Models for Complex Extreme Events

Anthony Davison

Joint with
Peiman Asadi, Raphaël de Fondeville, Sebastian Engelke,

Raphaël Huser, Simone Padoan, Mathieu Ribatet,
Mariá Süveges, Emeric Thibaud

Funding: Swiss National Science Foundation

Opening

- Classical extreme value theory/methods now widely used in climate science
- Multivariate ideas also starting to be used
- Major advances in more complex modelling over the past decade
- Goal of talk:
  - Recall some basics
  - Overview recent developments
- Applications illustrative—aim to show a toolkit, not build something with it

Basics

Back to basics

- Extreme value theory is based on **limiting** models for tails of distributions
  - Generalised extreme-value distribution (GEV) applies for maxima of an infinite sample
  - Generalized Pareto distribution (GPD) applies for peaks over an 'infinite' threshold
- In practice fitted to finite samples (e.g., seasonal maxima/minima), so fit may extrapolate badly
  - Shape parameter ξ_m depends on sample size m?—**penultimate approximation**
  - Typically unstable, need a lot of data—but in climate contexts may have this

How $\widehat{\xi}$ depends on record length for rainfall

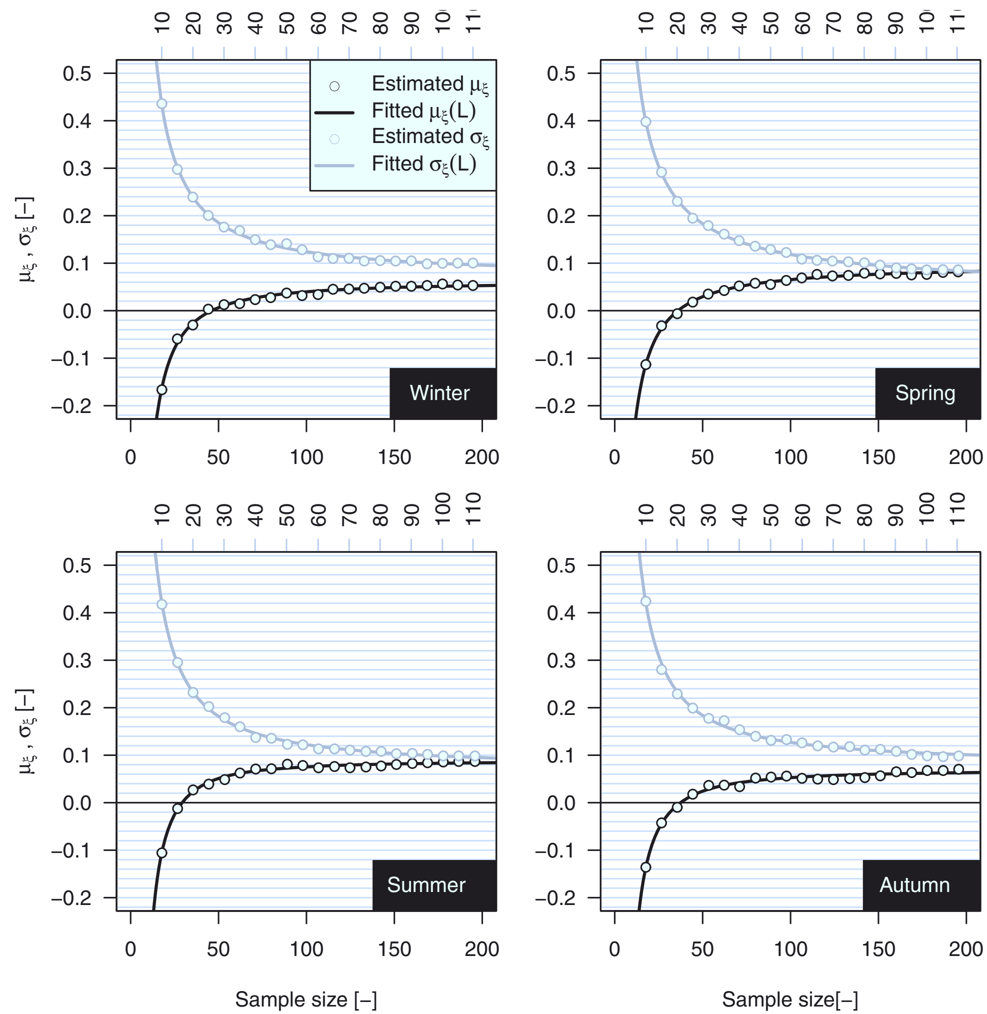
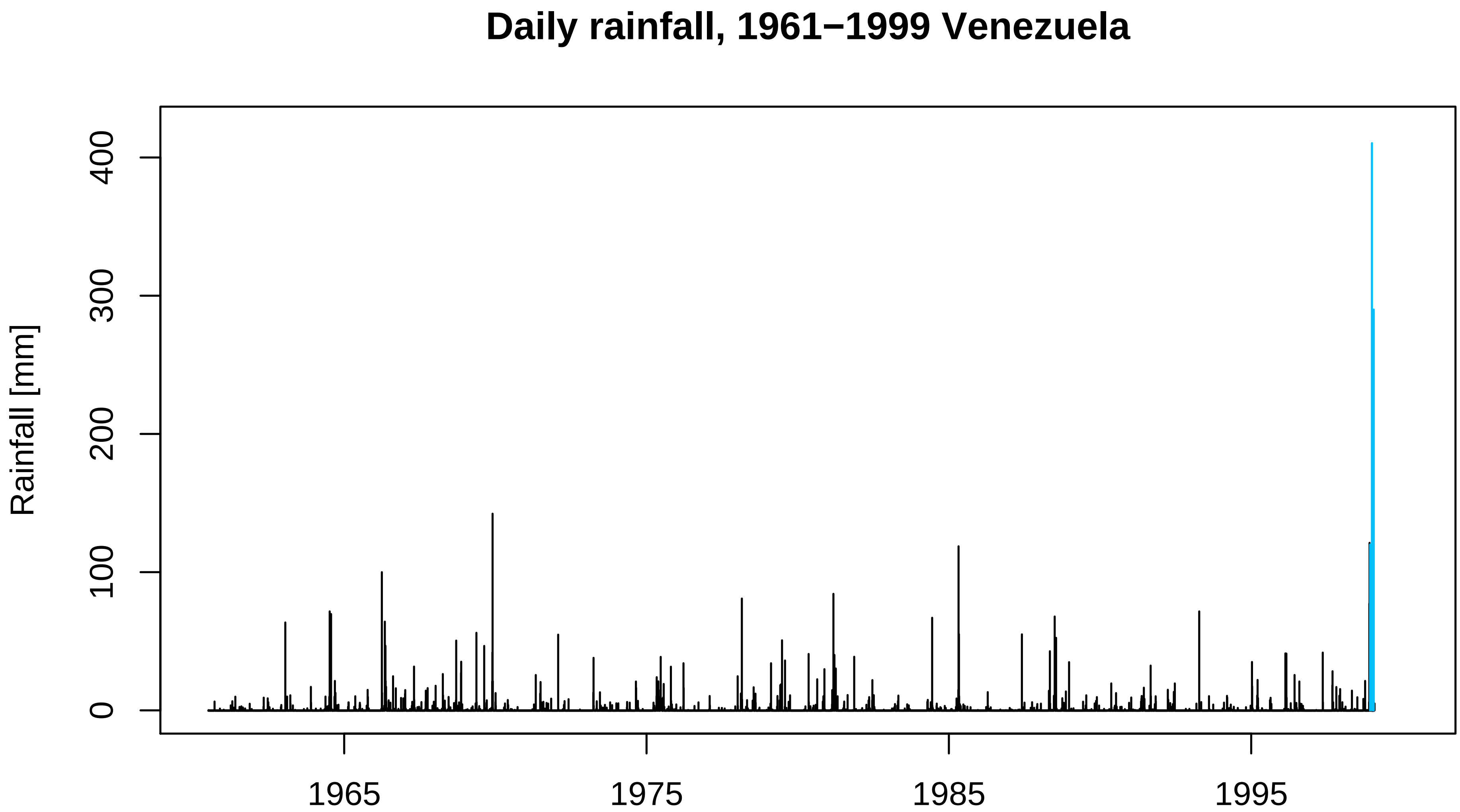


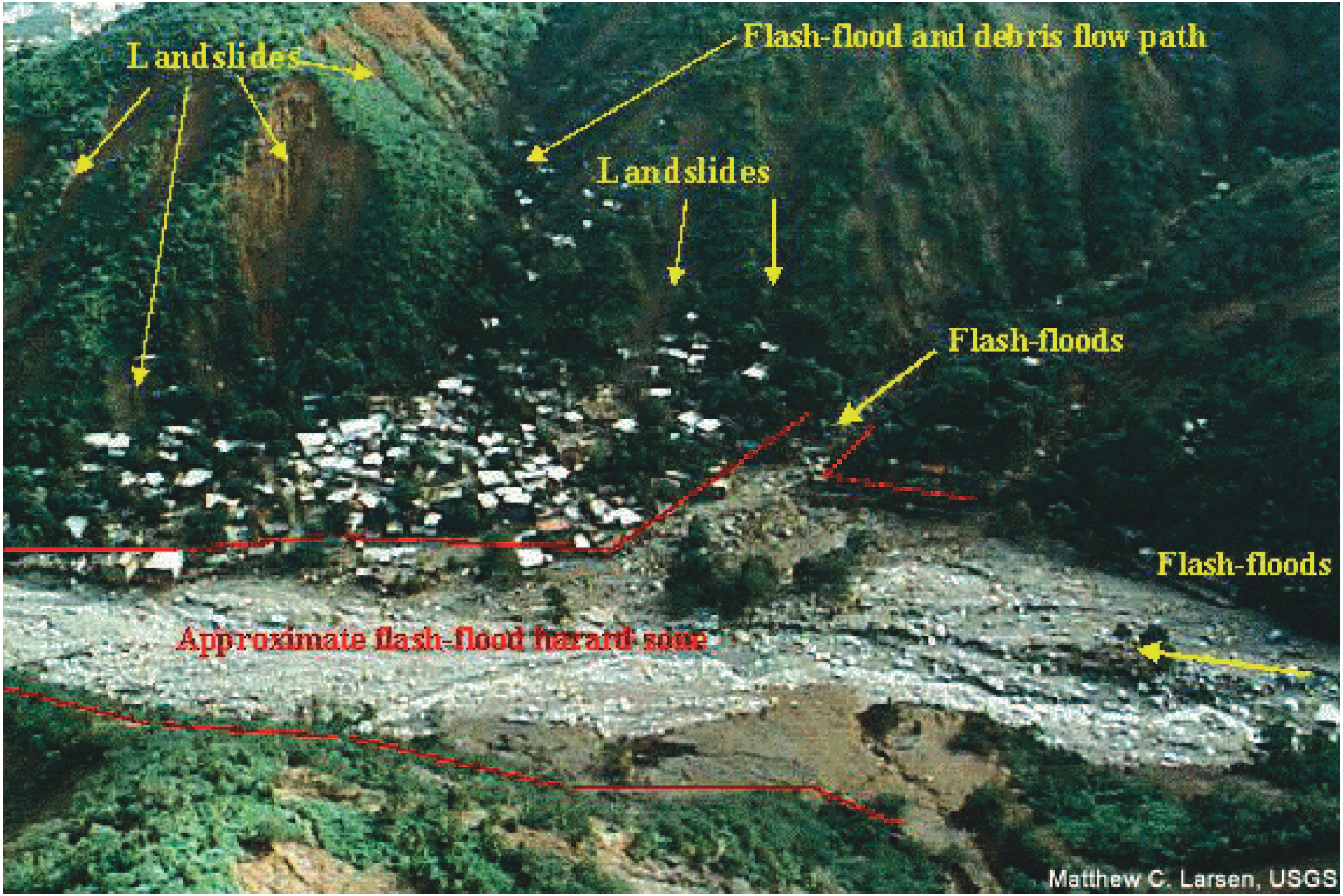
Figure 9. Mean and standard deviation of GP shape parameter versus sample size (record length) the 1900–2011 sample.

