Chapter 1

Modelling spatial extremes using max-stable processes

Abstract

This chapter introduces the theory related to the statistical modelling of spatial extremes. The presented modelling strategy heavily relies on max-stable processes which are asymptotically justified stochastic processes for pointwise maxima. After a brief reminder on the finite dimensional extreme value theory, these max-stable processes are introduced and their main properties are stressed. Their use in concrete situation is also illustrated with the help of the `R` package `SpatialExtremes`.

1.1 Introduction

Although observed at finite number of weather stations, many environmental processes such as precipitation or temperature are distributed continuously in space and the statistical modelling of such processes is usually referred to geostatistics. This specific branch of statistics dates back to the 50's/60's and is still a highly active field of research where the latest developments propose frameworks for handling massive gridded data sets with possibly non-stationary spatial dependence structures.

Most of the statistical models used in geostatistics rely on Gaussian processes but, as far as extreme events are of concern, it is well known that Gaussianity is far from being a sensible assumption–such claims date back to Sibuya [1960], and more suitable processes need to be considered. We will see in the next section why max-stable processes are especially relevant stochastic processes for modelling spatial extremes. But before going into their theoretical foundations, it is worthwhile to detail the special features of extreme events and their modelling.

Most often the aim of an extreme value analysis is to estimate the probability of lying in a critical set A_{crit}. For example in an univariate context, one could be interested in estimating the probability that a random variable Z exceeds a critical threshold $z_{\text{crit}} \in \mathbb{R}$, i.e.,

$$\Pr(Z \in A_{\text{crit}}) = \Pr(Z \geq z_{\text{crit}}).$$

The above equation appears to be a quite standard statistical problem but it is actually not the case as our focus is on the tail of the distribution and not its bulk. In other words only a few observations, or even none, belong to our critical set A_{crit} and hence our problem differs considerably to what statisticians are usually used to, i.e., drawing conclusions from a large enough number of observations. How one could reasonably estimate such probabilities if no data are available? Answering this question is the essence of the extreme value theory and readers interested in this topic should refer to Coles [2001] for a practical introduction and de Haan and Fereira [2006] for a complete and technical one.

Clearly to allow for such extrapolation in the tails of the distribution, an assumption has to be made to be able to estimate $\Pr(Z \in A_{\text{crit}})$ properly based on less extreme observations. Depending on the data considered, this assumption could be either a max-stable or a threshold-stable property but since this chapter focuses only on max-stable processes, we will restrict our attention on the former.

Definition 1.1.1. Max-stable distributions. A random variable Z is said to be max-stable if there exist sequences $\{a_n > 0 \colon n \geq 1\}$ and $\{b_n \in \mathbb{R} \colon n \geq 1\}$ such that, for all $n \geq 1$,

$$\frac{\max_{i=1,\dots,n} Z_i - b_n}{a_n} \stackrel{\text{d}}{=} Z, \tag{1.1}$$

where $Z_1, Z_2, \dots$ are independent copies of Z and the notation $\stackrel{\mathrm{d}}{=}$ means equality in distribution.

A random vector $Z = (Z_1, \dots, Z_k)$, $k \geq 2$, is said to be max-stable if (1.1) holds with component-wise algebra, e.g., $Z + a = (Z_1 + a, \dots, Z_k + a)$, $\max(X, Y) = \{\max(X_1, Y_1), \dots, \max(X_k, Y_k)\}$.

Assuming max-stability is not just an act of faith, but instead relies on mathematical arguments. Let $X_1, \dots, X_{nm}$, $n, m \geq 1$, be independent copies of X, then clearly

$$\max_{i=1,\dots,nm} X_i = \max\left\{\max(X_1, \dots, X_m), \dots, \max(X_{(n-1)m+1}, \dots, X_{nm})\right\}.$$

If in the above equation we let $m \to \infty$ and we suppose that, under an appropriate affine normalization, the left hand side converges to a non degenerate random variable/vector Z, then the n maxima appearing in the right hand side converge to $Z_1, \dots, Z_n$ respectively where $Z_1, \dots, Z_n$ are independent copies of Z. In other words, provided that it is non degenerate, Z is max-stable.

The above discussion is actually a rough sketch of the proof of the following theorem.

Theorem 1.1.1. *[de Haan and Fereira, 2006] Let $X_1, X_2, \dots$ be a sequence of independent copies of a random variable/vector X. If there exist normalizing sequences $\{c_n > 0 \colon n \geq 1\}$ and $\{d_n \colon n \geq 1\}$ such that*

$$\frac{\max_{i=1,\dots,n} X_i - d_n}{c_n} \longrightarrow Z, \qquad n \to \infty,$$

in distribution, then, provided it is non degenerate, the random variable/vector Z has a max-stable distribution.

From a modelling point of view, Theorem 1.1.1 plays a similar role to that of the central limit theorem when one assume a normal distribution for sample means except that the sample mean is now substituted for the sample maximum. For example given $X_1, X_2, \dots, X_n$ of say, daily data, one would form blocks of size $m < n$, e.g, annual block $m = 365$, and, provided the block size m is large enough, assume that these maxima are max-stable distributed.

If we now know that max-stable distributions are sensible candidates for modelling block maxima and hence characterising the distribution of extremes, there is still an open question. Apart from the rather theoretical Definition 1.1.1, is it possible to have a better representation of max-stable distributions? The answer to this question is both yes and no. We can say yes because, as the following two theorems will state, it is possible to have a precise characterisation of max-stable distributions. On the other hand we can say no as the necessary and sufficient conditions for being a max-stable distribution are too weak to yield to a single and unique statistical model to consider—at least for max-stable random vectors.

Theorem 1.1.2. *Univariate maxima. Let Z be a non degenerate max-stable random variable. Then Z has an generalized extreme value distribution, i.e., for all $z \in \mathbb{R}$,*

$$\Pr(Z \leq z) = \exp\left\{-\left(1 + \xi\frac{z - \mu}{\sigma}\right)^{-1/\xi}\right\}, \qquad 1 + \xi\frac{z - \mu}{\sigma} > 0, \tag{1.2}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$ are the location, scale and shape parameters respectively. The special case $\xi = 0$ is obtained by continuity and writes

$$\Pr(Z \leq z) = \exp\left\{-\exp\left(-\frac{z - \mu}{\sigma}\right)\right\}, \qquad z \in \mathbb{R}. \tag{1.3}$$

Due to the above theorem, the marginal distributions of the random vector $Z = (Z_1, \dots, Z_k)$, $k \geq 2$, are generalized extreme value distributions with possibly different location, scale and shape parameters. Without loss of generality, it is more convenient to treat these marginal distributions as fixed. A common choice is to assume unit Fréchet margins, i.e., $\mu = \sigma = \xi = 1$ so that $\Pr(Z_1 \leq z) = \exp(-1/z)$, $z > 0$. Note that this restriction is only assumed to ease the theoretical development, but in concrete application this assumption will be removed.

Theorem 1.1.3. *Multivariate maxima [de Haan and Fereira, 2006] Let $Z = (Z_1, \dots, Z_k)$, $k \geq 2$, be a non degenerate max-stable random vector with unit Fréchet margins. Then for all $z = (z_1, \dots, z_k) \in (0, \infty)^k$,*

$$\Pr(Z \leq z) = \exp\{-V(z_1, \dots, z_k)\}, \tag{1.4}$$

where V is a positive homogeneous function of order -1, i.e., $V(\lambda z_1, \dots, \lambda z_k) = \lambda^{-1} V(z_1, \dots, z_k)$ for all $\lambda > 0$.

The V function is sometimes referred to as the exponent function and satisfies

$$V(z) = \int_{S_k} \max_{j=1,\dots,k} \frac{w_j}{z_j} dH(w), \tag{1.5}$$

where H is a finite measure, called the spectral measure, defined on the simplex $S_k = \{w \in [0,1]^k \colon \sum_{j=1}^k w_j = 1\}$ and such that

$$\int_{S_k} w_j \, dH(w) = 1, \qquad j = 1, \dots, k.$$

1.2 Max-stable processes

This section introduces the theory for spatial extremes which is actually not much more complicated than what we saw with Theorem 1.1.3. The main difference with multivariate extremes is that extremes are now defined continuously in a spatial domain $\mathcal{X} \subset \mathbb{R}^d$, $d \geq 1$. Although one can relax a bit these assumptions, we will assume, to ease the theory, that $\mathcal{X}$ is a compact subset of $\mathbb{R}^d$ and that all the stochastic processes we will consider have continuous sample paths.

