J. Appl. Prob. **54,** 797–810 (2017)
*doi:*10.1017/jpr.2017.35

ON GENERALIZED MAX-LINEAR MODELS IN MAX-STABLE RANDOM FIELDS

MICHAEL FALK * AND
MAXIMILIAN ZOTT,* ** *University of Würzburg*

Abstract

In practice, it is not possible to observe a whole max-stable random field. Therefore, we propose a method to reconstruct a max-stable random field in $C([0, 1]^k)$ by interpolating its realizations at finitely many points. The resulting interpolating process is again a max-stable random field. This approach uses a *generalized max-linear model*. Promising results have been established in the $k = 1$ case of Falk *et al.* (2015). However, the extension to higher dimensions is not straightforward since we lose the natural order of the index space.

Keywords: Multivariate extreme value distribution; max-stable random field; D-norm; max-linear model; stochastic interpolation

2010 Mathematics Subject Classification: Primary 60G70
Secondary 62M40

1. Introduction and preliminaries

Dombry *et al.* [10] derived an algorithm to sample from the regular conditional distribution of a max-stable random field $\boldsymbol{\eta}$, say, given the marginal observations $\eta_{s_1} = z_1, \dots, \eta_{s_d} = z_d$ for some $z_1, \dots, z_d$ from the state space and d locations $s_1, \dots, s_d$. Clearly, this concerns the *distribution* of $\boldsymbol{\eta}$ and the derived distributional parameters.

Different to that, we try to *reconstruct* $\boldsymbol{\eta}$ from the observations $\eta_{s_1}, \dots, \eta_{s_d}$. This is carried out by a *generalized max-linear model* in such a way that the interpolating process $\hat{\boldsymbol{\eta}}$ is again a (standard) max-stable random field.

As our approach is deterministic, once the observations $\eta_{s_1} = z_1, \dots, \eta_{s_d} = z_d$ are given, a proper way to measure the performance of our approach is the *mean squared error* (MSE). Convergence of the pointwise MSE as well as the integrated MSE (IMSE) is established if the set of grid points $s_1, \dots, s_d$ becomes dense in the index space.

A *max-stable random process* with index set T is a family of random variables $\boldsymbol{\xi} = (\xi_t)_{t\in T}$ with the property that there are functions $a_n\colon T \to \mathbb{R}_0^+$ and $b_n\colon T \to \mathbb{R}$, $n \in \mathbb{N}$, such that

$$\left(\max_{i=1,\dots,n}\left(\frac{\xi_t^{(i)} - b_n(t)}{a_n(t)}\right)\right)_{t\in T} \overset{\mathrm{D}}{=} \boldsymbol{\xi},$$

where $\boldsymbol{\xi}^{(i)} = (\xi_t^{(i)})_{t\in T}$, $i = 1, \dots, n$, are independent copies of $\boldsymbol{\xi}$ and '$\overset{\mathrm{D}}{=}$' denotes equality in distribution. We obtain a max-stable random vector (RV) on $\mathbb{R}^d$ by setting $T = \{1, \dots, d\}$. Different to that, we obtain a max-stable process with continuous sample paths on some compact metric space S, if we set $T = S$ and require that the sample paths $\boldsymbol{\xi}(\omega)\colon S \to \mathbb{R}$

Received 28 April 2016; revision received 23 February 2017.
* Postal address: Institute of Mathematics, University of Würzburg, Emil-Fischer-Str. 30, 97074 Würzburg, Germany.
** Email address: maximilian.zott@mailbox.org

realize in $C(S) = \{g \in \mathbb{R}^S\colon g \text{ continuous}\}$, and that the norming functions a_n and b_n are also continuous. Max-stable RVs and processes have been investigated intensely over the last decades. For detailed reviews of max-stable RVs and processes, see, for example, [2], [6] (discussed in [5], [13], [20], and [21]; rejoinder at [7]), [8], [12], and [19]. Max-stable RVs and processes are of enormous interest in extreme value theory since they are the only possible limit of linearly standardized maxima of independent and identically distributed RVs or processes.

Clearly, the univariate margins of a max-stable random process are max-stable distributions on the real line. A max-stable random object $\boldsymbol{\xi} = (\xi_t)_{t\in T}$ is commonly called *simple max-stable* in the literature if each univariate margin is unit Fréchet distributed, i.e. $\mathbb{P}(\xi_t \le x) = \exp(-x^{-1}),\ x > 0,\ t \in T$. Different to that, we call a random process $\boldsymbol{\eta} = (\eta_t)_{t\in T}$ *standard max-stable* if all univariate marginal distributions are standard negative exponential, i.e. $\mathbb{P}(\eta_t \le x) = \exp(x),\ x \le 0,\ t \in T$. The transformation to simple/standard margins does not cause any problems, neither in the case of RVs (see, for example, [9] or [19]), nor in the case of random fields with continuous sample paths (see, for example, [14]).

It is well known (see, for example, [9], [12], and [18]) that an RV $(\eta_1, \dots, \eta_d)$ is a *standard max-stable RV* if and only if there exists an RV $(Z_1, \dots, Z_d)$ and some number $c \ge 1$ with $Z_i \in [0, c]$ almost surely (a.s.) and $\mathbb{E}(Z_i) = 1,\ i = 1, \dots, d$, such that, for all $\boldsymbol{x} = (x_1, \dots, x_d) \le \boldsymbol{0} \in \mathbb{R}^d$,

$$\mathbb{P}(\eta_1 \le x_1, \dots, \eta_d \le x_d) = \exp(-\|\boldsymbol{x}\|_D) := \exp\Big(-\mathbb{E}\Big(\max_{i=1,\dots,d}(|x_i|Z_i)\Big)\Big).$$

The condition $Z_i \in [0, c]$ a.s. can be weakened to $\mathbb{P}(Z_i \ge 0) = 1$. Note that $\|\cdot\|_D$ defines a norm on $\mathbb{R}^d$, called the *D-norm* with *generator* $\boldsymbol{Z}$. The D means dependence: we have independence of the margins if and only if $\|\cdot\|_D$ is equal to the norm $\|\boldsymbol{x}\|_1 = \sum_{i=1}^d |x_i|$, which is generated by $(Z_1, \dots, Z_d)$ being a random permutation of the vector $(d, 0 \dots, 0)$. We have complete dependence of the margins if and only if $\|\cdot\|_D$ is the maximum-norm $\|\boldsymbol{x}\|_\infty = \max_{1\le i\le d}|x_i|$, which is generated by the constant vector $(Z_1, \dots, Z_d) = (1, \dots, 1)$. We refer the reader to [12, Section 4.4] for further details of D-norms.

