Probabilities of concurrent extremes

C. Dombry[†], M. Ribatet[‡], S. Stoev[*]

[†]Department of Mathematics, University of Franche-Comté, Besançon, FRANCE
[‡]Department of Mathematics, University of Montpellier, Montpellier, FRANCE
[*]Department of Statistics, University of Michigan, Ann Arbor, USA

1. Introduction

Definition 1. The process $\{\eta(x)\colon x \in \mathcal{X}\}$ is said to be max-stable if for all $n \geq 1$ there exist continuous normalizing functions $a_n(\cdot) > 0$ and $b_n(\cdot) \in \mathbb{R}$ such that

$$\left\{ \frac{\max_{i=1,\dots,n} \eta_i(x) - b_n(x)}{a_n(x)} : x \in \mathcal{X} \right\} \stackrel{\mathrm{d}}{=} \{\eta(x)\colon x \in \mathcal{X}\}$$

where $\eta_1, \dots, \eta_n$ are independent copies of the process $\{\eta(x)\colon x \in \mathcal{X}\}$.

Remark. Throughout this talk we will assume that $\mathcal{X} \subset \mathbb{R}^d$, $d \geq 1$, is compact and that all stochastic processes have continuous sample paths.

Theorem 1. *(de Haan and Fereira, 2006)*
Let $\{X_i(x) \colon x \in \mathcal{X}, i \geq 1\}$ *be a sequence of independent copies of a stochastic process* $\{X(x) \colon x \in \mathcal{X}\}$*. If there exist sequences of normalizing functions* $\{c_n(x) > 0 \colon x \in \mathcal{X}, n \geq 1\}$ *and* $\{d_n(x) \in \mathbb{R} \colon x \in \mathcal{X}, n \geq 1\}$ *then, provided the limiting process is non degenerate,*

$$\left\{ \frac{\max_{i=1,\dots,n} X_i(x) - d_n(x)}{c_n(x)} \colon x \in \mathcal{X} \right\} \stackrel{\mathrm{d}}{\longrightarrow} \{\eta(x) \colon x \in \mathcal{X}\},$$

as $n \to \infty$*, it has to be a max-stable process.*

- The finite dimensional distributions are multivariate extreme value distributions and, in particular, $\eta(x) \sim \text{GEV}$, $x \in \mathcal{X}$.
- If $\{\eta(x) \colon x \in \mathcal{X}\}$ has unit Fréchet margins, i.e., $\Pr\{\eta(x) \leq z\} = \exp(-1/z)$, $z > 0$, we say that it is a simple max-stable process.

Theorem 2. *(de Haan, 1984; Penrose, 1992)*
Any simple max-stable process $\{\eta(x) \colon x \in \mathcal{X}\}$ *can be represented as follows*

$$\{\eta(x) \colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \left\{ \max_{\varphi \in \Phi} \varphi(x) \colon x \in \mathcal{X} \right\},$$

where Φ *is a Poisson point process on* $\mathbb{C}_0 = \mathbb{C}\{\mathcal{X}, [0, \infty)\} \setminus \{0\}$ *with intensity measure*

$$\Lambda(A) = \int_0^\infty \Pr(\zeta Y \in A) \zeta^{-2} d\zeta, \qquad A \subset \mathbb{C}_0 \text{ Borel set},$$

and where $\{Y(x) \colon x \in \mathcal{X}\}$ *is a non negative stochastic process such that* $\mathbb{E}\{Y(x)\} = 1$, $x \in \mathcal{X}$ *and* $\mathbb{E}\left\{\sup_{x \in \mathcal{X}} Y(x)\right\} < \infty$.

Popcorn time. . .

- The way the atoms of Φ contribute to $\{\eta(x) \colon x \in \mathcal{X}\}$ at locations $x_1, \ldots, x_k$ defines a hitting scenario.

