

Clustering of high values in random fields

Luísa Pereira¹ · Ana Paula Martins¹ ·
Helena Ferreira¹

Received: 19 October 2015 / Revised: 30 January 2017 / Accepted: 9 March 2017/
Published online: 30 March 2017
© Springer Science+Business Media New York 2017

Abstract The asymptotic results that underlie applications of extreme random fields often assume that the variables are located on a regular discrete grid, identified with \mathbb{Z}^2 , and that they satisfy stationarity and isotropy conditions. Here we extend the existing theory, concerning the asymptotic behavior of the maximum and the extremal index, to non-stationary and anisotropic random fields, defined over discrete subsets of \mathbb{R}^2 . We show that, under a suitable coordinatewise mixing condition, the maximum may be regarded as the maximum of an approximately independent sequence of submaxima, although there may be high local dependence leading to clustering of high values. Under restrictions on the local path behavior of high values, criteria are given for the existence and value of the spatial extremal index which plays a key role in determining the cluster sizes and quantifying the strength of dependence between exceedances of high levels. The general theory is applied to the class of max-stable random fields, for which the extremal index is obtained as a function of well-known tail dependence measures found in the literature, leading to a simple estimation method for this parameter. The results are illustrated with non-stationary Gaussian and 1-dependent random fields. For the latter, a simulation and estimation study is performed.

✉ Luísa Pereira
lpereira@ubi.pt

Ana Paula Martins
amartins@ubi.pt

Helena Ferreira
helenaf@ubi.pt

¹ Universidade da Beira Interior, Centro de Matemática e Aplicações (CMA-UBI), Avenida Marquês d' Ávila e Bolama, 6200-001 Covilhã, Portugal

Keywords Random field · Max-stable process · Extremal dependence · Spatial extremal index

AMS 2000 Subject Classifications 60G60 · 60G70

1 Introduction

Extremes of variables like wind, temperature and precipitation can affect anybody at any place. The potential consequences include increases in severe windstorms, flooding, wildfires, crop failure, population displacements and increased mortality. Apart from their direct impacts, these events will also have indirect effects such as increased costs for strengthening infrastructure or higher insurance premiums. When the interest lies in the study of variables measured at specifically-located monitors, such as the variables mentioned above, as well as air pollution, soil porosity or hydraulic conductivity, among others, spatial modeling is necessary, so random fields constitute an active area of current research.

The treatment of spatial and temporal dependence in random fields has been influenced by the multivariate Gaussian model, where dependence is characterized by the covariance structures. However, this model excludes all the situations of marginal distributions with heavier tails than the Gaussian distribution, leaving aside a huge set of problems related to rare events. Extreme Value Theory plays an important role in these situations.

A considerable amount of work has been done in extending results of Extreme Value Theory to random fields which have \mathbb{Z}^2 as their parameter space. Although their lack of easy separation of past and future, a general version of the classical Extreme Types Theorem was given, by replacing a single global dependence restriction by several assumptions, each dealing with one coordinate direction, for which past-future separation is considered (Leadbetter and Rootzén 1998; Pereira and Ferreira 2006, among others). Recall that in a random field with high local dependence, an exceedance is likely to have neighboring exceedances, resulting in a clustering of exceedances, which leads to a compounding of events in the limiting point process of exceedances (Ferreira and Pereira 2012). For example, an unusually stormy day at a particular location may be followed by another one at the same or a neighboring location. This type of dependence among spatial extremes can be summarized through the spatial extremal index of the random field. Under local restrictions on the oscillations of the values of the random field, Ferreira and Pereira (2008) and Pereira and Ferreira (2006) compute the extremal index from the joint distribution of a finite number of variables.

The aforementioned results assumed that the variables are located on a regular grid (Fig. 1, on left), identified with \mathbb{Z}^2 , and sometimes that they satisfy stationarity and isotropy conditions. This is a big restriction for the majority of the applications since usually spatial data are not regularly spaced (Fig. 1, on right), stationary and dependence is anisotropic, due to the presence of a main direction of dependence.

An usual way to deal with spatial observations is to interpolate the results to a regular longitude/latitude grid. However, the options for avoiding the reduction in

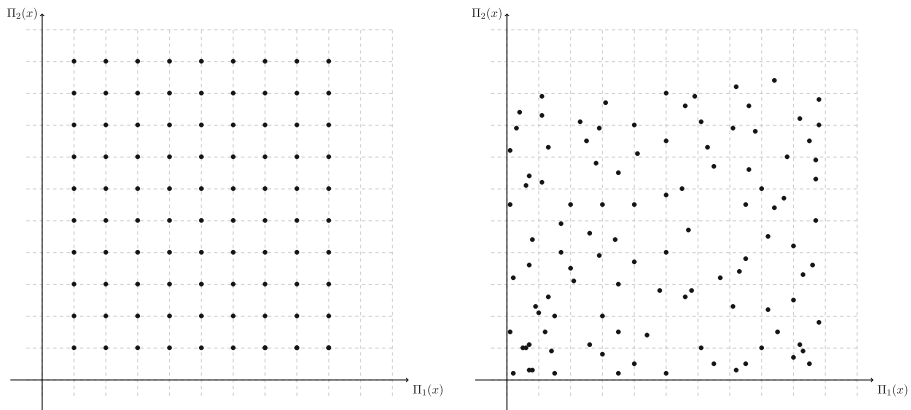


Fig. 1 Regular Grid on *left* and Random Pattern on *right*

extremes are very limited. The interpolation typically aims to assign some sort of “most likely value” to the grid points based on the surrounding stations. Since the most likely scenario is never the most extreme scenario this means that extremes are under represented.

To overcome this problem, in this paper we extend the existing theory, concerning the asymptotic behavior of the maximum and the spatial extremal index (Pereira and Ferreira 2006; Ferreira and Pereira 2008; Leadbetter and Rootzén 1998), to non-stationary and anisotropic random fields, $\mathbf{Z} = \{Z(x) : x \in S\}$, where $S = \bigcup_{n \geq 1} A_n$ and $A = \{A_n\}_{n \geq 1}$ is an increasing sequence of sets of isolated points of \mathbb{R}^2 , subject to mixing conditions.

In Section 2, we define a coordinatewise-mixing condition under which we prove that the maximum of $Z(x)$, $x \in A_n$, $n \geq 1$, behaves asymptotically as independent maxima over a family of disjoint subsets of A_n and we give criteria for the existence and value of the spatial extremal index, which plays a key role in determining the cluster sizes and quantifying the strength of dependence between exceedances of high values, as mentioned above.

Section 3 is devoted to the study of the behavior of maxima when clustering of high values of \mathbf{Z} is allowed, but we restrict the local path behavior of high values through an extension to random fields of the $D''(u_n)$ -condition found in Leadbetter and Nandagopalan (1988). Surprisingly, these smooth oscillation conditions allow us to obtain a simple method for computing the extremal index of the sequence $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n \geq 1}$, denoted by θ_A , as the limit of a sequence of tail dependence coefficients. To our knowledge is the first time that is established a connection between the spatial extremal index and well-known tail dependence coefficients found in literature (Li 2009; Sibuya 1959).

The natural class of models for spatial extremes are the max-stable random fields. Under suitable conditions, max-stable random fields are asymptotically justified models for maxima of independent replications of random fields, since they extend the generalized extreme value of univariate extreme value theory to the spatial

setting (de Haan 1984). They have been widely applied to real data in environmental, atmospheric and geological sciences.

Given the usefulness of max-stable random fields for modeling extreme events in the spatial domain, in Section 4 we apply the general theory to this class of random fields. We obtain the spatial extremal index as a function of the extremal coefficient defined in Schlather and Tawn (2003) (see also Smith 1990). This coefficient allows us to summarize the dependence structure of a max-stable random field and is widely applied as a dependence measure in a range of practical applications. In this section we also introduce the notions of local and regional extremal indices and relate them with θ_A .

Based upon these relations a simple estimator for θ_A is given in Section 5 and its performance is analyzed with an anisotropic and non-stationary 1-dependent max-stable random field.

Conclusions are drawn in Section 6 and the proofs and auxiliary results are collected in the appendices.

Throughout the paper we will assume, without loss of generality, that the variables $Z(x)$, $x \in S$, have common distribution F , being \bar{F} the corresponding survival function. We will denote the maximum and the minimum of $Z(x)$ over $B \subset S$ by $\bigvee_{x \in B} Z(x)$ and $\bigwedge_{x \in B} Z(x)$, respectively.

2 Asymptotic spatial independence

In this section, we show that, under a suitable mixing condition, the maximum of random fields defined over discrete subsets of \mathbb{R}^2 , may be regarded as the maximum of an approximately independent sequence of submaxima, even though there may be high local dependence leading to clustering of high values.

The dependence structure used here is a coordinatewise mixing condition, which restricts dependence by limiting $|P(\bigvee_{x \in C \cup D} Z(x) \leq u_n) - P(\bigvee_{x \in C} Z(x) \leq u_n)P(\bigvee_{x \in D} Z(x) \leq u_n)|$ with the two index sets $C, D \subset A_n$ being “separated” from each other by a certain distance l_n along some direction.

The results are obtained through an extension of the methodology in Pereira and Ferreira (2006), for extremes on a regular grid, relying on the novelty of irregularly occurring extremes in space.

Throughout we shall say that the pair (I, J) is in $\mathcal{S}(\pi_i(A_n), l_n)$ if $I \subset S$ and $J \subset S$ are subsets of consecutive values of $\pi_i(A_n)$ separated by at least l_n values of $\pi_i(A_n)$, where π_i , $i = 1, 2$, denote the cartesian projections. The cardinality of the sets A_n and $\pi_i(A_n)$, $i = 1, 2$, will be denoted by $\sharp A_n = f(n)$, $\sharp \pi_i(A_n) = f_i(n)$, $i = 1, 2$, and we will assume that $f(n) \rightarrow +\infty$ as $n \rightarrow +\infty$.

Definition 2.1 Let $\{u_n\}_{n \geq 1}$ be a sequence of real numbers. If there exist sequences of positive integers $l = \{l_n\}_{n \geq 1}$ and $k = \{k_n\}_{n \geq 1}$ satisfying

$$l_n \rightarrow +\infty, \quad k_n \rightarrow +\infty, \quad k_n l_n \frac{f_i(n)}{f(n)} \rightarrow 0, \quad \text{as } n \rightarrow +\infty, \quad \text{for each } i = 1, 2, \quad (2.1)$$

and $k_n^2 \alpha(l_n, u_n) \xrightarrow{n \rightarrow +\infty} 0$, with

$$\alpha(l_n, u_n) = \sup \left| P \left(\bigvee_{x \in C \cup D} Z(x) \leq u_n \right) - P \left(\bigvee_{x \in C} Z(x) \leq u_n \right) P \left(\bigvee_{x \in D} Z(x) \leq u_n \right) \right|,$$

where the supremum is taken over sets C and D such that $C, D \in \mathcal{S}(\pi_i(A_n), l_n)$, for some $i \in \{1, 2\}$, we say that the sequence \mathbf{Z}_A satisfies condition $D(u_n, k_n, l_n)$.

Under $D(u_n, k_n, l_n)$ -condition the asymptotic independence of maxima over disjoint sets of locations is guaranteed, from Lemma A.1 (see Appendix A), and therefore we can prove that asymptotically the distribution of the maximum of \mathbf{Z} over A_n , $n \geq 1$, coincides with the distribution of the maximum of \mathbf{Z} over a union of conveniently chosen disjoint subsets of A_n , see Lemma A.2 in Appendix A.