Let S be a compact metric space. A standard max-stable process $\boldsymbol{\eta} = (\eta_t)_{t\in S}$ with sample paths in $\bar{C}^-(S) := \{g \in C(S)\colon g \le 0\}$ is, in what follows, called a *standard max-stable process* (SMSP). Denote further by $E(S)$ the set of those bounded functions $f \in \mathbb{R}^S$ that have only a finite number of discontinuities and define $\bar{E}^-(S) := \{f \in E(S)\colon f \le 0\}$. We know from [14] that a process $\boldsymbol{\eta} = (\eta_t)_{t\in S}$ with sample paths in $C(S)$ is an SMSP if and only if there exists a stochastic process $\boldsymbol{Z} = (Z_t)_{t\in S}$ realizing in $\bar{C}^+(S) := \{g \in C(S)\colon g \ge 0\}$ and some $c \ge 1$, such that $Z_t \le c$ a.s., $\mathbb{E}(Z_t) = 1,\ t \in S$, and

$$\mathbb{P}(\boldsymbol{\eta} \le f) = \exp(-\|f\|_D) := \exp\Big(-\mathbb{E}\Big(\sup_{t\in S}(|f(t)|Z_t)\Big)\Big), \qquad f \in \bar{E}^-(S).$$

Note that $\|\cdot\|_D$ defines a norm on the function space $E(S)$, again called the *D-norm* with *generator process* $\boldsymbol{Z}$. The functional D-norm is topologically equivalent to the sup-norm $\|f\|_\infty = \sup_{t\in S}|f(t)|$, which is itself a D-norm by setting $Z_t = 1,\ t \in S$; see [1] for details.

At first it might seem unusual to consider the function space $E(S)$. The reason for this is that a suitable choice of the function $f \in \bar{E}^-(S)$ allows the incorporation of the finite-dimensional marginal distributions by the relation $\mathbb{P}(\boldsymbol{\eta} \le f) = \mathbb{P}(\eta_{t_i} \le x_i,\ 1 \le i \le d)$.

The condition $\mathbb{P}(\sup_{t\in S} Z_t \le c) = 1$ can be weakened to

$$\mathbb{E}\Big(\sup_{t\in S} Z_t\Big) < \infty; \tag{1}$$

see [8, Corollary 9.4.5].

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

2. Generalized max-linear models

2.1. The model and some examples

In this section we will approximate a given SMSP with sample paths in $\bar{C}^-([0,1]^k)$, where k is some integer, by using a generalized max-linear model for the interpolation of a finite dimensional marginal distribution. The parameter space $[0,1]^k$ is chosen for convenience and could be replaced by any compact metric space S.

Let, in what follows, $\boldsymbol{\eta} = (\eta_t)_{t\in[0,1]^k}$ be an SMSP with generator $\boldsymbol{Z} = (Z_t)_{t\in[0,1]^k}$ and D-norm $\|\cdot\|_D$. Choose pairwise different points $s_1,\dots,s_d \in [0,1]^k$ and obtain a standard max-stable RV $(\eta_{s_1},\dots,\eta_{s_d})$ with generator $(Z_{s_1},\dots,Z_{s_d})$ and D-norm $\|\cdot\|_{D_{1,\dots,d}}$, i.e.

$$\mathbb{P}(\eta_{s_1} \le x_1,\dots,\eta_{s_d} \le x_d) = \exp\Big(-\mathbb{E}\Big(\max_{i=1,\dots,d}(|x_i|Z_{s_i})\Big)\Big) =: \exp(-\|\boldsymbol{x}\|_{D_{1,\dots,d}}),$$

$\boldsymbol{x} = (x_1,\dots,x_d) \le \boldsymbol{0}$. Our aim is to find another SMSP that interpolates the above RV.

Take functions $g_i \in \bar{C}^+([0,1]^k)$, $i = 1,\dots,d$, with the property

$$\|(g_1(t),\dots,g_d(t))\|_{D_{1,\dots,d}} = 1 \quad \text{for all } t \in [0,1]^k. \tag{2}$$

Then the stochastic process $\hat{\boldsymbol{\eta}} = (\hat{\eta}_t)_{t\in[0,1]^k}$ that is generated by the *generalized max-linear model*

$$\hat{\eta}_t := \max_{i=1,\dots,d} \frac{\eta_{s_i}}{g_i(t)}, \qquad t \in [0,1]^k, \tag{3}$$

defines an SMSP with generator

$$\hat{Z}_t = \max_{i=1,\dots,d}(g_i(t)Z_{s_i}), \qquad t \in [0,1]^k, \tag{4}$$

due to (2); see [11] for details. The $\|\cdot\|_{D_{1,\dots,d}} = \|\cdot\|_1$ case leads to the regular *max-linear model*; see [22].

If we want $\hat{\boldsymbol{\eta}}$ to interpolate $(\eta_{s_1},\dots,\eta_{s_d})$ then we only have to demand

$$g_i(s_j) = \delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \ne j, \end{cases} \qquad 1 \le i,\ j \le d. \tag{5}$$

Recall that η_{s_i} is negative with probability 1. We call $\hat{\boldsymbol{\eta}}$ the *discretized version* of $\boldsymbol{\eta}$ with grid $\{s_1,\dots,s_d\}$ and weight functions $g_1,\dots,g_d$, when the weight functions satisfy both (2) and (5).

Example 1. In the one-dimensional $k = 1$ case, the weight functions g_i can be chosen as follows. Take a grid $0 := s_1 < s_2 < \dots < s_{d-1} < s_d =: 1$ of the interval $[0,1]$ and denote by $\|\cdot\|_{D_{i-1,i}}$ the D-norm pertaining to $(\eta_{s_{i-1}},\eta_{s_i})$, $i = 2,\dots,d$. Set

$$g_1(t) := \begin{cases} \dfrac{s_2 - t}{\|(s_2 - t, t)\|_{D_{1,2}}}, & t \in [0, s_2], \\ 0 & \text{otherwise}, \end{cases}$$

$$g_i(t) := \begin{cases} \dfrac{t - s_{i-1}}{\|(s_i - t, t - s_{i-1})\|_{D_{i-1,i}}}, & t \in [s_{i-1}, s_i], \\ \dfrac{s_{i+1} - t}{\|(s_{i+1} - t, t - s_i)\|_{D_{i,i+1}}}, & t \in [s_i, s_{i+1}], \\ 0 & \text{otherwise}, \end{cases} \qquad i = 2,\dots,d-1,$$

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

and

$$g_d(t) := \begin{cases} \dfrac{t - s_{d-1}}{\|(s_d - t, t - s_{d-1})\|_{D_{d-1,d}}}, & t \in [s_{d-1}, 1], \\ 0 & \text{otherwise.} \end{cases}$$

This model has been studied intensely in [11]. The functions $g_1, \dots, g_d$ are continuous and satisfy conditions (2) and (5), so they provide an interpolating generalized max-linear model on $C[0, 1]$.

Example 2. Choose pairwise different points $s_1, \dots, s_d \in [0, 1]^k$ and an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^k$. Define

$$\tilde{g}_i(t) := \min_{j \neq i}(\|t - s_j\|), \qquad t \in [0, 1]^k, \ i = 1, \dots, d.$$

In order to normalize, set

$$g_i(t) := \frac{\tilde{g}_i(t)}{\|(\tilde{g}_1(t), \dots, \tilde{g}_d(t))\|_{D_{1,\dots,d}}}, \qquad t \in [0, 1]^k, \ i = 1, \dots, d.$$

The functions g_i are well defined since the denominator never vanishes. Suppose that there is $t \in [0, 1]^k$ with $\tilde{g}_1(t) = \cdots = \tilde{g}_d(t) = 0$. Then $\min_{j \neq i}(\|t - s_j\|) = 0$ for all $i = 1, \dots, d$. Now fix $i \in \{1, \dots, d\}$. There is $j \neq i$ with $t = s_j$. But on the other hand, we also have $\min_{k \neq j}(\|t - s_k\|) = 0$ which implies that there is $k \neq j$ with $t = s_k = s_j$, which is a contradiction.