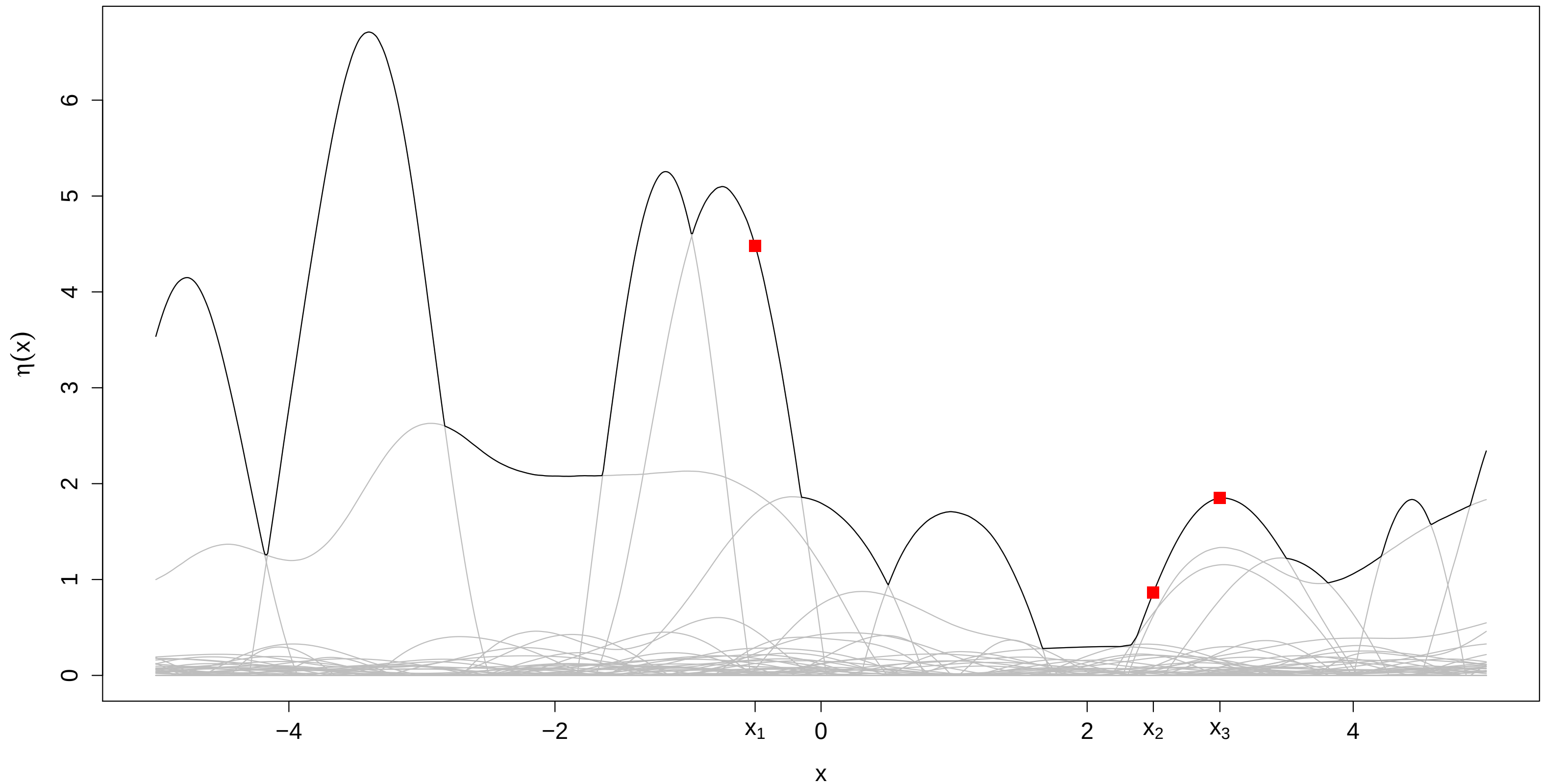

Figure 1: *Illustration of the notion of a hitting scenario. Here the hitting scenario is* $\tau = \{\{x_1\}, \{x_2, x_3\}\}$.

Hitting scenario

- The way the atoms of Φ contribute to $\{\eta(x) \colon x \in \mathcal{X}\}$ at locations $x_1, \dots, x_k$ defines a hitting scenario.

