Adv. Appl. Prob. **46,** 478–495 (2014)
Printed in Northern Ireland
© *Applied Probability Trust* 2014

STATISTICAL INFERENCE FOR MAX-STABLE PROCESSES BY CONDITIONING ON EXTREME EVENTS

SEBASTIAN ENGELKE,* *Université de Lausanne and Georg-August-Universität Göttingen*
ALEXANDER MALINOWSKI,** *** *Universität Mannheim and Georg-August-Universität Göttingen*
MARCO OESTING **** AND
MARTIN SCHLATHER,** ***** *Universität Mannheim*

Abstract

In this paper we provide the basis for new methods of inference for max-stable processes ξ on general spaces that admit a certain incremental representation, which, in important cases, has a much simpler structure than the max-stable process itself. A corresponding peaks-over-threshold approach will incorporate all single events that are extreme in some sense and will therefore rely on a substantially larger amount of data in comparison to estimation procedures based on block maxima. Conditioning a process η in the max-domain of attraction of ξ on being *extremal*, several convergence results for the increments of η are proved. In a similar way, the shape functions of mixed moving maxima (M3) processes can be extracted from suitably conditioned single events η. Connecting the two approaches, transformation formulae for processes that admit both an incremental and an M3 representation are identified.

Keywords: Extreme value statistics; incremental representation; max-stable process; mixed moving maxima; peaks-over-threshold

2010 Mathematics Subject Classification: Primary 60G70
Secondary 62G32; 62E20

1. Introduction

The joint extremal behavior at multiple locations of some random process $\{\eta(t) : t \in T\}$, T an arbitrary index set, can be captured via its limiting *max-stable process*, assuming that the latter exists and is nontrivial everywhere. Then, for independent copies η_i of η, $i \in \mathbb{N}$, the functions $b_n : T \to \mathbb{R}$ and $c_n : T \to (0, \infty)$ can be chosen such that the convergence

$$\xi(t) = \lim_{n\to\infty} c_n(t)\Big(\max_{i=1}^{n} \eta_i(t) - b_n(t)\Big), \qquad t \in T, \tag{1.1}$$

Received 13 September 2012; revision received 29 June 2013.
* Postal address: Université de Lausanne, UNIL-Dorigny, Bâtiment Extranef, 1015 Lausanne, Switzerland. Email address: sebastian.engelke@unil.ch
** Postal address: Institut für Mathematik, Universität Mannheim, A5, 6, 68131 Mannheim, Germany.
*** Email address: malinows@math.uni-goettingen.de
**** Postal address: INRA, UMR 518 Math. Info. Appli., Rue Claude Bernard, 75005 Paris, France. Email address: marco.oesting@agroparistech.fr
***** Email address: schlather@math.uni-mannheim.de

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

holds in the sense of finite-dimensional distributions. The process ξ is said to be *max-stable* and η is in its max-domain of attraction (MDA). The theory of max-stable processes is mainly concerned with the dependence structure, while the marginals are usually assumed to be known. Even for finite-dimensional max-stable distributions, the space of possible dependence structures is uncountably infinite dimensional, and parametric models are required to find a balance between flexibility and analytical tractability [8], [26].

A general construction principle for max-stable processes was provided in [7] and [31]. Let $\sum_{i\in\mathbb{N}} \delta_{(U_i,S_i)}$ be a Poisson point process (PPP) on $(0,\infty)\times\mathcal{S}$ with intensity measure $u^{-2}\,\mathrm{d}u\cdot\nu(\mathrm{d}s)$, where $(\mathcal{S},\mathfrak{S})$ is an arbitrary measurable space and ν is a positive measure on $\mathcal{S}$. Furthermore, let $f\colon\mathcal{S}\times T\to[0,\infty)$ be a nonnegative function with $\int_{\mathcal{S}} f(s,t)\nu(\mathrm{d}s)=1$ for all $t\in T$. Then the process $\xi(t)=\max_{i\in\mathbb{N}} U_i f(S_i,t)$, $t\in T$, is max-stable and has standard Fréchet margins. In this paper we investigate two specific choices for f and $(\mathcal{S},\mathfrak{S},\nu)$, and consider processes that admit one of the resulting representations. First, let $\{W(t)\colon t\in T\}$ be a nonnegative stochastic process with $\mathbb{E}W(t)=1,\ t\in T$, and $W(t_0)=1$ almost surely (a.s.) for some point $t_0\in T$. The latter condition means that $W(t)$ simply describes the multiplicative increment of W with respect to the location t_0. For $(\mathcal{S},\mathfrak{S},\nu)$ being the canonical probability space for the sample paths of W and with $f(w,t)=w(t),\ w\in\mathcal{S}$, $t\in T$, we refer to

$$\xi(t)=\max_{i\in\mathbb{N}} U_i W_i(t), \qquad t\in T, \tag{1.2}$$

as the *incremental representation* of ξ, where the $\{W_i\}_{i\in\mathbb{N}}$ are independent copies of W. Since T is an arbitrary index set, the above definition covers multivariate extreme value distributions, i.e. $T=\{t_1,\dots,t_k\}$, as well as max-stable random fields, i.e. $T=\mathbb{R}^d$. For the second specification, let $\{F(t)\colon\ t\in\mathbb{R}^d\}$ be a stochastic process with sample paths in the space $C(\mathbb{R}^d)$ of nonnegative continuous functions such that

$$\mathbb{E}\int_{\mathbb{R}^d} F(t)\,\mathrm{d}t=1. \tag{1.3}$$

With $S_i=(T_i,F_i)$, $i\in\mathbb{N}$, in $\mathcal{S}=\mathbb{R}^d\times C(\mathbb{R}^d)$, intensity measure $\nu(\mathrm{d}t\times\mathrm{d}g)=\mathrm{d}t\mathbb{P}_F(\mathrm{d}g)$, and $f((t,g),s)=g(s-t),\ (t,g)\in\mathcal{S}$, we obtain the large and frequently used class of *mixed moving maxima (M3) processes*

$$\xi(t)=\max_{i\in\mathbb{N}} U_i F_i(t-T_i), \qquad t\in\mathbb{R}^d. \tag{1.4}$$

These processes are max-stable and stationary on $\mathbb{R}^d$ (see, for instance, [32]). The function F is called the *shape function of* ξ and can also be deterministic (e.g. in the case of the Smith process). In Smith's 'rainfall-storm' interpretation [31], U_i and T_i are the strength and center point of the ith storm, respectively, and $U_iF_i(t-T_i)$ represents the corresponding amount of rainfall at location t. In this case, $\xi(t)$ is the process of extremal precipitation. Throughout the paper, we will assume that ξ admits representation (1.2) or (1.4).

When independent and identically distributed (i.i.d.) realizations $\eta_1,\dots,\eta_n$ of η in the MDA of a max-stable process ξ are observed, a classical approach for the challenging problem of parametric inference on ξ is based on generating (approximate) realizations of ξ out of the data $\eta_1,\dots,\eta_n$ via componentwise block maxima and applying maximum likelihood (ML) estimation afterwards. A clear drawback of this method is that it ignores all information on large values that is contained in the order statistics below the within-block maximum.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

Furthermore, ML estimation needs to evaluate multivariate densities while, for many max-stable models, only the bivariate densities are known in closed form. Thus, composite likelihood approaches have been proposed [6], [25].

In this paper we present an alternative approach of inference that aims at using more of the data by extracting realizations of the processes W and F, respectively, which uniquely determine the max-stable process ξ. It is based on a *peaks-over-threshold* (POT) procedure known from univariate [11], [21] and multivariate [4], [17], [28] extreme value theory, and recently also applied in a functional setting [10], [15]. Contrasting the aggregation of multiple extreme events in ML estimation, in our approach, all *single extreme events*, i.e. those of $\eta_1, \ldots, \eta_n$ that are extreme in some sense, can be used to estimate W and F. The specification of a *single extreme event* will depend on the respective representations (1.2) and (1.4). From a statistical point of view, a further major advantage of this method is that W and F often admit a much simpler structure than the corresponding max-stable process itself. This facilitates parametric inference since full multivariate densities are available and, thus, using composite likelihood can be avoided. In [13], for instance, this concept is applied to derive estimators for the class of Brown–Resnick processes [3], [20], which have the form (1.2) by construction. With $a(n)$ being a sequence with $\lim_{n\to\infty} a(n) = \infty$ and η being in the MDA of a Brown–Resnick process, the convergence in distribution

$$\left(\frac{\eta(t_1)}{\eta(t_0)}, \ldots, \frac{\eta(t_k)}{\eta(t_0)} \;\middle|\; \eta(t_0) > a(n)\right) \xrightarrow{\mathrm{D}} (W(t_1), \ldots, W(t_k)) \tag{1.5}$$

for $t_0, t_1, \ldots, t_k \in T$, $k \in \mathbb{N}$, is established, where W is the corresponding log-Gaussian random field.

In the literature, similar approaches based on conditioning on extreme values are considered in various frameworks. For instance, Cooley *et al.* [5] proposed an approximation of the conditional distribution given large observations via the angular measure. Basrak and Segers [1] and Meinguet and Segers [23] respectively considered multivariate time series $(X_t)_{t\in\mathbb{Z}}$ and time series in general Banach spaces, rescaled and conditioned on $\|X_0\|$ being large. They provided equivalent conditions for the existence of the corresponding tail processes and its spectral decomposition, whereas in this paper, we explicitly calculate the limiting processes in a different framework. We generalize the convergence result (1.5) in two different aspects. Arbitrary nonnegative processes $\{W(t) \colon t \in T\}$ with $\mathbb{E}W(t) = 1$, $t \in T$, are considered, and both convergence of the conditional increments of η in the sense of finite-dimensional distributions and weak convergence in continuous function spaces are shown (Theorems 2.1 and 2.2). Similarly, in Section 3 we establish the main result for M3 processes (Theorem 3.1) by considering realizations of η around their (local) maxima. Since one and the same max-stable process ξ might admit both representations (1.2) and (1.4), we provide formulae for switching between them in Section 4. In Section 5 we focus on statistical applications of our results in spatial extreme value theory and give several important cases where explicit formulae of the estimators can be derived.