The functions g_i, $i = 1, \dots, d$, are clearly functions in $\bar{C}^+([0, 1]^k)$ that also satisfy conditions (2) and (5) as can be seen as follows. We have, for $t \in [0, 1]^k$,

$$\begin{aligned} &\|(g_1(t), \dots, g_d(t))\|_{D_{1,\dots,d}} \\ &\quad = \left\| \left(\frac{\tilde{g}_1(t)}{\|(\tilde{g}_1(t), \dots, \tilde{g}_d(t))\|_{D_{1,\dots,d}}}, \dots, \frac{\tilde{g}_d(t)}{\|(\tilde{g}_1(t), \dots, \tilde{g}_d(t))\|_{D_{1,\dots,d}}} \right) \right\|_{D_{1,\dots,d}} \\ &\quad = \frac{\|(\tilde{g}_1(t), \dots, \tilde{g}_d(t))\|_{D_{1,\dots,d}}}{\|(\tilde{g}_1(t), \dots, \tilde{g}_d(t))\|_{D_{1,\dots,d}}} \\ &\quad = 1, \end{aligned}$$

which is condition (2). Note, moreover, that $\tilde{g}_i(s_j) = 0$ if $i \neq j$. But this implies condition (5):

$$\begin{aligned} g_i(s_j) &= \frac{\tilde{g}_i(s_j)}{\|(\tilde{g}_1(s_j), \dots, \tilde{g}_d(s_j))\|_{D_{1,\dots,d}}} \\ &= \frac{\tilde{g}_i(s_j)}{(0, \dots, 0, \tilde{g}_j(s_j), 0, \dots, 0)_{D_{1,\dots,d}}} \\ &= \frac{\tilde{g}_i(s_j)}{\tilde{g}_j(s_j)\|(0, \dots, 0, 1, 0, \dots, 0)\|_{D_{1,\dots,d}}} \\ &= \frac{\tilde{g}_i(s_j)}{\tilde{g}_j(s_j)} \\ &= \delta_{ij} \end{aligned}$$

by the fact that a D-norm of each unit vector in $\mathbb{R}^d$ is 1. Thus, we have found an interpolating generalized max-linear model on $C([0, 1]^k)$.

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

2.2. The MSE of the discretized version

We start this section with a result that applies to bivariate standard max-stable RVs in general.

Lemma 1. *Let* (X_1, X_2) *be standard max-stable with generator* (Z_1, Z_2) *and D-norm* $\|\cdot\|_D$.

(i) *We have*

$$\mathbb{E}(X_1 X_2) = \int_0^\infty \frac{1}{\|(1,u)\|_D^2}\,\mathrm{d}u,$$

(ii) $\mathbb{E}(|Z_1 - Z_2|) = 2(\|(1,1)\|_D - 1)$.

Proof. (i) See [11, Lemma 5].

(ii) The assertion follows from the general identity $\max(a,b) = \frac{1}{2}(a + b + |a-b|)$. □

Let $\hat{\boldsymbol{\eta}} = (\hat{\eta}_t)_{t\in[0,1]^k}$ be the discretized version of $\boldsymbol{\eta} = (\eta_t)_{t\in[0,1]^k}$ with grid $\{s_1, \dots, s_d\}$ and weight functions $g_1, \dots, g_d$. In order to calculate the MSE of $\hat{\eta}_t$, we need the following lemma.

Lemma 2. *Let* $\hat{\mathbf{Z}} = (\hat{Z}_t)_{t\in[0,1]^k}$ *be the generator of* $\hat{\boldsymbol{\eta}}$ *defined in (4). For each* $t \in [0,1]^k$, *the RV* $(\eta_t, \hat{\eta}_t)$ *is standard max-stable with generator* $(Z_t, \hat{Z}_t)$ *and D-norm*

$$\|(x,y)\|_{D_t} = \mathbb{E}(\max(|x|Z_t, |y|\hat{Z}_t)) = \|(x, g_1(t)y, \dots, g_d(t)y)\|_{D_{t,s_1,\dots,s_d}},$$

where $\|\cdot\|_{D_{t,s_1,\dots,s_d}}$ *is the D-norm pertaining to* $(\eta_t, \eta_{s_1}, \dots, \eta_{s_d})$.

Proof. As $\mathbf{Z} = (Z_t)_{t\in[0,1]^k}$ is a generator of $\boldsymbol{\eta}$, we have, for $x, y \le 0$,

$$\begin{aligned}\mathbb{P}(\eta_t \le x, \hat{\eta}_t \le y) &= \mathbb{P}(\eta_t \le x, \eta_{s_1} \le g_1(t)y, \dots, \eta_{s_d} \le g_d(t)y)\\ &= \exp(-\mathbb{E}(\max(|x|Z_t, |y|\max(g_1(t)Z_{s_1}, \dots, g_d(t)Z_{s_d}))))\\ &= \exp(-\mathbb{E}(\max(|x|Z_t, |y|\hat{Z}_t))).\end{aligned}$$

Then the assertion follows from the fact that $\hat{Z}_t \ge 0$ and $\mathbb{E}(\hat{Z}_t) = 1$. □

We can now use the preceding lemmas to compute the MSE.

Proposition 1. *The MSE of* $\hat{\eta}_t$ *is given by*

$$\mathrm{MSE}(\hat{\eta}_t) := \mathbb{E}((\eta_t - \hat{\eta}_t)^2) = 2\bigg(2 - \int_0^\infty \frac{1}{\|(1,u)\|_{D_t}^2}\,\mathrm{d}u\bigg), \qquad t \in [0,1]^k.$$

Proof. Due to Lemma 2, $(\eta_t, \hat{\eta}_t)$ is standard max-stable. Therefore, Lemma 1(i) and the fact that $\mathbb{E}(\eta_t) = \mathbb{E}(\hat{\eta}_t) = -1$ and $\mathrm{var}(\eta_t) = \mathrm{var}(\hat{\eta}_t) = 1$ yield

$$\mathrm{MSE}(\hat{\eta}_t) = \mathbb{E}(\eta_t^2) - 2\mathbb{E}(\eta_t\hat{\eta}_t) + \mathbb{E}(\hat{\eta}_t^2) = 4 - 2\int_0^\infty \frac{1}{\|(1,u)\|_{D_t}^2}\,\mathrm{d}u. \qquad □$$

Lemma 3. *The MSE of* $\hat{\eta}_t$ *satisfies* $\mathrm{MSE}(\hat{\eta}_t) \le 6\mathbb{E}(|Z_t - \hat{Z}_t|)$, $t \in [0,1]^k$.