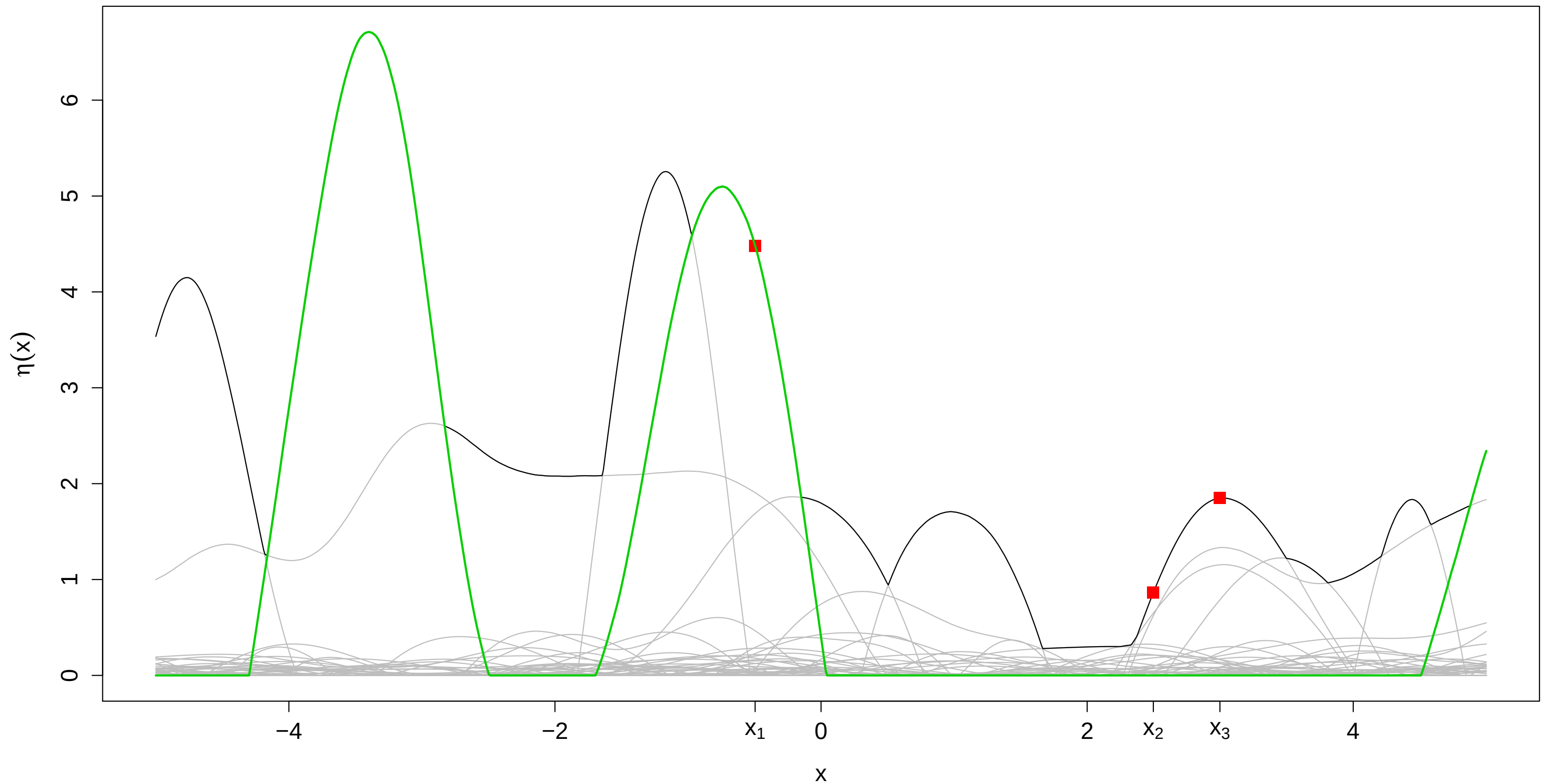

Figure 1: *Illustration of the notion of a hitting scenario. Here the hitting scenario is* $\tau = \{\{x_1\}, \{x_2, x_3\}\}$.

- The way the atoms of Φ contribute to $\{\eta(x) \colon x \in \mathcal{X}\}$ at locations $x_1, \ldots, x_k$ defines a hitting scenario.

