

A Comparative Tour through the Simulation Algorithms for Max-Stable Processes

Marco Oesting and Kirstin Strokorb

Abstract. Being the max-analogue of α -stable stochastic processes, max-stable processes form one of the fundamental classes of stochastic processes. With the arrival of sufficient computational capabilities, they have become a benchmark in the analysis of spatiotemporal extreme events. Simulation is often a necessary part of inference of certain characteristics, in particular for future spatial risk assessment. In this article, we give an overview over existing procedures for this task, put them into perspective of one another and use some new theoretical results to make comparisons with respect to their properties.

Key words and phrases: Spectral representation, threshold stopping, extremal functions, error assessment.

1. INTRODUCTION

The severe consequences of extreme events such as strong windstorms, heavy precipitation or heat waves emphasize the need of appropriate statistical models for these types of events. To adequately assess the associated risk, not only the intensity, but also the spatial or spatiotemporal extent of extremes has to be taken into account. Motivated by the central limit theorem, classical geostatistics typically applies Gaussian processes to model the bulk of the distribution and the dependence structure of continuous variables. Being the natural analogue to Gaussian or, more generally, α -stable processes for maxima, max-stable processes are frequently used to model spatial and spatiotemporal extremes, specifically in environmental applications (cf. Davison, Padoan and Ribatet, 2012, Davison, Huser and Thibaud, 2018 and references therein).

The prevalence of max-stable processes as a benchmark is justified by the fact that they arise as the only possible location-scale max-limits of stochastic processes in the following way: Let X_1, X_2, X_3, \dots be independent copies of a real-valued process $X = \{X(\mathbf{x})\}_{\mathbf{x} \in S}$ on some locally compact metric space S . If there exist suitable location-scale norming sequences $a_n = a_n(\mathbf{x}) > 0$

and $b_n = b_n(\mathbf{x}) \in \mathbb{R}$, such that the law of

$$\left\{ \max_{i=1, \dots, n} \frac{X_i(\mathbf{x}) - b_n(\mathbf{x})}{a_n(\mathbf{x})} \right\}_{\mathbf{x} \in S}$$

converges in distribution to a stochastic process $Z = \{Z(\mathbf{x})\}_{\mathbf{x} \in S}$, then the resulting limit process Z necessarily satisfies a stability property with respect to the maximum operation. More precisely,

$$\left\{ \max_{i=1, \dots, n} \frac{Z_i(\mathbf{x}) - d_n(\mathbf{x})}{c_n(\mathbf{x})} \right\}_{\mathbf{x} \in S} \\ = \{Z(\mathbf{x})\}_{\mathbf{x} \in S} \quad \text{in distribution}$$

for independent copies Z_1, Z_2, Z_3, \dots of Z and appropriate location-scale norming sequences $c_n = c_n(\mathbf{x}) > 0$ and $d_n = d_n(\mathbf{x}) \in \mathbb{R}$. In this sense, the process Z is *max-stable* and the process X lies in its *max-domain of attraction*.

Over the last decades, max-stable processes have gained attention from several research perspectives. In the 1980s and early 1990s, they have mainly been studied from a probabilistic angle, resulting, for instance, in a full characterization of the class of sample-continuous simple max-stable processes; see de Haan (1984), Norberg (1986) and Giné, Hahn and Vatan (1990), among others. This work has been complemented by a precise description of the corresponding max-domain of attraction in de Haan and Lin (2001). Since the early 2000s, methods for statistical inference have been developed and, in parallel, suitable models for subclasses of max-stable processes have been introduced. Important examples for such models comprise Gaussian extreme value processes (Smith, 1990), extremal Gaussian processes (Schlather, 2002), Brown–Resnick processes (Kablichko, Schlather and de Haan,

Marco Oesting is Junior Professor of Computational Statistics, Stuttgart Center for Simulation Science & Institute for Stochastics and Applications, University of Stuttgart, 70569 Stuttgart, Germany (e-mail: marco.oesting@mathematik.uni-stuttgart.de). Kirstin Strokorb is Senior Lecturer, School of Mathematics, Cardiff University, Cardiff CF24 4AG, UK (e-mail: strokorbk@cardiff.ac.uk).

2009) and extremal- t processes (Opitz, 2013) providing a generalization of extremal Gaussian processes. With these flexible models and tools at hand, max-stable processes have become attractive for practitioners from various areas, in particular from environmental sciences.

However, a serious drawback is that most probabilistic properties of max-stable processes are analytically intractable. Therefore, simulation is often a necessary part of inference of certain characteristics, in particular for future spatial risk assessment. Meanwhile, starting from a general idea coined by Schlather (2002), a number of approaches to the simulation of max-stable processes have emerged: They include both approximate (Oesting, Kabluchko and Schlather, 2012, Oesting and Strokorb, 2018) and exact (Dieker and Mikosch, 2015, Dombry, Engelke and Oesting, 2016, Oesting, Schlather and Zhou, 2018, Liu et al., 2019) simulation procedures, some of them focusing on the particularly difficult problem of simulating within the subclass of Brown–Resnick processes. A first overview over some of these methods has been given in Oesting, Ribatet and Dombry (2016). The present article extends and updates this overview. New theoretical results allow to put the different methods into perspective of one another and to make comparisons with respect to their theoretical properties and their performance in numerical experiments.

Our text is structured as follows. To illustrate the need for efficient and accurate simulation of max-stable processes in an application, Section 2 describes a worked example addressing questions that involve the (joint) distribution of areal maxima over certain time horizons. Subsequently, we review existing simulation approaches in Section 3. Besides generic algorithms for the simulation of arbitrary max-stable processes, we also present more specific procedures that have been developed for the important and popular subclass of Brown–Resnick processes (Section 4). In Section 5, we provide new theoretical results that allow us to evaluate and compare the simulation approaches with respect to their efficiency and their accuracy. The results of a numerical study comparing the performance of the generic algorithms in a wide range of scenarios are reported in Section 6. We conclude with a discussion in Section 7 including further practical advice.

Full reproducible code of the analysis of Section 2 and the numerical study of Section 6, all of which were carried out in the statistical software environment R (R Core Team, 2020), are available at <https://github.com/Oesting/Comparative-tour-through-simulation-algorithms-for-max-stable-processes>.

Marginal Standardization

For convenience, let us recall a structural result at the very start. We shall assume throughout that all marginal distributions of a max-stable process $Z = \{Z(\mathbf{x})\}_{\mathbf{x} \in S}$ are

nondegenerate (i.e., not concentrated on a single value). This implies that the marginal distributions of Z are necessarily *Generalized Extreme Value (GEV) distributions* (Jenkinson, 1955 and von Mises, 1936)

$$\mathbb{P}(Z(\mathbf{x}) \leq z) = G_{\xi(\mathbf{x})} \left(\frac{z - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \right),$$

$$G_{\xi(\mathbf{x})}(z) = \exp\left(-\left(1 + \xi(\mathbf{x})z\right)_+^{-1/\xi(\mathbf{x})}\right),$$

with shape, location and scale parameters $\xi(\mathbf{x}), \mu(\mathbf{x}) \in \mathbb{R}$ and $\sigma(\mathbf{x}) > 0$, a result that goes back to Fréchet (1927) and Fisher and Tippett (1928) and was first rigorously proved in Gnedenko (1943). As the max-stability property is preserved under marginal transformations within the class of GEV distributions, attention is often drawn to max-stable processes with fixed marginal distributions such as the class of so-called *simple max-stable processes*, the ones that have standard unit Fréchet marginal distributions, that is, $\mathbb{P}(Z(\mathbf{x}) \leq z) = \exp(-1/z)$ for all $z > 0$ and $\mathbf{x} \in S$. To make this precise, if Z is a general max-stable process with shape function ξ , location function μ and scale function σ , then

$$\left\{ \left(1 + \xi(\mathbf{x}) \frac{Z(\mathbf{x}) - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \right)^{1/\xi(\mathbf{x})} \right\}_{\mathbf{x} \in S}$$

is a simple max-stable process and, conversely, if Z is a simple max-stable process, then

$$\left\{ \sigma(\mathbf{x}) \frac{Z(\mathbf{x})^{\xi(\mathbf{x})} - 1}{\xi(\mathbf{x})} + \mu(\mathbf{x}) \right\}_{\mathbf{x} \in S}$$

is a max-stable process with general GEV marginal distributions. Thus, the simple max-stable process associated with a general max-stable process in this way is a natural object to encapsulate its dependence structure (similar to the disentanglement of a multivariate distribution into marginal distributions and a copula). In addition, if we are concerned about the simulation of a max-stable process with given marginal distributions, this amounts to the simulation of a simple max-stable process and transforming it thereafter to the desired marginal distributions.

2. DATA EXAMPLE

In statistical practice, max-stable processes are used to approximate pointwise block maxima taken over sufficiently long time periods. In this data example, we consider daily maximum summer temperatures from 1990 to 2019 that were measured at 18 inland stations in the Netherlands and are freely available from <http://projects.knmi.nl/klimatologie/daggegevens/selectie.cgi> (cf. Figure 1) to answer questions about ...

(Q1) ... the distribution of the maximum inland temperature over a period of 2 weeks in summer,

(Q2) ... the probability that three subregions (SW, SE, NE, cf. Figure 4) experience a joint exceedance of 38°C within the same period of 14 days in summer.

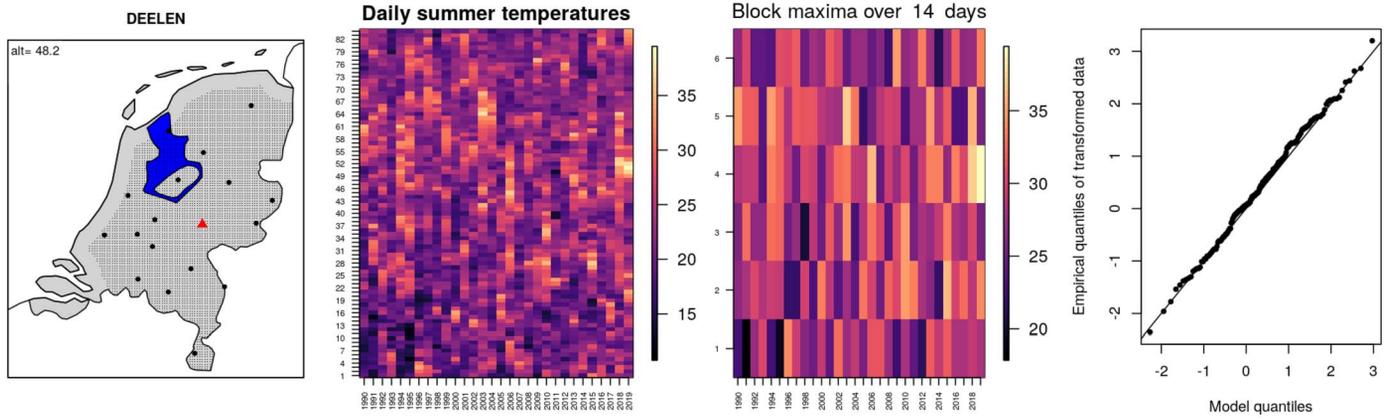


FIG. 1. Daily maximum summer temperatures on 84 summer days (5 June–27 August) from 1990 to 2019 at Deelen and their 14 day block maxima. Their location-scale transformation and the model quantiles arise from the marginal estimates as described in Section 2.

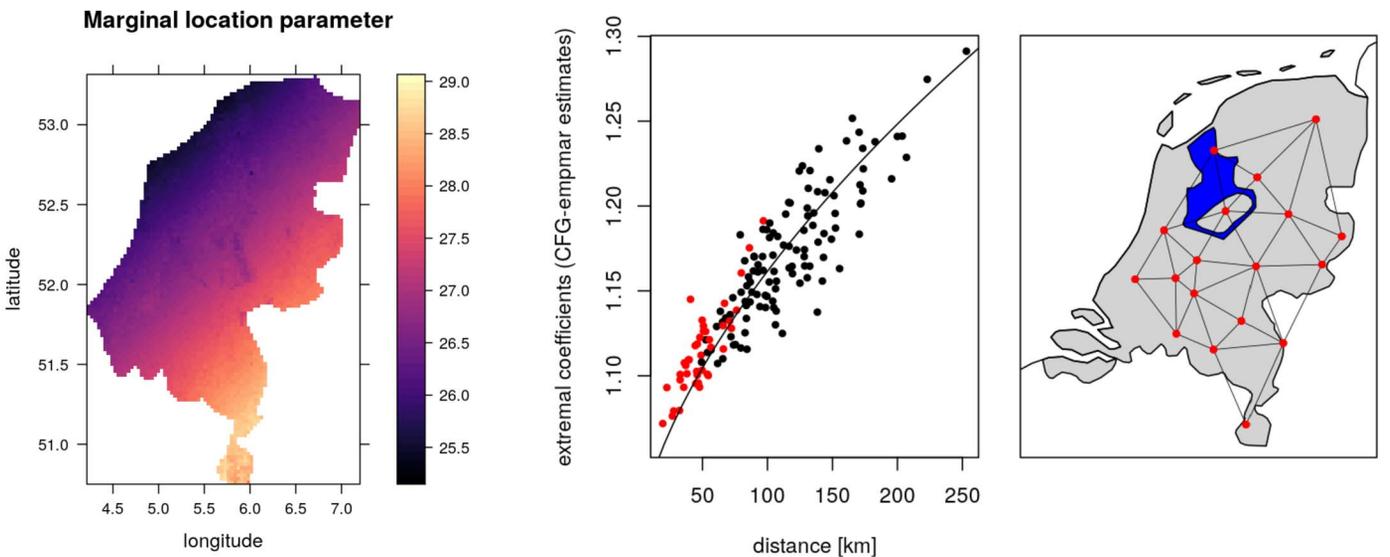


FIG. 2. Left: Location function of the marginal GEV fit evaluated on the entire inland grid. Middle: Nonparametric estimates of bivariate extremal coefficients corresponding to station pairs as estimated by the nonparametric procedure of Capéraà, Fougères and Genest (1997) (dots) and theoretical counterpart of the fitted isotropic spatial max-stable model (line). Red dots indicate pairs that have been included in the procedure of Einmahl et al. (2016) to estimate the spatial model. Right: Such pairs are connected in this graph.

To be more precise, when we say “inland,” we mean every location in the Netherlands that is 15 km away from the coast.¹ In this study, it is represented by a grid $S = \{y_j\}_{j=1}^{4712}$ of 4712 inland locations with (approximate) grid distance 2.5 km as displayed in Figures 1 (left) and 2 (right). As regards the temperature data, exemplarily, Figure 1 shows the daily maximum temperatures on 84 summer days (5 June to 27 August) from 1990 to 2019 as well as their 14 day block maxima at one of the 18 stations, Deelen. Considering temporal dependence, we did

¹Coastline data were obtained from `rnaturalearth` (South, 2017) and the package `geosphere` (Hijmans, 2019) used to compute coast distances. The data at stations in this region exhibit a more homogeneous behaviour compared to including stations closer to the coast.

not detect strong evidence against independence among 14 day block maxima, while spatial dependence seems to be of paramount importance, especially in the extreme values (as we shall see confirmed below). Hence, our working assumption is that these 14 day block maxima can be considered as 180 independent samples² (30 years with 6 blocks/year) of a max-stable process $\{Z(\mathbf{x})\}_{\mathbf{x} \in S}$ on the inland part S of the Netherlands measured at 18 sites $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{18} \in S$ as displayed in Figures 1 (left) and 2 (right).