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

Proof. We have

$$
\begin{aligned}
2 - \int_0^\infty & \frac{1}{\|(1,u)\|_{D_t}^2}\,\mathrm{d}u \\
&= \int_0^\infty \frac{1}{\|(1,u)\|_\infty^2}\,\mathrm{d}u - \int_0^\infty \frac{1}{\|(1,u)\|_{D_t}^2}\,\mathrm{d}u \\
&= \int_0^\infty (\|(1,u)\|_{D_t} - \|(1,u)\|_\infty)\frac{\|(1,u)\|_{D_t} + \|(1,u)\|_\infty}{\|(1,u)\|_{D_t}^2\|(1,u)\|_\infty^2}\,\mathrm{d}u \\
&= \int_0^1 (\|(1,u)\|_{D_t} - 1)\frac{\|(1,u)\|_{D_t} + 1}{\|(1,u)\|_{D_t}^2}\,\mathrm{d}u + \int_1^\infty (\|(1,u)\|_{D_t} - u)\frac{\|(1,u)\|_{D_t} + u}{u^2\|(1,u)\|_{D_t}^2}\,\mathrm{d}u \\
&\le 3\int_0^1 (\|(1,u)\|_{D_t} - 1)\,\mathrm{d}u + 2\int_1^\infty \frac{\|(1/u,1)\|_{D_t} - 1}{u^2}\,\mathrm{d}u \\
&=: 3I_1 + 2I_2.
\end{aligned}
$$

Since every D-norm is monotone, we have

$$
\|(1,u)\|_{D_t} \le \|(1,1)\|_{D_t}, \qquad u \in [0,1], \qquad \left\|\left(\frac{1}{u},1\right)\right\|_{D_t} \le \|(1,1)\|_{D_t}, \qquad u > 1,
$$

and, thus, by Lemma 1(ii),

$$
I_1 + I_2 \le \|(1,1)\|_{D_t} - 1 + (\|(1,1)\|_{D_t} - 1)\int_1^\infty u^{-2}\,\mathrm{d}u = \mathbb{E}(|Z_t - \hat{Z}_t|). \qquad \square
$$

Remark 1. The upper bound $\mathbb{E}(|Z_t - \hat{Z}_t|)$ in Lemma 3 becomes small if the distance between t and its nearest neighbor s_j, say, in the grid $\{s_1, \dots, s_d\}$ becomes small, which can be seen as follows. The triangle inequality implies that

$$
|Z_t - \hat{Z}_t| \le |Z_t - Z_{s_j}| + \Big|Z_{s_j} - \max_{i=1,\dots,d}(g_i(t)Z_{s_i})\Big|.
$$

From the condition $g_i(s_j) = \delta_{ij}$, we obtain the representation

$$
Z_{s_j} = \max_{i=1,\dots,d}(g_i(s_j)Z_{s_i})
$$

and, thus,

$$
\begin{aligned}
\Big|Z_{s_j} - \max_{i=1,\dots,d}(g_i(t)Z_{s_i})\Big| &= \Big|\max_{i=1,\dots,d}(g_i(s_j)Z_{s_i}) - \max_{i=1,\dots,d}(g_i(t)Z_{s_i})\Big| \\
&\le \max_{i=1,\dots,d}(|g_i(t) - g_i(s_j)|Z_{s_i})
\end{aligned}
$$

by elementary arguments. As a consequence, we obtain

$$
\begin{aligned}
\mathbb{E}(|Z_t - \hat{Z}_t|) &\le \mathbb{E}(|Z_t - Z_{s_j}|) + \mathbb{E}\Big(\max_{i=1,\dots,d}(|g_i(t) - g_i(s_j)|Z_{s_i})\Big) \\
&= \mathbb{E}(|Z_t - Z_{s_j}|) + \|(|g_1(t) - g_1(s_j)|, \dots, |g_d(t) - g_d(s_j)|)\|_{D_{1,\dots,d}} \\
&\le \mathbb{E}(|Z_t - Z_{s_j}|) + \max_{i=1,\dots,d}|g_i(t) - g_i(s_j)|\|(1,\dots,1)\|_{D_{1,\dots,d}} \\
&\to 0, \qquad |t - s_j| \to 0,
\end{aligned}
$$

by the fact that each D-norm $\|\cdot\|_D$ is monotone, i.e. $\|\boldsymbol{x}\|_D \le \|\boldsymbol{y}\|_D$ if $\boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{y} \in \mathbb{R}^d$, and by the continuity of the functions $g_1, \dots, g_d$ and $\boldsymbol{Z}$.

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

Example 3. Choose as a generator process $\boldsymbol{Z} = (Z_t)_{t\in[0,1]^k}$ of a D-norm

$$Z_t := \exp\Big(X_t - \frac{\sigma^2(t)}{2}\Big), \qquad t \in [0,1]^k,$$

where $(X_t)_{t\in\mathbb{R}^k}$ is a continuous zero mean Gaussian process with stationary increments $\sigma^2(t) := \mathbb{E}(X_t^2)$ and $X_0 = 0$. This model was originally created by Brown and Resnick [4], and developed by Kabluchko *et al.* [17] for max-stable random fields $\boldsymbol{\vartheta} = (\vartheta_t)_{t\in[0,1]^k}$ with Gumbel margins, i.e. $\mathbb{P}(\vartheta_t \le x) = \exp(-\mathrm{e}^{-x})$, $x \in \mathbb{R}$. The transformation to an SMSP $(\eta_t)_{t\in[0,1]^k}$ is straightforward by setting $\eta_t := -\exp(-\vartheta_t)$, $t \in [0,1]^k$.

Explicit formulae for the corresponding D-norm

$$\|f\|_D = \mathbb{E}\Big(\sup_{t\in[0,1]^k} (|f(t)|Z_t)\Big), \qquad f \in E([0,1]^k),$$

are only available for bivariate $\|\cdot\|_{D_{t_1,t_2}}$ and trivariate $\|\cdot\|_{D_{t_1,t_2,t_3}}$ D-norms pertaining to the RVs (η_{t_1}, η_{t_2}) and $(\eta_{t_1}, \eta_{t_2}, \eta_{t_3})$, respectively; see [15]. In the bivariate case, we have, for $(x_1, x_2) \in \mathbb{R}^2$,

$$\begin{aligned}\|(x_1,x_2)\|_{D_{t_1,t_2}} = {} & |x_1|\Phi\bigg(\frac{\sigma(|t_1-t_2|)}{2} + \frac{1}{\sigma(|t_1-t_2|)}\log\Big(\frac{|x_1|}{|x_2|}\Big)\bigg) \\ & + |x_2|\Phi\bigg(\frac{\sigma(|t_1-t_2|)}{2} + \frac{1}{\sigma(|t_1-t_2|)}\log\Big(\frac{|x_2|}{|x_1|}\Big)\bigg),\end{aligned}$$

where Φ denotes the standard normal distribution function and the absolute value $|t_1 - t_2|$ is meant componentwise; see [16, Remark 24].