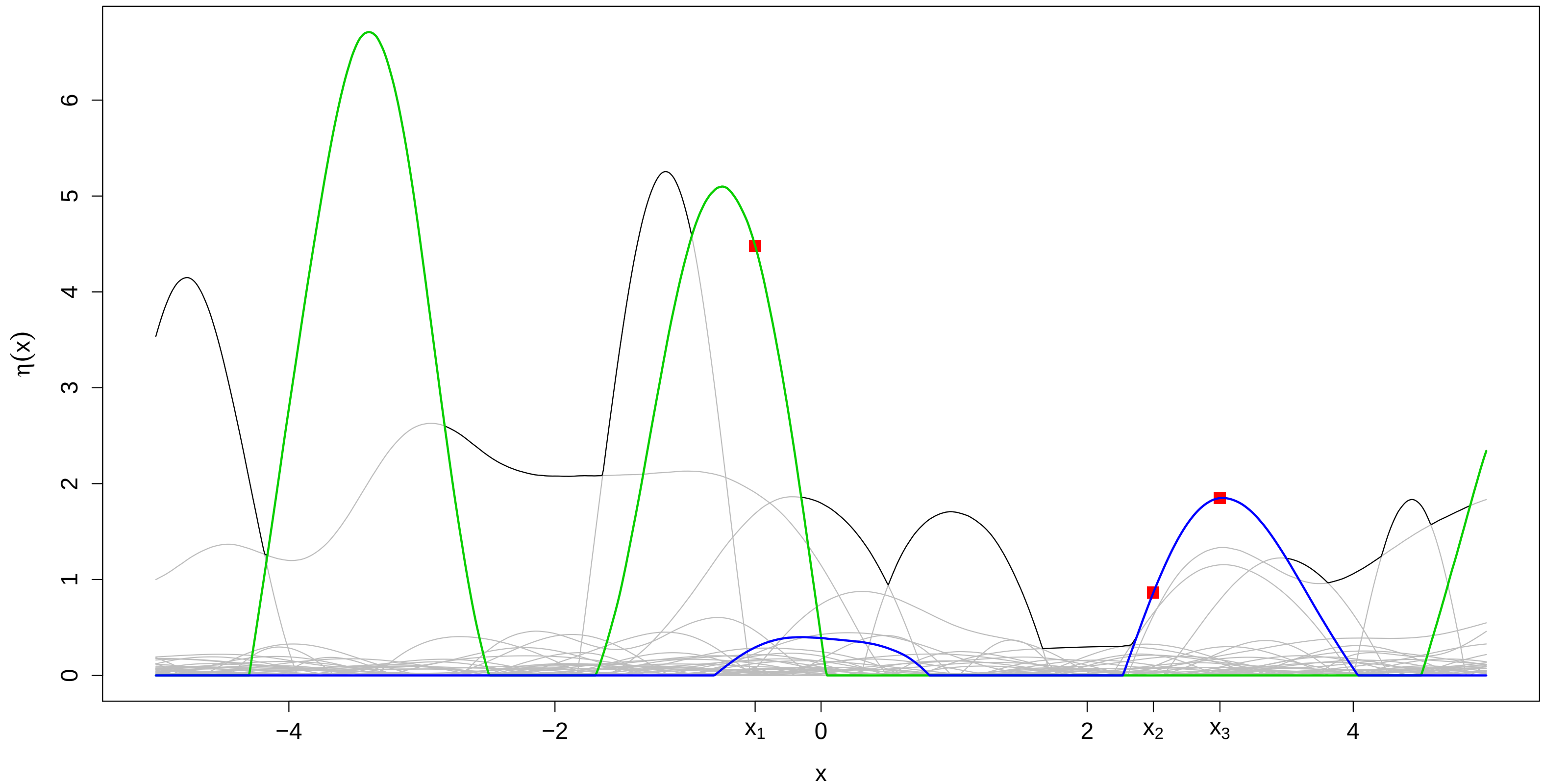

Figure 1: *Illustration of the notion of a hitting scenario. Here the hitting scenario is* $\tau = \{\{x_1\}, \{x_2, x_3\}\}$.

2. Probabilities of concurrent extremes

Definition 2. Let $\{X_i(x)\colon x \in \mathcal{X}, i = 1, \ldots, n\}$ be independent copies of a stochastic process $\{X(x)\colon x \in \mathcal{X}\}$—with continuous margins. Extremes are said sample concurrent at locations $(x_1, \ldots, x_k)$, $k \geq 2$, if there exists $\ell \in \{1, \ldots, n\}$ such that

$$\max_{i=1,\ldots,n} X_i(x_j) = X_\ell(x_j), \qquad \text{for all } j = 1, \ldots, k.$$

The associated sample concurrence probability is

$$p_n(x_1, \ldots, x_k) = \Pr\{\text{sample concurrence occurs at } (x_1, \ldots, x_k)\}.$$

- It is not difficult to show that

$$p_n(x_1, \ldots, x_k) = n\mathbb{E}\left[F\left\{X(x_1), \ldots, X(x_k)\right\}^{n-1}\right].$$

Definition 3. Let $\{\eta(x) \colon x \in \mathcal{X}\}$ be a (simple) max-stable process with spectral characterization $\eta(\cdot) = \max_{\varphi \in \Phi} \varphi(\cdot)$. Extremes are said extremal concurrent at location $(x_1, \ldots, x_k)$, $k \geq 2$, if there exists $\ell \geq 1$ such that

$$\eta(x_j) = \varphi_\ell(x_j), \qquad \text{for all } j = 1, \ldots, k.$$

The associated extremal concurrence probability is

$$p(x_1, \ldots, x_k) = \Pr\{\text{extremal concurrence occurs at } (x_1, \ldots, x_k)\}.$$

Definition 3. Let $\{\eta(x) \colon x \in \mathcal{X}\}$ be a (simple) max-stable process with spectral characterization $\eta(\cdot) = \max_{\varphi \in \Phi} \varphi(\cdot)$. Extremes are said extremal concurrent at location $(x_1, \dots, x_k)$, $k \geq 2$, if there exists $\ell \geq 1$ such that

$$\eta(x_j) = \varphi_\ell(x_j), \qquad \text{for all } j = 1, \dots, k.$$

The associated extremal concurrence probability is

$$p(x_1, \dots, x_k) = \Pr\{\text{extremal concurrence occurs at } (x_1, \dots, x_k)\}.$$

Remark. Extremal concurrence for $(x_1, \dots, x_k) \iff$ the hitting scenario is $\tau = \{x_1, \dots, x_k\}$.

Connection between the two types of concurrence

Proposition 1. *Let* $\{X(x) \colon x \in \mathcal{X}\}$ *be a stochastic process that belongs to the max-domain of attraction of some max-stable process* $\{\eta(x) \colon x \in \mathcal{X}\}$*. Then for all* $x_1, \dots, x_k \in \mathcal{X}$*,* $k \geq 2$*,*

$$p_n(x_1, \dots, x_k) \longrightarrow p(x_1, \dots, x_k), \qquad n \to \infty.$$

Remark. Actually we can show a bit more than the above:

the sample hitting scenario converges to the extremal one.