²Missing values only occur at two sites, Stavoren and Arcen, where we lack knowledge about the first 29 summer days in 1990. The corresponding 14 day block maxima have been removed from the analysis.

Estimation

Following the paradigm of separating estimation of the marginal and dependence structure, a standard approach in Extreme Value Analysis, we first fitted a marginal GEV distribution $G(z; \mu, \sigma, \xi)$ to the 14 day block maxima via maximization of the independence likelihood. Here, while assuming the shape parameter ξ and the scale parameter σ to be constant, geographical information on the measurement stations serve as covariates in the location parameter $\mu = \mu_0 + \mu_1 \text{longitude} + \mu_2 \text{latitude} + \mu_3 \text{altitude}$. This maximum likelihood estimator for the GEV parameters is consistent and asymptotically normal for $\xi > -0.5$ (cf. Chandler and Bate, 2007 and Bücher and Segers, 2017) and is, for instance, readily available in the R package `extRemes` (Gilleland and Katz, 2016). In our case, the estimated GEV distributions are of Weibull-type ($\hat{\xi} \approx -0.27 < 0$). Exemplarily, Figure 1 (right) displays a QQ-plot of the site-specific location-scale-standardized block maxima against the model quantiles of $G(x; \hat{\xi})$. The marginal GEV-fit has been found reasonable at almost all sites. Figure 2 (left) displays the estimated location function $\hat{\mu}$, when evaluated on the inland grid. Altitude information on the inland grid stems from the ASTER Global Digital Elevation Model V2 (U.S./Japan ASTER Science Team) and was accessed with `geonames` (Rowlingson, 2019).

Subsequently, what remains to be modeled, is the dependence structure of Z , or equivalently, the associated simple max-stable process after marginal standardization as described in Section 1. For simplicity, we decided for an isotropic Brown–Resnick process with variogram $\gamma(\mathbf{h}) = \|\mathbf{h}/s\|^\alpha$. Details about this process can be found in Section 4 here. In addition to the estimation of the six marginal parameters, this amounts to the estimation of two further parameters to account for the spatial dependence, a smoothness parameter $\alpha \in (0, 2)$ and a spatial scale $s > 0$. An approach for this task, for which consistency and asymptotic normality have been established even under max-domain of attraction conditions, is the M-estimator of spatial tail dependence (Einmahl et al., 2016). We chose it here as a generic approach with good large sample properties that is not specifically tailored to the class of Brown–Resnick processes. Alternatively, one might consider likelihood based methods (Huser, Davison and Genton, 2016) or more bespoke methods such as the Peaks-over-threshold-inspired estimation of Engelke et al. (2015) that was proposed specifically for Brown–Resnick processes. The M-estimator relies on a selection of bivariate distributions only, that is, pairs among the measurement stations, and a number k of upper order statistics to be taken into account for the estimation. Our choice of pairs is displayed in Figure 2 (right) and we used $k = 50$. In the related package `tailDepFun` (Kiriliouk, 2016),

we used the option `iterate=T` to improve and update the (internal) distance weight matrix.

A typical sanity check after fitting a spatial max-stable model to station data is a comparison of pairwise non-parametric estimates of bivariate extremal coefficients $\theta(\mathbf{x}_i, \mathbf{x}_j)$ (cf. (5.2) for $K = \{\mathbf{x}_i, \mathbf{x}_j\}$ in Section 5.1) with the theoretical extremal coefficient function $\theta(\|\mathbf{x} - \mathbf{y}\|)$ of the estimated spatial model as displayed in Figure 2 (middle). Our bivariate nonparametric estimates are based on the procedure of Capéraà, Fougères and Genest (1997) as implemented in `evd` (Stephenson, 2002). The very low extremal coefficients, even for large distances, hint already at a strong extremal dependence among high temperatures.

Simulation

Finally, we draw 60,000 independent samples $Z_i = \{Z_i(\mathbf{y}_j)\}_{j=1}^{4712}$ from the fitted max-stable model $Z_i \sim Z$ ($i = 1, \dots, 60,000$) on the inland grid $(\mathbf{y}_j)_{j=1}^{4712}$, which can be interpreted in this context as 60,000 (grid-pointwise) temperature maxima of a 14-day summer interval. All simulations were carried out using the exact extremal functions methodology from Dombry, Engelke and Oesting (2016) (cf. Section 3.3 below). Normal random variables therein were generated with the fast pseudo random number generator from `dqrng` (Stubner, 2019). Figure 3 shows the first six simulations and a histogram of the corresponding 60,000 inland maxima, the latter being our data-driven answer to (Q1).

Since each summer consists of six such blocks in our setup, one may see this as sampling 10,000 times from a summer that is represented by the data. In this sense, the $(1 - 1/(6\ell))$ -empirical quantile constitutes a return level estimate for ℓ years. Figure 3 (right) contains these estimates for $\ell \in \{10, 100, 1000\}$. Since our estimated model has Weibull margins, the maximum upper endpoint across the grid is finite and gives us an estimate for the maximum summer temperature across the entire inland. According to this model it would be 43.1°C .

With regard to (Q2), we would like to draw attention to Figure 4, which shows a scatterplot of the 60,000 joint areal maxima

$$\left\{ \left(\max_{\mathbf{y}_j \in \text{SW}} Z_i(\mathbf{y}_j), \max_{\mathbf{y}_j \in \text{SE}} Z_i(\mathbf{y}_j), \max_{\mathbf{y}_j \in \text{NE}} Z_i(\mathbf{y}_j) \right) \right\}_{i=1}^{60,000}$$

across the three marked regions. In particular, the plot depicts that in our model dependence among high temperatures increases. In other words, particularly high temperatures are more likely to be experienced across a wider range of the inland. So our purely data-driven model would be in line with a (physical) theory that supports a single cause in such cases, a theory in which particularly large values may arise only during a wide-spread heat-wave. Counting the joint exceedances of 38°C in all three

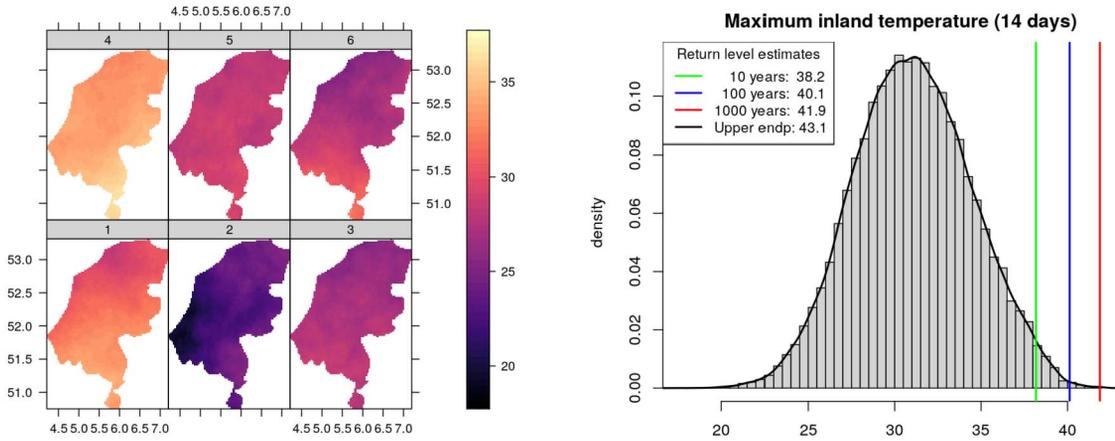


FIG. 3. Left: The first six out of 60,000 simulations from the estimated max-stable process on the inland grid after transformation to the estimated GEV marginal distributions. Right: Histogram of the corresponding 60,000 inland maxima. Return levels need to be interpreted with caution (cf. Section 2).

regions (marked red in Figure 4), leads to an estimate of $\hat{p} = 0.0046$ for a joint exceedance during a 14 day summer interval, our data-driven answer to Q2. Referring to an entire summer, one can interpret this as an estimate of $1 - (1 - \hat{p})^6 \approx 6\hat{p} \approx 2.75\%$ for the probability of a summer, in which 38°C is hit simultaneously in all three regions during at least one of the six 14-day summer periods.

Caution!

All results in this section need to be interpreted with caution. First of all, one needs to account for the uncertainty inherent in the estimated characteristics, such as the ones asked for in (Q1) and (Q2). Provided that the fitted model was perfectly correct, these could be estimated with arbitrary precision by the use of a sufficiently large number of simulations. In practice, however, as summarized in Davison and Huser (2015), there are different types of uncertainty in the model fit, including uncertainties related to taking measurements, choice of model class and parameter estimation, for instance.

The vigilant reader will have also noted, that already phrasing the initial questions (Q1) and (Q2) in this way comes along with several assumptions about the data and the underlying processes. Both (Q1) and (Q2) express that we assume to see a more or less homogeneous behavior of maximum temperatures within a given summer and otherwise independent and identically distributed summers. All the more, we would like to stress that our (imperfect) answers defy easy interpretation in a climatological context that is far beyond the scope of this article.

Instead, one may understand this section as a simple analysis of the current climate in the sense of the studied questions and time frame. We deliberately draw attention to the spatial rather than temporal aspects here, which is in line with the core contents of this article. And we hope that there is no doubt left about the critical role of the simulations in this data example to provide answers to questions that otherwise would be difficult to address at all.

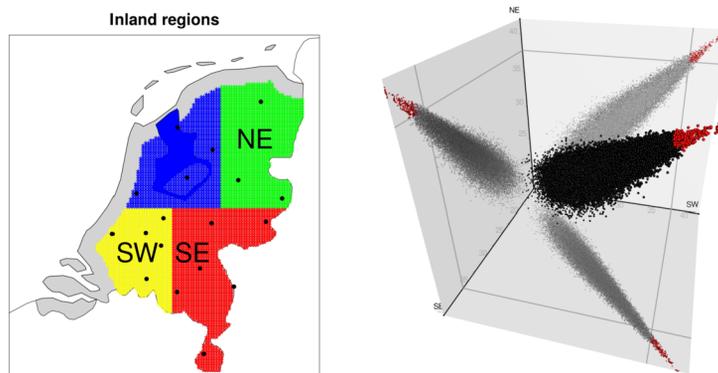


FIG. 4. Three inland regions (SW, SE, NE) and a scatterplot of simulated 60,000 joint areal maxima across these regions. Points (and their bivariate projections) are marked red if all three areal maxima exceed 38°C.

3. A SURVEY OF SIMULATION APPROACHES

In this section, we will give an overview over existing algorithms for the simulation of a simple max-stable process $\{Z(\mathbf{x})\}_{\mathbf{x} \in K}$ on a compact domain K . Almost all simulation approaches are based on the fact that any sample-continuous simple max-stable process Z possesses a *spectral representation*

$$(3.1) \quad \{Z(\mathbf{x})\}_{\mathbf{x} \in K} = \left\{ \max_{j \in \mathbb{N}} \Gamma_j^{-1} V_j(\mathbf{x}) \right\}_{\mathbf{x} \in K} \quad \text{in distribution,}$$

where $\{\Gamma_j\}_{j \in \mathbb{N}}$ are the arrival times of a unit rate Poisson process on $(0, \infty)$ with independent markings $\{V_j\}_{j \in \mathbb{N}}$ that are distributed according to a nonnegative sample-continuous stochastic process $\{V(\mathbf{x})\}_{\mathbf{x} \in K}$, the so-called *spectral process* (de Haan, 1984, Giné, Hahn and Vatan, 1990, Penrose, 1992). Since Z possesses standard unit Fréchet margins, the spectral process V satisfies $\mathbb{E}\{V(\mathbf{x})\} = 1$ for all $\mathbf{x} \in K$. Conversely, any continuous stochastic process V that satisfies the moment condition $\mathbb{E}\{V(\mathbf{x})\} = 1$ for all $\mathbf{x} \in K$ gives rise to a max-stable process Z via (3.1). The representation (3.1) is illustrated in Figure 5.

It is important to note that the law of the max-stable process Z in (3.1) does not uniquely determine the law of the spectral process V . Instead, a different spectral process V' in (3.1) may result in the same max-stable process Z . More precisely, two spectral processes V and V' generate the same max-stable process (in distribution) if and only if

$$\mathbb{E}\left\{ \max_{k=1}^n a_k V(\mathbf{x}_k) \right\} = \mathbb{E}\left\{ \max_{k=1}^n a_k V'(\mathbf{x}_k) \right\}$$

for all $n \in \mathbb{N}$, $a_k > 0$ ($k = 1, \dots, n$), $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset K$ (de Haan, 1978). We will call the processes V and V' *equivalent spectral processes* in this case. For instance, multiplying V with a positive random variable with unit expectation and independent of V yields the same max-stable

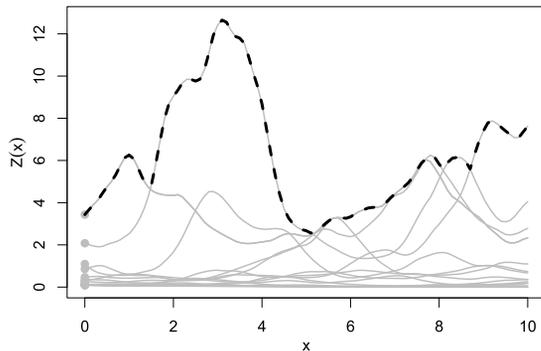


FIG. 5. Illustration of the spectral representation (3.1). Grey points represent the reciprocal arrival times Γ_j^{-1} , $j \in \mathbb{N}$, while grey lines correspond to the processes $\Gamma_j^{-1} V_j(\cdot)$, $j \in \mathbb{N}$. The resulting max-stable process Z is marked by the black dashed line.

process Z . In practice, the choice of spectral process V can have a major effect on the accuracy and efficiency of a certain simulation algorithm. As a starting point, however, we assume that a max-stable process Z is given by a specific choice of the spectral process V despite the fact that there may be other more convenient equivalent spectral processes V' for Z .

3.1 Threshold Stopping—the General Idea

The first simulation algorithm was introduced by Schlather (2002) and is motivated by the fact that the points $\{\Gamma_j\}_{j \in \mathbb{N}}$ are the arrival times of a renewal process with standard exponential interarrival times. In particular, the points $\{\Gamma_j\}_{j \in \mathbb{N}}$ are ordered: $\Gamma_1 < \Gamma_2 < \dots$ almost surely. Therefore, we would expect that the contribution of the process $\Gamma_j^{-1} V_j(\cdot)$ to the maximum in (3.1) becomes smaller and smaller as j gets large and at some point negligible. In other words, the distribution of Z can be approximated by the pointwise maximum

$$Z^{(T)}(\mathbf{x}) = \max_{j=1, \dots, T} \Gamma_j^{-1} V_j(\mathbf{x}), \quad \mathbf{x} \in K,$$

where T is a sufficiently large, but finite number. Instead of picking a deterministic number T , it turns out that an appropriately defined random stopping time T results in more accurate approximations. Typically, a threshold dependent stopping time

$$(3.2) \quad T = T_\tau = \min \left\{ j \in \mathbb{N} : \Gamma_{j+1}^{-1} \tau < \inf_{\mathbf{x} \in K} Z^{(j)}(\mathbf{x}) \right\}$$

is chosen, where $\tau > 0$ is a prescribed threshold reflecting an upper bound for the maximal contribution of the spectral process V to the maximum in (3.1). A precise description of the sampling procedure is given by Algorithm 1 below.

