Biometrika (2013), **100**, 1, *pp.* 111–124
doi: 10.1093/biomet/ass067
Advance Access publication 29 November 2012

Conditional simulation of max-stable processes

By C. Dombry, F. Éyi-Minko

Laboratory of Mathematics and Application, Université de Poitiers, Téléport 2, BP 30179, F-86962 Futuroscope-Chasseneuil cedex, France
clement.dombry@math.univ-poitiers.fr frederic.eyi.minko@math.univ-poitiers.fr

and M. Ribatet

Department of Mathematics, Université Montpellier 2, 4 place Eugéne Bataillon, 34095 cedex 2 Montpellier, France
mathieu.ribatet@math.univ-montp2.fr

Summary

Since many environmental processes are spatial in extent, a single extreme event may affect several locations, and the spatial dependence must be taken into account in an appropriate way. This paper proposes a framework for conditional simulation of max-stable processes and gives closed forms for the regular conditional distributions of Brown–Resnick and Schlather processes. We test the method on simulated data and present applications to extreme rainfall around Zurich and extreme temperatures in Switzerland. The proposed framework provides accurate conditional simulations and can handle problems of realistic size.

Some key words: Conditional simulation; Markov chain Monte Carlo; Max-stable process; Precipitation; Regular conditional distribution; Temperature.

1. Introduction

Max-stable processes arise naturally when studying extremes of stochastic processes and therefore play a major role in the statistical modelling of spatial extremes (Buishand et al., 2008; Padoan et al., 2010; Davison et al., 2012). Although a different spectral characterization of max-stable processes exists (de Haan, 1984), for our purposes the most useful representation is

$$Z(x) = \max_{i \geqslant 1} \zeta_i Y_i(x), \quad x \in \mathbb{R}^d \tag{1}$$

(Schlather, 2002), where $\{\zeta_i\}_{i \geqslant 1}$ are the points of a Poisson process on $(0, \infty)$ with intensity $\mathrm{d}\Lambda(\zeta) = \zeta^{-2}\,\mathrm{d}\zeta$ and the Y_i are independent replicates of a nonnegative stochastic process Y such that $E\{Y(x)\} = 1$ for all $x \in \mathbb{R}^d$. It is well known that Z is a max-stable process on $\mathbb{R}^d$ with unit Fréchet margins (Schlather, 2002; de Haan & Fereira, 2006, p. 307). Although (1) takes the pointwise maximum over an infinite number of points ζ_i and processes Y_i, it is possible to get approximate realizations from Z (Schlather, 2002; Oesting et al., 2012).

Based on (1), several parametric max-stable models have been proposed (Brown & Resnick, 1977; Schlather, 2002; Kabluchko et al., 2009; Davison et al., 2012) that share the same finite-dimensional distribution functions

$$\mathrm{pr}\{Z(x_1) \leqslant z_1, \ldots, Z(x_k) \leqslant z_k\} = \exp\left[-E\left\{\max_{j=1,\ldots,k} \frac{Y(x_j)}{z_j}\right\}\right],$$

where $k \in \mathbb{N}$, $z_1, \ldots, z_k > 0$ and $x_1, \ldots, x_k \in \mathbb{R}^d$.

Except in the case of the Smith model (Genton et al., 2011), only the bivariate cumulative distribution functions are explicitly known. To bypass this impediment to inference based on max-stable processes, de Haan & Pereira (2006) proposed a semiparametric estimator and Padoan et al. (2010) suggested the use of the maximum pairwise likelihood estimator.

Similar to the variogram in classical geostatistics, the extremal coefficient function

$$\theta(x_1 - x_2) = -z \log \mathrm{pr}\{Z(x_1) \leqslant z,\, Z(x_2) \leqslant z\}$$

(Schlather & Tawn, 2003; Cooley et al., 2006) is widely used to summarize the spatial dependence of extremes for stationary processes. It takes values in the interval $[1, 2]$; the lower bound indicates complete dependence and the upper bound complete independence.

The past decade has seen many advances in the geostatistics of extremes, and software packages have been developed in R (R Development Core Team, 2012) and made available to practitioners (Wang, 2010; Ribatet, 2011; Schlather, 2012). However, an important tool still missing is a method for conditional simulation of max-stable processes. In classical geostatistics based on Gaussian models, conditional simulation is well established (Chilès & Delfiner, 1999) and provides a framework for assessing the distribution of a Gaussian random field given values observed at fixed locations; for example, conditional simulations of Gaussian processes have been used to model land topography (Mandelbrot, 1982).

Conditional simulation of max-stable processes is a long-standing problem (Davis & Resnick, 1989, 1993). Wang & Stoev (2011) provided a first solution, but their framework is limited to processes having a discrete spectral measure and thus may be too restrictive to appropriately model spatial dependence in complex situations.

Based on recent developments in understanding the regular conditional distribution of max-infinitely divisible processes, the aim of this paper is to provide a method for conditional simulation of max-stable processes with continuous spectral measures. More formally, for a study region $\mathcal{X} \subset \mathbb{R}^d$, our goal is to derive an algorithm to sample from the regular conditional distribution of $Z \mid \{Z(x_1) = z_1, \ldots, Z(x_k) = z_k\}$ for some $z_1, \ldots, z_k > 0$ and k conditioning locations $x_1, \ldots, x_k \in \mathcal{X}$.

2. Conditional simulation of max-stable processes

2·1. *General framework*

This section reviews some key results from an unpublished paper by the first author, with a particular emphasis on max-stable processes. Our aim is to give a more practical interpretation of the results from a simulation perspective. With this in mind, we recall two key results and propose a procedure for generating conditional realizations of max-stable processes.

Let $\mathbb{R}^{\mathcal{X}}$ be the space of real-valued functions on $\mathcal{X} \subset \mathbb{R}^d$, and let $\Phi = \{\varphi_i\}_{i \geqslant 1}$ be a Poisson point process on $\mathbb{R}^{\mathcal{X}}$ where $\varphi_i(x) = \zeta_i Y_i(x)$ $(i = 1, 2, \ldots)$ with ζ_i and Y_i as in (1). We write $f(x) = \{f(x_1), \ldots, f(x_k)\}$ for all random functions $f : \mathcal{X} \to \mathbb{R}$ and $x = (x_1, \ldots, x_k) \in \mathcal{X}^k$. It is not difficult to show that for all Borel sets $A \subset \mathbb{R}^k$, the Poisson process $\{\varphi_i(x)\}_{i \geqslant 1}$ defined on $\mathbb{R}^k$ has intensity measure

$$\Lambda_x(A) = \int_0^\infty \mathrm{pr}\{\zeta Y(x) \in A\} \zeta^{-2} \, \mathrm{d}\zeta.$$

The point process Φ is said to be regular if the intensity measure Λ_x has an intensity function λ_x, that is, $\Lambda_x(\mathrm{d}z) = \lambda_x(z) \, \mathrm{d}z$ for all $x \in \mathcal{X}^k$.

The first key result asserts that, provided the Poisson process Φ is regular, the intensity function λ_x and the conditional intensity function

$$\lambda_{s|x,z}(u) = \frac{\lambda_{(s,x)}(u,z)}{\lambda_x(z)}, \quad (s,x) \in \mathcal{X}^{m+k},\ u \in \mathbb{R}^m,\ z \in (0,+\infty)^k$$

drive how the conditioning terms $\{Z(x_j) = z_j\}$ $(j = 1, \ldots, k)$ are met; see Steps 1 and 2 in Theorem 1.