The underlying idea to obtain the asymptotic distribution of the maximum of \mathbf{Z} over A_n , $n \geq 1$, is to subdivide A_n into k_n^2 disjoint subsets, $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, using the following construction method of the family $\mathcal{B}_n = \{B_n^{(s,t)} : s, t = 1, \dots, k_n\}$:

- build $\pi_1^{-1}(I^{(s)}) \cap A_n$, $s = 1, \dots, k_n$, with $I^{(s)}$, $s = 1, \dots, k_n$, abutting subsets of k_n consecutive values of $\pi_1(A_n)$, maximally chosen for the condition

$$\sum_{x \in \pi_1^{-1}(I^{(s)}) \cap A_n} P(Z(x) > u_n) \leq \frac{1}{k_n} \sum_{x \in A_n} P(Z(x) > u_n);$$

- for each $s = 1, \dots, k_n$, build $\pi_2^{-1}(J^{(s,t)}) \cap \pi_1^{-1}(I^{(s)}) \cap A_n$, $t = 1, \dots, k_n$, with $J^{(s,t)}$, $t = 1, \dots, k_n$, contiguous subsets of $\pi_2(\pi_1^{-1}(I^{(s)}) \cap A_n)$ and maximally chosen such that

$$\sum_{x \in \pi_2^{-1}(J^{(s,t)}) \cap \pi_1^{-1}(I^{(s)}) \cap A_n} P(Z(x) > u_n) \leq \frac{1}{k_n^2} \sum_{x \in A_n} P(Z(x) > u_n).$$

Figure 2 illustrates one possible set of disjoint blocks $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, with $\bigcup_{s,t} B_n^{(s,t)} = A_n$ constructed through the previous method, for a particular A_n , $n \in \mathbb{N}$.

Remark 1 In the above block construction, the quantity $P(Z(x) > u_n)$ factorizes on both sides and therefore the inequalities can be written as constraints on the cardinality of the sets under consideration. However, we decided to write it without such simplifications in order to uncover the original idea which can be extended to unequal marginal distributions F_i .

As a consequence of Lemmas A.1 and A.2 we can now state the following result concerning the asymptotic independence of maxima over $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, which extends Proposition 2.1 of Pereira and Ferreira (2006).

Proposition 2.1 Suppose that the sequence \mathbf{Z}_A satisfies condition $D(u_n, k_n, l_n)$ for a sequence of real numbers $\{u_n\}_{n \geq 1}$ such that

$$\{f(n)\overline{F}(u_n)\}_{n \geq 1} \text{ is bounded.} \quad (2.2)$$

Then, for each n , there exists a family \mathcal{B}_n of k_n^2 disjoint subsets of A_n , $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, with $\sharp B_n^{(s,t)} \sim \frac{f(n)}{k_n^2}$ and such that

$$P\left(\bigvee_{x \in A_n} Z(x) \leq u_n\right) - \prod_{s,t} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n\right) \xrightarrow{n \rightarrow +\infty} 0.$$

The way spatial extreme events interact is also of interest in spatial statistics. For example, an unusually stormy day at a particular location may be followed by another one at the same or a neighboring location. This type of dependence among spatial extremes can be summarized through the spatial extremal index of the sequence \mathbf{Z}_A .

Definition 2.2 The sequence \mathbf{Z}_A has spatial extremal index θ_A if, for each $\tau > 0$ and any sequence of real numbers $\{u_n(\tau)\}_{n \geq 1}$ satisfying

$$E\left(\sum_{x \in A_n} \mathbb{I}_{\{Z(x) > u_n(\tau)\}}\right) \xrightarrow{n \rightarrow +\infty} \tau, \quad (2.3)$$

where \mathbb{I}_A denotes the indicator function of the event A , it holds that

$$\lim_{n \rightarrow +\infty} P\left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau)\right) = \exp(-\theta_A \tau).$$

The extremal index of \mathbf{Z}_A is the key parameter to relate the limiting distributions of $\bigvee_{x \in A_n} Z(x)$ and $\bigvee_{x \in A_n} \widehat{Z}(x)$, where $\widehat{\mathbf{Z}}_A = \{\widehat{Z}(x), x \in A_n\}_{n \geq 1}$ is a sequence of independent random variables having the same distribution function F as each variable of the sequence \mathbf{Z}_A . In fact, if $f(n)$ is the number of locations on A_n and there exists a sequence of real numbers $\{u_n(\tau)\}_{n \geq 1}$ satisfying (2.3), then

$$\begin{aligned} P\left(\bigvee_{x \in A_n} \widehat{Z}(x) \leq u_n(\tau)\right) &= F^{f(n)}(u_n(\tau)) \xrightarrow{n \rightarrow +\infty} \exp\left(-\lim_{n \rightarrow +\infty} f(n)\overline{F}(u_n(\tau))\right) \\ &= \exp(-\tau) \end{aligned}$$

and

$$P\left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau)\right) \xrightarrow{n \rightarrow +\infty} \exp\left(-\lim_{n \rightarrow +\infty} \theta f(n)\overline{F}(u_n(\tau))\right) = \exp(-\theta_A \tau).$$

Consequently, the following equalities hold:

1. $\lim_{n \rightarrow +\infty} P \left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau) \right) = \lim_{n \rightarrow +\infty} P \left(\bigvee_{x \in B_n} \widehat{Z}(x) \leq u_n(\tau) \right),$ with $\sharp B_n = \theta_A f(n) \leq \sharp A_n$, i.e., for the sequence of real levels $\{u_n(\tau)\}_{n \geq 1}$, $\bigvee_{x \in A_n} Z(x)$ behaves asymptotically as the maximum of less than $f(n)$ independent variables.
2. $\lim_{n \rightarrow +\infty} P \left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau) \right) = \lim_{n \rightarrow +\infty} P \left(\bigvee_{x \in A_n} \widehat{Z}(x) \leq v_n(\theta\tau) \right)$ where $v_n(\theta\tau) \cong F^{-1} \left(1 - \frac{\theta\tau}{f(n)} \right) > u_n(\tau) \cong F^{-1} \left(1 - \frac{\tau}{f(n)} \right)$, i.e., $\bigvee_{x \in A_n} Z(x)$ behaves asymptotically as the maximum of the same number of independent variables but relatively to a higher level than $u_n(\tau)$.

The following result gives a convenient existence criterion for the extremal index of \mathbf{Z}_A and follows from Proposition 2.1: it depends on the local behavior of exceedances over $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, namely, on the limiting mean number of exceedances of u_n by $\bigvee_{x \in B_n^{(s,t)}} Z(x)$, $s, t = 1, \dots, k_n$.

Proposition 2.2 *Suppose that the sequence \mathbf{Z}_A satisfies condition $D(u_n(\tau), k_n, l_n)$, where $\{u_n(\tau)\}_{n \geq 1}$ is a sequence of real numbers satisfying (2.3) and \mathcal{B}_n is a family of subsets of A_n satisfying the conditions of Proposition 2.1. Then, there exists the spatial extremal index, θ_A , if and only if the following limit exists*

$$\lim_{n \rightarrow +\infty} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau) \right),$$

and, in this case, we have

$$\theta_A = \lim_{n \rightarrow +\infty} \frac{1}{f(n) \overline{F}(u_n(\tau))} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau) \right).$$

Note that if $\{u_n(\tau)\}_{n \geq 1}$ is a sequence of real numbers for which (2.3) holds, then it also satisfies condition (2.2).

We are now in conditions to prove that the expected number of exceedances of the level $u_n(\tau)$ on the blocks $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, with at least one exceedance, converges to the reciprocal of the extremal index θ_A . We can verify that the greater the clustering tendency of high threshold exceedances (several exceedances on $B_n^{(s,t)}$) the smaller θ_A will be. For isolated exceedances of $u_n(\tau)$, we have $\theta_A = 1$.

Proposition 2.3 *Suppose that the sequence \mathbf{Z}_A satisfies condition $D(u_n(\tau), k_n, l_n)$, where $\{u_n(\tau)\}_{n \geq 1}$ is a sequence of real numbers satisfying (2.3) and \mathcal{B}_n is a family*

of subsets of A_n satisfying the conditions of Proposition 2.1. If the sequence \mathbf{Z}_A has spatial extremal index, θ_A , then

$$\theta_A = \lim_{n \rightarrow +\infty} \frac{1}{k_n^2} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} E^{-1} \left(\sum_{x \in B_n^{(s,t)}} \mathbb{I}_{\{Z(x) > u_n(\tau)\}} \left| \sum_{x \in B_n^{(s,t)}} \mathbb{I}_{\{Z(x) > u_n(\tau)\}} > 0 \right. \right).$$

If

$$\lim_{n \rightarrow +\infty} E \left(\sum_{x \in B_n^{(s,t)}} \mathbb{I}_{\{Z(x) > u_n(\tau)\}} \left| \sum_{x \in B_n^{(s,t)}} \mathbb{I}_{\{Z(x) > u_n(\tau)\}} > 0 \right. \right) = 1,$$

uniformly in $s, t \in \{1, \dots, k_n\}$, then $\theta_A = 1$.

3 Local spatial dependence

The asymptotic behavior of the maximum of non-stationary and anisotropic random fields, defined over discrete subsets of \mathbb{R}^2 , subject to restrictions on the local path behavior of high values is now analyzed.

Criteria are given for the existence and value of the spatial extremal index, which plays a key role in determining the cluster sizes and quantifying the strength of dependence between exceedances of high levels. To attain this goal, we first introduce a condition for modeling local mild oscillations of the random field. This condition is an extension to random fields of the $D''(u_n)$ -condition found in Leadbetter and Nandagopalan (1988).

Throughout this section \mathcal{B}_n will denote a family of subsets of A_n in the conditions of Proposition 2.1.

Definition 3.1 If $V(x)$ is a finite set of neighbors of a point $x \in A$ and $\mathcal{V} = \{V(x) : x \in A\}$, then the sequence \mathbf{Z}_A verifies condition $D''(u_n, \mathcal{B}_n, \mathcal{V})$ if

$$k_n^2 \sup_{B_n^{(s,t)} \in \mathcal{B}_n} \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n \geq \bigvee_{y \in V(x)} Z(y), \bigvee_{y \in B_n^{(s,t)} - V(x)} Z(y) > u_n \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Although the choice of the family \mathcal{V} of neighborhoods can be conditioned by the nature of the practical problems under study, here we will illustrate the modeling with a natural choice based on the cardinal directions. Therefore, in what follows the initials N, E, S, W will represent, respectively, the cardinal directions North, East, South and West. The family of neighborhoods of $x \in A$ along directions E and N, will be denoted by $\mathcal{V}_{E,N}^{p,q,r}$, $p, q, r \in \mathbb{Z}_0^+$, and defined as

$$\mathcal{V}_{E,N}^{p,q,r} = \{V(x) : x \in A\},$$

where

$$V(x) = \{y : (a_1(\pi_1(x)) \leq \pi_1(y) \leq a_p(\pi_1(x)) \wedge a_{-q}(\pi_2(x)) < \pi_2(y) \leq a_r(\pi_2(x))) \\ \vee (\pi_1(x) = \pi_1(y) \wedge a_1(\pi_2(x)) \leq \pi_2(y) \leq a_r(\pi_2(x)))\}, x \in A\},$$

and, for each $z \in A_n$,

$$a_1(\pi_1(z)), \dots, a_r(\pi_1(z))$$

are the points after $\pi_1(z)$, in ascending order, and

$$a_{-q}(\pi_2(z)), \dots, a_{-2}(\pi_2(z)), a_{-1}(\pi_2(z)), a_0(\pi_2(z)) = \pi_2(z), a_1(\pi_2(z)), a_2(\pi_2(z)), \dots, a_r(\pi_2(z))$$

are the points before and after $\pi_2(z)$, also in ascending order.