Definition 1.2.1. Max-stable processes A stochastic process $\{Z(x) \colon x \in \mathcal{X}\}$ is said to be max-stable if there exist sequences of continuous functions $\{a_n(x) > 0 \colon x \in \mathcal{X}, n \geq 1\}$ and $\{b_n(x) \in \mathbb{R} \colon x \in \mathcal{X}, n \geq 1\}$ such that, for all $n \geq 1$,

$$\left\{ \frac{\max_{i=1,\dots,n} Z_i(x) - b_n(x)}{a_n(x)} \colon x \in \mathcal{X} \right\} \stackrel{\mathrm{d}}{=} \{Z(x) \colon x \in \mathcal{X}\}, \tag{1.6}$$

where $\{Z_i(x) \colon x \in \mathcal{X}, i \geq 1\}$ is a sequence of independent copies of $\{Z(x) \colon x \in \mathcal{X}\}$.

Recall that due to our assumption the above equality in distribution is meant in the sense of all finite dimensional distributions.

Theorem 1.2.1. *[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 normalizing sequences of functions $\{c_n(x) > 0 \colon x \in \mathcal{X}, n \geq 1\}$ and $\{d_n(x) \colon x \in \mathcal{X}, n \geq 1\}$ such that*

$$\left\{ \frac{\max_{i=1,\dots,n} X_i(x) - d_n(x)}{c_n(x)} \colon x \in \mathcal{X} \right\} \longrightarrow \{Z(x) \colon x \in \mathcal{X}\}, \qquad n \to \infty, \tag{1.7}$$

then, provided it is non degenerate, the stochastic process $\{Z(x) \colon x \in \mathcal{X}\}$ is max-stable.

Similarly to what we say in the introduction, the above theorem is a justification for considering max-stable processes when modelling pointwise maxima, and thus spatial extremes.

Throughout this chapter we will illustrate the theory using the `SpatialExtremes` R package [Ribatet, 2015] and will apply our results on the Swiss rainfall data set provided within this package. Figure 1.1 plots the locations of the 79 weather stations that recorded maximum daily rainfall amounts over the years 1962–2008 during the summer season June–August.

1.2.1 Spectral characterization

If from an inferential point of view, the process is likely to be observed at a finite number of locations, fitting max-stable processes is similar to fitting multivariate extreme value distributions. However since we are working with spatial processes, most often one would like to be able to get predictions at unobserved locations which differs significantly from multivariate extremes analysis. To allow for such predictions, we need to be able to have an extension of Theorem 1.1.3 that is valid over the entire set $\mathcal{X}$. This extension is known as the spectral representation of max-stable processes.

Theorem 1.2.2. *Spectral representation of max-stable process [de Haan, 1984; Penrose, 1992] Let $\{Z(x) \colon x \in \mathcal{X}\}$ be a max-stable process with unit Fréchet margins, i.e., $\Pr\{Z(x) \leq z\} = \exp(-1/z)$ for all $x \in \mathcal{X}$ and $z > 0$. Then*

$$\{Z(x) \colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \left\{ \max_{i \geq 1} \zeta_i Y_i(x) \colon x \in \mathcal{X} \right\}, \tag{1.8}$$

where $\{\zeta_i \colon i \geq 1\}$ is a Poisson point process on $(0, \infty)$ with intensity measure $d\Lambda(\zeta) = \zeta^{-2} d\zeta$ and $\{Y_i(x) \colon x \in \mathcal{X}\}_{i \geq 1}$ a sequence of independent copies of a non negative stochastic process $\{Y(x) \colon x \in \mathcal{X}\}$ such that $\mathbb{E}\{Y(x)\} = 1$ for all $x \in \mathcal{X}$.
Depending on the context it is sometimes more convenient to write the above representation as

$$\{Z(x) \colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \left\{ \max_{i \geq 1} \varphi_i(x) \colon x \in \mathcal{X} \right\},$$

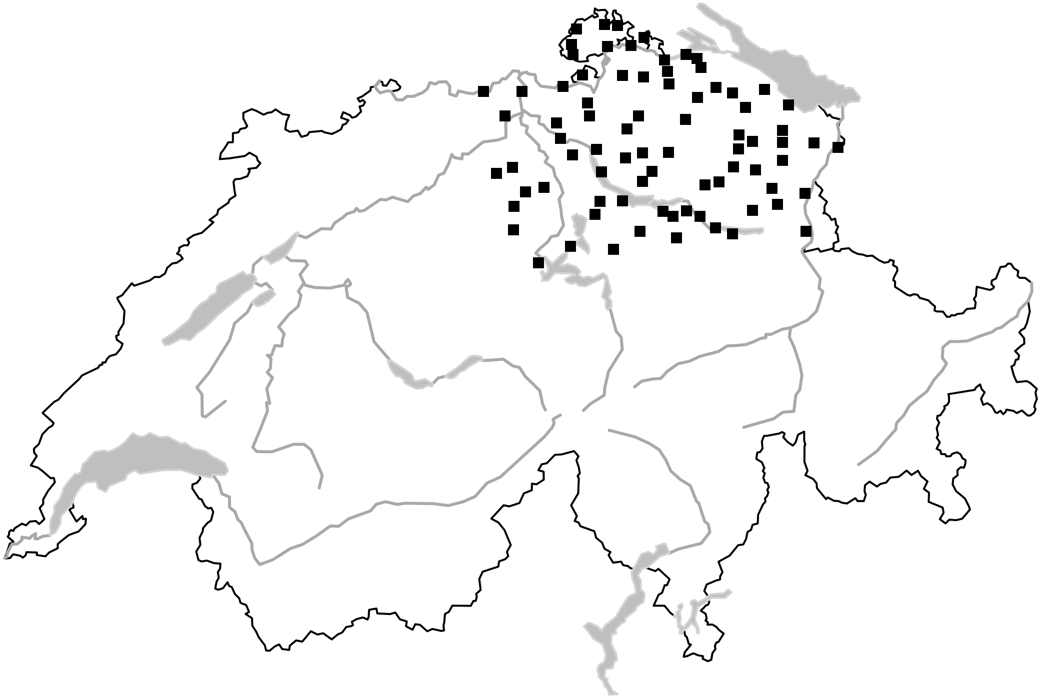

Figure 1.1: *The Swiss rainfall data set. Spatial distribution of the 79 weather stations that recorded maximum daily rainfall amounts during the summer season (June–August) over the years 1962–2008.*

where $\Phi = \{\varphi_i(x) = \zeta_i Y_i(x) \colon x \in \mathcal{X}, i \geq 1\}$ *is now a Poisson point process on* $\mathbb{C}_0$*, the space of non negative continuous functions, with intensity measure*

$$\Lambda(A) = \int_0^\infty \Pr(\zeta Y \in A) \zeta^{-2} d\zeta,$$

for all Borel set $A \subset \mathbb{C}_0$*. The function* $\{\varphi_i \colon i \geq 1\}$ *are usually referred to as spectral functions.*

For the readers not comfortable with point processes one can interpret (1.8) in a more pragmatical way [Smith, 1990]. Suppose we are concerned about rainfall storms, then we can interpret $\zeta_i Y_i(x)$ as the quantity of rain falling at location x for, say, the i–th day. In particular a given storm $\{\zeta Y(x) \colon x \in \mathcal{X}\}$ can be thought as a positive function defined on $\mathcal{X}$ which factorizes into two components:

- a storm intensity which is driven by ζ–the largest ζ is, the most severe the storm will be;
- and a storm areal impact which is controlled by $\{Y(x) \colon x \in \mathcal{X}\}$—when $\{Y(x) \colon x \in \mathcal{X}\}$ depicts strong spatial dependence, then the storm is likely to impact the whole study region $\mathcal{X}$.

Remark. It is important to keep in mind that different processes in (1.8), for instance Gaussian and Student processes, can yield the same max-stable process.

Typically the process $\{Z(x) \colon x \in \mathcal{X}\}$ is not observed over the whole set $\mathcal{X}$ but rather at a finite number of weather stations $x_1, \dots, x_k \in \mathcal{X}$, $k \geq 1$. Consequently the random vector $\{Z(x_1), \dots, Z(x_k)\}$ is necessarily max-stable and the spectral representation of Theorem 1.2.2 should be consistent with the one of Theorem 1.1.3 for multivariate maxima. Using standard computations for Poisson point processes, it is not difficult to show that for all $z_1, \dots, z_k > 0$

$$\Pr\{Z(x_1) \leq z_1, \dots, Z(x_k) < z_k\} = \exp\left[-\mathbb{E}\left\{\max_{j=1,\dots,k} \frac{Y(x_j)}{z_j}\right\}\right], \tag{1.9}$$

where $\{Y(x)\colon x \in \mathcal{X}\}$ is the process appearing in (1.8). Consequently the exponent function which now depends on the spatial location $x = (x_1, \ldots, x_k)$ is

$$V_x(z_1, \ldots, z_k) = \mathbb{E}\left\{\max_{j=1,\ldots,k} \frac{Y(x_j)}{z_j}\right\},$$

and is easily seen to be homogeneous of order -1, as required.