2. Incremental representation

Throughout this section, we suppose that $\{\xi(t) \colon t \in T\}$, T an arbitrary index set, is normalized to standard Fréchet margins and admits a representation

$$\xi(t) = \max_{i\in\mathbb{N}} U_i V_i(t), \qquad t \in T, \tag{2.1}$$

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

where $\sum_{i\in\mathbb{N}} \delta_{U_i}$ is a PPP on $(0, \infty)$ with intensity $u^{-2}\,\mathrm{d}u$, which we call the *Fréchet Poisson process* in the following. The $\{V_i\}_{i\in\mathbb{N}}$ are independent copies of a nonnegative stochastic process $\{V(t) : t \in T\}$ with $\mathbb{E}V(t) = 1$, $t \in T$. Note that (2.1) is slightly less restrictive than representation (1.2) in that we do not require that $V(t_0) = 1$ a.s. for some $t_0 \in T$.

For $k \in \mathbb{N}$ and $t_0, \ldots, t_k \in T$, the vector $\Xi = (\xi(t_0), \ldots, \xi(t_k))$ follows a $(k+1)$-variate extreme value distribution and its distribution function G is given by

$$G(\boldsymbol{x}) = \exp[-\mu([\boldsymbol{0}, \boldsymbol{x}]^C)], \qquad \boldsymbol{x} \in \mathbb{R}^{k+1}, \tag{2.2}$$

where μ is the *exponent measure* of G [26, Proposition 5.8], which is defined on $E = [0, \infty)^{k+1} \setminus \{\boldsymbol{0}\}$, and $[\boldsymbol{0}, \boldsymbol{x}]^C = E \setminus [\boldsymbol{0}, \boldsymbol{x}]$.

The following convergence result provides the theoretical foundation for statistical inference. Indeed, it shows that, by conditioning a process η in the MDA of ξ on an extreme event at one location t_0, the incremental process V can be extracted asymptotically.

Theorem 2.1. *Let $\{\eta(t) : t \in T\}$ be nonnegative and in the MDA of some max-stable process ξ that admits representation (2.1), and suppose that η is normalized such that (1.1) holds with $c_n(t) = 1/n$ and $b_n(t) = 0$ for $n \in \mathbb{N}$ and $t \in T$. Let $a(n) \to \infty$ as $n \to \infty$. For $k \in \mathbb{N}$ and $t_0, \ldots, t_k \in T$, we have the following convergence in distribution on $\mathbb{R}^{k+1}$:*

$$\left(\frac{\eta(t_0)}{a(n)}, \frac{\eta(t_1)}{\eta(t_0)}, \ldots, \frac{\eta(t_k)}{\eta(t_0)} \;\middle|\; \eta(t_0) > a(n) \right) \xrightarrow{\mathrm{D}} (Z, \Delta\tilde{\boldsymbol{V}}) \quad \text{as } n \to \infty. \tag{2.3}$$

Here the distribution of $\Delta\tilde{\boldsymbol{V}}$ is given by

$$\mathbb{P}(\Delta\tilde{\boldsymbol{V}} \in \mathrm{d}\boldsymbol{z}) = \mathbb{P}(\Delta\boldsymbol{V} \in \mathrm{d}\boldsymbol{z})\mathbb{E}(V(t_0) \mid \Delta\boldsymbol{V} = \boldsymbol{z}), \qquad \boldsymbol{z} \geq \boldsymbol{0},$$

where $\Delta\boldsymbol{V}$ denotes the vector of increments $(V(t_1)/V(t_0), \ldots, V(t_k)/V(t_0))$ with respect to t_0, following the convention that $x/0 = \infty$ for all $x \geq 0$, and Z is an independent Pareto variable.

Remark 2.1. The choice of norming functions in Theorem 2.1 is without loss of generality: if ξ has standard Fréchet margins, any process η satisfying the convergence in (1.1) can be normalized such that the norming functions in (1.1) are $c_n(t) = 1/n$ and $b_n(t) = 0$, $n \in \mathbb{N}$, $t \in T$ [26, Proposition 5.10].

Proof of Theorem 2.1. For $\boldsymbol{X} = (\eta(t_0), \ldots, \eta(t_k))$, which is in the MDA of the random vector $\Xi = (\xi(t_0), \ldots, \xi(t_k))$, it follows from [26, Proposition 5.17] that

$$\lim_{m\to\infty} m\mathbb{P}\left(\frac{\boldsymbol{X}}{m} \in B\right) = \mu(B) \tag{2.4}$$

for all elements B of the Borel σ-algebra $\mathcal{B}(E)$ of E bounded away from $\{\boldsymbol{0}\}$ with $\mu(\partial B) = 0$, where μ is defined by (2.2). For $s_0 > 0$ and $\boldsymbol{s} = (s_1, \ldots, s_k) \in [0, \infty)^k$, we consider the sets $A_{s_0} = (s_0, \infty) \times [0, \infty)^k$, $s_0 \geq 1$, and $B_{\boldsymbol{s}} = \{\boldsymbol{x} \in [0, \infty)^{k+1} : (x^{(1)}, \ldots, x^{(k)}) \leq x^{(0)}\boldsymbol{s}\}$ for $\boldsymbol{s}$ satisfying $\mathbb{P}(\Delta\tilde{\boldsymbol{V}} \in \partial[\boldsymbol{0}, \boldsymbol{s}]) = 0$. Since $B_{\boldsymbol{s}}$ satisfies $B_{\boldsymbol{s}} = cB_{\boldsymbol{s}}$ for any $c > 0$, we obtain

$$\begin{aligned}
\mathbb{P}&\left(\eta(t_0) > s_0 a(n), \left(\frac{\eta(t_1)}{\eta(t_0)}, \ldots, \frac{\eta(t_k)}{\eta(t_0)}\right) \leq \boldsymbol{s} \;\middle|\; \eta(t_0) > a(n)\right) \\
&= \frac{a(n)\mathbb{P}(\boldsymbol{X}/a(n) \in B_{\boldsymbol{s}} \cap A_{s_0})}{a(n)\mathbb{P}(\boldsymbol{X}/a(n) \in A_1)} \\
&\to \frac{\mu(B_{\boldsymbol{s}} \cap A_{s_0})}{\mu(A_1)} \quad \text{as } n \to \infty,
\end{aligned} \tag{2.5}$$

where the convergence follows from (2.4), if $\mu(\partial(B_{\boldsymbol{s}} \cap A_{s_0})) = 0$.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

By definition, $\mu(A_1) = -\log \mathbb{P}(\xi(t_0) \le 1) = 1$. At the same time, μ is the intensity measure of the PPP $\sum_{i\in\mathbb{N}} \delta_{(U_i V_i(t_0),\dots,U_i V_i(t_k))}$ on E, and, hence, for $\boldsymbol{s} \in [0,\infty)^k$ with $\mathbb{P}(\Delta\tilde{\boldsymbol{V}} \in \partial[\boldsymbol{0},\boldsymbol{s}]) = 0$,

$$\begin{aligned}
\mu(B_{\boldsymbol{s}} \cap A_{s_0}) &= \int_0^\infty u^{-2} \int_{[s_0 u^{-1},\infty)} \mathbb{P}(V(t_0) \in \mathrm{d}y,\ \Delta\boldsymbol{V} \le \boldsymbol{s})\,\mathrm{d}u \\
&= \int_0^\infty u^{-2} \int_{[\boldsymbol{0},\boldsymbol{s}]} \int_{[s_0 u^{-1},\infty)} \mathbb{P}(V(t_0) \in \mathrm{d}y \mid \Delta\boldsymbol{V} = \boldsymbol{z})\mathbb{P}(\Delta\boldsymbol{V} \in \mathrm{d}\boldsymbol{z})\,\mathrm{d}u \\
&= \int_{[\boldsymbol{0},\boldsymbol{s}]} \int_{[0,\infty)} y s_0^{-1} \mathbb{P}(V(t_0) \in \mathrm{d}y \mid \Delta\boldsymbol{V} = \boldsymbol{z})\mathbb{P}(\Delta\boldsymbol{V} \in \mathrm{d}\boldsymbol{z}) \\
&= s_0^{-1} \int_{[\boldsymbol{0},\boldsymbol{s}]} \mathbb{E}(V(t_0) \mid \Delta\boldsymbol{V} = \boldsymbol{z})\mathbb{P}(\Delta\boldsymbol{V} \in \mathrm{d}\boldsymbol{z}) \\
&= \mathbb{P}(Z > s_0)\mathbb{P}(\Delta\tilde{\boldsymbol{V}} \le \boldsymbol{s}).
\end{aligned} \tag{2.6}$$

Relations (2.6) and (2.5) imply (2.3), which completes the proof.

Remark 2.2. If $V(t_0)$ is stochastically independent of the increments $\Delta\boldsymbol{V}$, we simply have $\mathbb{P}(\Delta\tilde{\boldsymbol{V}} \in \mathrm{d}\boldsymbol{z}) = \mathbb{P}(\Delta\boldsymbol{V} \in \mathrm{d}\boldsymbol{z})$. This is particularly the case if ξ admits a representation (1.2), which shows that (1.5) is indeed a special case of Theorem 2.1.