The second key result states that, conditionally on $Z(x) = z$, the Poisson process Φ can be decomposed into two independent point processes, say $\Phi = \Phi^- \cup \Phi^+$, where

$$\Phi^- = \{\varphi \in \Phi\colon \varphi(x_i) < z_i,\ i = 1, \ldots, k\}, \quad \Phi^+ = \bigcup_{i=1}^{k} \{\varphi \in \Phi\colon \varphi(x_i) = z_i\}.$$

Before introducing a procedure for obtaining conditional realizations of max-stable processes, we define some notation and point out some connections with the pioneering work of Wang & Stoev (2011).

A function $\varphi \in \Phi^+$ such that $\varphi(x_i) = z_i$ for some $i \in \{1, \ldots, k\}$ is called an extremal function associated to x_i and is denoted by $\varphi^+_{x_i}$. It is easy to show that there exists almost surely a unique extremal function associated to x_i. Although $\Phi^+ = \{\varphi^+_{x_1}, \ldots, \varphi^+_{x_k}\}$ almost surely, it could happen that a single extremal function contributes to the random vector $Z(x)$ at several locations x_i, e.g., $\varphi^+_{x_1} = \varphi^+_{x_2}$. To take such repetitions into account, we define a random partition $\theta = (\theta_1, \ldots, \theta_\ell)$ of the set $\{x_1, \ldots, x_k\}$ into $\ell = |\theta|$ blocks, and define extremal functions $(\varphi^+_1, \ldots, \varphi^+_\ell)$ such that $\Phi^+ = \{\varphi^+_1, \ldots, \varphi^+_\ell\}$ and $\varphi^+_j(x_i) = z_i$ if $x_i \in \theta_j$ while $\varphi^+_j(x_i) < z_i$ if $x_i \notin \theta_j$ $(i = 1, \ldots, k;\ j = 1, \ldots, \ell)$. Wang & Stoev (2011) call the partition θ the hitting scenario. The set of all possible partitions of $\{x_1, \ldots, x_k\}$, denoted by $\mathscr{P}_k$, identifies all possible hitting scenarios.

From a simulation perspective, the fact that Φ^- and Φ^+ are independent given $Z(x) = z$ is especially convenient and suggests a three-step procedure for sampling from the conditional distribution of Z given $Z(x) = z$.

THEOREM 1. *Suppose that the point process Φ is regular and let $(x, s) \in \mathcal{X}^{k+m}$. For $\tau = (\tau_1, \ldots, \tau_\ell) \in \mathscr{P}_k$ and $j = 1, \ldots, \ell$, define $I_j = \{i : x_i \in \tau_j\}$, $x_{\tau_j} = (x_i)_{i \in I_j}$, $z_{\tau_j} = (z_i)_{i \in I_j}$, $x_{\tau_j^c} = (x_i)_{i \notin I_j}$ and $z_{\tau_j^c} = (z_i)_{i \notin I_j}$. Consider the following three-step procedure.*

Step 1. Draw a random partition $\theta \in \mathscr{P}_k$ with distribution

$$\pi_x(z, \tau) = \mathrm{pr}\{\theta = \tau \mid Z(x) = z\} = \frac{1}{C(x,z)} \prod_{j=1}^{|\tau|} \lambda_{x_{\tau_j}}(z_{\tau_j}) \int_{\{u_j < z_{\tau_j^c}\}} \lambda_{x_{\tau_j^c}|x_{\tau_j}, z_{\tau_j}}(u_j)\, \mathrm{d}u_j,$$

where the normalization constant is

$$C(x,z) = \sum_{\tilde{\tau} \in \mathscr{P}_k} \prod_{j=1}^{|\tilde{\tau}|} \lambda_{x_{\tilde{\tau}_j}}(z_{\theta_j}) \int_{\{u_j < z_{\tilde{\tau}_j^c}\}} \lambda_{x_{\tilde{\tau}_j^c}|x_{\tilde{\tau}_j}, z_{\tilde{\tau}_j}}(u_j)\, \mathrm{d}u_j.$$

Step 2. Given $\tau = (\tau_1, \ldots, \tau_\ell)$, draw ℓ independent random vectors $\varphi_1^+(s), \ldots, \varphi_\ell^+(s)$ from the distribution

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

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

$$C_j = \int 1_{\{u < z_{\tau_j^c}\}} \lambda_{(s, x_{\tau_j^c}) | x_{\tau_j}, z_{\tau_j}}(v, u)\, \mathrm{d}u\, \mathrm{d}v,$$

and define the random vector $Z^+(s) = \max_{j=1,\ldots,\ell} \varphi_j^+(s)$.

Step 3. Independently draw a Poisson point process $\{\zeta_i\}_{i \geqslant 1}$ on $(0, \infty)$ with intensity $\zeta^{-2}\,\mathrm{d}\zeta$ and $\{Y_i\}_{i \geqslant 1}$ independent copies of Y, and define the random vector $Z^-(s) = \max_{i \geqslant 1} \zeta_i Y_i(s) 1_{\{\zeta_i Y_i(x) < z\}}$.

Then the random vector $\tilde{Z}(s) = \max\{Z^+(s), Z^-(s)\}$ *follows the conditional distribution of* $Z(s)$ *given* $Z(x) = z$.

The corresponding conditional cumulative distribution function is

$$\mathrm{pr}\{Z(s) \leqslant a \mid Z(x) = z\} = \left\{ \sum_{\tau \in \mathscr{P}_k} \pi_x(z, \tau) \prod_{j=1}^{|\tau|} F_{\tau, j}(a) \right\} \frac{\mathrm{pr}\{Z(s) \leqslant a, Z(x) \leqslant z\}}{\mathrm{pr}\{Z(x) \leqslant z\}}, \qquad (2)$$

where

$$F_{\tau, j}(a) = \mathrm{pr}\{\varphi_j^+(s) \leqslant a \mid Z(x) = z,\ \theta = \tau\} = \frac{\int_{\{u < z_{\tau_j^c}, v < a\}} \lambda_{(s, x_{\tau_j^c}) | x_{\tau_j}, z_{\tau_j}}(v, u)\, \mathrm{d}u\, \mathrm{d}v}{\int_{\{u < z_{\tau_j^c}\}} \lambda_{x_{\tau_j^c} | x_{\tau_j}, z_{\tau_j}}(u)\, \mathrm{d}u}.$$

The first term on the right-hand side of (2) comes from Steps 1 and 2, while the ratio is a consequence of Step 3. It is clear from (2) that the conditional random field $Z \mid \{Z(x) = z\}$ is not max-stable.

2·2. *Distribution of the extremal functions*

In this section we derive closed forms for the intensity function $\lambda_x(z)$ and the conditional intensity function $\lambda_{s|x,z}(u)$ for two widely used max-stable processes: the Brown–Resnick (Brown & Resnick, 1977; Kabluchko et al., 2009) and the Schlather (2002) processes. Details of the derivations are given in the Appendix.