1.2.2 Dependence structure

From the previous section we can guess that different choices for the process $\{Y(x)\colon x \in \mathcal{X}\}$ are likely to give different spatial dependence structures as well. However because the dependence structure of the random vector $\{Z(x_1), \ldots, Z(x_k)\}$ is completely characterized by the exponent function V_x, they all share a common property. Since V_x is homogeneous of order -1 we have for all $z > 0$

$$\Pr\{Z(x_1) \leq z, \ldots, Z(x_k) \leq z\} = \exp\{-V_x(z, \ldots, z)\} = \exp\left\{-\frac{V_x(1, \ldots, 1)}{z}\right\},$$

and hence the quantity $\theta_x = V_x(1, \ldots, 1)$, called the k–variate extremal coefficient [Smith, 1990; Schlather and Tawn, 2003], provides a measure of dependence for the random vector $\{Z(x_1), \ldots, Z(x_k)\}$ that is especially designed for extremes. Although the extremal coefficient does not fully characterize the dependence structure of $\{Z(x_1), \ldots, Z(x_k)\}$, it is a useful dependence summary and also has the advantage of being independent of the level z since

$$\theta_x = -z \log \Pr\left\{\max_{j=1,\ldots,k} Z(x_j) \leq z\right\} = -\log \Pr\left\{\max_{j=1,\ldots,k} Z(x_j) \leq 1\right\} \in [1, k].$$

In the above equation, the lower bound coincides with the complete dependence case and the upper bound to independence.

Since we are interested in spatial extremes, one could be tempted to use classical measure of spatial dependence. It is not a good idea. Indeed most of the usual measures of spatial dependence such as the semi-variogram

$$\gamma(h) = \frac{1}{2}\mathrm{Var}\{Z(x) - Z(x+h)\}, \qquad x, x+h \in \mathcal{X},$$

are not adapted to extreme values since extremes usually exhibit heavy tails and in such situations even low order moments of $Z(x)$ can be infinite. For instance with unit Fréchet margins $\mathbb{E}\{Z(x)\} = \infty$. Paralleling the role of the semi-variogram, Schlather and Tawn [2003] introduce, as a special case of the extremal coefficient, the extremal coefficient function,

$$\begin{aligned} \mathbb{R}^d &\longrightarrow [1, 2] \\ h &\longmapsto \theta(h) = \theta_{(o,h)}, \end{aligned} \tag{1.10}$$

where $o \in \mathbb{R}^d$ denotes the origin.

Interestingly, the extremal coefficient function has a one-to-one relationship with the F-madogram [Cooley et al., 2006]

$$\nu_F(h) = \frac{1}{2}\mathbb{E}\left[|F\{Z(x)\} - F\{Z(x+h)\}|\right],$$

where $F(z) = \Pr\{Z(x) \leq z\}$. Indeed using the fact that $|a - b| = 2\max(a, b) - a - b$, it is not difficult to show that

$$\theta(h) = \frac{1 + 2\nu_F(h)}{1 - 2\nu_F(h)}. \tag{1.11}$$

From a statistical point of view, (1.11) is especially convenient since it suggests a simple empirical estimator by substituting $\nu_F(h)$ with its empirical counterpart and hence provides an estimator that is very similar to the empirical variogram except that it is designed for max-stable data.

More recently Dombry et al. [2015] propose a new summary measure of spatial dependence: extremal concurrence probabilities. Using the spectral characterization (1.8), we can say that extremes are concurrent at location $x_1, \ldots, x_k \in \mathcal{X}$ if

$$Z(x_j) = \varphi_\ell(x_j), \qquad j = 1, \ldots, k,$$

for some $\ell \geq 1$, i.e., the values that take the process $\{Z(x)\colon x \in \mathcal{X}\}$ at locations $x_1, \ldots, x_k$ are obtained from a single spectral function φ_ℓ. Dombry et al. [2015] define the new dependence measure as the probability of being concurrent

$$p(x_1, \ldots, x_k) = \Pr\left\{\text{for some } \ell \geq 1\colon Z(x_j) = \varphi_\ell(x_j), j = 1, \ldots, k\right\}.$$

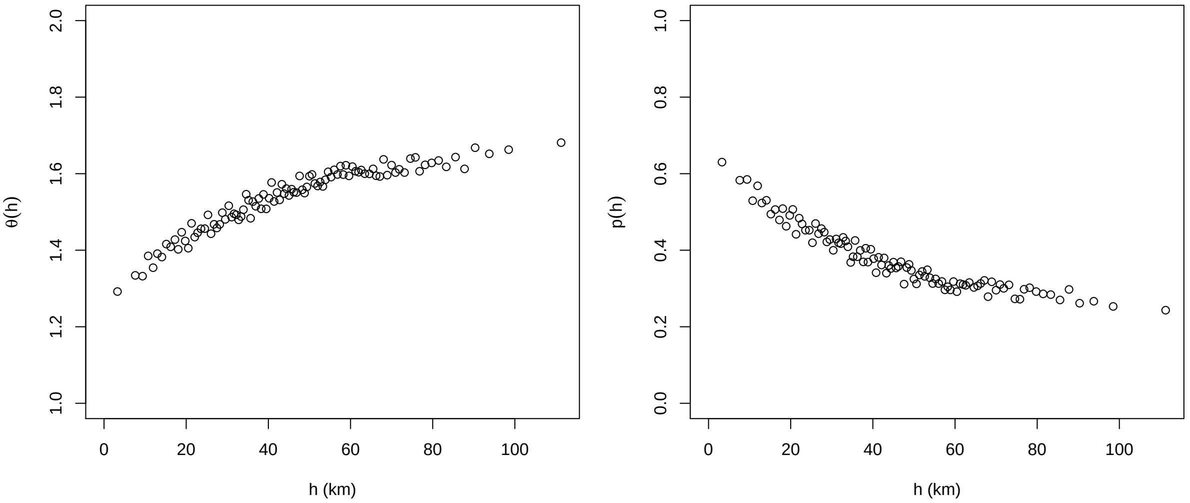

Figure 1.2: *Estimation of the spatial dependence for the Swiss rainfall data. Left: Pairwise extremal coefficient estimates. Right: Pairwise extremal concurrence probabilities.*

As expected $p(x_1, \ldots, x_k) \in [0, 1]$ where the lower bound corresponds to independence and the upper one complete dependence. Compared to extremal coefficients, concurrence probabilities are more intuitive and, as probability measures, more interpretable—using the extremal coefficient it is difficult to tell if $\theta(x_1, x_2) = 1.8$ really differs from the independent case $\theta(x_1, x_2) = 2$.

Surprisingly it can be shown that the pairwise concurrence probability coincides with the Kendall's τ, i.e.,

$$p(x_1, x_2) = \mathbb{E}\left[\text{sign}\{Z(x_1) - Z_*(x_1)\}\text{sign}\{Z(x_2) - Z_*(x_2)\}\right],$$

where $\{Z_*(x) \colon x \in \mathcal{X}\}$ is an independent copy of $\{Z(x) \colon x \in \mathcal{X}\}$. This result suggests a simple estimator

$$\hat{p}(x_1, x_2) = \frac{2}{n(n-1)} \sum_{1 \leq j < j \leq n} \text{sign}\{Z_i(x_1) - Z_j(x_1)\}\text{sign}\{Z_i(x_2) - Z_j(x_2)\},$$

where $\{Z_i(x) \colon x \in \mathcal{X}, i = 1, \ldots, n\}$ are n independent copies of $\{Z(x) \colon x \in \mathcal{X}\}$.

Paralleling the extremal coefficient function, one can define an extremal concurrence function

$$\begin{aligned} \mathbb{R}^d &\longrightarrow [0, 1] \\ h &\longmapsto p(h) = p(o, h), \end{aligned}$$

where $o \in \mathbb{R}^d$ denotes the origin.