In Fig. 3 we find an illustration of a neighborhood $V(x) \in \mathcal{V}_{E,N}^{p,q,r}$, $p, q, r \in \mathbb{Z}_0^+$.

As a consequence of Lemma B.1 (see Appendix B) the extremal index of \mathbf{Z}_A , θ_A , can be viewed, as $n \rightarrow +\infty$, as the mean of tail dependence coefficients of the form

$$\lambda_n^{(\tau)}(V(x), x) = P \left(\bigvee_{y \in V(x)} Z(y) > u_n(\tau) \middle| Z(x) > u_n(\tau) \right), \quad (3.1)$$

which are the tail dependence coefficients of Li (2009). Observe also that if $\sharp V(x) = 1$, then $\lambda_n^{(\tau)}(V(x), x)$ is the traditional upper tail dependence coefficient introduced far back in the sixties (Sibuya 1959; Oliveira 1962/63).

Proposition 3.1 *Let $\{u_n(\tau)\}_{n \geq 1}$ be a sequence of real numbers satisfying (2.3). If sequence \mathbf{Z}_A verifies conditions $D''(u_n(\tau), \mathcal{B}_n, \mathcal{V})$ and $D(u_n(\tau), k_n, l_n)$ then the spatial extremal index of \mathbf{Z}_A , θ_A , exists if and only if there exists*

$$\lambda_A = \lim_{n \rightarrow +\infty} \frac{1}{f(n)} \sum_{x \in A_n} \lambda_n^{(\tau)}(V(x), x),$$

and, in this case, we have

$$\theta_A = 1 - \lambda_A.$$

Remark 2 The extremal index θ_A does not depend on the choice of the sequence $\{u_n(\tau)\}_{n \geq 1}$, provided it satisfies (2.3).

Note that some models can verify condition D'' only for certain types of neighborhoods. Nevertheless, there exist models, as we shall see further on, that verify condition $D''(u_n, \mathcal{B}_n, \mathcal{V})$, for all \mathcal{V} . A particular case of such models are those that verify a local dependence restriction that leads to isolated exceedances, which we shall denominate condition $D'(u_n, \mathcal{B}_n)$ and define as follows:

Definition 3.2 The sequence \mathbf{Z}_A verifies condition $D'(u_n, \mathcal{B}_n)$ if

$$\sum_{B_n^{(s,t)} \in \mathcal{B}_n} \sum_{x, y \in B_n^{(s,t)}} P(Z(x) > u_n, Z(y) > u_n) \xrightarrow{n \rightarrow +\infty} 0.$$

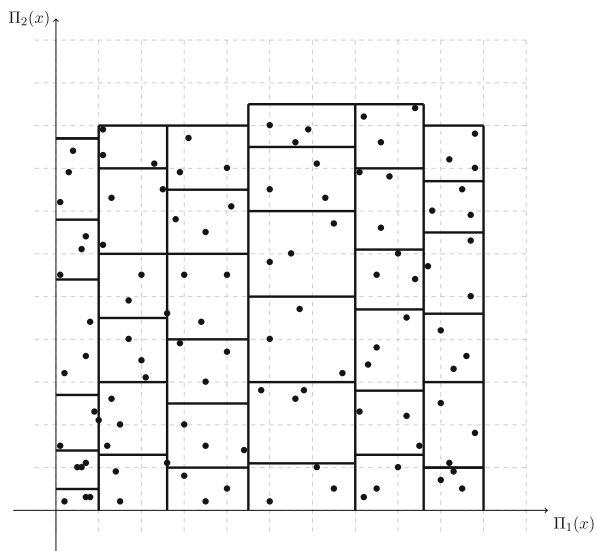


Fig. 2 Example of a set of disjoint blocks $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, with $\bigcup_{s,t} B_n^{(s,t)} = A_n$, for a particular A_n , $n \in \mathbb{N}$

This dependence condition, which bounds the probability of more than one exceedance of u_n over a block $B_n^{(s,t)}$ with approximately $\frac{f(n)}{k_n^2}$ elements, together with condition $D(u_n, k_n, l_n)$ lead to an unit extremal index.

Proposition 3.2 *Let $\{u_n(\tau)\}_{n \geq 1}$ be a sequence of real numbers satisfying (2.3). Suppose \mathbf{Z}_A satisfies conditions $D'(u_n(\tau), B_n)$ and $D(u_n(\tau), k_n, l_n)$. Then,*

$$P \left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau) \right) \xrightarrow{n \rightarrow +\infty} \exp(-\tau).$$

We now present a class of Gaussian random fields that verifies the conditions established in Proposition 3.2.

Example 3.1 Let $\mathbf{Z} = \{Z(x) : x \in S\}$ be a standard Gaussian random field on $S \subset \mathbb{R}^2$ with correlations $r_{x,y}$, $x, y \in S$, such that

$$\delta = \sup_{x,y \in S} |r_{x,y}| < 1,$$

where $S = \bigcup_{n \geq 1} A_n$ and $\{A_n\}_{n \geq 1}$ is an increasing sequence of sets of isolated points of \mathbb{R}^2 , satisfying, as $n \rightarrow +\infty$,

$$f(n) \rightarrow +\infty, \quad k_n l_n \frac{f_i(n)}{f(n)} \rightarrow 0, \quad \frac{f_i(n)}{f(n)} \rightarrow 0, \quad i = 1, 2,$$

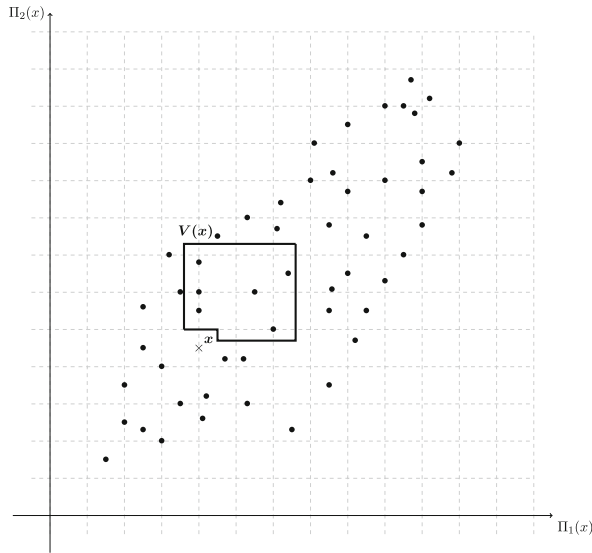


Fig. 3 Example of a neighborhood $V(x) \in \mathcal{V}_{E,N}^{p,q,r}$, with $p = 10$, $q = 1$ and $r = 11$

with $\{k_n\}_{n \geq 1}$ and $\{l_n\}_{n \geq 1}$ sequences of integer numbers verifying $k_n \rightarrow +\infty$ and $l_n \rightarrow +\infty$. We will show that under the following correlation condition,

$$\sup_{x,y \in A_n} |r_{x,y}| \leq r_n = o\left(\frac{1}{\log f(n)}\right),$$

the sequence $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n \geq 1}$ verifies conditions $D(u_n(\tau), k_n, l_n)$ and $D'(u_n(\tau), \mathcal{B}_n)$ with $u_n(\tau) = -\frac{\log \tau}{af(n)} + b_{f(n)}$, where $a_n = (2 \log n)^{1/2}$ and $b_n = (2 \log n)^{1/2} - \frac{1}{2}(2 \log n)^{-1/2}(\log \log n + \log 4\pi)$.

By Corollary 4.2.9 of Leadbetter et al. (1983), we have

$$\begin{aligned} & \sum_{x,y \in B_n^{(s,t)}} P(Z(x) > u_n(\tau), Z(y) > u_n(\tau)) \\ & \leq \sum_{x,y \in B_n^{(s,t)}} \left| P(Z(x) > u_n(\tau), Z(y) > u_n(\tau)) - (1 - \Phi(u_n(\tau)))^2 \right| + \sum_{x,y \in B_n^{(s,t)}} (1 - \Phi(u_n(\tau)))^2 \\ & \leq \sum_{x,y \in B_n^{(s,t)}} K |r_{x,y}| \exp\left(\frac{-u_n^2(\tau)}{1 + |r_{x,y}|}\right) + \frac{f(n)}{k_n^2} (1 - \Phi(u_n(\tau)))^2, \end{aligned}$$

where Φ denotes the standard Gaussian distribution function and K is a constant depending on δ .

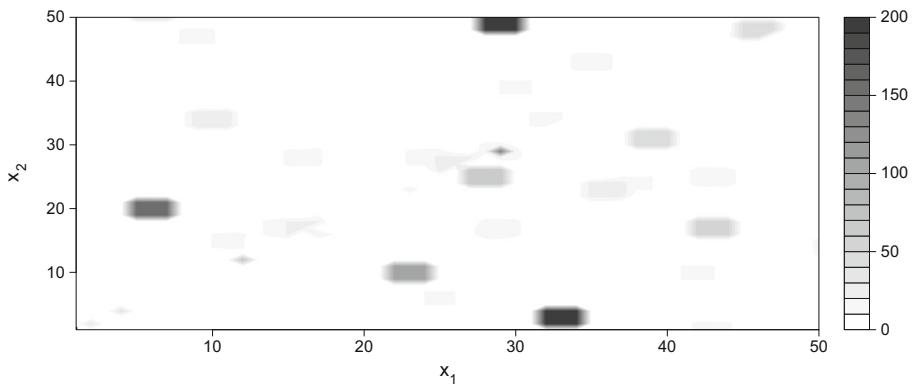


Fig. 4 Simulation of the random field $Z(x)$ defined in Eq. 3.2

Now, since $\sup_{x,y \in A_n} |r_{x,y}| \leq r_n = o\left(\frac{1}{\log f(n)}\right)$, by Lemma 4.3.2 of Leadbetter et al. (1983), we obtain

$$\begin{aligned} & \sum_{B_n^{(s,t)}} \sum_{x,y \in B_n^{(s,t)}} K |r_{x,y}| \exp\left(\frac{-u_n^2(\tau)}{1 + |r_{x,y}|}\right) + f(n)(1 - \Phi(u_n(\tau)))^2 \\ & \leq \sum_{x,y \in A_n} K |r_{x,y}| \exp\left(\frac{-u_n^2(\tau)}{1 + |r_{x,y}|}\right) + o(1) \\ & = o(1), \end{aligned}$$

proving that \mathbf{Z}_A verifies condition $D'(u_n(\tau), B_n)$. Condition $D(u_n(\tau), k_n, l_n)$ follows from Corollary 4.2.4 of Leadbetter et al. (1983).

For other related results concerning Gaussian random fields we refer the readers to Piterbarg (1996), Adler (1981), Berman (1992), Choi (2010), and Pereira (2010).

Proposition 3.1, which states that the spatial extremal index θ_A is asymptotically equal to the mean of tail dependence coefficients is illustrated in the following example with a 1-dependent random field.