If the spectral process V is uniformly bounded, we can choose τ large enough to satisfy $\sup_{\mathbf{x} \in K} V(\mathbf{x}) < \tau$ almost surely. Clearly, in this case, (3.2) implies that for all $\mathbf{x} \in K$ and $j > T_\tau$

$$\Gamma_j^{-1} V_j(\mathbf{x}) < \Gamma_j^{-1} \tau \leq \Gamma_{T_\tau+1}^{-1} \tau < Z^{(T_\tau)}(\mathbf{x}).$$

Consequently,

$$\{Z^{(T_\tau)}(\mathbf{x})\}_{\mathbf{x} \in K} = \{Z^{(\infty)}(\mathbf{x})\}_{\mathbf{x} \in K} \quad \text{almost surely,}$$

and a sample from the finite maximum $Z^{(T_\tau)}$ can be seen as an exact sample from Z , since the distribution of $Z^{(\infty)}$ equals the distribution of Z by (3.1).

If, by contrast, $\mathbb{P}\{\sup_{\mathbf{x} \in K} V(\mathbf{x}) > \tau\} > 0$, there is a positive probability that $Z^{(T_\tau)}(\mathbf{x}) \neq Z^{(\infty)}(\mathbf{x})$ for some $\mathbf{x} \in K$. In this case, samples from $Z^{(T_\tau)}$ only provide approximations to the process of interest Z .

REMARK. The stopping time in (3.2) is almost surely finite, since we assumed the max-stable process Z to be sample-continuous. Indeed, sample-continuity implies

Algorithm 1 Threshold Stopping Algorithm

```

1 input: domain  $K$ , threshold  $\tau$ 
2 output: one (approximate/exact) max-stable process realization  $z$  on  $K$ 
3 begin
4   set  $z(\mathbf{x}) = 0$  for all  $\mathbf{x} \in K$ 
5   simulate  $\Gamma \sim \text{Exp}(1)$ 
6   while  $\tau/\Gamma \geq \inf_{\mathbf{x} \in K} z(\mathbf{x})$ 
7     simulate  $v \sim V$ 
8     set  $z(\mathbf{x}) = \max\{\Gamma^{-1}v(\mathbf{x}), z(\mathbf{x})\}$  for all  $\mathbf{x} \in K$ 
9     simulate  $E \sim \text{Exp}(1)$ 
10    set  $\Gamma = \Gamma + E$ 
11  end while
12  return  $z$ 
13 end

```

that $\inf_{\mathbf{x} \in K} Z(\mathbf{x}) > 0$ almost surely and only a finite number of functions $\{\Gamma_j^{-1}V_j(\mathbf{x})\}_{\mathbf{x} \in K, j \in \mathbb{N}}$, contributes to the maximum in (4.1); see Dombry and Eyi-Minko (2012), Theorem 2.2, and de Haan and Ferreira (2006), Corollary 9.4.4, respectively. Therefore, the infimum of the $Z^{(j)}$'s on the right-hand side in (3.2) exceeds 0 after a finite number of steps j almost surely, while the inverses of the Γ_j 's on the left-hand side tend to 0 with probability one. Consequently, the stopping time in (3.2) is almost surely finite.

3.2 Threshold Stopping—Normalizing Spectral Processes

As discussed above, Threshold Stopping Algorithm 1 provides exact realizations of the max-stable process Z if the spectral process V is almost surely bounded. If this is not the case for V , it can still often be transformed into an equivalent spectral process V' , which is uniformly bounded and generates the same max-stable process Z in the sense of (3.1). Two such procedures have been studied in further detail, both of which transform a given spectral process V into an equivalent spectral process $V^{\|\cdot\|}$, which is normalized w.r.t. some norm $\|\cdot\|$, that is, it satisfies

$$(3.3) \quad \|V^{\|\cdot\|}\| = \theta_{\|\cdot\|} \quad \text{almost surely for some } \theta_{\|\cdot\|} > 0.$$

The constant $\theta_{\|\cdot\|}$ is uniquely determined by

$$(3.4) \quad \theta_{\|\cdot\|} = \mathbb{E}\|V\| = \lim_{u \rightarrow \infty} u\mathbb{P}\{\|Z\| > u\}.$$

and does not depend on the choice of the starting spectral process V .

REMARK. More generally, Dombry and Ribatet (2015) show that for each sample-continuous simple max-stable process Z and each nonnegative measurable 1-homogeneous functional ℓ on $C(K, [0, \infty))$, there exists a spectral process V^ℓ for Z in the sense of (3.1) that is uniquely characterized by the property $\ell(V^\ell) = \theta_\ell$ a.s. for some constant $\theta_\ell > 0$. The constant θ_ℓ is necessarily given by $\theta_\ell = \mathbb{E}\ell(V) = \lim_{u \rightarrow \infty} u\mathbb{P}\{\ell(Z) > u\}$ and

we may call V^ℓ the ℓ -normalized spectral process of Z . If V is an arbitrary spectral process for the simple max-stable process Z , the equivalent ℓ -normalized spectral process V^ℓ can be obtained by a measure transform $V^\ell = \theta_\ell Y / \ell(Y)$, where $\mathbb{P}\{Y \in dv\} = \ell(v) / \theta_\ell \mathbb{P}\{V \in dv\}$. The ℓ -normalized spectral process V^ℓ characterizes extremes of stochastic processes also in terms of threshold exceedances instead of maxima. Let X be a sample-continuous process in the max-domain of attraction of Z . Then, as $u \uparrow \infty$, the conditional distribution of $u^{-1}X$ given that $\ell(X) > u$ converges weakly to the distribution of the product $P \cdot \theta_\ell^{-1}V^\ell$, where P is a standard Pareto random variable independent of the process V^ℓ . Thus, the resulting limit process, the so-called ℓ -Pareto process (Dombry and Ribatet, 2015), is fully described by the ℓ -normalized spectral process V^ℓ and being able to effectively simulate from V^ℓ has important implications beyond the max-stable context. In what follows, with the max-stable simulation context in view, we consider the cases $\ell = \|\cdot\|_\infty$ (when K is compact) and $\ell = \|\cdot\|_1$ and ℓ being the evaluation at a single point in K (when K is a finite set) and discuss how the resulting processes V^ℓ are related to each other.

Sup-normalization. The first transformation of this type was introduced in Oesting, Schlather and Zhou (2018) who proposed a normalization w.r.t. the supremum norm $\|f\|_\infty = \sup_{\mathbf{x} \in K} f(\mathbf{x})$ for $f \geq 0$. Starting from an arbitrary sample-continuous spectral process V , the unique equivalent sup-normalized spectral process $V^{\|\cdot\|_\infty}$, which satisfies $\sup_{\mathbf{x} \in K} V^{\|\cdot\|_\infty}(\mathbf{x}) = \theta_{\|\cdot\|_\infty}$ almost surely, can be obtained as the normalization

$$(3.5) \quad V^{\|\cdot\|_\infty}(\mathbf{x}) = \frac{\theta_{\|\cdot\|_\infty} Y(\mathbf{x})}{\sup_{\mathbf{x}' \in K} Y(\mathbf{x}')}$$

of a stochastic process Y with transformed law

$$\begin{aligned} & \mathbb{P}\{Y \in A\} \\ &= \frac{1}{\theta_{\|\cdot\|_\infty}} \int_A \sup_{\mathbf{x} \in K} v(\mathbf{x}) \mathbb{P}\{V \in dv\}, \quad A \subset C(K). \end{aligned}$$

As the spectral processes V and $V^{\|\cdot\|_\infty}$ are equivalent, we can use the sup-normalized spectral process for simulation. By construction, all the sample paths of $V^{\|\cdot\|_\infty}$ are bounded by $\theta_{\|\cdot\|_\infty}$ almost surely, and by [Resnick and Roy \(1991\)](#), sample-continuity of Z already implies that $\theta_{\|\cdot\|_\infty}$ is finite (see also [de Haan and Ferreira, 2006](#), Theorem 9.6.1). Therefore, the output of Algorithm 1 with threshold $\tau = \theta_{\|\cdot\|_\infty}$ is an exact realization of the max-stable process Z , when the sup-normalized spectral representation $V^{\|\cdot\|_\infty}$ is used therein.

By (3.5), simulation of $V^{\|\cdot\|_\infty}$ can be based on simulation of the transformed process Y . While [Oesting, Schlather and Zhou \(2018\)](#) suggest an approximating Markov Chain Monte Carlo (MCMC) algorithm for this task, more recently, [de Fondeville and Davison \(2018\)](#) present a relation that allows for exact simulation of Y via rejection sampling provided that a simulation procedure for the normalized spectral process $V^{\|\cdot\|}$ for an arbitrary norm is given. We refer to Section 4.3 for further efficiency improvements when V is log-Gaussian. The resulting process $Y/\|Y\|_\infty$ has then the law of $V^{\|\cdot\|_\infty}/\theta_{\|\cdot\|_\infty}$. While analytic expressions for the normalizing constant $\theta_{\|\cdot\|_\infty}$ are usually not available, it can still be estimated in the course of the simulation procedure, for example, via the relation $\theta_{\|\cdot\|_\infty} = \mathbb{E}\|V\|_\infty$, and can subsequently be used for an ex post normalization. The constant $\theta_{\|\cdot\|_\infty}$ is also known as *extremal coefficient*; cf. (5.2).

Sum-normalization. The second transformation of type (3.3) uses a normalization w.r.t. the ℓ_1 -norm $\|f\|_1 = \sum_{k=1}^N f(\mathbf{x}_k)$ for $f \geq 0$ on a finite domain $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$. It has been proposed by [Dieker and Mikosch \(2015\)](#) for the special case of Brown–Resnick processes (see Section 4) and extended to a more general framework by [Dombry, Engelke and Oesting \(2016\)](#). The starting point for the construction of this representation is the fact that, given the distribution of Z , for each $k \in \{1, \dots, N\}$, there is a unique equivalent spectral process $V^{(k)}$ of Z with the property $V^{(k)}(\mathbf{x}_k) = 1$ almost surely. Its law is given by

$$(3.6) \quad P_k(A) = \int_{[0, \infty)^K} v(\mathbf{x}_k) \mathbf{1}\{v \in v(\mathbf{x}_k)A\} \mathbb{P}\{V \in dv\},$$

for $A \subset [0, \infty)^K = [0, \infty)^{\{\mathbf{x}_1, \dots, \mathbf{x}_N\}}$, where V is an arbitrary spectral process of Z . [Dombry, Engelke and Oesting \(2016\)](#) show that the unique equivalent sum-normalized spectral process $V^{\|\cdot\|_1}$, which satisfies $\sum_{k=1}^N V^{\|\cdot\|_1}(\mathbf{x}_k) = \theta_{\|\cdot\|_1} = N$ almost surely, is then given by

$$V^{\|\cdot\|_1} = N \frac{Y}{\|Y\|_1}, \quad \text{where } Y \sim \frac{1}{N} \sum_{k=1}^N P_k.$$

Since $V^{\|\cdot\|_1}$ satisfies $\|V^{\|\cdot\|_1}\|_1 = N$ almost surely, and each component of a vector is bounded by its ℓ_1 -norm,

that is, $V^{\|\cdot\|_1}(\mathbf{x}) \leq \|V^{\|\cdot\|_1}\|_1 = N$ for all $\mathbf{x} \in K$, the sum-normalized spectral process $V^{\|\cdot\|_1}$ can be used as spectral process for Algorithm 1 resulting in exact realizations for the threshold $\tau = N$. [Dombry, Engelke and Oesting \(2016\)](#) also explicitly calculate the laws P_k , $k = 1, \dots, N$, and thus, verify that they can be easily sampled for many popular max-stable models such as Brown–Resnick processes or extremal- t processes.

3.3 Extremal Functions

A simulation procedure that essentially differs from the previously considered threshold stopping algorithm is the extremal functions approach, which was also introduced in [Dombry, Engelke and Oesting \(2016\)](#). Instead of simulating the elements of the Poisson point process $\Phi = \{\varphi_j\}_{j \in \mathbb{N}}$ with $\varphi_j(\cdot) = \Gamma_j^{-1} V_j(\cdot)$ in an ascending order w.r.t. Γ_j until a stopping criterion is fulfilled, the idea is to simulate only the so-called *extremal functions* ([Dombry and Eyi-Minko, 2013](#)) that definitely contribute to the final maximum in (3.1) on the finite domain $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, that is, all the functions $\varphi \in \Phi$ such that

$$(3.7) \quad \varphi(\mathbf{x}_k) = \max_{j \in \mathbb{N}} \varphi_j(\mathbf{x}_k)$$

for some $k \in \{1, \dots, N\}$. It can be shown that, for each $k \in \{1, \dots, N\}$, with probability one, there is exactly one (extremal) function $\varphi \in \Phi$ satisfying (3.7), which we denote by $\varphi_+^{(k)}$ in the following. The algorithm subsequently simulates the processes

$$(3.8) \quad Z_+^{(k)}(\cdot) = \max_{j=1, \dots, k} \varphi_+^{(j)}(\cdot).$$

By construction, the process $Z_+^{(k)}$ is exact at $\mathbf{x}_1, \dots, \mathbf{x}_k$, that is,

$$\{Z_+^{(k)}(\mathbf{x}_i)\}_{i=1, \dots, k} = \{Z(\mathbf{x}_i)\}_{i=1, \dots, k} \quad \text{in distribution.}$$

In particular, the final process $Z_+^{(N)}$ has the same distribution as the desired max-stable process Z on the entire domain $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$.

The theory behind this procedure stems from [Dombry and Eyi-Minko \(2013\)](#) who show—by using Slivnyak–Mecke type arguments—that, for each $k = 1, \dots, N$, the point process $\Phi \setminus \{\varphi_+^{(1)}, \dots, \varphi_+^{(k)}\}$ is conditionally independent from $\{\varphi_+^{(1)}, \dots, \varphi_+^{(k)}\}$ conditional on $Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_k)$. Based on this result, the law of the collection of all N needed extremal functions $\{\varphi_+^{(k)}\}_{k=1, \dots, N}$ can be described by an iterative procedure:

- Initially, $\varphi_+^{(1)} \sim \Gamma_1^{-1} V^{(1)}$.
- Subsequently, the conditional law of the next extremal function $\varphi_+^{(k)}$, when the previous extremal functions $\varphi_+^{(1)}, \dots, \varphi_+^{(k-1)}$ are already given, can be described as follows:

(i) Either $\varphi_+^{(k)}$ is the unique argmax of the evaluation functional at \mathbf{x}_k in a Poisson point process $\Phi^{(k)}$ on $[0, \infty)^K$, whose intensity measure equals the intensity measure of Φ restricted to the set

$$\{\varphi \in [0, \infty)^K : \varphi(\mathbf{x}_i) < Z_+^{(k-1)}(\mathbf{x}_i) \\ \text{for } i < k \text{ and } \varphi(\mathbf{x}_k) > Z_+^{(k-1)}(\mathbf{x}_k)\}$$

(ii) Or, in the event that $\Phi^{(k)}$ is empty, $\varphi_+^{(k)}$ is one of the previous extremal functions.