Clearly pairwise concurrence probability gives information about the areal extent of a storm occuring at some fixed location $x \in \mathcal{X}$. It is possible to go a bit further by considering the *concurrence cell* for location $x \in \mathcal{X}$

$$C(x) = \{s \in \mathcal{X} \colon \text{extremes are concurrent at locations } s \text{ and } x\}. \tag{1.12}$$

Clearly $C(x)$ is a random set and its expected area satisfies

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

where $1_{\{\cdot\}}$ is the indicator function.

Figure 1.2 plots the empirical pairwise extremal coefficient and extremal concurrence probability estimates for the Swiss rainfall data. Both dependence measures suggests that there is still some fair amount of dependence after 100km. This figure was produced using the following code

```
> coord <- coord[,-3]##remove elevation as a spatial coordinate
>
> par(mfrow = c(1,2), mar = c(4,5,0.5,0.5))
> fmadogram(rain, coord, which = "ext", n.bins = 100)
> concprob(rain, coord, n.bins = 100)
> concarea(rain, coord)
```

1.2.3 Models

Based on the spectral representation of Theorem 1.2.2, different max-stable models can be obtained using different distributional assumptions for $\{Y(x)\colon x \in \mathcal{X}\}$. For instance a first choice could be

$$\{Y(x)\colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \left\{\sqrt{2\pi}\max\{0, \varepsilon(x)\}\colon x \in \mathcal{X}\right\},$$

where $\{\varepsilon(x)\colon x \in \mathcal{X}\}$ is a standard Gaussian process with correlation function $\rho(\cdot)$. This specific choice corresponds to the Schlather process also known as the extremal Gaussian process [Schlather, 2002]. Note that the scaling factor $\sqrt{2\pi}$ is necessary to ensure that $\mathbb{E}\{Y(x)\} = 1$ for all $x \in \mathcal{X}$ as required.

A generalization of this model is the extremal–t model [Davison et al., 2012; Opitz, 2013; Ribatet, 2013] which assumes

$$\{Y(x)\colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \{c_\nu \max\{0, \varepsilon(x)\}^\nu\colon x \in \mathcal{X}\}, \qquad c_\nu = \sqrt{\pi}2^{-(\nu-2)/2}\Gamma\left(\frac{\nu+1}{2}\right),$$

where $\{\varepsilon(x)\colon x \in \mathcal{X}\}$ is as in the Schlather model, Γ is the gamma function, $\nu \geq 1$.

Another popular model is the Brown–Resnick model [Brown and Resnick, 1977; Kabluchko et al., 2009] for which

$$\{Y(x)\colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \{\exp\{\varepsilon(x) - \gamma(x)\}\colon x \in \mathcal{X}\},$$

where $\{\varepsilon(x)\colon x \in \mathcal{X}\}$ is a Gaussian process with stationary increments and semi-variogram $\gamma(\cdot)$. Typically one uses $\gamma(h) = (h/\lambda)^\alpha$, $\lambda > 0$ and $0 < \alpha \leq 2$, i.e., the process $\{\varepsilon(x)\colon x \in \mathcal{X}\}$ is a fractional Brownian motion. Interestingly the special case $\alpha = 2$ corresponds to another well known max-stable model: the Smith model also known as the Gaussian extreme value model [Smith, 1990]. This model was initially introduced using an alternative representation to the one of Theorem 1.2.2 based on a mixed moving maxima representation [de Haan, 1984; Kabluchko et al., 2009]. More precisely the Gaussian extreme value model was originally defined by

$$\{Z(x)\colon x \in \mathcal{X}\} \stackrel{\mathrm{d}}{=} \left\{\max_{i\geq 1} \zeta_i \varphi(x - U_i; \Sigma)\colon x \in \mathcal{X}\right\},$$

where $\varphi(\cdot; \Sigma)$ is the d–variate density of a centered Normal distribution with covariance matrix Σ and $\{(\zeta_i, U_i)\colon i \geq 1\}$ is a Poisson point process on $(0, \infty) \times \mathbb{R}^d$ with intensity measure $\mathrm{d}\Lambda(\zeta, u) = \zeta^{-2}\mathrm{d}\zeta\mathrm{d}u$.

Among the above max-stable models, only two are of particular interest: the Brown–Resnick and extremal–t models. First because these two models generalize the Gaussian extreme value model and the Schlather processes respectively so that one have at the end more flexible models. More precisely it is well known that the Gaussian extreme value model provide too smooth sample paths that are unrealistic in many concrete application and that the Schlather model would never reach independence. Using their generalized version, these drawbacks are cancelled.

1.3 Inference

Fitting a max-stable process to spatial extremes data is rather complicated as two types of spatial structures needs to be taken into account:

- the spatial dependence;
- and any changes in space of marginal distributions—combined eventually with some temporal trends.

We will treat these two points in turn and start with the spatial dependence only—hence treating the margins as fixed and to be unit Fréchet.

Suppose we have observed a single realization $z = (z_1, \ldots, z_k)$ of a simple max-stable process $\{Z(x)\colon x \in \mathcal{X}\}$ at locations $x = (x_1, \ldots, x_k) \in \mathcal{X}^k$. From (1.4) we deduce that the density of $\{Z(x_1), \ldots, Z(x_k)\}$ is [Ribatet, 2013]

$$f(z_1, \ldots, z_k) = \exp\{-V_x(z_1, \ldots, z_k)\} \sum_{\tau \in \mathscr{P}_k} w(\tau), \tag{1.13}$$

where $\mathscr{P}_k$ denotes the set of all possible partitions of the set $\{x_1, \ldots, x_k\}$, $\tau = (\tau_1, \ldots, \tau_\ell)$, $|\tau| = \ell$ is the size of the partition τ and

$$w(\tau) = (-1)^{|\tau|} \prod_{j=1}^{|\tau|} \frac{\partial^{|\tau_j|}}{\partial z_{\tau_j}} V_x(z_1, \ldots, z_k),$$

where $\partial^{|\tau_j|}/\partial z_{\tau_j}$ denotes the mixed partial derivatives with respect to the element of τ_j.

The cardinality of $\mathscr{P}_k$ increases dramatically as k gets larger. For instance when $k = 10$ we have around 116000 different partitions and it is very common in concrete application that we have more than 50 weather stations. Hence most often the density cannot be evaluated and standard estimation procedures such as the maximum likelihood estimator cannot be used. As we will see in the next section, it is however possible to used likelihood based approaches.

1.3.1 Composite likelihood

Composite likelihoods is a likelihood based approach that is particularly relevant when the likelihood is either intractable or too time consuming to evaluate. Roughly speaking, composite likelihoods consist of a combination of valid likelihood entities [Lindsay, 1988].

Definition 1.3.1. Let Y be a random vector in $\mathbb{R}^k$, $k \geq 2$, with probability density function $f(y;\theta)$ where $\theta \in \mathbb{R}^p$, $p \geq 1$, is an unknown parameter vector. Let $\{\mathcal{A}_i : i \in I\}$, $I \subset \mathbb{N}$, be a set of marginal or conditional events for Y and let $\{w_i : i \in I\}$ be a set of non-negative weights. If $y_1, \ldots, y_n$ are n independent realizations of Y, the corresponding composite log likelihood is

$$\ell_c(\theta; y) = \sum_{m=1}^{n} \sum_{i \in I} w_i \log f(y_m \in \mathcal{A}_i; \theta). \tag{1.14}$$

Widely used examples of composite likelihoods are the *independence likelihood* which considers univariate densities, i.e.,

$$\sum_{m=1}^{n} \sum_{i=1}^{k} w_i \log f(y_{m,i}; \theta),$$

and the *pairwise likelihood* which considers bivariate densities, i.e.,

$$\sum_{m=1}^{n} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} w_{i,j} \log f(y_{m,i}, y_{m,j}; \theta),$$

where $y_m = (y_{m,1}, \ldots, y_{m,k})$, $m = 1, \ldots, n$.

Remark. Note that the usual likelihood, sometimes referred to as the full likelihood, is a composite likelihood. However apart from the trivial cases, e.g., the independent and full likelihoods, composite likelihoods are usually not valid likelihoods.

Assuming the same regularity conditions as the ones ensuring the asymptotic normality of the maximum likelihood estimator combined with the additional assumption that the parameter θ is identifiable from the densities appearing in (1.14), the maximum composite likelihood estimator converges in distribution:

$$\sqrt{n}\{H(\theta_0)J(\theta_0)^{-1}H(\theta_0)\}^{1/2}(\hat{\theta}_c - \theta_0) \longrightarrow N(0, \mathrm{Id}_p), \qquad n \to \infty,$$

where $M^{1/2}$ denotes a matrix square root, Id_p denotes the $p \times p$ identity matrix, $H(\theta_0) = -\mathbb{E}\{\nabla^2 \ell_c(\theta_0; Y)\}$ and $J(\theta_0) = \mathrm{Var}\{\ell_c(\theta_0; Y)\}$.