Remark 2.3. Engelke *et al.* [13] considered Hüsler–Reiss distributions (cf. [18] and [19]) and obtained their limiting results by conditioning on certain extremal events $A \subset E$. They showed that various choices of A are sensible in the Hüsler–Reiss case, leading to different limiting distributions of the increments of η. In the case where ξ is a Brown–Resnick process and $A = (1,\infty) \times [0,\infty)^k$, the assertions of Theorem 2.1 and [13, Theorem 3.3] coincide.

Example 2.1. (*Extremal Gaussian process [29].*) A commonly used class of stationary yet nonergodic max-stable processes on $\mathbb{R}^d$ is defined by

$$\xi(t) = \max_{i\in\mathbb{N}} U_i Y_i(t), \qquad t \in \mathbb{R}^d,$$

where $\sum_{i\in\mathbb{N}} \delta_{U_i}$ is a Fréchet Poisson process, $Y_i(t) = \max(0, \tilde{Y}_i(t))$, and the $\tilde{Y}_i$ are i.i.d., stationary, centered Gaussian processes with $\mathbb{E}(\max(0,\tilde{Y}_i(t))) = 1$ for all $t \in \mathbb{R}^d$ [2], [29]. Note that, in general, a $t_0 \in \mathbb{R}^d$ such that $Y_i(t_0) = 1$ a.s. does not exist, i.e. the process admits representation (2.1) but not representation (1.2).

While the Hüsler–Reiss distribution is already given by the incremental representation (1.2), cf. [19], other distributions can be suitably rewritten, provided that the respective exponent measure μ is known.

Proposition 2.1. *Let $\Xi = (\xi(t_0),\dots,\xi(t_k))$ be a max-stable process on $T = \{t_0,\dots,t_k\}$ with standard Fréchet margins, and suppose that its exponent measure μ is concentrated on $(0,\infty) \times [0,\infty)^k$. Define a random vector $\boldsymbol{W}$ via*

$$\mathbb{P}(\boldsymbol{W} \le \boldsymbol{s}) = \mu(B_{\boldsymbol{s}} \cap A), \qquad \boldsymbol{s} \in [0,\infty)^k, \tag{2.7}$$

where $A = (1,\infty) \times [0,\infty)^k$ and $B_{\boldsymbol{s}} = \{\boldsymbol{x} \in [0,\infty)^{k+1} \colon (x^{(1)},\dots,x^{(k)}) \le x^{(0)}\boldsymbol{s}\}$. Then $\boldsymbol{W}$ is the incremental process of Ξ in (1.2).

Proof. With the measurable transformation

$$\tau\colon (0,\infty) \times [0,\infty)^k \to (0,\infty) \times [0,\infty)^k, \qquad (x_0,\dots,x_k) \mapsto \left(x_0, \frac{x_1}{x_0},\dots,\frac{x_k}{x_0}\right),$$

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

$\tau(B_{\boldsymbol{s}} \cap A) = (1, \infty) \times [\boldsymbol{0}, \boldsymbol{s}]$ and $\mu^\tau(\cdot) = \mu(\tau^{-1}((1, \infty) \times \cdot))$ defines a probability measure on $[0, \infty)^k$ because $\mu^\tau([0, \infty)^k) = \mu(A) = 1$. Since

$$\mu(B_{\boldsymbol{s}} \cap A) = \mu(\tau^{-1}((1, \infty) \times [\boldsymbol{0}, \boldsymbol{s}])) = \mu^\tau([\boldsymbol{0}, \boldsymbol{s}]),$$

the random vector $\boldsymbol{W}$ in (2.7) is well defined and has law μ^τ.

By the definition of the exponent measure, we have $\Xi \stackrel{\mathrm{D}}{=} \max_{i\in\mathbb{N}} \mathbf{X}_i$, where $\Pi = \sum_{i\in\mathbb{N}} \delta_{\mathbf{X}_i}$ is a PPP on E with intensity measure μ. Then, the transformed point process $\tau\Pi = \sum_{i\in\mathbb{N}} \delta_{(X_i^{(0)}, X_i^{(1)}/X_i^{(0)}, \dots, X_i^{(k)}/X_i^{(0)})}$ has intensity measure

$$\tilde{\mu}((c, \infty) \times [\boldsymbol{0}, \boldsymbol{s}]) = \mu(B_{\boldsymbol{s}} \cap ((c, \infty) \times [0, \infty)^k)) = c^{-1}\mu(B_{\boldsymbol{s}} \cap A)$$

for any $c > 0$ and $\boldsymbol{s} \in [0, \infty)^k$, where we use the homogeneity property $c^{-1}\mu(\mathrm{d}\boldsymbol{x}) = \mu(\mathrm{d}(c\boldsymbol{x}))$. Thus, $\tau\Pi$ has the same intensity as $\sum_{i\in\mathbb{N}} \delta_{(U_i, \boldsymbol{W}_i)}$, where $\sum_{i\in\mathbb{N}} \delta_{U_i}$ is a Fréchet Poisson process and $\boldsymbol{W}_i$, $i \in \mathbb{N}$, are i.i.d. vectors with law μ^τ. Hence,

$$\Xi \stackrel{\mathrm{D}}{=} \max_{i\in\mathbb{N}} \tau^{-1}\left(\left(X_i^{(0)}, \frac{X_i^{(1)}}{X_i^{(0)}}, \dots, \frac{X_i^{(k)}}{X_i^{(0)}}\right)\right) \stackrel{\mathrm{D}}{=} \max_{i\in\mathbb{N}} \tau^{-1}((U_i, \boldsymbol{W}_i)) = \max_{i\in\mathbb{N}} U_i \boldsymbol{W}_i,$$

which completes the proof.

Example 2.2. If $\mathbb{P}(V(t_0) = 0) = 0$ in construction (2.1), the exponent measure μ of any finite-dimensional vector $\Xi = (\xi(t_0), \dots, \xi(t_k))$, $t_0, \dots, t_k \in T$, $k \in \mathbb{N}$, satisfies the condition $\mu(\{0\} \times [0, \infty)^k) = 0$. Thus, by Proposition 2.1 and the proof of Theorem 2.1, the incremental representation of Ξ according to (1.2) is given by $\Xi = \max_{i\in\mathbb{N}} U_i \cdot (1, \Delta\tilde{\boldsymbol{V}}_i)^\top$, where $\Delta\tilde{\boldsymbol{V}}_i$, $i \in \mathbb{N}$, are independent copies of $\Delta\tilde{\boldsymbol{V}}$.

Example 2.3. (*Symmetric logistic distribution; cf. [16].*) For $T = \{t_0, \dots, t_k\}$, the symmetric logistic distribution is given by

$$\mathbb{P}(\xi(t_0) \le x_0, \dots, \xi(t_k) \le x_k) = \exp[-(x_0^{-q} + \dots + x_k^{-q})^{1/q}] \tag{2.8}$$

for $x_0, \dots, x_k > 0$ and $q > 1$. Hence, the density of the exponent measure equals

$$\mu(\mathrm{d}x_0, \dots, \mathrm{d}x_k) = \left(\sum_{i=0}^{k} x_i^{-q}\right)^{1/q-k-1} \left(\prod_{i=1}^{k} (iq - 1)\right) \prod_{i=0}^{k} x_i^{-q-1}\, \mathrm{d}x_0 \cdots \mathrm{d}x_k.$$

Applying Proposition 2.1, the incremental process W in (1.2) is given by

$$\mathbb{P}(W(t_1) \le s_1, \dots, W(t_k) \le s_k) = \left(1 + \sum_{i=1}^{k} s_i^{-q}\right)^{1/q-1}.$$

While Theorem 2.1 relates to convergence in the sense of finite-dimensional distributions, an analogous result can be formulated for weak convergence on suitable function spaces. In particular, we consider processes with continuous sample paths. For a Borel set $U \subset \mathbb{R}^d$, we denote by $C(U)$ and $C^+(U)$ the space of nonnegative and strictly positive continuous functions on U, respectively, equipped with the topology of uniform convergence on compact sets.

Theorem 2.2. *Let $\{\eta(t) \colon t \in \mathbb{R}^d\}$ be a $C^+(\mathbb{R}^d)$-valued process in the MDA of a max-stable process $\{\xi(t) \colon t \in \mathbb{R}^d\}$ as in (1.2) in the sense of weak convergence on $C(\mathbb{R}^d)$. Without loss of generality, assume that $n^{-1}\max_{i=1}^n \eta_i(\cdot) \xrightarrow{\mathrm{D}} \xi(\cdot)$ as $n \to \infty$. Let W be the incremental* *process from (1.2), and let Z be a Pareto random variable, independent of W. Then, for any sequence $a(n) \to \infty$ as $n \to \infty$, we have the following weak convergence on $(0, \infty) \times C(\mathbb{R}^d)$:*

$$\left(\frac{\eta(t_0)}{a(n)}, \frac{\eta(\cdot)}{\eta(t_0)} \,\middle|\, \eta(t_0) > a(n)\right) \xrightarrow{\mathrm{D}} (Z, W(\cdot)).$$

Proof. Analogously to [33, Theorem 5], weak convergence of a sequence of probability measures P_n, $n \in \mathbb{N}$, to some probability measure P on $C(\mathbb{R}^d)$ is equivalent to weak convergence of $P_n r_j^{-1}$ to $P r_j^{-1}$ on $C([-j, j]^d)$ for all $j \geq 1$, where $r_j : C(\mathbb{R}^d) \to C([-j, j]^d)$ denotes the restriction to the cube $[-j, j]^d$. Hence, it suffices to consider the processes η and ξ restricted to a compact set $K \subset \mathbb{R}^d$ and to show weak convergence on $(0, \infty) \times C(K)$.

