
ON EXTREME VALUE ANALYSIS OF A SPATIAL PROCESS

Authors: LAURENS DE HAAN
– Erasmus University Rotterdam and University Lisbon,
The Netherlands and Portugal
ldehaan@few.eur.nl

CHEN ZHOU
– Tinbergen Institute, Erasmus University Rotterdam,
The Netherlands
zhou@few.eur.nl

Abstract:

- One common way to deal with extreme value analysis in spatial statistics is by using the max-stable process. By employing a representation of simple max-stable processes in de Haan and Ferreira ([3]), we propose a stationary max-stable process as a model of the dependence structure in two-dimensional spatial problems. We calculate its two-dimensional marginal distributions, which creates the opportunity to estimate the dependence parameter. The model is used in Buishand, de Haan and Zhou ([1]) for a spatial rainfall problem.

Key-Words:

- *spatial extremes; max-stable process.*

AMS Subject Classification:

- 62M30, 60G70.

1. INTRODUCTION

Problems of spatial statistics connected with high values of the spatial process need to be dealt with using extreme value theory (EVT), since the dependence between locations at high levels may differ from the dependence at moderate levels.

A case in point is the estimation of high quantiles of the total rainfall in a certain area. Engineers often need extreme rainfall statistics for the design of structures for flood protection. The observed rainfall data is only available on a few fixed monitoring stations. In order to study the high quantiles of the total rainfall, it is necessary to model the extreme rainfall process with dependence.

Considering the dependence structure, Cooley, Nychka and Naveau ([2]) used a Bayesian hierarchical model: locally the extreme rainfall is modeled by a one-dimensional EVT distribution and the parameters of this distribution follow some spatial dependence model.

A different way of introducing dependence is via a max-stable process. The mathematical setting of a spatial model for extreme rainfall is as follows. Consider independent replications of a stochastic process with continuous sample paths

$$\{X_n(t)\}_{t \in \mathbb{R}},$$

$n = 1, 2, \dots$. Suppose that the process is in the domain of attraction of a max-stable process, that is, there are sequences of continuous functions $a_n > 0$ and b_n such that as $n \rightarrow \infty$

$$(1.1) \quad \left\{ \frac{\max_{1 \leq i \leq n} X_i(t) - b_n(t)}{a_n(t)} \right\}_{t \in \mathbb{R}} \xrightarrow{w} \{\tilde{\eta}(t)\}_{t \in \mathbb{R}}$$

in C -space. Necessary and sufficient conditions have been given by de Haan and Lin ([4]). The limit process $\{\tilde{\eta}(t)\}$ is a max-stable process. Without loss of generality we can assume that the marginal distribution of $\tilde{\eta}$ can be written as

$$\exp \left\{ - (1 + \gamma(t)x)^{-1/\gamma(t)} \right\}$$

for all x with $1 + \gamma(t)x > 0$ where the function γ is continuous.

Buishand, de Haan and Zhou ([1]) simulated extreme rainfall from a max-stable process. Combining simulations of extreme rainfall with resampling from the non-extreme observations, an overview on the total rainfall can be generated. This is a novel solution for problems connected to both spatial statistics and extreme value analysis.

A major difficulty in the above methodology is to find a reasonable model for the max-stable process. With a suitable standardization, we can restrict ourselves to discussing the standardized process, called simple max-stable,

$$\{\eta(t)\} := \left\{ (1 + \gamma(t) \tilde{\eta}(t))_+^{1/\gamma(t)} \right\},$$

whose marginal distribution functions are all standard Fréchet: $\exp(-1/x)$, $x > 0$.

For application, it would be nice to have a stationary simple max-stable process. There are two different representations of stationary simple max-stable processes in literature. We consider one of them as follows, see Corollary 9.4.5, de Haan and Ferreira ([3]).