Example 3.2 Let $\{Y(x) : x \in \mathbb{Z}^2\}$ be an independent and identically distributed random field with common distribution function $F_Y(y) = \exp\left(\frac{-y^{-1}}{9}\right)$, $y > 0$, and define

$$Z(x) = \begin{cases} 9Y(x) & , \pi_1(x) = \pi_2(x) \\ \bigvee_{y \in E(x)} Y(y) & , \pi_1(x) \neq \pi_2(x), \end{cases} \quad (3.2)$$

where

$$E(x) = \{y : a_{-1}(\pi_1(x)) \leq \pi_1(y) \leq a_1(\pi_1(x)) \wedge a_{-1}(\pi_2(x)) \leq \pi_2(y) \leq a_1(\pi_2(x))\}.$$

Figure 4 shows a simulation of the random field defined in Eq. (3.2). We will calculate the extremal index of sequence $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n \geq 1}$, where

$$A_n = \{x \in \mathbb{Z}^2 : -n \leq \pi_1(x) \leq n, (-n) \vee (\pi_1(x) - n) \leq \pi_2(x) \leq (\pi_1(x) + n) \wedge n\}$$

and

$$f(n) = 3n^2 + 3n + 1, \quad f_i(n) = 2n + 1, \quad i = 1, 2, \quad \frac{f_i(n)}{f(n)} \xrightarrow{n \rightarrow +\infty} 0, \quad \bigcup_{n \geq 1} A_n = \mathbb{Z}^2,$$

by two different methods.

Note that the random field $\mathbf{Z} = \{Z(x) : x \in \bigcup_{n \geq 1} A_n\}$ is anisotropic and non-stationary with common unit Fréchet distribution, $F(y) = \exp(-y^{-1})$, $y > 0$.

Let $\widehat{\mathbf{Z}} = \{\widehat{Z}(x) : x \in \mathbb{Z}^2\}$ be the associated random field of \mathbf{Z} , i.e., $\widehat{Z}(x)$, $x \in \mathbb{Z}^2$, are independent and identically distributed random variables having unit Fréchet distribution.

For a sequence of real numbers $\{u_n(\tau)\}_{n \geq 1}$ verifying (2.3), that is $u_n(\tau) = \frac{f(n)}{\tau}$, we have

$$\lim_{n \rightarrow +\infty} P \left(\bigvee_{x \in A_n} \widehat{Z}(x) \leq u_n(\tau) \right) = \lim_{n \rightarrow +\infty} \left(\exp \left(-\frac{f(n)}{\tau} \right)^{-1} \right)^{f(n)} = \exp(-\tau).$$

On the other hand, for each $D_n = \{x \in A_n : \pi_1(x) = \pi_2(x)\}$ and $u_n(\tau) = \frac{f(n)}{\tau}$, it holds

$$P \left(\bigvee_{x \in D_n} Z(x) > u_n(\tau) \right) \leq \sharp D_n \bar{F}(u_n(\tau)) = \frac{2n+1}{f(n)} f(n) \bar{F}(u_n(\tau)) \xrightarrow{n \rightarrow +\infty} 0,$$

and consequently

$$\begin{aligned} P \left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau) \right) &= P \left(\bigvee_{x \in A_n - D_n} Z(x) \leq u_n(\tau) \right) + o(1) \\ &= P \left(\bigvee_{x \in A_{n+1}} Y(x) \leq u_n(\tau) \right) + o(1) \\ &= (\exp(-u_n(\tau)^{-1}/9))^{3(n+1)^2 + 3(n+1) + 1} \\ &\xrightarrow{n \rightarrow +\infty} \exp \left(-\frac{\tau}{9} \right), \end{aligned}$$

from which we conclude that $\theta_A = \frac{1}{9}$.

Let us now consider Proposition 3.1 for the computation of the extremal index of \mathbf{Z}_A .

Sequence \mathbf{Z}_A verifies condition $D(u_n(\tau), k_n, l_n)$, for all sequences of integer numbers $\{l_n\}_{n \geq 1}$ and $\{k_n\}_{n \geq 1}$ satisfying (2.2), since $\alpha_n(l_n, u_n(\tau)) = 0$ for $l_n \geq 2$.

Considering $\mathcal{V} = \mathcal{V}_{S,W}^{1,1,0}$ the family of neighborhoods

$$V(x) = \{y : (a_{-1}(\pi_2(x)) \leq \pi_2(y) < \pi_2(x) \wedge a_{-1}(\pi_1(x)) \leq \pi_1(y) \leq \pi_1(x)) \vee (\pi_2(y) = \pi_2(x) \wedge a_{-1}(\pi_1(x)) \leq \pi_1(y) < \pi_1(x))\},$$

sequence \mathbf{Z}_A also verifies $D''(u_n(\tau), \mathcal{B}_n, \mathcal{V})$ -condition since, for each $x \notin D_n$,

$$\begin{aligned} P \left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y), \bigvee_{y \in B_n^{(s,t)} - V(x)} Z(y) > u_n(\tau) \right) \\ \leq P(Y(a) > u_n(\tau), Y(b) > u_n(\tau)) = \bar{F}_Y^2(u_n(\tau)), \end{aligned}$$

for pairs of different locations a and b , and consequently

$$\begin{aligned} k_n^2 \sup_{B_n^{(s,t)} \in \mathcal{B}_n} \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y), \bigvee_{y \in B_n^{(s,t)} - V(x)} Z(y) > u_n(\tau) \right) \\ \leq k_n^2 \frac{f(n)}{k_n^2} \bar{F}_Y^2(u_n(\tau)) + o(1) \\ = f(n)(1 - \exp(-\tau/9 f(n)))^2 = o(1). \end{aligned}$$

We can then apply Proposition 3.1 to obtain

$$\begin{aligned} \theta_A &= 1 - \lambda_A \\ &= 1 - \lim_{n \rightarrow +\infty} P \left(\bigvee_{y \in V(x)} Z(y) > u_n(\tau) \mid Z(x) > u_n(\tau) \right) \\ &= 1 - \lim_{n \rightarrow +\infty} \sum_{x \in E(x) - \{(a_{-1}(\pi_1(x)), a_{-1}(\pi_2(x)))\}} \frac{P(Y(x) > u_n(\tau))}{P(Z(x) > u_n(\tau))} = \frac{1}{9}. \end{aligned}$$

4 Consequences for max-stable random fields

Max-stable random fields are very useful models for spatial extremes since under suitable conditions they are asymptotically models for maxima of independent replications of random fields. Furthermore, all finite dimensional distributions of a max-stable process are multivariate extreme value distributions.

Within these random fields, it is important to identify dependence among extremes. In particular detecting asymptotic independence is fundamental and recently some authors have proposed measures of extreme dependence/independence with associated tests. With this in mind we compute, in this section, the extremal index of the class of max-stable random fields, as a function of well known extremal dependence coefficients found in literature, which will provide immediate estimators for θ_A .

One convenient way to summarize the dependence structure of a max-stable random field $\mathbf{Z} = \{Z(x) : x \in S\}$, $S \subseteq \mathbb{R}^2$, with marginal distribution F , is through the

extremal coefficient, ϵ_I , of Smith (1990) and Schlather and Tawn (2003), defined by the relation

$$P\left(\bigvee_{x \in I} Z(x) \leq y\right) = F^{\epsilon_I}(y), \quad y \in \mathbb{R}, \quad I \subseteq \mathbb{R}^2,$$

which measures the extremal dependence between the variables indexed by the set I . Its simple interpretation as the effective number of independent variables indexed in I from which the maximum is drawn has led to its use as a dependence measure in a range of practical applications. With values ranging from 1 to $\sharp I$, corresponding the lower and upper bounds to complete dependence and independence. Observe that if $I = \{x\}$, $x \in \mathbb{R}^2$, the extremal coefficient $\epsilon_I = 1$.

Another way to access the amount of extremal dependence of a random field is through a particular case of the tail dependence function introduced in Ferreira and Ferreira (2012), defined as

$$\Lambda_U^{(I_1|I_2)}(y_1, y_2) = \lim_{t \rightarrow +\infty} P\left(\bigvee_{x \in I_1} F(Z(x)) > 1 - \frac{y_1}{t} \mid \bigvee_{x \in I_2} F(Z(x)) > 1 - \frac{y_2}{t}\right), \quad (y_1, y_2) \in \mathbb{R}_+^2, \quad (4.1)$$

provided the limit exists. This limit always exists for max-stable random fields.

The function $\Lambda_U^{(I_1|I_2)}(y_1, y_2)$ is a measure of the probability of occurring extreme values for the maximum of the variables indexed in a region $I_1 \subseteq \mathbb{R}_+^2$ given that the maximum of the variables indexed in another region I_2 , with $I_1 \cap I_2 = \emptyset$, assumes an extreme value too.

Considering $(y_1, y_2) = (1, 1)$ in Eq. 4.1, we obtain a tail dependence coefficient

$$\Lambda_U^{(I_1|I_2)}(1, 1) = \lim_{t \rightarrow +\infty} P\left(\bigvee_{x \in I_1} Z(x) > 1 - \frac{1}{t} \mid \bigvee_{x \in I_2} Z(x) > 1 - \frac{1}{t}\right),$$

which is related with the extremal coefficients in the following way

$$\Lambda_U^{(I_1|I_2)}(1, 1) = \frac{\epsilon_{I_1} + \epsilon_{I_2} - \epsilon_{I_1 \cup I_2}}{\epsilon_{I_2}},$$

as proved in Proposition 2.2 of Ferreira and Ferreira (2012). Note that if $I_1 = V(x)$, $I_2 = \{x\}$, $x \in \mathbb{R}_+^2$, $\Lambda_U^{(I_1|I_2)}(1, 1)$ is the upper tail dependence coefficient $\lambda_n^{(\tau)}(V(x), x)$ in Eq. 3.1 with $\sharp V(x) = 1$. On the other hand, if $I_1 = V(x)$ and $I_2 = \{x\}$, $x \in \mathbb{R}_+^2$, we obtain a connection between the dependence structure of the sequence of max-stable random fields $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n \geq 1}$ and the limit of the sequence $\{\lambda_n^{(\tau)}(V(x), x)\}_{n \geq 1}$, with $\lambda_n^{(\tau)}(V(x), x)$ defined in Eq. 3.1, as shown in the next result.

We assume, without loss of generality that for each $x \in A_n$, $n \geq 1$, $Z(x)$ has a unit Fréchet distribution.

Proposition 4.1 *Let $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n \geq 1}$ be a sequence of max-stable random fields with unit Fréchet margins. Then*

$$\lim_{n \rightarrow +\infty} (1 - \lambda_n^{(\tau)}(V(x), x)) = \epsilon_{V(x) \cup \{x\}} - \epsilon_{V(x)},$$

which does not depend on the choice of τ .

By combining the previous result with Proposition 3.1 we are now able to conclude that if the extremal index of \mathbf{Z}_A exists, then it can be computed from the difference between extremal coefficients at $V(x) \cup \{x\}$ and $V(x)$, i.e., $\epsilon_{V(x) \cup \{x\}} - \epsilon_{V(x)}$. We shall denote $\epsilon_{V(x) \cup \{x\}} - \epsilon_{V(x)}$ simply by $\theta_{\{x\}}$, $x \in A_n$, and name them local extremal indices.