The restricted Poisson point process $\Phi^{(k)}$ in (i) can be conveniently simulated by iteratively generating Poisson points $\varphi_j^{(k)} = \Gamma_j^{-1} V_j$ from the original Poisson point process Φ , where we choose the spectral processes V_j as independent copies of the special spectral process $V^{(k)}$ (cf. (3.6)). This ensures that $\varphi_j^{(k)}(\mathbf{x}_k) = \Gamma_j^{-1} V_j(\mathbf{x}_k) = \Gamma_j^{-1}$ a.s. Thereby, similar to a threshold stopping procedure, the potential new value at \mathbf{x}_k is running through the descending values $\Gamma_1^{-1} > \Gamma_2^{-1} > \dots$. However, since we are ultimately interested in $\Phi^{(k)}$ (and not Φ), we test each time if $\varphi_j^{(k)}(\mathbf{x}_i) < Z_+^{(k-1)}(\mathbf{x}_i)$ for all $i < k$. If that happens, we have found our new argmax within $\Phi^{(k)}$, and hence extremal function $\varphi_+^{(k)} = \varphi_j^{(k)}$. Otherwise, we will arrive at a Γ_j^{-1} that falls below $Z_+^{(k-1)}(\mathbf{x}_k)$ for some $j \geq 1$. This corresponds to the event that $\Phi^{(k)}$ is empty.

The entire procedure is summarized by Algorithm 2 below.

3.4 Summary of Generic Simulation Approaches

The aforementioned simulation approaches to obtain a simple max-stable process Z are generic in the sense that they are not tailored to a specific class of max-stable processes. Figure 6 provides a quick overview. We would

like to draw attention to the fact that each of these approaches relies on the ability to simulate from a (family of) specific spectral process(es) V for the max-stable process Z ; cf. also Table 2. Considering a finite domain $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, the extremal functions approach needs the simulation of spectral processes $V^{(1)}, \dots, V^{(N)}$ (i.e., simulation from the N measures in (3.6)) readily available. The sum-normalized and sup-normalized approaches are by definition based on the availability of the spectral processes $V^{\|\cdot\|_1}$ and $V^{\|\cdot\|_\infty}$, respectively. As explained in Section 3.2, availability of $V^{(1)}, \dots, V^{(N)}$ guarantees availability of $V^{\|\cdot\|_1}$ at a negligible additional computational cost (drawing a point from the finite domain $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$). Further, if any normalized $V^{\|\cdot\|}$ is available (for instance, the sum-normalized spectral process $V^{\|\cdot\|_1}$), then it can also be used to simulate the sup-normalized spectral functions $V^{\|\cdot\|_\infty}$ via rejection sampling up to a multiplicative constant (de Fondeville and Davison, 2018); cf. Section 3.2. Rejection sampling is usually costly; the computational cost can be reduced in some special cases; cf. Section 4.3.

4. SPECIALTIES FOR BROWN–RESNICK PROCESSES

Among several popular max-stable processes, the class of Brown–Resnick processes stands out as a parsimonious spatial model that has become a benchmark in spatial extremes. Let $\{W(\mathbf{x})\}_{\mathbf{x} \in K}$ be a centered Gaussian process with variance $\{\sigma^2(\mathbf{x})\}_{\mathbf{x} \in K}$. Then the max-stable process $\{Z(\mathbf{x})\}_{\mathbf{x} \in K}$ that is associated to the spectral process

$$(4.1) \quad V(\mathbf{x}) = \exp\left(W(\mathbf{x}) - \frac{\sigma^2(\mathbf{x})}{2}\right), \quad \mathbf{x} \in K$$

Algorithm 2 Extremal Functions Algorithm

```

1  input: domain  $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ 
2  output: one (exact) max-stable process realization  $z$  on  $K$ 
3  begin
4      set  $z(\mathbf{x}_k) = 0$  for  $k = 1, \dots, N$ 
5      for  $k = 1, \dots, N$ 
6          simulate  $\Gamma \sim \text{Exp}(1)$ 
7          while  $\Gamma^{-1} > z(\mathbf{x}_k)$ 
8              simulate  $v \sim P_k$ 
9              if  $k = 1$  or  $\Gamma^{-1} v(\mathbf{x}_i) < z(\mathbf{x}_i)$  for all  $i < k$ 
10                 set  $z(\mathbf{x}_i) = \max\{\Gamma^{-1} v(\mathbf{x}_i), z(\mathbf{x}_i)\}$  for all  $i \geq k$ 
11             end if
12             simulate  $E \sim \text{Exp}(1)$ 
13             set  $\Gamma = \Gamma + E$ 
14         end while
15     end for
16     return  $z$ 
17 end

```

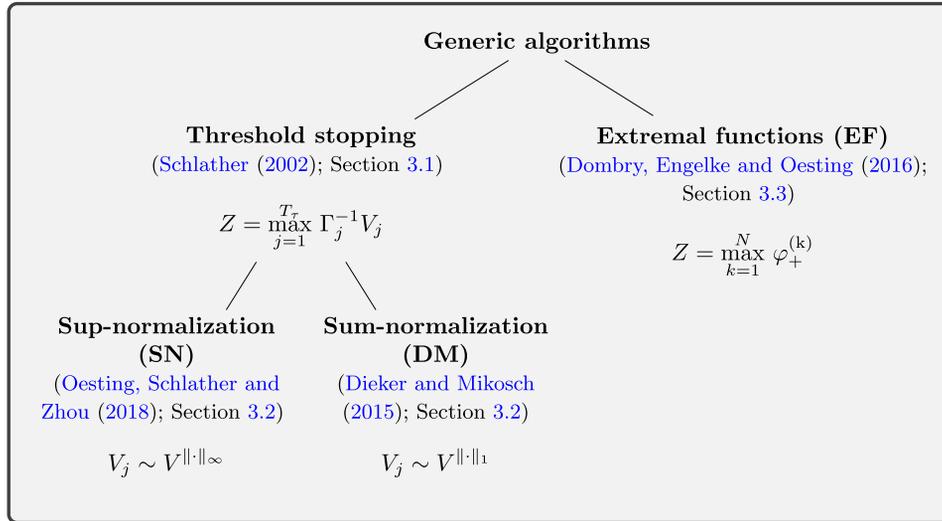


FIG. 6. Overview over generic simulation approaches to obtain a max-stable process Z from one of its spectral processes V .

via (3.1) has unit Fréchet marginal distributions and its law depends only on the *variogram*

$$\begin{aligned} \gamma : K \times K &\rightarrow [0, \infty), \\ \gamma(\mathbf{x}, \mathbf{y}) &= \mathbb{E}(W(\mathbf{x}) - W(\mathbf{y}))^2 \end{aligned}$$

(Kabluchko, 2011). In particular, for $K \subset \mathbb{R}^d$, the max-stable process Z is stationary if the variogram γ depends only on $\mathbf{x} - \mathbf{y}$ and by slight abuse of notation we write $\gamma(\mathbf{x} - \mathbf{y}) = \gamma(\mathbf{x}, \mathbf{y})$ in this case. The stationary process Z has first been introduced in Kabluchko, Schlather and de Haan (2009) in this generality and is now widely known as *Brown–Resnick process*. In practice, among unbounded variograms on \mathbb{R}^d , it is almost exclusively the variogram family $\gamma(\mathbf{x} - \mathbf{y}) = \|(\mathbf{x} - \mathbf{y})/s\|^\alpha$, $s > 0$, $\alpha \in (0, 2)$ of fractional Brownian sheets (fBS) that is considered in applications.

4.1 Threshold Stopping Based on Gaussian Mixtures

The first attempts of simulating a Brown–Resnick process Z were based on threshold stopping with a log-Gaussian spectral process V as in (4.1) satisfying $W(\mathbf{x}_o) = 0$ for some $\mathbf{x}_o \in K$. Typically, the origin $\mathbf{o} \in \mathbb{R}^d$ belongs to the simulation domain K and $\mathbf{x}_o = \mathbf{o}$. We refer to such spectral processes $V = V^{(\text{orig})}$ as the *original* spectral representation of the Brown–Resnick process Z . Since log-Gaussian processes do not have an almost sure upper bound, such a threshold stopping procedure based on $V^{(\text{orig})}$ cannot be exact. Instead, a typical phenomenon is that the threshold stopping procedure works well in a neighborhood of \mathbf{x}_o , where the variance of the underlying Gaussian process is small, but a simulation bias appears in those parts of the domain where the variance is large. To avoid this phenomenon, Oesting, Kabluchko and

Schlather (2012) introduced a uniformly distributed *random shift* in the spectral process

$$(4.2) \quad \begin{aligned} V^{(\text{shift})}(\mathbf{x}) \\ = V^{(\text{orig})}(\mathbf{x} - \mathbf{S}), \quad \mathbf{x} \in K, \mathbf{S} \sim \text{Unif}(K). \end{aligned}$$

Note that the superimposed homogeneity comes however at the cost of increasing the variance of the spectral process even further in most situations.

More recently, Oesting and Strokorb (2018) explain how a variance reduction of the Gaussian process W in (4.1) with fixed variogram γ can lead to faster and more accurate simulations based on the threshold stopping procedure. Specifically, when W is chosen such that the maximal variance $\sup_{\mathbf{x} \in K} \text{Var}(W(\mathbf{x}))$ is minimal among all Gaussian processes on K with variogram γ , we call the corresponding spectral process in (4.1) *minimal variance spectral process* $V^{(\text{minvar})}$. Table 1 lists the corresponding minimal Gaussian processes on the d -dimensional hyperrectangle $[-\mathbf{R}, \mathbf{R}] \subset \mathbb{R}^d$ for the variogram family $\gamma(\mathbf{h}) = \|\mathbf{h}/s\|^\alpha$, $\alpha \in (0, 2)$, $s > 0$. For $d \geq 2$ and $\alpha \in (0, 1)$, the minimal representation is unknown. However, also in this case the modified Gaussian process

$$W(\mathbf{x}) = W^{(\text{orig})}(\mathbf{x}) - 2^{-d} \sum_{\mathbf{v} \in \text{Ex}([- \mathbf{R}, \mathbf{R}])} W^{(\text{orig})}(\mathbf{v}),$$

where $\text{Ex}([- \mathbf{R}, \mathbf{R}])$ is the set of vertices of the simulation domain $[- \mathbf{R}, \mathbf{R}] \subset \mathbb{R}^d$, has a substantially reduced maximal variance compared to the original process $W^{(\text{orig})}$ and should be preferred.

4.2 Record Breakers

An exact simulation procedure for Brown–Resnick processes, which is specifically tailored to this class, is the record breakers approach by Liu et al. (2019). It is based on the original spectral representation (3.1) with V being

TABLE 1

Gaussian process $W^{(\text{minvar})}$ with variogram $\gamma(\mathbf{h}) = \|\mathbf{h}/s\|^\alpha$, $\alpha \in (0, 2)$, $s > 0$ that minimize the maximal variance on the hyperrectangle $[-\mathbf{R}, \mathbf{R}] \subset \mathbb{R}^d$. The process is either given by its covariance $C^{(\text{minvar})}$ or built from the original representation $W^{(\text{orig})}$

	$d = 1$	$d \geq 2$
$\alpha \in (0, 1]$	$C^{(\text{minvar})}(x, y) = 2^{-1} s^{-\alpha} (\Gamma(\frac{2-\alpha}{2}) \Gamma(\frac{1+\alpha}{2}) \Gamma(\frac{1}{2})^{-1} - x - y ^\alpha)$	unknown
$\alpha \in [1, 2)$	$W^{(\text{minvar})}(\mathbf{x}) = W^{(\text{orig})}(\mathbf{x}) - 2^{-d} \sum_{\mathbf{v} \in \text{Ex}([- \mathbf{R}, \mathbf{R}])} W^{(\text{orig})}(\mathbf{v})$	

a log-Gaussian random field as in (4.1). Let $a, c \in (0, 1)$ and $C > 0$ and consider the three random times

$$N_X = \sup \left\{ n \in \mathbb{N} : \max_{i=1, \dots, N} V_n(\mathbf{x}_i) > n^a \exp(C) \right\},$$

$$N_\Gamma = \sup \{ n \in \mathbb{N} : \Gamma_n \leq cn \},$$

$$N_a = \sup \left\{ n \in \mathbb{N} : nc \leq \frac{\Gamma_1 n^a \exp(C)}{\min_{i=1, \dots, N} V_1(\mathbf{x}_i)} \right\}$$

$$= \left\lfloor \left(\frac{\Gamma_1 \exp(C)}{c \min_{i=1, \dots, N} V_1(\mathbf{x}_i)} \right)^{1/(1-a)} \right\rfloor.$$

From the definition of N_X , N_Γ and N_a , it is easily checked that

$$\{Z^{\{\max\{N_X, N_\Gamma, N_a\}}(\mathbf{x})}\}_{\mathbf{x} \in K}$$

$$= \{Z^{(\infty)}(\mathbf{x})\}_{\mathbf{x} \in K} \quad \text{in distribution.}$$

While N_a can be obtained directly from Γ_1 and V_1 , simulation of the random times N_X and N_Γ is more sophisticated. For the simulation of N_Γ , Liu et al. (2019) make use of the fact that $\{\Gamma_n - cn\}_{n \in \mathbb{N}}$ is a random walk with positive drift. An algorithm is provided that subsequently samples the times when the random walk crosses zero. Due to its positive drift the process will finally stay positive. To obtain N_X , all so-called *record-breaking* times $\eta_1 < \eta_2 < \dots$, that is, all $\eta \in \mathbb{N}$ such that

$$\max_{i=1, \dots, N} V_\eta(\mathbf{x}_i) > n^a \exp(C),$$

are subsequently simulated. The finiteness of all moments of $\max_{i=1, \dots, N} V(\mathbf{x}_i)$ implies that the number of record-breaking times is almost surely finite. Consequently, the record-breakers algorithm requires an almost surely finite number of simulations of log-Gaussian processes V_i to obtain a realization of the Brown–Resnick process Z .

REMARK. On a practical note, Liu et al. (2019) also provide guidance on how to choose the parameters $a, c \in (0, 1)$ and $C > 0$. However, there is an additional parameter $\delta \in (0, 1)$ involved, where the practical implications of this choice and its interplay with the other parameters are still open.

4.3 Generic Approaches

Besides these approaches that are rather specific to the class of Brown–Resnick processes, general procedures such as simulation based on normalized spectral processes (Section 3.2) or the extremal functions approach (Section 3.3) can be used; cf. also Figure 6. Considering a finite domain $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, the distribution P_k from (3.6) is the distribution of the stochastic process $\{\exp(W^{(\text{orig})}(\mathbf{x} - \mathbf{x}_k))\}_{\mathbf{x} \in K}$ (Dombry, Engelke and Oesting, 2016). Therefore, the extremal functions approach as in Algorithm 2 is readily available for Brown–Resnick processes. Further, this implies that the sum-normalized spectral process is of the form

$$V^{\|\cdot\|_1}(\mathbf{x})$$

$$(4.3) \quad = N \frac{\exp(W^{(\text{orig})}(\mathbf{x} - \mathbf{S}))}{\sum_{k=1}^N \exp(W^{(\text{orig})}(\mathbf{x} - \mathbf{x}_k))}, \quad \mathbf{x} \in K,$$

where \mathbf{S} is uniformly distributed on $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ and independent of $W^{(\text{orig})}$, which has been demonstrated already earlier in Dieker and Mikosch (2015). Finally, the simple representation (4.3) of the sum-normalized spectral functions can also be used to simulate the sup-normalized spectral functions via rejection sampling (de Fondeville and Davison, 2018). Modifications of the last approach to reduce the rejection rate and alternative MCMC procedures have recently been proposed by Oesting, Schlather and Schillings (2019).

