METMA V International Workshop on Spatio-Temporal Modelling
Santiago de Compostela
30 Jun - 2 Jul 2010

Asymptotic joint distributions of locations of largest values of a stationary random field

Pereira, L.[1]

[1] *Departamento de Matemática, Universidade da Beira Interior, Portugal; lpereira@ubi.pt*

Abstract. *The study of the relationship between extreme values of dependent random fields and their locations has important practical applications, for instance, when dealing with censored data. In this paper we study the limiting distribution of the joint locations of the largest order statistics generated by a stationary random field with extremal index as well as the asymptotic behavior of the location, of a high level exceedance, nearest of the origin and the location of the maximum.*

Keywords. *Random field; Location of extremes; Point process of exceedances.*

1 Introduction

Let $\mathbf{X}=\{X_{\mathbf{n}}\}_{\mathbf{n}\geq\mathbf{1}}$ be a random field on $\mathbb{N}^2$, where $\mathbb{N}$ is the set of all positive integers. For $\mathbf{i}=(i_1,i_2)$ and $\mathbf{j}=(j_1,j_2)$ the inequality $\mathbf{i}\leq\mathbf{j}$ means $i_s\leq j_s,\ s=1,2,$ and $\mathbf{n}=(n_1,n_2)\to\infty$ means $n_s\to\infty$, $s=1,2$. For a family of real levels $\{u_{\mathbf{n}}\}_{\mathbf{n}\geq\mathbf{1}}$ and a subset $\mathbf{I}$ of the rectangle of points $\mathbf{R_n}=\{1,\ldots,n_1\}\times\{1,\ldots,n_2\}$, we will denote the event $\{X_{\mathbf{i}}\leq u_{\mathbf{n}}:\mathbf{i}\in\mathbf{I}\}$ by $\{M_{\mathbf{n}}^{(1)}(\mathbf{I})\leq u_{\mathbf{n}}\}$ or simply by $\{M_{\mathbf{n}}^{(1)}\leq u_{\mathbf{n}}\}$ when $\mathbf{I}=\mathbf{R_n}$. For each $i=1,2$, we say the pair $\mathbf{I}\subset\mathbb{N}^2$ and $\mathbf{J}\subset\mathbb{N}^2$ is in $S_i(l)$ if the distance between $\Pi_i(\mathbf{I})$ and $\Pi_i(\mathbf{J})$ is greater or equal to l, where $\Pi_i, i=1,2$, denote the cartesian projections. The distance between sets $\mathbf{I}$ and $\mathbf{J}$ of $\mathbb{N}^2$, is the minimum of distances $d(\mathbf{i},\mathbf{j})=\max\{|i_s-j_s|:s\in\{1,2\}\}$, $\mathbf{i}\in\mathbf{I}$ and $\mathbf{j}\in\mathbf{J}$. We shall assume that $\mathbf{X}$ is a stationary random field and that there are sequences of constants $\{a_{\mathbf{n}}>0\}_{\mathbf{n}\geq\mathbf{1}}$ and $\{b_{\mathbf{n}}\}_{\mathbf{n}\geq\mathbf{1}}$ such that, as $\mathbf{n}\longrightarrow\infty$, for each $x\in\mathbb{R}$,

$$P(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}}^{(1)}-b_{\mathbf{n}})\leq x)\to H(x), \tag{1}$$

where H is a nondegenerate distribution function.

If $\mathbf{X}$ is a random field of independent and identically distributed random variables or if it satisfies the coordinatewise-mixing condition (Cw-mixing) from Leadbetter and Rootzén [2], then $\mathbf{X}$ verifies the

Extremal Types Theorem, *id est*, H is Gumbel, Weibull or a Fréchet distribution. However, in the last case the rate of convergence is slower than in the i.i.d. case and can be quantified using the extremal index, θ, of the random field.