Theorem 3. *Let $\{Y(x) \colon x \in \mathcal{X}\}$ and $\{\tilde{Y}(x) \colon x \in \mathcal{X}\}$ be two independent copies of the process appearing in the spectral characterization. Then for all $x_1, \ldots, x_k \in \mathcal{X}$, $k \geq 2$,*

$$p(x_1, \ldots, x_n) = \mathbb{E}\left[\frac{1}{V\{Y(x_1), \ldots, Y(x_k)\}}\right].$$

Remark. Typically closed forms won't be available but the above equation suggests a (simple) Monte Carlo estimator.

Expression using the max-stable process

Proposition 2. *Let* $\{\tilde{\eta}(x) \colon x \in \mathcal{X}\}$ *be an independent copy of* $\{\eta(x) \colon x \in \mathcal{X}\}$. *Then for all* $x_1, \ldots, x_k \in \mathcal{X}$, $k \geq 2$,

$$p(x_1, \ldots, x_k) = \sum_{r=1}^{k} (-1)^{r+1} \sum_{\substack{J \subseteq \{1,\ldots,k\} \\ |J|=r}} \mathbb{E}\left[V\{\eta(x_j) \colon j \in J\}\right].$$

Remark. As expected the extremal concurrence probability does not depend on a specific spectral representation but only on the distribution of $\{\eta(x) \colon x \in \mathcal{X}\}$.

- The above expression simplify a lot when $k = 2$.
- In particular we can define an extremal concurrence probability function $h\colon \mapsto p(o, h)$.

Proposition 3. *Let* $p\colon h \mapsto p(o, h)$ *be the extremal concurrence probability function of a stationary max-stable process* $\{\eta(x)\colon x \in \mathcal{X}\}$*. Then*

i) $h \mapsto p(h)$ *is positive semidefinite,*
ii) $h \mapsto p(h)$ *is not differentiable at the origin unless* $p \equiv 1$*.*
iii) *If* $d \geq 1$ *and* η *is isotropic, then* $h \mapsto p(h)$ *has at most one jump at the origin and is continuous elsewhere.*
iv) $2 - p(h_1 + h_2) \leq \{2 - p(h_1)\}\{2 - p(h_2)\}$*.*

Proposition 4. *For all $x_1, x_2 \in \mathcal{X}$,*

$$p(x_1, x_2) = 0 \iff \eta(x_1) \text{ and } \eta(x_2) \text{ are independent},$$
$$p(x_1, x_2) = 1 \iff \eta(x_1) \text{ and } \eta(x_2) \text{ are completely dependent},$$

Theorem 4. *Let $\{\tilde{\eta}(x) \colon x \in \mathcal{X}\}$ be an independent copy of $\{\eta(x) \colon x \in \mathcal{X}\}$. Then for all $x_1, x_2 \in \mathcal{X}$,*

$$p(x_1, x_2) = \mathbb{E}\left[\textit{sign}\{\eta(x_1) - \tilde{\eta}(x_1)\}\textit{sign}\{\eta(x_2) - \tilde{\eta}(x_2)\}\right],$$

i.e., it is the Kendall's τ.

3. Inference

Sample concurrence estimator

- Suppose we have observed $n = m \times \ell$ independent copies of a stochastic process $\{X(x) \colon x \in \mathcal{X}\}$ at locations $x_1, \ldots, x_k$ and that $X(\cdot)$ belongs to the max-domain of attraction of a max-stable process.
- As usual we partition the data into non overlapping blocks of size m, i.e., the r-th block corresponds to

$$X_{r(m-1)+1}(\cdot), \ldots, X_{r \times m}(\cdot).$$

- And check whether sample concurrence arises in each block, leading to the estimator

$$\hat{p}_m(x_1, \ldots, x_k) = \frac{1}{\ell} \sum_{r=1}^{\ell} 1_{\{\text{sample concurrence in the } r\text{-th block}\}}.$$

- Such a "blocking estimator" usually gives rise to a bias/variance trade-off. As the block size m increases,
  - the bias $p_m - p$ decreases;
  - while the variance $\frac{m}{n} p_m(1 - p_m)$ increases.
- To get the optimal block size m_* we need to compute $p_m - p$ which is usually intractable—apart from specific situations.

Proposition 5. *If $\{X(x) \colon x \in \mathcal{X}\}$ is max-stable, then*

$$p_m - p = \sum_{r=2}^{k} \frac{\Pr(|\Theta| = r)}{m^{r-1}}, \qquad m \geq 1.$$

In particular $0 \leq p_m - p \leq (1 - p)/m$ *and* $p_m - p \sim m^{-r} c_r$, $r \geq 1$ *and* $c_r > 0$.