As the process ξ is max-stable and $\eta \in \mathrm{MDA}(\xi)$, similarly to the case of multivariate max-stable distributions (cf. Theorem 2.1),

$$\lim_{u \to \infty} u \mathbb{P}\left(\frac{\eta}{u} \in B\right) = \mu(B) \tag{2.9}$$

for any Borel set $B \subset C(K)$ bounded away from 0^K, i.e. $\inf\{\sup_{s \in K} f(s) \colon f \in B\} > 0$, and with $\mu(\partial B) = 0$ [8, Corollary 9.3.2], where μ is the *exponent measure* of ξ, defined by

$$\begin{aligned}
&\mathbb{P}(\xi(s) \leq x_j,\ s \in K_j,\ j = 1, \dots, m) \\
&\qquad = \exp\left[-\mu\left(\left\{f \in C(K) \colon \sup_{s \in K_j} f(s) > x_j \text{ for some } j \in \{1, \dots, m\}\right\}\right)\right]
\end{aligned}$$

for $x_j \geq 0$ and $K_j \subset K$ compact. Thus, μ equals the intensity measure of the PPP $\sum_{i \in \mathbb{N}} \delta_{U_i W_i(\cdot)}$. For $z > 0$ and a Borel set $D \subset C(K)$, we consider the sets

$$A_z = \{f \in C(K) \colon f(t_0) > z\}, \qquad B_D = \left\{f \in C(K) \colon \frac{f(\cdot)}{f(t_0)} \in D\right\},$$

and $A = A_1$. Again, $cB_D = B_D$ for any $c > 0$. Then, as $W(t_0) = 1$ a.s., we have $\mu(A_z) = \int_z^\infty u^{-2}\,\mathrm{d}u = z^{-1}$. For $s_0 \geq 1$ and any Borel set $D \subset C(K)$ with $\mathbb{P}(W \in \partial D) = 0$, by (2.9), analogously to (2.5), we obtain

$$\begin{aligned}
\mathbb{P}\left\{\frac{\eta(t_0)}{a(n)} > s_0, \frac{\eta(\cdot)}{\eta(t_0)} \in D \,\middle|\, \eta(t_0) > a(n)\right\} &\xrightarrow{n \to \infty} \frac{\mu(B_D \cap A_{s_0})}{\mu(A)} \\
&= \int_{s_0}^\infty u^{-2} \mathbb{P}\{uW(\cdot) \in B_D\}\,\mathrm{d}u \\
&= s_0^{-1} \mathbb{P}\{W(\cdot) \in D\},
\end{aligned}$$

which is the joint distribution of Z and $W(\cdot)$. This completes the proof.

3. Mixed moving maxima representation

Let

$$\Pi_0 = \sum_{i \in \mathbb{N}} \delta_{(U_i, T_i, F_i)}$$

be the PPP on $(0, \infty) \times \mathbb{R}^d \times C(\mathbb{R}^d)$ with intensity $u^{-2}\,\mathrm{d}u\,\mathrm{d}t\,\mathbb{P}_F(\mathrm{d}f)$ corresponding to the M3 process (1.4). In the sequel, we denote M3 processes by

$$M(t) = \max_{i \in \mathbb{N}} U_i F_i(t - T_i), \qquad t \in \mathbb{R}^d. \tag{3.1}$$

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

For $t_0, \dots, t_k \in \mathbb{R}^d$, $s_0, \dots, s_k \geq 0$, and $k \in \mathbb{N}$, the marginal distributions of M are

$$\mathbb{P}(M(t_l) \leq s_l,\ 0 \leq l \leq k) = \exp\biggl[-\int\int_{\mathbb{R}^d} \max_{l=0}^{k} \frac{f(t_l - t)}{s_l}\,\mathrm{d}t\,\mathbb{P}_F(\mathrm{d}f)\biggr]. \tag{3.2}$$

In Section 2 we recovered the incremental process W from processes in the MDA of a max-stable process with representation (1.2) or (2.1). In the case of M3 processes, although an incremental representation might still exist, the shape function F from (3.1) usually has a much simpler structure than the corresponding incremental process W (see also Section 4). In all practically relevant M3 models, the shape function F has its unique maximum at the origin (e.g. the Gaussian and t-extreme value process of Smith [31], M3 models in [9], and the M3 representation of Brown–Resnick processes [12], [20]). Hence, to reconstruct the distribution of F from realizations of η, we consider η in the neighborhood of its own (local) maximum. If the latter is large enough, a single shape function becomes visible in a certain surrounding area. Thus, conditioning on the maximum provides an appropriate framework for statistical inference; for further discussion and examples, we refer the reader to Section 5.2. A related approach of reconstructing the parameter functions in a time-series M3 model is considered in [22].

Let $\{\eta(t) : t \in \mathbb{R}^d\}$ be strictly positive and in the MDA of an M3 process M in the sense of weak convergence in $C(\mathbb{R}^d)$. Let η be normalized such that the norming functions in (1.1) are $c_n(t) = 1/n$ and $b_n(t) = 0$, $n \in \mathbb{N}$, $t \in \mathbb{R}^d$. Suppose further that the shape function F of M satisfies

$$F(\mathbf{0}) = \lambda \quad \text{a.s.}, \qquad F(t) \in [0, \lambda) \quad \text{for all } t \in \mathbb{R}^d \setminus \{\mathbf{0}\} \text{ a.s.}$$

for some $\lambda > 0$. Note that F is sample continuous as M is (cf. [27]). Under these assumptions, there is an analogous result to Theorem 2.2.

Theorem 3.1. *Let $Q, K \subset \mathbb{R}^d$ be compact with ∂Q a Lebesgue null set, and let*

$$\tau_Q\colon\ C(Q) \to \mathbb{R}^d, \qquad f \mapsto \inf\Bigl(\arg\max_{t\in Q} f(t)\Bigr),$$

where 'inf' *is understood in the lexicographic sense. Then, under the above assumptions, for any Borel set $B \subset C(K)$ with $\mathbb{P}(F/\lambda \in \partial B) = 0$, and any sequence $a(n)$ with $a(n) \to \infty$ as $n \to \infty$, we have*

$$\begin{aligned}\lim_{\substack{\{\mathbf{0}\}\in L \nearrow \mathbb{R}^d\\ \text{compact}}} \limsup_{n\to\infty} \mathbb{P}\biggl(\frac{\eta(\tau_Q(\eta|_Q) + \cdot)}{\eta(\tau_Q(\eta|_Q))} \in B \biggm| \max_{t\in Q} \eta(t) = \max_{t\in Q\oplus L} \eta(t),\ \max_{t\in Q} \eta(t) \geq a(n)\biggr) \\ = \mathbb{P}\biggl(\frac{F(\cdot)}{\lambda} \in B\biggr),\end{aligned}$$

where '$\oplus$' *denotes morphological dilation.*

The same result holds if we replace $\limsup_{n\to\infty}$ *by* $\liminf_{n\to\infty}$.

Remark 3.1. In order to understand Theorem 3.1 intuitively, we assume that η is the M3 process M itself and that the shape function F has a compact support $K \subset \mathbb{R}^d$. Then, for $L \supset K$, the statement $\max_{t\in Q\oplus L} \eta(t) = \max_{t\in Q} \eta(t)$ implies that $\max_{t\in Q} \eta(t)$ is given by $U_{i_0} F_{i_0}(0) = U_{i_0}\lambda$ for some $i_0 \in \mathbb{N}$. More precisely, we have $(U_{i_0}, T_{i_0}, F_{i_0}) = (\max_{t\in Q} \eta(t)/\lambda$, $\arg\max_{t\in Q} \eta(t), F_{i_0}(t))$. If $\max_{t\in Q} \eta(t)$ is large, and, hence, so is U_{i_0}, it is very likely that $\eta(\cdot)$ coincides with $U_{i_0} F_{i_0}(\cdot - T_{i_0})$ in a neighborhood of T_{i_0}. This allows us to reconstruct F_{i_0} directly from the observed sample path.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

Proof of Theorem 3.1. First, we consider a fixed compact set $L \subset \mathbb{R}^d$ large enough such that $K \cup \{\mathbf{0}\} \subset L$ and define

$$A_L = \Big\{ f \in C(Q \oplus L) \colon \max_{t \in Q} f(t) \geq 1,\ \max_{t \in Q} f(t) = \max_{t \in Q \oplus L} f(t) \Big\},$$
$$C_B = \Big\{ f \in C(Q \oplus L) \colon \frac{f(\tau_Q(f|_Q) + \cdot)}{f(\tau_Q(f|_Q))} \in B \Big\}$$

for any Borel set $B \subset C(K)$. Note that C_B is invariant with respect to multiplication by any positive constant. Thus, we obtain

$$\begin{aligned} \mathbb{P}\Big\{ \frac{\eta(\tau_Q(\eta|_Q) + \cdot)}{\eta(\tau_Q(\eta|_Q))} \in B \,\Big|\, \max_{t \in Q} \eta(t) = \max_{t \in Q \oplus L} \eta(t) \geq a(n) \Big\} \\ = \mathbb{P}\Big\{ \frac{\eta}{a(n)} \in C_B \,\Big|\, \frac{\eta}{a(n)} \in A_L \Big\} \\ = \frac{a(n)\mathbb{P}\{\eta/a(n) \in C_B,\ \eta/a(n) \in A_L\}}{a(n)\mathbb{P}\{\eta/a(n) \in A_L\}}. \end{aligned} \tag{3.3}$$

By [8, Corollary 9.3.2] and [26, Proposition 3.12] we have

$$\mu(D^o) \leq \liminf_{u \to \infty} u\mathbb{P}\Big(\frac{\eta}{u} \in D^o\Big) \leq \limsup_{u \to \infty} u\mathbb{P}\Big(\frac{\eta}{u} \in \bar{D}\Big) \leq \mu(\bar{D}), \qquad D \in C(Q \oplus L),$$

where D is bounded away from $0^{Q \oplus L}$. Here μ is the intensity measure of the PPP $\sum_{i \in \mathbb{N}} \delta_{U_i F_i(\cdot - T_i)}$ restricted to $C(Q \oplus L)$. Thus, by adding or removing the boundary, we see that all the limit points of (3.3), as $n \to \infty$, lie in the interval

$$\Big[\frac{\mu(C_B \cap A_L) - \mu(\partial(C_B \cap A_L))}{\mu(A_L) + \mu(\partial A_L)}, \frac{\mu(C_B \cap A_L) + \mu(\partial(C_B \cap A_L))}{\mu(A_L) - \mu(\partial A_L)} \Big]. \tag{3.4}$$