Accordingly Choi [1], we shall say that $\mathbf{X}$ has extremal index θ, $0 \leq \theta \leq 1$, if for each $\tau > 0$ there exists $\{u_{\mathbf{n}}^{(\tau)}\}_{\mathbf{n}\geq\mathbf{1}}$ such that, as $\mathbf{n} \longrightarrow \infty$, $n_1 n_2 P(X_{\mathbf{1}} > u_{\mathbf{n}}^{(\tau)}) \longrightarrow \tau$ and $P(M_{\mathbf{n}} \leq u_{\mathbf{n}}^{(\tau)}) \longrightarrow \exp(-\theta\tau)$.

When $\mathbf{X}$ has extremal index θ, then in (1) we have $H(x) = G^{\theta}(x)$ if and only if , as $\mathbf{n} \longrightarrow \infty$, $P(\widehat{M}_{\mathbf{n}}^{(1)} \leq a_{\mathbf{n}}x + b_{\mathbf{n}}) \rightarrow G(x)$, where $\widehat{M}_{\mathbf{n}}^{(1)} = \max\{\widehat{X}_{\mathbf{i}} : \mathbf{i} \in \mathbf{R}_{\mathbf{n}}\}$ and $\{\widehat{X}_{\mathbf{n}}\}$ is the associated i.i.d. random field with the same marginal distribution, F, as the original random field $\mathbf{X}$.

The study of the relationship between extreme values and their locations has important practical applications, for instance, when dealing with censored data. In Pereira [3] it was shown that the normalized location of the maximum of a stationary random field with extremal index $\theta \in (0,1]$ satisfying a long range dependence for each coordinate at a time, converges to a uniform variable on $[0,1]^2$ and is asymptotically independent of the height of the maximum. In this paper we study the joint locations of the k largest maxima of a stationary random field with extremal index $\theta \in (0,1]$ satisfying a slight generalization of $\Delta^*(u_{\mathbf{n}})-$condition of Ferreira and Pereira [4]. Furthermore, results concerning the location of a high level exceedance nearest of the site (0,0) are also presented.

2 Joint asymptotic behavior of locations of high values

The coordinatewise-mixing condition suitable in the present setting, is a slight generalization of $\Delta^*\left(u_{\mathbf{n}}(x)\right)$ condition of Ferreira and Pereira [4] which enable us to deal with the joint behavior of largest values, in disjoint rectangles, and their locations.

Definition 2.1 *Let $\mathbf{X}$ be a stationary random field and $\{u_{\mathbf{n}}(x_1)\}_{\mathbf{n}\geq\mathbf{1}}$, $\{u_{\mathbf{n}}(x_2)\}_{\mathbf{n}\geq\mathbf{1}}$ be sequences of real numbers. The coordinatewise-mixing condition $\Delta_2^*(u_{\mathbf{n}}(x_1), u_{\mathbf{n}}(x_2))$ is said to hold for $\mathbf{X}$ if there exist sequences of integer valued constants $\{k_{n_i}\}_{n_i\geq 1}$, $\{l_{n_i}\}_{n_i\geq 1}$, $i = 1,2$, such that, as $\mathbf{n} \rightarrow \infty$, we have*

$$(k_{n_1}, k_{n_2}) \longrightarrow \infty, \ \left(\frac{k_{n_1} l_{n_1}}{n_1}, \frac{k_{n_2} l_{n_2}}{n_2}\right) \longrightarrow \mathbf{0}, \ \left(k_{n_1}\Delta_{\mathbf{n},l_{n_1}}^{*(1)}, k_{n_1}k_{n_2}\Delta_{\mathbf{n},l_{n_2}}^{*(2)}\right) \longrightarrow \mathbf{0}, \tag{2}$$

where $\Delta_{\mathbf{n},l_{n_i}}^{*(i)}, i = 1,2$, are the components of the mixing coefficient defined as follows:

$$\Delta_{\mathbf{n},l_{n_1}}^{*(1)} = \sup\{|P(A\cap B) - P(A)P(B)| : A \in \mathcal{B}_{\mathbf{I}_1}(u_{\mathbf{n}}(x_1)), B \in \mathcal{B}_{\mathbf{I}_2}(u_{\mathbf{n}}(x_2))\}, \tag{3}$$