The above convergence implies that doing inference from the maximum composite likelihood estimator is much like using the maximum likelihood estimator except that standard errors and model selection need extra care due to model misspecification.

Using the `SpatialExtremes` package [Ribatet, 2015], one can fit a simple max-stable process to the Swiss rainfall data using the following code

```
> ## First convert data to unit Frechet margins using the empirical CDF
> data.frech <- apply(rain, 2, gev2frech, emp = TRUE)
> fit <- fitmaxstab(data.frech, coord, "powexp", nugget = 0)
> fit
        Estimator: MPLE
            Model: Schlather
         Weighted: FALSE
   Pair. Deviance: 1136875
              TIC: 1137468
Covariance Family: Powered Exponential

Estimates
```

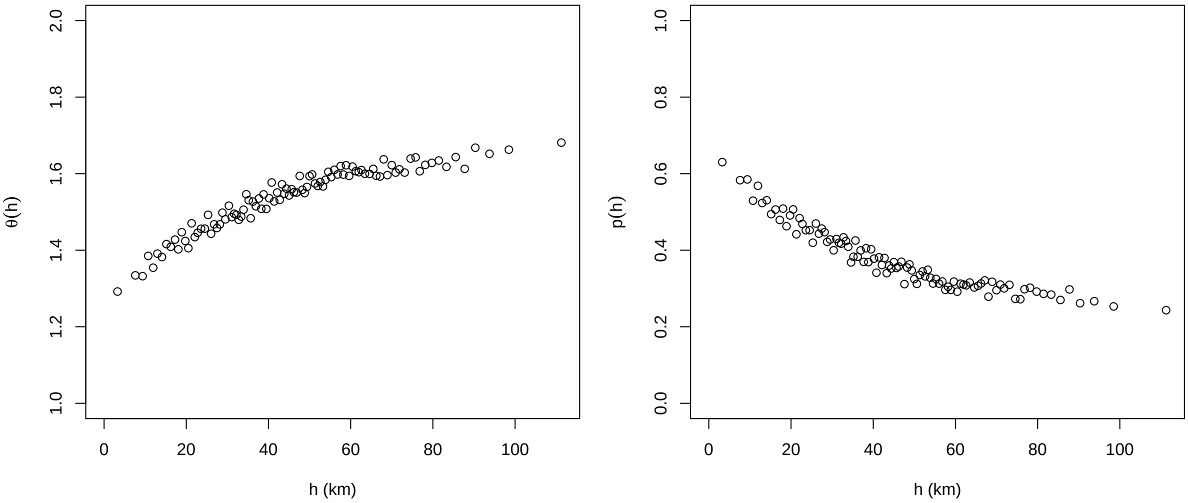

Figure 1.3: *Comparison of the fitted extremal coefficient and concurrence probability functions to their respective pairwise estimates for the Swiss rainfall data set. The data were first empirically transformed to unit Fréchet margins prior to fitting a simple max-stable process. Left: Extremal coefficients. Right: Concurrence probabilities.*

```
  Marginal Parameters:
  Assuming unit Frechet.

  Dependence Parameters:
  range   smooth
38.4402   0.8528

Standard Errors
 range  smooth
 8.713   0.119

Asymptotic Variance Covariance
        range     smooth
range   75.90902  -0.91383
smooth  -0.91383   0.01417

Optimization Information
  Convergence: successful
  Function Evaluations: 71
```

The above code fit a Schlather model with powered exponential correlation function, i.e., $\rho(h) = \exp\{-(h/\lambda)^\kappa\}$ by maximizing the pairwise likelihood. To have a graphical assessment of the goodness of fit of the fitted model we can compare the fitted extremal coefficient / concurrence probability functions to their corresponding pairwise estimates. Figure 1.3 provides such a plot and we can see that a Schlather model with a powered exponential correlation function seems to be able to reproduce the observed spatial dependence. This figure was obtained using the following code

```
> par(mfrow = c(1, 2), mar = c(4, 5, 0.5, 0.5))
> fmadogram(fitted = fit, which = "ext", n.bins = 100)
> concprob(fitted = fit, n.bins = 100)
```

1.3.2 Taking care of the margins

So far we treat the marginal distribution of $\{Z(x) \colon x \in \mathcal{X}\}$ as fixed and set to unit Fréchet margins. Clearly in concrete situations this assumption is unrealistic and there is a pressing need to allow for the marginal distribution is space, i.e., the generalized extreme value parameters μ, σ and ξ vary in space. Two strategies are possible:

1. fit trend surfaces for μ, σ and ξ based on some relevant covariates while ignoring the spatial dependence. Then given the estimated trend surfaces, transform the observations to unit Fréchet margins

and fit a simple max-stable process. This two step procedure has been widely used with copula but typically underestimates the uncertainty.

2. embed the trend surfaces directly into the likelihood of the max-stable process. This approach appears more natural but typically yields rough likelihood surfaces and might introduce some bias for the dependence structure if the trend surfaces are poor.

Trend surfaces are just functions defined on $\mathcal{X}$ that depends on some covariates, i.e., $x \mapsto \mu(x)$, $x \mapsto \sigma(x)$ and $x \mapsto \xi(x)$. For example if one defines linear trend surfaces, the location parameter could vary in space as follows

$$\mu(x) = \beta_0 + \beta_1 \text{lon}(x) + \beta_2 \text{lat}(x),$$

where lon(x) and lat(x) are the longitude and latitude for location x and $\beta_0, \beta_1, \beta_2$ unknown coefficients that need to be estimated.

Using these trend surfaces, the pairwise likelihood resembles (1.14) except that extra terms appear to take into account the pointwise transformations of a generalized extreme value random variable with location $\mu(x)$, scale $\sigma(x)$ and shape $\xi(x)$ to a unit Fréchet random variable, i.e.,

$$\sum_{m=1}^{n} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} w_{i,j} \log f(z_{m,i}, z_{m,j}; \theta) + \log J(y_{m,i}) + \log J(y_{m,j}),$$

where $J(y_{m,i})$ and $J(y_{m,j})$ are Jacobian terms of the form

$$J(y_{m,i}) = \frac{1}{\sigma(x_i)} \left(1 + \xi(x) \frac{y_{m,i} - \mu(x_i)}{\sigma(x_i)}\right)^{1/\xi(x_i)},$$

and

$$z_{m,i} = \left(1 + \xi(x_i) \frac{y_{m,i} - \mu(x_i)}{\sigma(x_i)}\right)^{1/\xi(x_i)}.$$