All simple max-stable process in $C^+(\mathbb{R})$ (the positive continuous functions on \mathbb{R}) can be generated in the following way. Consider a Poisson point process on $(0, +\infty]$ with mean measure dr/r^2 . Let $\{Z_i\}_{i=1}^\infty$ be a realization of this point process. Further consider i.i.d. stochastic processes V, V_1, V_2, \dots in $C^+(\mathbb{R})$ with $EV(s) = 1$ for all $s \in \mathbb{R}$ and $E \sup_{s \in I} V(s) < \infty$ for all compact interval I . Let the point process and the sequence V, V_1, V_2, \dots be independent. Then

$$(1.2) \quad \{\eta(s)\}_{s \in \mathbb{R}} \stackrel{d}{=} \left\{ \max_{i \geq 1} Z_i V_i(s) \right\}_{s \in \mathbb{R}}$$

is a simple max-stable process. Conversely each simple max-stable process has such a representation.

We use this result in a two-dimensional context and propose the following model

$$(1.3) \quad \eta(s_1, s_2) := \max_{i \geq 1} Z_i \exp \left\{ W_{1i}(\beta s_1) + W_{2i}(\beta s_2) - \beta(|s_1| + |s_2|)/2 \right\}$$

for $(s_1, s_2) \in \mathbb{R}^2$. The processes $W_{11}, W_{21}, W_{12}, W_{22}, W_{13}, W_{23}, \dots$ are independent copies of double-sided Brownian motions W defined as follows. Take two independent Brownian motions B_1 and B_2 . Then

$$(1.4) \quad W(s) := \begin{cases} B_1(s), & s \geq 0; \\ B_2(-s), & s < 0. \end{cases}$$

The positive constant β reflects the amount of spatial dependence at high levels of local observation: “ β small” means strong dependence and “ β large” means weak dependence. For this model, we shall prove that the dependence between extreme observations at two locations depends only on the distance between the locations.

The process η satisfies the requirements as follows:

$$E \exp\left\{W_1(\beta s_1) + W_2(\beta s_2) - \beta(|s_1| + |s_2|)/2\right\} = 1 \quad \text{for } (s_1, s_2) \in \mathbb{R}^2,$$

and

$$E \sup_{\substack{a_1 \leq s_1 \leq b_1 \\ a_2 \leq s_2 \leq b_2}} \exp\left\{W_1(\beta s_1) + W_2(\beta s_2) - \beta(|s_1| + |s_2|)/2\right\} < \infty$$

for all $a_1 < b_1, a_2 < b_2$ real.

Meanwhile, the one-dimensional marginal distribution functions of (1.3) are all $e^{-1/x}$, $x > 0$. Notice that only a one-dimensional Poisson point process is used in η . Thus, this process is easy to simulate.

Similar to de Haan and Pereira ([5]), in order to use this model in studying spatial extremes, we have to prove that the process η is shift stationary and we have to calculate the two-dimensional marginal distributions.

Since the two-dimensional process η is a combination of two one-dimensional processes, for the stationarity it is sufficient to prove the same for the one-dimensional version, i.e. that the process

$$(1.5) \quad \eta'(s) := \max_{i \geq 1} Z_i \exp\left\{W_{1i}(\beta s_1) - \beta|s_1|/2\right\}$$

is stationary. This follows from the fact that the process η' can be obtained as the limit of the pointwise maximum of i.i.d. Ornstein–Uhlenbeck processes (cf. e.g. Example 9.8.2, de Haan and Ferreira ([3])). The stationarity follows from the stationarity of the Ornstein–Uhlenbeck process.

It remains to calculate the two-dimensional marginal distributions. This is done in Section 2.

2. THE TWO-DIMENSIONAL MARGINAL DISTRIBUTION OF η

The two-dimensional marginal distribution of η' in (1.5) is calculated in de Haan and Ferreira ([3]), section 9.8. We state it as the following proposition.

Proposition 2.1. *Suppose $\{\eta'(s)\}_{s \in \mathbb{R}}$ is defined as in (1.5). Then for $x, y \in \mathbb{R}$ and $s_1, s_2 \in \mathbb{R}$,*

$$\begin{aligned} -\log P\left(\eta'(s_1) \leq e^x, \eta'(s_2) \leq e^y\right) &= \\ &= e^{-x} \Phi\left(\frac{\sqrt{|s_1 - s_2|}}{2} + \frac{-x + y}{\sqrt{|s_1 - s_2|}}\right) + e^{-y} \Phi\left(\frac{\sqrt{|s_1 - s_2|}}{2} + \frac{x - y}{\sqrt{|s_1 - s_2|}}\right). \end{aligned}$$

This is useful in similar calculation for the two-dimensional process η . Besides Proposition 2.1, we need the following Lemma.