- Proposition 5 suggests that the optimal block size satisfies

$$m_* \sim \left\{ \frac{2rc_r^2 n}{p(1-p)} \right\}^{1/(2r+1)}, \qquad \mathrm{MSE}(\hat{p}_{m_*}) \propto n^{-2r/(2r+1)}.$$

Proposition 6. *If* $p \in (0,1)$ *and* $m \sim \lambda n^{1/(2r+1)}$, $\lambda \in (0, \infty)$, *then*

$$\sqrt{\frac{n}{m}}(\hat{p}_m - p) \longrightarrow N\left\{ \frac{c_r}{\lambda^{r+1/2}}, p(1-p) \right\}, \qquad n \to \infty.$$

Remark. When $k = 2$, we can get a unbiased estimator

$$\tilde{p}_m = \frac{m\hat{p}_m - 1}{m - 1},$$

because from Proposition 5 we have $p_m - p = (1-p)/m$.

- Suppose we have observed n independent copies of a max-stable process $\{\eta(x) \colon x \in \mathcal{X}\}$ at locations $x_1, \ldots, x_k$.
- Bivariate extremal concurrence probabilities are estimated using Kendall's τ, i.e.,

$$\hat{p}(x_1, x_2) = \frac{2}{n-1} \sum_{1 \leq i < j \leq n} \mathsf{sign}\{\eta_i(x_1) - \eta_j(x_1)\} \mathsf{sign}\{\eta_i(x_2) - \eta_j(x_2)\},$$

which is an unbiased and asymptotically efficient estimator.

Remark. When $k \geq 3$, we were not able to find any unbiased estimator.

4. Application

Temperature extremes in continental USA

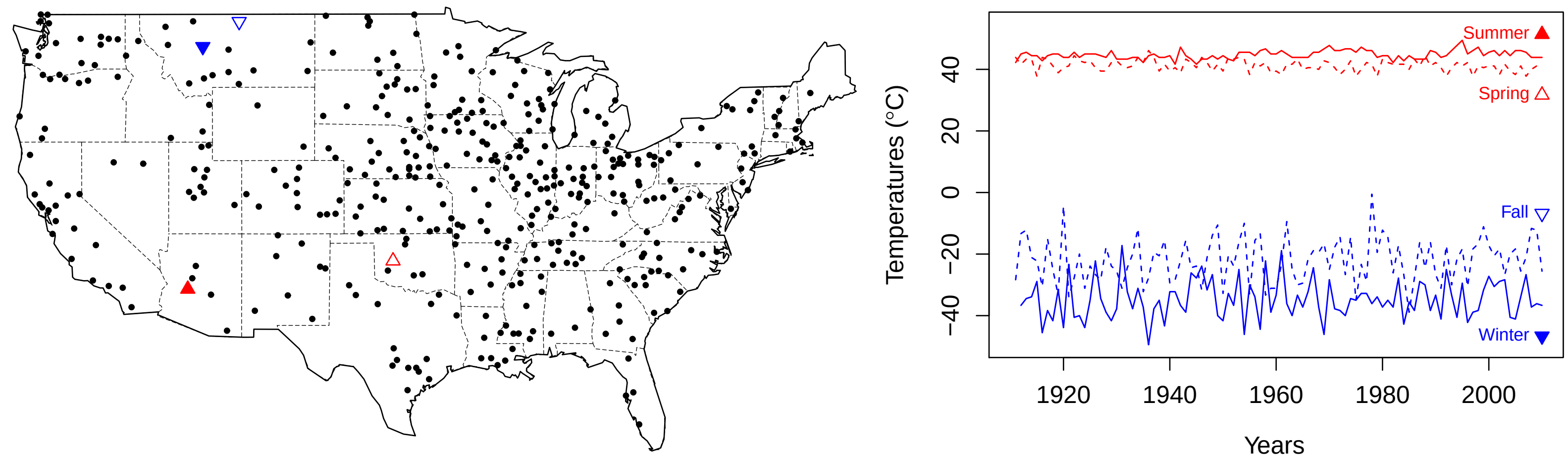

Figure 2: *Left: Spatial distribution of the 424 weather stations. The triangles indicate the selected stations for the analysis—upward: daily maxima, downward: daily minima. Right: The seasonal extrema time series of the selected stations.*

- The data are freely available from http://cdiac.ornl.gov/.
- To avoid any seasonal effect, we will work with seasonal extremes.