From Serinaldi and Kilsby (2014), *Water Resources Research*

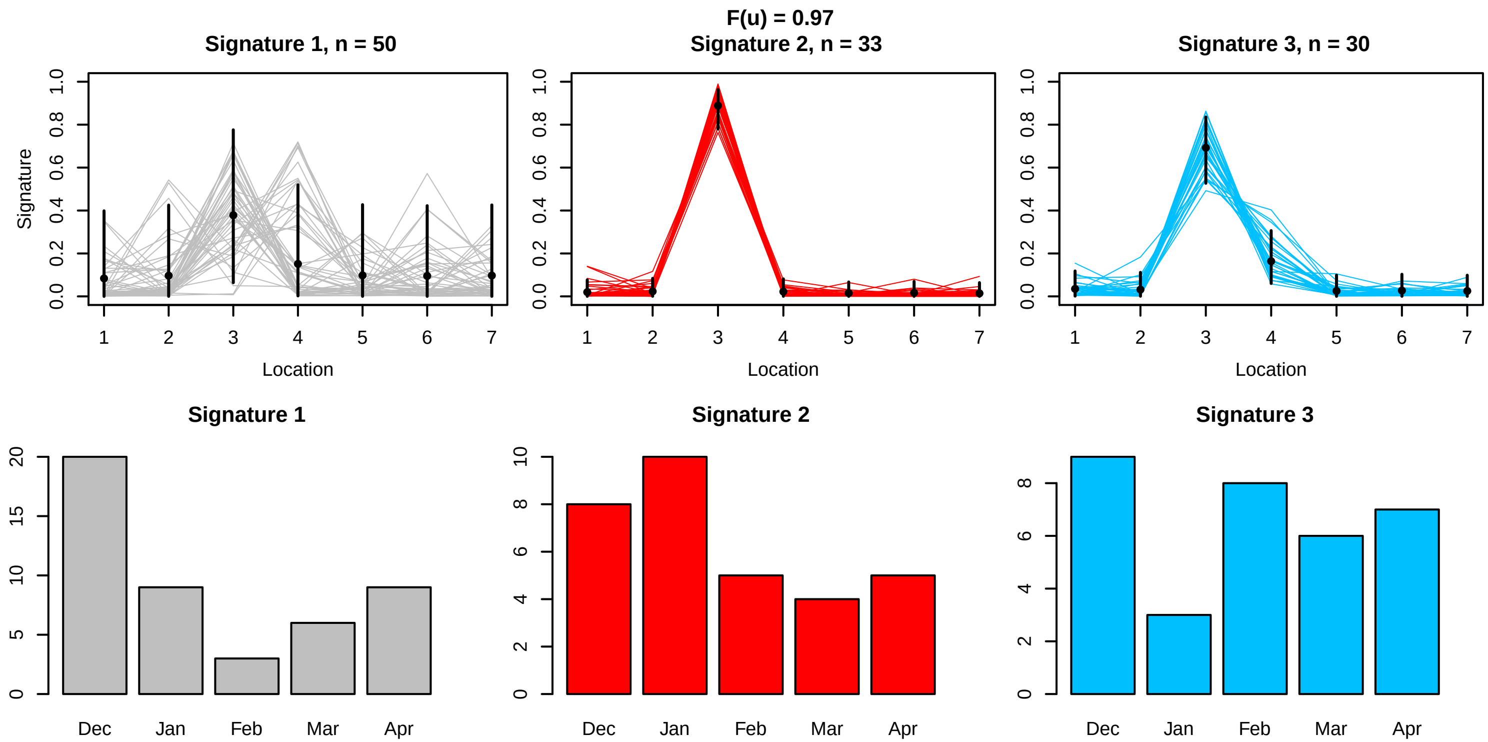
The Vargas event and mixtures



Tanaguarena afterwards

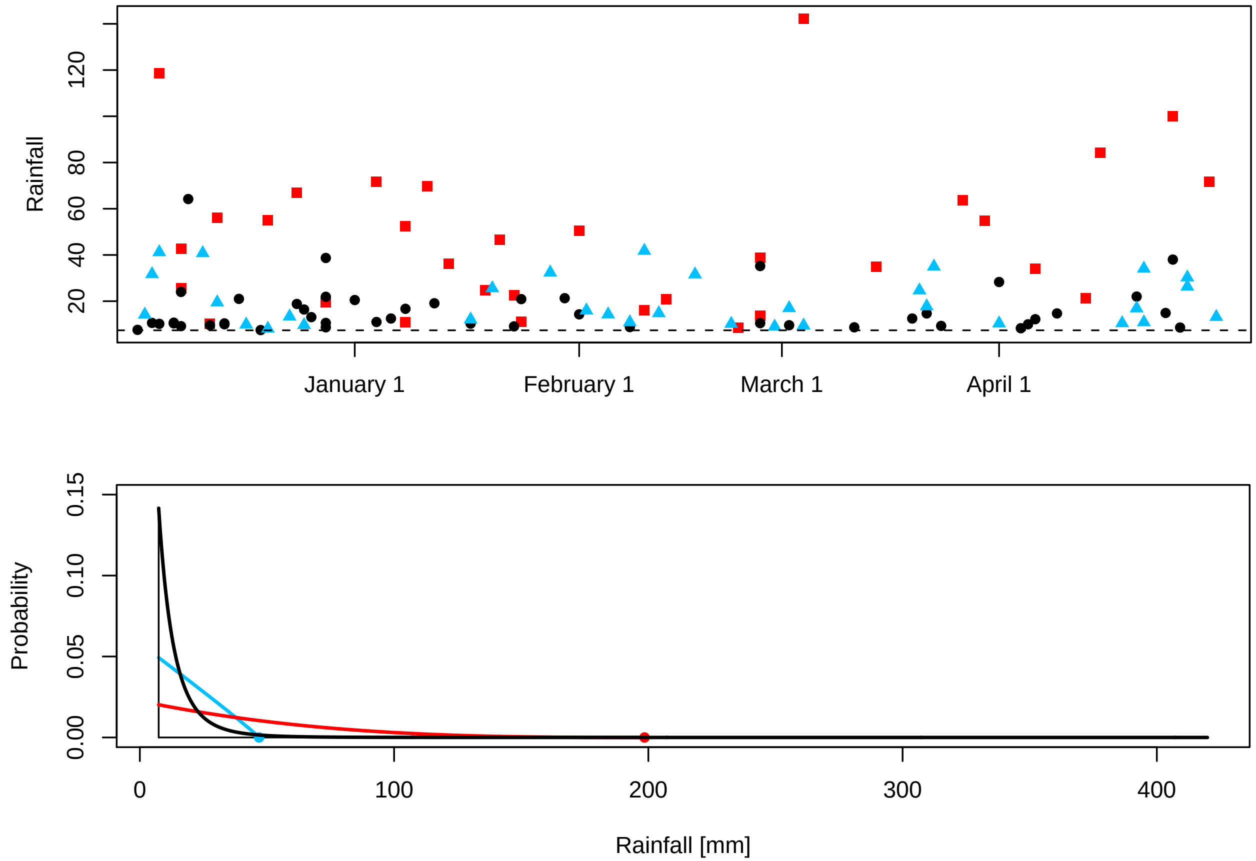


Mixture model



Top: Normalized clusters (coloured lines) of daily precipitation for the Venezuela data, observed before December 1999, and the fitted signatures (black dots). The black lines indicate the 0.025 and 0.975 quantiles of the fitted Dirichlet component.
Bottom: Frequency per month of the different signatures observed before December 1999 having peaks above u corresponding to $F(u) = 0.97$. Signature 1 is plotted in black, signature 2 in red, signature 3 in blue.

Mixture model



Top: peaks of the clusters observed before December 1999 and used for the mixture fit using $F(u) = 0.97$. Bottom: GPD densities corresponding to the peaks of the three cluster types. In both panels, signature 1 is plotted in black, signature 2 in red, signature 3 in blue.

- Easiest to describe multivariate/complex extremes on a standard scale
  - Consider bivariate maxima (M_X, M_Y), and suppose that each individually has a GEV distribution
  - Then back-transformation using the respective GEVs gives that

$$Z_1 = \{1 + \xi_X(M_X - \eta_X)/\sigma_X\}^{1/\xi_X}, \quad Z_2 = \{1 + \xi_Y(M_Y - \eta_Y)/\tau_Y\}^{1/\xi_Y}$$

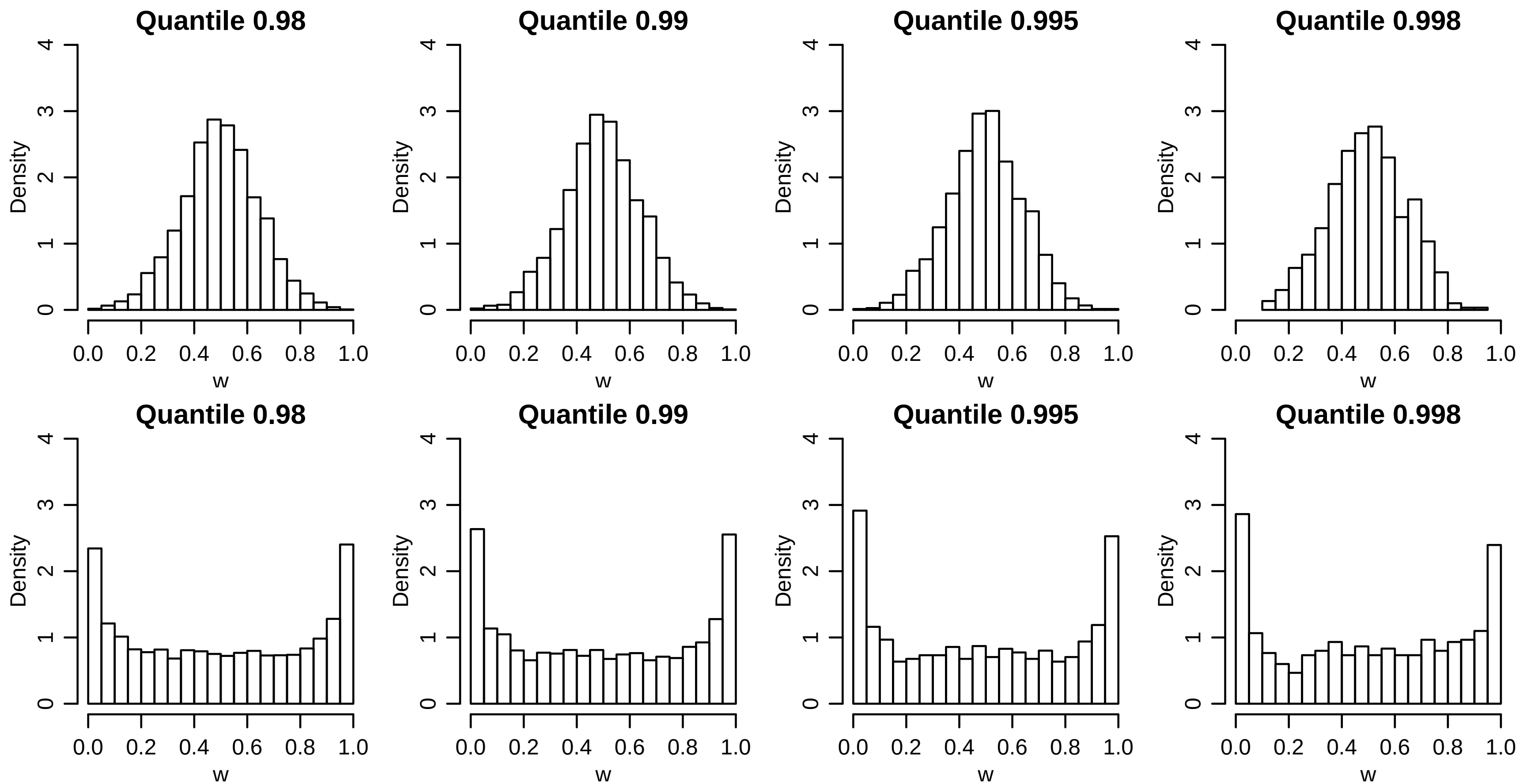
    have unit Fréchet marginal distributions, $\mathrm{P}(Z \le z) = \exp(-1/z)$, for $z > 0$.
  - Likewise for bivariate exceedances over high thresholds u_X, u_Y.
- In both cases, if a joint limiting distribution for (M_X, M_Y) exists, then

$$\mathrm{P}(Z_1 \le z_1, Z_2 \le z_2) = \exp\{-V(z_1, z_2)\}, \quad z_1, z_2 > 0,$$

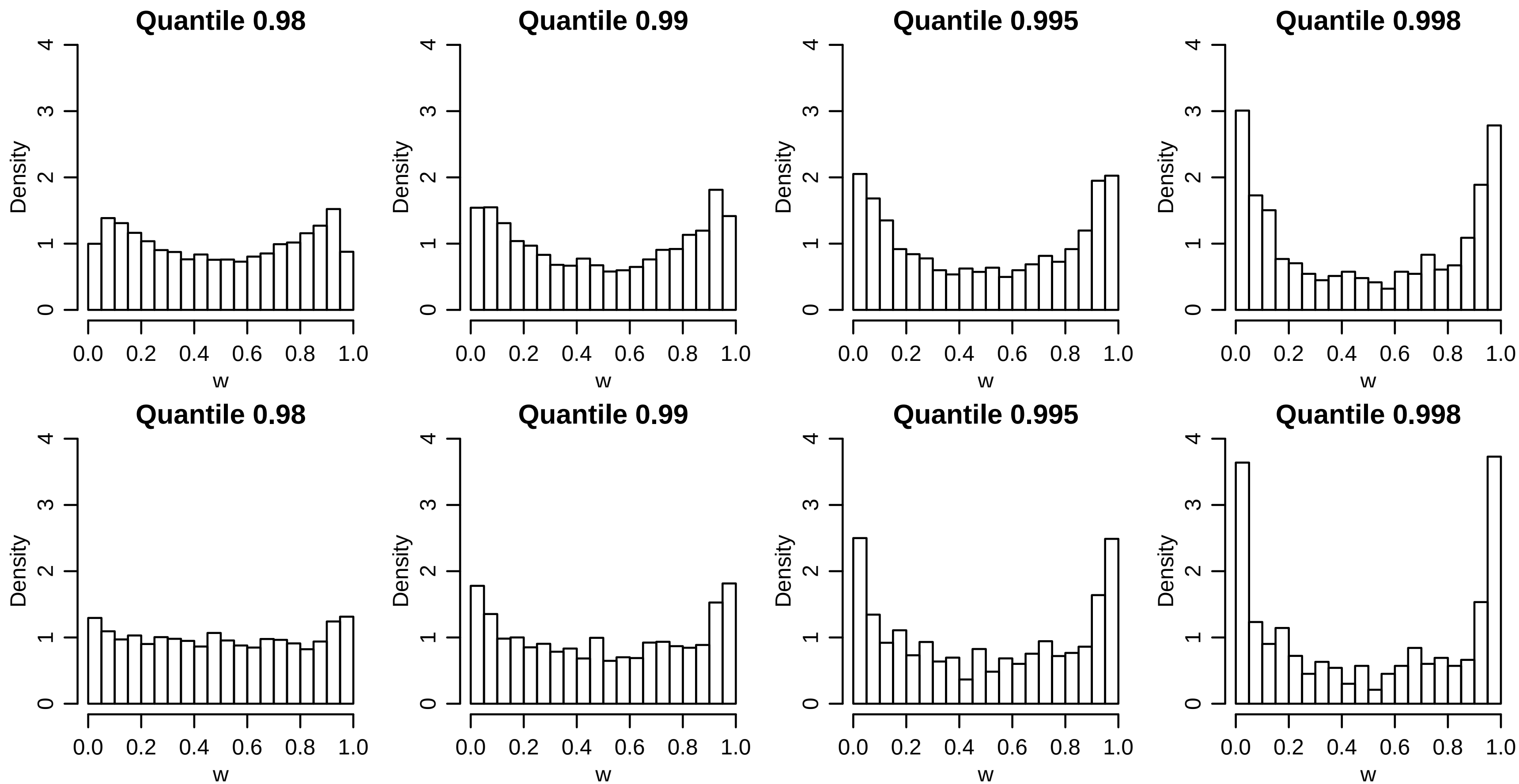
  where V is called an **exponent measure**.
- For large $R = Z_1 + Z_2$, the variables R and $W = Z_1/R$ are (approximately) independent, so we can extrapolate to rarer events than those observed.

Logistic dependence

Histograms of $W = Z_1/(Z_1 + Z_2)$ given $R > r_0$ for pairs (Z_1, Z_2) from the logistic dependence function, with r_0 corresponding to the 0.98, 0.99, 0.995, 0.998 quantiles of Z. Above: $\alpha = 0.3$. Below: $\alpha = 0.7$.



Histograms of $W = Z_1/(Z_1 + Z_2)$ given $R > r_0$ for pairs (Z_1, Z_2), with r_0 corresponding to the 0.98, 0.99, 0.995, 0.998 quantiles of Z. Above: normal data with $\rho = 0.8$. Below: successive pairs of Eskdalemuir observations.



Asymptotic dependence and independence

- Classical models for multivariate extremes (logistic, bilogistic, ...) show
  - **asymptotic dependence**: degree of dependence does not vary with the severity of the event.
- Data can (often?) show
  - **asymptotic independence**: dependence decreases as severity increases.
- Multivariate Gaussian models (e.g., ARMA) are asymptotically independent
- Asymptotic independence models can be constructed from asymptotic dependence models by **inversion**, but few encompass both
- Conclusion:

Asymptotic dependence and independence

- Classical models for multivariate extremes (logistic, bilogistic, . . .) show
  - **asymptotic dependence**: degree of dependence does not vary with the severity of the event.
- Data can (often?) show
  - **asymptotic independence**: dependence decreases as severity increases.
- Multivariate Gaussian models (e.g., ARMA) are asymptotically independent
- Asymptotic independence models can be constructed from asymptotic dependence models by **inversion**, but few encompass both
- Conclusion: **if extrapolation from univariate extremes is dodgy, extrapolation from multivariate extremes is even dodgier**

Asymptotic dependence and independence

- Classical models for multivariate extremes (logistic, bilogistic, . . .) show
  - **asymptotic dependence**: degree of dependence does not vary with the severity of the event.
- Data can (often?) show
  - **asymptotic independence**: dependence decreases as severity increases.
- Multivariate Gaussian models (e.g., ARMA) are asymptotically independent
- Asymptotic independence models can be constructed from asymptotic dependence models by **inversion**, but few encompass both
- Conclusion: **if extrapolation from univariate extremes is dodgy, extrapolation from multivariate extremes is even dodgier**

Why do it? Extrapolation!

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Motivations for modelling extremes:

- pointwise maps of return levels
  - joint probabilities not of interest but
  - estimation may be aided by smoothing, 'borrowing strength'
  - but how many 'independent' station-years?
  - Approaches: Bayesian hierarchical models ('non-extremal'), estimating functions (Jun Yan, Thursday?)

Why do it? Extrapolation!

Motivations for modelling extremes:

- pointwise maps of return levels
  - joint probabilities not of interest but
  - estimation may be aided by smoothing, ‘borrowing strength’
  - but how many ‘independent’ station-years?
  - Approaches: Bayesian hierarchical models (‘non-extremal’), estimating functions (Jun Yan, Thursday?)
- estimation of probability of rare complex events
  - joint probabilities essential
  - max-stable processes/exceedances

Why do it? Extrapolation!

Motivations for modelling extremes:

- pointwise maps of return levels
  - joint probabilities not of interest but
  - estimation may be aided by smoothing, ‘borrowing strength’
  - but how many ‘independent’ station-years?
  - Approaches: Bayesian hierarchical models (‘non-extremal’), estimating functions (Jun Yan, Thursday?)
- estimation of probability of rare complex events
  - joint probabilities essential
  - max-stable processes/exceedances
- detection/attribution

Why do it? Extrapolation!

Motivations for modelling extremes:

- pointwise maps of return levels
  - joint probabilities not of interest but
  - estimation may be aided by smoothing, 'borrowing strength'
  - but how many 'independent' station-years?
  - Approaches: Bayesian hierarchical models ('non-extremal'), estimating functions (Jun Yan, Thursday?)
- estimation of probability of rare complex events
  - joint probabilities essential
  - max-stable processes/exceedances
- detection/attribution
- short-range forecasting
  - statistics of extremes not very relevant? (à revoir)

Max-stable processes

Max-stable processes

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- The GEV distribution is **max-stable**: maxima of independent GEV variables are also GEV—in fact, this is the defining property of the GEV distribution, and allows extrapolation to rare events.
- For the unit Fréchet distribution, this means that if $Z, Z_1, \ldots, Z_n \stackrel{\text{iid}}{\sim} \exp(-1/z)$, then for any n,

$$\max\{Z_1, \ldots, Z_n\} \stackrel{D}{=} nZ.$$

- For space/space-time problems we need a process analogue of the GEV, i.e., we seek a process $Z(x)$ such that if $Z_1(x), \ldots, Z_n(x) \stackrel{\text{iid}}{\sim} Z(x)$, then

$$\max\{Z_1(x), \ldots, Z_n(x)\} \stackrel{D}{=} nZ(x), \quad x \in \mathcal{X},$$

  where $\mathcal{X}$ represents a space/space-time domain of interest (e.g., the Rhine watershed within Switzerland over the years 2020–2100).
- In the process case we first transform the process so that its marginal distributions are standard Fréchet at every x.