The Brown–Resnick process corresponds to the case $Y(x) = \exp\{W(x) - \gamma(x)\}$ $(x \in \mathbb{R}^d)$ in (1), where W is a centred Gaussian process with stationary increments and semivariogram γ, such that $W(o) = 0$ almost surely, where o denotes the origin of $\mathbb{R}^d$. For $x \in \mathcal{X}^k$, provided the covariance matrix Σ_x of the random vector $W(x)$ is positive definite, the intensity function is

$$\lambda_x(z) = C_x \exp\left(-\frac{1}{2} \log z^{\mathrm{T}} Q_x \log z + L_x \log z \right) \prod_{i=1}^k z_i^{-1}, \quad z \in (0, \infty)^k,$$

with $1_k = (1)_{i=1,\ldots,k}$, $\gamma_x = \{\gamma(x_i)\}_{i=1,\ldots,k}$,

$$Q_x = \Sigma_x^{-1} - \frac{\Sigma_x^{-1} 1_k 1_k^{\mathrm{T}} \Sigma_x^{-1}}{1_k^{\mathrm{T}} \Sigma_x^{-1} 1_k}, \quad L_x = \left(\frac{1_k^{\mathrm{T}} \Sigma_x^{-1} \gamma_x - 1}{1_k^{\mathrm{T}} \Sigma_x^{-1} 1_k} 1_k - \gamma_x \right)^{\mathrm{T}} \Sigma_x^{-1},$$

$$C_x = (2\pi)^{(1-k)/2} |\Sigma_x|^{-1/2} (1_k^{\mathrm{T}} \Sigma_x^{-1} 1_k)^{-1/2} \exp\left\{ \frac{1}{2} \frac{(1_k^{\mathrm{T}} \Sigma_x^{-1} \gamma_x - 1)^2}{1_k^{\mathrm{T}} \Sigma_x^{-1} 1_k} - \frac{1}{2} \gamma_x^{\mathrm{T}} \Sigma_x^{-1} \gamma_x \right\}.$$

For all $(s, x) \in \mathcal{X}^{m+k}$ and $(u, z) \in (0, \infty)^{m+k}$, provided the covariance matrix $\Sigma_{(s,x)}$ is positive definite, the conditional intensity function corresponds to a multivariate lognormal probability density function

$$\lambda_{s|x,z}(u) = (2\pi)^{-m/2} |\Sigma_{s|x}|^{-1/2} \exp\left\{ -\frac{1}{2} (\log u - \mu_{s|x,z})^{\mathrm{T}} \Sigma_{s|x}^{-1} (\log u - \mu_{s|x,z}) \right\} \prod_{i=1}^{m} u_i^{-1}$$

where $\mu_{s|x,z} \in \mathbb{R}^m$ and $\Sigma_{s|x}$ are, respectively, the mean and covariance matrix of the underlying normal distribution, given by

$$\mu_{s|x,z} = \left\{ L_{(s,x)} J_{m,k} - \log z^{\mathrm{T}} \tilde{J}_{m,k}^{\mathrm{T}} Q_{(s,x)} J_{m,k} \right\} \Sigma_{s|x}, \quad \Sigma_{s|x}^{-1} = J_{m,k}^{\mathrm{T}} Q_{(s,x)} J_{m,k}$$

with

$$J_{m,k} = \begin{pmatrix} \mathrm{I}_m \\ 0_{k,m} \end{pmatrix}, \quad \tilde{J}_{m,k} = \begin{pmatrix} 0_{m,k} \\ \mathrm{I}_k \end{pmatrix},$$

where I_k denotes the $k \times k$ identity matrix and $0_{m,k}$ the $m \times k$ null matrix.

The Schlather process considers the case of $Y(x) = (2\pi)^{1/2} \max\{0, \varepsilon(x)\}$ $(x \in \mathbb{R}^d)$ in (1), where ε is a standard Gaussian process with correlation function ρ. The associated point process Φ is not regular, and it is more convenient to consider the equivalent representation where $Y(x) = (2\pi)^{1/2} \varepsilon(x)$ $(x \in \mathbb{R}^d)$. For $x \in \mathcal{X}^k$, provided the covariance matrix Σ_x of the random vector $\varepsilon(x)$ is positive definite, the intensity function is

$$\lambda_x(z) = \pi^{-(k-1)/2} |\Sigma_x|^{-1/2} a_x(z)^{-(k+1)/2} \Gamma\left(\frac{k+1}{2} \right), \quad z \in \mathbb{R}^k,$$

where $a_x(z) = z^{\mathrm{T}} \Sigma_x^{-1} z$.

For $(s, x) \in \mathcal{X}^{m+k}$ and $(u, z) \in \mathbb{R}^{m+k}$, provided the covariance matrix $\Sigma_{(s,x)}$ is positive definite, the conditional intensity function $\lambda_{s|x,z}(u)$ corresponds to the density of a multivariate Student distribution with $k+1$ degrees of freedom, location parameter $\mu = \Sigma_{s:x} \Sigma_x^{-1} z$ and scale matrix

$$\tilde{\Sigma} = \frac{a_x(z)}{k+1} \left(\Sigma_s - \Sigma_{s:x} \Sigma_x^{-1} \Sigma_{x:s} \right), \quad \Sigma_{(s,x)} = \begin{pmatrix} \Sigma_s & \Sigma_{s:x} \\ \Sigma_{x:s} & \Sigma_x \end{pmatrix}.$$

3. Markov chain Monte Carlo sampler

Section 2 introduced a procedure for obtaining realizations from the regular conditional distribution of max-stable processes. This sampling scheme amounts to sampling from a discrete distribution whose state space corresponds to all possible partitions of the set of conditioning points; see Step 1 in Theorem 1. Hence, even for a moderate number k of conditioning locations,

the state space $\mathscr{P}_k$ will be very large and the distribution $\pi_x(z, \cdot)$ cannot be computed exactly. A Gibbs sampler is especially convenient for Monte Carlo sampling from $\pi_x(z, \cdot)$.

For $\tau \in \mathscr{P}_k$, let τ_{-j} be the restriction of τ to the set $\{x_1, \ldots, x_k\} \setminus \{x_j\}$. Our goal is to simulate from the conditional distribution

$$\mathrm{pr}(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \tag{3}$$

where $\theta \in \mathscr{P}_k$ is a random partition which follows the target distribution $\pi_x(z, \cdot)$.

Since the number of possible updates is always less than k, a combinatorial explosion is avoided. Indeed, for $\tau \in \mathscr{P}_k$ of size ℓ, the number of partitions $\tau^* \in \mathscr{P}_k$ such that $\tau^*_{-j} = \tau_{-j}$ for some $j \in \{1, \ldots, k\}$ is

$$b^+ = \begin{cases} \ell, & \{x_j\} \text{ is a partitioning set of } \tau, \\ \ell + 1, & \{x_j\} \text{ is not a partitioning set of } \tau, \end{cases}$$

since the point x_j may be re-allocated to any partitioning set of τ_{-j} or to a new one.

As an example, consider the set $\{x_1, x_2, x_3\}$ and let $\tau = (\{x_1, x_2\}, \{x_3\})$. Then the possible partitions τ^* such that $\tau^*_{-2} = \tau_{-2}$ are $(\{x_1, x_2\}, \{x_3\})$, $(\{x_1\}, \{x_2\}, \{x_3\})$ and $(\{x_1\}, \{x_2, x_3\})$, while there exist only two partitions such that $\tau^*_{-3} = \tau_{-3}$, namely $(\{x_1, x_2\}, \{x_3\})$ and $(\{x_1, x_2, x_3\})$.