This Brown–Resnick model could, in particular, be used for the generalized max-linear model in dimension $k = 1$ as in Example 1, since in this case the approximation $\hat{\boldsymbol{\eta}}$ of $\boldsymbol{\eta}$ only uses bivariate D-norms $\|\cdot\|_{t_1,t_2}$.

3. A generalized max-linear model based on kernels

3.1. The model

There is the need for the definition of d functions $g_1, \ldots, g_d$ satisfying certain constraints in the *ordinary* generalized max-linear model with $d = d(n)$ tending to ∞ as the grid $s_1, \ldots, s_d$ becomes dense in the index set. For the kernel approach introduced in this section, this is reduced to the choice of just one kernel and a bandwidth. And in this case we can establish convergence to 0 of MSE and IMSE as the grid becomes dense, essentially without further conditions. This approach was briefly mentioned in [11] and is evaluated here.

There are disadvantages. The interpolation is not an exact one at the grid points, i.e. $\hat{\eta}_{s_j} \neq \eta_{s_j}$. This is due to the fact that the generated functions do not satisfy the condition $g_i(s_j) = \delta_{ij}$ exactly, but only in the limit as h tends to 0; see Lemma 4. The choice of an optimal bandwidth, which is statistical folklore in kernel density estimation, is still an open problem here.

Again, throughout this section, let $\boldsymbol{\eta} = (\eta_t)_{t\in[0,1]^k}$ be an SMSP with generator $\boldsymbol{Z} = (Z_t)_{t\in[0,1]^k}$ and denote by $\|\cdot\|_{s_1,\ldots,s_d}$ the D-norm pertaining to $(\eta_{s_1}, \ldots, \eta_{s_d})$.

Let $K\colon [0,\infty) \to [0,1]$ be a continuous and strictly monotonically decreasing function (kernel) with the two properties

$$K(0) = 1, \qquad \lim_{x\to\infty} \frac{K(ax)}{K(bx)} = 0, \quad 0 \le b < a. \tag{6}$$

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

The exponential kernel $K_e(x) = \exp(-x)$, $x \geq 0$, is a typical example. Choose an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^k$ and a grid of pairwise different points $\{s_1, \dots, s_d\}$ in $[0, 1]^k$. Set, for $i = 1, \dots, d$ and the bandwidth $h > 0$,

$$g_{i,h}(t) := \frac{K(\|t - s_i\|/h)}{\|(K(\|t - s_1\|/h), \dots, K(\|t - s_d\|/h))\|_{D_{s_1,\dots,s_d}}}, \qquad t \in [0, 1]^k.$$

Define, for $i = 1, \dots, d$,

$$N(s_i) := \{t \in [0, 1]^k : \|t - s_i\| \leq \|t - s_j\|,\ j \neq i\}, \tag{7}$$

which is the set of those points $t \in [0, 1]^k$ that are closest to the grid point s_i.

Lemma 4. *We have, for arbitrary $t \in [0, 1]^k$ and $1 \leq i \leq d$,*

$$g_{i,h}(t) \to \begin{cases} 1 & \text{if } t = s_i, \\ 0 & \text{if } t \notin N(s_i), \end{cases} \qquad h \downarrow 0,$$

as well as $g_{i,h}(t) \leq 1$.

Proof. The convergence $g_{i,h}(s_i) \to 1$, $h \downarrow 0$, follows from the fact that $K(0) = 1$ and that the D-norm of a unit vector is 1. The fact that an arbitrary D-norm is bounded below by the sup-norm together with the monotonicity of K implies that for $t \in [0, 1]^k$,

$$g_{i,h}(t) \leq \frac{K(\|t - s_i\|/h)}{\max_{1\leq j\leq d} K(\|t - s_j\|/h)} = \frac{K(\|t - s_i\|/h)}{K(\min_{1\leq j\leq d} \|t - s_j\|/h)} \leq 1.$$

Note that $K(\|t - s_i\|/h)/K(\min_{1\leq j\leq d} \|t - s_j\|/h) \to 0$, $h \downarrow 0$, if $t \notin N(s_i)$ by the required growth condition on the kernel K in (6). □

From the above lemma we see that, in particular, $g_{i,h}(s_j) \to \delta_{ij}$, $h \downarrow 0$, which is close to condition (5). Obviously, the functions $g_{i,h}$ are constructed in such a way that condition (2) holds exactly. Therefore, we obtain the generalized max-linear model

$$\hat{\eta}_{t,h} = \max_{i=1,\dots,d} \frac{\eta_{s_i}}{g_{i,h}(t)}, \qquad t \in [0, 1]^k,$$

which does not interpolate $(\eta_{s_1}, \dots, \eta_{s_d})$ exactly, but $\hat{\eta}_{s_i,h}$ converges to η_{s_i} as $h \downarrow 0$. Note that the limit functions $\lim_{h\downarrow 0} g_{i,h}$ are not necessarily continuous: For instance, there may be $t_0 \in [0, 1]^k$ with $\|t_0 - s_1\| = \cdots = \|t_0 - s_d\|$. Then $t_0 \in \partial N(s_1)$ and $\lim_{h\downarrow 0} g_{1,h}(t_0) = 1/\|(1, \dots, 1)\|_{D_{1,\dots,d}}$, but $\lim_{h\downarrow 0} g_{1,h}(t) = 0$ for all $t \notin N(s_1)$ due to Lemma 4.

3.2. Convergence of the MSE

In this section we investigate a sequence of kernel-based generalized max-linear models, where the diameter of the grids decreases. We analyze under which conditions the IMSE of $(\hat{\eta}_{t,h})_{t\in[0,1]^k}$ converges to 0. We start with a general result on generator processes.

Lemma 5. *Let $(Z_t)_{t\in[0,1]^k}$ be a generator of an SMSP and ε_n, $n \in \mathbb{N}$, be a null sequence. Then*

$$\mathbb{E}\Big(\sup_{\|t-s\|\leq\varepsilon_n} |Z_t - Z_s|\Big) \to 0, \qquad n \to \infty,$$

where $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^k$.

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

Proof. The paths of $(Z_t)_{t\in[0,1]^k}$ are continuous, so they are also uniformly continuous. Therefore, $\sup_{\|t-s\|\le\varepsilon_n}|Z_t-Z_s|\to 0$, $n\to\infty$. Furthermore,

$$\sup_{\|t-s\|\le\varepsilon_n}|Z_t-Z_s|\le 2\sup_{t\in[0,1]^k}Z_t$$

with $\mathbb{E}(\sup_{t\in[0,1]^k}Z_t)<\infty$ due to property (1) of a generator. The assertion now follows from the dominated convergence theorem. □