- Let $W(x)$ be a non-negative random process with $\mathrm{E}\{W(x)\} = 1$ ($x \in \mathcal{X}$), and let

$$Z(x) = \sup_j R_j W_j(x), \quad x \in \mathcal{X}, \tag{1}$$

with $\{R_j\}$ a Poisson process on $\mathbb{R}_+$ of rate $\mathrm{d}r/r^2$ and $\{W_j\}$ replicates of W.

- Then

$$\mathrm{P}\{Z(x) \le z(x), x \in \mathcal{X}\} = \exp\left(-\mathrm{E}\left[\sup_{x\in\mathcal{X}}\left\{\frac{W(x)}{z(x)}\right\}\right]\right) = \exp\left[-V\{z(x)\}\right],$$

say, and this gives:

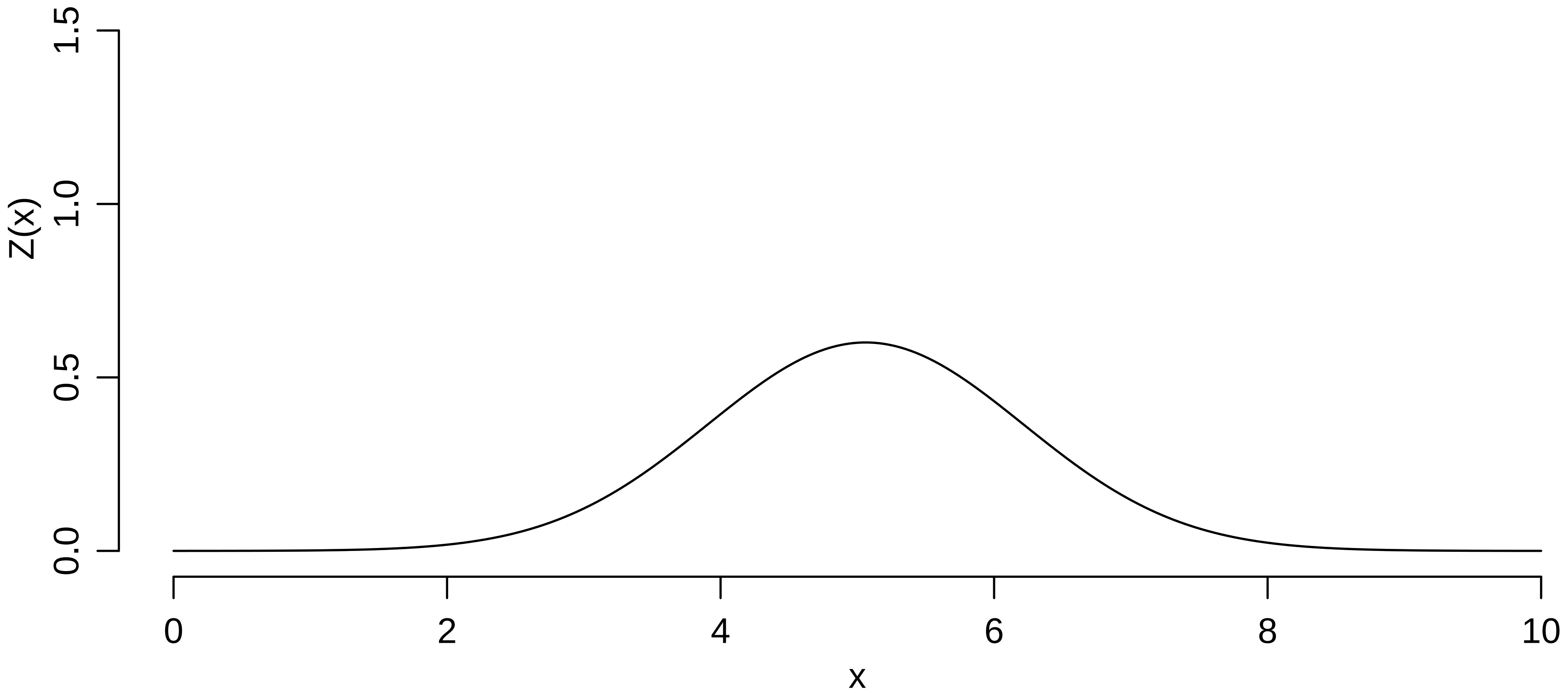
  – a **max-stable process** $\{Z(x) : x \in \mathcal{X}\}$, i.e., there exist functions $\{b_n(x)\}$ and $\{a_n(x)\} > 0$ such that

$$Z(x) \stackrel{D}{=} \max_{j=1}^{n}\left\{\frac{Z_j(x) - b_n(x)}{a_n(x)}\right\}, \quad x \in \mathcal{X}.$$

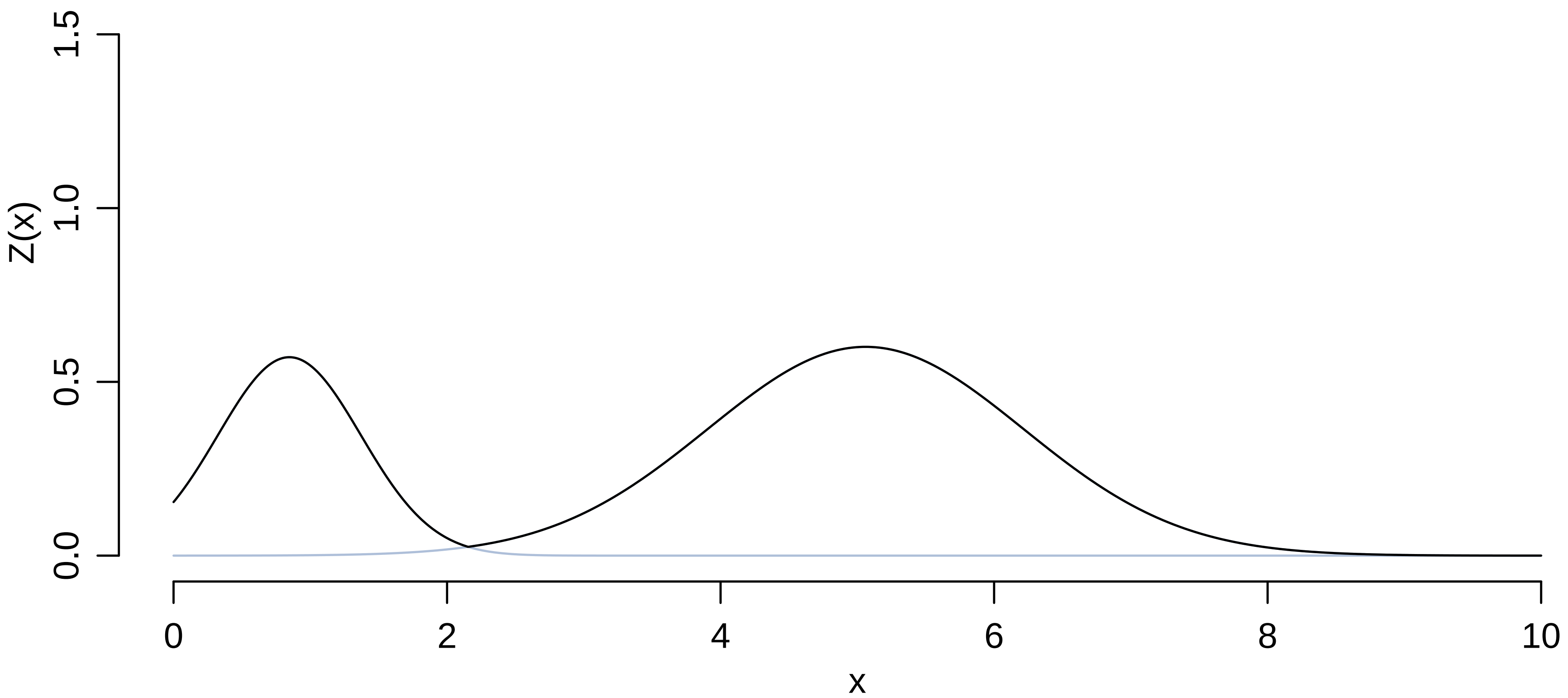
  – unit Fréchet margins at each $x \in \mathcal{X}$.

- In fact any max-stable process can be written using the **spectral representation** (1).
- Example: random Smith model . . .