To illustrate our purposes we fit the Swiss rainfall data set using trend surfaces for the location and scale parameters that depend on longitude and latitude while the shape parameter is supposed to be constant over space. A preliminary model selection study indicates that these trend surfaces were a good compromise between flexibility and parsimony.

```
> loc.form <- y ~ lon + lat
> scale.form <- y ~ lon + lat
> shape.form <- y ~ 1
>
> fit <- fitmaxstab(rain, coord[,-3], "powexp", nugget = 0, loc.form, scale.form,
+                   shape.form)
Computing appropriate starting values
Starting values are defined
Starting values are:
      range      smooth   locCoeff1   locCoeff2   locCoeff3 scaleCoeff1
27.79525896  0.96491664 20.65574265  0.06577214 -0.15917825  8.96763099
scaleCoeff2 scaleCoeff3 shapeCoeff1
 0.01818016 -0.04757241  0.18757914
> fit
        Estimator: MPLE
            Model: Schlather
         Weighted: FALSE
   Pair. Deviance: 2243829
              TIC: 2254631
Covariance Family: Powered Exponential

Estimates
  Marginal Parameters:
    Location Parameters:
locCoeff1  locCoeff2  locCoeff3
 18.80554    0.06775   -0.15713
       Scale Parameters:
```

```
scaleCoeff1  scaleCoeff2  scaleCoeff3
    4.32475      0.02249     -0.04074
       Shape Parameters:
shapeCoeff1
     0.1952
  Dependence Parameters:
  range   smooth
29.1018   0.9733

Standard Errors
      range       smooth    locCoeff1    locCoeff2    locCoeff3
   9.395412     0.175926     6.714423     0.008977     0.016350
scaleCoeff1  scaleCoeff2  scaleCoeff3  shapeCoeff1
   4.774224     0.007255     0.011630     0.052612

Asymptotic Variance Covariance
             range       smooth      locCoeff1   locCoeff2   locCoeff3
range         8.827e+01  -1.466e+00   2.681e+00   1.044e-02  -2.466e-02
smooth       -1.466e+00   3.095e-02  -2.814e-03  -1.350e-04   2.227e-04
locCoeff1     2.681e+00  -2.814e-03   4.508e+01  -4.660e-02  -4.528e-02
locCoeff2     1.044e-02  -1.350e-04  -4.660e-02   8.059e-05  -3.518e-05
locCoeff3    -2.466e-02   2.227e-04  -4.528e-02  -3.518e-05   2.673e-04
scaleCoeff1   7.044e-01  -9.693e-03   1.168e+01  -1.521e-02  -2.567e-03
scaleCoeff2   1.393e-02  -1.437e-04  -1.223e-02   2.941e-05  -2.774e-05
scaleCoeff3  -1.914e-02   1.408e-04  -9.871e-03  -1.830e-05   7.796e-05
shapeCoeff1   2.504e-01  -2.515e-03   3.322e-02  -2.619e-05  -7.647e-05
             scaleCoeff1  scaleCoeff2  scaleCoeff3  shapeCoeff1
range         7.044e-01    1.393e-02   -1.914e-02    2.504e-01
smooth       -9.693e-03   -1.437e-04    1.408e-04   -2.515e-03
locCoeff1     1.168e+01   -1.223e-02   -9.871e-03    3.322e-02
locCoeff2    -1.521e-02    2.941e-05   -1.830e-05   -2.619e-05
locCoeff3    -2.567e-03   -2.774e-05    7.796e-05   -7.647e-05
scaleCoeff1   2.279e+01   -2.740e-02   -1.230e-02    2.517e-02
scaleCoeff2  -2.740e-02    5.263e-05   -3.411e-05    2.207e-05
scaleCoeff3  -1.230e-02   -3.411e-05    1.353e-04   -9.392e-05
shapeCoeff1   2.517e-02    2.207e-05   -9.392e-05    2.768e-03

Optimization Information
  Convergence: successful
  Function Evaluations: 750
```

Similarly to what we did in Section 1.3.1, we can check if the fitted max-stable process is able to reproduce appropriately the observed spatial dependence. Figure 1.4 does such a plot using the exact same lines as we did earlier. Compared to Figure 1.3, we can see that the fitted model underestimates slightly the spatial dependence. Indeed when using trend surfaces it is common that if these trend surfaces are not accurate enough to capture the spatial variability of the marginal parameters, this will induce underestimation of the spatial dependence. It is therefore important to take care about building relevant trend surfaces including any relevant covariable.

1.4 Simulation

Sooner or later one would typically be interested in a Monte Carlo experiment to assess the distribution of some useful quantity. For instance, in a spatial extreme context, such a quantity could be

$$I = \frac{1}{|A|} \int_A Z(x) \mathrm{d}x,$$

where $A \subset \mathcal{X}$ is a sub region of particular importance of area $|A|$. Most often the distribution of the random variable I will be unknown but numerical simulations help us in characterizing it from a Monte Carlo sample

$$\left\{ \frac{1}{|A|} \int_A Z_i(x) \mathrm{d}x \colon i = 1, \ldots, N \right\},$$

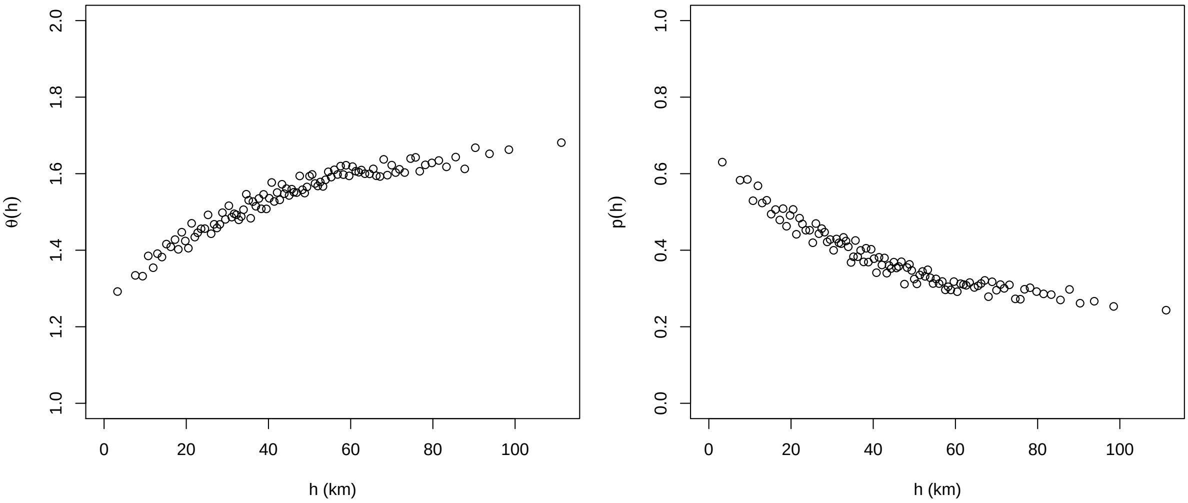

Figure 1.4: *Comparison of the fitted extremal coefficient and concurrence probability functions to their respective pairwise estimates for Swiss rainfall data set using trend surfaces. Left: Extremal coefficients. Right: Concurrence probabilities.*

where $\{Z_i(x) \colon x \in A, i = 1, \ldots, N\}$ are independent copies of $\{Z(x) \colon x \in A\}$.

When we are talking about simulation, we typically have two different kind of simulations:

1. unconditional simulations where we want to sample from a given distribution without any further constraints ;
2. conditional simulations where we want to sample from a given distribution subject to some prescribed constraints.

1.4.1 Unconditional simulation

Although one could use (1.7) to get realization from a max-stable process, this approach is rarely used in practice since the rate of convergence to the limiting process $\{Z(x) \colon x \in \mathcal{X}\}$ is usually very slow. A better strategy relies on the spectral characterization (1.8) of $\{Z(x) \colon x \in \mathcal{X}\}$. To this aim we need to be able to simulate points of a Poisson process on $(0, \infty)$ with intensity $\mathrm{d}\Lambda(\zeta) = \zeta^{-2}\mathrm{d}\zeta$ and independent copies of a non negative stochastic process $\{Y(x) \colon x \in \mathcal{X}\}$ such that $\mathbb{E}\{Y(x)\} = 1$.

If the latter is generally not complicated as it usually consists in, up to a transformation, simulating Gaussian processes, the former is less standard. The intensity measure $\mathrm{d}\Lambda(\zeta) = \zeta^{-2}\mathrm{d}\zeta$ diverges for any set of the form $(0, a)$, $a > 0$, and thus implies that 0 is an accumulation point for $\{\zeta_i \colon i \geq 1\}$. For simulation purposes, this feature is highly convenient as it suggests that only a finite number of random functions would be necessary to get a realization from $\{Z(x) \colon x \in \mathcal{X}\}$. Moreover due to the independence between $\{\zeta_i \colon i \geq 1\}$ and $\{Y_i(x) \colon x \in \mathcal{X}, i \geq 1\}$, it is, as noted by Schlather [2002], more efficient to reorder the points $\{\zeta_i \colon i \geq 1\}$ into decreasing order, i.e., $\{\zeta_{(1)} > \zeta_{(2)} > \cdots\}$ so that

$$\left\{\zeta_{(1)} > \zeta_{(2)} > \cdots\right\} \stackrel{\mathrm{d}}{=} \left\{\left(\sum_{j=1}^{i} E_j\right)^{-1} \colon i \geq 1\right\}, \qquad E_1, E_2, \ldots \stackrel{\mathrm{iid}}{\sim} \mathrm{Exp}(1).$$

If there exist some positive constant $C < \infty$ such that $\Pr\{Y(x) \leq C\} = 1$ for all $x \in \mathcal{X}$, then

$$\frac{Y_i(x)}{\sum_{j=1}^{i} E_j} \leq \frac{C}{\sum_{j=1}^{i} E_j} \longrightarrow 0, \qquad i \to \infty,$$

almost surely. The above equation states that, as expected, only a finite number of points ζ_i and stochastic processes $\{Y_i(x) \colon x \in \mathcal{X}\}$ will contribute to the pointwise maxima. It also suggests the following stopping rule: keep on generating points (E_n, Y_n) until

$$\max_{i=1,\ldots,n} \frac{Y_i(x)}{\sum_{j=1}^{i} E_j} \geq \frac{C}{\sum_{j=1}^{i} E_j}, \qquad x \in \mathcal{X}.$$

Remark. It may happen that the assumption $\Pr\{Y(x) \leq C\} = 1$, $x \in \mathcal{X}$, is not satisfied. In such situations you can define a pseudo uniform bound $C < \infty$ such that $\Pr\{Y(x) > C\} = \varepsilon$, for some small enough $\varepsilon > 0$. For example with a Schlather process one could set $C = 4\sqrt{2\pi}$.