Concurrence probability maps

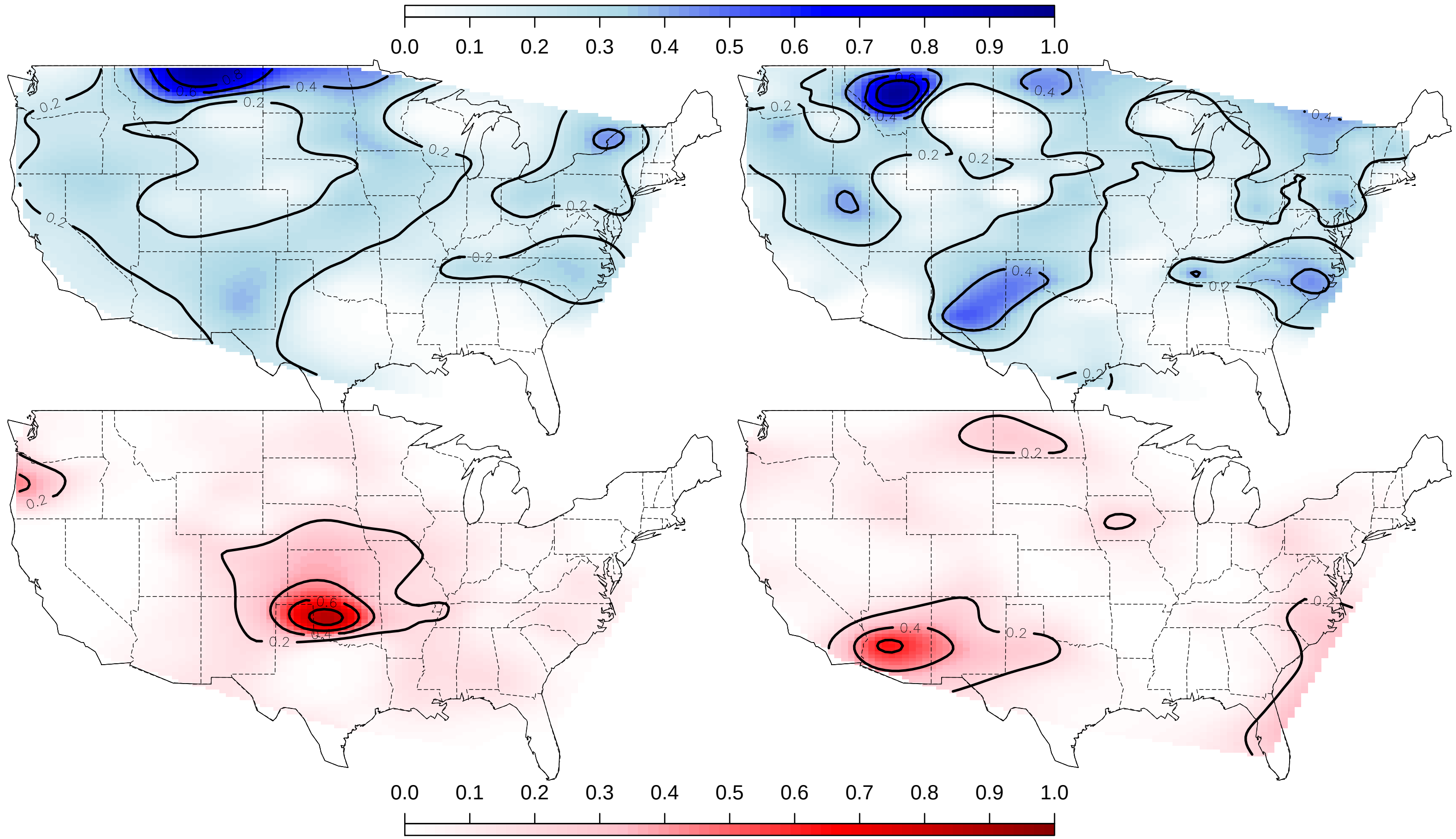

Figure 3: *Maps of the extremal concurrence probability for the four selected stations. Top left: Fall (Sep., Oct. Nov.), top right: Winter (Dec., Jan., Feb.), bottom left: Spring (Mar., Apr., May) and bottom right: Summer (June, July, Aug.).*

Concurrence cells

- For each $x \in \mathcal{X}$ we define the concurrence cell of x as the random set

$$C(x) = \{s \in \mathcal{X} \colon x \text{ and } s \text{ are concurrent}\}.$$

- Clearly we have

$$\mathbb{E}\{|C(x)|\} = \mathbb{E}\left[\int_{\mathcal{X}} 1_{\{s \in C(x)\}} \mathrm{d}s\right] = \int_{\mathcal{X}} p(x, s) \mathrm{d}s.$$