Let $\mathcal{G}_n:=\{s_{1,n},\dots,s_{d(n),n}\}$, $n\in\mathbb{N}$, be a set of distinct points in $[0,1]^k$ with the property

$$\text{for all } n\in\mathbb{N}, \text{ for all } t\in[0,1]^k, \text{ there exists } s_{i,n}\in\mathcal{G}_n\colon \|t-s_{i,n}\|\le\varepsilon_n,$$

where $\varepsilon_n\to 0$, $n\to\infty$. Define, for instance, $\mathcal{G}_n$ in such a way that

$$\varepsilon_n:=\max_{i=1,\dots,d}\sup_{s,t\in N(s_{i,n})}\|s-t\|\to 0,\qquad n\to\infty,$$

with $N(s_{i,n})$ as defined in (7). Clearly, $d:=d(n)\to\infty$, $n\to\infty$. Denote by $\|\cdot\|_{D^{(n)}_{s_1,\dots,s_d}}$ the D-norm pertaining to $\eta_{s_{1,n}},\dots,\eta_{s_{d,n}}$. Furthermore, let $\hat{\boldsymbol{\eta}}_n=(\hat\eta_{t,n})_{t\in[0,1]^k}$ be the kernel-based discretized version of $\boldsymbol{\eta}$ with grid $\mathcal{G}_n$, that is,

$$\hat\eta_{t,n}=\max_{i=1,\dots,d}\frac{\eta_{s_{i,n}}}{g_{i,n}(t)},\qquad t\in[0,1]^k,$$

where, for $i=1,\dots,d$,

$$g_{i,n}(t)=\frac{K(\|t-s_{i,n}\|/h_n)}{\|(K(\|t-s_{1,n}\|/h_n),\dots,K(\|t-s_{d,n}\|/h_n))\|_{D^{(n)}_{s_1,\dots,s_d}}},\qquad t\in[0,1]^k,$$

with $K\colon[0,\infty)\to[0,1]$, is the continuous and strictly decreasing kernel function satisfying condition (6), and h_n, $n\in\mathbb{N}$, is some positive sequence. We have already seen in Lemma 4 that $g_{i,n}(t)\in[0,1]$, $t\in[0,1]^k$, $n\in\mathbb{N}$. Furthermore, we have the following result.

Lemma 6. *Choose $t\in[0,1]^k$. There is a sequence $i(n)$, $n\in\mathbb{N}$, such that $t\in\bigcap_{n\in\mathbb{N}}N(s_{i(n),n})$. Define $g_{i(n),n}$ and ε_n as above for $n\in\mathbb{N}$. Then*

$$\lim_{n\to\infty}g_{i(n),n}(t)=1\quad \textit{if } \varepsilon_n\to 0,\ h_n\to 0,\ \varepsilon_n/h_n\to\infty,\ n\to\infty.$$

Proof. Let $t\in[0,1]^k$ and choose a sequence $i(n)$, $n\in\mathbb{N}$, as above. Set, for simplicity, $s_{i(n),n}=:s_{i,n}$ and $g_{i(n),n}=:g_{i,n}$. We have

$$\begin{aligned}
1&\ge g_{i,n}(t)\\
&=\frac{K(\|t-s_{i,n}\|/h_n)}{\mathbb{E}(\max_{j=1,\dots,d}K(\|t-s_{j,n}\|/h_n)Z_{s_{j,n}})}\\
&\ge\bigg(\frac{\mathbb{E}(\max_{j\colon\|s_{j,n}-t\|\ge 2\varepsilon_n}K(\|t-s_{j,n}\|/h_n)Z_{s_{j,n}})}{K(\|t-s_{i,n}\|/h_n)}\\
&\qquad+\frac{\mathbb{E}(\max_{j\colon\|s_{j,n}-t\|<2\varepsilon_n}K(\|t-s_{j,n}\|/h_n)Z_{s_{j,n}})}{K(\|t-s_{i,n}\|/h_n)}\bigg)^{-1}\\
&=:(A_{i,n}(t)+B_{i,n}(t))^{-1}.
\end{aligned}$$

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

From $t \in N(s_{i,n})$, we conclude that $\|t - s_{i,n}\| \le \varepsilon_n$. Hence, we have, due to (1) and the properties of the kernel function K,

$$0 \le A_{i,n}(t) \le \frac{K(2\varepsilon_n/h_n)}{K(\varepsilon_n/h_n)} \mathbb{E}\Big(\sup_{t\in[0,1]^k} Z_t\Big) \to 0, \qquad n \to \infty,$$

since $\varepsilon_n/h_n \to \infty$, $n \to \infty$, by assumption. Furthermore, $t \in N(s_{i,n})$ and the fact that K is decreasing implies that

$$\max_{j:\ \|s_{j,n}-t\|<2\varepsilon_n} K\bigg(\frac{\|t - s_{j,n}\|}{h_n}\bigg) = K\bigg(\frac{\|t - s_{i,n}\|}{h_n}\bigg).$$

Thus,

$$\begin{aligned}
1 &\le B_{i,n}(t) \\
&= \frac{1}{K(\|t - s_{i,n}\|/h_n)} \bigg(\mathbb{E}\bigg(\max_{j:\ \|s_{j,n}-t\|<2\varepsilon_n} K\bigg(\frac{\|t - s_{j,n}\|}{h_n}\bigg) Z_{s_{j,n}} \\
&\qquad\qquad - \max_{j:\ \|s_{j,n}-t\|<2\varepsilon_n} K\bigg(\frac{\|t - s_{j,n}\|}{h_n}\bigg) Z_{s_{i,n}}\bigg)\bigg) + 1 \\
&\le \frac{\mathbb{E}(\max_{j:\ \|s_{j,n}-t\|<2\varepsilon_n} K(\|t - s_{j,n}\|/h_n)|Z_{s_{j,n}} - Z_{s_{i,n}}|)}{K(\|t - s_{i,n}\|/h_n)} + 1 \\
&\le \mathbb{E}\Big(\max_{j:\ \|s_{j,n}-t\|<2\varepsilon_n} |Z_{s_{j,n}} - Z_{s_{i,n}}|\Big) + 1 \\
&\le \mathbb{E}\Big(\sup_{\|r-s\|<3\varepsilon_n} |Z_r - Z_s|\Big) + 1 \\
&\to 1, \qquad n \to \infty,
\end{aligned}$$

due to Lemma 5. Note that $\|s_{j,n} - t\| < 2\varepsilon_n$ and $t \in N(s_{i,n})$ imply that $\|s_{j,n} - s_{i,n}\| < 3\varepsilon_n$. This completes the proof. □

We have now gathered the tools to prove convergence of the MSE to 0.

Theorem 1. *Define $\hat{\boldsymbol{\eta}}_n$ and ε_n as above for $n \in \mathbb{N}$. Then, for every $t \in [0, 1]^k$,*

$$\mathrm{MSE}(\hat{\eta}_{t,n}) \to 0, \qquad \mathrm{IMSE}(\hat{\eta}_{t,n}) := \int_{[0,1]^k} \mathrm{MSE}(\hat{\eta}_{t,n})\, dt \to 0, \qquad n \to \infty,$$

if $\varepsilon_n \to 0$, $h_n \to 0$, $\varepsilon_n/h_n \to \infty$, $n \to \infty$.

