVT object.