Lemma 2.1. *Suppose N is normally distributed with mean 0, variance u , then with non-random constants $a > 0$ and b ,*

$$(2.1) \quad E e^{N-u/2} \Phi(aN+b) = \Phi\left(\frac{au+b}{\sqrt{a^2u+1}}\right).$$

Proof: Suppose N_1 is standard normally distributed, and independent of N , then we have

$$E e^{N-u/2} 1_{N_1 \leq aN+b} = E_N E\left(e^{N-u/2} 1_{N_1 \leq aN+b} \mid N\right) = E e^{N-u/2} \Phi(aN+b),$$

which is the left side of (2.1). By Fubini's Theorem, it can be recalculated in the following way

$$\begin{aligned} E e^{N-u/2} 1_{N_1 \leq aN+b} &= E_{N_1} E\left(e^{N-u/2} 1_{N_1 \leq aN+b} \mid N_1\right) \\ &= E_{N_1} \int_{\frac{N_1-b}{a}}^{\infty} e^{t-u/2} \frac{1}{\sqrt{2\pi u}} e^{-\frac{t^2}{2u}} dt \\ &= E_{N_1} \int_{\frac{N_1-b}{a}}^{\infty} \frac{1}{\sqrt{2\pi u}} e^{-\frac{(t-u)^2}{2u}} dt \\ &= E_{N_1} \left(1 - \Phi\left(\frac{N_1-b}{a\sqrt{u}} - \sqrt{u}\right)\right). \end{aligned}$$

By a similar trick — introducing a standard normal variable N_2 independent of N_1 , the calculation can be finished to prove the lemma.

$$\begin{aligned} E_{N_1} \left(1 - \Phi\left(\frac{N_1-b}{a\sqrt{u}} - \sqrt{u}\right)\right) &= E_{N_1} E\left(1_{N_2 \geq \frac{N_1-b}{a\sqrt{u}} - \sqrt{u}} \mid N_1\right) \\ &= E_{N_1, N_2} 1_{N_2 \geq \frac{N_1-b}{a\sqrt{u}} - \sqrt{u}} \\ &= P\left(N_2 \geq \frac{N_1-b}{a\sqrt{u}} - \sqrt{u}\right) \\ &= \Phi\left(\frac{au+b}{\sqrt{a^2u+1}}\right). \end{aligned}$$

We remark that the last calculation is similar to that of Lemma 2.1 in Gupta, González-Farías and Domínguez-Molina ([6]). \square

The lemma can be used to derive the two-dimensional marginal distributions as follows. As in the proof of Proposition 2.1 (cf. de Haan and Ferreira ([3]),

Section 9.8), we have

$$\begin{aligned}
(2.2) \quad & -\log P\left(\eta(u_1, u_2) \leq e^x, \eta(v_1, v_2) \leq e^y\right) = \\
& = E \max\left(e^{W_1(\beta u_1) + W_2(\beta u_2) - (|\beta u_1| + |\beta u_2|)/2 - x}, e^{W_1(\beta v_1) + W_2(\beta v_2) - (|\beta v_1| + |\beta v_2|)/2 - y}\right) \\
& = E_{W_1} E\left(\max\left(e^{W_1(\beta u_1) + W_2(\beta u_2) - (|\beta u_1| + |\beta u_2|)/2 - x}, \right. \right. \\
& \quad \left. \left. e^{W_1(\beta v_1) + W_2(\beta v_2) - (|\beta v_1| + |\beta v_2|)/2 - y}\right) \mid W_1\right) \\
& = E e^{-x + W_1(\beta u_1) - \beta|u_1|/2} \\
& \quad \cdot \Phi\left(\frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x + W_1(\beta u_1) - W_1(\beta v_1) - \beta|u_1|/2 + \beta|v_1|/2}{\sqrt{\beta|u_2 - v_2|}}\right) \\
& + E e^{-y + W_1(\beta v_1) - \beta|v_1|/2} \\
& \quad \cdot \Phi\left(\frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{x - y + W_1(\beta v_1) - W_1(\beta u_1) - \beta|v_1|/2 + \beta|u_1|/2}{\sqrt{\beta|u_2 - v_2|}}\right).
\end{aligned}$$

Now we can calculate the two parts in (2.2) separately. Without losing generality, we only focus on the first part.