The distribution (3) has nice properties. For all $\tau^* \in \mathscr{P}_k$ such that $\tau^*_{-j} = \tau_{-j}$, we have

$$\mathrm{pr}(\theta = \tau^* \mid \theta_{-j} = \tau_{-j}) = \frac{\pi_x(z, \tau^*)}{\sum_{\tilde{\tau} \in \mathscr{P}_k} \pi_x(z, \tilde{\tau}) 1_{\{\tilde{\tau}_{-j} = \tau_{-j}\}}} \propto \frac{\prod_{j=1}^{|\tau^*|} w_{\tau^*, j}}{\prod_{j=1}^{|\tau|} w_{\tau, j}} \tag{4}$$

where

$$w_{\tau, j} = \lambda_{x_{\tau_j}}(z_{\tau_j}) \int_{\{u < z_{\tau_j^c}\}} \lambda_{x_{\tau_j^c} | x_{\tau_j}, z_{\tau_j}}(u) \, \mathrm{d}u.$$

Since many factors cancel out on the right-hand side of (4), the Gibbs sampler is convenient.

The most computationally demanding part of (4) is the evaluation of the integral

$$\int_{\{u < z_{\tau_j^c}\}} \lambda_{x_{\tau_j^c} | x_{\tau_j}, z_{\tau_j}}(u) \, \mathrm{d}u.$$

For the Brown–Resnick and Schlather processes, we follow Genz (1992) and compute these probabilities using a separation-of-variables method which transforms the original integration problem to the unit hypercube. A quasi-Monte Carlo scheme and antithetic variables are used to improve efficiency.

Since it may not be obvious how to implement a Gibbs sampler whose target distribution has support $\mathscr{P}_k$, the remainder of this section gives practical details. For any fixed locations $x_1, \ldots, x_k \in \mathcal{X}$, we first describe how each partition of $\{x_1, \ldots, x_k\}$ is stored. To illustrate this, consider the set $\{x_1, x_2, x_3\}$ and the partition $(\{x_1, x_2\}, \{x_3\})$. This partition is defined as $(1, 1, 2)$, indicating that x_1 and x_2 belong to the partitioning set labelled 1 and x_3 belongs to the partitioning set labelled 2. There are several equivalent ways of denoting this partition; for instance, one could write $(2, 2, 1)$ or $(1, 1, 3)$. Since there is a one-to-one mapping between $\mathscr{P}_k$ and the set

$$\mathscr{P}^*_k = \left\{ (a_1, \ldots, a_k) \colon i \in \{2, \ldots, k\}, \ 1 = a_1 \leqslant a_i \leqslant \max_{1 \leqslant j < i} a_j + 1, \ a_i \in \mathbb{Z} \right\},$$

we shall restrict our attention to the partitions that live in $\mathscr{P}_k^*$; going back to our example, we see that (1, 1, 2) is valid but (2, 2, 1) and (1, 1, 3) are not.

For $\tau \in \mathscr{P}_k^*$ of size ℓ, let $r_1 = \sum_{i=1}^k 1_{\{\tau_i = a_j\}}$ and $r_2 = \sum_{i=1}^k 1_{\{\tau_i = b\}}$; in other words, r_1 and r_2 are the numbers of conditioning locations that belong to the partitioning sets a_j and b, where $b \in \{1, \ldots, b^+\}$ with

$$b^+ = \begin{cases} \ell, & r_1 = 1, \\ \ell + 1, & r_1 \neq 1. \end{cases}$$

Then the conditional probability distribution (4) satisfies

$$\mathrm{pr}(\tau_j = b \mid \tau_i = a_i,\ i \neq j) \propto \begin{cases} 1, & b = a_j, \quad (5a) \\ w_{\tau^*,b}/(w_{\tau,b} w_{\tau,a_j}), & r_1 = 1, r_2 \neq 0, b \neq a_j\ , \quad (5b) \\ w_{\tau^*,b} w_{\tau^*,a_j}/(w_{\tau,b} w_{\tau,a_j}), & r_1 \neq 1, r_2 \neq 0, b \neq a_j, \quad (5c) \\ w_{\tau^*,b} w_{\tau^*,a_j}/w_{\tau,a_j}, & r_1 \neq 1, r_2 = 0, b \neq a_j, \quad (5d) \end{cases}$$

where $\tau^* = (a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_k)$. Although τ^* may not belong to $\mathscr{P}_k^*$, it corresponds to a unique partition of $\mathscr{P}_k$ and we can use the bijection $\mathscr{P}_k \to \mathscr{P}_k^*$ to recode τ^* into an element of $\mathscr{P}_k^*$. The event $\{r_1 = 1, r_2 = 0, b \neq a_j\}$ is missing from (5a)–(5d), because $\{r_1 = 1, r_2 = 0\}$ implies $\tau^* = \tau$, where the equality has to be understood in terms of elements of $\mathscr{P}_k$, and this case is already covered by (5a).

Once these conditional weights have been computed, the Gibbs sampler proceeds by updating each element of τ successively. We use a random scan implementation (Liu et al., 1995). One iteration selects an element of τ at random, say $p = (p_1, \ldots, p_k)$, according to a given distribution, and then updates this element. Throughout this paper we will use the uniform random scan Gibbs sampler for which the selection distribution is a discrete uniform distribution, i.e., $p = (k^{-1}, \ldots, k^{-1})$.

4. Simulation study

This section shows how we checked whether our algorithm is able to produce realistic conditional simulations of Brown–Resnick and Schlather processes. For each model, we consider three different sample path properties, as summarized in Table 1. These configurations were chosen so that the spatial dependence structures are similar to those of the applications given in § 5.

In order to check whether our sampling procedure is accurate, given a single conditional event $\{Z(x) = z\}$ for each configuration, we generated 1000 conditional realizations with standard Gumbel margins. Figure 1 shows the pointwise sample quantiles obtained from the 1000 simulated paths and compares them to unit Gumbel quantiles. As expected, the conditional sample paths inherit the regularity driven by the shape parameter κ, and there is less variability in regions close to conditioning locations. Since the Brown–Resnick processes considered are ergodic (Kabluchko & Schlather, 2010), for regions far away from any conditioning location the sample quantiles converge to those of a standard Gumbel distribution, indicating that the conditional event has no influence. This is not the case for the nonergodic Schlather process. Most of the time, the sample paths used to get the conditional events belong to the 95% pointwise confidence intervals, which confirms that our sampling procedure seems to be accurate.