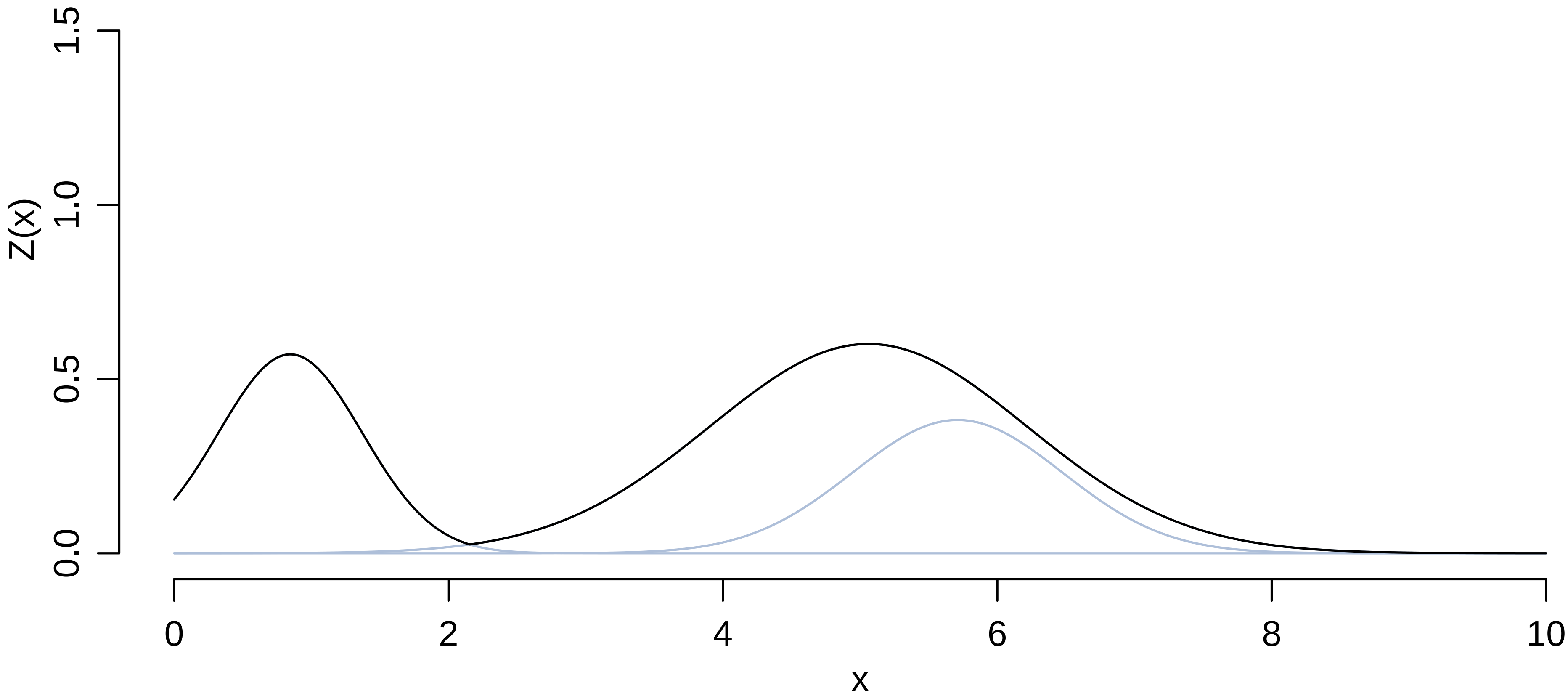
Construction of a max-stable process



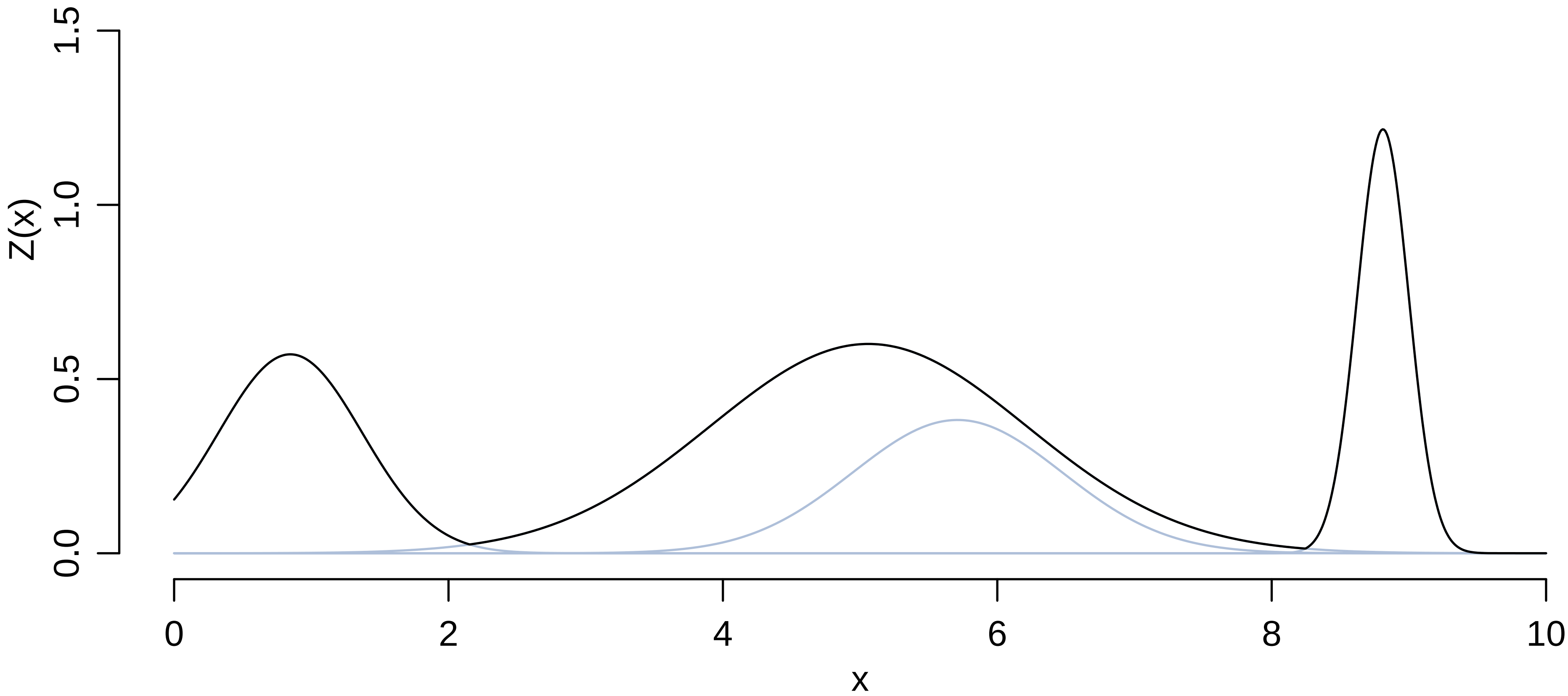
Construction of a max-stable process



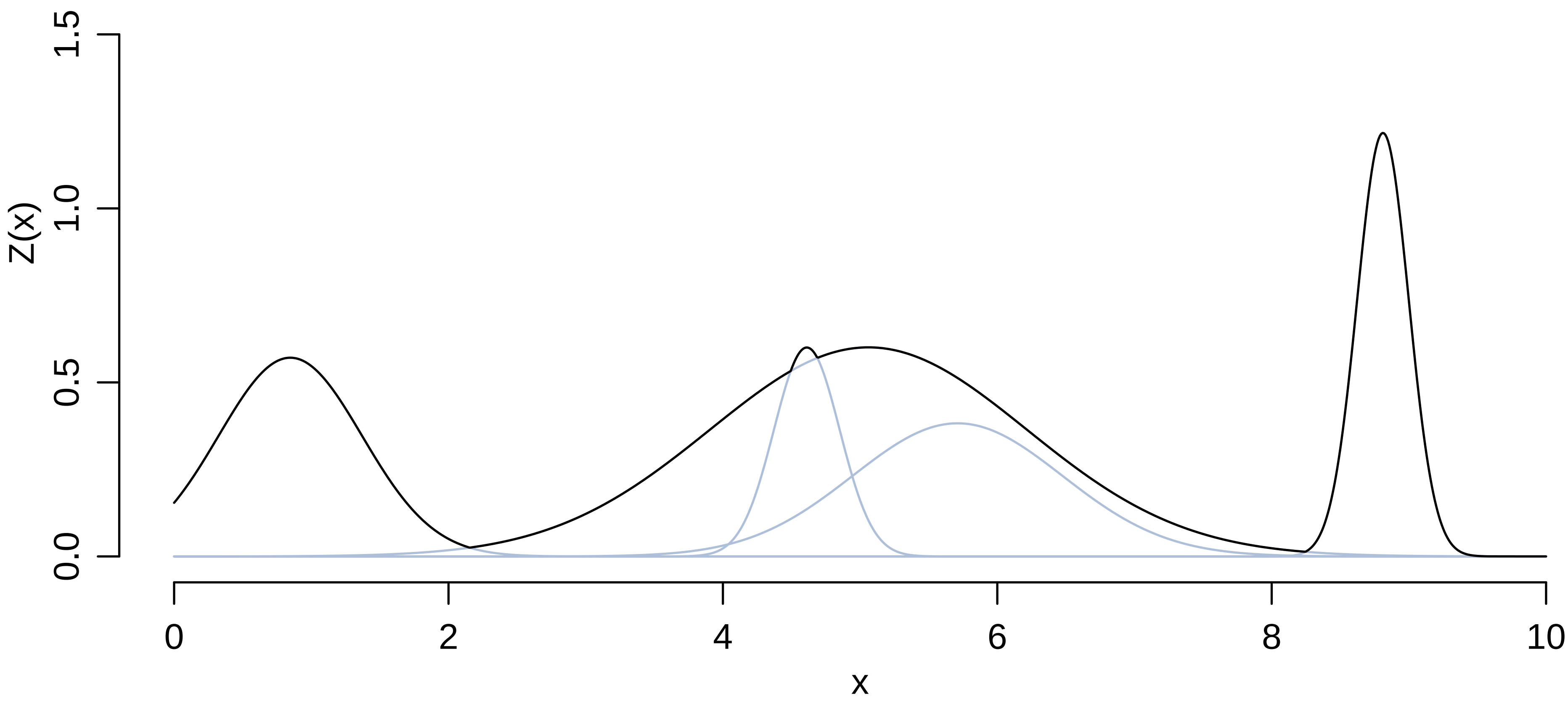
Construction of a max-stable process



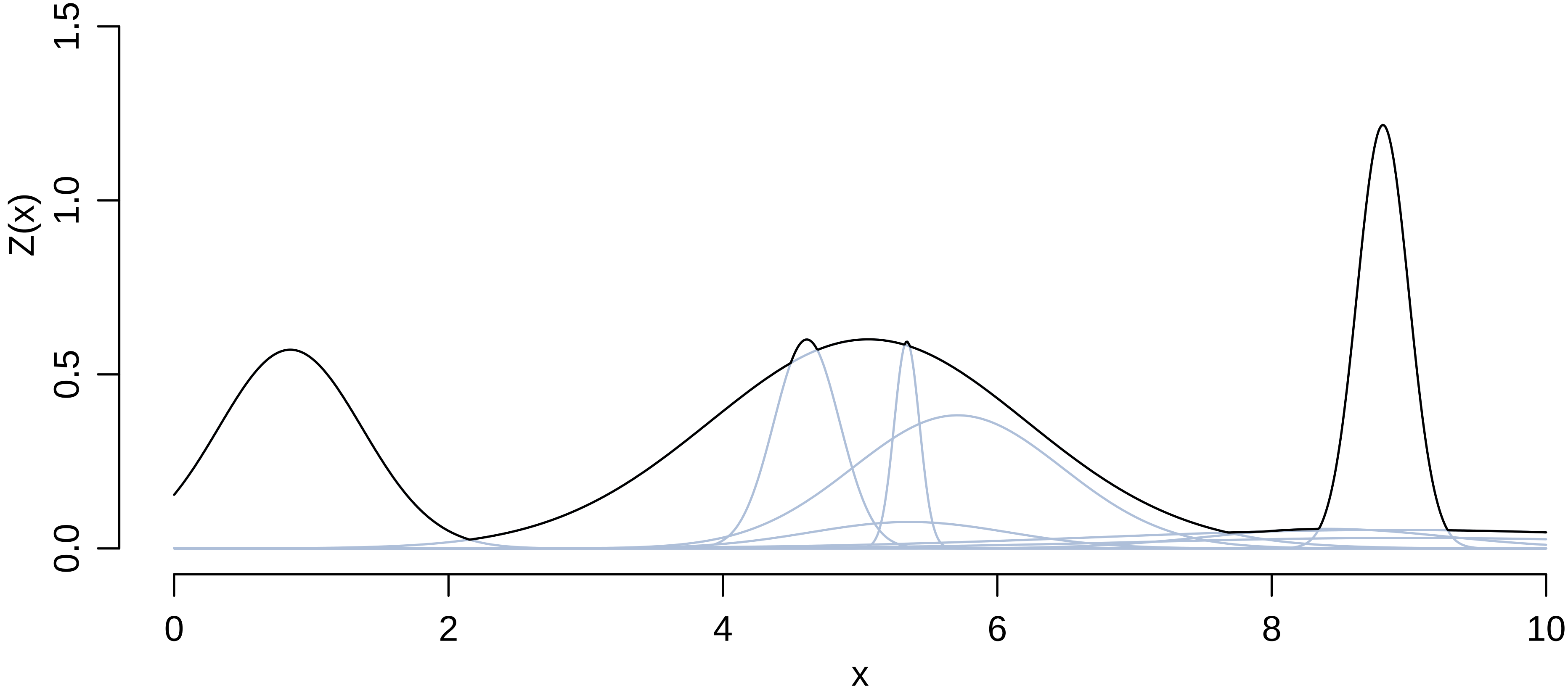
Construction of a max-stable process



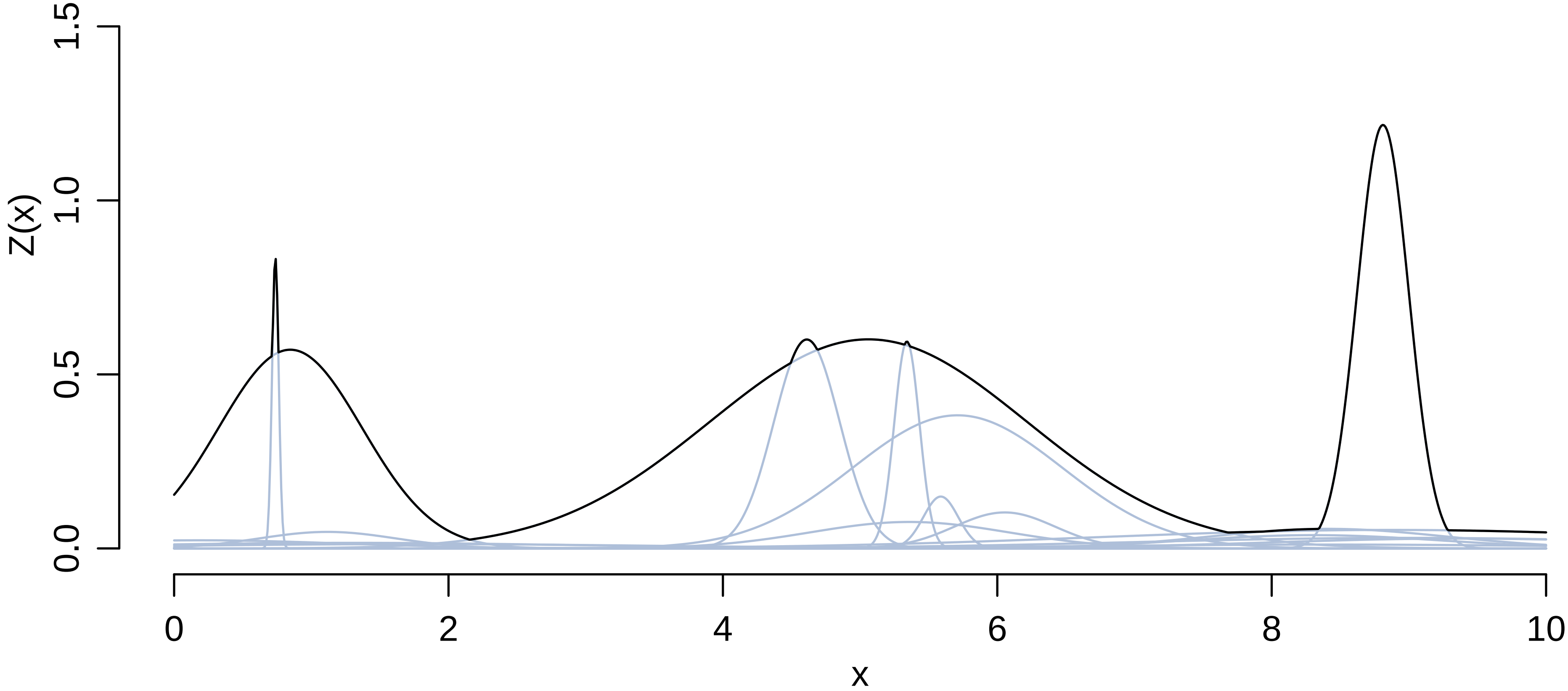
Construction of a max-stable process



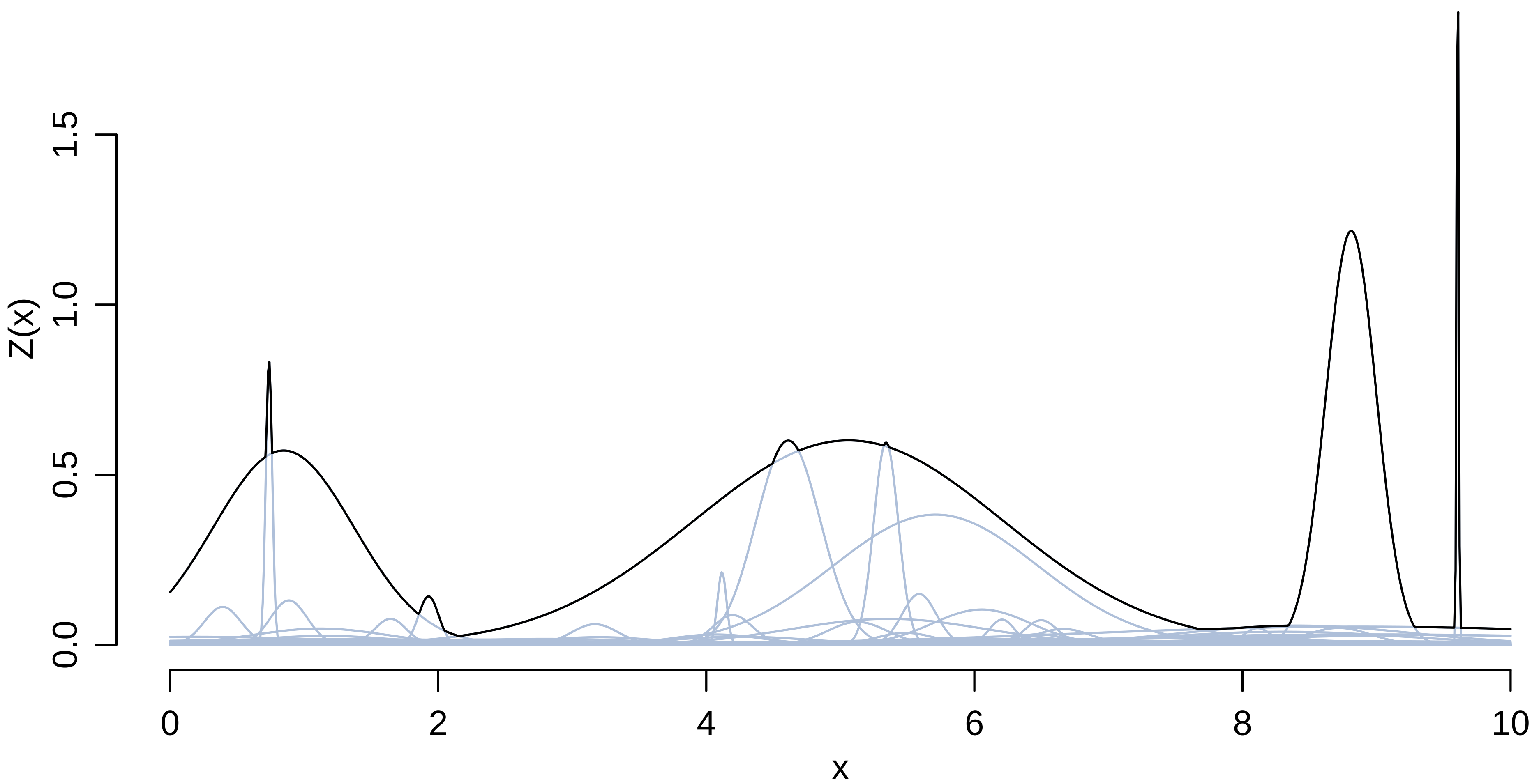
Construction of a max-stable process



Construction of a max-stable process



Max-stable process



Comments

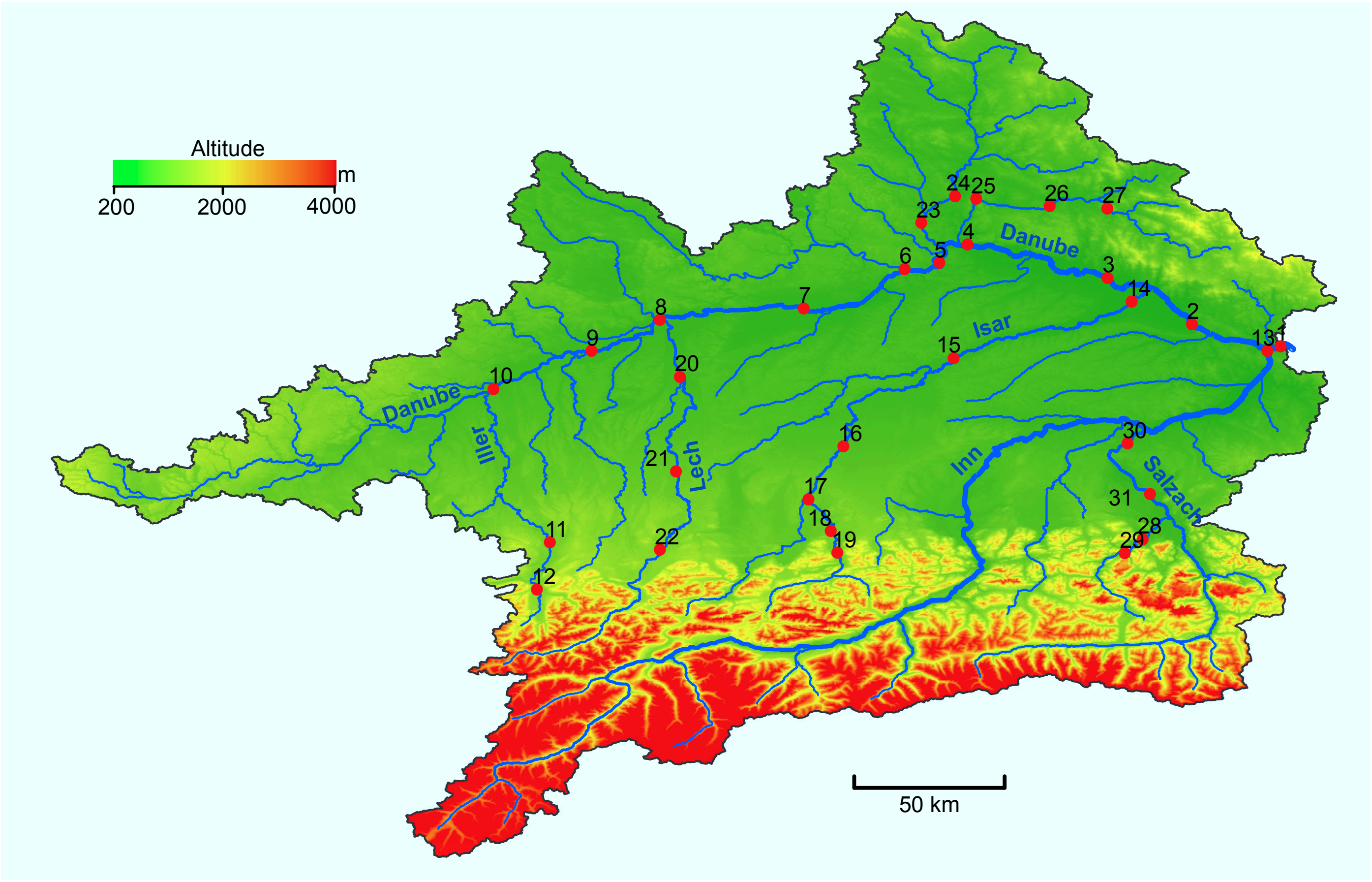
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Numerous max-stable models now exist, some more 'realistic' than others
- Particularly flexible example is the Brown–Resnick process, which takes

$$W(x) = \exp\{\varepsilon(x) - \gamma(x)\},$$

where $\varepsilon(x)$ is a stationary or intrinsically stationary Gaussian process with semi-variance or semivariogram $\gamma(x)$—can use panoply of functions γ from spatial statistics, or can invent your own.