Case 1: $0 \leq u_1 \leq v_1$.

In this case $e^{-x + W_1(\beta u_1) - \beta|u_1|/2}$ is independent of the other part. Hence,

$$\begin{aligned}
& E e^{-x + W_1(\beta u_1) - \beta|u_1|/2} \\
& \quad \cdot \Phi\left(\frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x + W_1(\beta u_1) - W_1(\beta v_1) - \beta|u_1|/2 + \beta|v_1|/2}{\sqrt{\beta|u_2 - v_2|}}\right) = \\
& = e^{-x} E \Phi\left(\frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x - (W_1(\beta v_1) - W_1(\beta u_1) - \beta(v_1 - u_1)/2)}{\sqrt{\beta|u_2 - v_2|}}\right) \\
& = e^{-x} P\left(N \leq \frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x - (W_1(\beta v_1) - W_1(\beta u_1) - \beta(v_1 - u_1)/2)}{\sqrt{\beta|u_2 - v_2|}}\right) \\
& = e^{-x} \Phi\left(\frac{\sqrt{\beta|u_2 - v_2| + \beta(v_1 - u_1)}}{2} + \frac{y - x}{\sqrt{\beta|u_2 - v_2| + \beta(v_1 - u_1)}}\right).
\end{aligned}$$

Case 2: $0 \leq v_1 < u_1$.

Note that $E e^{W_1(\beta v_1) - \beta v_1/2} = 1$ and $W_1(\beta v_1)$ is independent of $W_1(\beta u_1) - W_1(\beta v_1)$,

we have

$$\begin{aligned}
& E e^{-x+W_1(\beta u_1)-\beta|u_1|/2} \\
& \cdot \Phi\left(\frac{\sqrt{\beta|u_2-v_2|}}{2} + \frac{y-x+W_1(\beta u_1)-W_1(\beta v_1)-\beta|u_1|/2+\beta|v_1|/2}{\sqrt{\beta|u_2-v_2|}}\right) = \\
& = e^{-x} E e^{W_1(\beta u_1)-W_1(\beta v_1)-\beta(u_1-v_1)/2} \\
& \cdot \Phi\left(\frac{\sqrt{\beta|u_2-v_2|}}{2} + \frac{y-x+W_1(\beta u_1)-W_1(\beta v_1)-\beta|u_1|/2+\beta|v_1|/2}{\sqrt{\beta|u_2-v_2|}}\right).
\end{aligned}$$

Since $W_1(\beta u_1) - W_1(\beta v_1)$ is normally distributed with mean 0, variance $\beta(u_1 - v_1)$, we can apply Lemma 2.1 with the constants $a = 1/\sqrt{\beta|u_2 - v_2|}$, $u = \beta(u_1 - v_1)$ and

$$b = \frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x - \beta u_1/2 + \beta v_1/2}{\sqrt{\beta|u_2 - v_2|}}.$$

The final result is

$$\begin{aligned}
& E e^{-x+W_1(\beta u_1)-\beta|u_1|/2} \\
& \cdot \Phi\left(\frac{\sqrt{\beta|u_2-v_2|}}{2} + \frac{y-x+W_1(\beta u_1)-W_1(\beta v_1)-\beta|u_1|/2+\beta|v_1|/2}{\sqrt{\beta|u_2-v_2|}}\right) = \\
& = e^{-x} \Phi\left(\frac{\sqrt{\beta|u_2-v_2|+\beta(u_1-v_1)}}{2} + \frac{y-x}{\sqrt{\beta|u_2-v_2|+\beta(u_1-v_1)}}\right).
\end{aligned}$$

Case 3: $v_1 < u_1 < 0$ and $u_1 \leq v_1 < 0$.