REMARK. Furthermore, Ho and Dombry (2019) show that, conditional on the component $k^* \in \{1, \dots, N\}$ where the maximum is assumed, the distribution of the vector $(V^{\|\cdot\|_\infty}(\mathbf{x}_k))_{k=1, \dots, N} / V^{\|\cdot\|_\infty}(\mathbf{x}_{k^*})$ equals the distribution of a log-Gaussian vector conditional on not exceeding one—a fact that can be used for its simulation. As efficient sampling from such a conditional distribution is not straightforward in high dimension and the calculation of the distribution of k^* involves the inversion of several matrices of sizes $N \times N$ and $(N - 1) \times (N - 1)$ as well as evaluations of $(N - 1)$ -dimensional Gaussian distribution functions; however, this approach is limited to small or moderate N in practice.

5. DESIRABLE PROPERTIES

Simulation algorithms are supposed to provide results in an efficient and accurate way. In this section, we investigate the performance of the algorithms above with respect to these two aspects from a theoretical angle. While the efficiency of an algorithm can be characterized in terms of its computational complexity, we measure its accuracy in terms of distributional properties of the approximation error. The proofs for this section are postponed to Appendix A.

5.1 Efficiency

Apart from the record breakers approach (Section 4.2), which is tailored to the class of Brown–Resnick processes, all other simulation algorithms reviewed in this manuscript are based on the simulation of standard Poisson points Γ_j on the positive real line and associated spectral processes V_j on the simulation domain K only. Hence, if $c_V(K)$ denotes the computational complexity of simulating a single spectral process V on the domain K and $N_V(K)$ denotes the total number of spectral processes V_j to be simulated in such a simulation algorithm, then the law of the product $N_V(K) \cdot c_V(K)$ describes the computational complexity of the entire procedure. As the second factor $c_V(K)$ inevitably depends on the simulation technique used to generate samples from the specific spectral function V , we focus our analysis mainly on the first factor $N_V(K)$ henceforth.

In case of the Threshold Stopping Algorithm 1, the random number $N_V(K)$ of spectral processes to be simulated coincides with the stopping time $T = T_\tau$ from (3.2). Its expected value can be bounded as follows.

PROPOSITION 5.1.

(a) *The expected stopping time of the Threshold Stopping Algorithm 1 is bounded from below by*

$$(5.1) \quad \mathbb{E}(N_V(K)) = \mathbb{E}(T_\tau) \geq \tau \mathbb{E} \left\{ 1 / \inf_{x \in K} Z(x) \right\}.$$

(b) *Equality in (5.1) holds if and only if $\sup_{x \in K} V(x) \leq \tau$ almost surely.*

The lower bound in (5.1) is finite for sample-continuous Z by Theorem 2.2 in Dombry and Eyi-Minko (2012). It should be relatively sharp in most practically relevant situations, while an ad hoc rough upper bound is given by

$$\mathbb{E}(N_V(K)) = \mathbb{E}(T_\tau) \leq 1 + \mathbb{E} \left\{ \tau / \inf_{x \in K} V(x) - 1 \right\}_+.$$

Another interpretation of Proposition 5.1 (b) is that equality in (5.1) holds if and only if the threshold stopping algorithm produces exact samples of the max-stable process Z ; cf. Section 3.1. Equality in (5.1) in this situation was already proved by a different technique in Oesting, Ribatet and Dombry (2016). Naturally, the following respective results for the normalized spectral representations (Section 3.3) can be recovered as special cases.

COROLLARY 5.2.

(a) (Oesting, Schlather and Zhou, 2018) *The expected stopping time of the Threshold Stopping Algorithm 1 with sup-normalized representation $V = V^{\|\cdot\|_\infty}$ and threshold $\tau = \theta_{\|\cdot\|_\infty}$ is*

$$\mathbb{E}(N_{V^{\|\cdot\|_\infty}}(K)) = \mathbb{E}T_{\theta_{\|\cdot\|_\infty}} = \theta_{\|\cdot\|_\infty} \mathbb{E} \left\{ 1 / \inf_{x \in K} Z(x) \right\}.$$

(b) (Dombry, Engelke and Oesting, 2016) *The expected stopping time of the Threshold Stopping Algorithm 1 with sum-normalized representation $V = V^{\|\cdot\|_1}$ and threshold $\tau = N$ is*

$$\mathbb{E}(N_{V^{\|\cdot\|_1}}(K)) = \mathbb{E}T_N = N \mathbb{E} \left\{ 1 / \inf_{x \in K} Z(x) \right\}.$$

An interesting observation is that the expressions for $\mathbb{E}(N_V(K)) = \mathbb{E}T_\tau$ above depend on the law of the spectral process V used only via the law of the resulting max-stable process Z . In particular, if V is any spectral process for Z , the constant

$$(5.2) \quad \begin{aligned} \theta_{\|\cdot\|_\infty} &= \mathbb{E} \left\{ \sup_{x \in K} V(x) \right\} \\ &= -\log \mathbb{P} \left\{ \sup_{x \in K} Z(x) \leq 1 \right\} \end{aligned}$$

is usually known as *extremal coefficient* of Z on K . For $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, it ranges between 1 and N and can be interpreted as the effective number of independent variables in the set $\{Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_N)\}$. In view of Corollary 5.2, being able to effectively simulate from a sup-normalized spectral process is therefore a worthwhile endeavor. What is however unclear in this general setting, is how the computational complexities $c_{V^{\|\cdot\|_1}}(K)$ and $c_{V^{\|\cdot\|_\infty}}(K)$ of obtaining a single realization of either $V^{\|\cdot\|_1}$ or $V^{\|\cdot\|_\infty}$ relate to one another. Here, a more effective simulation technique for $V^{\|\cdot\|_1}$ rather than for $V^{\|\cdot\|_\infty}$ may outweigh the reduction of the factor $\mathbb{E}(N_V(K)) = \mathbb{E}(T_\tau)$ by using $V = V^{\|\cdot\|_\infty}$ instead of $V = V^{\|\cdot\|_1}$; see Section 4.3 for related references for the case of Brown–Resnick processes. This is a general trade-off associated with the choice of the spectral process V for threshold stopping algorithms. Should it be easy to simulate Z from V (low $N_V(K)$) or should be easy to simulate V itself (low $c_V(K)$)? For $V = V^{\|\cdot\|_1}$ and $V = V^{\|\cdot\|_\infty}$, we can at least trace $\mathbb{E}(N_V(K))$ analytically as just discussed.

What is more, Dombry, Engelke and Oesting (2016) show that the expected number of simulated spectral processes in the Extremal Functions Algorithm 2 neither depends on the law of Z nor on the geometry of the domain K .

PROPOSITION 5.3 (Dombry, Engelke and Oesting, 2016). *The expected number of spectral processes to be simulated in the Extremal Functions Algorithm 2 in order to obtain an exact sample of Z on the set $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ equals N , that is, $\mathbb{E}N_V(K) = N$ for this algorithm.*

TABLE 2

The expected number $\mathbb{E}(N_V(K))$ of spectral functions V_j to be simulated to obtain an exact sample of a max-stable process Z on a set $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ for three generic simulation algorithms. Each method relies on the ability to simulate from specific spectral functions V

Method/Reference	Spectral fcts. V	$\mathbb{E}(N_V(K))$	
SN	Sup-normalized threshold stopping (Oesting, Schlather and Zhou, 2018, Section 3.2)	$V^{\ \cdot\ _\infty}$	$\theta_{\ \cdot\ _\infty} \mathbb{E}\{1/\inf_{\mathbf{x} \in K} Z(\mathbf{x})\}$
DM	Sum-normalized threshold stopping (Dieker and Mikosch, 2015, Section 3.2)	$V^{\ \cdot\ _1}$	$N \mathbb{E}\{1/\inf_{\mathbf{x} \in K} Z(\mathbf{x})\}$
EF	Extremal functions (Dombry, Engelke and Oesting, 2016, Section 3.3)	$V^{(1)}, \dots, V^{(N)}$	N

Table 2 summarizes these findings on the efficiency of the three generic exact simulation algorithms considered in this section. Since the max-stable process Z to be simulated has standard Fréchet margins, we have

$$(5.3) \quad \mathbb{E}\left\{1/\inf_{\mathbf{x} \in K} Z(\mathbf{x})\right\} \geq 1,$$

and equality holds if and only if Z is almost surely constant on K . Hence, apart from this exceptional case, the expected number of simulated spectral processes $\mathbb{E}(N_V(K))$ for the extremal functions approach is always smaller than the corresponding number for the sum-normalized approach. According to the results in Dombry, Engelke and Oesting (2016) (cf. also Section 3.4), the spectral processes involved in the two approaches are very closely related to each other, that is, their complexities $c_V(K)$ are almost identical. Thus, in terms of computational complexity, the extremal functions approach is always preferable to the sum-normalized approach if exact samples are desired.

REMARK. For the record breakers approach, Liu et al. (2019) show that the expected number $\mathbb{E}(N_V(\{\mathbf{x}_1, \dots, \mathbf{x}_N\}))$ of spectral processes $V_j \sim V^{(\text{orig})}$ to be simulated in order to produce an exact sample of a Brown–Resnick process Z lies in $o(N^\varepsilon)$ for any $\varepsilon > 0$. The result is however difficult to compare with Table 2 as it holds for fixed $K \supset \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ only. For instance, the corresponding result for the sup-normalized threshold stopping could be phrased as $\mathbb{E}(N_V(\{\mathbf{x}_1, \dots, \mathbf{x}_N\})) \in \mathcal{O}(1)$ for $V = V^{\|\cdot\|_\infty}$. This exacerbates meaningful comparisons.

5.2 Accuracy

While simulation via normalized spectral functions with appropriate thresholds or the extremal functions approach produce exact samples from the distribution of a max-stable process, these algorithms can be computationally expensive. Therefore, also the analysis of nonexact simulation algorithms is of interest.

Threshold stopping. Our main focus lies on the potentially nonexact Threshold Stopping Algorithm 1 in what follows. As explained in Section 3.1, such an algorithm can be nonexact if the threshold τ is exceeded by the spectral process V on K with positive probability. Naturally, decreasing the threshold τ reduces the computational cost. But at the same time, the simulation is more likely to be less accurate as well. To make this specific, let us define the *simulation error* as the deviation of the resulting finite approximation $Z^{(T_\tau)}$ from the exact sample $Z = Z^{(\infty)}$. The following proposition provides a very general description of the distribution of the simulation error.

PROPOSITION 5.4. *For any measurable function $f : C(K) \times K \rightarrow [0, \infty]$, we have*

$$\begin{aligned} & \mathbb{P}\{|Z(\mathbf{x}) - Z^{(T_\tau)}(\mathbf{x})| > f(Z^{(T_\tau)}, \mathbf{x}) \text{ for some } \mathbf{x} \in K\} \\ &= 1 - \mathbb{E}_{Z^{(T_\tau)}} \left\{ \exp \left(-\mathbb{E}_V \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z^{(T)}(\mathbf{x}) + f(Z^{(T_\tau)}, \mathbf{x})} \right. \right. \right. \\ & \quad \left. \left. \left. - \sup_{\mathbf{x} \in K} \frac{\tau}{Z^{(T_\tau)}(\mathbf{x})} \right\}_+ \right) \right\}, \end{aligned}$$

where the spectral process V and the stopped process $Z^{(T_\tau)}$ are stochastically independent.

Specifically, by setting $f(Z, \mathbf{x}) = \varepsilon$ or $f(Z, \mathbf{x}) = \varepsilon Z(\mathbf{x})$, Proposition 5.4 entails the probability that an absolute error of size larger than ε occurs

$$\begin{aligned} \mathcal{P}_{\tau, \varepsilon}^{(\text{abs})} &= \mathbb{P}\left\{ \sup_{\mathbf{x} \in K} |Z(\mathbf{x}) - Z^{(T_\tau)}(\mathbf{x})| > \varepsilon \right\} \\ &= 1 - \mathbb{E}_{Z^{(T_\tau)}} \left\{ \exp \left(-\mathbb{E}_V \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z^{(T_\tau)}(\mathbf{x}) + \varepsilon} \right. \right. \right. \\ & \quad \left. \left. \left. - \sup_{\mathbf{x} \in K} \frac{\tau}{Z^{(T_\tau)}(\mathbf{x})} \right\}_+ \right) \right\}, \end{aligned}$$

and the probability that a relative error of size larger than ε occurs

$$\begin{aligned} \mathcal{P}_{\tau, \varepsilon}^{(\text{rel})} &= \mathbb{P} \left\{ \sup_{\mathbf{x} \in K} \frac{|Z(\mathbf{x}) - Z^{(T_\tau)}(\mathbf{x})|}{Z^{(T_\tau)}(\mathbf{x})} > \varepsilon \right\} \\ &= 1 - \mathbb{E}_{Z^{(T_\tau)}} \left\{ \exp \left(-\mathbb{E}_V \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{(1 + \varepsilon)Z^{(T_\tau)}(\mathbf{x})} \right. \right. \right. \\ &\quad \left. \left. \left. - \sup_{\mathbf{x} \in K} \frac{\tau}{Z^{(T_\tau)}(\mathbf{x})} \right\} \right) \right\}. \end{aligned}$$

Both error occurrence probabilities are increasing as the error size ε goes to zero. For $\varepsilon = 0$, they coincide with the probability that a simulation error occurs at all

$$\begin{aligned} \mathcal{P}_\tau &= \mathbb{P} \left\{ \sup_{\mathbf{x} \in K} |Z(\mathbf{x}) - Z^{(T_\tau)}(\mathbf{x})| > 0 \right\} \\ (5.4) \quad &= 1 - \mathbb{E}_{Z^{(T_\tau)}} \left\{ \exp \left(-\mathbb{E}_V \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z^{(T_\tau)}(\mathbf{x})} \right. \right. \right. \\ &\quad \left. \left. \left. - \sup_{\mathbf{x} \in K} \frac{\tau}{Z^{(T_\tau)}(\mathbf{x})} \right\} \right) \right\}, \end{aligned}$$

which can serve as a benchmark. In the notation of Section 3.3, an approximation error occurs precisely when the finite approximation $Z^{(T_\tau)}$ does not involve all extremal functions $\varphi \in \Phi_+ = \{\varphi_+^{(1)}, \dots, \varphi_+^{(N)}\}$. The situation gets worse, the larger the number of missing extremal functions

$$M_\tau = \left| \left\{ \Gamma_j^{-1} V_j \in \Phi_+ : \Gamma_j^{-1} \tau \leq \inf_{\mathbf{x} \in K} \max_{k < j} \Gamma_k^{-1} V_k(\mathbf{x}) \right\} \right|.$$

is. The expected number of missing extremal functions $\mathbb{E}(M_\tau)$ is a natural upper bound for the error probability \mathcal{P}_τ , that is, $\mathcal{P}_\tau \leq \mathbb{E}(M_\tau)$.

PROPOSITION 5.5. *The expected number of missing extremal functions $\mathbb{E}(M_\tau)$ in the finite approximation $Z^{(T_\tau)}$ of the max-stable random field Z is bounded by*

$$(5.5) \quad \mathbb{E}(M_\tau) \leq \mathbb{E} \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z(\mathbf{x})} - \sup_{\mathbf{x} \in K} \frac{\tau}{Z(\mathbf{x})} \right\}_+$$

where the max-stable process Z and the spectral process V are stochastically independent.