Proof. Denote by

$$\hat{Z}_{t,n} = \max_{j=1,\dots,d} (g_{j,n}(t) Z_{s_{j,n}}), \qquad t \in [0, 1]^k,$$

the generator of $\hat{\boldsymbol{\eta}}_n$. Choose $t \in [0, 1]^k$ and a sequence $i := i(n)$, $n \in \mathbb{N}$, such that $t \in \bigcap_{n\in\mathbb{N}} N(s_{i,n})$. We have, by Lemma 3, Lemma 6, and the continuity of $\boldsymbol{Z}$,

$$\begin{aligned}
\mathrm{MSE}(\hat{\eta}_{t,n}) &\le 6\mathbb{E}(|Z_t - \hat{Z}_{t,n}|) \\
&\le 6\mathbb{E}(|Z_t - Z_{s_{i,n}}|) + 6\mathbb{E}(|Z_{s_{i,n}} - g_{i,n}(t) Z_{s_{i,n}}|) + 6\mathbb{E}(|g_{i,n}(t) Z_{s_{i,n}} - \hat{Z}_{t,n}|) \\
&= 6\mathbb{E}(|Z_t - Z_{s_{i,n}}|) + 12(1 - g_{i,n}(t)) \\
&\to 0, \qquad n \to \infty;
\end{aligned}$$

recall that $g_{i,n}(t) Z_{s_{i,n}} \le \hat{Z}_{t,n}$.

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

Next we establish convergence of the IMSE. The sets $N(s_{i,n})$, as defined in (7), are typically not disjoint, but the intersections $N(s_{i,n}) \cap N(s_{j,n})$, $i \neq j$, have Lebesgue measure 0 on $[0, 1]^k$. Clearly, $\bigcup_{i=1}^d N(s_{i,n}) = [0, 1]^k$. Therefore, applying Lemma 3 yields

$$\begin{aligned}
\mathrm{IMSE}(\hat{\eta}_{t,n}) &= \sum_{i=1}^d \int_{N(s_{i,n})} \mathrm{MSE}(\hat{\eta}_{t,n})\,\mathrm{d}t \\
&\leq 6 \sum_{i=1}^d \int_{N(s_{i,n})} \mathbb{E}(|Z_t - \hat{Z}_{t,n}|)\,\mathrm{d}t \\
&\leq 6 \bigg(\sum_{i=1}^d \int_{N(s_{i,n})} \mathbb{E}(|Z_t - Z_{s_i,n}|)\,\mathrm{d}t + \sum_{i=1}^d \int_{N(s_{i,n})} |1 - g_{i,n}(t)| \mathbb{E}(Z_{s_i,n})\,\mathrm{d}t \\
&\qquad + \sum_{i=1}^d \int_{N(s_{i,n})} \mathbb{E}(|g_{i,n}(t) Z_{s_i,n} - \hat{Z}_{t,n}|)\,\mathrm{d}t \bigg) \\
&=: 6(S_{1,n} + S_{2,n} + S_{3,n})
\end{aligned}$$

due to Lemma 3. From Lemma 5, we conclude that

$$\begin{aligned}
S_{1,n} &= \sum_{i=1}^d \int_{N(s_{i,n})} \mathbb{E}(|Z_t - Z_{s_{i,n}}|)\,\mathrm{d}t \\
&\leq \sum_{i=1}^d \int_{N(s_{i,n})} \mathbb{E}\Big(\sup_{\|r-s\| \leq \varepsilon_n} |Z_r - Z_s| \Big)\,\mathrm{d}t \\
&= \int_{[0,1]^k} \mathbb{E}\Big(\sup_{\|r-s\| \leq \varepsilon_n} |Z_r - Z_s| \Big)\,\mathrm{d}t \\
&= \mathbb{E}\Big(\sup_{\|r-s\| \leq \varepsilon_n} |Z_r - Z_s| \Big) \\
&\to 0, \qquad n \to \infty.
\end{aligned}$$

Define

$$A_n := \frac{K(2\varepsilon_n/h_n)}{K(\varepsilon_n/h_n)} \mathbb{E}\Big(\sup_{t \in [0,1]^k} Z_t \Big), \qquad B_n := \mathbb{E}\Big(\sup_{\|r-s\| < 3\varepsilon_n} |Z_r - Z_s| \Big) + 1.$$

As we have seen in the proof of Lemma 6, we have, for $t \in N(s_{i,n})$,

$$1 \geq g_{i,n}(t) \geq (A_n + B_n)^{-1} \to 1,$$

and, therefore,

$$\begin{aligned}
S_{2,n} &= \sum_{i=1}^d \int_{N(s_{i,n})} (1 - g_{i,n}(t))\,\mathrm{d}t \\
&\leq \sum_{i=1}^d \int_{N(s_{i,n})} 1 - (A_n + B_n)^{-1}\,\mathrm{d}t
\end{aligned}$$

https://doi.org/10.1017/jpr.2017.35 Published online by Cambridge University Press

$$\begin{aligned} &= \int_{[0,1]^k} 1 - (A_n + B_n)^{-1}\,\mathrm{d}t \\ &= 1 - (A_n + B_n)^{-1} \\ &\to 0, \qquad n \to \infty. \end{aligned}$$

Finally, we have, by the same argument as above,

$$S_{3,n} = \sum_{i=1}^{d} \int_{N(s_{i,n})} \mathbb{E}(\hat{Z}_{t,n} - g_{i,n}(t) Z_{s_i,n})\,\mathrm{d}t = S_{2,n}, \to 0, \qquad n \to \infty,$$

which completes the proof. □

Remark 2. Given a grid $s_1, \dots, s_{d(n)}$ with pertaining ε_n, the bandwidth $h_n := \varepsilon_n^2$ would, for example, satisfy the required growth conditions entailing convergence of MSE and IMSE to 0. But it would clearly be desirable to provide some details on how to choose the bandwidth in an optimal way, which is, for example, statistical folklore in kernel density estimation. In our setup, however, this is an open problem that requires future work.

4. Discretized versions of copula processes

Next we transfer the model established in Section 2 to copula processes that are in a sense close to max-stable processes. A *copula process* $\boldsymbol{U} = (U_t)_{t\in[0,1]^k}$ is a stochastic process with continuous sample paths, such that each RV U_t is uniformly distributed on the interval [0, 1]. We say that $\boldsymbol{U}$ is in the *functional domain of attraction* of an SMSP $\boldsymbol{\eta} = (\eta_t)_{t\in[0,1]^k}$, if

$$\lim_{n\to\infty} \mathbb{P}(n(\boldsymbol{U} - 1) \le f)^n = \mathbb{P}(\boldsymbol{\eta} \le f) = \exp(-\|f\|_D), \qquad f \in \bar{E}^-([0,1]^k). \tag{8}$$

Define, for any $t \in [0,1]^k$ and $n \in \mathbb{N}$,

$$Y_t^{(n)} := n\Big(\max_{i=1,\dots,n} U_t^{(i)} - 1\Big),$$

with $\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \dots$ being independent copies of $\boldsymbol{U}$. Now choose again pairwise different points $s_1, \dots, s_d \in [0,1]^k$ and functions $g_1, \dots, g_d \in \bar{C}^+([0,1]^k)$ with the properties (2) and (5). Condition (8) implies weak convergence of the finite-dimensional distributions of $\boldsymbol{Y}^{(n)} = (Y_t^{(n)})_{t\in[0,1]^k}$, i.e.