where $\mathcal{B}_{\mathbf{I}_j}(u_{\mathbf{n}})$, $j = 1,2$, denotes the σ-field generated by the events $\{X_{\mathbf{i}} \leq u_{\mathbf{n}} : \mathbf{i} \in \mathbf{I}_j\}$, $j = 1,2$, and the supremum is taken over pairs $\mathbf{I}_1$ and $\mathbf{I}_2$ in $S_1(l_{n_1})$,

$$\Delta_{\mathbf{n},l_{n_2}}^{*(2)} = \sup\{|P(A\cap B) - P(A)P(B)| : A \in \mathcal{B}_{\mathbf{I}_1}(u_{\mathbf{n}}(x_1)), B \in \mathcal{B}_{\mathbf{I}_2}(u_{\mathbf{n}}(x_2))\}, \tag{4}$$

where the supremum is taken over pairs $\mathbf{I}_1$ and $\mathbf{I}_2$ in $S_2(l_{n_2})$.

For $x_1 = x_2$, condition $\Delta_2^*(u_{\mathbf{n}}(x_1), u_{\mathbf{n}}(x_2))$ reduces to $\Delta^*(u_{\mathbf{n}})$ condition of Ferreira and Pereira [4]. Moreover, taking in (3) and (4) the events $A = \{M_{\mathbf{n}}^{(1)}(\mathbf{I}_1) \leq u_{\mathbf{n}}(x_1)\}$ and $B = \{M_{\mathbf{n}}^{(1)}(\mathbf{I}_2) \leq u_{\mathbf{n}}(x_2)\}$ we obtain the condition $\Delta_2\left(u_{\mathbf{n}}(x_1),u_{\mathbf{n}}(x_2)\right)$ defined in Pereira [3]. Furthermore, if we assume that $x_1=x_2$, condition $\Delta_2\left(u_{\mathbf{n}}(x_1),u_{\mathbf{n}}(x_2)\right)$ is the Cw-mixing condition of Leadbetter and Rootzén [2].

If in (3) and (4) we consider $u_{\mathbf{n}}(x_i),u_{\mathbf{n}}(x_j)$ each being any choice of the r values $u_{\mathbf{n}}(x_1)$, $u_{\mathbf{n}}(x_2),\ldots,u_{\mathbf{n}}(x_r)$, we say that $\mathbf{X}$ verifies the cordinatewise-mixing condition $\Delta_r^*\left(u_{\mathbf{n}}(x_1),\ldots,u_{\mathbf{n}}(x_r)\right)$.

Under $\Delta^*\left(u_{\mathbf{n}}\right)-$condition, Ferreira and Pereira [4] characterized the possible distributional limits for the sequence of normalized point processes of exceedances of $u_{\mathbf{n}}(x)$, by $\mathbf{X}$, defined on $[0,1]^2$ and for $\mathbf{n}\geq\mathbf{1}$ by

$$S_{\mathbf{n}}\left[\mathbf{X},u_{\mathbf{n}}(x)\right](\mathbf{B})=\sum_{\mathbf{i}\leq\mathbf{n}}\mathbb{I}_{\{X_{\mathbf{i}}>u_{\mathbf{n}}(x)\}}\delta_{\left(\frac{i_1}{n_1},\frac{i_2}{n_2}\right)}(\mathbf{B}),\quad \mathbf{B}\subseteq[0,1]^2,$$

where $\delta_{(a_1,a_2)}$ denotes the Dirac measure at $\mathbf{a}=(a_1,a_2)\in\mathbb{R}^2$.