REMARK. Oesting, Schlather and Zhou (2018) showed that

$$\mathbb{E}|\Phi_+| = \mathbb{E} \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z(\mathbf{x})} \right\}.$$

In view of (5.1) and $\mathbb{E}(|\Phi_+|) \leq \mathbb{E}(T_\tau) + \mathbb{E}(M_\tau)$, we believe that inequality (5.5) provides a relatively sharp bound for the error term $\mathbb{E}(M_\tau)$. In particular, it is sharper than the bound

$$\mathbb{E}(M_\tau) \leq \mathbb{E} \left\{ \left(\sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z(\mathbf{x})} \right) \mathbf{1}_{\{\sup_{\mathbf{x} \in K} V(\mathbf{x}) > \tau\}} \right\}$$

in the proof of Proposition 10.4.2 in Oesting, Ribatet and Dombry (2016). A significantly simplified (though less sharp) version of (5.5) is obtained by

$$\begin{aligned} \mathbb{E}(M_\tau) &\leq \mathbb{E} \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x}) - \tau}{Z(\mathbf{x})} \right\}_+ \\ &\leq \mathbb{E} \left\{ \sup_{\mathbf{x} \in K} (V(\mathbf{x}) - \tau) \right\}_+ \mathbb{E} \left\{ 1 / \inf_{\mathbf{x} \in K} Z(\mathbf{x}) \right\}. \end{aligned}$$

For both, the error bound (5.5) and the exact error (5.4), it is generally difficult to provide analytic expressions. The precise terms can however be assessed for finite $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ via simulation; see Appendix B for details. The assessment is based on the observations that all the extremal functions of a max-stable process can be simulated via the Extremal Functions Algorithm 2 and that the potentially relevant nonextremal functions can be simulated independently once the extremal functions and the process Z are known. This allows us to check which of these functions would have been taken into account by the Threshold Stopping Algorithm 1. Hence, we can compare the approximation $Z^{(T_\tau)}$ and the exact realization Z and identify the missing extremal functions.

Extremal functions. Besides the threshold stopping algorithm, also the Extremal Functions Algorithm 2 may include a simulation error if not all extremal functions $\varphi_+^{(1)}, \dots, \varphi_+^{(N)}$ are taken into account. Considering the approximation $Z_+^{(n)}$ of $Z = Z_+^{(N)}$ on $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ after the n th step of the extremal functions algorithm as given in (3.8) for some $n \leq N$ yields the following analogies to Propositions 5.4 and 5.5.

PROPOSITION 5.6. *For any measurable function $f : C(K) \times K \rightarrow [0, \infty]$, we have*

$$\begin{aligned} &\mathbb{P} \{ |Z(\mathbf{x}) - Z_+^{(n)}(\mathbf{x})| > f(Z_+^{(n)}, \mathbf{x}) \text{ for some } \mathbf{x} \in K \} \\ &= 1 - \mathbb{E}_{Z_+^{(n)}} \left\{ \exp \left(-\mathbb{E}_V \left\{ \max_{i=1, \dots, n} \frac{V(\mathbf{x}_i)}{Z_+^{(n)}(\mathbf{x})} \right. \right. \right. \\ &\quad \left. \left. \left. - \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z_+^{(n)}(\mathbf{x}) + f(Z_+^{(n)}, \mathbf{x})} \right\} \right) \right\}, \end{aligned}$$

where the spectral process V and the process $Z_+^{(n)}$ are stochastically independent.

PROPOSITION 5.7. *The expected number $\mathbb{E}(M_+^{(n)})$ of missing extremal functions $M_+^{(n)}$ after n steps of the extremal functions algorithm can be computed as*

$$(5.6) \quad \begin{aligned} \mathbb{E}(M_+^{(n)}) &= \mathbb{E} \left\{ \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z(\mathbf{x})} \right\} \\ &\quad - \mathbb{E} \left\{ \max_{i=1, \dots, n} \frac{V(\mathbf{x}_i)}{Z(\mathbf{x}_i)} \right\}, \end{aligned}$$

where the max-stable process Z and the spectral process V are stochastically independent.

6. COMPARATIVE NUMERICAL STUDY

In order to gain further insights on the comparative performance of the different approaches to max-stable process simulation, specifically as we deviate from the exact setting, the absence of analytic expressions makes it necessary to conduct a broader numerical study. We focus on comparing the three generic and potentially exact methods from Table 2/Figure 6 (DM/EF/SN) when applied to generate (approximate or exact) samples from the widely used classes of:

(i) *Brown–Resnick processes* (Kablichko, Schlather and de Haan, 2009) with spectral representation (4.1) and fBS variogram $\gamma(\mathbf{h}) = 2v\|\mathbf{h}\|^\alpha$, where $v > 0$ and $\alpha \in (0, 2)$.

(ii) *extremal- t processes* (Opitz, 2013) with spectral representation

$$(6.1) \quad V(\mathbf{x}) = \frac{\sqrt{\pi}2^{1-\nu/2}}{\Gamma((\nu+1)/2)}W(\mathbf{x})_+^\nu, \quad \mathbf{x} \in K,$$

where W is a standard Gaussian random field with exponential correlation function $\rho(\mathbf{h}) = \text{Cov}(W(\mathbf{x} + \mathbf{h}), W(\mathbf{x})) = \exp(-\|\mathbf{h}/s\|)$ with scale $s > 0$ and the parameter $\nu > 0$ influences the degrees of freedom of the underlying multivariate t -distribution in the dependence structure, cf. Opitz (2013).

Both classes of processes are based on underlying Gaussian random fields W and all algorithms considered to obtain an (approximate or exact) sample $Z^{(\text{simulated})}$ of the associated max-stable process Z are based on repeated sampling from W . Therefore, the number $N_W(K)$ of Gaussian processes $W_j \sim W$ needed for one sample of Z constitutes a natural measure for the algorithms' efficiency and we call the number $\mathbb{E}(N_W(K))$ that is needed on average the *mean time* in this context.

For exact simulation, it is possible to derive the precise mean times of each algorithm from the theoretical considerations of Section 5; see Table 3. While sampling from the spectral processes $V^{(1)}, V^{(2)}, \dots, V^{(N)}$ and $V^{\|\cdot\|_1}$ is straightforward for Brown–Resnick and extremal- t processes and involves only one Gaussian process simula-

tion for each sample of the respective spectral process (see Dombry, Engelke and Oesting, 2016, Dieker and Mikosch, 2015 or Sections 3.2 and 4.3), we choose to use the rejection sampling algorithm proposed by de Fondeville and Davison (2018) based on sum-normalized spectral processes as proposals to obtain (exact) samples from the sup-normalized spectral function $V^{\|\cdot\|_\infty}$. In this case, we need to take into account the average acceptance rate $\theta_{\|\cdot\|_\infty}/N$; see Table 3. In view of (5.3), this shows already that for exact simulation the EF algorithm should be preferred over the DM approach and the SN approach according to the mean time $\mathbb{E}(N_W(K))$.

The main purpose of our study is now to investigate the relative efficiency of the algorithms as we vary their accuracy in a reasonable range. As a simulation domain, we consider the 501 equi-distantly spaced points $K = \{-1, -0.996, \dots, 1\}$ in the interval $[-1, 1]$. In the Brown–Resnick case, we consider the parameter scenarios that arise from choosing α in $\{0.2, 0.6, 1.0, 1.4, 1.8\}$ (rather noisy to rather smooth) and variance parameter ν in $\{0.5, 1, 2\}$. The extremal- t scenarios consist of $\nu \in \{1, 2, 4\}$ and scale parameter $s \in \{0.5, 1, 2\}$. Further, we prespecify error probabilities $\mathcal{P} = \mathbb{P}(Z^{(\text{simulated})} \neq Z) \in \{0, 0.01, 0.05, 0.1\}$ that we are willing to tolerate, while observing the corresponding times $N_W(K)$ and estimating the mean time $\mathbb{E}(N_W(K))$ based on 50,000 simulations in each case. For the threshold stopping approaches DM and SN, the error probability \mathcal{P} coincides with the benchmark error term \mathcal{P}_τ in (5.4). That is, for these algorithms we need to select the threshold τ appropriately in order ensure \mathcal{P} assumes the right value. For the EF approach, we deviate from accuracy by fixing an appropriate equi-distantly spaced subset of locations in the simulation domain K . The appropriate thresholds and subsets for given error probability \mathcal{P} were also found simulation-based.

The results of our study for Brown–Resnick processes are reported in Figures 7 and 8 and for extremal- t processes in Figure 9. Some of the observations are as expected. The mean time increases in each scenario with lower tolerable error probability. It also increases in the Brown–Resnick case with higher variance and as the

TABLE 3

The expected number $\mathbb{E}(N_W(K))$ of Gaussian processes W_j to be simulated to obtain an exact sample of the associated Brown–Resnick or extremal- t process Z on a set $K = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ for the three generic simulation algorithms from Table 2, where the SN algorithm is based on rejection sampling (de Fondeville and Davison, 2018). The ratio $c_V(K)/c_W(K)$ represents the number of samples from W that are needed to obtain a sample from V

Method	Spectral functions V	$c_V(K)/c_W(K)$	Mean time $\mathbb{E}(N_W(K))$
SN	$V^{\ \cdot\ _\infty}$	$N/\theta_{\ \cdot\ _\infty}$	$N\mathbb{E}\{1/\inf_{\mathbf{x} \in K} Z(\mathbf{x})\}$
DM	$V^{\ \cdot\ _1}$	1	$N\mathbb{E}\{1/\inf_{\mathbf{x} \in K} Z(\mathbf{x})\}$
EF	$V^{(1)}, \dots, V^{(N)}$	1	N

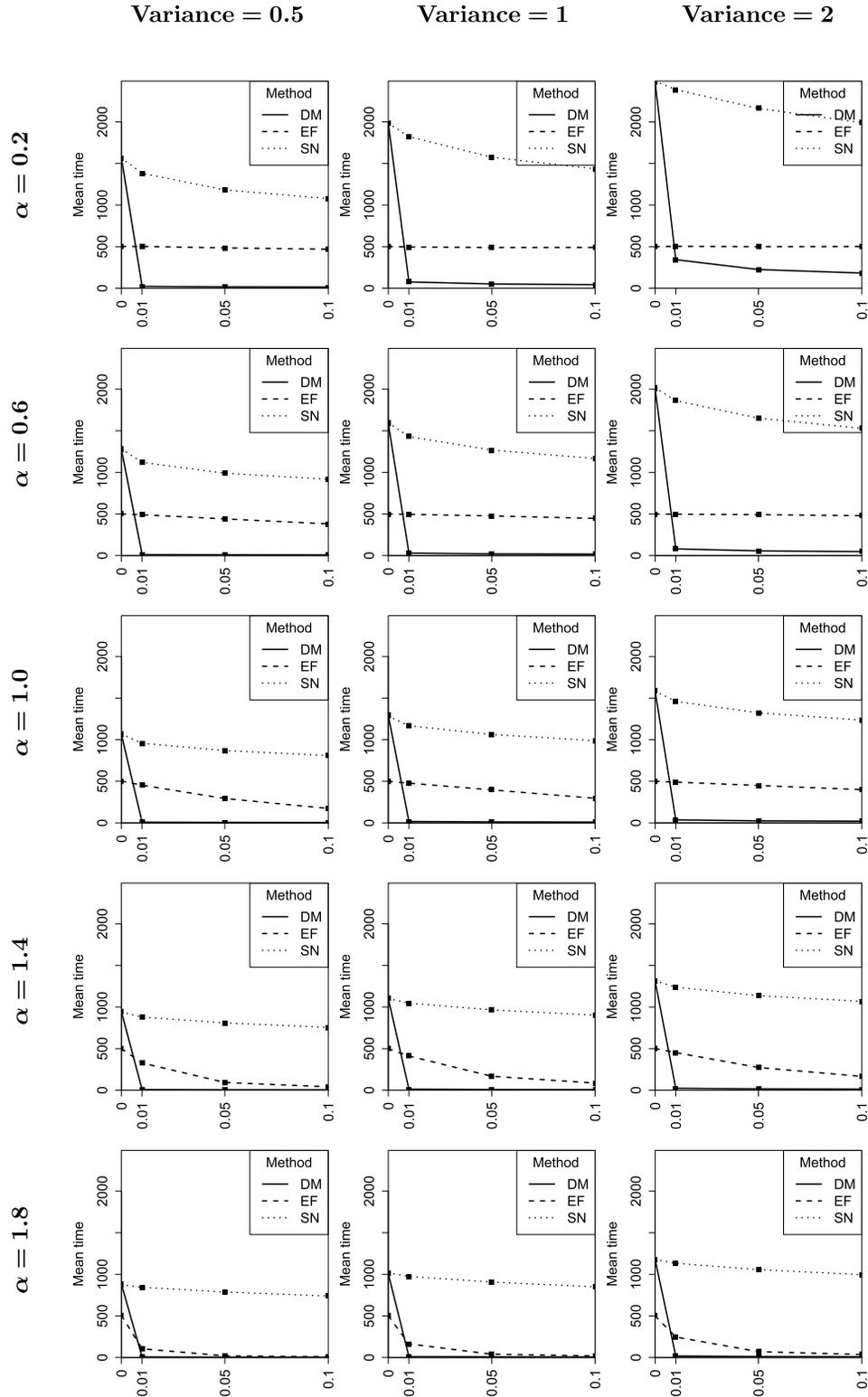


FIG. 7. *Brown–Resnick process simulation in 15 scenarios: Displayed are the mean times of three generic simulation algorithms (DM/EF/SN) for a given tolerated error probability ranging from 0 (“exact simulation”) to 0.1; see Section 6 and Figure 8 for further details.*

processes’ roughness increases due to smaller α . In the extremal- t case, the mean time increases as the scale gets smaller and as the degrees of freedom parameter ν increases. However, our main interest lies in the relative

performance of the three algorithms DM, EF and SN. And while for exact simulation ($\mathcal{P} = 0$), the theoretical dominance of the EF approach can be confirmed, the DM approach seems to be uniformly best once we allow for a

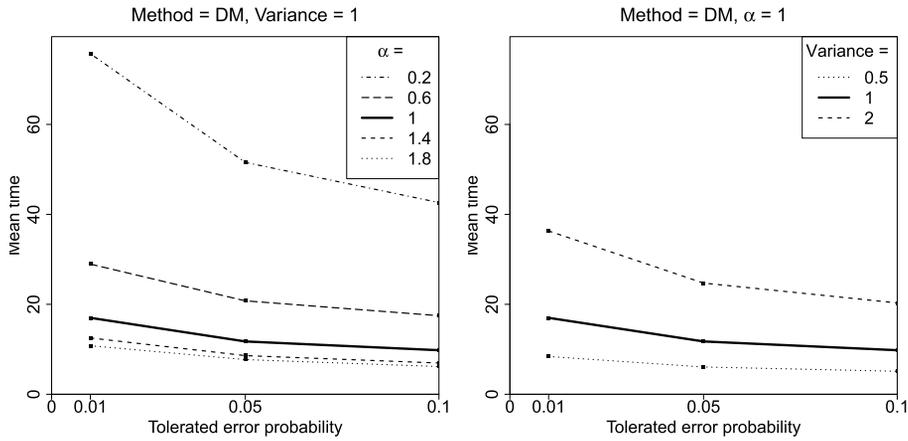


FIG. 8. A closer look at the mean times of the Dieker–Mikosch (DM) algorithm for Brown–Resnick process simulation. The plots complement Figure 7 revealing the scale and variability of the mean times of the DM algorithm, which may seem reduced to zero in Figure 7.

small tolerable error probability $\mathcal{P} \geq 0.01$. We anticipate a critical value $\mathcal{P}_{\text{critical}}$ closer to zero for the tolerable error probability at which the EF approach will start to perform better than the DM algorithm. The SN approach—

in the form considered here—cannot match up with either the EF or DM algorithm, chiefly because sampling from the sup-normalized spectral process is costly for Brown–Resnick and extremal- t processes. However, in Section 7