$$(Y_{s_1}^{(n)}, \dots, Y_{s_d}^{(n)}) \xrightarrow{\mathrm{D}} (\eta_{s_1}, \dots, \eta_{s_d}),$$

where '$\xrightarrow{\mathrm{D}}$' denotes convergence in distribution. As before, we can define the *discretized version* $\hat{\boldsymbol{Y}}^{(n)} = (\hat{Y}_t^{(n)})_{t\in[0,1]^k}$ of $\boldsymbol{Y}^{(n)}$ with grid $\{s_1, \dots, s_d\}$ and weight functions $g_1, \dots, g_d$ to be

$$\hat{Y}_t^{(n)} := \max_{i=1,\dots,d} \frac{Y_{s_i}^{(n)}}{g_i(t)}, \qquad t \in [0,1]^k.$$

Elementary calculations show that (8) implies that

$$\lim_{n\to\infty} \mathbb{P}(\hat{\boldsymbol{Y}}^{(n)} \le f) = \mathbb{P}(\hat{\boldsymbol{\eta}} \le f), \qquad f \in \bar{E}^-([0,1]^k),$$

where $\hat{\boldsymbol{\eta}}$ is the discretized version of $\boldsymbol{\eta}$ as defined in (3). Also, it is not difficult to see that, for each $t \in [0,1]^k$,

$$(Y_t^{(n)}, \hat{Y}_t^{(n)}) \xrightarrow{\mathrm{D}} (\eta_t, \hat{\eta}_t),$$

where $(\eta_t, \hat{\eta}_t)$ is the standard max-stable RV from Lemma 2. Now applying the continuous mapping theorem, we obtain

$$(Y_t^{(n)} - \hat{Y}_t^{(n)})^2 \xrightarrow{\mathrm{D}} (\eta_t - \hat{\eta}_t)^2.$$

It remains to prove uniform integrability of the sequence on the left-hand side in order to obtain the next result.

Proposition 2. *Let $t \in [0,1]^k$. Then*

$$\mathrm{MSE}(\hat{Y}_t^{(n)}) = \mathbb{E}((Y_t^{(n)} - \hat{Y}_t^{(n)})^2) \to \mathrm{MSE}(\hat{\eta}_t), \qquad n \to \infty.$$

Proof. Fix $t \in [0,1]^k$. It remains to show that the sequence $X_t^{(n)} := (Y_t^{(n)} - \hat{Y}_t^{(n)})^2$ is uniformly integrable. A sufficient condition for uniform integrability is

$$\sup_{n \in \mathbb{N}} \mathbb{E}((X_t^{(n)})^2) < \infty;$$

see [3, Section 3]. Clearly, for every $n \in \mathbb{N}$,

$$\mathbb{E}((X_t^{(n)})^2) \le \mathbb{E}((Y_t^{(n)})^4) + \mathbb{E}((\hat{Y}_t^{(n)})^4).$$

It is easy to verify that the RV $Y_t^{(n)}$ has the density $(1 + x/n)^{n-1}$ on $[-n, 0]$. Therefore,

$$\mathbb{E}((Y_t^{(n)})^4) = \int_{-n}^{0} x^4 \left(1 + \frac{x}{n}\right)^{n-1} \mathrm{d}x = \frac{24 n^5 (n-1)!}{(n+4)!} \le 24.$$

Moreover, setting $c := \min_{i=1,\dots,d} g_i(t) > 0$,

$$|\hat{Y}_t^{(n)}| = \min_{i=1,\dots,d} \frac{|Y_{s_i}^{(n)}|}{g_i(t)} \le \frac{|Y_{s_1}^{(n)}|}{c},$$

and, hence,

$$\mathbb{E}((\hat{Y}_t^{(n)})^4) \le \frac{24}{c^4},$$

which completes the proof. □

Acknowledgements

We are grateful to two anonymous referees for their careful reading of the manuscript. The paper has benefitted greatly from their critical remarks.

References

[1] Aulbach, S., Falk, M. and Hofmann, M. (2013). On max-stable processes and the functional D-norm. *Extremes* **16,** 255–283.

[2] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley, Chichester.

[3] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
[4] Brown, B. M. and Resnick, S. I. (1977). Extreme values of independent stochastic processes. *J. Appl. Prob.* **14,** 732–739.
[5] Cooley, D. and Sain, S. R. (2012). Discussion of 'Statistical modeling of spatial extremes' by A. C. Davison, S. A. Padoan and M. Ribatet. *Statist. Sci.* **27,** 187–188.
[6] Davison, A. C., Padoan, S. A. and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27,** 161–186.
[7] Davison, A. C., Padoan, S. A. and Ribatet, M. (2012). Rejoinder: 'Statistical modeling of spatial extremes'. *Statist. Sci.* **27,** 199–201.
[8] De Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York. See http://people.few.eur.nl/ldehaan/EVTbook.correction.pdf and http://home.isa.utl.pt/~anafh/corrections.pdf for corrections and extensions.
[9] De Haan, L. and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitsth.* **40,** 317–337.
[10] Dombry, C., Éyi-Minko, F. and Ribatet, M. (2013). Conditional simulation of max-stable processes. *Biometrika* **100,** 111–124.
[11] Falk, M., Hofmann, M. and Zott, M. (2015). On generalized max-linear models and their statistical interpolation. *J. Appl. Prob.* **52,** 736–751.
[12] Falk, M., Hüsler, J. and Reiss, R.-D. (2011). *Laws of Small Numbers: Extremes and Rare Events*, 3rd edn. Birkhäuser, Basel.
[13] Gabda, D., Towe, R., Wadsworth, J. and Tawn, J. (2012). Discussion of 'Statistical modeling of spatial extremes' by A. C. Davison, S. A. Padoan and M. Ribatet. *Statist. Sci.* **27,** 189–192.
[14] Giné, E., Hahn, M. G. and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Prob. Theory Relat. Fields* **87,** 139–165.
[15] Huser, R. and Davison, A. C. (2013). Composite likelihood estimation for the Brown–Resnick process. *Biometrika* **100,** 511–518.
[16] Kabluchko, Z. (2009). Spectral representations of sum- and max-stable processes. *Extremes* **12,** 401–424.
[17] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Prob.* **37,** 2042–2065.
[18] Pickands, J., III (1981). Multivariate extreme value distributions. *Bull. Inst. Internat. Statist.* **49,** 859–878, 894–902.
[19] Resnick, S. I. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
[20] Segers, J. (2012). Nonparametric inference for max-stable dependence. *Statist. Sci.* **27,** 193–196.
[21] Shaby, B. and Reich, B. J. (2012). Discussion of 'Statistical modeling of spatial extremes' by A. C. Davison, S. A. Padoan and M. Ribatet. *Statist. Sci.* **27,** 197–198.
[22] Wang, Y. and Stoev, S. A. (2011). Conditional sampling for spectrally discrete max-stable random fields. *Adv. Appl. Prob.* **43,** 461–483.