We note that A_L is closed and the set

$$\begin{aligned} A_L^* = \Big\{ f \in C(Q \oplus L) \colon \tau_Q(f|_Q) \in Q^o, \\ \max_{t \in Q} f(t) > \max\{1, f(t)\} \text{ for all } t \in Q \oplus L \setminus \{\tau_Q(f|_Q)\} \Big\} \end{aligned}$$

is in the interior of A_L. Hence, with $(Q \oplus L)^c = \mathbb{R}^d \setminus (Q \oplus L)$, we can assess

$$\begin{aligned} \mu(\partial A_L) &\leq \mu\Big(\Big\{ f \in C(Q \oplus L) \colon \max_{t \in Q} f(t) = 1 \Big\}\Big) \\ &\quad + \mu\Big(\Big(\{ f \in C(Q \oplus L) \colon \tau_Q(f|_Q) \in \partial Q\} \\ &\qquad \cup \Big\{ f \in C(Q \oplus L) \colon \arg\max_{t \in Q \oplus L} f(t) \text{ is not unique} \Big\}\Big) \\ &\qquad \cap \Big\{ f \in C(Q \oplus L) \colon \max_{t \in Q} f(t) = \max_{t \in Q \oplus L} f(t) \geq 1 \Big\}\Big) \\ &\leq 0 + \int_{\partial Q} \int_{\lambda^{-1}}^{\infty} u^{-2}\, \mathrm{d}u\, \mathrm{d}t_0 + \int_{(Q \oplus L)^c} \int_{\lambda^{-1}}^{\infty} u^{-2} \mathbb{P}\Big(u \max_{t_0 \in Q} F(t_0 - x) \geq 1\Big)\, \mathrm{d}u\, \mathrm{d}x. \end{aligned} \tag{3.5}$$

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

Here $\mu(\{f \in C(Q \oplus L)\colon \max_{t\in Q} f(t) = 1\}) = 0$ holds as $\max_{t\in Q} M(t)$ is Fréchet distributed (cf. [9, Lemma 9.3.4]). Since ∂Q is a Lebesgue null set, the second term on the right-hand side of (3.5) also vanishes. Thus,

$$\mu(\partial A_L) \leq \int_{(Q\oplus L)^c} \int_{\lambda^{-1}}^{\infty} u^{-2}\mathbb{P}\Big(u \max_{t_0\in Q} F(t_0 - x) \geq 1\Big)\, \mathrm{d}u\, \mathrm{d}x =: c(L). \tag{3.6}$$

Now, let $B \subset C(K)$ be a Borel set such that $\mathbb{P}(F/\lambda \in \partial B) = 0$. The set

$$C_B^* = \left\{ f \in C(Q \oplus L)\colon \ \arg\max_{f\in Q} f(t) \text{ is unique}, \ \frac{f(\tau_Q(f|_Q) + \cdot)}{f(\tau_Q(f|_Q))} \in B^o \right\}$$

is in the interior of C_B and the closure of C_B is a subset of

$$\begin{aligned} C_B^* &\cup \left\{ f \in C(Q \oplus L)\colon \ \arg\max_{t\in Q} f(t) \text{ is not unique} \right\} \\ &\cup \left\{ f \in C(Q \oplus L)\colon \ \frac{f(\tau_Q(f|_Q) + \cdot)}{f(\tau_Q(f|_Q))} \in \partial B \right\}. \end{aligned}$$

Thus, by (3.6), we can assess

$$\begin{aligned} \mu(\partial(C_B \cap A_L)) &\leq \mu(\partial A_L) + \mu(\partial C_B \cap A_L) \\ &\leq c(L) + \int_{\mathbb{R}^d \setminus (Q\oplus L)} \int_{\lambda^{-1}}^{\infty} u^{-2}\mathbb{P}\Big(u \max_{t_0\in Q} F(t_0 - x) \geq 1\Big)\, \mathrm{d}u\, \mathrm{d}x \\ &\quad + \int_Q \int_{\lambda^{-1}}^{\infty} u^{-2}\mathbb{P}\Big(\frac{F}{\lambda} \in \partial B\Big)\, \mathrm{d}u\, \mathrm{d}t \\ &= \ 2c(L). \end{aligned} \tag{3.7}$$

Furthermore, we obtain

$$\begin{aligned} &\mu(C_B \cap A_L) \\ &\quad = \int_Q \int_{\lambda^{-1}}^{\infty} u^{-2}\mathbb{P}\Big(\frac{F(\cdot)}{\lambda} \in B\Big)\, \mathrm{d}u\, \mathrm{d}t_0 \\ &\quad\quad + \int_{(Q\oplus L)^c} \int_{\lambda^{-1}}^{\infty} u^{-2}\mathbb{P}\Big(u \max_{t_0\in Q} F(t_0 - x) \geq 1, \ \frac{F((\tau_Q(F(\cdot - x)|_Q)) + \cdot - x)}{\max_{t_0\in Q} F(t_0 - x)} \in B, \\ &\quad\quad\quad \frac{F(t - x)}{\max_{t_0\in Q} F(t_0 - x)} \leq 1 \text{ for all } t \in Q \oplus L\Big)\, \mathrm{d}u\, \mathrm{d}x. \end{aligned} \tag{3.8}$$

The second term in (3.8) is positive and can be bounded from above by $c(L)$. Setting $B = C(K)$, $\mu(A_L)$ can be expressed in an analogous way. Now, we substitute in the results of (3.6), (3.7), and (3.8) into (3.4) to reveal that all the limit points of (3.3) are in the interval

$$\left[\frac{\lambda|Q|\mathbb{P}(F(\cdot)/\lambda \in B) - 2c(L)}{\lambda|Q| + 2c(L)}, \frac{\lambda|Q|\mathbb{P}(F(\cdot)/\lambda \in B) + 3c(L)}{\lambda|Q| - c(L)} \right].$$

Finally, we note that $c(L) \leq \int_{(Q\oplus L)^c} \mathbb{E}(\max_{t_0\in Q} F(t_0 - x))\, \mathrm{d}x$, which vanishes for $L \nearrow \mathbb{R}^d$ as the integral $\int_{\mathbb{R}^d} \mathbb{E}(\max_{t_0\in Q} F(t_0 - x))\, \mathrm{d}x$ is finite because of the sample continuity of M (cf. [27]). This completes the proof.

Remark 3.2. Whilst Theorem 3.1 requires full sample paths, for applications, it might be more natural to consider a discrete grid as index space. If both the shape functions and the processes η and M are restricted to $\mathbb{Z}^d$, an analogous result to Theorem 3.1 holds, where $Q = \{t_0\} \subset \mathbb{Z}^d$

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

and the process is evaluated at a set of k points $t_1, \dots, t_k \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$, i.e.

$$\lim_{\substack{\{\mathbf{0}\}\in L\nearrow\mathbb{Z}^d \\ \text{compact}}} \limsup_{n\to\infty} \mathbb{P}\bigg(\frac{\eta(t_0+t_i)}{\eta(t_0)} \in B_i,\ 1\le i\le k \,\bigg|\, \eta(t_0) = \max_{t\in L}\eta(t_0+t) \ge a(n)\bigg)$$
$$= \mathbb{P}\bigg(\frac{F(t_i)}{\lambda} \in B_i,\ 1\le i\le k\bigg)$$

for any Borel sets $B_1, \dots, B_k \subset [0,\infty)$ such that

$$\mathbb{P}\bigg(\bigg(\frac{F(t_1)}{\lambda}, \dots, \frac{F(t_k)}{\lambda}\bigg) \in \partial(B_1\times\cdots\times B_k)\bigg) = 0.$$

Example 3.1. Let $\{F(t) : t \in \mathbb{R}^d\}$ be a random shape function as in (1.3). For $c, \varepsilon, \kappa > 0$, let

$$\Pi_{c,\varepsilon} = \sum_{i\in\mathbb{N}} \delta_{(U_i,T_i,F_i)}$$

be a PPP on $(0,\infty)\times\mathbb{R}^d\times C(\mathbb{R}^d)$ with intensity $c\mathbf{1}_{\{u\ge\varepsilon\}}u^{-2}\,\mathrm{d}u\,\mathrm{d}t\mathbb{P}_F(\mathrm{d}f)$. Then, by tightness considerations, it can be shown that the process $\tilde{M} = \tilde{M}_{c,\varepsilon,\kappa}$, defined by

$$\tilde{M}(\cdot) = \kappa \vee \max_{(u,t,f)\in\Pi_{c,\varepsilon}} uf(\cdot - t),$$

is in the MDA of M in the sense of weak convergence on $C(\mathbb{R}^d)$ and, thus, satisfies the assumptions of Theorem 3.1.

The important feature of $\tilde{M}$ is that, in contrast to the M3 process M itself, it is constructed from a (locally) *finite* number of shape functions. Thus, it can be perceived as (a composition of) observable environmental events, as, for instance, in Smith's [31] 'rainfall-storm' interpretation. This example therefore has the potential for various practical applications.

4. Switching between the different representations

In the previous sections we analyzed processes that admit an incremental representation (1.2) or (2.1) and processes of M3 type as in (1.4). We will now show that, under certain assumptions, we can switch from one representation to the other. For simulation of and statistical inference for max-stable processes, these conversion formulae allow choosing the more appropriate representation.