Upper Danube Basin



Asadi, Davison, Engelke (2016) *Annals of Applied Statistics*

Extremal dependence on river network

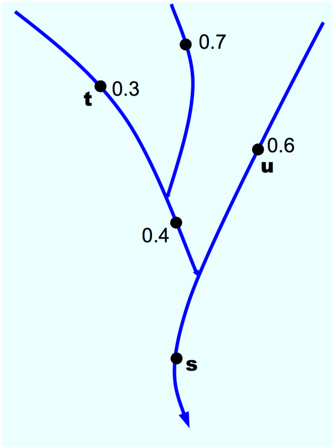
- Sources of dependence between data at locations t_1 and t_2 on the network T:
  - **flow-dependence**; t_2 is downstream of t_1, or vice versa
  - **'geo'-dependence**: the same events may impact nearby watersheds
- Overall semi-variogram

$$\gamma(s,t) = \lambda_{\mathrm{RIV}}\{1 - C_{\mathrm{RIV}}(s,t)\} + \lambda_{\mathrm{EUC}}\,\gamma_{\mathrm{EUC}}(s,t), \quad s,t \in T,$$

where $\lambda_{\mathrm{RIV}}, \lambda_{\mathrm{EUC}} > 0$.

- Flow-dependence in terms of shortest river distance $d(\cdot,\cdot)$:

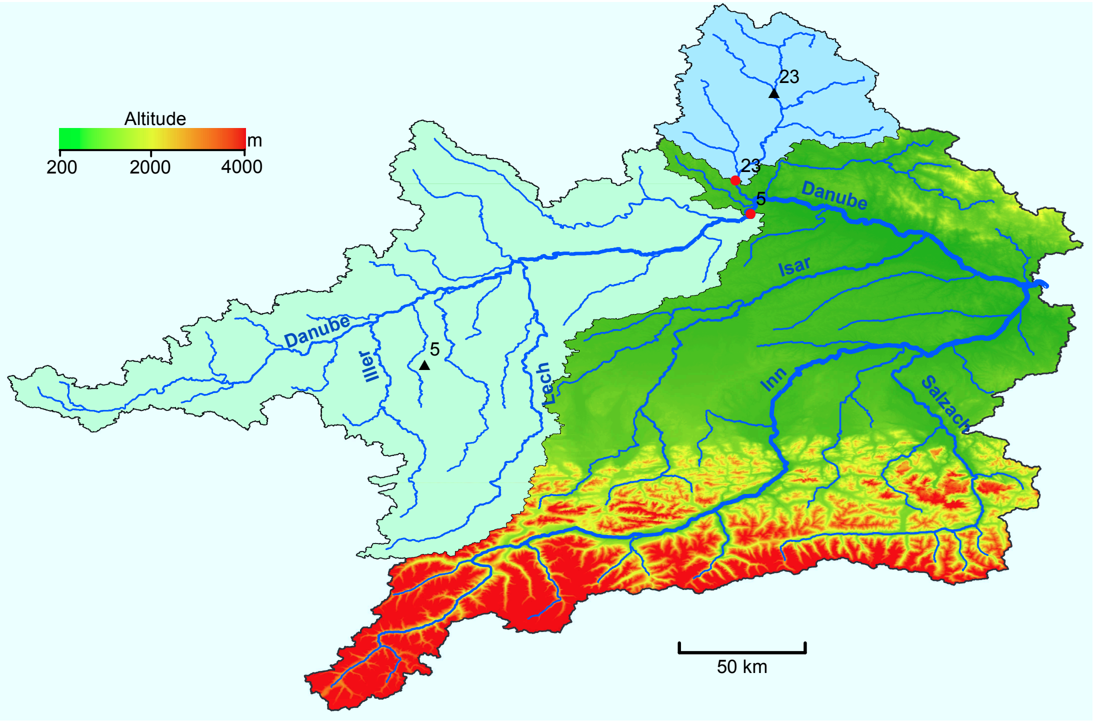
$$\begin{aligned} C_{\mathrm{RIV}}(s,u) &= C_1\{d(s,u)\} \times \sqrt{0.6}, \\ C_{\mathrm{RIV}}(s,t) &= C_1\{d(s,t)\} \times \sqrt{0.4 \times 0.3}, \\ C_{\mathrm{RIV}}(u,t) &= 0, \\ C_1(h) &= \exp(-h/\theta), \quad \theta > 0. \end{aligned}$$



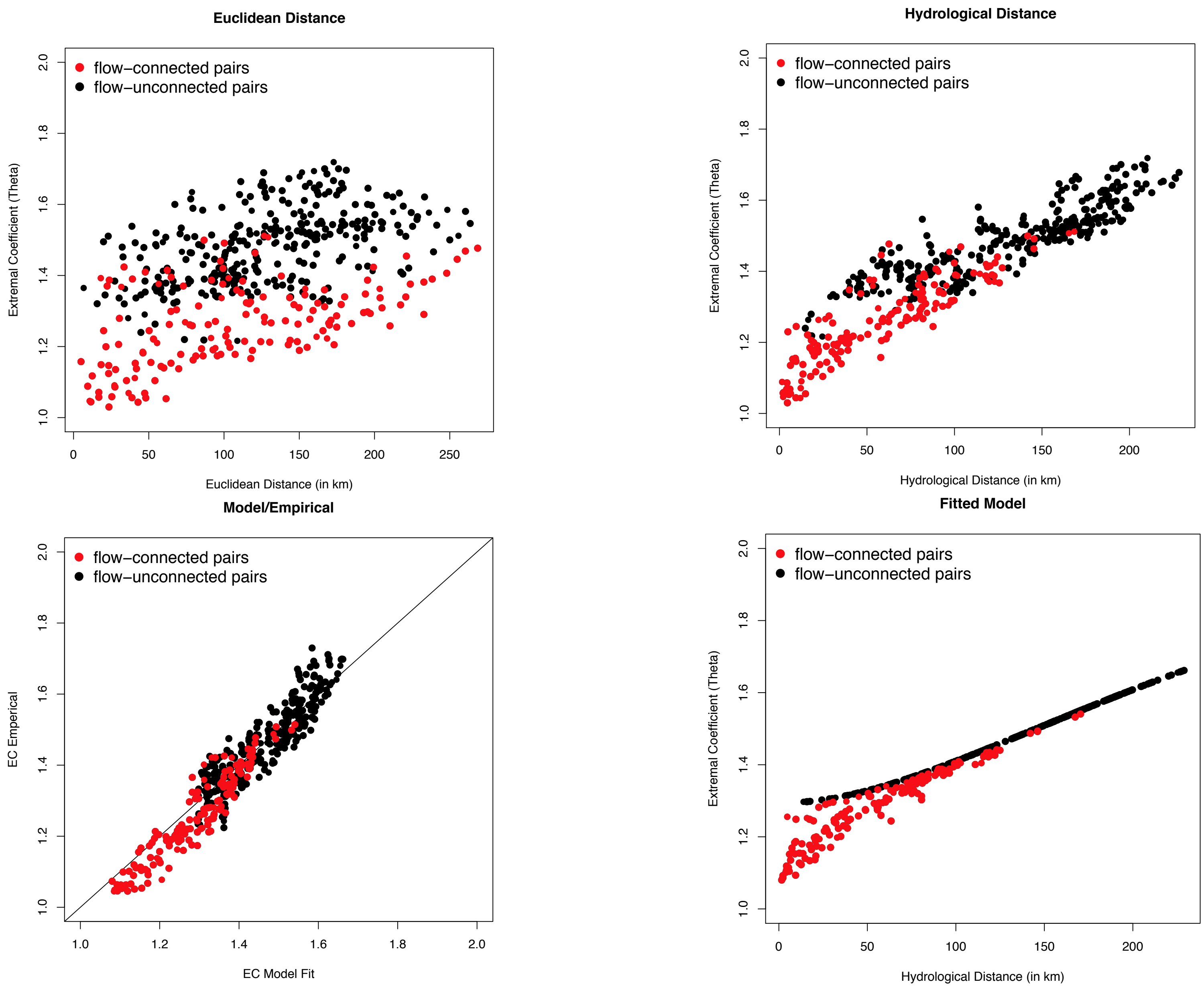
Extremal dependence on river network

- Introduce **hydrological location** of each station, as $h(s) \in \mathbb{R}^2$ as centroid of its sub-catchment, and define dependence measure

$$\gamma_{\mathrm{EUC}}(s,t) = \|h(s) - h(t)\|^{\alpha}, \quad \alpha \in (0,2].$$



Extremal coefficients



Comments

- Numerous max-stable models now exist, some more ‘realistic’ than others
- Particularly flexible example is the Brown–Resnick process, which takes

$$W(x) = \exp\{\varepsilon(x) - \gamma(x)\},$$

where $\varepsilon(x)$ is a stationary or intrinsically stationary Gaussian process with semi-variance or semivariogram $\gamma(x)$—can use panoply of functions γ from spatial statistics, or can invent your own

Comments

- Numerous max-stable models now exist, some more ‘realistic’ than others
- Particularly flexible example is the Brown–Resnick process, which takes

  $$W(x) = \exp\{\varepsilon(x) - \gamma(x)\},$$

  where $\varepsilon(x)$ is a stationary or intrinsically stationary Gaussian process with semi-variance or semivariogram $\gamma(x)$—can use panoply of functions γ from spatial statistics, or can invent your own
- Can simulate from these models both unconditionally, and conditionally on observed extremes

Comments

- Numerous max-stable models now exist, some more 'realistic' than others
- Particularly flexible example is the Brown–Resnick process, which takes

$$W(x) = \exp\{\varepsilon(x) - \gamma(x)\},$$

  where $\varepsilon(x)$ is a stationary or intrinsically stationary Gaussian process with semi-variance or semivariogram $\gamma(x)$—can use panoply of functions γ from spatial statistics, or can invent your own
- Can simulate from these models both unconditionally, and conditionally on observed extremes
- **BUT** likelihood inference is awkward, because
  - for D variables, all 2^D derivatives of V are needed, and the number of terms in the likelihood is monstrous
  - for the B–R process, the derivatives involve multivariate normal integrals, which are slow to compute

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Suppose we have independent (annual) maxima observed at $\mathcal{D} = \{x_1, \ldots, x_D\} \subset \mathcal{X}$ for n years, so the data for each year have joint distribution

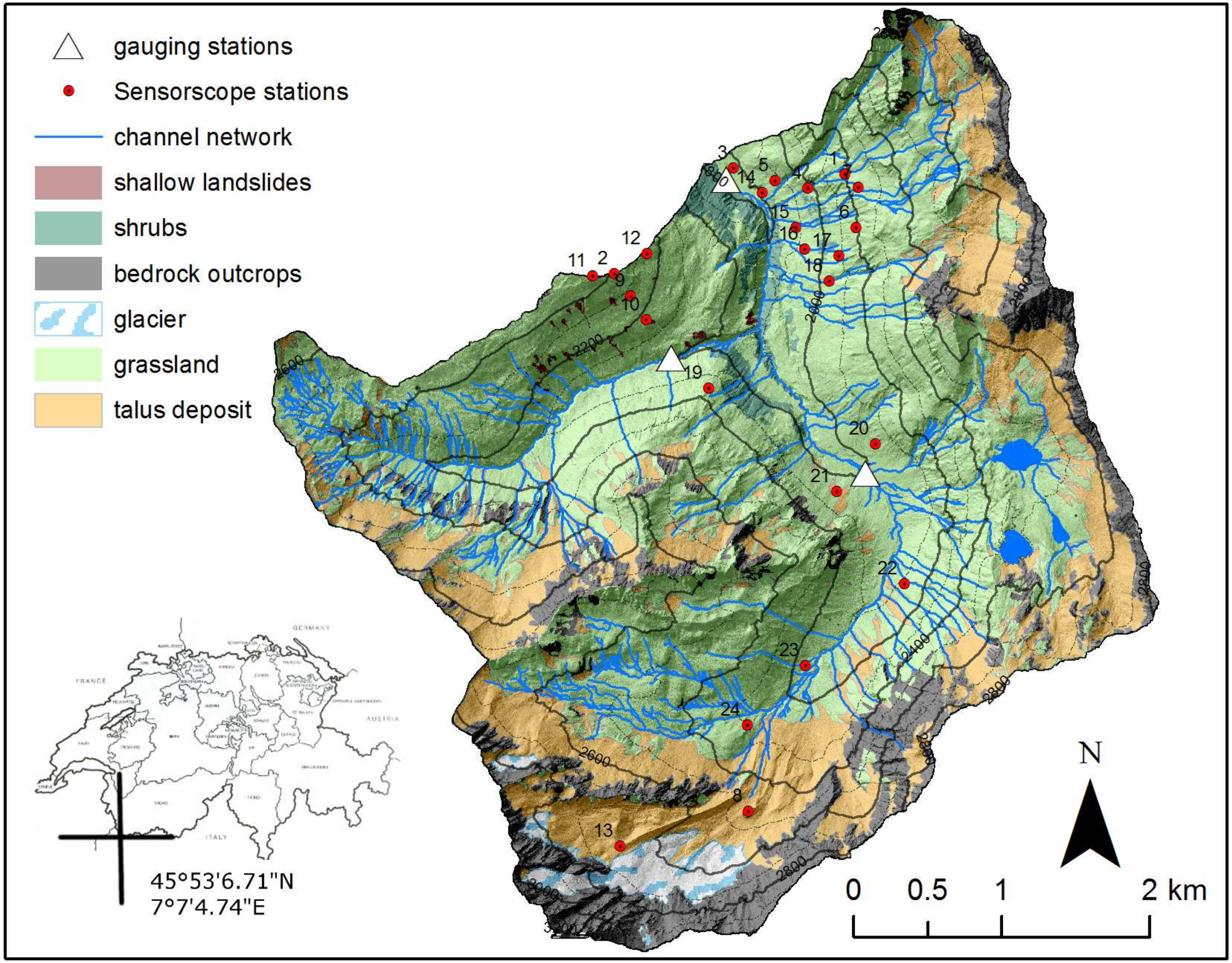
$$\mathrm{P}\{Z(x_1) \leq z_1, \ldots, Z(x_D) \leq z_D\} = \exp\left\{-V(z_1, \ldots, z_D)\right\}, \quad z_1, \ldots, z_D > 0.$$

- The formulation of the model using its CDF means that to compute the likelihood function we must differentiate e^{-V} with respect to $z_1, \ldots, z_D$, leading to combinatorial explosion:

$$-V_1 e^{-V}, \quad (V_1 V_2 - V_{12}) e^{-V}, \quad (-V_1 V_2 V_3 + V_{12} V_3 [3] - V_{123}) e^{-V}, \quad \ldots,$$

  with about 10^5 terms for $D = 10$. Clearly this is infeasible for realistic applications, so we need to avoid this, by
  - using a composite (usually a pairwise) likelihood; or
  - using the Stephenson–Tawn approach, using the timing of events to inform us which term of the partition should enter the likelihood;
  - using threshold exceedances
- Bayesian inference is possible (Emeric Thibaud's poster)
- In any case we must compute (many) derivatives of V, and sometimes integrate them ...