If $\Delta(u_{\mathbf{n}}(x))$ holds and $\{S_{\mathbf{n}}\left[\mathbf{X},u_{\mathbf{n}}(x)\right]\}_{\mathbf{n}\geq\mathbf{1}}$, converges in distribution to some point process $\mathbf{S}$, then $\mathbf{S}$ is necessarily a compound Poisson point process.The Poisson rate ν can be obtained from the limit of zero exceedances, $\nu=-\log\lim_{\mathbf{n}\to\infty}P(M_{\mathbf{n}}\leq u_{\mathbf{n}})$, and the multiplicity distribution π is the limit of cluster size distribution $\pi_{\mathbf{n}}$, that is, $\pi(j)=\lim_{\mathbf{n}\to\infty}\pi_{\mathbf{n}}(j)=\lim_{\mathbf{n}\to\infty}P\left(\sum_{\mathbf{i}\leq\mathbf{r_n}}\mathbb{I}_{\{X_{\mathbf{i}}>u_{\mathbf{n}}\}}=j\,\middle|\sum_{\mathbf{i}\leq\mathbf{r_n}}\mathbb{I}_{\{X_{\mathbf{i}}>u_{\mathbf{n}}\}}>0\right)$, $j\in\mathbb{N}$, for some $\mathbf{r_n}=\left(\frac{n_1}{k_{n_1}},\frac{n_2}{k_{n_2}}\right)$ and $\mathbf{k_n}=(k_{n_1},k_{n_2})$ satisfying (2).

The next result establishes that if $\mathbf{I_i}$, $\mathbf{i}\leq\mathbf{k_n}$, are disjoint rectangles of $[0,1]^2$, then the $k_{n_1}k_{n_2}$ random variables $S_{\mathbf{n}}\left[\mathbf{X},u_{\mathbf{n},\mathbf{i}}\right](\mathbf{I_i})$, $\mathbf{i}\leq\mathbf{k_n}$, where, for each $\mathbf{i}$, $u_{\mathbf{n},\mathbf{i}}$ is any one of $u_{\mathbf{n}}(x_1),u_{\mathbf{n}}(x_2)$, are asymptotically independent and provides their distributional limit.

Proposition 2.1 *Let $\{X_{\mathbf{n}}\}_{\mathbf{n}\geq\mathbf{1}}$ be a stationary random field with extremal index $0<\theta\leq1$, $\{a_{\mathbf{n}}>0\}_{\mathbf{n}\geq\mathbf{1}}$ and $\{b_{\mathbf{n}}\}_{\mathbf{n}\geq\mathbf{1}}$ sequences of constants such that, as $\mathbf{n}\longrightarrow\infty$, $P(M_{\mathbf{n}}^{(1)}\leq a_{\mathbf{n}}x+b_{\mathbf{n}})\to G^{\theta}(x)$, with a non-degenerate distribution function G, and $\mathbf{I_i}$, $\mathbf{i}\leq\mathbf{k_n}$ be disjoint rectangles of $[0,1]^2$. If, for each $x_1,x_2\in\mathbb{R}$, and $u_{\mathbf{n}}(x_i)=a_{\mathbf{n}}x_i+b_{\mathbf{n}}$, $i=1,2$, $\mathbf{X}$ satisfies the condition $\Delta_2(u_{\mathbf{n}}(x_1),u_{\mathbf{n}}(x_2))$, then*
1. For any nonnegative integers $s_{\mathbf{i}}$, $\mathbf{i}\leq\mathbf{k_n}$,

$$d_{\mathbf{n}}=P\left(\bigcap_{\mathbf{i}\leq\mathbf{k_n}}\left\{S_{\mathbf{n}}\left[\mathbf{X},u_{\mathbf{n},\mathbf{i}}\right](\mathbf{I_i})=s_{\mathbf{i}}\right\}\right)-\prod_{\mathbf{i}\leq\mathbf{k_n}}P\left(S_{\mathbf{n}}\left[\mathbf{X},u_{\mathbf{n},\mathbf{i}}\right](\mathbf{I_i})=s_{\mathbf{i}}\right)\xrightarrow[\mathbf{n}\to\infty]{}0,$$

where, for each $\mathbf{i}$, $u_{\mathbf{n},\mathbf{i}}$ is any one of $u_{\mathbf{n}}(x_1),u_{\mathbf{n}}(x_2)$.
2. For each $k=1,2,\ldots,n_1n_2$,

$$\lim_{\mathbf{n}\to\infty}P\left(S_{\mathbf{n}}\left[\mathbf{X},u_{\mathbf{n}}(x)\right](\mathbf{I_1})\leq k-1\right)=G^{\theta|\mathbf{I_1}|}(x)\left(1+\sum_{i=1}^{k-1}\sum_{j=i}^{k-1}\frac{\left(-\log G^{\theta|\mathbf{I_1}|}(x)\right)^i}{i!}\pi^{*^i}(j)\right),$$

where π^{*^i} is the ith convolution of the multiplicity distribution π. We use the convention $\sum_{j=s}^{t}a_j=0$ if $t<s$.