4.1. Incremental representation of mixed moving maxima processes

We distinguish between M3 processes with strictly positive shape functions, for which an incremental representation (1.2) exists, and general nonnegative shape functions, for which only the weaker representation (2.1) can be obtained.

4.1.1. *Mixed moving maxima processes with positive shape functions.*

Theorem 4.1. *Let M be an M3 process on $\mathbb{R}^d$ as in (1.4) with a continuous shape function F with $F(t) > 0$ for all $t \in \mathbb{R}^d$. Then M admits a representation (1.2) with $t_0 = 0$ and incremental process W given by*

$$\mathbb{P}(W\in L) = \int_{C^+(\mathbb{R}^d)}\int_{\mathbb{R}^d} \mathbf{1}_{\{f(\cdot-t)/f(-t)\in L\}} f(-t)\,\mathrm{d}t\mathbb{P}_F(\mathrm{d}f), \qquad L\in\mathcal{B}(C^+(\mathbb{R}^d)), \quad (4.1)$$

which defines a proper probability distribution because of condition (1.3).

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

Proof. We consider the two PPPs on $(0, \infty) \times C^+(\mathbb{R}^d)$,

$$\Pi_1 = \sum_{i \in \mathbb{N}} \delta_{(U_i F_i(-T_i), F_i(\cdot - T_i)/F_i(-T_i))},$$

as a transformation of Π_0 from Section 3, and $\Pi_2 = \sum_{i\in\mathbb{N}} \delta_{(U_i', W_i(\cdot))}$, with W_i, $i \in \mathbb{N}$, being independent copies of W, and with $\sum_{i\in\mathbb{N}} \delta_{U_i'}$ being a Fréchet Poisson process (as defined at the beginning of Section 2). Then the intensity measures of Π_1 and Π_2 satisfy

$$\begin{aligned}
\mathbb{E}\Pi_1([z,\infty)\times L) &= \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \int_0^\infty u^{-2} \mathbf{1}_{\{uf(-t)\geq z\}} \mathbf{1}_{\{f(\cdot - t)/f(-t) \in L\}} \,\mathrm{d}u \,\mathrm{d}t \mathbb{P}_F(\mathrm{d}f) \\
&= z^{-1} \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{\{f(\cdot - t)/f(-t)\in L\}} f(-t) \,\mathrm{d}t \mathbb{P}_F(\mathrm{d}f) \\
&= z^{-1}\mathbb{P}(W \in L) \\
&= \mathbb{E}\Pi_2([z,\infty)\times L)
\end{aligned}$$

for $L \in \mathcal{B}(C^+(\mathbb{R}^d))$ and $z > 0$, and, hence, $\Pi_1 \stackrel{\mathrm{D}}{=} \Pi_2$. The assertion follows from the fact that M is uniquely determined by Π_1 via the relation $M(\cdot) = \max_{(v,g)\in\Pi_1} vg(\cdot)$.

While the definition of W in (4.1) is rather implicit, in the following, we provide an explicit construction of the incremental process W, which can also be used for simulation. To this end, let the distribution of $(S, G) \in \mathbb{R}^d \times C^+(\mathbb{R}^d)$ be given by

$$\begin{aligned}
\mathbb{P}((S,G) \in (B \times L)) &= \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{\{s\in B\}} \mathbf{1}_{\{f \in L\}} \frac{f(-s)}{\int f(r)\,\mathrm{d}r} \,\mathrm{d}s \left(\int f(r)\,\mathrm{d}r \right) \mathbb{P}_F(\mathrm{d}f) \\
&= \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{\{s\in B\}} \mathbf{1}_{\{f\in L\}} f(-s) \,\mathrm{d}s \mathbb{P}_F(\mathrm{d}f) \qquad (4.2)
\end{aligned}$$

for $B \in \mathcal{B}^d$ and $L \in \mathcal{B}(C^+(\mathbb{R}^d))$. In other words, $\mathbb{P}_G(\mathrm{d}f) = (\int f(r)\,\mathrm{d}r)\mathbb{P}_F(\mathrm{d}f)$ and, conditional on $\{G = f\}$, the density function of the shift S is proportional to $f(-\cdot)$. Putting $W(\cdot) = G(\cdot - S)/G(-S)$, (4.1) is satisfied.

Example 4.1. (*M3 representation of Brown–Resnick processes; cf. [20].*) We consider the following two special cases of M3 processes.

1. Let $\Sigma \in \mathbb{R}^{d\times d}$ be a positive definite matrix, and let the shape function be given by $F(t) = (2\pi)^{-d/2}|\Sigma|^{-1/2} \exp[-\frac{1}{2}t^\top \Sigma^{-1} t]$, $t \in \mathbb{R}^d$. Then M becomes the well-known Smith process. At the same time, by (4.2), we have $S \sim N(0, \Sigma)$ and $G \equiv F$. Thus,

   $$Y(t) = \exp\big[-\tfrac{1}{2}(t-S)^\top\Sigma^{-1}(t-S) + \tfrac{1}{2}S^\top\Sigma^{-1}S\big] = \exp\big[-\tfrac{1}{2}t^\top\Sigma^{-1}t + t^\top\Sigma^{-1}S\big].$$

   Since $\mathbb{E}(t^\top\Sigma^{-1}S)^2 = t^\top\Sigma^{-1}t$, M is equivalent to the Brown–Resnick process in (5.1) with variogram $\gamma(h) = h^\top\Sigma^{-1}h$.

2. For the one-dimensional Brown–Resnick process ξ in (5.1) with variogram $\gamma(h) = |h|$, i.e. Y is the exponential of a standard Brownian motion with drift $-|t|/2$, Engelke *et al.* [12] recently showed that the M3 representation is given by $\{F(t) \colon t \in \mathbb{R}\} = \{Y(t) \mid Y(s) \leq 1 \text{ for all } s \in \mathbb{R} \colon t \in \mathbb{R}\}$, i.e. the shape function is the exponential of a conditionally negative drifted Brownian motion. Given these two representations, it follows that the law of F, reweighted by $\int F(t)\,\mathrm{d}t$ and randomly shifted with density $F(-\cdot)/\int F(t)\,\mathrm{d}t$, coincides with the law of Y.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

4.1.2. *Mixed moving maxima processes with finitely supported shape functions.* Let M be an M3 process on $\mathbb{R}^d$ as in (1.4). In contrast to Section 4.1.1, where the shape functions are required to take positive values, here we allow for arbitrary shape functions with values in $[0, \infty)$. The main difference is that, now, the density of the shift variable (in the following referred to as R) cannot be given by the shape function itself any more.

Theorem 4.2. *The M3 process M as in (1.4) allows for an incremental representation of the form (2.1), with incremental processes V_i, given by*

$$V_i(\cdot) = \frac{F_i(\cdot - R_i)}{g(R_i)}.$$

Here R_i, $i \in \mathbb{N}$, are i.i.d. copies of a random vector R with arbitrary density g satisfying $g(t) > 0$ for all $t \in \mathbb{R}^d$, and F_i, $i \in \mathbb{N}$, are i.i.d. copies of the random shape function F.

Proof. With $\sum_{i\in\mathbb{N}} \delta_{U_i}$ a Fréchet Poisson process, we consider the process

$$\tilde{M}(t) = \max_{i\in\mathbb{N}} \frac{U_i F_i(t - R_i)}{g(R_i)}, \qquad t \in \mathbb{R}^d,$$

which is clearly of the form (2.1). Then,

$$\mathbb{P}(\tilde{M}(t_l) \le s_l,\ 0 \le l \le k) = \exp\biggl(-\int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} \max_{l=0}^{k} \frac{f(t_l - t)}{g(t)s_l} g(t)\, \mathrm{d}t \mathbb{P}_F(\mathrm{d}f)\biggr).$$

The right-hand side coincides with the marginal distribution of M, which is given by (3.2). This concludes the proof.

For $t_0 = 0$, the conditional increments $\Delta\tilde{\boldsymbol{V}}$ in Theorem 2.1 are given by

$$\begin{aligned}\mathbb{P}(\Delta\tilde{\boldsymbol{V}} \in \mathrm{d}\boldsymbol{z}) &= \int_0^\infty y\mathbb{P}(V(0) \in \mathrm{d}y,\ \Delta\boldsymbol{V} \in \mathrm{d}\boldsymbol{z}) \\ &= \int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} f(t)\mathbf{1}_{\{(f(t_l+t)/f(t))_{l=1}^k \in \mathrm{d}\boldsymbol{z}\}}\, \mathrm{d}t\mathbb{P}_F(\mathrm{d}f).\end{aligned}$$

The asymptotic conditional increments of $\eta \in \mathrm{MDA}(M)$ can hence be seen as a convolution of the shape function's increments with a random shift, whose density is given by the shape function itself. The distribution is, in particular, independent of the choice of the density g in Theorem 4.2.

Remark 4.1. In Section 4.1.1 we consider the subclass of M3 processes with strictly positive shape functions and provided an incremental representation as in (1.2), which is nicely related to the conditional increments of η due to the property $W(0) = 1$. Section 4.1.2 applies to arbitrary M3 processes, i.e. has less restrictive assumptions, but it only yields an incremental representation as in (2.1), for which the incremental process V does not directly represent the conditional increments of η.

4.2. Mixed moving maxima representation of the incremental construction

In this section we provide an explicit formula to derive the shape function of an M3 process if only the incremental process in representation (2.1) is given. For instance, simulation methods based on the M3 construction are usually much faster than those based on the incremental representation (cf. [12] and [29]).