These two cases are similar to Case 1 and 2 respectively. The final results are all the same as follows.

$$\begin{aligned}
& E e^{-x+W_1(\beta u_1)-\beta|u_1|/2} \\
& \cdot \Phi\left(\frac{\sqrt{\beta|u_2-v_2|}}{2} + \frac{y-x+W_1(\beta u_1)-W_1(\beta v_1)-\beta|u_1|/2+\beta|v_1|/2}{\sqrt{\beta|u_2-v_2|}}\right) = \\
& = e^{-x} \Phi\left(\frac{\sqrt{\beta|u_2-v_2|+\beta|u_1-v_1|}}{2} + \frac{y-x}{\sqrt{\beta|u_2-v_2|+\beta|u_1-v_1|}}\right).
\end{aligned}$$

Case 4: u_1 and v_1 have different signs.

In this case $W_1(\beta u_1)$ and $W_1(\beta v_1)$ are independent, we can calculate the expectation with respect to $W_1(\beta v_1)$ first, then with respect to $W_1(\beta u_1)$.

$$\begin{aligned}
& E e^{-x+W_1(\beta u_1)-\beta|u_1|/2} \\
& \cdot \Phi\left(\frac{\sqrt{\beta|u_2-v_2|}}{2} + \frac{y-x+W_1(\beta u_1)-W_1(\beta v_1)-\beta|u_1|/2+\beta|v_1|/2}{\sqrt{\beta|u_2-v_2|}}\right) = \\
& = e^{-x} E e^{W_1(\beta u_1)-\beta|u_1|/2} \Phi\left(\frac{\sqrt{\beta|u_2-v_2|+\beta|v_1|}}{2} + \frac{y-x+W_1(\beta u_1)-\beta|u_1|/2}{\sqrt{\beta|u_2-v_2|+\beta|v_1|}}\right).
\end{aligned}$$

Now we can again apply Lemma 2.1 with the constants $a = 1/\sqrt{\beta|u_2 - v_2| + \beta|v_1|}$, $u = \beta|u_1|$ and

$$b = \frac{\sqrt{\beta|u_2 - v_2| + \beta|v_1|}}{2} + \frac{y - x - \beta|u_1|/2}{\sqrt{\beta|u_2 - v_2| + \beta|v_1|}}$$

to get that

$$\begin{aligned} & E e^{-x+W_1(\beta u_1)-\beta|u_1|/2} \\ & \cdot \Phi\left(\frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x + W_1(\beta u_1) - W_1(\beta v_1) - \beta|u_1|/2 + \beta|v_1|/2}{\sqrt{\beta|u_2 - v_2|}}\right) = \\ & = e^{-x} \Phi\left(\frac{\sqrt{\beta|u_2 - v_2| + \beta(|u_1| + |v_1|)}}{2} + \frac{y - x}{\sqrt{\beta|u_2 - v_2| + \beta(|u_1| + |v_1|)}}\right). \end{aligned}$$

Notice that due to the different signs of u_1 and v_1 , $|u_1 - v_1| = |u_1| + |v_1|$.

By defining $h = |u_1 - v_1| + |u_2 - v_2|$, all these cases can be combined together as

$$\begin{aligned} & E e^{-x+W_1(\beta u_1)-\beta|u_1|/2} \\ & \cdot \Phi\left(\frac{\sqrt{\beta|u_2 - v_2|}}{2} + \frac{y - x + W_1(\beta u_1) - W_1(\beta v_1) - \beta|u_1|/2 + \beta|v_1|/2}{\sqrt{\beta|u_2 - v_2|}}\right) = \\ & = e^{-x} \Phi\left(\frac{\sqrt{\beta h}}{2} + \frac{y - x}{\sqrt{\beta h}}\right). \end{aligned}$$

Symmetrically, the second part of (2.2) can be simplified as

$$e^{-y} \Phi\left(\frac{\sqrt{\beta h}}{2} + \frac{x - y}{\sqrt{\beta h}}\right).$$

Combining these two parts, we get the following theorem about the two-dimensional marginal distribution of η .