It now seems natural to combine these local extremal indices, defined on a finite number of points of a region contained in A_n , through an average, in order to obtain a regional extremal index.

Definition 4.1 Let \mathbf{Z}_A be a sequence of max-stable random fields with unit Fréchet margins. The extremal index of \mathbf{Z}_A over a region $R \subset A_n$, with $\sharp R < +\infty$, is defined as

$$\theta_R = \frac{1}{\sharp R} \sum_{x \in R} \theta_{\{x\}}.$$

We are now ready to state the main result of this section, which asserts that if \mathbf{Z}_A verifies conditions $D(u_n(\tau), k_n, l_n)$ and $D''(u_n(\tau), \mathcal{B}_n)$ then, for large n , its spatial extremal index, θ_A , can be viewed as the mean of local extremal indices. This result assumes even greater importance in the following section, where it motivates an estimation procedure for the spatial extremal index θ_A .

Proposition 4.2 Suppose that the sequence of max-stable random fields \mathbf{Z}_A has marginal unit Fréchet distribution and verifies conditions $D(u_n(\tau), k_n, l_n)$ and $D''(u_n(\tau), \mathcal{B}_n)$. If

$$1 - \lambda_n^{(\tau)}(V(x), x) \xrightarrow{n \rightarrow +\infty} \theta_{\{x\}}, \text{ uniformly in } x,$$

then

$$\theta_A = \lim_{n \rightarrow +\infty} \frac{1}{f(n)} \sum_{x \in A_n} \theta_{\{x\}}.$$

The following example illustrate Proposition 4.2 with an anisotropic and non-stationary random field.

Example 4.1 Let $\{Y(x) : x \in \mathbb{Z}^2\}$ be an independent and identically distributed random field with common distribution function $F_Y(y) = \exp\left(\frac{-y^{-1}}{3}\right)$, $y > 0$ and define the random field $\{Z(x), x \in (E \cup F) \times \mathbb{Z}\}$, with $E = \{4k : k \in \mathbb{Z}\}$ and $F = \{4k + 3 : k \in \mathbb{Z}\}$, in the following way

$$Z(x) = \begin{cases} \bigvee_{y \in U_1(x)} Y(y) & , x \in E \times \mathbb{Z} \\ \bigvee_{y \in U_2(x)} Y(y) & , x \in F \times \mathbb{Z}, \end{cases} \quad (4.2)$$

where

$$U_1(x) = \{x, (a_1(\pi_1(x)), a_1(\pi_2(x))), (a_1(\pi_1(x)), a_{-1}(\pi_2(x)))\}$$

and

$$U_2(x) = \{(a_{-1}(\pi_1(x)), a_{-1}(\pi_2(x))), (a_1(\pi_1(x)), \pi_2(x)), (a_{-1}(\pi_1(x)), a_1(\pi_2(x)))\}.$$

The random field $\mathbf{Z} = \{Z(x) : x \in (E \cup F) \times \mathbb{Z}\}$ is anisotropic and non-stationary with common marginal distribution $F(x) = \exp(-x^{-1})$, $x \in (E \cup F) \times \mathbb{Z}$. A simulation of this random field is given in Fig. 5.

Let us consider

$$A_n = (E_n \cup F_n) \times G_n = (\{4k : -n \leq k \leq n\} \cup \{4k+3 : -n \leq k \leq n\}) \times \{-4n, \dots, 4n+3\},$$

and consequently,

$$f(n) = 8(2n+1)^2, \quad f_1(n) = 2(2n+1), \quad f_2(n) = 4(2n+1), \quad \frac{f_i(n)}{f(n)} \xrightarrow{n \rightarrow +\infty} 0, \quad \bigcup_{n \geq 1} A_n = (E \cup F) \times \mathbb{Z}.$$

The sequence $\mathbf{Z}_A = \{Z(x) : x \in A_n\}_{n \geq 1}$ verifies condition $D(u_n, k_n, l_n)$, for all sequences of integer numbers $\{l_n\}_{n \geq 1}$ and $\{k_n\}_{n \geq 1}$ satisfying (2.1), since $\alpha_n(l_n, u_n) = 0$ with $l_n \geq 2$, as well as condition $D''(u_n(\tau), \mathcal{B}_n, \mathcal{V})$ where $\{u_n(\tau)\}_{n \geq 1}$ is a sequence of real numbers satisfying (2.3) and \mathcal{V} is the family of neighborhoods $V(x) = \{(\pi_1(x), a_1(\pi_2(x))), (\pi_1(x), a_2(\pi_2(x)))\}$. In fact, for each $x \in (E_n \times G_n) \cup (F_n \times G_n)$, it holds

$$\begin{aligned} P \left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y), \quad \bigvee_{y \in B_n^{(s,t)} - V(x)} Z(y) > u_n(\tau) \right) \\ = P \left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y), \quad \bigvee_{y \in (B_n^{(s,t)} - V(x)) \wedge \pi_2(y) > \pi_2(x)} Z(y) > u_n(\tau) \right) \\ + P \left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y), \quad \bigvee_{y \in (B_n^{(s,t)} - V(x)) \wedge \pi_2(y) = \pi_2(x)} Z(y) > u_n(\tau) \right) \\ \leq \bar{F}_Z^2(u_n(\tau)) \frac{f(n)}{k_n^2} + \bar{F}_Y^2(u_n(\tau)) \frac{1}{k_n^2}. \end{aligned}$$

Therefore, for any family \mathcal{B}_n in the conditions of the previously established results, we obtain

$$\begin{aligned} k_n^2 \sup_{B_n^{(s,t)}} \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y), \quad \bigvee_{y \in B_n^{(s,t)} - V(x)} Z(y) > u_n(\tau) \right) \\ \leq \frac{(f(n) \bar{F}_Z(u_n(\tau)))^2}{k_n^2} + \frac{(f(n) \bar{F}_Y(u_n(\tau)))^2}{k_n^2} = o(1). \end{aligned}$$

Now, from the definition of the spatial extremal index of \mathbf{Z}_A we get $\theta_A = \frac{1}{2}$, since, for $u_n(\tau) = f(n)/\tau = 8(2n+1)^2/\tau$,

$$\begin{aligned} P\left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau)\right) &= P\left(\bigvee_{x \in E_n \times G_n} Z(x) \leq u_n(\tau)\right) P\left(\bigvee_{x \in F_n \times G_n} Z(x) \leq u_n(\tau)\right) \\ &= (F_Y^3(u_n(\tau)))^{\frac{f(n)}{2}} \xrightarrow{n \rightarrow +\infty} (\exp(-\tau))^{\frac{1}{2}}. \end{aligned}$$

On the other hand, from Proposition 4.2,

$$\theta_A = 1 - \lim_{n \rightarrow +\infty} \frac{1}{f(n)} \sum_{x \in A_n} \frac{P(Z(x) > u_n(\tau), Z(\pi_1(x), a_1(\pi_2(x))) > u_n(\tau) \vee Z(\pi_1(x), a_2(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))}.$$

If $x \in E_n \times G_n$,

$$\begin{aligned} &\frac{P(Z(x) > u_n(\tau), Z(\pi_1(x), a_1(\pi_2(x))) > u_n(\tau) \vee Z(\pi_1(x), a_2(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))} \\ &= \frac{P(Z(x) > u_n(\tau), Z(\pi_1(x), a_1(\pi_2(x))) \leq u_n(\tau), Z(\pi_1(x), a_2(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))} \\ &= \frac{P(Y(a_1(\pi_1(x)), a_1(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))}, \end{aligned}$$

otherwise, if $x \in F_n \times G_n$,

$$\begin{aligned} &\frac{P(Z(x) > u_n(\tau), Z(\pi_1(x), a_1(\pi_2(x))) > u_n(\tau) \vee Z(\pi_1(x), a_2(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))} \\ &= \frac{P(Z(x) > u_n(\tau), Z(\pi_1(x), a_1(\pi_2(x))) > u_n(\tau), Z(\pi_1(x), a_2(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))} \\ &\quad + \frac{P(Z(x) > u_n(\tau), Z(\pi_1(x), a_1(\pi_2(x))) > u_n(\tau), Z(\pi_1(x), a_2(\pi_2(x))) \leq u_n(\tau))}{P(Z(x) > u_n(\tau))} \\ &= \frac{P(Y(a_{-1}(\pi_1(x)), a_1(\pi_2(x))) > u_n(\tau))}{P(Z(x) > u_n(\tau))} + \frac{P(Y(a_{-1}(\pi_1(x)), \pi_2(x)) > u_n(\tau))}{P(Z(x) > u_n(\tau))}. \end{aligned}$$

Thus,

$$\begin{aligned} \theta_A &= 1 - \lim_{n \rightarrow +\infty} \left(\frac{\bar{F}_Y(u_n(\tau))}{\bar{F}_Z(u_n(\tau))} \frac{(2n+1) \times 4(2n+1)}{f(n)} + 2 \frac{\bar{F}_Y(u_n(\tau))}{\bar{F}_Z(u_n(\tau))} \frac{(2n+1) \times 4(2n+1)}{f(n)} \right) \\ &= 1 - \lim_{n \rightarrow +\infty} \left(\frac{f(n) \bar{F}_Y(u_n(\tau))}{f(n) \bar{F}_Z(u_n(\tau))} \times \frac{1}{2} + \frac{2f(n) \bar{F}_Y(u_n(\tau))}{f(n) \bar{F}_Z(u_n(\tau))} \times \frac{1}{2} \right) \\ &= 1 - \left(\frac{\tau/3}{\tau} \times \frac{1}{2} + \frac{2\tau/3}{\tau} \times \frac{1}{2} \right) = \frac{1}{2}, \end{aligned}$$

as expected.

Although, in practical applications the conditions $D(u_n(\tau), k_n, l_n)$ and $D''(u_n(\tau), \mathcal{B}_n)$ are not easy to verify, the results of this section highlight the importance of $\theta_{\{x\}}$ in the study of locally occurring large observations in clusters. In a

region $R \subseteq A_n$, the smaller the values of $\theta_{\{x\}}$, $x \in R$, the greater the propensity for clustering.

Note that, beyond the interpretation of the inverse proportionality between the value of θ_A and the propensity for clustering, small values of $\theta_{\{x\}}$ indicate a strong dependence between $Z(x)$ and $\{Z(y), y \in V(x)\}$.

5 Estimation of the spatial extremal index

For max-stable random fields, verifying the mixing conditions previously defined, Proposition 4.2 enables the estimation of the spatial extremal index θ_A from the local extremal indices $\theta_{\{x\}}$, $x \in A_n$, since, for n large and integer,

$$\theta_A \approx \frac{1}{f(n)} \sum_{x \in A_n} \theta_{\{x\}}. \quad (5.1)$$

By plugging estimators of the local extremal indices, $\hat{\theta}_{\{x\}}$, $x \in A_n$, into Eq. 5.1, one can generate an estimator for θ_A .

It is now clear that the estimation of the spatial extremal index basically sums up to the estimation of the local extremal indices $\theta_{\{x\}}$, $x \in A_n$, which, due to Propositions 2.1 and 2.2, are simply the difference between the extremal coefficients at $V(x) \cup \{x\}$ and $V(x)$, where $V(x)$ is a neighbourhood of $x \in A_n$.