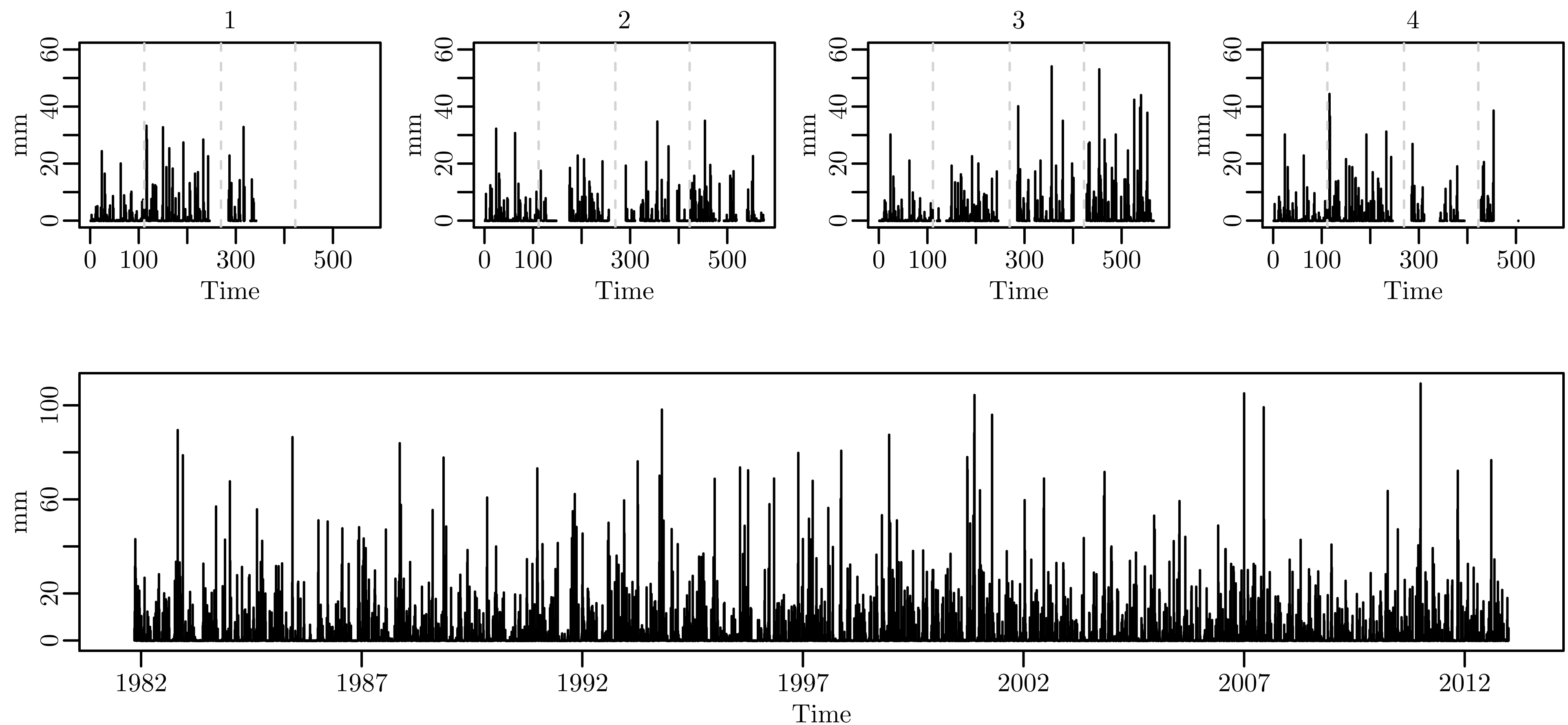
Val Ferret



gauging stations
Sensorscope stations
channel network
shallow landslides
shrubs
bedrock outcrops
glacier
grassland
talus deposit
45°53'6.71"N
7°7'4.74"E
N
0 0.5 1 2 km

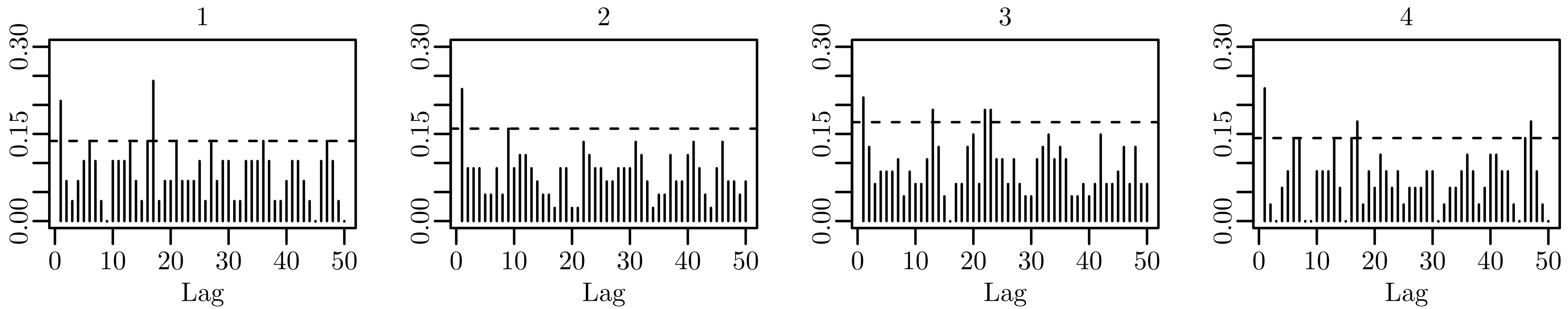
Sensorscope station

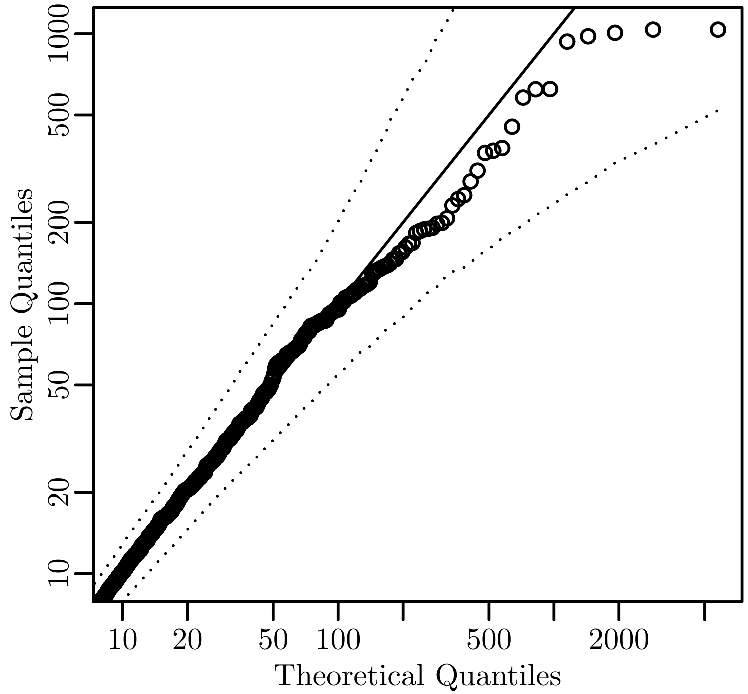
Val Ferret, daily rainfall data



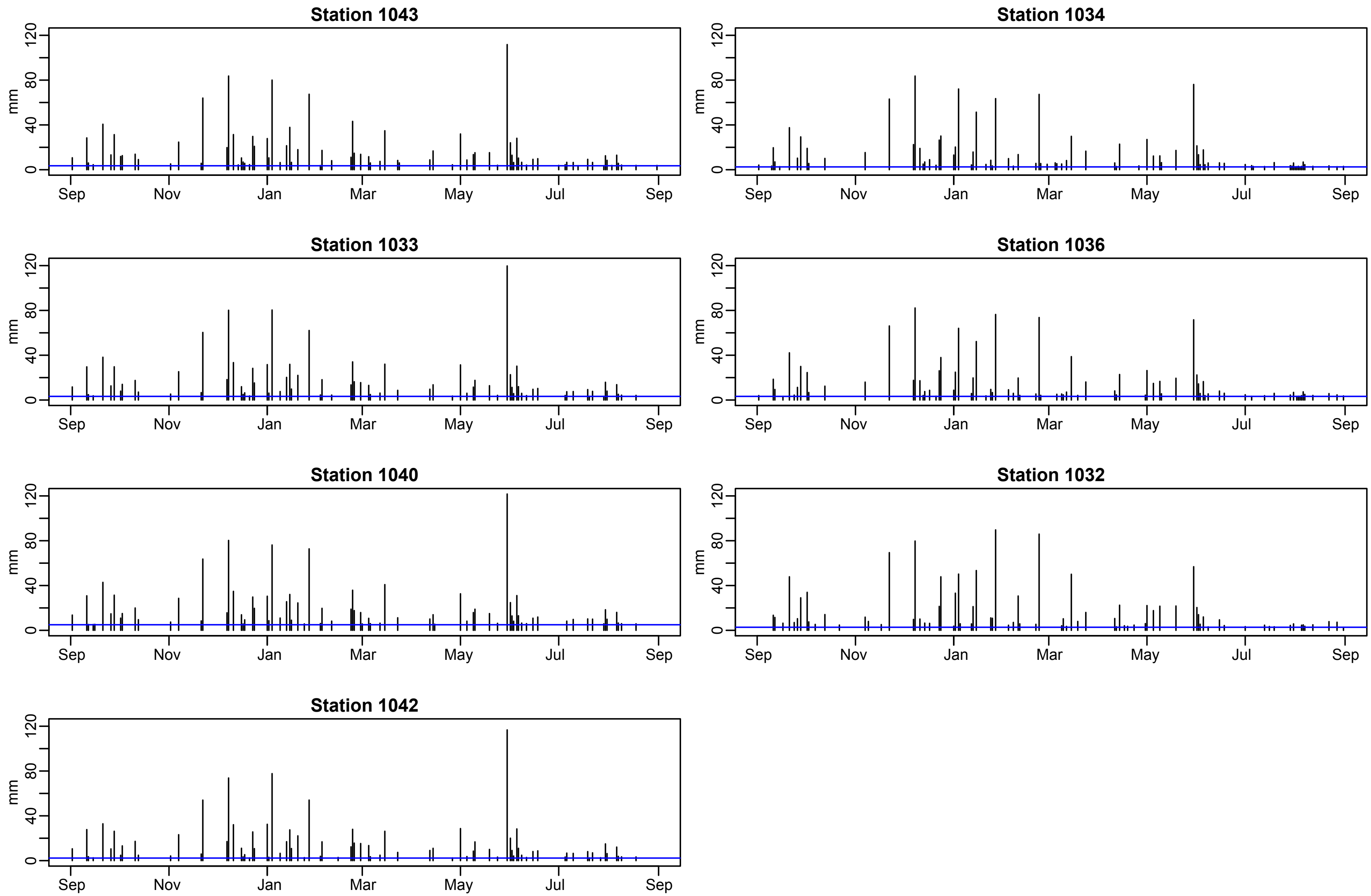
Top: Daily cumulative rainfall totals for 575 days in summers 2009 to 2012, recorded by Sensorscope stations 1–4. White spaces correspond to missing data.
Bottom: Daily cumulative rainfall totals for 31 years in summers 1982 to 2012, recorded by MétéoSuisse at the Grand St-Bernard.

Val Ferret, daily rainfall extremograms, marginal model fit





Thresholding daily rain data

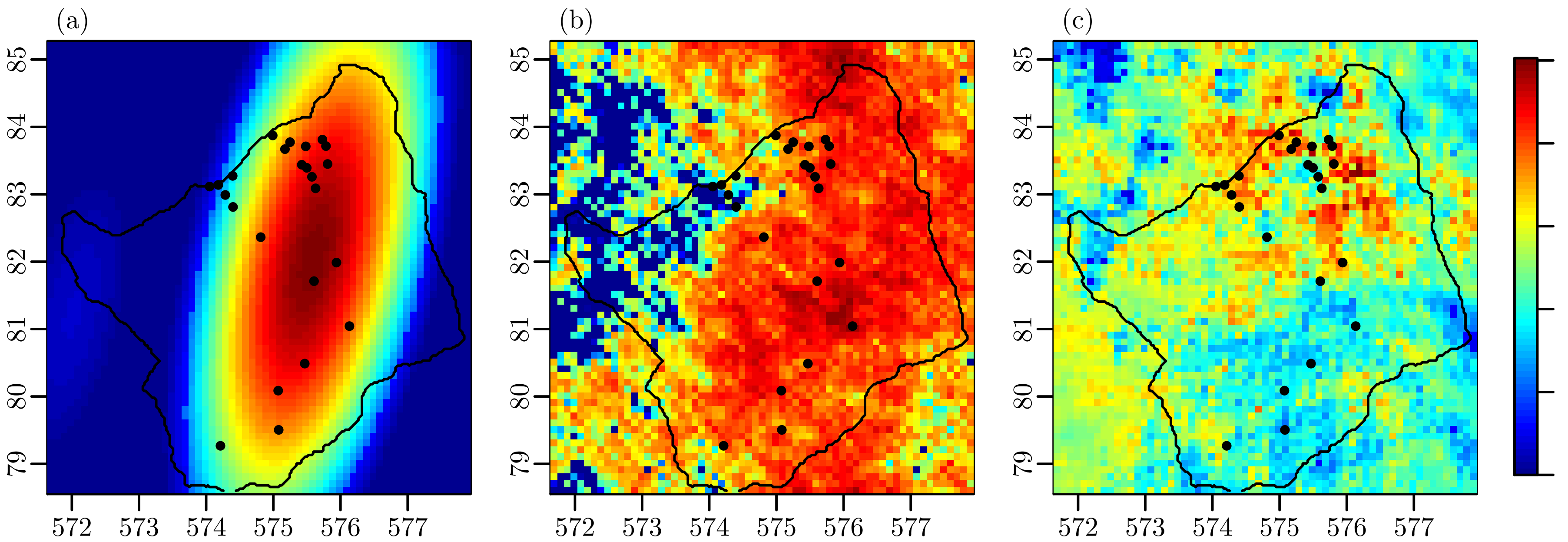


Results for Val Ferret data

	τ	τ_{GSB}	ξ	Dependence parameters			$\widehat{\ell}^*$	CLIC*
Max-stable models								
Smith	7.48	4.69	0.11	0.48	2.94	0.57	−6122	12264
Schlather	7.34	5.17	0.09		10.63	0.52	−6035	12087
B–R	6.89	5.46	0.10		5.82	0.46	−6036	12092
Asymptotic independence models								
Schlather	7.85	6.42	0.04		186	0.60	−6023	12065
B–R	7.79	6.48	0.04		113	0.57	−6027	12074
GC	7.53	6.21	0.06		103	0.50	−6029	12118

- For the Smith model, $\theta = (\tau, \tau_{\text{GSB}}, \xi, \Sigma_{11}, \Sigma_{22}, \Sigma_{12})$, where Σ is the variance matrix of the underlying bivariate Gaussian density.
- For the Brown–Resnick (B–R), Schlather, and Gaussian copula (GC) models, $\theta = (\tau, \tau_{\text{GSB}}, \xi, \lambda, \kappa)$. The units of τ, τ_{GSB} and the range parameter are respectively mm, mm and km.
- The likelihood maximisation fails for the inverted Smith model.

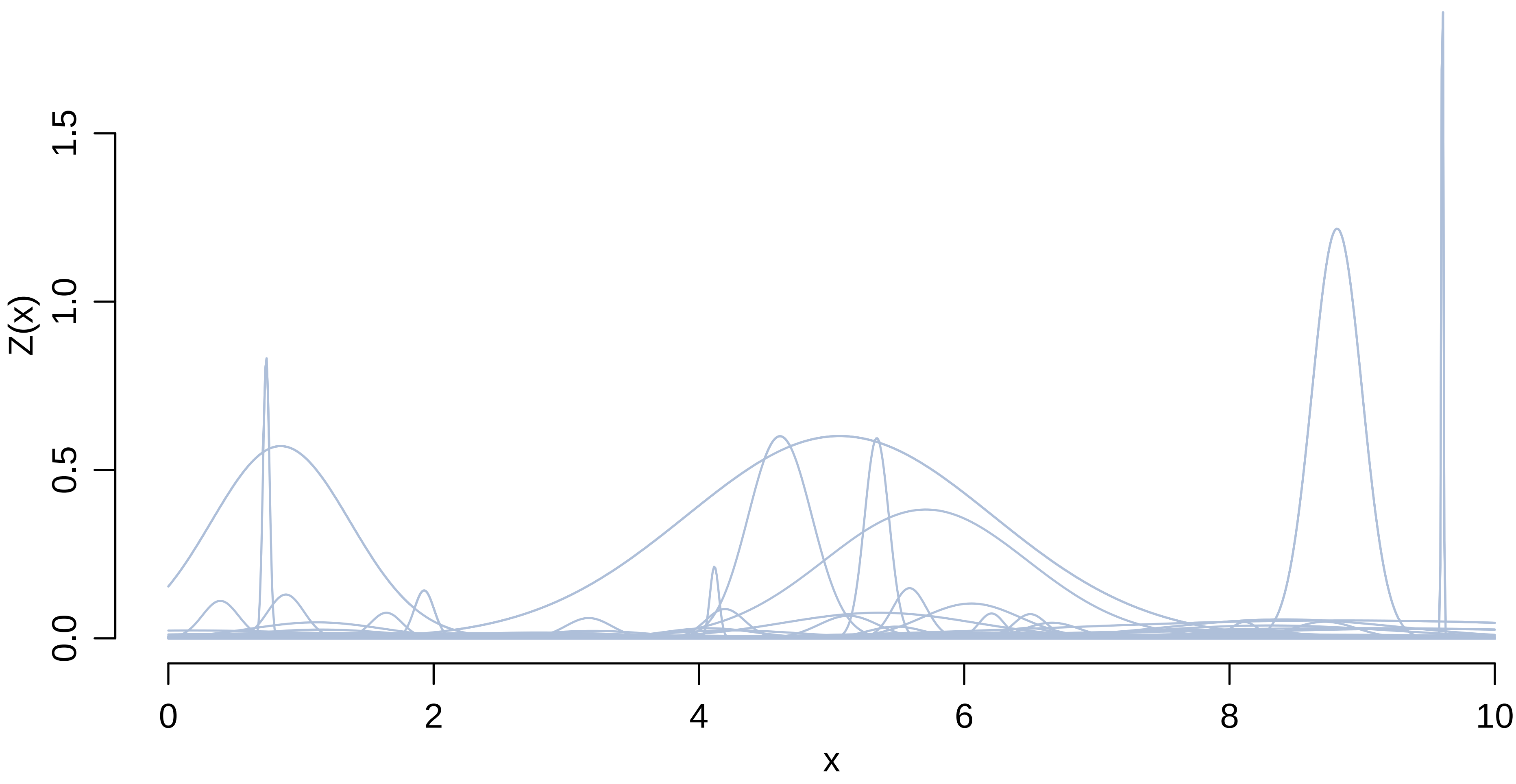
Simulated extreme rain at Val Ferret



Simulation of max-stable random fields, on the original data scale (mm), from the fitted (a) Smith and (b) Schlather models and (c) an inverted max-stable process based on the Schlather model. Black dots show the locations of the 24 stations.