Table 2 gives timings for conditional simulations of max-stable processes on a 50×50 grid with a varying number of conditioning locations. Due to the combinatorial complexity of the partition set $\mathscr{P}_k$, computation times increase rapidly with respect to k, the number of conditioning

Table 1. *Sample path properties of the max-stable models. For the Brown–Resnick model the variogram parameters are set to ensure that the extremal coefficient function satisfies* $\theta(115) = 1{\cdot}7$*, while for the Schlather model the correlation function parameters are set to ensure that* $\theta(100) = 1{\cdot}5$

	Brown–Resnick: $\gamma(h) = (h/\lambda)^\kappa$			Schlather: $\rho(h) = \exp\{-(h/\lambda)^\kappa\}$		
	γ_1: Very wiggly	γ_2: Wiggly	γ_3: Smooth	ρ_1: Very wiggly	ρ_2: Wiggly	ρ_3: Smooth
λ	25	54	69	208	144	128
κ	0·5	1·0	1·5	0·5	1·0	1·5

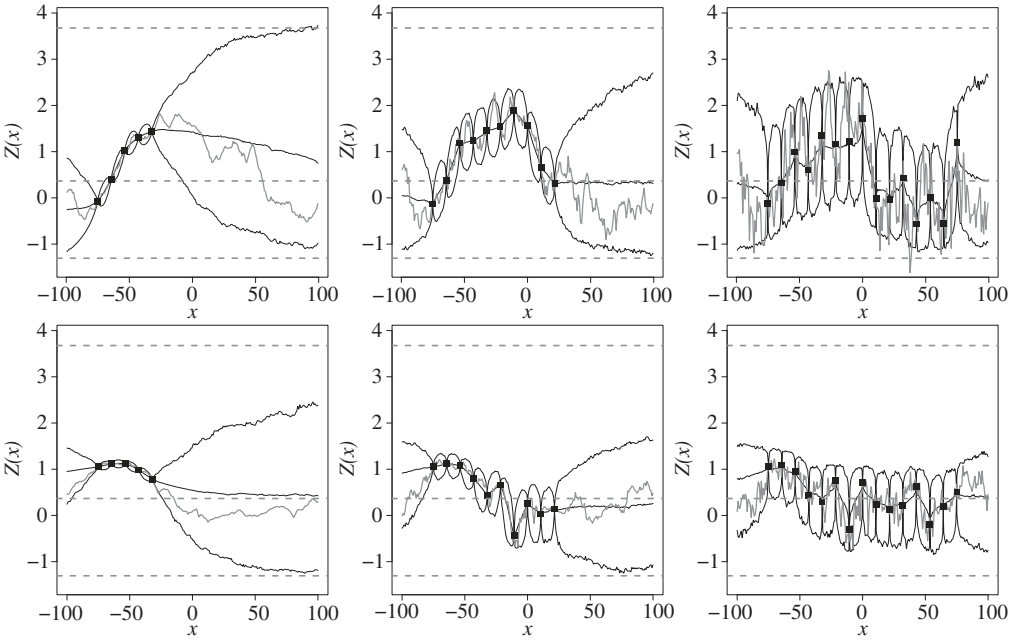

Fig. 1. Pointwise sample quantiles estimated from 1000 conditional simulations of max-stable processes with standard Gumbel margins, where the number of conditioning locations is $k = 5$, 10 or 15. The top row shows results for the Brown–Resnick models with semivariograms γ_3, γ_2 and γ_1 from left to right. The bottom row shows results for the Schlather models with correlation functions ρ_3, ρ_2 and ρ_1 from left to right. In each panel, the solid black lines represent the pointwise 0·025, 0·5 and 0·975 sample quantiles, and the dashed grey lines represent the corresponding quantiles of a standard Gumbel distribution. The squares show the conditioning points $\{(x_i, z_i)\}_{i=1,\ldots,k}$. The solid grey lines show the simulated paths used to obtain the conditioning events.

Table 2. *Timings for conditional simulations of max-stable processes on a* 50×50 *grid defined on the square* $[0, 100 \times 2^{1/2}]^2$ *for a varying number k of conditioning locations uniformly distributed over the region. The times, in seconds, are mean values over* 100 *conditional simulations; standard deviations are reported in parentheses*

	Brown–Resnick: $\gamma(h) = (h/25)^{0{\cdot}5}$				Schlather: $\rho(h) = \exp\{-(h/208)^{0{\cdot}50}\}$			
	Step 1	Step 2	Step 3	Overall	Step 1	Step 2	Step 3	Overall
$k = 5$	0·21 (0·01)	49 (11)	1·4 (0·1)	50 (11)	1·4 (0·02)	1·9 (0·7)	0·9 (0·3)	4·2 (0·8)
$k = 10$	8 (2)	76 (18)	1·4 (0·1)	85 (19)	12 (4)	2·4 (0·8)	1·0 (0·3)	15 (4)
$k = 25$	95 (38)	117 (30)	1·4 (0·1)	214 (61)	86 (42)	4 (1)	1·0 (0·3)	90 (43)
$k = 50$	583 (313)	348 (391)	1·5 (0·1)	931 (535)	367 (222)	62 (113)	1·0 (0·3)	430 (262)

Conditional simulations with $k = 5$ do not use a Gibbs sampler.

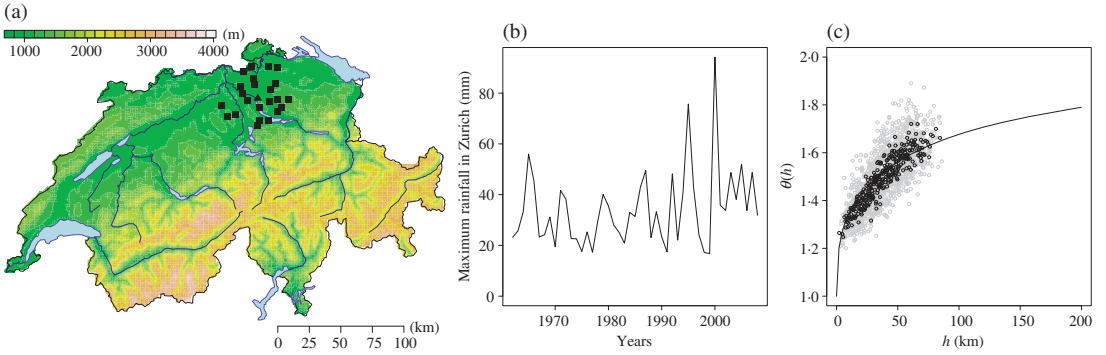

Fig. 2. (a) Map of Switzerland showing the stations of the 24 rainfall gauges used for the analysis, with an insert giving the altitude; the station marked with a triangle corresponds to Zurich. (b) Summer maximum rainfall in Zurich for 1962–2008. (c) Comparison of the pairwise extremal coefficient estimates for the 51 original weather stations and the extremal coefficient function derived from a fitted Brown–Resnick process having semivariogram $\gamma(h) = (h/\lambda)^{\kappa}$; the grey points are pairwise estimates, the black points are binned estimates, and the curve is the fitted extremal coefficient function.

points. It is, however, reassuring that the algorithm is tractable when $k \leqslant 50$, as this covers many practical situations.

5. Applications

5·1. *Extreme precipitation around Zurich*

The data considered here were previously analysed by Davison et al. (2012), who showed that Brown–Resnick processes are among the best models for these data. The data consist of maximum summer rainfall measurements for the years 1962–2008 at 51 weather stations in the plateau region of Switzerland, provided by the national meteorological service, MeteoSuisse. To ensure strong dependence between the conditioning locations, we took the 24 weather stations located at most 30 km from Zurich as conditioning locations, and set as the conditional values the rainfall amounts recorded in 2000, the year of the largest precipitation event recorded in the region between 1962 and 2008; see Fig. 2. The greatest distance between conditioning locations was around 55 km, and the shortest was just over 4 km.