The Proposition 2.1 will lead to the asymptotic joint distributions of various quantities of interest, such as, the locations of the k largest maxima, as $\mathbf{n}\to\infty$.

For the stationary random field $\mathbf{X}$, let $M_{\mathbf{n}}^{(i)}$ the ith largest maxima, $\Im_i=\{\mathbf{j}\in\mathbb{N}^2:X_{\mathbf{j}}=M_{\mathbf{n}}^{(i)}\}$, $\mathcal{P}_i=\{\mathbf{j}\in\Im_i:\forall\mathbf{j}'\in\Im_i,d(\mathbf{j},\mathbf{1})\leq d(\mathbf{j}',\mathbf{1})\}$, $\mathcal{Q}_i=\{\mathbf{j}\in\mathcal{P}_i:\forall\mathbf{j}'\in\mathcal{P}_i,j_1\leq j_1'\}$ and $\mathcal{R}_i=\{\mathbf{j}\in\mathcal{Q}_i:\forall\mathbf{j}'\in\mathcal{Q}_i,j_2\leq j_2'\}$.

We define the location of $M_{\mathbf{n}}^{(i)}$, $L_{\mathbf{n}}^{(i)}$, as follows

$$L_{\mathbf{n}}^{(i)} = \begin{cases} \mathbf{j}^{(1)} & \text{if } \mathcal{P}_i = \{\mathbf{j}^{(1)}\} \\ \mathbf{j}^{(2)} & \text{if } |\mathcal{P}_i| > 1 \text{ and } \mathcal{Q}_i = \{\mathbf{j}^{(2)}\} \\ \mathbf{j}^{(3)} & \text{if } |\mathcal{Q}_i| > 1 \text{ and } \mathcal{R}_i = \{\mathbf{j}^{(3)}\} \end{cases}.$$

Proposition 2.2 *Assume that $\Delta_2^*(u_{\mathbf{n}}(x_1), u_{\mathbf{n}}(x_2))$ holds for all $x_1, x_2 \in \mathbb{R}$. Then, for each $\varepsilon_1, \varepsilon_2 \in (0,1]$,*

$$\lim_{\mathbf{n}\to\infty} P\left(\bigcap_{i=1}^{k}\left\{L_{\mathbf{n}}^{(i)} \in [1, n_1\varepsilon_1]\times[1,n_2\varepsilon_2]\cap\mathbb{N}^2\right\}\right) = \varepsilon_1\varepsilon_2 - \sum_{s=1}^{k-1}(\varepsilon_1\varepsilon_2)^s(1-\varepsilon_1\varepsilon_2)\sum_{t=s}^{k-1}\pi^{*^s}(t)$$

Now we study the asymptotic behavior of the location of the exceedance of the levels $\{u_{\mathbf{n}}(x)\}_{\mathbf{n}\geq\mathbf{1}}$ by the stationary random field $\mathbf{X}$ nearest to the site (0,0).
Let $\Im_i^* = \left\{\mathbf{j}\in\mathbb{N}^2 : X_{\mathbf{j}} > u_{\mathbf{n}}(x)\right\}$, $\mathcal{P}_i^* = \{\mathbf{j}\in\Im_i^* : \forall \mathbf{j}'\in\Im_i^*, d(\mathbf{j},\mathbf{1}) \leq d(\mathbf{j}',\mathbf{1})\}$, $\mathcal{Q}_i{}^* = \{\mathbf{j}\in\mathcal{P}_i^* : \forall\mathbf{j}'\in\mathcal{P}_i^*, j_1\leq j_1'\}$ and $\mathcal{R}_i{}^* = \{\mathbf{j}\in\mathcal{Q}_i{}^* : \forall\mathbf{j}'\in\mathcal{Q}_i{}^*, j_2\leq j_2'\}$. We define the location of the exceedance of the levels $\{u_{\mathbf{n}}(x)\}_{\mathbf{n}\geq\mathbf{1}}$ nearest to the site (0,0), $L_{\mathbf{n}}^*$, as follows