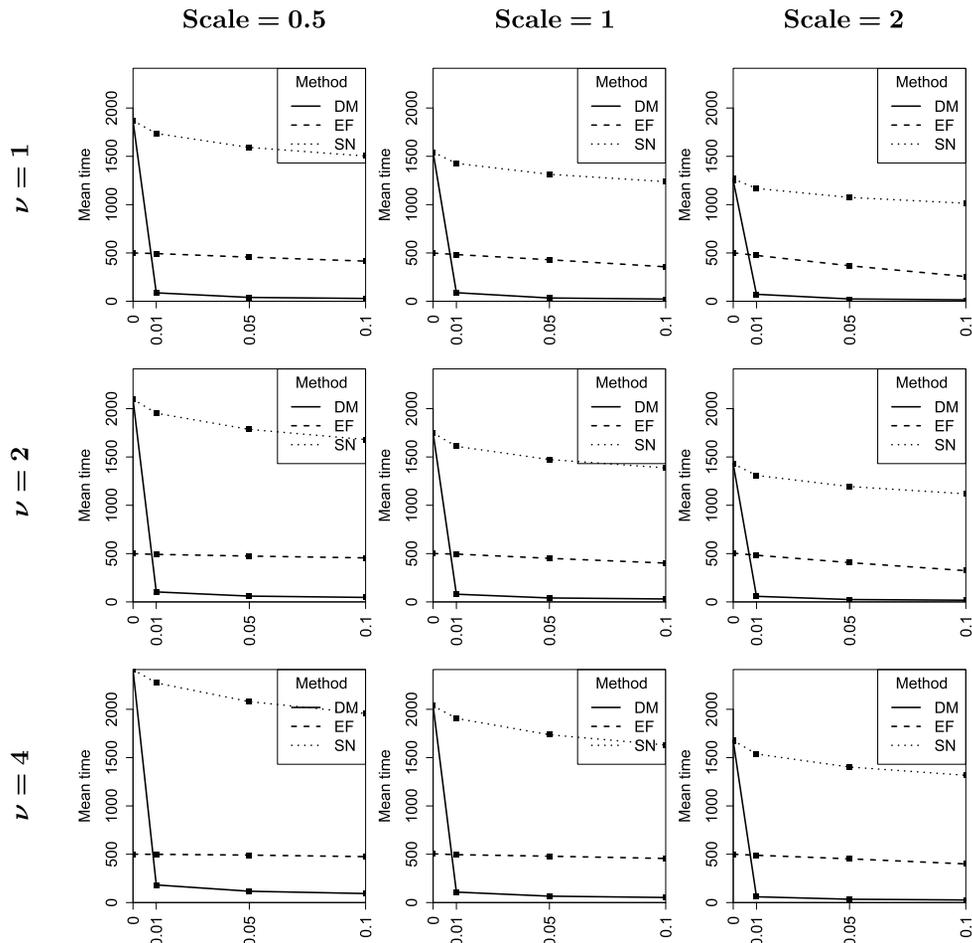


FIG. 9. Extremal- t -process simulation in 9 different parameter scenarios: Displayed are the mean times of three generic simulation algorithms (DM/EF/SN) for a given tolerated error probability ranging from 0 (“exact simulation”) to 0.1; see Section 6 for further details.

we will point the reader to modifications of the SN approach and situations, in which it can be very valuable again.

7. DISCUSSION

The simulation of max-stable processes has become an important task as part of spatial risk assessment, specifically in the environmental sciences. The last decade saw several new approaches to the simulation of max-stable processes. The present article provides an overview and compares the generic approaches according to their efficiency in relation to their accuracy. Moving from accurate simulation to tolerating small errors is a major issue of practical concern due to the inherently large computational costs for simulating a max-stable process. We contextualize known theoretical results on the efficiency in the exact setting (cf. Tables 2 and 3), while adding some new point process based results on the efficiency and accuracy for the approximate setting (cf. Section 5). An at first sight surprising observation of our numerical study is that the Dieker–Mikosch (DM) approach—despite being uniformly worse than the extremal functions (EF) approach in the exact setting—significantly outperforms all generic approaches, once we allow for a small tolerable error probability. That said, this finding is in line with the computational results in Oesting and Strokorb (2018) and may be attributed to the DM approach’s probabilistic homogeneity of spectral functions. In other words, compared to other algorithms, the DM approach “converges” enormously fast to a stochastic process, which is an accurate sample of a max-stable process with very high probability. However, the algorithm fails to be certain and seeks this confirmation for a very long time.

Further, our numerical study might create the impression that the threshold stopping approach using sup-normalized spectral functions (SN) is not worth considering anymore. We would like to correct that impression by emphasizing that the success of this approach depends largely on the ability to efficiently simulate from the sup-normalized spectral process $V^{\|\cdot\|_\infty}$, which is a research question in its own and of independent interest in other contexts; see also Remark in Section 3.2. In fact, the motivation of de Fondeville and Davison (2018) for introducing the generic rejection sampling approach was not to use it for max-stable process simulation, but to obtain accurate samples from associated Pareto-processes to be readily available for threshold-based inference. Spatiotemporal threshold based inference on extremes is currently an active area of research. For Brown–Resnick processes or extremal- t processes sampling from $V^{\|\cdot\|_\infty}$ is hard and choosing a generic rejection sampling approach for this task is not particularly helpful, which explains the poor performance of the SN approach in our study. Alternatives include MCMC approaches (Oesting, Schlather

and Zhou, 2018, Oesting, Schlather and Schillings, 2019) and for the class of Brown–Resnick processes modified rejection sampling (Oesting, Schlather and Schillings, 2019) or using the ansatz of Ho and Dombry (2019). For other classes of max-stable processes, the SN approach may well be the most efficient way of exact simulation. For instance, Oesting, Schlather and Zhou (2018) show that the sup-normalized process $V^{\|\cdot\|_\infty}$ can be easily simulated for a broad subclass of mixed moving maxima processes including Gaussian extreme value models (Smith, 1990). Simulation procedures for several of these max-stable models are implemented in the R package `RandomFields` (Schlather et al., 2020).

We conclude with some practical advice. First of all, we recommend to use exact simulation of max-stable processes, whenever it is feasible. The EF algorithm is designed for this purpose and from our perspective it is the generic approach to use as long as the number of points N in the simulation domain does not get too large. Should exact simulation from the sup-normalized spectral process $V^{\|\cdot\|_\infty}$ not be too costly, for example, for mixed moving maxima processes, then the SN approach can be a worthwhile alternative. In view of the comparison in Table 2, it may even reduce the computational cost significantly, when a large number N of points in a fixed domain K is considered. For approximate simulation, the DM approach seems to perform best—at least we could not detect a single scenario during our extensive numerical studies in which this was not the case. Unfortunately, the efficiency comes at the price of not knowing when to stop the DM algorithm. We therefore recommend to employ at least additional checks to ensure the obtained sample exhibits reasonable characteristics. Alternatively, the SN approach also lends itself as an approximate approach by means of more efficient MCMC techniques to obtain samples from the sup-normalized spectral process $V^{\|\cdot\|_\infty}$ as discussed above.

Finally, we would like to mention that for specific classes of max-stable processes, such as Brown–Resnick processes, specific approaches tailored to this class, such as Liu et al. (2019), may be worth considering, even though meaningful comparisons in terms of efficiency and accuracy seem difficult to achieve (cf. Remark in Section 5.1), and it is unclear how an approximate version would look like in this case.

APPENDIX A: PROOFS

The proofs given in this section rely on the fact that the pairs $\{(U_i, V_i)\}_{i \in \mathbb{N}} = \{(\Gamma_i^{-1}, V_i)\}_{i \in \mathbb{N}}$ form a Poisson point process Π on $(0, \infty) \times C(K)$ with intensity $u^{-2} du \mathbb{P}\{V \in dv\}$. We make extensive use of the Slivnyak–Mecke equation (see, e.g., equation (4.1) in Møller (2003)), which we recall here for convenience for our situation. To this end,

let \mathcal{N} denote the set of locally finite simple counting measures $(0, \infty) \times C(K)$, whose σ -algebra is generated by evaluations on Borel subsets of $(0, \infty) \times C(K)$. With a slight (but common and convenient) abuse of notation by identifying simple counting measures with their induced sets, we have

$$(A.1) \quad \begin{aligned} & \mathbb{E} \left\{ \sum_{(u,v) \in \Pi} f((u,v), \Pi \setminus \{(u,v)\}) \right\} \\ &= \int_{(0,\infty) \times C(K)} \mathbb{E}_{\Pi} \{ f((u,v), \Pi) \} \\ & \quad \times u^{-2} du \mathbb{P}\{V \in dv\}. \end{aligned}$$

for any nonnegative measurable function $f : ((0, \infty) \times C(K)) \times \mathcal{N} \rightarrow [0, \infty)$.

PROOF OF PROPOSITION 5.1. By definition of T_{τ} , we have

$$\begin{aligned} & \mathbb{E}(T_{\tau}) \\ &= \mathbb{E} \left\{ (u,v) \in \Pi : u\tau > \inf_{x \in K} \bigvee_{(\tilde{u}, \tilde{v}) \in \Pi, \tilde{u} > u} \tilde{u}\tilde{v}(\mathbf{x}) \right\} \\ &\geq \mathbb{E} \left\{ (u,v) \in \Pi : u > \inf_{x \in K} \frac{\bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u,v)\}} \tilde{u}\tilde{v}(\mathbf{x})}{\tau} \right\}, \end{aligned}$$

with equality if and only if $\sup_{x \in K} V(\mathbf{x}) \leq \tau$ almost surely. Then the Slivnyak–Mecke equation (A.1) can be applied to the right-hand side to obtain

$$\begin{aligned} & \mathbb{E} \left\{ (u,v) \in \Pi : u > \inf_{x \in K} \frac{\bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u,v)\}} \tilde{u}\tilde{v}(\mathbf{x})}{\tau} \right\} \\ &= \int_{C(K)} \int_{C(K)} \int_0^{\infty} \mathbf{1}_{\{\inf_{x \in K} \frac{z(\mathbf{x})}{v(\mathbf{x})} < u\}} u^{-2} du \\ & \quad \times \mathbb{P}\{V \in dv\} \mathbb{P}\{Z \in dz\} \\ &= \mathbb{E} \left\{ \sup_{x \in K} \frac{\tau}{Z(\mathbf{x})} \right\} \end{aligned}$$

as desired. \square

PROOF OF PROPOSITION 5.4. Since any condition on $Z^{(T_{\tau})}$ can be rewritten in terms of the restricted point process $\Pi(\cdot \cap ((\tau^{-1} \inf_{x \in K} Z^{(T_{\tau})}(\mathbf{x}), \infty) \times C(K)))$, we have that, conditional on $Z^{(T_{\tau})}$, the restricted point process

$$\Pi(\cdot \cap ((0, \tau^{-1} \inf_{x \in K} Z^{(T_{\tau})}(\mathbf{x})) \times C(K)))$$

is a Poisson point process with intensity measure $u^{-2} du \mathbb{P}(V \in dv)$. Consequently,

$$\begin{aligned} & \mathbb{P}\{|Z(\mathbf{x}) - Z^{(T_{\tau})}(\mathbf{x})| \\ & \quad > f(Z^{(T_{\tau})}(\mathbf{x}), \mathbf{x}) \text{ for some } \mathbf{x} \in K \mid Z^{(T)}\} \\ &= \mathbb{P}\{\exists(u,v) \in \Pi : u\tau < \inf_{x \in K} Z^{(T_{\tau})}(\mathbf{x}), \end{aligned}$$

$$\begin{aligned} & uv(\mathbf{x}) > Z^{(T_{\tau})}(\mathbf{x}) + f(Z^{(T_{\tau})}(\mathbf{x}), \mathbf{x}) \text{ for some } \mathbf{x} \in K \} \\ &= 1 - \exp\left(-\mathbb{E}_V \left\{ \int_{\inf_{x \in K} \frac{Z^{(T_{\tau})}(\mathbf{x})}{V(\mathbf{x})}}^{\frac{Z^{(T_{\tau})}(\mathbf{x})}{\tau}} u^{-2} du \right\}\right) \\ &= 1 - \exp\left(-\mathbb{E}_V \left\{ \sup_{x \in K} \frac{V(\mathbf{x})}{Z^{(T_{\tau})}(\mathbf{x}) + f(Z^{(T_{\tau})}(\mathbf{x}), \mathbf{x})} \right. \right. \\ & \quad \left. \left. - \sup_{x \in K} \frac{\tau}{Z^{(T_{\tau})}(\mathbf{x})} \right\}_+\right). \end{aligned}$$

Taking the expectation with respect to $Z^{(T_{\tau})}$ completes the proof. \square

PROOF OF PROPOSITION 5.5. Let $\Pi_+ = \{(u,v) \in \Pi : uv(\cdot) \in \Phi_+\}$. Then we can rewrite

$$M_{\tau} = \left| \left\{ (u,v) \in \Pi_+ : u\tau \leq \inf_{x \in K} \bigvee_{(\tilde{u}, \tilde{v}) \in \Pi, \tilde{u} > u} \tilde{u}\tilde{v}(\mathbf{x}) \right\} \right|$$

and hence,

$$\begin{aligned} & \mathbb{E}(M_{\tau}) \\ &= \mathbb{E} \left\{ (u,v) \in \Pi : uv(\mathbf{x}) \right. \\ & \quad \left. > \bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u,v)\}} \tilde{u}\tilde{v}(\mathbf{x}) \text{ for some } \mathbf{x} \in K, \right. \\ & \quad \left. u\tau \leq \inf_{x \in K} \bigvee_{(\tilde{u}, \tilde{v}) \in \Pi, \tilde{u} > u} \tilde{u}\tilde{v}(\mathbf{x}) \right\} \\ &\leq \mathbb{E} \left\{ (u,v) \in \Pi : \inf_{x \in K} \frac{\bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u,v)\}} \tilde{u}\tilde{v}(\mathbf{x})}{v(\mathbf{x})} \right. \\ & \quad \left. < u \leq \inf_{x \in K} \frac{\bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u,v)\}} \tilde{u}\tilde{v}(\mathbf{x})}{\tau} \right\}. \end{aligned}$$

Applying the Slivnyak–Mecke formula (A.1) gives

$$\begin{aligned} \mathbb{E}(M_{\tau}) &\leq \int_{C(K)} \int_{C(K)} \int_0^{\infty} \mathbf{1}_{\{\inf_{x \in K} \frac{z(\mathbf{x})}{v(\mathbf{x})} < u \leq \inf_{x \in K} \frac{z(\mathbf{x})}{\tau}\}} \\ & \quad \times u^{-2} du \mathbb{P}\{V \in dv\} \mathbb{P}\{Z \in dz\}. \end{aligned}$$

which is equivalent to inequality (5.5). \square

PROOF OF PROPOSITION 5.6. Analogously to the proof of Proposition 5.4, we obtain

$$\begin{aligned} & \mathbb{P}\{|Z(\mathbf{x}) - Z^{(n)}(\mathbf{x})| \\ & \quad > f(Z^{(n)}(\mathbf{x}), \mathbf{x}) \text{ for some } \mathbf{x} \in K \mid Z^{(n)}\} \\ &= \mathbb{P}\{\exists(u,v) \in \Pi : uv(\mathbf{x}_i) < Z^{(n)}(\mathbf{x}_i), i = 1, \dots, n, \\ & \quad \times uv(\mathbf{x}) > Z^{(n)}(\mathbf{x}) + f(Z^{(n)}(\mathbf{x}), \mathbf{x}) \text{ for some } \mathbf{x} \in K \} \\ &= 1 - \exp\left(-\mathbb{E}_V \left\{ \int_{\inf_{x \in K} \frac{Z^{(n)}(\mathbf{x}_i)}{V(\mathbf{x}_i)}}^{\min_{i=1, \dots, n} \frac{Z^{(n)}(\mathbf{x}_i)}{V(\mathbf{x}_i)}} u^{-2} du \right\}\right) \end{aligned}$$

$$= 1 - \exp\left(-\mathbb{E}_V \left\{ \max_{i=1, \dots, n} \frac{V(\mathbf{x}_i)}{Z^{(n)}(\mathbf{x}_i)} - \sup_{\mathbf{x} \in K} \frac{V(\mathbf{x})}{Z^{(n)}(\mathbf{x}) + f(Z^{(n)}, \mathbf{x})} \right\}\right).$$