The following lemma has been proved in [20] and it states sufficient conditions for the existence of an M3 representation.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

Lemma 4.1. (cf. Theorem 14 of [20].) *Let $\sum_{i\in\mathbb{N}} \delta_{U_i}$ be a Fréchet Poisson process, and let W_i, $i \in \mathbb{N}$, be i.i.d. copies of a nonnegative, sample-continuous process $\{W(t),\ t \in \mathbb{R}^d\}$, satisfying $\lim_{\|t\|\to\infty} W(t) = 0$ a.s., $\mathbb{E}W(t) = 1$ for all $t \in \mathbb{R}^d$, and $\mathbb{E}(\max_{t\in K} W(t)) < \infty$ for any compact set $K \subset \mathbb{R}^d$. Furthermore, let W be such that the process $\xi(\cdot) = \max_{i\in\mathbb{N}} U_i W_i(\cdot)$ is stationary. Then the random variables $\tau_i = \inf\{\arg\sup_{t\in\mathbb{R}^d} W_i(t)\}$ and $\gamma_i = \sup_{t\in\mathbb{R}^d} W_i(t)$ are well defined. Furthermore, $\sum_{i\in\mathbb{N}} \delta_{(U_i\gamma_i,\tau_i,W_i(\cdot+\tau_i)/\gamma_i)}$ is a PPP on $(0,\infty)\times\mathbb{R}^d\times C(\mathbb{R}^d)$ with intensity measure $\Psi(\mathrm{d}u,\ \mathrm{d}t,\ \mathrm{d}f) = cu^{-2}\,\mathrm{d}u\,\mathrm{d}t\mathbb{P}_{\tilde{F}}(\mathrm{d}f)$ for some $c > 0$ and some probability measure $\mathbb{P}_{\tilde{F}}$.*

Theorem 4.3. *Under the assumptions of Lemma 4.1, ξ has an M3 representation with $\mathbb{P}_F(\mathrm{d}f) = \mathbb{P}_{\tilde{F}}(c\,\mathrm{d}f)$ being the probability measure of the shape function F. The constant $c > 0$ is given by $c^{-1} = \int_{\mathbb{R}^d}\int_{C(\mathbb{R}^d)} f(t)\mathbb{P}_{\tilde{F}}(\mathrm{d}f)\,\mathrm{d}t$ and the probability measure $\mathbb{P}_{\tilde{F}}$ is defined by*

$$\mathbb{P}_{\tilde{F}}(A) = \frac{\int_0^\infty y\mathbb{P}(W(\cdot+\tau)/y \in A,\ \tau\in K \mid \gamma = y)\mathbb{P}_\gamma(\mathrm{d}y)}{\int_0^\infty y\mathbb{P}(\tau\in K \mid \gamma = y)\mathbb{P}_\gamma(\mathrm{d}y)}, \qquad A \in \mathcal{B}(C(\mathbb{R}^d)),$$

for any compact set $K \subset \mathbb{R}^d$. Here $\mathbb{P}_\gamma$ is the probability measure belonging to γ, and τ and γ are defined as τ_i and γ_i, respectively, with W_i being replaced by W.

Proof. From Lemma 4.1, it follows that the PPP

$$\Phi_0 = \sum_{i\in\mathbb{N}} \delta_{(U_i\gamma_i/c,\tau_i,c\cdot W_i(\cdot+\tau_i)/\gamma_i)}$$

has intensity $\Psi_0(\mathrm{d}u,\ \mathrm{d}t,\ \mathrm{d}f) = u^{-2}\,\mathrm{d}u \times \mathrm{d}t \times \mathbb{P}_F(\mathrm{d}f)$, where $\mathbb{P}_F(\mathrm{d}f) = \mathbb{P}_{\tilde{F}}(c\,\mathrm{d}f)$. Hence, Φ_0 is of the same type as Π_0 from the beginning of Section 3 and

$$\xi(t) = \max_{(y,s,f)\in\Phi_0} yf(t-s), \qquad t\in\mathbb{R}^d,$$

is an M3 representation. The integrability condition (1.3) follows from the fact that ξ has standard Fréchet marginals. Thus, $\int_{\mathbb{R}^d}\int_{C(\mathbb{R}^d)} cf(t)\mathbb{P}_{\tilde{F}}(\mathrm{d}f)\,\mathrm{d}t = 1$. In order to calculate $\mathbb{P}_{\tilde{F}}$, let $A \in \mathcal{B}(C(\mathbb{R}^d))$ and $K \in \mathcal{B}^d$ be compact. Lemma 4.1 implies that $\Psi([1,\infty)\times K\times A) = c|K|\mathbb{P}_{\tilde{F}}(A)$. Thus,

$$\mathbb{P}_{\tilde{F}}(A) = \frac{\Psi([1,\infty)\times K\times A)}{\Psi([1,\infty)\times K\times C(\mathbb{R}^d))}, \tag{4.3}$$

and both the numerator and denominator are finite. The numerator equals

$$\begin{aligned}\int_0^\infty u^{-2}\int_{u^{-1}}^\infty \mathbb{P}\bigg(\frac{W(\cdot+\tau)}{\gamma}\in A,\ \tau\in K\ \bigg|\ \gamma = y\bigg)\mathbb{P}_\gamma(\mathrm{d}y)\,\mathrm{d}u \\ = \int_0^\infty y\mathbb{P}\bigg(\frac{W(\cdot+\tau)}{y}\in A,\ \tau\in K\ \bigg|\ \gamma = y\bigg)\mathbb{P}_\gamma(\mathrm{d}y).\end{aligned}$$

Thus, by (4.3),

$$\mathbb{P}_{\tilde{F}}(A) = \frac{\int_0^\infty y\mathbb{P}(W(\cdot+\tau)/y \in A,\ \tau\in K \mid \gamma = y)\mathbb{P}_\gamma(\mathrm{d}y)}{\int_0^\infty y\mathbb{P}(\tau\in K \mid \gamma = y)\mathbb{P}_\gamma(\mathrm{d}y)},$$

which completes the proof.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

5. Outlook: statistical applications

In univariate extreme value theory, a standard method for estimating the extreme value parameters fits all data exceeding a high threshold to a certain Poisson point process. This peaks-over-threshold approach has been generalized to the multivariate setting in [4], [17], and [28], and also to spatial processes [15]. Conditioning on the event that at least one component of a random vector is large, the recent contribution [14] analyzes the asymptotic distribution of exceedance counts of stationary sequences.

Here, we have suggested conditioning a stochastic process $\{\eta(t)\colon t \in T\}$ in the MDA of a max-stable process $\{\xi(t)\colon t \in T\}$ such that it converges to the incremental processes W in (1.2) or to the shape functions F in (1.4). In this section we provide several examples how these theoretical results can be used for statistical inference. The approach is a multivariate peaks-over-threshold method for max-stable processes, though the definition of extreme events differs from that in [14] and [28].

In the sequel, suppose that $\eta_1, \dots, \eta_n$, $n \in \mathbb{N}$, are independent observations of the random process η, already normalized to standard Pareto margins.

5.1. Incremental representation

For a max-stable process ξ that admits an incremental representation (1.2), the statistical merit of the convergence results in Theorem 2.1 and Theorem 2.2 is the 'de-convolution' of U and W, which allows us to substitute estimation of ξ by estimation of the process W. As only the single *extreme* events converge to W, we define the index set of extremal observations as

$$I_1(n) = \{i \in \{1, \dots n\}\colon \eta_i(t_0) > a(n)\}$$

for some fixed $t_0 \in T$. The set $\{\eta_i(\cdot)/\eta_i(t_0)\colon i \in I_1(n)\}$ then represents a collection of independent random variables that approximately follow the distribution of W. Thus, once the representation in (1.2) is known, both parametric and nonparametric estimation for the process W is feasible.

Note that, as usual in extreme value statistics, there is a trade-off between the number $I_1(n)$ of observations that is used for estimation and the threshold $a(n)$ which has to be large in order that the limit results approximately hold.

Example 5.1. (*Symmetric logistic distribution; cf. Example 2.3.*) The dependence parameter $q \geq 1$ of the symmetric logistic distribution (2.8) can be estimated by perceiving the conditional increments of η in the MDA as realizations of W and maximizing the likelihood

$$\begin{aligned}&\mathbb{P}(W(t_1) \in \mathrm{d}s_1, \dots, W(t_k) \in \mathrm{d}s_k \mid q)\\ &\quad = \bigg(1 + \sum_{i=1}^k s_i^{-q}\bigg)^{1/q-(k+1)} \bigg(\prod_{i=1}^k (iq-1)\bigg) \prod_{i=0}^k s_i^{-q-1}\,\mathrm{d}s_1 \cdots \mathrm{d}s_k.\end{aligned}$$

Example 5.2. (*Brown–Resnick processes; cf. [3] and [20].*) Let $\{Y(t)\colon t \in \mathbb{R}^d\}$ be a centered Gaussian process with stationary increments and $Y(t_0) = 0$ for some $t_0 \in \mathbb{R}^d$. Let $\gamma(t) = \mathbb{E}(Y(t) - Y(0))^2$, $t \in \mathbb{R}^d$, denote the variogram of Y. Then, with a Fréchet Poisson process $\sum_{i\in\mathbb{N}} \delta_{U_i}$ and independent copies Y_i of Y, $i \in \mathbb{N}$, the *Brown–Resnick process for the variogram* γ is given by

$$\xi(t) = \max_{i\in\mathbb{N}} U_i \exp\big[Y_i(t) - \tfrac{1}{2}\gamma(t - t_0)\big], \qquad t \in \mathbb{R}^d. \tag{5.1}$$

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

Its distribution depends only on γ. Here W from representation (1.2) is the log-Gaussian process. Hence, standard estimation procedures for Gaussian processes can be applied for statistical inference. Engelke *et al.* [13] explicitly constructed several new estimators of the variogram γ based on the incremental representation, which also covers Hüsler–Reiss distributions.