Ferreira and Ferreira (2012) proposed a nonparametric estimator for the extremal coefficient based on the following relation

$$\epsilon_C = \frac{E(M(C))}{1 - E(M(C))}, \quad \text{with} \quad M(C) = \bigvee_{x \in C} F(Z(x)),$$

that holds from the fact that $\epsilon_C = l(1, \dots, 1)$, where $l(1, \dots, 1)$ denotes the $\sharp C$ -variate stable tail dependence function evaluated at point $(1, \dots, 1)$. This estimator naturally considers sample means, to estimate the previous expectations, and is defined as

$$\hat{\epsilon}_C = \frac{\overline{M(C)}}{1 - \overline{M(C)}} \quad (5.2)$$

where $\overline{M(C)}$ denotes the sample mean,

$$\overline{M(C)} = \frac{1}{k} \sum_{i=1}^k \bigvee_{y \in C} \hat{F}(Z^{(i)}(y))$$

and \hat{F} , is the (modified) empirical distribution function of F ,

$$\hat{F}(u) = \frac{1}{k+1} \sum_{i=1}^k \mathbb{I}_{\{Z^{(i)}(y) \leq u\}},$$

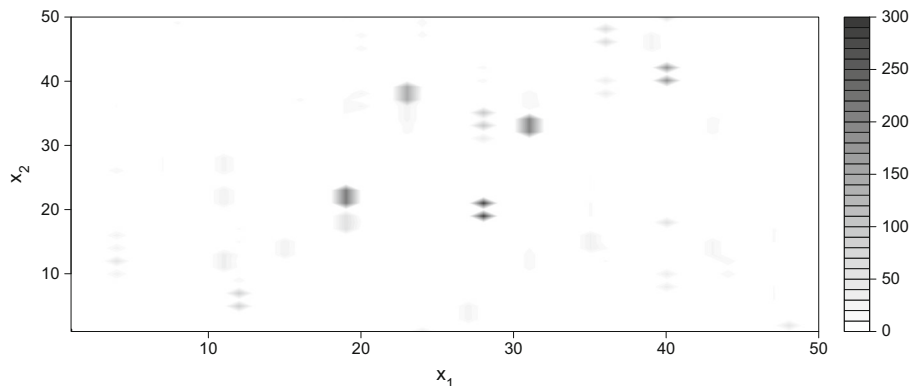
where $Z^{(i)}(y)$, $i = 1, \dots, k$, are k independent replications of $Z(y)$.

Table 1 Estimated mean values and mean square errors of estimator $\hat{\theta}_A$ defined in Eq. 5.2, for the random field of Example 4.1

$f(n)$	k					
	100		500		1000	
	$E[\bullet]$	$MSE[\bullet]$	$E[\bullet]$	$MSE[\bullet]$	$E[\bullet]$	$MSE[\bullet]$
$f(1)=72$	0.5247	6.9e-4	0.5122	1.7e-4	0.5101	1.2e-4
$f(10)=3528$	0.5240	5.8e-4	0.5130	1.7e-4	0.5096	9.3e-5
$f(20)=13448$	0.5238	5.7e-4	0.5129	1.7e-4	0.5095	9.0e-5
$f(30)=29768$	0.5238	5.7e-4	0.5129	1.7e-4	0.5095	9.0e-5

Properties such as consistency and asymptotic normality hold for estimator (5.2), even in this case of unknown marginal distribution F , as shown in Ferreira and Ferreira (2012).

Other estimators for the extremal coefficient have been proposed since Smith (1990) presented his *raw estimates* of the extremal coefficients for 2 variables. Nevertheless, most studies only focus on the estimation of the pairwise dependence structure of stationary and isotropic max-stable random fields, that is the extremal coefficient function $\epsilon(h)$, where h is the distance between two locations, and exploit its relation with known dependence measures of geostatistics, namely the normalized madogram (see Schlather and Tawn 2003; Cooley et al. 2006 and references therein). This is justified by the fact that characterizing dependence in higher orders quickly becomes unwieldy. Schlather and Tawn (2003) also estimate the multivariate extremal coefficient, they use a censored likelihood approach and propose different strategies for building self-consistent estimators in this context, for which numerical solutions can be given in relatively small dimensions.

**Fig. 5** Simulation of the random field $Z(x)$ defined in Eq. 4.2

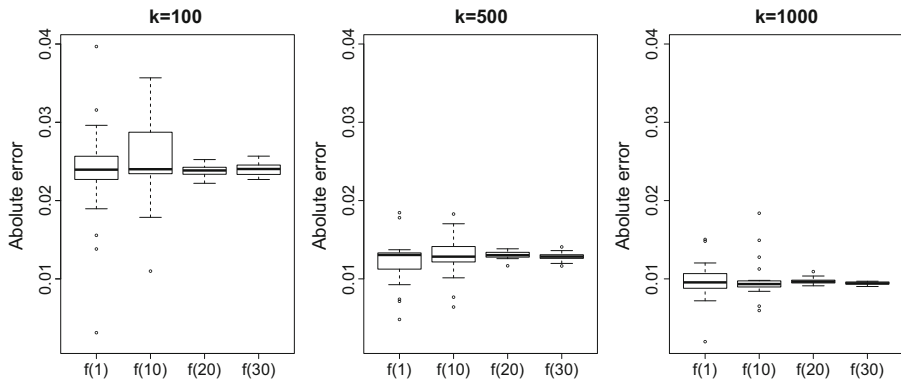


Fig. 6 Box plots of the absolute errors of 20 estimates of the spatial extremal index obtained with estimator (5.2) for $k = 100, 500, 1000$ and $f(n)$, $n = 1, 10, 20, 30$, when considering the random field of Example 4.1

Several parametric and semi-parametric estimators for the multivariate stable tail dependence function can be found in the literature, and since the extremal coefficient is related to this function, as previously mentioned, they can also be viewed as estimators for ϵ_C . For a survey, see Krajina (2010). To illustrate the usefulness of Proposition 4.2 we shall consider in our simulation study Ferreira and Ferreira's simple nonparametric estimator (5.2), that just evolves sample means, to estimate the extremal coefficients and consequently estimate the local extremal indices. A full comparison with other estimators of the extremal index in the literature is available in Ferreira (2015).

We then propose the following estimator the spatial extremal index θ_A

$$\hat{\theta}_A = \frac{1}{f(n)} \sum_{x \in A_n} \hat{\epsilon}_{\{x\} \cup V(x)} - \hat{\epsilon}_{V(x)} \quad , x \in A_n, \quad (5.3)$$

where $\hat{\epsilon}_\bullet$ are the estimators given in Eq. 5.2.

Estimator (5.3) of the spatial extremal index is consistent, given the consistency of the estimators $\hat{\epsilon}_{\{x\} \cup V(x)}$ and $\hat{\epsilon}_{V(x)}$, as previously stated. We analyze, in what follows, its finite sample behaviour on simulated data from the anisotropic and non-stationary random field $\{Z(x), x \in (E \cup F) \times \mathbb{Z}\}$ considered in Example 4.1. For each simulation experiment $k = 100, 500$ and 1000 independent random fields were simulated, each experiment was repeated 20 times and the mean (E) together with the mean square error (MSE) calculated.

The results from the experiments are shown in Table 1 and in Fig. 6 where the box plots of the 20 absolute errors for $k = 100, 500, 1000$ in the four groups defined by $f(n)$, $n = 1, 10, 20, 30$ are plotted.

As we can see from the values reported in Table 1, the estimator $\hat{\theta}_A$ has quite a good performance, with biases around 0.02 for small values of k and around 0.01 for bigger values of k . The values of n considered have a small effect on the bias, nevertheless the variance decreases with n and with k , as well as the absolute errors.

Despite the good performance of the estimator (5.3), for the spatial extremal index θ_A , and its ease of implementation, it has the drawback of the validation of the mixing conditions, which can be cumbersome to verify.

6 Conclusion

In this paper we establish existence criteria for the extremal index of a nonstationary and anisotropic random field, defined on \mathbb{R}^2 . Under restrictions on the local path behavior of exceedances, that allow clustering of high values, we obtain the extremal index as the limit of a sequence of upper tail dependence coefficients. For the particular case of max-stable random fields, we prove that the extremal index can be obtained as a function of extremal dependence coefficients. Based on this relation we give a simple estimator of the extremal index and we analyze its performance with an anisotropic and nonstationary 1-dependent random field. The simulation study results show the good performance of the proposed estimator.

Acknowledgments We are grateful to the referees for their detailed comments and important suggestions which significantly improved this paper. This research was supported by National Foundation of Science and Technology through UID/MAT/00212/2013.

Appendix A: Auxiliary results and proofs for section 2

Lemma A.1 *Suppose that the sequence \mathbf{Z}_A satisfies condition $D(u_n, k_n, l_n)$ for a sequence of real numbers $\{u_n\}_{n \geq 1}$ such that*

$$\{f(n)\bar{F}(u_n)\}_{n \geq 1} \text{ is bounded.}$$

If $I^{(s)}$, $s = 1, \dots, k_n$, are disjoint subsets of $\pi_1(A_n)$ and $J^{(t)}$, $t = 1, \dots, k_n$, are disjoint subsets of $\pi_2(A_n)$, and

$$B_n^{(s,t)} = \pi_2^{-1}(J^{(t)}) \cap \pi_1^{-1}(I^{(s)}) \cap A_n, \quad s, t = 1, \dots, k_n,$$

then

$$\left| P \left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n \right) - \prod_{s,t} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n \right) \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Proof If all the subsets $I^{(s)}$ and $J^{(t)}$, $s, t = 1, \dots, k_n$, have less than l_n elements, the result is trivial, since both terms in the difference tend to 1. On the other hand, if some $I^{(s_i)}$ has less than l_n consecutive elements of $\pi_1(A_n)$, we can eliminate $B_n^{(s_i,t)}$

in the family $\{B_n^{(s,t)} : s, t = 1, \dots, k_n\}$. Indeed, since $\sharp(\pi_1^{-1}(a) \cap A_n) \leq f_2(n)$, it follows that

$$\begin{aligned} & P\left(\bigvee_{x \in \bigcup_{s,t,s \neq s_i} B_n^{(s,t)}} Z(x) \leq u_n\right) - P\left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n\right) \\ & \leq P\left(\bigvee_{x \in \bigcup_t B_n^{(s_i,t)}} Z(x) > u_n\right) \\ & \leq l_n \bigvee_{a \in I^{(s_i)}} \frac{\sharp(\pi_1^{-1}(a) \cap A_n)}{f(n)} f(n) \bar{F}(u_n) = o\left(\frac{1}{k_n}\right) \end{aligned}$$

and, by applying the inequality $|\prod_i a_i - \prod_i b_i| \leq \sum_i |a_i - b_i|$, we have

$$\begin{aligned} & \prod_{s,t,s \neq s_i} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n\right) - \prod_{s,t} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n\right) \\ & = \prod_{s,t,s \neq s_i} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n\right) \left(1 - \prod_{t=1}^{k_n} P\left(\bigvee_{x \in B_n^{(s_i,t)}} Z(x) \leq u_n\right)\right) \\ & \leq 1 - \prod_{t=1}^{k_n} P\left(\bigvee_{x \in B_n^{(s_i,t)}} Z(x) \leq u_n\right) \\ & \leq \sum_{t=1}^{k_n} P\left(\bigvee_{x \in B_n^{(s_i,t)}} Z(x) > u_n\right) \\ & \leq k_n l_n \bigvee_{a \in I^{(s_i)}} \frac{\sharp(\pi_1^{-1}(a) \cap A_n)}{f(n)} f(n) \bar{F}(u_n) = o(1). \end{aligned}$$

With similar arguments, we can eliminate in the family $\{B_n^{(s,t)}, s, t = 1, \dots, k_n\}$ the subsets $J^{(t_i)}$, $t_i \in \{1, \dots, k_n\}$, that have less than l_n elements.