A Brown–Resnick process having semivariogram $\gamma(h) = (h/\lambda)^{\kappa}$ was fitted using the maximum pairwise likelihood estimator introduced by Padoan et al. (2010) to simultaneously estimate the marginal parameters and the spatial dependence parameters λ and κ. As in Davison et al. (2012), the marginal parameters are $\eta(x)$, $\sigma(x)$ and $\xi(x)$, respectively the location, scale and shape parameters of the generalized extreme value distribution, described by $\eta(x) = \beta_{0,\eta} + \beta_{1,\eta}\mathrm{lon}(x) + \beta_{2,\eta}\mathrm{lat}(x)$, $\sigma(x) = \beta_{0,\sigma} + \beta_{1,\sigma}\mathrm{lon}(x) + \beta_{2,\sigma}\mathrm{lat}(x)$ and $\xi(x) = \beta_{0,\xi}$ where $\mathrm{lon}(x)$ and $\mathrm{lat}(x)$ denote the longitude and latitude of the stations at which the data are observed. The maximum pairwise likelihood estimates and their standard errors for λ and κ are, respectively, 38 (14) and 0·69 (0·07); these give a practical extremal range, i.e., the distance h_+ such that $\theta(h_+) = 1{\cdot}7$, of around 115 km; see Fig. 2(c).

Table 3 shows the distribution of the partition size estimated from a Markov chain of length 15 000. Around 65% of the time, the summer maxima observed at the 24 conditioning locations belonged to a single extremal function, i.e., only one storm event, and around 30% of the time the maxima were associated with two different storms. Since the simulated Markov chain keeps a trace of all the simulated partitions, we looked at the partitions of size two and saw that around

Table 3. *Distribution of the partition size for the rainfall and temperature data estimated from simulated Markov chains of lengths* 15 000 *and* 10 000*, respectively*

	Rainfall							Temperature				
Partition size	1	2	3	4	5	6	7–24	1	2	3	4	5–16
Frequency (%)	66·2	28·0	4·8	0·5	0·2	0·2	$< 0{\cdot}05$	2·47	21·55	64·63	10·74	0·61

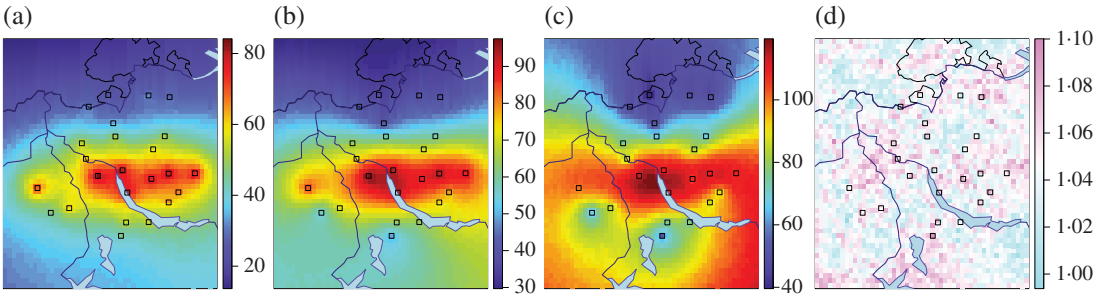

Fig. 3. (a)–(c) Maps on a 50×50 grid of the pointwise 0·025, 0·5 and 0·975 sample quantiles for rainfall (mm) obtained from 10 000 conditional simulations of Brown–Resnick processes having semivariogram $\gamma(h) = (h/38)^{0{\cdot}69}$. (d) Plot of the ratios of the widths of the pointwise confidence intervals with and without taking estimation uncertainties into account. The squares show the conditioning locations.

65% of the time, at least one of the four northern conditioning locations was impacted by one extremal function, while the remaining 20 locations were always influenced by another one.

In Fig. 3, panels (a)–(c) show the pointwise 0·025, 0·5 and 0·975 sample quantiles obtained from 10 000 conditional simulations of our fitted Brown–Resnick process. The conditional median provides a point estimate for the rainfall at an ungauged location, and the 0·025 and 0·975 conditional quantiles give a 95% pointwise confidence interval. As indicated by Fig. 1, the shape parameter κ has a major impact on the regularity of paths and on the width of the confidence interval. The value $\hat{\kappa} \approx 0{\cdot}69$ corresponds to very wiggly sample paths and wider confidence intervals. To assess the effect of parameter uncertainties on conditional simulations, the ratios of the widths of the confidence intervals with and without parameter uncertainty are plotted in Fig. 3(d). The uncertainties were taken into account by sampling from the asymptotic distribution of the maximum composite likelihood estimator and drawing one conditional simulation for each realization. These ratios show no clear spatial pattern, and the width of the confidence interval is increased by at most 10%.

5·2. *Extreme temperatures in Switzerland*

The data considered here are annual maximum temperatures recorded at 16 sites in Switzerland during the period 1961–2005; see Fig. 4. Following Davison & Gholamrezaee (2012), we fit a Schlather process with an isotropic powered exponential correlation function and trend surfaces $\eta(x) = \beta_{0,\eta} + \beta_{1,\eta}\,\mathrm{alt}(x)$, $\sigma(x) = \beta_{0,\sigma}$ and $\xi(x) = \beta_{0,\xi} + \beta_{1,\xi}\,\mathrm{alt}(x)$, where $\eta(x)$, $\sigma(x)$ and $\xi(x)$ are the location, scale and shape parameters of the generalized extreme value distribution at location x and $\mathrm{alt}(x)$ denotes the altitude above mean sea level in kilometres. The spatial dependence parameter estimates and their standard errors are $\hat{\lambda} = 260$ (149) and $\hat{\kappa} = 0{\cdot}52$ (0·12), which yield a fitted extremal coefficient function similar to our test case ρ_3 in § 4.

In 2003, western Europe experienced a heat wave, believed to be the most severe one recorded since 1540 (Beniston, 2004). Switzerland was seriously affected by this event: the nationwide record temperature of 41·5°C was recorded that year in Grono, Graubünden, near Lugano. For

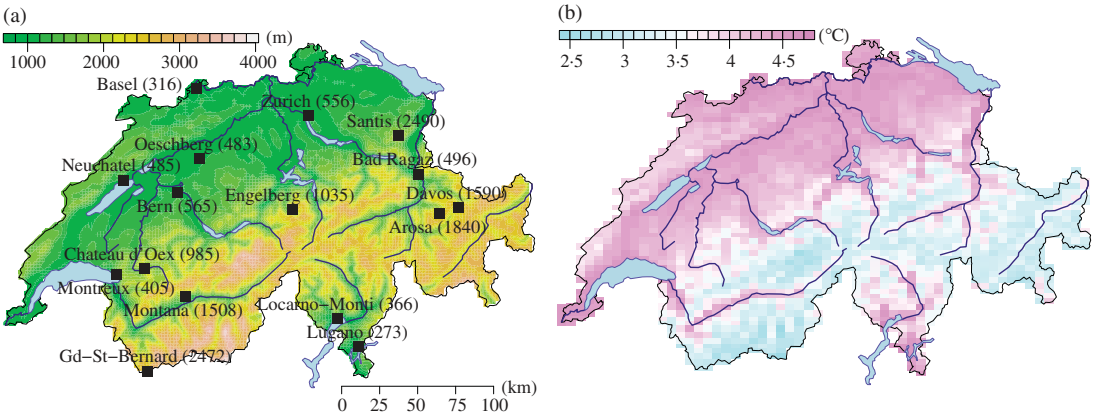

Fig. 4. (a) Topographical map of Switzerland showing the sites and altitudes, in metres above sea level, of 16 weather stations for which annual maximum temperature data were available. (b) Map of temperature anomalies, in degrees Celsius, i.e., the differences between the pointwise medians obtained from 10 000 conditional simulations and the unconditional medians estimated from the fitted Schlather process.

our analysis, we condition on the maximum temperatures observed in 2003. Based on the fitted Schlather model, we simulate a Markov chain of effective length 10 000, with a burn-in period of length 500 and a thinning lag of 100 iterations. The distribution of the partition size estimated from these Markov chains is shown in Table 3. Around 90% of the time, the conditional realizations were a consequence of at most three extremal functions. Since our original observations were not summer maxima but maximum daily values, a close inspection of the time series in year 2003 reveals that the hottest temperatures occurred between the 11th and 13th of August, corroborating to some extent the distribution shown in Table 3.