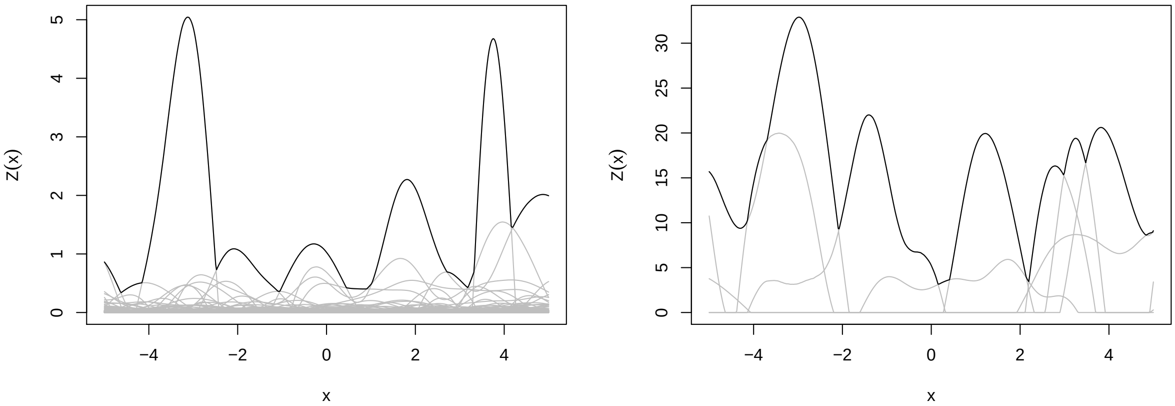

Figure 1.5: *Illustration of the two different strategies for simulating max-stable processes. Left: the one based on uniform random variables with $N = 500$. Right: The one based on reordering using a pseudo uniform bound $C = 4\sqrt{2\pi}$..*

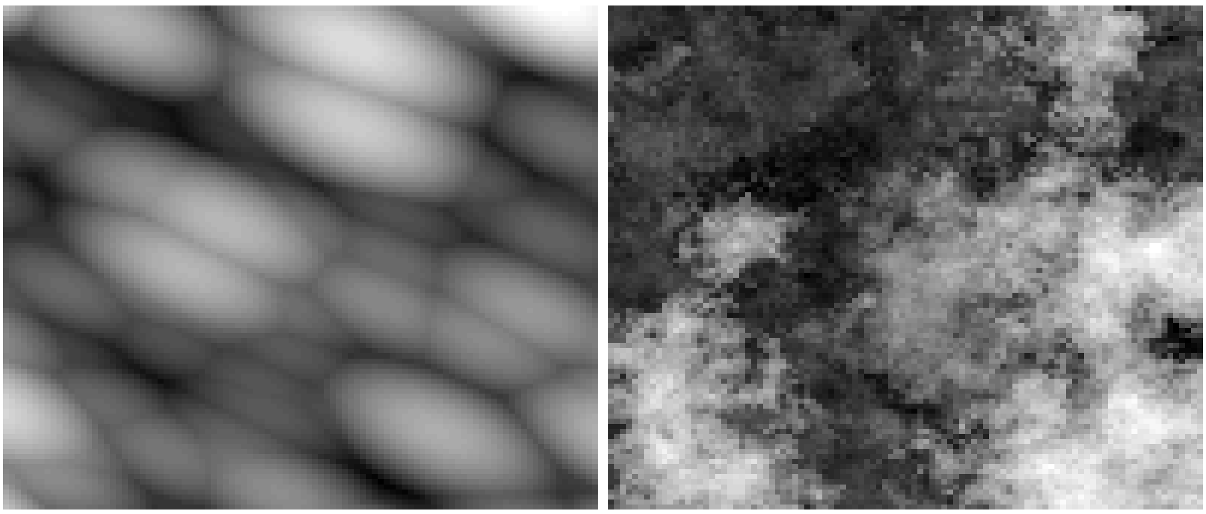

Figure 1.6: *Simulation of two max-stable processes on a grid. Left: Smith model. Right: Extremal-t model. A log-scale was used for a better display.*

Figure 1.5 illustrates the two different strategies for simulating a Schlather process with correlation function $\rho(h) = \exp(-h^2)$. For the first approach we compute the pointwise maxima based on $N = 500$ random functions $\varphi_i(x) = \zeta_i Y_i(x)$ while using the stopping rule only 5 random functions were generated.

The `SpatialExtremes` package implements the second strategy—although more refined strategies are required for sampling from Brown–Resnick processes. Figure 1.6 plots two realizations of max-stable process on $\mathcal{X} = [-5, 5] \times [-5, 5]$ and was produced using the code

```
> x <- y <- seq(-5, 5, length = 100)
> Z1 <- rmaxtab(1, cbind(x, y), "gauss", cov11 = 2, cov12 = 0.5, cov22 = 3,
+               grid = TRUE)
> Z2 <- rmaxstab(1, cbind(x, y), "tpowexp", DoF = 3, nugget = 0, range = 3,
+                smooth = 1, grid = TRUE)
>
> par(mfrow = c(1,2), mar = rep(0,0))
> image(x,y, Z1, col = grey(0:63/63))
> image(x,y, Z2, col = grey(0:63/63))
```

1.4.2 Conditional simulation

Performing conditional simulations amount to obtaining independent realizations from a given distribution with the additional feature that each realization should satisfy some constraint. Although other constraints are possible, we will restrict our attention to the most natural ones, i.e., simulate from a simple max-stable process subject to

$$Z(x_1) = z_1, \ldots, Z(x_k) = z_k, \qquad x_1, \ldots, x_k \in \mathcal{X}, z_1, \ldots, z_k \in (0, \infty), k \geq 1.$$

This type of constraint is of greatest importance since it enables to characterize the behaviour of the process $\{Z(x) \colon x \in \mathcal{X}\}$ at some new locations $s \in \mathcal{X}$ given that we have observed some known values at locations $x = (x_1, \ldots, x_k)$.

Algorithms to get such conditional simulations were recently proposed. Wang and Stoev [2011] were the first to work on this topic and propose a procedure to get conditional realization from max-linear model, i.e., max-stable processes with a discrete spectral measure. Later Dombry and Éyi-Minko [2013] were able to derive explicit formulas for the conditional distribution of (regular) max-stable processes and Dombry et al. [2013] derive a framework to sample from these conditional distributions. This section summarizes their results.

The algorithm of Dombry et al. [2013] is a three step procedure:

Step 1 Sample a random partition $\vartheta \in \mathscr{P}_k$, i.e., a hitting scenario, according to the discrete distribution

$$\Pr\{\vartheta = \tau \mid Z(x) = z\} \propto \prod_{j=1}^{|\tau|} \int_{\{u_j < z_{\tau_j^c}\}} \lambda_{(x_{\tau_j}, x_{\tau_j^c})}(z_{\tau_j}, u_j) \mathrm{d}u_j$$

where $\tau \in \mathscr{P}_k$.

Step 2 Given $\vartheta = \tau$ of size ℓ; sample independently the extremal function $\varphi_1^+, \ldots, \varphi_\ell^+$ from the distribution

$$\Pr\left\{\varphi_j^+(s) \in \mathrm{d}v \mid Z(x) = z, \vartheta = \tau\right\} \propto \int 1_{\{u < z_{\tau_j^c}\}} \lambda_{(s, x_{\tau_j^c}) | x_{\tau_j}, z_{\tau_j}}(v, u) \mathrm{d}u \mathrm{d}v,$$

and let

$$\left\{Z^+(s) \colon s \in \mathcal{X}\right\} = \left\{\max_{j=1,\ldots,\ell} \varphi_j^+(s) \colon s \in \mathcal{X}\right\}.$$

Step 3 Independently from the previous steps, draw a "thinned" max-stable process

$$\left\{Z^-(s) \colon s \in \mathcal{X}\right\} \stackrel{\mathrm{d}}{=} \left\{\max_{\varphi \in \Phi} \varphi(s) 1_{\{\varphi(x) < z\}} \colon s \in \mathcal{X}\right\},$$

where the Poisson point process Φ is as in Theorem 1.2.2.