To conclude, let us then assume that all the subsets $I^{(s)}$ and $J^{(t)}$, $s, t = 1, 2, \dots, k_n$, have more than l_n elements. Start by eliminating in each $I^{(s)}$ and $J^{(t)}$, respectively, the sets $I^{*(s)}$ and $J^{*(t)}$ with the highest l_n values. The resulting sets $\bar{I}^{(s)}$, $s = 1, \dots, k_n$, belong to $\mathcal{S}(\pi_1(A_n), l_n)$, $s = 1, \dots, k_n$, and $\bar{J}^{(t)}$, $t = 1, \dots, k_n$, belong to $\mathcal{S}(\pi_2(A_n), l_n)$. Now consider

$$\bar{B}_n^{(s,t)} = \pi_1^{-1}(\bar{I}^{(s)}) \cap \pi_2^{-1}(\bar{J}^{(t)}) \cap A_n,$$

and note that

$$\begin{aligned}
 & \left| P \left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n \right) - \prod_{s,t} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n \right) \right| \\
 & \leq \left| P \left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n \right) - P \left(\bigvee_{x \in \bigcup_{s,t} \bar{B}^{(s,t)}} Z(x) \leq u_n \right) \right| \\
 & + \left| P \left(\bigvee_{x \in \bigcup_{s,t} \bar{B}^{(s,t)}} Z(x) \leq u_n \right) - \prod_{s,t} P \left(\bigvee_{x \in \bar{B}^{(s,t)}} Z(x) \leq u_n \right) \right| \\
 & + \left| \prod_{s,t} P \left(\bigvee_{x \in \bar{B}^{(s,t)}} Z(x) \leq u_n \right) - \prod_{s,t} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n \right) \right|, \\
 & = o(1) + k_n^2 \alpha(l_n, u_n) = o(1),
 \end{aligned}$$

since

$$P \left(\bigvee_{x \in \bigcup_{s,t} \bar{B}_n^{(s,t)}} Z(x) \leq u_n \right) - P \left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n \right) \leq \bigvee_{i=1}^2 2k_n l_n \frac{f_i(n)}{f(n)} f(n) \bar{F}(u_n) = o(1),$$

condition $D(u_n, k_n, l_n)$ holds for \mathbf{Z}_A and

$$\begin{aligned}
 & \prod_{s,t} P \left(\bigvee_{x \in \bar{B}_n^{(s,t)}} Z(x) \leq u_n \right) - \prod_{s,t} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n \right) \\
 & \leq \sum_s \sum_t P \left(\bigvee_{x \in \bar{B}_n^{(s,t)}} Z(x) \leq u_n \right) - P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n \right) \\
 & \leq \sum_s \sum_t \left(l_n \sharp \pi_2(B_n^{(s,t)}) + l_n \sharp \pi_1(B_n^{(s,t)}) \right) \bar{F}(u_n) \\
 & \leq l_n \bar{F}(u_n) (k_n f_1(n) + k_n f_2(n)) \\
 & \leq \bigvee_{i=1}^2 2k_n l_n \frac{f_i(n)}{f(n)} f(n) \bar{F}(u_n) = o(1).
 \end{aligned}$$

□

Lemma A.2 *If the sequence \mathbf{Z}_A satisfies condition $D(u_n, k_n, l_n)$, with $\{u_n\}_{n \geq 1}$ verifying (2.2) then, for each n , there exists a family \mathcal{B}_n of k_n^2 disjoint subsets of A_n , $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, with $\sharp B_n^{(s,t)} \sim \frac{f(n)}{k_n^2}$ and such that*

$$P\left(\bigvee_{x \in A_n} Z(x) \leq u_n\right) - P\left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n\right) \xrightarrow{n \rightarrow +\infty} 0.$$

Proof First, we prove that the disjoint subsets $B_n^{(s,t)}$, $s, t = 1, \dots, k_n$, constructed through the method described and illustrated in Section 2, have approximately $\frac{f(n)}{k_n^2}$ elements.

1. Let us consider, in $\pi_1(A_n)$, the set $I^{(1)}$ of the first elements maximally constructed such that

$$\sum_{x \in \pi_1^{-1}(I^{(1)}) \cap A_n} P(Z(x) > u_n) \leq \frac{1}{k_n} \sum_{x \in A_n} P(Z(x) > u_n),$$

and in $\pi_1(A_n) - I^{(1)}$ the maximal set, $I^{(2)}$, of the first elements such that

$$\sum_{x \in \pi_1^{-1}(I^{(2)}) \cap A_n} P(Z(x) > u_n) \leq \frac{1}{k_n} \sum_{x \in A_n} P(Z(x) > u_n).$$

Similarly, we obtain k_n disjoint subsets of A_n ,

$$\pi_1^{-1}(I^{(1)}) \cap A_n, \dots, \pi_1^{-1}(I^{(k_n)}) \cap A_n.$$

2. For each of the previous subsets, let us consider an analogous decomposition using projection π_2 . Hence, we consider the set $J^{(s,1)}$, in $\pi_2(A_n)$, that consists of the first elements maximally chosen such that

$$\sum_{x \in \pi_2^{-1}(J^{(s,1)}) \cap \pi_1^{-1}(I^{(s)}) \cap A_n} P(Z(x) > u_n) \leq \frac{1}{k_n^2} \sum_{x \in A_n} P(Z(x) > u_n).$$

Using the same technique, we obtain the following k_n subsets of $\pi_1^{-1}(I^{(s)}) \cap A_n$,

$$B_n^{(s,t)} = \pi_2^{-1}(J^{(s,t)}) \cap \pi_1^{-1}(I^{(s)}) \cap A_n, \quad t = 1, \dots, k_n.$$

Next, we will prove that $\sharp B_n^{(s,t)} \sim \frac{f(n)}{k_n^2}$. Since

$$\sum_{x \in \pi_2^{-1}(J^{(s,t)}) \cap \pi_1^{-1}(I^{(s)}) \cap A_n} P(Z(x) > u_n) \leq \frac{1}{k_n^2} \sum_{x \in A_n} P(Z(x) > u_n)$$

we obtain

$$\sharp B_n^{(s,t)} \leq \frac{1}{k_n^2} f(n),$$

and, from the maximality criteria used in the construction of $J^{(s,t)}$, it follows that

$$\sharp B_n^{(s,t)} \bar{F}(u_n) + o\left(\frac{1}{k_n l_n}\right) > \frac{1}{k_n^2} f(n) \bar{F}(u_n).$$

Therefore,

$$1 \geq \frac{\sharp B_n^{(s,t)}}{\frac{f(n)}{k_n^2}} \geq 1 - \frac{k_n^2}{f(n)} o\left(\frac{1}{k_n l_n}\right),$$

which allows us to conclude that $\sharp B_n^{(s,t)} \sim \frac{f(n)}{k_n^2}$, since $\frac{k_n^2}{f(n)} \xrightarrow{n \rightarrow +\infty} 0$.

Now,

$$\begin{aligned} 0 &\leq P\left(\bigvee_{x \in A_n} Z(x) \leq u_n\right) - P\left(\bigvee_{x \in \bigcup_{s,t} B_n^{(s,t)}} Z(x) \leq u_n\right) \\ &\leq \sum_{x \in A_n - \bigcup_{s,t} B_n^{(s,t)}} P(Z(x) > u_n) \\ &= \frac{\sharp(A_n - \bigcup_{s,t} B_n^{(s,t)})}{f(n)} f(n) \bar{F}(u_n) \end{aligned}$$

and

$$\begin{aligned} \frac{f(n) - \sharp \bigcup_{s,t} B_n^{(s,t)}}{f(n)} &= \frac{f(n) - \sharp \sum_{s,t} B_n^{(s,t)}}{f(n)} \\ &= 1 - \frac{1}{k_n^2} \sum_{s,t} \frac{\sharp B_n^{(s,t)}}{\frac{f(n)}{k_n^2}} \leq 1 - \bigwedge_{s,t} \frac{\sharp B_n^{(s,t)}}{\frac{f(n)}{k_n^2}} = o(1), \end{aligned}$$

which concludes the proof. \square

Proof of Proposition 2.2 From Proposition 2.1, we have

$$\begin{aligned} &P\left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau)\right) \\ &= \prod_{B_n^{(s,t)} \in \mathcal{B}_n} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) \leq u_n(\tau)\right) + o(1) \\ &= \exp\left\{-\sum_{B_n^{(s,t)} \in \mathcal{B}_n} \frac{-\ln\left(1 - P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau)\right)\right)}{P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau)\right)} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau)\right)\right\} + o(1). \end{aligned}$$

Then, if $\lim_{n \rightarrow +\infty} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau) \right)$ exists, the result follows from the definition of θ_A . \square

Proof of Proposition 2.3 Follows from Proposition 2.2, since

$$\begin{aligned} \theta_A &= \lim_{n \rightarrow +\infty} \frac{1}{f(n) \overline{F}(u_n(\tau))} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau) \right) \\ &= \lim_{n \rightarrow +\infty} \frac{1}{k_n^2} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} \frac{P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau) \right)}{\frac{f(n)}{k_n^2} \overline{F}(u_n(\tau))} \\ &= \lim_{n \rightarrow +\infty} \frac{1}{k_n^2} \sum_{B_n^{(s,t)} \in \mathcal{B}_n} \frac{P \left(\sum_{x \in B_n^{(s,t)}} \mathbb{I}_{\{Z(x) > u_n(\tau)\}} > 0 \right)}{E \left(\sum_{x \in B_n^{(s,t)}} \mathbb{I}_{\{Z(x) > u_n(\tau)\}} \right)}. \end{aligned}$$

\square

Appendix B: auxiliary results and proofs for section 3

Lemma B.1 Let $\{u_n\}_{n \geq 1}$ be a sequence of real numbers and suppose that sequence Z_A satisfies condition $D''(u_n, \mathcal{B}_n, \mathcal{V})$. Then, for each $B_n^{(s,t)} \in \mathcal{B}_n$, we have

$$P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n \right) = \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n > \bigvee_{y \in V(x)} Z(y) \right) + o \left(\frac{1}{k_n^2} \right).$$

Proof We can write

$$\begin{aligned} &P \left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n \right) \\ &= P \left(\bigcup_{x \in B_n^{(s,t)}} \left\{ Z(x) > u_n \geq \bigvee_{\{y \in B_n^{(s,t)} : \pi_1(x) > \pi_1(y) \vee (\pi_1(y) = \pi_1(x) \wedge \pi_2(y) > \pi_2(x))\}} Z(y) \right\} \right) \\ &= \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n \geq \bigvee_{\{y \in B_n^{(s,t)} : \pi_1(x) > \pi_1(y) \vee (\pi_1(y) = \pi_1(x) \wedge \pi_2(y) > \pi_2(x))\}} Z(y) \right) \\ &= \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n \geq \bigvee_{y \in V(x)} Z(y) \right) + o \left(\frac{1}{k_n^2} \right), \end{aligned}$$

since, by $D''(u_n, \mathcal{B}_n, \mathcal{V}_{E,N}^{p,q,r})$ -condition, we have

$$\begin{aligned} & \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n \geq \bigvee_{y \in V(x)} Z(y) \right) \\ & - \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n \geq \bigvee_{\{y \in B_n^{(s,t)} : \pi_1(x) > \pi_1(y) \vee (\pi_1(y) = \pi_1(x) \wedge \pi_2(y) > \pi_2(x))\}} Z(y) \right) \\ & \leq \sup_{B_n^{(s,t)} \in \mathcal{B}_n} \sum_{x \in B_n^{(s,t)}} P \left(Z(x) > u_n \geq \bigvee_{y \in V(x)} Z(y), \bigvee_{y \in B_n^{(s,t)} - V(x)} Z(y) > u_n \right) = o\left(\frac{1}{k_n^2}\right). \end{aligned}$$