Figure 4(b) shows the spatial distribution of temperature anomalies, i.e., the differences between the pointwise conditional medians obtained from 10 000 conditional simulations and the pointwise unconditional medians estimated from the fitted Schlather model. As expected, the largest deviations occur in the plateau region of Switzerland, while appreciably smaller values are found in the Alps. The differences range from 2·5°C to 4·75°C and are consistent with the values reported by climatologists for mean values (Beniston, 2004).

Acknowledgement

M. Ribatet was partly funded by the MIRACCLE-GICC and McSim ANR projects. The authors thank MeteoSuisse and Dr S. A. Padoan for providing the precipitation data, Prof. A. C. Davison and Dr M. M. Gholamrezaee for providing the temperature data, and Dr M. Oesting for pointing out an error in the computations for the Brown–Resnick model.

Appendix

Brown–Resnick model

For all $x \in \mathcal{X}^k$ and Borel sets $A \subset \mathbb{R}^k$,

$$\Lambda_x(A) = \int_0^\infty \mathrm{pr}[\zeta \exp\{W(x) - \gamma(x)\} \in A]\zeta^{-2}\,\mathrm{d}\zeta = \int_0^\infty \int_{\mathbb{R}^k} 1_{[\zeta \exp\{y-\gamma(x)\}\in A]} f_x(y)\,\mathrm{d}y\,\zeta^{-2}\,\mathrm{d}\zeta,$$

where f_x denotes the density of the random vector $W(x)$, i.e., a centred Gaussian random vector with covariance matrix Σ_x and variance $2\gamma(x)$. The change of variables $z = \zeta \exp\{y - \gamma(x)\}$ and $r = \log \zeta$ yields

$$\Lambda_x(A) = \int_{-\infty}^{\infty} \int_A f_x\{\log z - r + \gamma(x)\} \prod_{i=1}^{k} z_i^{-1} \, \mathrm{d}z \exp(-r) \, \mathrm{d}r = \int_A \lambda_x(z) \, \mathrm{d}z,$$

with

$$\lambda_x(z) = \prod_{i=1}^{k} z_i^{-1} \int_{-\infty}^{\infty} f_x\{\log z - r + \gamma(x)\} \exp(-r) \, \mathrm{d}r.$$

Since

$$f_x\{\log z - r + \gamma(x)\} \exp(-r) = (2\pi)^{-k/2} |\Sigma_x|^{-1/2} \exp\left\{-\frac{1}{2} P(r)\right\}$$

with

$$P(r) = r^2 \, 1_k^{\mathrm{T}} \, \Sigma_x^{-1} 1_k - 2r[1_k^{\mathrm{T}} \, \Sigma_x^{-1}\{\log z + \gamma(x)\} - 1] + \{\log z + \gamma(x)\}^{\mathrm{T}} \, \Sigma_x^{-1}\{\log z + \gamma(x)\},$$

standard computations for Gaussian integrals give

$$\lambda_x(z) = C_x \exp\left(-\frac{1}{2} \log z^{\mathrm{T}} Q_x \log z + L_x \log z\right) \prod_{i=1}^{k} z_i^{-1}.$$

The conditional intensity function is

$$\lambda_{s|x,z}(u) = \frac{C_{(s,x)}}{C_x} \exp\left\{-\frac{1}{2} \log(u, z)^{\mathrm{T}} Q_{(s,x)} \log(u, z) + L_{(s,x)} \log(u, z) + \frac{1}{2} \log z^{\mathrm{T}} Q_x \log z - L_x \log z\right\} \prod_{i=1}^{m} u_i^{-1}$$

and, since $\log(u, z) = J_{m,k} \log u + \tilde{J}_{m,k} \log z$, it is not difficult to show that

$$\lambda_{s|x,z}(u) = \frac{C_{(s,x)}}{C_x} \exp\left\{-\frac{1}{2}(\log u - \mu_{s|x,z})^{\mathrm{T}} \Sigma_{s|x}^{-1} (\log u - \mu_{s|x,z})\right\} \prod_{i=1}^{m} u_i^{-1}.$$

Finally, the relation $C_{(s,x)}/C_x = (2\pi)^{-m/2} |\Sigma_{s|x}|^{-1/2}$ is a simple consequence of the normalization $\int \lambda_{s|x,z}(u) \, \mathrm{d}u = 1$.

Schlather model

For all $x \in \mathcal{X}^k$ and Borel sets $A \subset \mathbb{R}^k$,

$$\Lambda_x(A) = \int_0^{\infty} \mathrm{pr}\{(2\pi)^{1/2} \zeta \varepsilon(x) \in A\} \zeta^{-2} \, \mathrm{d}\zeta = \int_0^{\infty} \int_{\mathbb{R}^k} 1_{\{(2\pi)^{1/2} \zeta y \in A\}} f_x(y) \, \mathrm{d}y \, \zeta^{-2} \, \mathrm{d}\zeta,$$

where f_x denotes the density of the random vector $\varepsilon(x)$, i.e., a centred Gaussian random vector with covariance matrix Σ_x. The change of variable $z = (2\pi)^{1/2}\zeta y$ gives

$$\begin{aligned}
\Lambda_x(A) &= (2\pi)^{-k/2} \int_0^\infty \int_A f_x\left\{\frac{z}{(2\pi)^{1/2}\zeta}\right\} \zeta^{-(k+2)}\,\mathrm{d}z\,\mathrm{d}\zeta \\
&= (2\pi)^{-k}|\Sigma_x|^{-1/2} \int_0^\infty \int_A \exp\left(-\frac{1}{4\pi\zeta^2}\, z^{\mathrm{T}}\,\Sigma_x^{-1} z\right) \zeta^{-(k+2)}\,\mathrm{d}z\,\mathrm{d}\zeta \\
&= (2\pi)^{-k}|\Sigma_x|^{-1/2} \int_A \frac{2\pi}{z^{\mathrm{T}}\,\Sigma_x^{-1} z}\, E(X^{k-1})\,\mathrm{d}z, \quad X \sim \mathrm{Wei}\left\{\left(\frac{4\pi}{z^{\mathrm{T}}\,\Sigma_x^{-1} z}\right)^{1/2},\ 2\right\} \\
&= (2\pi)^{-k}|\Sigma_x|^{-1/2} \int_A \frac{2\pi}{z^{\mathrm{T}}\,\Sigma_x^{-1} z} \left(\frac{4\pi}{z^{\mathrm{T}}\,\Sigma_x^{-1} z}\right)^{(k-1)/2} \Gamma\left(\frac{k+1}{2}\right)\mathrm{d}z \\
&= \int_A \lambda_x(z)\,\mathrm{d}z,
\end{aligned}$$

where $\lambda_x(z) = \pi^{-(k-1)/2}|\Sigma_x|^{-1/2} a_x(z)^{-(k+1)/2}\,\Gamma\{(k+1)/2\}$ and $a_x(z) = z^{\mathrm{T}}\Sigma_x^{-1}z$.