Threshold exceedances

Individual events



Brown–Resnick model

- The individual events of the max-stable process are (say)

$$Q(x) = RW(x), \quad x \in \mathcal{X},$$

  where $R \sim$ a Poisson process with intensity $1/r^2$ on $r > 0$, and $W(x)$ is log-Gaussian, so we can get an explicit intensity for the Poisson 'event' $Q(x)$ observed to take values $z_1, \ldots, z_D$ at points $x_1, \ldots, x_D$

- This intensity may have to be applied to values that are not extreme, for which the extremal model will be poor.

- If $z_d > u$ for a subset $\mathcal{C}$ of $\{1, \ldots, D\}$, with $C = |\mathcal{C}|$, and that $z_d < u$ for the remaining subset $\mathcal{C}'$, we end up with a censored likelihood contribution

$$\frac{1}{z_1^2 z_2 \cdots z_C} \phi_{C-1}(\log \tilde{z}_{\mathcal{C}}; \tilde{\Omega}_{\mathcal{C},\mathcal{C}}) \Phi_{D-C}\left(\tilde{\mu}_{\mathcal{C}'|\mathcal{C}}; \tilde{\Omega}_{\mathcal{C}'|\mathcal{C}}\right),$$

  where ϕ_k and Φ_k denote the k-dimensional normal density and distribution functions, and for $c, d \in \{1, \ldots, D\}$, $\tilde{\Omega}_{c,d} = \frac{1}{2}\{\Omega_{c,1} + \Omega_{1,d} - \Omega_{c,d}\}$ and $\log \tilde{z}_d = \log z_d - \log z_1 + \Omega_{d,1}/2$, and with $\mathcal{C} = \{2, \ldots, C\}$, and $\mathcal{C}' = \{C+1, \ldots, D\}$,

$$\mu_{\mathcal{C}'|\mathcal{C}} = (\log u - \log z_1 + \tfrac{1}{2}\Omega_{1,\mathcal{C}'}) - \tilde{\Omega}_{\mathcal{C}',\mathcal{C}} \tilde{\Omega}_{\mathcal{C},\mathcal{C}}^{-1} \tilde{z}_{\mathcal{C}}, \quad \tilde{\Omega}_{\mathcal{C}'|\mathcal{C}} = \tilde{\Omega}_{\mathcal{C}',\mathcal{C}'} - \tilde{\Omega}_{\mathcal{C}',\mathcal{C}} \tilde{\Omega}_{\mathcal{C},\mathcal{C}}^{-1} \tilde{\Omega}_{\mathcal{C},\mathcal{C}'}.$$

- Nasty but manageable for D not too big, extends to extremal-t processes.

Exceedances and risk functions

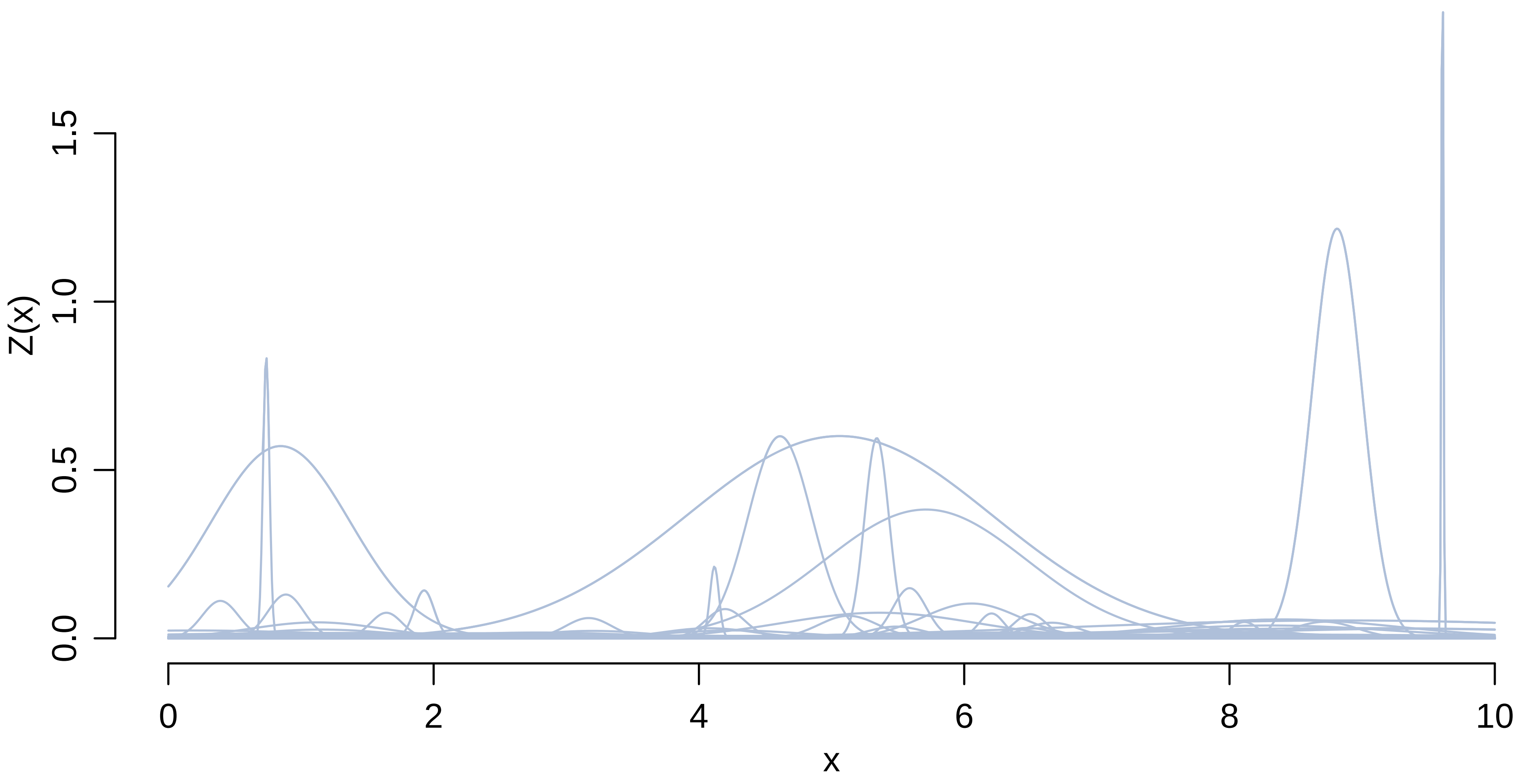
- Modelling threshold exceedances is widely used in (scalar) practice:
  - more flexible than using maxima
  - statistically more efficient, makes better use of data
- For scalar data, choosing rare events is easy: either they're big or they're small.
- For multivariate data, we need to say what 'direction' is extreme
- Do this via a scalar **risk function** f applied to the individual events $Q_j(x) = R_j W_j(x)$ of the max-stable process
  - Choose those events Q_j for which $f(Q_j)$ exceeds a threshold u
  - **Red**: extremes on $[0, 2]$, selected using risk function

$$f(Q) = \int_0^2 Q(x)\, dx$$

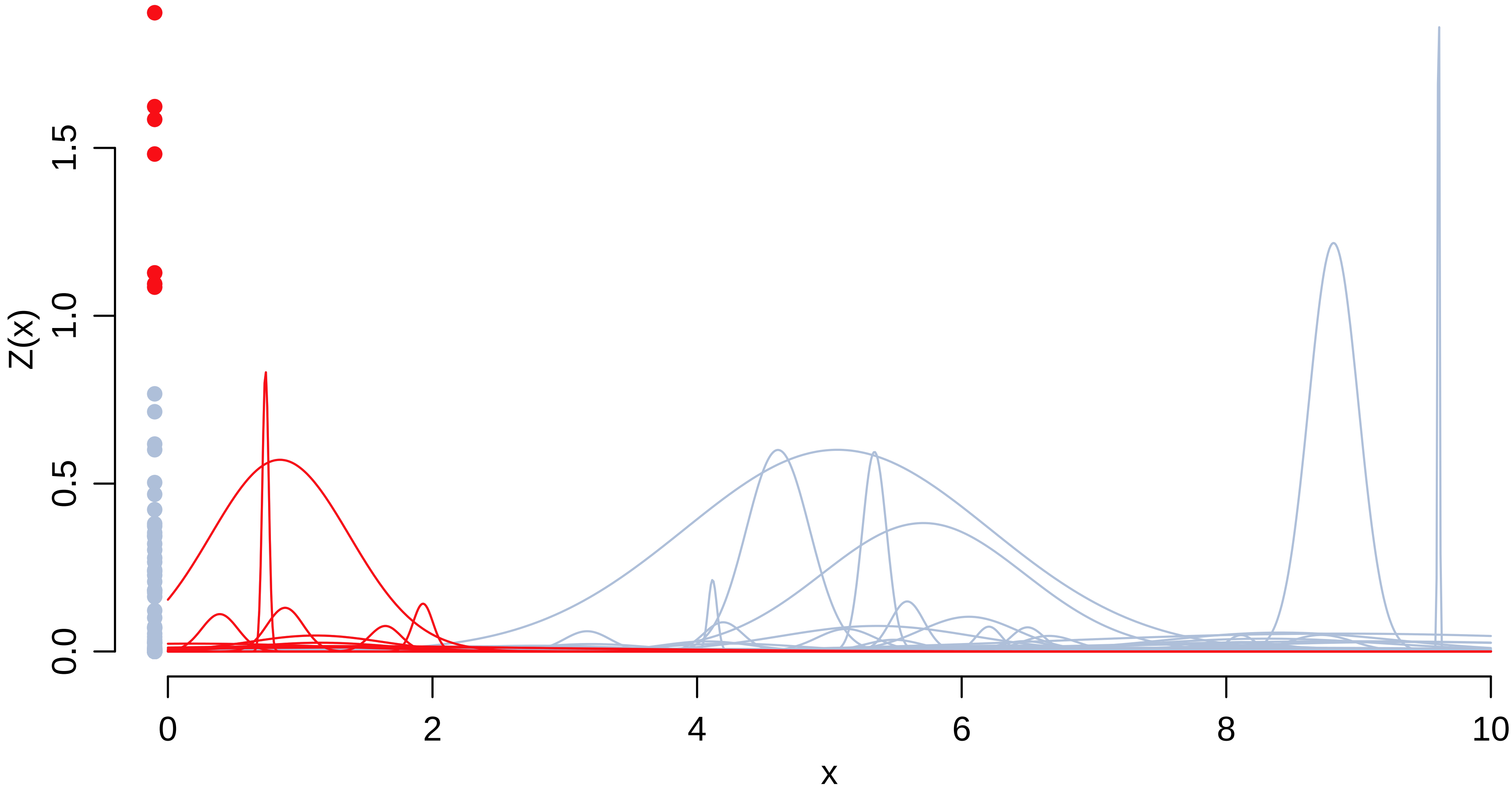
  - **Blue**: most intense events, selected using risk function