□

Remark 3 In the previous proof $\left\{ \bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n \right\}$ was decomposed as the union of events $\{Z(x) > u_n\}$, $x \in B_n^{(s,t)}$, where x is the location with the highest coordinates. Other decompositions could have been considered, for example, considering the events $\{Z(x) > u_n\}$ where $x \in B_n^{(s,t)}$ is the location with the lowest abscissa and biggest ordinate, would lead to the following decomposition

$$\left\{ \bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n \right\} = \bigcup_{x \in B_n^{(s,t)}} \left\{ Z(x) > u_n \geq \bigvee_{\{y \in B_n^{(s,t)} : \pi_1(y) < \pi_1(x) \vee (\pi_1(y) = \pi_1(x) \wedge \pi_2(y) > \pi_2(x))\}} Z(y) \right\}.$$

Now, denoting by $\mathcal{V}_{W,N}^{p,q,r}$, $p, q, r \in \mathbb{Z}$, the family of neighborhoods

$$\begin{aligned} V(x) = \{y : (a_{-p}(\pi_1(x)) \leq \pi_1(y) \leq a_1(\pi_1(x)) \wedge a_{-q}(\pi_2(x)) \leq \pi_2(y) \leq a_r(\pi_2(x))) \\ \vee (\pi_1(y) = \pi_1(x) \wedge a_1(\pi_2(y)) \leq \pi_2(y) \leq a_r(\pi_2(y)))\} \end{aligned}$$

we obtain an analogous result to Lemma 3.1, if we assume condition $D''(u_n, \mathcal{B}_n, \mathcal{V}_{W,N}^{p,q,r})$.

In fact, we can decompose the event $\left\{ \bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau) \right\}$ in eight different ways corresponding to the different forms of starting from x and getting around \mathbb{Z}^2 , along the directions $\{(a_k(\pi_1(x)), 0), k \in \{-1, 1\}\}$ and $\{(0, a_k(\pi_2(x))), k \in \{-1, 1\}\}$.

Proof of Proposition 3.1 From Lemma B.1 and the definition of the tail dependence coefficient $\lambda_n^{(\tau)}(V(x), x)$, it follows that

$$\begin{aligned}
 & P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau)\right) \\
 &= \sum_{x \in B_n^{(s,t)}} P\left(Z(x) > u_n(\tau) \geq \bigvee_{y \in V(x)} Z(y)\right) + o\left(\frac{1}{k_n^2}\right) \\
 &= \bar{F}(u_n(\tau)) \sum_{x \in B_n^{(s,t)}} P\left(\bigvee_{y \in V(x)} Z(y) \leq u_n(\tau) \mid Z(x) > u_n(\tau)\right) + o\left(\frac{1}{k_n^2}\right) \\
 &= \bar{F}(u_n) \sum_{x \in B_n^{(s,t)}} (1 - \lambda_n^{(\tau)}(V(x), x)) + o\left(\frac{1}{k_n^2}\right).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 & P\left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau)\right) \\
 &= \exp\left(-\sum_{B_n^{(s,t)} \in \mathcal{B}_n} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau)\right)\right) + o(1) \\
 &= \exp\left(-\sum_{B_n^{(s,t)} \in \mathcal{B}_n} \left(\bar{F}(u_n(\tau)) \sum_{x \in B_n^{(s,t)}} (1 - \lambda_n^{(\tau)}(V(x), x)) + o\left(\frac{1}{k_n^2} + o(1)\right)\right)\right) \\
 &= \exp\left(-\bar{F}(u_n(\tau)) \sum_{B_n^{(s,t)} \in \mathcal{B}_n} \sum_{x \in B_n^{(s,t)}} (1 - \lambda_n^{(\tau)}(V(x), x))\right) + o(1) \\
 &= \exp\left(-\bar{F}(u_n(\tau)) \sum_{x \in A_n} (1 - \lambda_n^{(\tau)}(V(x), x))\right) + o(1) \\
 &\xrightarrow{n \rightarrow +\infty} \exp(-\tau(1 - \lambda_A)),
 \end{aligned}$$

which concludes the proof. \square

Proof of Proposition 3.2 From Proposition 2.1 and condition $D'(u_n(\tau), \mathcal{B}_n)$ we have

$$\begin{aligned}
 & P\left(\bigvee_{x \in A_n} Z(x) \leq u_n(\tau)\right) \\
 &= \exp\left(-\sum_{B_n^{(s,t)} \in \mathcal{B}_n} P\left(\bigvee_{x \in B_n^{(s,t)}} Z(x) > u_n(\tau)\right)\right) + o(1) \\
 &= \exp\left(-\sum_{B_n^{(s,t)} \in \mathcal{B}_n} \sum_{x \in B_n^{(s,t)}} P(Z(x) > u_n(\tau)) + o(1)\right) + o(1) \\
 &= \exp(-f(n)P(Z(x) > u_n(\tau))o(1)) + o(1) \xrightarrow{n \rightarrow +\infty} \exp(-\tau),
 \end{aligned}$$

which proves the result. \square

Appendix C: proofs for section 4

Proof of Proposition 4.1 Since $F(y) = \exp(-y^{-1})$, then $u_n(\tau) = \frac{f(n)}{\tau}$. So, we have

$$\begin{aligned}
 & \lim_{n \rightarrow +\infty} \lambda_n^{(\tau)}(V(x), x) \\
 &= \lim_{n \rightarrow +\infty} P\left(\bigvee_{y \in V(x)} Z(y) > \frac{f(n)}{\tau} \middle| Z(y) > \frac{f(n)}{\tau}\right) \\
 &= \lim_{n \rightarrow +\infty} P\left(F\left(\bigvee_{y \in V(x)} Z(y)\right) > F\left(\frac{f(n)}{\tau}\right) \middle| F(Z(x)) > F\left(\frac{f(n)}{\tau}\right)\right) \\
 &= \lim_{n \rightarrow +\infty} P\left(\bigvee_{y \in V(x)} F(Z(y)) > 1 - \frac{\tau}{f(n)} \middle| F(Z(x)) > 1 - \frac{\tau}{f(n)}\right) \\
 &= \Lambda_U^{(V(x)|\{x\})} \\
 &= 1 + \epsilon_{V(x)} - \epsilon_{V(x) \cup \{x\}}
 \end{aligned}$$

\square

Proof of Proposition 4.2 Results from the fact that $\{1 - \lambda_n^{(\tau)}(V(x), x)\}_{n \geq 1}$ converges to $\theta_{\{x\}}$, uniformly in x , since this implies that for every $\epsilon > 0$, there exists a natural

number $m = m(\epsilon)$ such that, for every $n > m$,

$$\begin{aligned} \frac{1}{f(n)} \sum_{x \in A_n} (1 - \lambda_n^{(\tau)}(V(x), x)) &\leq \frac{1}{f(n)} \sum_{x \in A_n} (\theta_{\{x\}} + \epsilon) \\ &= \left(\frac{1}{f(n)} \sum_{x \in A_n} \theta_{\{x\}} \right) + \epsilon \end{aligned}$$

and

$$\begin{aligned} \frac{1}{f(n)} \sum_{x \in A_n} (1 - \lambda_n^{(\tau)}(V(x), x)) &\geq \frac{1}{f(n)} \sum_{x \in A_n} (\theta_{\{x\}} - \epsilon) \\ &= \left(\frac{1}{f(n)} \sum_{x \in A_n} \theta_{\{x\}} \right) - \epsilon. \end{aligned}$$

□

References

- Adler, R.A.: The Geometry of Random Fields. Wiley, New York (1981)
- Berman, S.: Sojourns and Extremes of Stochastic Processes. Taylor & Francis Ltd, USA (1992)
- Choi, H.: Almost sure limit theorem for stationary Gaussian random fields. *J. Korean Stat. Soc.* **39**, 475–482 (2010)
- Cooley, D., Naveau, P., Poncet, P.: Variograms for spatial max-stable random fields. *Lect. Notes Stat.* **187**, 373–390 (2006)
- de Haan, L.: A spectral representation for max-stable processes. *Ann. Probab.* **12**, 1194–1204 (1984)
- Ferreira, M.: Estimating the extremal index through the tail dependence concept. *Discussiones Mathematicae - Probab. Stat.* **35**, 61–74 (2015)
- Ferreira, H., Pereira, L.: How to compute the extremal index of stationary random fields. *Stat. Probab. Lett.* **78**(11), 1301–1304 (2008)
- Ferreira, M., Ferreira, H.: On extremal dependence of block vectors. *Kybernetika* **48**(5), 988–1006 (2012)
- Ferreira, H., Pereira, L.: Point processes of exceedances by random fields. *J. Stat. Plan. Infer.* **142**(3), 773–779 (2012)
- Krajina, A.: An M-Estimator of Multivariate Dependence Concepts. PhD Thesis. Tilburg, Tilburg University Press (2010)
- Leadbetter, M.R., Lindgren, G., Rootzén, H.: Extremes and Related Properties of Random Sequences and Processes. Springer, Berlin (1983)
- Leadbetter, M.R., Nandagopalan, S.: On exceedance point process for stationary sequences under mild oscillation restrictions. In: Hüslér, J., Reiss, D. (eds.) *Extreme Value Theory: Proceedings, Oberwolfach*, vol. 1987, pp. 69–80. Springer, New York (1988)
- Leadbetter, M.R., Rootzén, H.: On Extreme Values in Stationary Random Fields. *Stochastic Processes and Related Topics*, 275–285 Trends Math. Birkhauser, Boston (1998)
- Li, H.: Orthant tail dependence of multivariate extreme value distributions. *J. Multivar. Anal.* **100**(1), 243–256 (2009)
- Pereira, L., Ferreira, H.: Limiting crossing probabilities of random fields. *J. Appl. Probab.* **3**, 884–891 (2006)

- Pereira, L.: On the extremal behavior of a non-stationary normal random field. *J. Stat. Plan. Infer.* **140**(11), 3567–3576 (2010)
- Piterbarg, V.I.: *Asymptotics Methods in Theory of Gaussian Processes and Fields*. Translations of Mathematical Monographs, vol. 48. American Mathematical Society (1996)
- Schlather, M., Tawn, J.: A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90**, 139–156 (2003)
- Sibuya, M.: Bivariate extreme statistics. *Ann. Inst. Stat. Math.* **11**(2), 195–210 (1959)
- Smith, R.L.: *Max-Stable Processes and Spatial Extremes*. Preprint Univ North Carolina, USA (1990)
- Oliveira, J.T.: Structure theory of bivariate extremes: Extensions. *Estudos de Matemática, Estatística e Econometria* **7**, 165–195 (1962/63)