For all $u \in \mathbb{R}^m$, the conditional intensity function is

$$\lambda_{s|x,z}(u) = \pi^{-m/2}\frac{|\Sigma_{(s,x)}|^{-1/2}}{|\Sigma_x|^{-1/2}} \left\{\frac{a_{(s,x)}(u,z)}{a_x(z)}\right\}^{-(m+k+1)/2} a_x(z)^{-m/2}\frac{\Gamma\left(\frac{m+k+1}{2}\right)}{\Gamma\left(\frac{k+1}{2}\right)}.$$

We start by focusing on the ratio $a_{(s,x)}(u,z)/a_x(z)$. Since

$$\begin{pmatrix} \Sigma_s & \Sigma_{s:x} \\ \Sigma_{x:s} & \Sigma_x \end{pmatrix}^{-1} = \begin{pmatrix} (\Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s})^{-1} & -(\Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s})^{-1}\Sigma_{s:x}\Sigma_x^{-1} \\ -\Sigma_x^{-1}\Sigma_{x:s}(\Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s})^{-1} & \Sigma_x^{-1} + \Sigma_x^{-1}\Sigma_{x:s}(\Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s})^{-1}\Sigma_{s:x}\Sigma_x^{-1} \end{pmatrix},$$

straightforward algebra shows that

$$\frac{a_{(s,x)}(u,z)}{a_x(z)} = 1 + \frac{(u-\mu)^{\mathrm{T}}\tilde{\Sigma}^{-1}(u-\mu)}{k+1}, \quad \mu = \Sigma_{s:x}\Sigma_x^{-1}z, \quad \tilde{\Sigma} = \frac{a_x(z)}{k+1}(\Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s}).$$

We now simplify the ratio $|\Sigma_{(s,x)}|/|\Sigma_x|$. Using the fact that

$$\Sigma_{(s,x)} = \begin{pmatrix} \Sigma_s & \Sigma_{s:x} \\ \Sigma_{x:s} & \Sigma_x \end{pmatrix} = \begin{pmatrix} \mathrm{I}_m & \Sigma_{s:x} \\ 0_{k,m} & \Sigma_x \end{pmatrix} \begin{pmatrix} \Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s} & 0_{m,k} \\ \Sigma_x^{-1}\Sigma_{x:s} & \mathrm{I}_k \end{pmatrix}$$

together with some more algebra, we obtain

$$\frac{|\Sigma_{(s,x)}|}{|\Sigma_x|} = |\Sigma_s - \Sigma_{s:x}\Sigma_x^{-1}\Sigma_{x:s}| = \left\{\frac{k+1}{a_x(z)}\right\}^m |\tilde{\Sigma}|.$$

Using the previous two results, it is easily found that

$$\lambda_{s|x,z}(u) = \pi^{-m/2}(k+1)^{-m/2}|\tilde{\Sigma}|^{-1/2}\left\{1 + \frac{(u-\mu)^{\mathrm{T}}\tilde{\Sigma}^{-1}(u-\mu)}{k+1}\right\}^{-(m+k+1)/2} \frac{\Gamma\left(\frac{m+k+1}{2}\right)}{\Gamma\left(\frac{k+1}{2}\right)},$$

which corresponds to the density of a multivariate Student distribution with the stated parameters.

References

Beniston, M. (2004). The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations. *Geophys. Res. Lett.* **31**, L02202.

Brown, B. M. & Resnick, S. (1977). Extreme values of independent stochastic processes. *J. Appl. Prob.* **14**, 732–9.

Buishand, T. A., de Haan, L. & Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *Ann. Appl. Statist.* **2**, 624–42.

CHILÈS, J.-P. & DELFINER, P. (1999). *Geostatistics: Modelling Spatial Uncertainty*. New York: Wiley.
COOLEY, D., NAVEAU, P. & PONCET, P. (2006). Variograms for spatial max-stable random fields. In *Dependence in Probability and Statistics*, Ed. P. Bertail, P. Soulier, P. Doukhan, P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin & S. Zeger, pp. 373–90, vol. 187 of *Lecture Notes in Statistics*. New York: Springer.
DAVIS, R. & RESNICK, S. (1989). Basic properties and prediction of max-ARMA processes. *Adv. Appl. Prob.* **21**, 781–803.
DAVIS, R. & RESNICK, S. (1993). Prediction of stationary max-stable processes. *Ann. Appl. Prob.* **3**, 497–525.
DAVISON, A. C. & GHOLAMREZAEE, M. M. (2012). Geostatistics of extremes. *Proc. R. Soc. Lond.* A **468**, 581–608.
DAVISON, A. C., PADOAN, S. A. & RIBATET, M. (2012). Statistical modelling of spatial extremes. *Statist. Sci.* **7**, 161–86.
DE HAAN, L. (1984). A spectral representation for max-stable processes. *Ann. Prob.* **12**, 1194–204.
DE HAAN, L. & PEREIRA, T. T. (2006). Spatial extremes: Models for the stationary case. *Ann. Statist.* **34**, 146–68.
DE HAAN, L. & FEREIRA, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer.
GENTON, M. G., MA, Y. & SANG, H. (2011). On the likelihood function of Gaussian max-stable processes. *Biometrika* **98**, 481–8.
GENZ, A. (1992). Numerical computation of multivariate normal probabilities. *J. Comp. Graph. Statist.* **1**, 141–9.
KABLUCHKO, Z., SCHLATHER, M. & DE HAAN, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Prob.* **37**, 2042–65.
KABLUCHKO, Z. & SCHLATHER, M. (2010). Ergodic properties of max-infinitely divisible processes. *Stoch. Proces. Appl.* **120**, 281–95.
LIU, J., WONG, W. & KONG, A. (1995). Correlation structure and convergence rate of the Gibbs sampler with various scans. *J. R. Statist. Soc.* B **57**, 157–69.
MANDELBROT, B. (1982). *The Fractal Geometry of Nature*. New York: W. H. Freeman.
OESTING, M., KABLUCHKO, Z. & SCHLATHER, M. (2012). Simulation of Brown–Resnick processes. *Extremes* **15**, 89–107.
PADOAN, S. A., RIBATET, M. & SISSON, S. (2010). Likelihood-based inference for max-stable processes. *J. Am. Statist. Assoc.* **105**, 263–77.
R DEVELOPMENT CORE TEAM (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, ISBN 3-900051-07-0, http://www.R-project.org.
RIBATET, M. (2011). *SpatialExtremes: Modelling Spatial Extremes*. R package version 1.8-5.
SCHLATHER, M. (2002). Models for stationary max-stable random fields. *Extremes* **5**, 33–44.
SCHLATHER, M. (2012). *RandomFields: Simulation and Analysis of Random Fields*. R package version 2.0.54.
SCHLATHER, M. & TAWN, J. (2003). A dependence measure for multivariate and spatial extremes: Properties and inference. *Biometrika* **90**, 139–56.
WANG, Y. (2010). *maxLinear: Conditional Sampling for Max-Linear Models*. R package version 1.0.
WANG, Y. & STOEV, S. A. (2011). Conditional sampling for spectrally discrete max-stable random fields. *Adv. Appl. Prob.* **43**, 461–83.

[*Received December* 2011. *Revised September* 2012]