5.2. Mixed moving maxima representation

Similarly, if M is an M3 process as in (3.1), the convergence results of Theorem 3.1 can be used to estimate F (or $F_1 = F/\lambda$) on some compact domain K instead of estimating M directly. Here the index set T of the observed processes $\{\eta_i(t)\colon t \in T\}$, $i = 1, \ldots, n$, can be identified with $Q \oplus L$ from Theorem 3.1. The set L should be sufficiently large such that it is reasonable to assume that the components $\{U_i F_i(\cdot - S_i)\colon S_i \notin Q \oplus L\}$ hardly affect the process M on $Q \oplus K$ (that is, $\mu(C_B \cap A_L)/\mu(A_L) \approx \mathbb{P}(F(\cdot)/\lambda \in B)$ in the proof of Theorem 3.1). At the same time, a large set Q leads to a rich set of usable observations $\widetilde{F}_1^{(i)} = \eta_i(\tau_Q(\eta_i) + \cdot)/\eta_i(\tau_Q(\eta_i))$, $i \in I_2(n)$, where

$$I_2(n) = \Big\{ i \in \{1, \ldots, n\}\colon \max_{t \in Q} \eta_i(t) = \max_{t \in Q \oplus L} \eta_i(t) \geq a(n) \Big\}.$$

The resulting processes $\widetilde{F}_1^{(i)}$, $i \in I_2(n)$, can be interpreted as independent samples from an approximation to F_1. This approach can be expected to be particularly promising in the case of F having a simple distribution.

Our approach is flexible and versatile as it allows for parametric inference (Example 5.3 below gives estimators for the parametric family of processes considered in [9]), but also for nonparametric estimation methods (cf. Example 5.4 below).

Example 5.3. (*M3 processes with deterministic shape functions.*) Some examples of M3 processes have already been analyzed for statistical inference by de Haan and Pereira [9], who used normal, exponential, and t densities as shape functions:

$$F_1(t) = \exp\bigg[-\frac{\beta^2 t^2}{2}\bigg], \qquad \lambda = \frac{\beta}{\sqrt{2\pi}}, \tag{5.2}$$

$$F_1(t) = \exp[-\beta|t|], \qquad \lambda = \frac{\beta}{2}, \tag{5.3}$$

$$\text{and} \quad F_1(t) = (1 + \nu^{-1}\beta^2 t^2)^{-(\nu+1)/2}, \qquad \lambda = \frac{\beta\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}, \quad \nu > 0, \tag{5.4}$$

all parameterized by $\beta > 0$. De Haan and Pereira [9] introduced consistent and asymptotically normal estimators based on the interpretation of β as a dependence parameter.

From the samples $\widetilde{F}_1^{(i)}$, $i \in I_2(n)$, we obtain a new estimator for F_1:

$$\widehat{F}_1 = |I_2(n)|^{-1} \sum_{i \in I_2} \widetilde{F}_1^{(i)}.$$

Applying this estimator, β can be estimated by a least squares fit of (5.2)–(5.4) to $\widehat{F}_1$ at some locations $t_1, \ldots, t_m \in K$.

Example 5.4. (*Brown–Resnick processes.*) The M3 representation can also be employed for estimation of Brown–Resnick processes, although W has a much simpler form than F in this case (cf. [12] and [24]). A relation between the shape function F and the variogram γ of

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

the Brown–Resnick process can be obtained via the *extremal coefficient function* $\theta(\cdot)$. For a stationary, max-stable process ξ with identically distributed marginals, Schlather and Tawn [30] defined the extremal coefficient function θ via the relation $\mathbb{P}(\xi(0) \leq u,\ \xi(h) \leq u) = \mathbb{P}(\xi(0) \leq u)^{\theta(h)}$, $h \in \mathbb{R}^d$. For M3 processes, we have

$$\theta(h) = \mathbb{E} \int_{\mathbb{R}^d} \{F(t) \vee F(t+h)\}\,\mathrm{d}t = \frac{\mathbb{E} \int_{\mathbb{R}^d} \{F_1(t) \vee F_1(t+h)\}\,\mathrm{d}t}{\mathbb{E} \int_{\mathbb{R}^d} F_1(t)\,\mathrm{d}t} \tag{5.5}$$

and, at the same time, for Brown–Resnick processes [20],

$$\theta(h) = 2\Phi\big(\tfrac{1}{2}\sqrt{\gamma(h)}\big), \tag{5.6}$$

where Φ is the standard Gaussian distribution function. Identifying (5.5) with (5.6) and plugging in the samples $\widetilde{F}_1^{(i)}$, $i \in I_2(n)$, we obtain the variogram estimator

$$\widehat{\gamma}(h) = \left\{2\Phi^{-1}\left(\frac{\sum_{i \in I_2(n)} \int_{\widetilde{K}} \widetilde{F}_1^{(i)}(t) \vee \widetilde{F}_1^{(i)}(t+h)\,\mathrm{d}t}{2\sum_{i \in I_2(n)} \int_{\widetilde{K}} \widetilde{F}_1^{(i)}(t)\,\mathrm{d}t}\right)\right\}^2,$$

where $\widetilde{K}$ is a large set such that $\widetilde{K}$, $\widetilde{K} + h \subset K$.

Acknowledgements

The authors are grateful to an anonymous referee for valuable and detailed comments. The authors are indebted to Zakhar Kabluchko for useful suggestions and hints. S. Engelke was financially supported by Deutsche Telekom Stiftung and the Swiss National Science Foundation project 200021-134785. A. Malinowski was financially supported by the German Science Foundation (DFG), Research Training Group 1644 ‘Scaling problems in Statistics’. M. Oesting and M. Schlather were financially supported by Volkswagen Stiftung within the project ‘WEX-MOP’.

References

[1] Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stoch. Process. Appl.* **119,** 1055–1080. (Erratum: **121**(2011), 896–898.)

[2] Blanchet, J. and Davison, A. C. (2011). Spatial modeling of extreme snow depth. *Ann. Appl. Statist.* **5,** 1699–1725.

[3] Brown, B. M. and Resnick, S. I. (1977). Extreme values of independent stochastic processes. *J. Appl. Prob.* **14,** 732–739.

[4] Cai, J.-J., Einmahl, J. H. J. and de Haan, L. (2011). Estimation of extreme risk regions under multivariate regular variation. *Ann. Statist.* **39,** 1803–1826.

[5] Cooley, D., Davis, R. A. and Naveau, P. (2012). Approximating the conditional density given large observed values via a multivariate extremes framework, with application to environmental data. *Ann. Appl. Statist.* **6,** 1406–1429.

[6] Davison, A. C. and Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proc. R. Soc. London* **468,** 581–608.

[7] De Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Prob.* **12,** 1194–1204.

[8] De Haan, L. and Ferreira, A. (2006). *Extreme Value Theory*. Springer, New York.

[9] De Haan, L. and Pereira, T. T. (2006). Spatial extremes: models for the stationary case. *Ann. Statist.* **34,** 146–168.

[10] Dombry, C. and Ribatet, M. (2014). Functional regular variation, pareto processes and peaks over threshold. To appear in *Statist. Interface*.

[11] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Springer, Berlin.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press

[12] Engelke, S., Kabluchko, Z. and Schlather, M. (2011). An equivalent representation of the Brown–Resnick process. *Statist. Prob. Lett.* **81,** 1150–1154.
[13] Engelke, S., Malinowski, A., Kabluchko, Z. and Schlather, M. (2014). Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. To appear in *J. R. Statist. Soc. B*. Available at http://uk.arxiv.org/abs/1207.6886.
[14] Falk, M. and Tichy, D. (2012). Asymptotic conditional distribution of exceedance counts. *Adv. Appl. Prob.* **44,** 270–291.
[15] Ferreira, A. and de Haan, L. (2012). The generalized Pareto process; with a view towards application and simulation. Preprint. Available at http://uk.arxiv.org/abs/1203.2551.
[16] Gumbel, E. J. (1960). Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* **9,** 171–173.
[17] Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values. With discussions and reply by the authors. *J. R. Statist. Soc. Ser. B* **66,** 497–546.
[18] Hüsler, J. and Reiss, R.-D. (1989). Maxima of normal random vectors: between independence and complete dependence. *Statist. Prob. Lett.* **7,** 283–286.
[19] Kabluchko, Z. (2011). Extremes of independent Gaussian processes. *Extremes* **14,** 285–310.
[20] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Prob.* **37,** 2042–2065.
[21] Leadbetter, M. R. (1991). On a basis for 'peaks over threshold' modeling. *Statist. Prob. Lett.* **12,** 357–362.
[22] Meinguet, T. (2012). Maxima of moving maxima of continuous functions. *Extremes* **15,** 267–297.
[23] Meinguet, T. and Segers, J. (2010). Regularly varying time series in Banach spaces. Preprint. Available at http://uk.arxiv.org/abs/1001.3262.
[24] Oesting, M., Kabluchko, Z. and Schlather, M. (2012). Simulation of Brown–Resnick processes. *Extremes* **15,** 89–107.
[25] Padoan, S. A., Ribatet, M. and Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105,** 263–277.
[26] Resnick, S. I. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
[27] Resnick, S. I. and Roy, R. (1991). Random usc functions, max-stable processes and continuous choice. *Ann. Appl. Prob.* **1,** 267–292.
[28] Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli* **12,** 917–930.
[29] Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes* **5,** 33–44.
[30] Schlather, M. and Tawn, J. A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90,** 139–156.
[31] Smith, R. L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript. Available at http://www.unc.edu/∼rls/.
[32] Wang, Y. and Stoev, S. A. (2010). On the structure and representations of max-stable processes. *Adv. Appl. Prob.* **42,** 855–877.
[33] Whitt, W. (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Statist.* **41,** 939–944.

https://doi.org/10.1239/aap/1401369703 Published online by Cambridge University Press