Taking the expectation with respect to $Z^{(n)}$ completes the proof. \square

PROOF OF PROPOSITION 5.7. The expected number of missing extremal functions after the n th step can be expressed as

$$(A.2) \quad \begin{aligned} & \mathbb{E}(M_+^{(n)}) \\ &= \mathbb{E} \left| \left\{ (u, v) \in \Pi_+ : uv(\mathbf{x}_i) \right. \right. \\ & \quad \left. \left. < \bigvee_{(\tilde{u}, \tilde{v}) \in \Pi} \tilde{u}\tilde{v}(x_i) \text{ for all } i = 1, \dots, n \right\} \right| \end{aligned}$$

Consequently, we have that $\mathbb{E}(M_+^{(n)})$ equals

$$\begin{aligned} & \mathbb{E} \left| \left\{ (u, v) \in \Pi : \inf_{\mathbf{x} \in K} \frac{\bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u, v)\}} \tilde{u}\tilde{v}(\mathbf{x})}{v(\mathbf{x})} \right. \right. \\ & \quad \left. \left. < u < \min_{i=1}^n \frac{\bigvee_{(\tilde{u}, \tilde{v}) \in \Pi \setminus \{(u, v)\}} \tilde{u}\tilde{v}(\mathbf{x}_i)}{v(\mathbf{x}_i)} \right\} \right|. \end{aligned}$$

The Slivnyak–Mecke equation (A.1) can be applied to the right-hand side. Hence, we obtain that $\mathbb{E}(M_+^{(n)})$ equals

$$\begin{aligned} & \int_{C(K)} \int_{C(K)} \int_0^\infty \mathbf{1}_{\{\inf_{\mathbf{x} \in K} \frac{z(\mathbf{x})}{v(\mathbf{x})} < u \leq \min_{i=1, \dots, n} \frac{z(\mathbf{x}_i)}{v(\mathbf{x}_i)}\}} u^{-2} du \\ & \quad \times \mathbb{P}\{V \in dv\} \mathbb{P}\{Z \in dz\}. \end{aligned}$$

The latter coincides with (5.6). \square

APPENDIX B: SIMULATION ASSESSMENT OF THE ACCURACY

This section is a step-by-step description how one can assess the approximation error of the Threshold Stopping Algorithm 1 with threshold τ by simulation.

(1) Use the Extremal Functions Algorithm 2 to simulate Φ_+ . Denote the resulting functions by $\varphi_1, \dots, \varphi_k$ and set $Z(\mathbf{x}) = \max_{i=1, \dots, k} \varphi_i(\mathbf{x})$, $\mathbf{x} \in K$.

Note that only the products of the type $\varphi = UV$ are obtained this way, the components U and V with $(U, V) \in \Pi_+$ (as in the proof of Proposition 5.5) are unknown.

(2) Simulate the set $\Pi_+ = \{(U_i^+, V_i^+)\}_{i=1, \dots, k}$ conditional on $U_i^+ \cdot V_i^+ = \varphi_i$ for $i = 1, \dots, k$.

Here, the U_i^+ are independent with density

$$f_{U_i^+}(u) \propto u^{-2} \mathbb{P}(V \in d\varphi_i/u)$$

and $V_i^+ = (U_i^+)^{-1} \varphi_i$.

(3) Simulate the entire set $\Pi_{\min, -} = \{(U_i^-, V_i^-)\}_{i=1, \dots, l}$ of all nonextremal functions with $U_i^- > U_{\min} := \min_{i=1, \dots, k} U_i^+$.

These form a Poisson point process with intensity $u^{-2} du \cdot \mathbb{P}(V \in dv)$ restricted to the set $\{(u, v) \in [U_{\min}, \infty) \times C(K) : u \cdot v(\mathbf{x}) < Z(\mathbf{x}) \text{ for all } \mathbf{x} \in K\}$.

(4) Merge Π_+ and $\Pi_{\min, -}$ to the set $\Pi_{\min} = \{(U_i, V_i)\}_{i=1, \dots, k+l}$ labeling the points in such a way that $U_i > U_{i+1}$.

(5) Set $Z^{(j)}(\mathbf{x}) = \max_{i=1, \dots, j} U_i \cdot V_i(\mathbf{x})$ and define

$$T := \min \left\{ j \in \{1, \dots, k+l-1\} : \right. \\ \left. U_{j+1} \tau < \inf_{\mathbf{x} \in K} Z^{(j)}(\mathbf{x}) \right\},$$

that is, T equals the stopping time T_τ provided that $T_\tau < k+l$. Otherwise, we have $T = \infty$.

(6) If $T = \infty$, the stopping criterion does not apply before all the extremal functions are simulated, that is, $Z^{(T)} = Z$ and there is no error. Otherwise, that is, if $T < k+l$, there is an error. This can either be measured in terms of the absolute/relative deviation between Z and $Z^{(T)}$ or in terms of the number of missing extremal functions, that is, the cardinality of the set $\{i \in \{1, \dots, k\} : U_i^+ < U_T\}$.

Repeating this procedure, the average error is an unbiased estimator of the expected simulation error of the Threshold Stopping Algorithm 1.

ACKNOWLEDGMENTS

The new theoretical results for this manuscript were obtained during mutual visits of KS at the University of Siegen and MO at Cardiff University. MO and KS thank their hosting institutions for their generous hospitality. The authors are also very grateful for the thoughtful suggestions from the reviewing process. In particular, these comments resulted in the clarification of the marginal standardization and the inclusion of Section 2. This substantial revision was undertaken during summer/autumn 2020. MO thankfully acknowledges financial support by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2075—390740016 at the University of Stuttgart.

REFERENCES

- BÜCHER, A. and SEGERS, J. (2017). On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes* **20** 839–872. MR3737387 <https://doi.org/10.1007/s10687-017-0292-6>
- CAPÉRAA, P., FOUGÈRES, A.-L. and GENEST, C. (1997). A non-parametric estimation procedure for bivariate extreme value copulas. *Biometrika* **84** 567–577. MR1603985 <https://doi.org/10.1093/biomet/84.3.567>

- CHANDLER, R. E. and BATE, S. (2007). Inference for clustered data using the independence loglikelihood. *Biometrika* **94** 167–183. MR2367830 <https://doi.org/10.1093/biomet/asm015>
- DAVISON, A. C. and HUSER, R. (2015). Statistics of extremes. *Annu. Rev. Stat. Appl.* **2** 203–235.
- DAVISON, A. C., HUSER, R. and THIBAUD, E. (2018). Spatial extremes with application to climate and environmental exposure. In *Handbook of Environmental and Ecological Statistics* (A. E. Gelfand, M. Fuentes, J. A. Hoeting and R. L. Smith, eds.) CRC Press, Boca Raton, FL.
- DAVISON, A. C., PADOAN, S. A. and RIBATET, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27** 161–186. MR2963980 <https://doi.org/10.1214/11-STS376>
- DE FONDEVILLE, R. and DAVISON, A. C. (2018). High-dimensional peaks-over-threshold inference. *Biometrika* **105** 575–592. MR3842886 <https://doi.org/10.1093/biomet/asy026>
- DE HAAN, L. (1978). A characterization of multidimensional extreme-value distributions. *Sankhyā Ser. A* **40** 85–88. MR0545467
- DE HAAN, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. MR0757776
- DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer, New York. MR2234156 <https://doi.org/10.1007/0-387-34471-3>
- DE HAAN, L. and LIN, T. (2001). On convergence toward an extreme value distribution in $C[0, 1]$. *Ann. Probab.* **29** 467–483. MR1825160 <https://doi.org/10.1214/aop/1008956340>
- DIEKER, A. B. and MIKOSCH, T. (2015). Exact simulation of Brown–Resnick random fields at a finite number of locations. *Extremes* **18** 301–314. MR3351818 <https://doi.org/10.1007/s10687-015-0214-4>
- DOMBRY, C., ENGELKE, S. and OESTING, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317. MR3509888 <https://doi.org/10.1093/biomet/asw008>
- DOMBRY, C. and EYI-MINKO, F. (2012). Strong mixing properties of max-infinitely divisible random fields. *Stochastic Process. Appl.* **122** 3790–3811. MR2965926 <https://doi.org/10.1016/j.spa.2012.06.013>
- DOMBRY, C. and EYI-MINKO, F. (2013). Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab.* **18** 7. MR3024101 <https://doi.org/10.1214/EJP.v18-1991>
- DOMBRY, C. and RIBATET, M. (2015). Functional regular variations, Pareto processes and peaks over threshold. *Stat. Interface* **8** 9–17. MR3320385 <https://doi.org/10.4310/SII.2015.v8.n1.a2>
- EINMAHL, J. H. J., KIRILIOUK, A., KRAJINA, A. and SEGERS, J. (2016). An M -estimator of spatial tail dependence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 275–298. MR3453656 <https://doi.org/10.1111/rssb.12114>
- ENGELKE, S., MALINOWSKI, A., KABLUCHKO, Z. and SCHLATHER, M. (2015). Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 239–265. MR3299407 <https://doi.org/10.1111/rssb.12074>
- FISHER, R. A. and TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Cambridge Philos. Soc.* **24** 180–190.
- FRÉCHET, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.* **6** 93–116.
- GILLELAND, E. and KATZ, R. W. (2016). extRemes 2.0: An extreme value analysis package in R. *J. Stat. Softw.* **72** 1–39.
- GINÉ, E., HAHN, M. G. and VATAN, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87** 139–165. MR1080487 <https://doi.org/10.1007/BF01198427>
- GNEDENKO, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math. (2)* **44** 423–453. MR0008655 <https://doi.org/10.2307/1968974>
- HIJMANS, R. J. (2019). geosphere: Spherical Trigonometry. R package version 1.5-10.
- HO, Z. W. O. and DOMBRY, C. (2019). Simple models for multivariate regular variation and the Hüsler–Reiss Pareto distribution. *J. Multivariate Anal.* **173** 525–550. MR3954072 <https://doi.org/10.1016/j.jmva.2019.04.008>
- HUSER, R., DAVISON, A. C. and GENTON, M. G. (2016). Likelihood estimators for multivariate extremes. *Extremes* **19** 79–103. MR3454032 <https://doi.org/10.1007/s10687-015-0230-4>
- JENKINSON, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **81** 158–171.
- KABLUCHKO, Z. (2011). Extremes of independent Gaussian processes. *Extremes* **14** 285–310. MR2824498 <https://doi.org/10.1007/s10687-010-0110-x>
- KABLUCHKO, Z., SCHLATHER, M. and DE HAAN, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440 <https://doi.org/10.1214/09-AOP455>
- KIRILIOUK, A. (2016). tailDepFun: Minimum Distance Estimation of Tail Dependence Models. R package version 1.0.0.
- LIU, Z., BLANCHET, J. H., DIEKER, A. B. and MIKOSCH, T. (2019). On logarithmically optimal exact simulation of max-stable and related random fields on a compact set. *Bernoulli* **25** 2949–2981. MR4003570 <https://doi.org/10.3150/18-BEJ1076>
- MØLLER, J., ed. (2003). *Spatial Statistics and Computational Methods. Lecture Notes in Statistics* **173**. Springer, New York. MR2001383 <https://doi.org/10.1007/978-0-387-21811-3>
- NORBERG, T. (1986). Random capacities and their distributions. *Probab. Theory Related Fields* **73** 281–297. MR0855227 <https://doi.org/10.1007/BF00339941>
- OESTING, M., KABLUCHKO, Z. and SCHLATHER, M. (2012). Simulation of Brown–Resnick processes. *Extremes* **15** 89–107. MR2891311 <https://doi.org/10.1007/s10687-011-0128-8>
- OESTING, M., RIBATET, M. and DOMBRY, C. (2016). Simulation of max-stable processes. In *Extreme Value Modeling and Risk Analysis* 195–214. CRC Press, Boca Raton, FL. MR3644314
- OESTING, M., SCHLATHER, M. and SCHILLINGS, C. (2019). Sampling sup-normalized spectral functions for Brown–Resnick processes. *Stat* **8** e228. MR3978404 <https://doi.org/10.1002/sta4.228>
- OESTING, M., SCHLATHER, M. and ZHOU, C. (2018). Exact and fast simulation of max-stable processes on a compact set using the normalized spectral representation. *Bernoulli* **24** 1497–1530. MR3706800 <https://doi.org/10.3150/16-BEJ905>
- OESTING, M. and STROKORB, K. (2018). Efficient simulation of Brown–Resnick processes based on variance reduction of Gaussian processes. *Adv. in Appl. Probab.* **50** 1155–1175. MR3881113 <https://doi.org/10.1017/apr.2018.54>
- OPITZ, T. (2013). Extremal t processes: Elliptical domain of attraction and a spectral representation. *J. Multivariate Anal.* **122** 409–413. MR3189331 <https://doi.org/10.1016/j.jmva.2013.08.008>
- PENROSE, M. D. (1992). Semi-min-stable processes. *Ann. Probab.* **20** 1450–1463. MR1175271
- R CORE TEAM (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. R version 3.6.3.
- RESNICK, S. I. and ROY, R. (1991). Random usc functions, max-stable processes and continuous choice. *Ann. Appl. Probab.* **1** 267–292. MR1102320
- ROWLINGSON, B. (2019). geonames: Interface to the “Geonames” Spatial Query Web Service. R package version 0.999.

- SCHLATHER, M. (2002). Models for stationary max-stable random fields. *Extremes* **5** 33–44. MR1947786 <https://doi.org/10.1023/A:1020977924878>
- SCHLATHER, M., MALINOWSKI, A., OESTING, M., BOECKER, D., STROKORB, K., ENGELKE, S., MARTINI, J., BALLANI, F., MOREVA, O. et al. (2020). RandomFields: Simulation and Analysis of Random Fields. R package version 3.3.8.
- SMITH, R. L. (1990). Max-stable processes and spatial extremes. Unpublished Manuscript.
- SOUTH, A. (2017). rnaturalearth: World Map Data from Natural Earth. R package version 0.1.0.
- STEPHENSON, A. G. (2002). evd: Extreme value distributions. *R News* **2** 0.
- STUBNER, R. (2019). dqrng: Fast Pseudo Random Number Generators. R package version 0.2.1.
- U.S./JAPAN ASTER SCIENCE TEAM ASTER Global Digital Elevation Model, V2.
- VON MISES, R. (1936). La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcanique* **1** 141–160. Reproduced in: *Selected papers of Richard von Mises*, Amer. Math. Soc., Vol 2, 271-294 (1964).