$$L_{\mathbf{n}}^*(u_{\mathbf{n}}(x)) \equiv L_{\mathbf{n}}^* = \begin{cases} \mathbf{0} & \text{if } \Im_i^* = \emptyset \\ \mathbf{j}^{(1)} & \text{if } \mathcal{P}_i^* = \left\{\mathbf{j}^{(1)}\right\} \\ \mathbf{j}^{(2)} & \text{if } |\mathcal{P}_i^*| > 1 \text{ and } \mathcal{Q}_i{}^* = \left\{\mathbf{j}^{(2)}\right\} \\ \mathbf{j}^{(3)} & \text{if } |\mathcal{Q}_i{}^*| > 1 \text{ and } \mathcal{R}_i{}^* = \left\{\mathbf{j}^{(3)}\right\} \end{cases}.$$

Proposition 2.3 *Suppose that $\Delta_2^*(u_{\mathbf{n}}(x_1), u_{\mathbf{n}}(x_2))$ holds for all $x_1, x_2 \in \mathbb{R}$. Then, for each $\varepsilon_1, \varepsilon_2 \in (0,1]$,*

$$\lim_{\mathbf{n}\to\infty} P\left(L_{\mathbf{n}}^* \in [0, n_1\varepsilon_1]\times[0, n_2\varepsilon_2]\cap\mathbb{N}_{\mathbf{0}}^2\right) = G^{\theta}(x) + 1 - G^{\theta\varepsilon_1\varepsilon_2}(x).$$

The next result provides the asymptotic joint distribution of locations of the exceedance nearest to the site (0,0) and the high local maxima.

Proposition 2.4 *Suppose $\Delta_3^*(u_{\mathbf{n}}(x_1), u_{\mathbf{n}}(x_2), u_{\mathbf{n}}(x_3))$ holds for all $x_1, x_2, x_3 \in \mathbb{R}$. Then, for each $\varepsilon_1, \varepsilon_2 \in (0,1]$, $\delta_1, \delta_2 \in [0,1]$,*

$$\lim_{\mathbf{n}\to\infty} P(L_{\mathbf{n}}^* \in [0, n_1\delta_1]\times[0,n_2\delta_2]\cap\mathbb{N}_{\mathbf{0}}^2, L_{\mathbf{n}}^{(1)} \in [0,n_1\varepsilon_1]\times[0,n_2\varepsilon_2]\cap\mathbb{N}^2)$$
$$= \begin{cases} \varepsilon_1\varepsilon_2 G^{\theta}(x) & \text{if } \delta_1=\delta_2=0 \\ \varepsilon_1\varepsilon_2 - \frac{\varepsilon_1\varepsilon_2-\delta_1\delta_2}{1-\delta_1\delta_2}\left(G^{\theta\delta_1\delta_2}(x) - G^{\theta}(x)\right) & \text{if } \mathbf{0} < (\delta_1,\delta_2) \leq (\varepsilon_1,\varepsilon_2) < (1,1) \\ 1 & \text{if } \delta_1=\delta_2=\varepsilon_1=\varepsilon_2 \end{cases}.$$

References

[1] Choi, H. (2002). Central limit theory and extremes nof random fields. Phd Dissertation in Univ. of North Carolina at Chapel Hill.

[2] Leadbetter and Rootzén, H. (1998) On extreme values in stationary random fields. *Stochastic processes and related topics*, 275-285, Trends Math. Birkhauser Boston, Boston.

[3] Pereira, L. (2009) The asymptotic location of the maximum of a stationary random field. *Statistics* & *Probability Letters*, 79, 2166-2169.

[4] Ferreira, H. and Pereira, L. (2008). Point processes of exceedances by random fields. (Submitted)