$$f(Q) = \max Q(x)$$

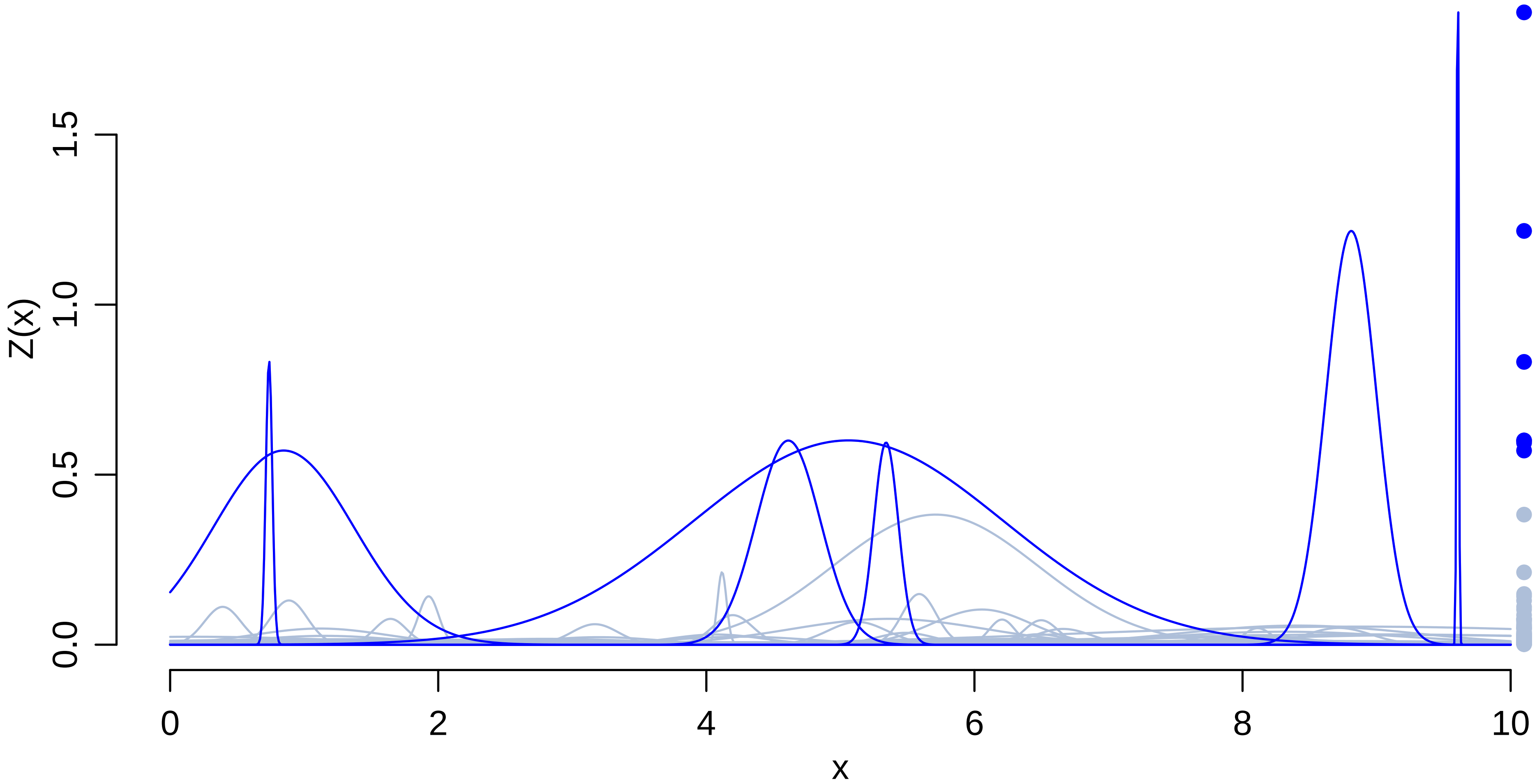
Individual events



Extremes in $[0, 2]$



Highest peaks anywhere

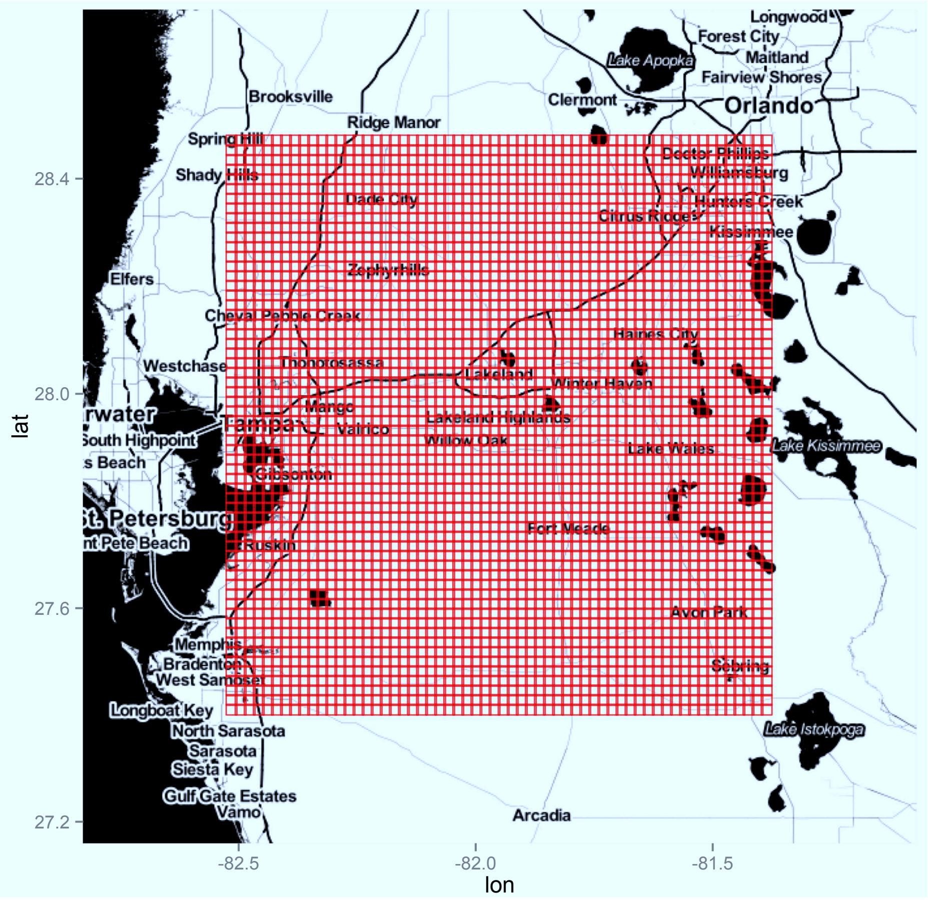


Inference

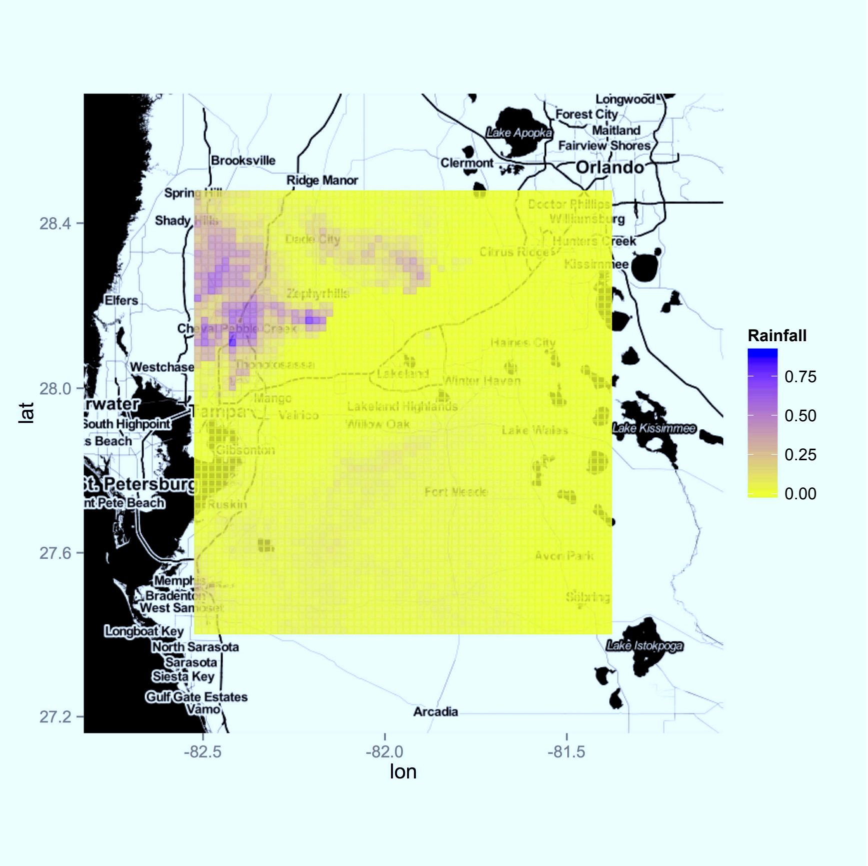
- Fitting for the 'exceedances' Q_j is (in principle) much easier than for the max-stable process $Z(x)$:
  - likelihoods can be constructed, but
  - they still involve lots of computationally expensive integrals
- Fixes
  - estimate the integrals using quasi-Monte Carlo or other methods,
  - avoid likelihood inference, using the gradient score
- Big problems ($D \approx 1000$s) feasible

Extreme rainfall over Florida

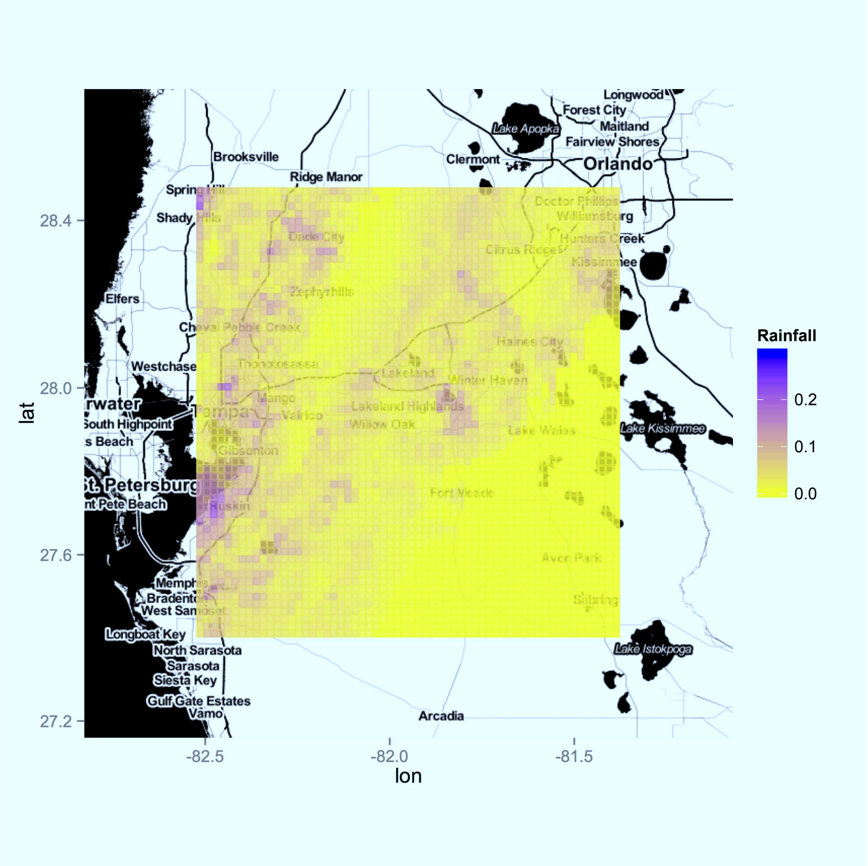
- 15-minute radar rainfall measurements over Florida from 1994–2010
- We focus on a 120 km$\times$ 120 km square south-west of Orlando and on the wet season, i.e., June to September.



Florida rainfall



Local rainfall



Spatially dispersed rainfall

Marginal modelling

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Generalized Pareto distributions are fitted for each of the locations s_i using exceedances over the 99 percentile.
- A model with common shape parameter $\xi_0 = 0.124$ is retained.
- Margins are then transformed to unit Fréchet:

$$X^*(s_i) = -1/\log \tilde{F}_i\{X(s_i)\},$$

where

$$\tilde{F}_i\{X(s_i)\} = \begin{cases} \widehat{F}_i\{X(s_i)\}, & X(s_i) \leqslant q_{99}(s_i), \\ 1 - G_{\{\xi_0, \widehat{\sigma}(s_i), q_{99}(s_i)\}}\{X(s_i)\}, & X(s_i) > q_{99}(s_i), \end{cases}$$

and

– $\widehat{F}_i$ is the empirical cumulative distribution function at location s_i,

– $G_{\{\xi_0, \widehat{\sigma}(s_i), q_{99}(s_i)\}}$ is the distribution function of a generalized Pareto random variable with shape ξ_0, scale $\widehat{\sigma}(s_i)$ and location $q_{99}(s_i)$.

Risks functionals

- We define two risk functionals

$$f_{\max}(X^*) = \left[\sum_{i=1}^{\ell} \{X^*(s_i)\}^{20}\right]^{1/20},$$
$$f_{\text{sum}}(X^*) = \left[\sum_{i=1}^{\ell} \{X^*(s_i)\}^{\xi_0}\right]^{1/\xi_0},$$

  where $\ell = 3600$ is the number of grid cells and ξ_0 is the shape parameter of the marginal model.
- $f_{\max}$ is a continuous and differentiable approximation of $\max_{i=1,\ldots,\ell} X^*(s_i)$ which satisfies the requirements for the gradient score.
- f_{sum} selects events with large spatial cover. The power ξ_0 approximately transforms the data X^* back to a scale where summing observations has a physical meaning.

Spatial model and parameter estimates

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Non-separable semi-variogram model

$$\gamma(s_i, s_j) = \left\| \frac{\Omega(s_i - s_j)}{\tau} \right\|^{\kappa}, \quad s_i, s_j \in [0, 120]^2, \quad i, j \in \{1, \ldots 3600\},$$

with $0 < \kappa \leqslant 2, \tau > 0$ and anisotropy matrix

$$\Omega = \begin{bmatrix} \cos\eta & -\sin\eta \\ a\sin\eta & a\cos\eta \end{bmatrix}, \quad \eta \in \left(-\frac{\pi}{2}; \frac{\pi}{2}\right], \quad a > 1.$$

- Fitted parameters obtained for both risk functionals with exceedances of $f_{\max}(X^*)$ and $f_{\text{sum}}(X^*)$ over the 99 quantile:

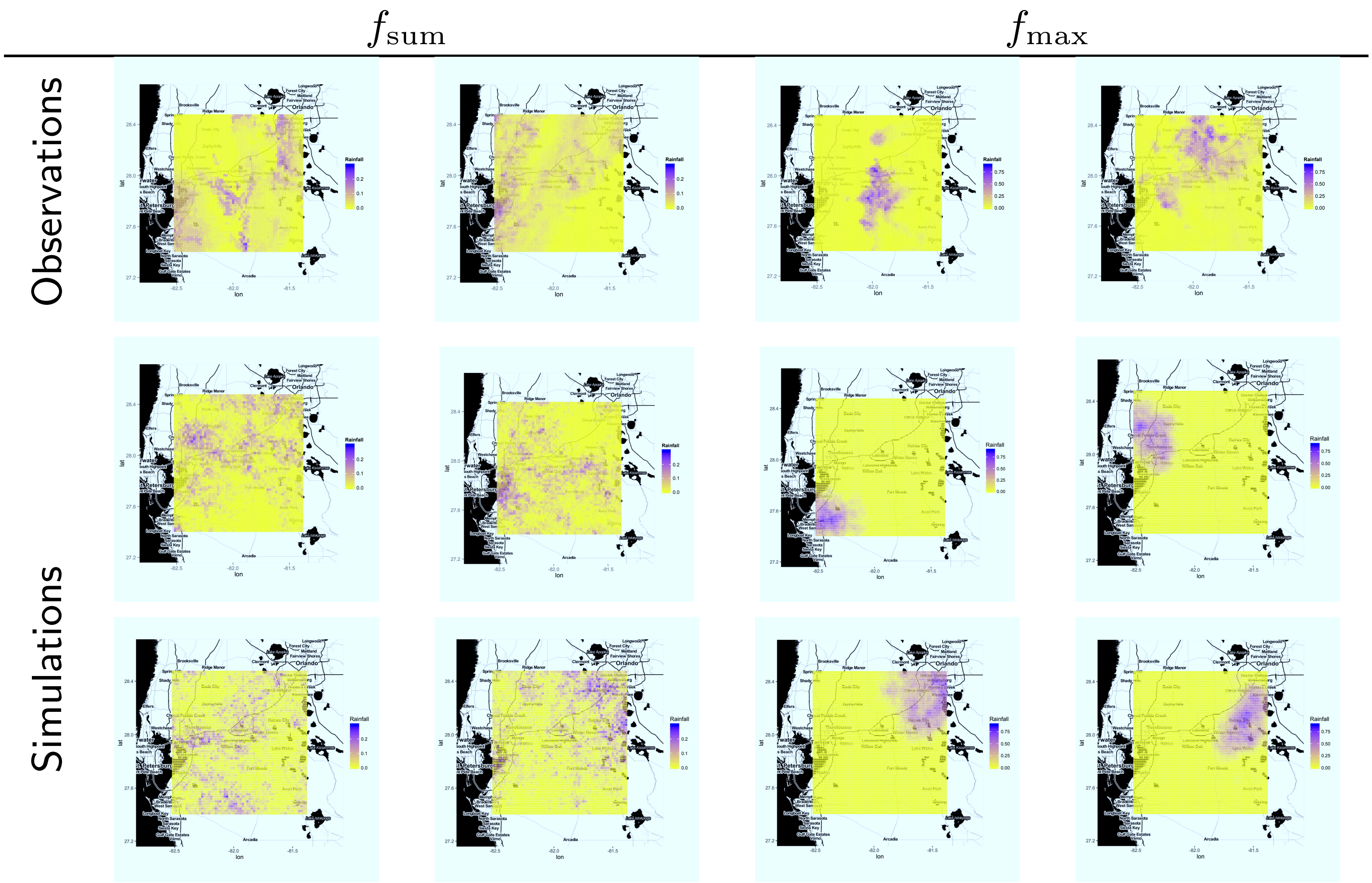
	κ	τ	η	a
$f_{\max}$	$1.192_{0.02}$	$9.06_{0.19}$	$0.08_{0.61}$	$1.008_{0.005}$
f_{sum}	$0.326_{0.007}$	$46.67_{0.018}$	$-0.30_{0.10}$	$1.064_{0.017}$

  – $f_{\max}$ estimates are quite smooth with a small scale, they capture high quantiles and induce a model similar to that in earlier work.

  – For f_{sum}, the semi-variogram is rougher but with a much larger scale, which is consistent with large-scale events.

  – Anisotropy does not seem significant.

Simulated extreme rainfall



15-minute cumulated rainfall (inches): observed (first row) and simulated (second and third rows) for the risk functionals f_{sum} (left) and f_{max} (right) with intensity equivalent to the 0.99 quantile.

Closing

Closing

- Care needed with extrapolation of scalar and multivariate extremes
- Basic ideas on maxima and point processes extend to spatial and space-time settings.
- Max-stable processes give asymptotic dependence models, but asymptotic independence models exist
- Can fit either using pairwise likelihood (can be inefficient), or, for 'exceedances', full likelihood ($D \leq 30$, say).
- Model-checking possible, using simulation from fitted models, and extensions of previous ideas (e.g., χ and $\overline{\chi}$).
- Current 'hot' research area: lots going on (e.g., threshold models, non-stationarity, gridded data, non-Euclidean spaces, ...).

Some reading

- Coles (2001), *Introduction to the Statistical Modeling of Extreme Values*, Springer
- de Haan and Ferreira (2006) *Extreme Value Theory: An Introduction*, Springer
- Davison and Huser (2015) *Annual Review of Statistics and its Applications*

- de Haan (1984) *Annals of Probability*
- Smith (1990) unpublished
- Schlather (2002) *Extremes*
- Süveges and Davison (2012) *European Physical Journal Special Topics*
- Davison, Padoan and Ribatet (2012) *Statistical Science*
- Wadsworth and Tawn (2012) *Biometrika*
- Thibaud, Mutzner and Davison (2013), *Water Resources Research*
- Huser and Davison (2014) *J. R. Statist. Soc., series B*
- Wadsworth and Tawn (2015) *Biometrika*
- Thibaud and Opitz (2015) *Biometrika*
- Asadi, Davison and Engelke (2016) *Annals of Applied Statistics*
- de Fondeville and Davison (2016) arXiv