The process $\{\max\{Z^-(s), Z^+(s)\} \colon s \in \mathcal{X}\}$ has the required conditional distribution.

The above algorithm is rather simple provided that we are able to get explicit formulas for $\lambda_x(\cdot)$ and $\lambda_{s|x,z}(u)$. Closed forms for Brown–Resnick processes are given in Dombry and Éyi-Minko [2013], while formulas for Schlather and extremal–t processes can be found in Dombry et al. [2013] and Ribatet [2013] respectively.

Figure 1.7 plots 5 conditional simulations from an extremal-t model. These conditional simulations were obtained using the following snippet

```
> ## Generate the conditional values from a Brown--Resnick process
> n.cond <- 5
> x <- runif(n.cond, -5, 5)
> z <- rmaxstab(1, x, "brown", range = 3, smooth = 1.5)
>
> ## Generate 2 x 5 conditional simulations
> n.sim <- 5
> s <- seq(-5, 5, length = 500)
> cond.sim <- condrmaxstab(n.sim, s, x, z, "brown", range = 3, smooth = 1.5)
> ## The same with much smoother sample paths
> cond.sim2 <- condrmaxstab(n.sim, s, x, z, "brown", range = 3, smooth = 1.995)
```

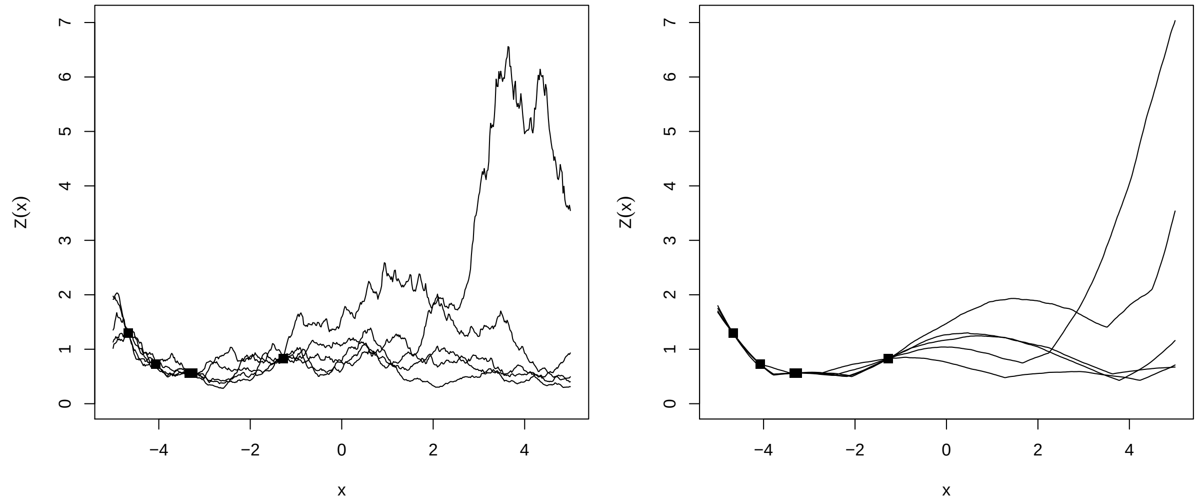

Figure 1.7: *Conditional simulation from a Brown–Resnick model with semi-variogram $\gamma(h) = (h/3)^{1.5}$ (left) and $\gamma(h) = (h/3)^2$ (right). The squares correspond to the conditional values.*

1.5 Discussion

In this chapter we introduce the basic theory on max-stable processes and their use for modelling spatial extremes. We also see how to use them in practice using the `SpatialExtremes` package [Ribatet, 2015]. Many recent developments were not covered in this chapter including efficient estimation of max-stable processes, e.g., full likelihood based approaches. The reason for such a major omission is that these type of approaches are extremely time consuming and reliable code implementing these approaches are still missing.

In this chapter we also focus on modelling pointwise maxima while, similarly to the univariate / multivariate cases, other representations of extremes are possible, e.g., threshold exceedances. As for efficient estimations, no code are yet available and we decided to not cover this topic but if the reader feels comfortable with the material introduced within this chapter, the move to threshold exceedances will hopefully be smooth.

To conclude, the modelling of spatial extremes using extreme value arguments is an extremely active area of research and progress is expected to be made in the next coming years. However so far the software implementation of these methodology is rather limited but may see huge advances in the near future. The use of max-stable processes for modelling spatial extremes has mainly be restricted to the statistical literature [Davison et al., 2012; Thibaud et al., 2015; Wadsworth and Tawn, 2014; Ribatet and Sedki, 2013]. However their use begins to emerge in other fields. For instance, Westra and Sisson [2011] analyse the impact of spatial dependence in identifying trends in precipitation extremes; Shang et al. [2011] fit a max-stable process whose marginal parameters depend on some climatic covariates and Gaume et al. [2013] use max-stable processes to predict extreme snowfall quantiles in the French Alps.

Although some theoretical aspects are missing, with this chapter, the authors try to provide a solid theoretical basis for those interested in modelling spatial extremes while mainly focusing on its practical implementation. Readers interested in the latest developments are invited to have a look at the following papers: for Bayesian hierarchical models using max-stable processes as the data layer [Ribatet et al., 2012; Reich and Shaby, 2012; Thibaud et al., 2015], for higher order composite likelihood and full likelihood approaches for max-stable processes [Genton et al., 2011; Huser and Davison, 2013; Wadsworth and Tawn, 2014].

Bibliography

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

Coles, S. (2001). *An Introduction to Statistical Modelling of Extreme Values*. Springer Series in Statistics. Springers Series in Statistics, London.

Cooley, D., Naveau, P., and Poncet, P. (2006). Variograms for spatial max-stable random fields. In Bertail, P., Soulier, P., Doukhan, P., Bickel, P., Diggle, P., Fienberg, S., Gather, U., Olkin, I., and Zeger, S., editors, *Dependence in Probability and Statistics*, volume 187 of *Lecture Notes in Statistics*, pages 373–390. Springer New York.

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.

Dombry, C., Ribatet, M., and Stoev, S. (2015). Probabilities of concurrent extremes. *Submitted*.

Gaume, J., Eckert, N., Chambon, G., Naaim, M., and Bel, L. (2013). Mapping extreme snowfalls in the French Alps using max-stable processes. *Water Resources Research*, 49:1–20.

Genton, M. G., Ma, Y., and Sang, H. (2011). On the likelihood function of Gaussian max-stable processes. *Biometrika*, 98(2):481–488.

Huser, R. and Davison, A. C. (2013). Composite likelihood estimation for the Brown–Resnick process. *Biometrika*, 100(2):511–518.

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.

Lindsay, B. (1988). *Composite likelihood methods*. Statistical Inference from Stochastic Processes. American Mathematical Society, Providence.

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.

Reich, B. and Shaby, B. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Annals Of Applied Statistics*, 6:1430–1451.

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

Ribatet, M. (2015). *SpatialExtremes: Modelling Spatial Extremes*. R package version 2.0-2.

Ribatet, M., Cooley, D., and Davison, A. (2012). Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statistica Sinica*, 22:813–845.

Ribatet, M. and Sedki, M. (2013). Extreme value copulas and max-stable processes. *Journal de la Société Française de Statistique*, 154(1):138–150.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5(1):33–44.

Schlather, M. and Tawn, J. (2003). A dependence measure for multivariate and spatial extremes: Properties and inference. *Biometrika*, 90(1):139–156.

Shang, H., Yan, J., and Zhang, X. (2011). El Nino-Southern Ocillation influence on winter maximum daily precipitation in California in a spatial model. *Water Resources Research*, 47.

Sibuya, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.*, 11:195–210.

Smith, R. L. (1990). Max-stable processes and spatial extreme. *Unpublished manuscript*.

Thibaud, E., Aalto, J., Cooley, D., Davison, A. C., and Heikkinen, J. (2015). Bayesian inference for the Brown–Resnick process, with an application to extreme low temperatures. *Submitted*.

Wadsworth, J. L. and Tawn, J. A. (2014). Efficient inference for spatial extreme value processes associated to log-gaussian random functions. *Biometrika*, 101(1):1–15.

Wang, Y. and Stoev, S. A. (2011). Conditional sampling for spectrally discrete max-stable random fields. *Advances in Applied Probability*, 443:461–483.

Westra, S. and Sisson, S. (2011). Dectection of non-stationarity in precipitation extremes using a max-stable process model. *Journal of Hydrology*, 406:119–128.