- This suggests plotting the spatial distribution of the pointwise expected concurrence cell area, i.e.,

$$\{(s, \mathbb{E}\{|C(s)|\}) \colon s \in \mathcal{X}\}.$$

Pre/Post industrialization periods

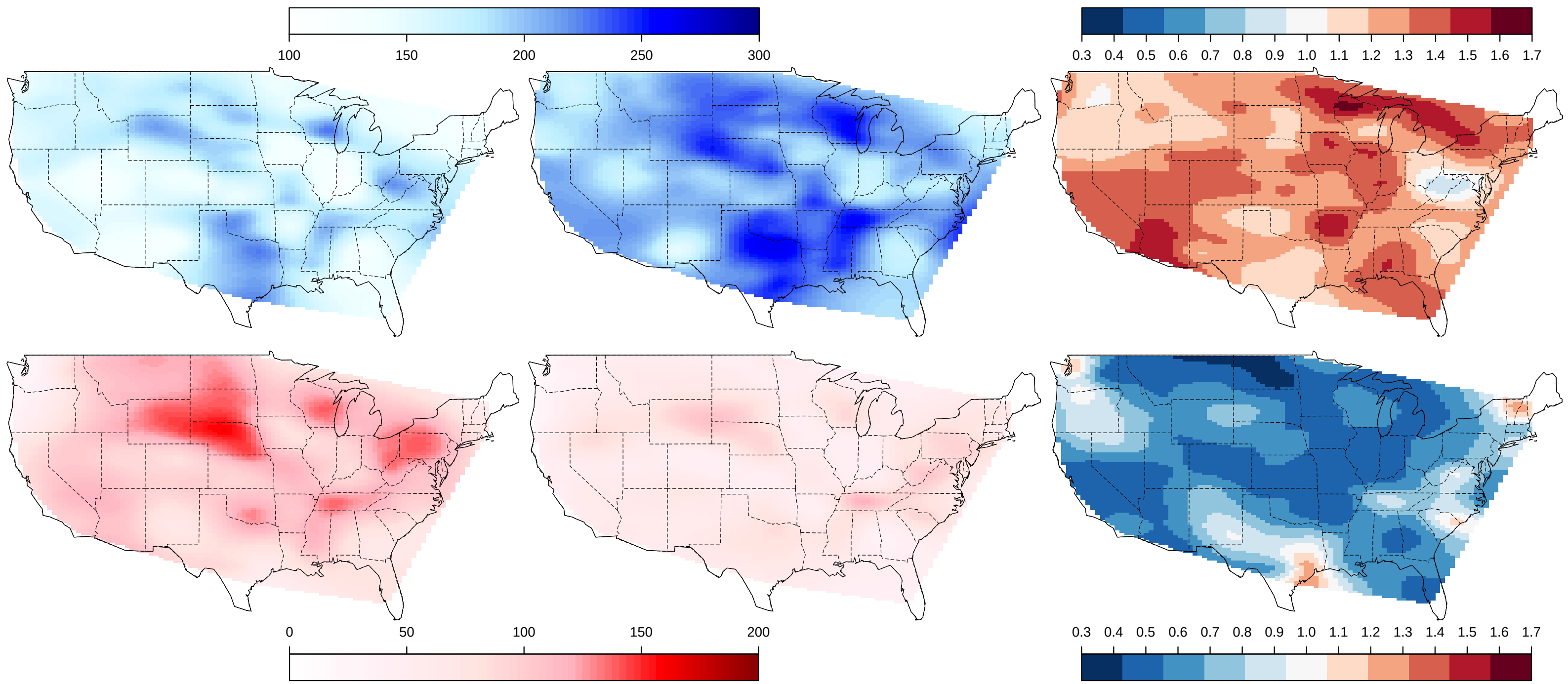

Figure 4: *Estimated spatial distribution of the expected extremal concurrence cell areas—in squared degree, i.e., around* 1000 *km*^{2}*. From left to right: 1910–1950, 1951–2010, and their ratio (1951–2010 at the numerator). Top: Winter minima, bottom: Summer maxima.*

THANK YOU FOR YOUR ATTENTION!

Dombry, C., Ribatet, M. and Stoev, S. (2015) Probabilities of concurrent extremes. (to be submitted). http://arxiv.org/abs/1503.05748

References

Brown, B. M. and Resnick, S. I. (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14:732–739.

Davison, A., Padoan, S., and Ribatet, M. (2012). Statistical modelling of spatial extremes. *Statistical Science*, 7(2):161–186.

de Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12(4):1194–1204.

de Haan, L. and Fereira, A. (2006). *Extreme value theory: An introduction*. Springer Series in Operations Research and Financial Engineering.

Dombry, C. and Éyi-Minko, F. (2013). Regular conditional distributions of max infinitely divisible random fields. *Electronic Journal of Probability*, 18(7):1–21.

Dombry, C., Éyi-Minko, F., and Ribatet, M. (2013). Conditional simulations of max-stable processes. *Biometrika*, 100(1):111–124.

Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Annals of Probability*, 37(5):2042–2065.

Opitz, T. (2013). Extremal-t process: Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis*, 122:409–413.

Penrose, M. D. (1992). Semi-min-stable processes. *Annals of Probability*, 20(3):1450–1463.

Ribatet, M. (2013). Spatial extremes: Max-stable processes at work. *Journal de la Société Française de Statistique*, 154(2):156–177.