Theorem 2.1. *Suppose the simple max-stable process η is defined in (1.3). Given any two coordinates (u_1, u_2) and (v_1, v_2) on \mathbb{R}^2 , denote the distance between them as $h := |u_1 - v_1| + |u_2 - v_2|$. Then the two-dimensional distribution function of $(\eta(u_1, u_2), \eta(v_1, v_2))$ is*

$$(2.3) \quad P\left(\eta(u_1, u_2) \leq e^x, \eta(v_1, v_2) \leq e^y\right) = \exp\left\{-\left(e^{-x} \Phi\left(\frac{\sqrt{\beta h}}{2} + \frac{y - x}{\sqrt{\beta h}}\right) + e^{-y} \Phi\left(\frac{\sqrt{\beta h}}{2} + \frac{x - y}{\sqrt{\beta h}}\right)\right)\right\},$$

where Φ is the standard normal distribution function and $x, y \in \mathbb{R}$.

Note that the two-dimensional marginal distribution depends on only h . It agrees with the shift stationarity discussed in Section 1.

Similar to de Haan and Pereira ([5]), Theorem 2.1 is useful in estimating β . By taking $x = y = 0$, we get that

$$P\left(\eta(u_1, u_2) \leq 1, \eta(v_1, v_2) \leq 1\right) = \exp\left\{-2\Phi\left(\frac{\sqrt{\beta h}}{2}\right)\right\}.$$

Consequently, we have that

$$\beta = \frac{4}{h} \left(\Phi^{\leftarrow} \left(-\frac{1}{2} \log P\left(\eta(u_1, u_2) \leq 1, \eta(v_1, v_2) \leq 1\right) \right) \right)^2.$$

Hence we can estimate β if we know how to estimate

$$L_{(u_1, u_2), (v_1, v_2)}(1, 1) := -\log P\left(\eta(u_1, u_2) \leq 1, \eta(v_1, v_2) \leq 1\right).$$

In fact, this problem has been solved by Huang and Mason (cf. Huang ([8]), Drees and Huang ([7])). Suppose we have i.i.d. observations of η as η_1, η_2, \dots . Write $\{\eta_{i,n}(s_1, s_2)\}_{i=1}^n$ for the order statistics at location (s_1, s_2) . Then the estimator

$$\hat{L}_{(u_1, u_2), (v_1, v_2)}^{(k)}(1, 1) := \frac{1}{k} \sum_{j=1}^n \mathbf{1}_{\{\eta_j(u_1, u_2) \geq \eta_{n-k+1, n}(u_1, u_2) \text{ or } \eta_j(v_1, v_2) \geq \eta_{n-k+1, n}(v_1, v_2)\}}$$

is consistent provided $k = k(n) \rightarrow \infty$, $k(n)/n \rightarrow 0$, $n \rightarrow \infty$. It is asymptotically normal under certain mild extra conditions.

Hence, from the two-dimensional marginal distribution, we can estimate β when we have the observation at two specific locations. An application of this method is in Buishand, de Haan and Zhou ([1]), Section 5.

ACKNOWLEDGMENTS

Research partially supported by FCT/POCI2010 and EXES–Project PTDC/MAT/64924/2006.

REFERENCES

- [1] BUISSHAND, T.A.; DE HAAN, L. and ZHOU, C. (2008). On spatial extremes: with application to a rainfall problem, *Ann. Appl. Statist.*, To appear.
- [2] COOLEY, D.; NYCHKA, D. and NAVEAU, P. (2007). Bayesian spatial modeling of extreme precipitation return levels, *J. Amer. Statist. Association*, **102**, 479, 824–840.
- [3] DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction*, Springer.
- [4] DE HAAN, L. and LIN, T. (2001). On convergence towards an extreme value distribution in $C[0,1]$, *Ann. Prob.*, **29**, 467–483.
- [5] DE HAAN, L. and PEREIRA, T.T. (2006). Spatial Extremes: Models for the stationary case, *Ann. Statist.*, **34**, 146–168.
- [6] GUPTA, A.; GONZÁLEZ-FARÍAS, G. and DOMÍNGUEZ-MOLINA, J. (2004). A multivariate skew normal distribution, *J. Multivariate Anal.*, **89**, 181–190.
- [7] DREES, H. and HUANG, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function, *J. Multivariate Anal.*, **64**, 25–47.
- [8] HUANG, X. (1992). *Statistics of Bivariate Extreme Values*, PhD Thesis, Tinbergen Institute.